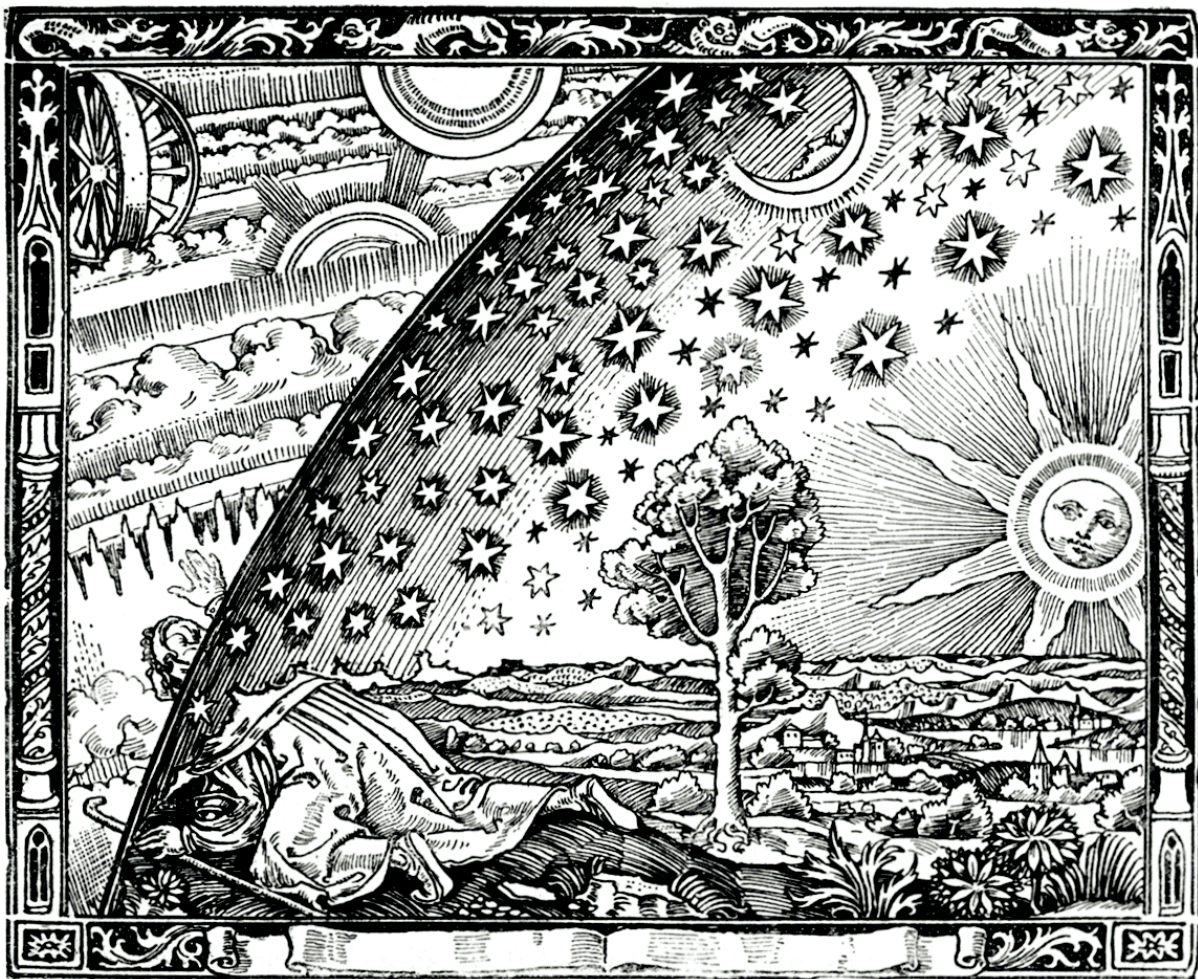


Traces, Enigmes, Problèmes :

Emergence et construction du sens

13èmes journées de Rochebrune :
Rencontres interdisciplinaires sur les systèmes
complexes naturels et artificiels

22 au 27 janvier 2006, Rochebrune, Megève



ENST 2006 S001
École Nationale des Télécommunications





ENST 2006 S001

Traces, Enigmes, Problèmes :

Emergence et construction du sens

13èmes journées de Rochebrune :
Rencontres interdisciplinaires sur les systèmes
complexes naturels et artificiels

22 au 27 janvier 2006, Rochebrune, Megève

Organisées avec le soutien

de l'European Conference on Artificial Life (ECAL)
de la Délégation Générale pour l'Armement (DGA)

Sous le patronage

du Centre National de la Recherche Scientifique (CNRS),
département Sciences et Technologies de l'Information, et de
la Communication (STIC), section 7 du Comité National.

École Nationale Supérieure des Télécommunications

Traces, Enigmes, Problèmes : Emergence et construction du sens

Les Journées de Rochebrune 2006

L'énigme, suppose que, partant de données a priori obscures, celui qui s'y trouve confronté peut la résoudre au terme d'une série d'opérations destinées à « donner, construire, faire émerger » un sens.

Ce qui est posé comme énigme, soit par l'émetteur soit par le récepteur d'un message, possède toujours deux faces, dont l'une est exposée et l'autre supposée « voilée ». La face voilée sollicite un travail de déchiffrement consistant à inventorier des éléments, naturels ou artificiels, que l'on pose devant soi. Le sens naîtra de notre « capacité à faire des liens », ce qui correspond à un comportement cognitif dit intelligent. D'une manière analogue, tout problème scientifique issu de la réalité provient d'une mise en forme intersubjective d'observations et de projets. Le problème ne se pose pas complet dans son énoncé, mais procède d'un ensemble de traces signifiantes et de modèles partiels, qui sont les premiers pas de la résolution, pas qui seront complétés par la recherche de nouveaux signes et par la mise en cohérence des modèles partiels. Le raisonnement abductif puis inductif sert à la fois d'amorce, de révision et d'explication, qui fonctionnent de façon évolutive dans la production d'un savoir. Dans une perspective de « cognition située », formuler une explication entraîne la production de traces écrites censées éclairer des pratiques intellectuelles. Dès lors, si l'on considère que les questions de connaissance (y compris « savante ») se posent relativement à un champ énigmatique, la construction du savoir s'amorce par un travail de repérage, de décodage, qui sollicite une activité cognitive complexe de modélisation incrémentale. Cette dernière pose la question de la rétroaction, de la trace de l'activité engendrée par le processus modélisateur même. Ainsi, toutes les disciplines visant à faire émerger, ou à construire une interprétation à partir d'un ensemble de données incomplètes ou dégradées sont concernées par la problématique de l'énigme et de la trace. De la même façon, toutes les démarches systémiques, qui vérifient les traces des rétroactions sur un ensemble finalisé et les interprètent dans le but de maintenir l'ensemble dans sa trajectoire, relèvent aussi de ce domaine.

Rapports aux différents champs disciplinaires

On proposera ci-dessous quelques exemples, non limitatifs, des disciplines qui peuvent être questionnées par la problématique des journées de Rochebrune 2006 :

- En **droit** et en **sociologie du droit** se pose le problème du code, de ses interprétations, de l'esprit et de la lettre. Mais aussi et surtout, de ce qui relève de l'instruction, de la quête de la « vérité » définie par la recherche d'explications cohérentes portant sur un ensemble de traces devenant des indices. A l'inverse se trouve posé le problème du secret.
- Pour les **sciences de l'ingénieur**, du **génie civil**, les **ergonomes** et les spécialistes du **design**, la question se pose exactement à l'inverse du jeu de l'obscurité : comment rendre un objet, un logiciel, une voiture, utilisables de façon « intuitive », sans qu'ils

paraissent à l'acquéreur totalement énigmatiques, dans leur apparence et leur fonctionnement.

- La question est la même dans les **espaces géographiques et urbains** où la signalétique doit permettre de décoder sans difficultés les signes nécessaires aux repérages.
- En **informatique** on identifie les problèmes suivants : codage/encodage, fichiers de log et entêtes de mails et/ou routage, annotations et surannotation de documents, interprétation des actions d'utilisateurs, profilage, cryptographie, cryptanalyse. Dans le domaine des multi-agents, les raisonnements pour la révision de connaissances relèvent d'émergences toujours partielles à partir de captures de traces dans l'environnement ou dans les messages échangés.
- En **mathématiques** : la théorie des nombres ouvre sur les tentatives de rendre fiables les codes et les chiffres. En arrière plan se trouve également convoqué tout un domaine de jeux et d'énigmes mathématiques. On peut ajouter ici la question de l'émergence des symboles utilisés, leur apparition, leur persistance. Comment un signe s'impose plutôt qu'un autre ?
- En **théories des jeux**, on rencontre le problème des connaissances partagées, des paradoxes de la rationalité complète (en relation avec les mathématiques et la logique), du processus d'apprentissage, de prise de décision, d'adaptation et d'émergence de croyances sociales en information incomplète et rationalité affaiblie.
- En **économie** : problème de la coordination et de l'interaction ; traces d'activité individuelles dans les mouvements de masses et d'opinions, prix, flux de biens et services. L'*homo cogitans* construit et ajuste les croyances qui conditionnent ses délibérations à partir des traces de l'action des autres. Eléments qui rejoignent, dans le monde de la **finance**, l'interprétation des signaux émis par les « marchés », les observateurs et les experts. En **gestion**, reconstruction du sens de l'activité par des systèmes d'information et des indicateurs synthétiques. En **management**, interprétation des attitudes et des comportements, en relation avec les formes d'organisation.
- La question de la trace et de l'énigme concerne également et au premier chef les disciplines en rapport avec la **sémiotique**, les **théories de la communication**, **l'histoire**, **l'archéologie**.
- En **linguistique**, **herméneutique**, **narratologie** et **paléographie** on analyse ou on propose des parcours interprétatifs selon les contextes, les relations fond/forme, les intentions communicatives déchiffrables (ou non), en fonction des choix de mise en page. L'histoire du déchiffrement c'est aussi celle des écritures anciennes. Elles ne cherchaient pas a priori à être indéchiffrables mais, pour bon nombre d'entre elles, nous en avons perdu les clefs.
- La trace comme amorce hypothétique est prégnante dans les **sciences de la terre et de l'Univers**, où la construction des hypothèses explicatives repose essentiellement sur la transformation des traces en indices de processus anciens, ou en signes de transformations à venir.
- En **physique**, par exemple, on cherche à faire le lien entre les théories physiques et les signes du temps.
- Les sciences de la « santé » sont essentiellement fondées, comme la **médecine**, sur l'interprétation de traces en symptômes (sémiologie médicale), de symptômes en causes, ce qui conduit à choisir des remèdes ...
- En **psychanalyse**, la question de l'énigme est corrélative de celle du rêve et de la connaissance de sa propre psyché, avec, en arrière-plan, la question de savoir qui est habilité à interpréter quoi et comment.

- Les **sciences humaines**, réflexives s'il en est, étudient également les activités cognitives dites « populaires » folklore, traditions, jeux d'énigmes, transmission cryptées des savoirs.

Enfin, en dehors de l'investigation scientifique, il ne faut pas omettre ce qui est la « trace » ultime de l'homme face à un univers énigmatique. Depuis la première empreinte de doigts imprégnés de peinture sur la paroi d'une grotte jusqu'aux expressions les plus élaborées, **l'art** (pictural, musical, architectural, poétique, etc.) n'est-il pas la forme sublimée de la trace produite, de la trace comme un défi au temps et à l'évanescence ?

L'objectif des Journées de Rochebrune

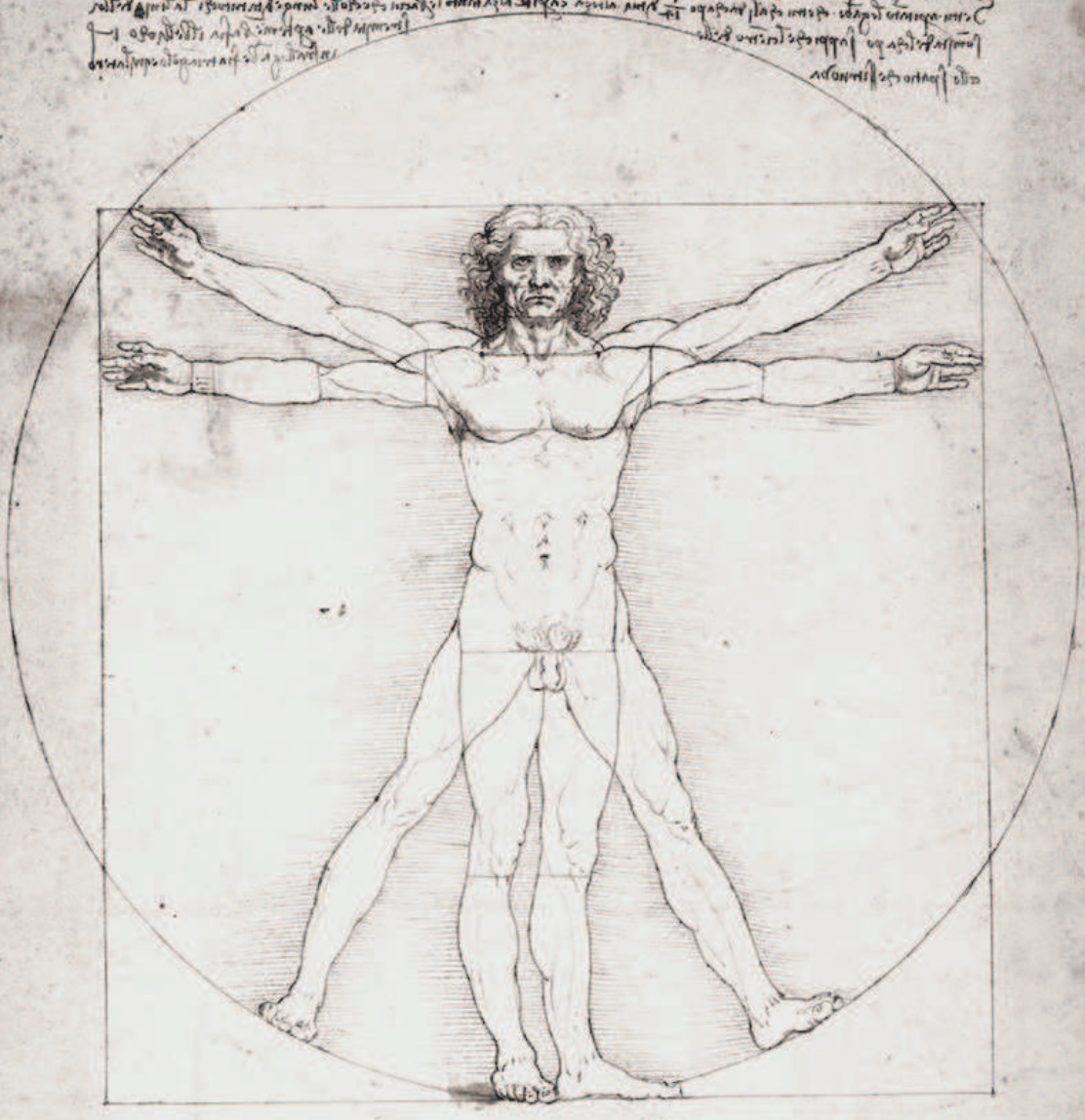
Depuis 1992, l'objectif des Journées De Rochebrune est d'offrir un espace d'échanges et de débats interdisciplinaires à tous ceux qui travaillent sur les systèmes complexes naturels ou artificiels. Il est demandé aux participants un effort tout particulier de vulgarisation dans la présentation de leurs travaux afin que le débat soit fructueux entre les champs disciplinaires.

Caen, le 18 décembre 2005

Organisateurs

n. 1.
A.

Handwritten text in Italian, likely a preface or introduction to the drawing, discussing the proportions of the human body and the golden ratio.



Handwritten text below the drawing, possibly a scale or measurement reference.

B.

Main body of handwritten text in Italian, providing detailed notes and measurements related to the drawing, including references to Vitruvius and the golden ratio.

Leonardo da Vinci

Programme de la semaine :

Dimanche 22 janvier

17h00-19h30 *Accueil des participants*

19h30-20h30 *Repas*

20h30-21h00 Ouverture de Rochebrune 2006 - Détails et programme

21h00-21h45 ..**S. Mauger** Oedipe et le sphinx : une question d'heuristique..... **11**

21h45-22h30 **Nadine Lucas**, Quelques textes entre l'énigme et la devinette **21**

Lundi 23 janvier

16h00-16h45 **F. Pfaänder & M. Jacomy**, Explorer et appréhender l'Internet..... **35**

16h45-17h30 **P. Beust & T. Roy**, Utiliser des traces de la dimension globale
d'un corpus pour l'accès au contenu des documents **47**

17h30-18h00 *Pause*

18h00-18h45 **P.S. Luquet & E. Giguët**, Rapports entre traces et signes dans
l'activité de langage : Exemples de décryptages linguistiques assisté par
ordinateur..... **61**

18h45-19h30 **A. Nicolle**, Un modèle des traces pour l'interaction entre les processus **75**

19h30-20h30 *Repas*

20h30-21h15 **F. Rousseaux & T. Bouaziz**, Construction et déconstruction des
connaissances par les informaticiens **89**

21h15-22h00 **Pek van Andel**, Traces et sérendipité

Mardi 24 janvier

16h00-16h45 **D. Bourcier**, Du roman policier à la criminologie: traces et
abduction **103**

16h45-17h30 **R. Cozien**, Traces légales **119**

17h30-18h00 *Pause*

| | | |
|-------------|---|------------|
| 18h00-18h45 | D. Battistelli, J.L. Minel & S. Schwer , Signification et Contextes..... | 129 |
| 18h45-19h30 | J. Cardot & S. Schwer , Sur les traces d'un temps fractal..... | 139 |
| 19h30-20h30 | <i>Repas</i> | |
| 20h30-22h00 | Atelier <i>Police Scientifique</i> animé par Roger Cozien | |

Mercredi 25 janvier

| | | |
|-------------|--|------------|
| 16h00-16h45 | M. Roche , Acquisition de la terminologie et définition des tâches à effectuer, deux principes indissociables | 151 |
| 16h45-17h30 | N. Ferrand & R. Raymond , De l'observation à l'intervention dans les processus de gestion ou faut-il verbaliser les sens interdits ? | 163 |
| 17h30-18h00 | <i>Pause</i> | |
| 18h00-18h45 | C. Champendal , Trace, signe et apprentissage..... | 177 |
| 18h45-19h30 | A. Mille & Y. Prié , Une théorie de la trace informatique pour faciliter l'adaptation dans la confrontation logique d'utilisation/logique de conception | 183 |
| 19h30-20h30 | <i>Repas</i> | |
| 20h30-21h15 | I. Yocaris , Signes et interprétation des signes dans la fiction postmoderne : le schéma de l'enquête dans le Nom de la Rose | 197 |
| 21h15-22h00 | F. Barbin , Opacité et traduction : le cas des récits étiologiques..... | 209 |

Jeudi 26 janvier

| | | |
|-------------|--|------------|
| 16h00-16h45 | A. Ciaccia , Demandes documentaires en contexte professionnel : Analyse cognitive de l'explicitation de la demande | 225 |
| 16h45-17h30 | T. Carron, J.C. Marty, J.M. Herauld & F. Laure , Préparer une expérimentation pédagogique observée | 237 |
| 17h30-18h00 | <i>Pause</i> | |
| 18h00-18h45 | F. Amblard, A. Penciolelli , L'énigme de l'articulation entre comportements individuels et collectifs dans les simulations multi-agents : problèmes et méthodes | 249 |
| 18h45-19h30 | M. & J. Nanard, J.-Y. Delort , Enigmes et traces. Le rôle de la décontextualisation et de la re-contextualisation dans le processus d'écriture lecture | 261 |
| 19h30-20h30 | <i>Repas</i> | |

20h30-23h00 Rochebrune 2007 : Thème, organisation etc ...

Vendredi 27 janvier

- 16h00-16h45 **M. Groccia**, La confrontation des systèmes sémio-linguistique
et musical en chanson : l'autre voie du sens..... **279**
- 16h45-17h30 **A. Banos**, Des trajectoires individuelles à la ville en mouvement :
reconstruire la mobilité urbaine quotidienne **289**
- 17h30-18h00 *Pause*
- 18h00-18h45 **S. Occelli**, Dealing with the not known as a leverage in innovative
policy making..... **301**
- 18h45-19h30 **S. Lardon**, Tracer laisse-t-il des traces ? ou Comment la
représentation spatiale de territoires de projets contribue à la
conception collaborative de projets de territoire..... **313**
- 19h30-20h30 *Repas*
- 20h30-21h15 **D. Badariotti**, Les processus morphodynamiques des villes :
énigmes et traces **325**
- 21h15-22h00 **O. Soubeyran**, Une histoire de faits : le bon, le brut et le truand
- 22h00 Clôture de Rochebrune 2006 - Discussions



Œdipe et le sphinx : une question d'heuristique.

Serge Mauger

GREYC CNRS UMR 6072 – ISLanD & Pôle ModesCoS de la MRSH
Université de Caen - Basse Normandie
14032 Caen Cedex, France
mauger@iut3.unicaen.fr

Pourquoi, lorsque la Chienne chantait ici ses vers, n'as-tu pas dit à ces gens le mot qui les aurait délivré ? Pourtant, l'énigme, ce n'était pas le premier venu qui pouvait la résoudre : il y fallait l'art du devin. ... Tandis que moi, je suis arrivé, ne sachant rien, moi, Œdipe, et j'ai fermé la bouche à la sphinge par ma seule intelligence, sans avoir rien appris des oiseaux.

Réplique d'Œdipe au devin Tiresias in *Œdipe Roi*. Sophocle. Trad. Maurice Véricel. Univers des lettres. Bordas.

Résumé

L'énigme du sphinx est le prototype des situations où le sujet interprétant est confronté à son incapacité à saisir le sens d'un message obscur. Face à cette impossibilité à comprendre d'emblée le texte auquel il se heurte, le « solveur d'énigme » subit d'abord la blessure de sa lenteur cognitive avant de retrouver, poussé par sa libido sciendi, la relation oblique que les mots entretiennent avec l'ordre du monde tel que son encyclopédie et sa mémoire le représentent. Les pages qui suivent proposent un parcours interprétatif qui pourrait être celui la méthode heuristique du personnage d'Œdipe au moment où, pour vaincre le sphinx, il cesse de se soumettre au muthos en le contraignant à emprunter les catégories du logos et de la démonstration logique.

1. Introduction

L'énigme a une histoire. Histoire elle-même amarrée à un mythe : celui d'Œdipe, qui libère la ville de Thèbes en répondant au sphinx¹. La question qui lui est posée est désormais inscrite au patrimoine mondial de la devinette :

« Quel est l'être qui marche tantôt à deux pattes, tantôt à trois pattes, tantôt à quatre pattes et qui, contrairement à la loi générale, est le plus faible quand il a le plus de pattes ? »

Quant à la réponse : « l'homme », elle est désormais considérée comme une évidence. Mais, au-delà de cette évidence, et de cette trivialité, se cachent un parcours interprétatif et une réalité complexe qui disent, implicitement, le triomphe du calcul du sens sur la révélation infuse. Le sphinx est un personnage venu du royaume de l'obscurité, du chaos, du monde inhumain, là où tout est confondu, indifférencié. La réponse est lumineuse, c'est-à-dire que soudain, comme dans les anamorphoses, on comprend les choses d'un certain point de

¹ Parfois appelé « la sphinge » ou « la chienne », d'où la traduction de Véricel (voir épigramme).

vue, selon un regard renouvelé et en retrouvant l'accord du monde. Elle met de l'ordre dans ce qui était obscur et s'impose à notre entendement. La manière dont on passe de l'un à l'autre, du chaos à la clarté, de la question obscure à la réponse évidente, suppose une heuristique qui n'est pas montrée clairement dans le mythe. Or, il faut bien qu'il y ait là quelque chose comme un « organon », une technique de recherche, un chemin à parcourir (« méthode ») pour expliquer la trouvaille. Au devin Tiresias (voir épigramme), le problème ne se pose pas : il est censé savoir, mais ne dit rien, et même ne peut rien dire, pour des raisons qui font un peu écho à la malédiction de Cassandre. Dans cet article, après avoir tenté de définir le phénomène de l'énigme comme défi angoissant à saisir le sens, nous proposons une interprétation de ce qui peut être sous-jacent à la « seule intelligence » dont Œdipe se réclame. Le développement exhaustif d'une telle étude excéderait la tâche qu'on peut raisonnablement s'assigner dans les pages qui suivent ; il s'agit cependant de replacer l'énigme dans son contexte culturel et d'analyser le processus interprétatif et de proposer une hypothèse d'heuristique mise en œuvre par le personnage du mythe.

2. L'énigme : un texte menaçant et qui « ne nous dit rien »

D'une manière générale, ce qui caractérise l'énigme, c'est que, soudain, une voix vient parasiter l'individu et que « ça ne lui dit rien », ou « ne lui rappelle rien ». Pratiquement, tout en ressemblant à du langage courant, « ça ne lui parle pas ». Dans le même temps ce « rien » est insupportable, parce que la parole obscure (« aïnigma ») dit aussi être une cohérence dont l'individu est privé, qu'il ressent comme une difficulté particulière, une déperdition agaçante de ses capacités cognitives, un manque à être, un défi à relever. L'énigme n'ouvre d'abord à aucune des possibilités de ce que Montaigne appelle « *entrer en conférence* ». On reformule aujourd'hui l'expression en disant : « entrer en dialogue », le dialogue étant le fond même de l'interprétation. Reste à savoir avec qui, ou plutôt avec quoi, se réalise cette traversée de la langue et ce rapport à l'autre.

2.1. Le substitut d'un combat

C'est que l'énigme est elle-même énigmatique. Celle du sphinx comme les autres. D'où vient-elle ? De quelle tradition relève-t-elle ? Comment finit-on par la résoudre. On sait que le jeu des énigmes était répandu dans tout le Moyen-Orient et tout particulièrement à la cour de Pharaons. Les Sibylles, comme celle de Cumès, ou encore la Pythie de Delphes, ne s'exprimaient pas autrement. La Reine de Sabbath de son côté défie Salomon d'interpréter des paroles obscures pour éprouver son intelligence. En Égypte, les affrontements entre deux armées étaient parfois remplacés par une sorte de « combat des chefs » où les généraux se défiaient à coup d'énigmes. Celui qui ne trouvait pas la solution était

condamné à avoir la langue coupée ; on la donnait à manger à un chien ou à un chat et il perdait ainsi, littéralement, « la parole », mutilé par là où le sujet habituellement se constitue socialement. L'expression « donner sa langue au chat » est restée. Ce qui caractérise donc parfois l'énigme, et qui est vrai de l'énigme du Sphinx comme de bien d'autres, c'est que la vie y est en jeu. Les jeunes Thébains, avant Œdipe, ont péri dévorés. Après lui, c'est le Sphinx qui est précipité dans le néant, tandis que le meurtrier de Laïos et futur père d'Antigone passe du statut d'inconnu à celui de héros qui restaure la cité. Enigmatique, surtout, est le fait qu'en dehors de l'explication qu'en donne le personnage lui-même lorsqu'il s'adresse au devin Tiresias (« ...*sans avoir rien appris des oiseaux*. »), la manière dont il trouve la solution n'est jamais donnée. Sophocle en parle au passé, comme d'un exploit accompli. Il en va de même chez Euripide, chez Homère et dans les reprises du mythe que l'on rencontre à travers les époques (cf. Gide ou Cocteau par exemple). On ne voit jamais Œdipe élaborer sa réponse, laquelle est réputée arriver tout d'un bloc et entièrement pertinente, achevée en même temps qu'énoncée. Ce qu'il faut cependant retenir, en relisant Sophocle, c'est qu'Œdipe identifie la force qu'il a eu de « *fermer la bouche à la Sphinge* » à son « *intelligence* ». Etymologiquement l'intelligence est une « capacité à faire des liens » pour produire du sens. Or, et parce que cette étymologie là est complexe et plurielle, l'intelligence partage également ses racines avec une capacité à se conduire « religieusement ». Non qu'il s'agisse de croyance ou de dogme, mais de participation à un système de valeurs et de références par quoi on s'inscrit dans le monde, lequel est le contraire du chaos, du moins en principe.

2.2. Contre la tératologie

Oedipe est d'abord menacé par une proposition formulée sur le mode interrogatif qui recouvre une assertion. Id est : « *il existe un être qui possède tantôt...tantôt...* ». Elle dit une réalité étrangère au sujet, réalité qui le rend étranger à lui-même, voire le confronte à une « inquiétante étrangeté ». L'encyclopédie qui est la sienne lui dit par ailleurs qu'en dehors des cas tératologiques, les animaux ont habituellement un nombre pair de pattes et qu'ils ne se métamorphosent pas sans cesse. Il n'y a qu'une alternative à cette situation de l'étrangement : ou cesser d'être (être dévoré), ou précipiter la chute du sphinx en rendant la question caduque, ce qui, dans le même temps nous permet d'accéder à ce surcroît d'être que nous confère la jouissance de l'entendement. Défié par l'énigme, Oedipe n'a d'autre recours que d'en affronter le texte, de le réapproprié à son encyclopédie par un retournement. De « séduit », détourné de son chemin habituel autant que fasciné ou angoissé, il doit se faire à son tour « séducteur » du sens. Le calcul de l'énigme (sur lequel l'antiquité ne s'était pas trompé, au point d'en avoir fait authentiquement le substitut d'un combat), est, du côté sémiotique la manifestation de la nécessité vitale où nous nous trouvons de ne pas rester séparé du monde par l'indécis, le non-sens.

Au point de départ, Oedipe est donc enlisé dans un enchevêtrement de mots qui, par l'horreur du non-sens génèrent l'angoisse de la rupture et de la « dés-intégration ». Il importe peu que le sphinx dise quelque chose qui, de son côté à lui, décrive un état du monde ou pas. Ce qu'il donne est une trace sonore dont Oedipe doit faire une texture, ou plus exactement un texte recouvrant un « vrai », qui, sans être une « vérité » doit permettre de se retrouver, de se reconstituer comme ego signifiant dans le miroir brouillé des mots que lui tend son adversaire. Or, ce qu'on demande aux miroirs c'est de nous donner la preuve qu'on se trouve bien là, au centre d'un territoire connu. On n'a pas souvent envie de les traverser. Se demander si l'on est là c'est aussi, comme le fait la marâtre de Blanche-neige, les interroger pour savoir si l'on est toujours en équation avec l'idée qu'on se fait de son propre rapport au monde. Narcissisme ou pas, plus le reflet est trouble, plus on cherche dans le miroir ce que l'on souhaite y voir, et il faut que cela fasse sens, même si c'est en ce même lieu que le désir est trahi.

2.3. L' autre comme moi-même

D'où l'appréhension et la haine nécessaire d'Oedipe face au sphinx, lequel joue d'abord à lui renvoyer le reflet de son incapacité. Le sphinx terrorise puisqu'il dit qu'il sera peut-être à jamais inaccessible, faisant sens à lui seul et sans nous. L'énigme, la devinette, est au départ plus forte que l'individu qui ne l'a pas encore résolue et qui mesure ainsi la perte de son image d'interprétant, le renvoie à une identité qu'il n'avait sans doute pas réellement. Le déchiffrement passe par cette phase d'angoisse où cesse d'être postulée comme évidente la similitude du « lecteur » à ce qu'il lit, où ce même lecteur n'est plus, comme le disait Baudelaire, le « semblable » ou le « frère » de l'auteur. Nous sommes face à l'énigme dans le manque à être comme dans le manque à savoir. La fracture ainsi provoquée ne redevient supportable que par le postulat de l'« intelligence », c'est-à-dire d'une réconciliation résultant d'un calcul compatible avec les normes du langage, retrouvant les liens avec les lieux communs, et, avant la lettre, quelque chose qui évoque les maximes conversationnelles de Grice. En l'absence de réponse, ou de la possibilité de la construire, le sujet est enlisé dans son incapacité à toute symbolisation, et ce, bien que l'énigme dise implicitement qu'elle n'est pas un paradoxe, pas une impasse logique, mais qu'au contraire, il faut, à travers son déchiffrement, retrouver la doxa.

Oedipe doit répondre à la béance créée par les paroles du sphinx en raison d'une sorte d'instinct de conservation qui l'engage lui-même et derrière lui toute la cité. Le combat est inévitable, mais, en l'espèce, les épées ne servent à rien. Il faut une impulsion à ne pas sombrer dans le précipice de l'absurde, dans lequel ce n'est pas tant le sphinx qui le pousserait mais où il se jetterait lui-même. Précipice de l'aphasie, de l'absence de toute nomination et gouffre de

l'impossibilité à trouver un « topos » qui permette le sens. La réponse « l'homme » le désigne donc lui-même comme seul lieu du possible de l'interprétation. « L'homme » n'est pas l'individu mais la récurrence des individus interprétants réconciliables avec les signes.

La clé et le "découvreur de la clé" ne font donc qu'un. Non pas l'un dans sa singularité mais dans sa démultiplication et sa capacité à retrouver le collectif. La dévoration qui menace se retourne alors en une « lecture absorbante ». Après tout, on « dévore » les livres sans trop bien savoir qui dévore quoi, mais en pensant que la lecture est "refuge" parce qu'elle est présence d'un autre. Oedipe, anéantissant le sphinx rencontre son double, son « *Je est un autre* » ou son « *Soi-même comme un autre*². » qui redevient « l'autre comme moi-même » par la prise en compte d'une chaîne qu'il pose comme signifiante pour la communauté parlante. Le « *deviens ce que tu es* » s'accroche au « *connais-toi toi-même* » parce que la cité grecque est là, fondatrice d'un ordre du monde (nomos) et de sa claire interprétation (logos).

Ainsi l'énigme ne cesse de dire la séparation et la nécessité d'une réconciliation avec l'univers sémiotique, par la réappropriation de l'ordre. Elle nous dit aussi l'impossibilité de ne pas retrouver le connu, en dépit des chaînes lexicales qui en démonte a priori l'apparence et mettent en crise le lexique voire la syntaxe. La question, dès lors, est aussi de savoir si le calcul interprétatif ne serait pas une fiction ou quelle part de fiction entre dans l'illusion du sens. Illusion qui est littéralement la capacité à entrer dans un jeu et par laquelle les éléments obscurs, d'abord déroutants, se réorganisent soudain en information.

3. Question d'heuristique

Que fait Oedipe que n'ont su faire les jeunes Thébains ? Rien d'autre, dirons-nous a priori, que d'accéder à la maîtrise du message en se l'appropriant. C'est à dire au fond en rompant avec « l'art du devin ». Ne parlant pas la langue des oiseaux, ni celle des dieux, Oedipe lit, il relie et calcule. Mieux, il doit construire un point de vue, se regarder dans l'énigme, s'y plonger. Cette immersion n'est pas une disparition du sujet dans le texte mais un transfert du texte obscur dans les paramètres de la langue. Oedipe organise son point de vue et ne se laisse pas ordonner par le mythe. En d'autres termes et pour retrouver la question sous-jacente de l'heuristique, Oedipe se trouve confronté au paradoxe de Ménon (Platon) : il ne lui est pas possible de trouver quoi que ce soit s'il ne sait pas d'abord ce qu'il cherche. Mais s'il sait quoi chercher sa recherche sera une anamnèse pour retrouver ce qu'il savait déjà, sans savoir qu'il le savait. La découverte ultime, participant de l'intelligence ne sera pas l'équivalent de la

² Paul Ricoeur

révélation qui soudain illumine le devin. Elle supposera au contraire quelque chose comme l'examen systématique et rationnel des données obscures. Sa quête est donc celle du passage de l'implicite à l'explicite. Il crée le sens en recolonisant par le « logos » le texte mythique. En réduisant l'étrangeté crypturale par son « intelligence », Oedipe réinvente le message.

3.1. Analyse formelle de l'énigme du sphinx

« *Quel est l'être qui tantôt marche à quatre pattes, tantôt sur deux pattes et tantôt sur trois?...* »

Le texte de l'énigme se pose sous la forme d'un énoncé qui lui même s'analyse en trois informations : « *quel est l'être qui* » suppose que « X »

- a) est un « être »
- b) existe assurément
- c) est unique dans son genre

et

au premier temps (« tantôt 1 ») il possède quatre pattes,
 au deuxième temps (« tantôt 2 ») il possède deux pattes
 au troisième temps (« tantôt 3 ») il possède trois pattes.

La double série du « tempo » et du « nombre de pattes » se présente sous la forme du tableau suivant :

| Tempo | Pattes |
|-------|--------|
| 1° | 4 |
| 2° | 2 |
| 3° | 3 |

Observons la corrélation entre les nombres ordinaux et les cardinaux : la série des ordinaux, sous « tempo », est en ordre; la série des cardinaux, sous « pattes » est en désordre. En d'autres termes : la stabilité du sème « tempo » est indiquée par celle, croissante, des ordinaux : 1° , 2° , 3° ; et la non stabilité du sème « pattes » est indiquée par le désordre décroissant et croissant des cardinaux : 4 > 2 < 3

La première partie de la série cardinale, 4 > 2, contient deux nombres, qui valent pour pattes, en relation avec l'ordre 1°, 2°. L'ordre cardinal est décroissant, tandis que le tempo est croissant. A la fin de la série, alors que le tempo reste inscrit dans une progression croissante (2° , 3°), le nombre de pattes ne continue pas à décroître, mais au contraire augmente (2 < 3). La confrontation des deux séries fait ainsi apparaître une brèche qui nous invite à déduire que 3, censé valoir lui aussi pour « pattes », n'est plus dans l'ordre au même titre que 4 et 2. La brèche, qui révèle le désordre, est la réversibilité dans la suite des cardinaux.

En partant de 4 nous passons d'abord à 2 (ordre), puis est donné le 3 qu'on attendrait, selon l'ordre décroissant des entiers naturels, normalement entre le 4 et le 2 (soit : 4, 3, 2).

Autrement dit, si l'on devait maintenir la série du sème « patte » pendant toute la série du tempo, le désordre décroissant /croissant des cardinaux ($4 > 2 < 3$) ferait de cet « être qui existe », un être irrationnel, ou, à tout le moins « anti-encyclopédique » et donc non existant, selon l'intelligence d'Oedipe.

3.2. La métaphore saisit le nomos

La troisième « patte » est donc une « patte non-patte ». Une « patte non-patte », est, soit contradictoire, soit un analogue de patte. Et c'est cet analogue qui ouvre le domaine de la métaphore, et, partant, de la clé de l'énigme. L'énoncé du sphinx ne pouvant tricher (c'est-à-dire « devant être vrai » et non contradictoire puisque l'énigme suppose une solution), la clé réside dans le transfert du raisonnement inductif par récurrence au raisonnement analogique. Le raisonnement analogique commande la métaphore et le changement de sème. Dans la série $4 > 2 < 3$, on est passé de « patte naturelle » à « analogue de patte » donc à « patte artificielle » : canne ou bâton. La question du sphinx se reformule donc en : « *Quel est l'animal qui peut contrarier l'ordre naturel des nombres ou l'ordre du monde au moyen d'un artefact ?* ». La réponse pointe au terme de ce calcul : l'homme ! « *La nature est pleine de merveilles. Mais l'homme est la merveille de la nature ... L'homme très intelligent...* », dit le choeur dans l'*Antigone* de Sophocle³.

La saisie métaphorique, d'une partie du message obscur, l'impossible mise en place selon l'« ordre du monde » de la troisième patte, est la porte ouverte à un élargissement des possibilités du pacte sémiotique. On quitte le muthos, (le récit qui, comme la parole pythique évoque par plans dièdres et pluralité des possibles) par la découverte de l'analogie, qui lui était consubstantielle. On se déplace en s'appuyant sur une « canne », qui est un artefact de patte, comme on se déplace avec des « pattes naturelles ». Dans le décodage de l'analogie, le terme substitué peut être assimilé, c'est-à-dire rendu semblable, « d'un certain point de vue » au mot qu'il remplace. Ainsi Oedipe replace-t-il les signes dans le champ du monde partagé, celui qui est mis en ordre selon le principe du nomos (la règle de la cité qui est conforme à celle du cosmos, lui-même contraire du chaos) et échappe au vertige de l'obscurité. Le texte de l'énigme, obscur et inquiétant au début, redevient par le calcul (intelligence) un lieu de sens socialisé, banal, et qui, dans le même temps perd de son pouvoir d'inquiéter. Il était, en tant qu'obscur, un stimulant pour l'intellect, déclencheur d'une *libido sciendi*, en même temps qu'un risque à assumer. Il est désormais intégré à la possibilité de vivre parce qu'on l'a domestiqué, qu'il est maintenant approprié définitivement à notre manière de voir : on ne peut poser la même énigme deux

³ Traduction Bonnard, TNP Jean Vilar.

fois à celui qui l'a décodée ! La force troublante de l'énigme a disparu à jamais, en même temps que le sphinx.

Comprendre un message obscur, ici une énigme, peut s'analyser comme apprendre à retrouver un nom et une situation du monde où l'on puisse obtenir l'expérience de l'objet auquel ce nom se réfère. Il n'y a dans l'expérience de l'énigme d'autre révélation prométhéenne que celle de l'équilibre de la langue et de la nomination, c'est-à-dire de l'expérience partagée. La nomination analytique est consubstantielle à la convention de l'énigme. Et l'énigme s'avère être le modèle de la relation entre le monde et le sujet. Elle dit qu'il ne saurait y avoir d'ambiguïté dans les classifications. A la question « *qu'est-ce que c'est que cet objet?* », on répond par un nom. A l'autre pôle correspond la question « *que signifie ce nom ?* », à laquelle on répond par une définition. Aussi longtemps que la relation entre les deux n'est pas faite, l'un et l'autre sont obscurs. Au départ, Oedipe courait le risque de considérer le message du sphinx comme un instrument de saisie littérale d'un « réel » délirant (= sorti du sillon conventionnel). Il l'interprète au contraire comme un appel à sa conscience inductive, il y applique un raisonnement analogique qui lui permet de découvrir que l'ordre du monde et le signe, par déplacement (trope), se manifestent sans véritable « im-pré-vu ». Aucune autre réponse que celle-ci ne peut être conformes aux catégories du savoir, même si l'on considère le monde comme un continuum, c'est-à-dire le contraire d'un prédécoupé ontologique.

4. Conclusion

Oedipe a su se déplacer dans les plis du langage pour arrêter le vertige d'un sens qui pouvait être téatologique ou plus simplement n'être pas rationnel. Pour toute perception où la raison se retrouve, toute interprétation, il faut un calcul et une activité mémoire, au sens psychique, comme peut-être aussi au sens informatique. On peut ici s'appuyer sur Husserl : « *chaque état de conscience possède un horizon qui est à mettre en perspective avec d'autres états.*»⁴ La mémoire est ici « modèle », lieu des attendus linguistiques, et condition de la nomination et de l'interprétation. Si l'énigme ne pouvait être référée au monde, organisé et en ordre, c'est qu'il n'y aurait pas de modèle à y retrouver. La formule du sphinx s'apparenterait alors à une suite sans modèle, un chaos. Mais, « *dire qu'une suite n'a pas de modèle est une contradiction mathématique. On peut tout au plus dire qu'elle n'a aucun modèle susceptible d'être recherché. (...) Il est entendu que l'ordre et le chaos ne sont pas des vérités objectives, mais, comme tant d'autre choses, déterminés par le point de vue de l'observateur* »⁵.

⁴ Husserl 1953, *Méditations cartésiennes*, Vrin, Paris

⁵ G. Spencer Brown, 1977, *Probability an scientific inference*, cité par Watzlawick, *How real is real*, trad *La réalité de la réalité*, Seuil, Paris.

Einstein disait également à propos de la science « *C'est la théorie qui décide de ce que nous sommes en mesure d'observer* ». Ce que saisit Oedipe, à défaut d'être une théorie du langage ou de l'énigme, c'est que les signes peuvent commuter entre eux et qu'il n'y a pas de temps dans l'énigme, pas plus que dans la logique. Le sphinx est le nom d'une menace que nous faisons peser sur nous-mêmes, lorsque nous croyons qu'en restant enfermés dans un solipsisme interprétatif, nous trouverons quelque chose. Derrière l'énigme, il n'y a aucun indéterminé, il n'y a que la langue et le social.

Eléments bibliographiques

- Baltrusaitis, J., 1969, *Anamorphoses*, Perrin, Paris.
- Barthes, R., 1970, *S/Z*, Editions du Seuil, Paris.
- Benvéniste, E., 1966. Problèmes de linguistique générale I et II, Gallimard, Paris.
- Bernasconi, M., 1964, *Histoire des énigmes*, « Que sais-je? » N° 1087, P.U.F. , Paris
- BnF, 1998 A, *L'aventure des écritures : Ecriture et sens caché, Ecritures non déchiffrées, Ecritures micrographiques* , site web, <http://www.fr/web-bnf/pedagos/dossiecr/in-secre.htm>.
- Bougnoux, D. (textes rassemblés et présentés par), 1993, *Sciences de l'information et de la communication*, Coll. Textes essentiels Larousse, Paris.
- Deledalle G., 1979, *Théorie et pratique du signe, Introduction à la sémiotique de C.S. Peirce*, Payot, Paris.
- Eco, U., 1963, *L'œuvre ouverte*, Seuil, Paris.
- Eco, U., 1983, *Lector in fabula*, Grasset, Paris.
- Eco, U., 1988, *Le signe*, Labor, Bruxelles.
- Freud, S., 1983, *L'inquiétante étrangeté et autres essais*, Gallimard, Paris.
- Greimas, 1970, *Du sens*, Seuil, Paris.
- Grice, H.P., 1975, *Speech acts, logic and conversation*, trad. in *Communications* n°30, pp. 41-58, Galimard, Paris.
- Leguern, M., 1973, *Sémantique de la Métaphore et de la Métonymie*, Coll. "Langue et langage", Larousse Université, Larousse, Paris.
- Panofsky, E., 1973 *La perspective comme forme symbolique*, Minuit, Paris
- Peirce, C.,S., 1978, *Ecrits sur le signe*, Seuil, Paris.
- Rastier, F., 1987, *Sémantique interprétative*, Coll Formes Sémiotiques, PUF, Paris.
- Rastier, F., 1996, *Problématique du signe et du texte*, *Intellectica* 1996 / 2 pp 11-52.
- Sophocle, 1973, *Tragédies*, Gallimard, Paris.

Quelques textes entre l'énigme et la devinette

Nadine Lucas
GREYC CNRS Université de Caen
Nadine.Lucas@info.unicaen.fr

Résumé. Trois textes à ressort énigmatique sont présentés et comparés à un texte explicatif. Les textes à devinette montrent des procédés exagérés de cohésion locale et des procédés discrets de marquage du fil de l'histoire. La compréhension implique un changement de repère dans l'ordre du texte.

Mots-clés : Linguistique, textes, comparaison, énigme

Les textes sont généralement écrits de telle sorte que le sens émerge sans difficulté pour le lecteur. Mais bien sûr, certains textes proposent des défis, plus ou moins difficiles à relever (Rastier, 2001). L'énigme partage certains traits d'ellipse avec les histoires drôles, d'étrangeté avec le fantastique, mais elle est plus subtile (Vanderdorpe, 1998). Parmi les récits énigmatiques, les plus célèbres sont *La lettre volée* de Poe et *A/Z* de Balzac. Trois textes seront proposés ici, avec des degrés différents d'accès au sens. Une histoire à devinette, un conte drôle traduit, un récit court ancien illustreront le masquage ou l'absence des marques ordinaires de repérage, par rapport à une histoire sans surprise. Ils sont présentés dans l'ordre d'étrangeté croissante.

Exemple 1. Tertulia *La Recherche* 315 (1998) p. 110.

TERTULIA*

Jorge Wagensberg

Oui mais... que faisait le petit dans le grand?

* La tertulia est une conversation typiquement espagnole, autour d'un café

Juin 1966. Le père Daniel se promène, comme à son habitude, sur ce qui va bientôt devenir la route BR-216, à Santa Maria (Etat du Rio Grande do Sul, Brésil). Les excavateurs fouillent des terres qui ont sommeillé pendant des centaines de millions d'années et recrachent parfois des restes de la vie du moyen Trias (230-240 millions d'années). Aujourd'hui la chance lui sourit. Son cœur se met à battre la chamade: devant lui se trouve un rynchosaure, un *Scaphonyx fisheri*, reptile de plus de un mètre de long, entier, parfait. Mais le cœur du père Daniel subit une nouvelle et forte émotion : entre deux côtes ventrales, il y a quelque chose d'autre. C'est un squelette minuscule, sans tête et aux côtes fines comme des cheveux. Sur le chemin du retour une seule question le hante : *Que faisait le petit dans le grand?* Daniel Cargnin insiste : *Je ne suis pas un scientifique. Je ne fais que chercher, et si je trouve, alors, je protège...* A première vue, il ne semble y avoir que deux possibilités. Soit le rynchosaure a avalé le petit animal peu de temps avant d'être enseveli à tout jamais. Impossible, le rynchosaure est un animal rigoureusement herbivore ! Soit le petit était un fœtus sur le point de naître. Impossible, le rynchosaure est ovipare ! Mais il est vrai que l'on n'a jamais trouvé d'œufs de rynchosaure... tiens, tiens !

Octobre 1967. A la clôture d'un congrès en Argentine, quelqu'un montre la photo de l'étrange fossile. Le lendemain, les cinquante participants de la réunion (au grand complet), se rendent à l'Institut de Géoscience de Porto Alegre pour examiner la merveille. Les professeurs tournent lentement autour de la table. Pour conclure à l'unanimité : il manque le crâne. Et sans le crâne, on ne pourra guère avancer. Le père Daniel s'engage à remuer ciel et terre (surtout cette dernière, bien entendu) pour le retrouver. *Je vais m'y mettre, promet-il, mais dites-moi, que faisait le petit dans le grand?*

Janvier 1968. Le père Daniel arrive à neuf heures du matin à l'endroit exact de sa découverte et commence à peigner le terrain grain par grain. Inouï. Il rentre chez lui, transporté de joie. J'ai réussi ! *Mais que faisait le petit dans le grand?*

Octobre 1975. Daniel Cargnin reçoit un tirage à part du très prestigieux *Journal of Paleontology*, accompagné d'une affectueuse dédicace manuscrite des auteurs, J.F. Bonaparte et M.C. Barberena. C'est un passeport pour la gloire. On y trouve la description de l'exemplaire unique d'un petit reptile, une nouvelle espèce de carnivore qui a tout ce qu'il faut pour s'ériger en prémonition du concept "mammifère" : il s'agit du tout nouvellement baptisé *Theryoherpeton cargini*. Le père Daniel accepte avec reconnaissance et émotion tous les détails : *cargini*. Quel honneur inattendu ! Lui qui ne recherche pas, mais trouve et protège ! *Mais s'il vous plaît, que quelqu'un vienne à mon aide...*

Septembre 1998. Je suis en face du père Daniel. Les médecins lui ont interdit toute émotion, il a mal aux genoux et sa vue faiblit. Rien de tout cela ne l'a empêché de me conduire, avec une précision algébrique et pendant deux jours, sur le gisement triasique de ses amours. Nous prenons ensemble un petit verre d'excellent "cachaça" lorsqu'il me raconte quelque chose de très important pour lui. Cela s'est passé il y a quelques années, sur un chemin creux où il se promenait tranquillement. Soudain son odorat fut assailli, offensé, par la puanteur émanant du cadavre d'une vache. Il pressait le pas en se pinçant le nez lorsqu'un fort battement d'aile noire surgit, justement, de l'intérieur de la charogne et lui fit se dresser les cheveux sur la tête. Ce fut alors, un millième de seconde plus tard, que son visage s'éclaira d'une lumière qui n'était pas celle de l'après-midi.

Exemple 2. Un conte populaire japonais¹

Le renard

Alors, voici. C'est une histoire qui remonte à très très longtemps, voyez.

Dans la montagne vivait un renard mage qui pouvait se métamorphoser². De temps en temps, il venait au village, dans le coin des Trois pins, et oh là là, il en faisait des farces aux gens !

Un beau jour... Les paysans de la campagne en étaient venus à parler de ce renard, bon. Là-dessus, hein. Voilà-t-y pas qu'un samouraï en voyage passe par là :

"En voilà des histoires pour un malheureux renard ! Moi, je m'en vais vous l'attraper, et lui mettre une raclée !"

Ayant dit, il remonta ses manches, et frappa sur ses gros bras poilus pour bien montrer sa force. Puis, il s'en alla, en roulant des mécaniques.

Il arriva aux Trois pins.

"Il est pas encore là, il est pas encore là ?"

Il attendit. Finalement,

"Ah ! Le voilà !"

Du côté de la montagne, venait une jolie jeune fille, toute seule, qui marchait en se pavanant gentiment.

La jeune fille, en voyant le samouraï, vint à ses côtés.

"Je m'en vais au village, mais la nuit va tomber, et j'ai peur ! Monseigneur ! S'il vous plaît, accompagnez-moi au village !" dit-elle d'une jolie voix sucrée.

"Qu'est-ce que tu me chantes là ! Sale renard !"

Le samouraï, roulant des yeux furibonds, se mit à lui crier dessus.

"Alors toi, sans arrêt tu tarabustes les gens de ce village, hein !. Mais moi, je t'ai reconnu, tu ne vas pas m'échapper."

et il l'attrape.

Là-dessus, bon. La jeune fille fit un tour sur elle-même et pffft, disparut. Et, à peine disparue, voilà que cette fois, c'est un jeune marchand qui apparaît.

"Je viens d'Edo, mais un voyage solitaire est une bien triste chose. Monseigneur, ne voudriez-vous pas m'accompagner, je vous prie ?"

"Quoi ?"

Le samouraï, estomaqué, lui jeta un regard haineux.

"Toi aussi, tu dois être le renard de tout à l'heure. Tu as cru me rouler, mais je ne mange pas de ce pain-là !"

Le renard, se voyant découvert, se changea cette fois en vieillard. À nouveau découvert, il se changea en grand-mère. La grand-mère découverte, il se changea en bonze. Le bonze découvert, le renard mage, l'air tout déconfit, finit par se changer en simple renard.

¹ *Nihon no waraibanashi*, [Contes drôles japonais] compil. Kawasaki. Tokyo, Dôshinsha, 1968. pp. 219-223

² Dans le folklore japonais, les renards ont la réputation de pouvoir changer de forme, spécialement de prendre forme humaine.

Le samouraï s'amusa fort.

"Wa ouh ha, ha ha ! T'as fini par montrer ta vraie forme, hein ? Sale renard ! Je m'en vais t'ôter la vie"

et, écartant les bras, il le poursuivit.

Le renard, affolé, s'enfuit, s'enfuit à perdre haleine. Il fuit par ici, il fuit par là, et finalement il plonge dans un buisson.

Tout de go le samouraï plonge dans le buisson à sa suite et hop, il lui attrape la queue.

"Ho hisse, ho hisse !"

il tira.

Le renard hurlait sans arrêt, jappant :

"Ouap, au secours ! Ouap, au secours !"

Mais il eut beau crier et pleurer, rien n'y fit. Le samouraï tirait, à deux mains, de toutes ses forces.

Sbong !

Avec un grand bruit, la queue sortit. Toute seule, on ne voyait le renard nulle part.

"Wa ouh ha, ha ha ! La queue du renard, que voilà un beau trophée !"

Là-dessus, à ce moment-là...

"Hé bien, Monseigneur ? Que faites-vous donc ?"

Se retournant, le samouraï vit un paysan qui le regardait d'un air furieux.

Le paysan agrippa subitement ce que le samouraï tenait pendant dans les mains et lui dit, très en colère :

"Non mais, t'as pas fini de faire des idioties ? Ça veut dire quoi, ça, d'arracher les raiforts de nos champs ?!"

Exemple 3. Un récit ancien (1665)³

Histoire du maréchal de Bassompierre et de la belle lingère

Je partis un soir de la Cour et veux dire une aventure qui me survint qui, pour n'être pas de grande conséquence, est néanmoins extrêmement agréable.

Il y avait cinq ou six mois, que toutes les fois que je passais sur le petit pont (car en ce temps-là le pont neuf n'était point bâti) qu'une belle femme lingère à l'enseigne des deux Anges, me faisait de grandes révérences, et m'accompagnait de la vue tant qu'elle pouvait. Et comme j'eus pris garde à son action, je la regardais aussi et la saluais avec plus de soin. Il advint que lorsque j'arrivai de Fontainebleau à Paris, passant sur le petit pont, dès qu'elle m'aperçut venir, elle se mit sur l'entrée de sa boutique, et me dit comme je passais, Monsieur, je suis votre servante. Je lui rendis son salut, et me retournant de temps en temps, je vis qu'elle me suivait de la vue, aussi longtemps qu'elle pouvait.

³ Extrait des *Mémoires de M. de Bassompierre contenant l'histoire de sa vie et de ce qui s'est fait de plus remarquable à la Cour de France pendant quelques années*, Cologne, 1665. pp. 160-164. (Version originale, citée par Montandon, *Les formes brèves* Paris, 1993, pp. 110-111)

J'avais mené un de mes laquais en poste, pour le renvoyer le soir même, avec des lettres pour Entragues, et pour une autre dame de Fontainebleau. Je le fis lors descendre et donner son cheval au postillon, pour le mener, et l'envoyais dire à cette jeune femme, que voyant la curiosité qu'elle avait de me voir, et me saluer, si elle désirait une plus particulière vue, j'offrais de la voir là où elle voudrait. Elle dit à ce laquais que c'était la meilleure nouvelle que l'on lui eût su apporter, et qu'elle irait où je voudrais, pourvu que ce fut à condition de coucher entre deux draps avec moi. J'acceptai le parti, et dis à ce laquais s'il connaissait quelque lieu où la mener, il me dit qu'il connaissait une maquerelle nommée Noiret, chez qui il la mènerait, et que si je voulais qu'il portât des draps, matelas et couvertes de mon logis, il m'y apprêterait un bon lit. Je le trouvai bon, et le soir y allai et y trouvai une très-belle femme âgée de vingt ans, qui était coiffée de nuit, n'ayant qu'une très-fine chemise sur elle, et une petite jupe de revêche verte et des mules aux pieds avec un peignoir sur elle, elle me plut bien fort, et me voulant jouer avec elle je ne lui sus faire résoudre, si je ne me mettais dans le lit avec elle: ce que je fis, et elle s'y étant jetée en un instant, je m'y mis incontinent après, pouvant dire n'avoir jamais vu femme plus jolie, ni qui m'eût donné plus de plaisir pour une nuit: laquelle finie, je lui demandai si je ne la pourrais pas voir encore une autre fois, et que je ne partirais que dimanche, dont cette nuit-là avait été celle du jeudi au vendredi. Elle me répondit qu'elle le souhaitait plus ardemment que moi; mais qu'il lui était impossible si je ne demeurais tout dimanche, et que la nuit du dimanche au lundi elle me verrait. Et comme je lui en faisais difficulté, elle me dit, je crois que maintenant que vous êtes las de cette nuit passée, vous avez le dessein de partir dimanche, mais quand vous vous serez reposé, vous serez bien aise de demeurer un jour davantage, pour me voir une nuit.

Enfin je fus aisé à persuader, et lui dis que je lui donnerais cette journée pour la voir la nuit du même lieu. Alors elle me répartit, Monsieur, je sais bien que je suis en un bordel infâme, où je suis venue de bon cœur pour vous voir, de qui je suis si amoureuse que pour jouir de vous, je crois que je vous l'eusse permis au milieu de la rue, plutôt que de m'en passer. Or, une fois n'est pas coutume, et forcée d'une passion, on peut venir une fois dans le bordel, mais ce serait être garce publique d'y retourner une deuxième fois. Je n'ai jamais connu que mon mari et vous, ou que je meure misérable, et n'ai pas dessein d'en connaître jamais d'autre. Mais que ne ferait-on point pour une personne que l'on aime, et pour un Bassompierre. C'est pourquoi je suis venue au bordel, mais ç'a été avec un homme qui a rendu ce bordel honorable par sa présence. Si vous me voulez voir une autre fois, ce sera chez une de mes tantes, qui se tient en la rue du Bourg-labbé proche des Halles auprès de la rue aux Ours, à la troisième porte du côté de la rue Saint-Martin, je vous y attendrai depuis dix heures jusqu'à minuit, et plus tard encore, laisserai la porte ouverte. À l'entrée il y a une petite allée que vous passerez vite, car la chambre de ma tante y répond, et trouverez un degré qui vous mènera à ce second étage.

Je pris le parti, et ayant fait partir le reste de mon train j'attendis le dimanche, pour voir cette jeune femme: je vins à dix heures et trouvai la porte qu'elle m'avait marquée, et de la

lumière bien grande, non seulement au second étage, mais au troisième et au premier encore, mais la porte était fermée: je frappai pour avertir de ma venue, mais j'ouïs une voix d'homme, qui me demanda qui j'étais. Je m'en retournais à la rue aux Ours, et étant retourné pour la deuxième fois ayant trouvé la porte ouverte, j'entrai jusques à ce second étage, où je trouvai que cette lumière était la paille du lit que l'on brûlait, et deux corps nus étendus sur la table de la chambre. Alors je me retirai bien étonné, et en sortant je rencontrai des corbeaux⁴, qui me demandèrent ce que je cherchais, et moi pour les faire écarter mis l'épée à la main, et passai outre en m'en revenant à mon logis un peu ému de ce spectacle inopiné. Je bus trois ou quatre verres de vin pur, qui est un remède d'Allemagne contre la peste, et m'endormis pour aller en Lorraine le lendemain matin ; et quelque diligence que j'aie su faire depuis, pour apprendre ce qu'était devenue cette femme, je n'en ai jamais rien su. J'ai été même aux deux Angès, où elle logeait, m'enquérir qui elle était, mais les locataires de ce logis-là ne m'ont dit autre chose, sinon qu'ils ne savaient point, qui était l'ancien locataire. Je vous ai voulu dire cette aventure bien qu'elle soit de personne de peu, mais elle était si jolie, que je l'ai regrettée et eusse désiré pour beaucoup de la pouvoir revoir.

Il est clair que l'impression de texte énigmatique est due à diverses raisons. Dans la première histoire, l'explication attendue est éludée. Dans le conte, le même principe joue, mais la compréhension est moins immédiate, du fait de l'éloignement culturel. La troisième histoire n'est pas nécessairement écrite comme énigme, même si elle est allusive, elle le devient du fait de l'éloignement historique.

Pour mieux apprécier la construction des textes énigmatiques, nous présentons un article de journal, assez complexe, mais sans énigme.

Exemple 4. Un texte explicatif à deux niveaux *Le Monde* 19 avril 1999

LES LOUEURS D'AVIONS DÉTIENNENT LE TIERS DE LA FLOTTE MONDIALE.

Par FRANCOIS BOSTNAVARON.

POURQUOI DIABLE le nouvel Airbus A-319 d'Air France, millième exemplaire de la famille A-320, fraîchement sorti de son atelier d'assemblage de Hambourg, dans le nord de l'Allemagne, est-il aller croiser dans l'espace aérien britannique pour rejoindre Toulouse, ce qui n'est pas a priori la route la plus courte ? Tout simplement parce que la compagnie aérienne nationale n'est pas propriétaire de cet Airbus et que la syndication bancaire qui a monté le plan de financement pour le loueur est de nationalité britannique. Pour des raisons fiscales, l'avion doit, au moins une fois dans sa vie, passer dans l'espace aérien britannique.

⁴ Corbeaux : croque-morts

Les sociétés de leasing sont devenues des acteurs incontournables du marché aéronautique. C'est grâce à la plus grande d'entre elles qu'Air France vient d'acquérir son nouvel appareil: International Lease Finance Corporation (ILFC), filiale de l'assureur américain AIG, est spécialisée dans le financement d'avions. Le portefeuille d'ILFC est valorisé à quelque 19 milliards de dollars (17,5 milliards d'euros) et représente environ quatre cents appareils, répartis dans plus de cent vingt compagnies aériennes.

ILFC a réalisé la première opération de leasing de l'histoire de l'aviation commerciale en 1973 en louant un DC8-51 à AeroMexico, avant de devenir le premier bailleur d'avions à offrir tous types d'appareils, Airbus ou Boeing. Aujourd'hui, ILFC est devenu le plus gros client d'Airbus Industrie, avec un total de deux cent quatre-vingt-quatre appareils. En novembre 1998, le loueur s'est prononcé pour le futur A-318, le plus petit appareil de la famille A-320, qu'il compte placer dans les compagnies du monde entier.

Le rôle du loueur est devenu essentiel pour la gestion des flottes des compagnies aériennes. Sur les deux cent huit appareils qui volent sous le pavillon d'Air France, la moitié est en pleine propriété, un quart est en crédit-bail et le dernier quart est en location pour des périodes allant de deux à sept ans, parfois plus. Tous les acteurs y trouvent leur intérêt. Les compagnies n'ont pas à inscrire l'avion dans leurs comptes, elles peuvent en changer plus fréquemment, et, en cas de surcapacités, s'en séparer plus facilement. Les avions en pleine propriété sont difficiles à revendre à bon prix, surtout si d'autres compagnies font de même en raison de la conjoncture.

Le constructeur y trouve aussi son compte, car le loueur est généralement son premier client. Air France, par exemple, n'aurait pas pu obtenir le nouvel Airbus si ILFC ne s'en était porté acquéreur bien avant lui: la compagnie a pris conscience de son besoin alors que l'A-319 n'était plus disponible à la date voulue.

UNE ACTIVITÉ LUCRATIVE

Quant au loueur lui-même, son métier est proche de celui d'un spéculateur boursier: il parie sur le succès futur, à deux ou trois ans, d'un nouveau modèle d'avion. Une activité lucrative: ILFC a réalisé un résultat avant impôts de 562 millions de dollars en 1998. Ce métier est appelé à se développer. Les loueurs, qui ont fait irruption sur le marché il y a vingt-cinq ans, détiennent en moyenne un tiers des avions, toutes flottes confondues. Sans eux, nombre de compagnies, comme Air Liberté, n'auraient pas pu prendre l'air.

À l'image des grandes chaînes hôtelières, qui estiment qu'il n'est plus nécessaire de posséder les murs d'un établissement pour l'exploiter, les compagnies s'interrogent désormais sur leur métier. Celui de transporteur aérien consiste-t-il à être propriétaire de sa flotte et à l'entretenir, ou à faire voyager ses clients, dans les meilleures conditions possibles, au meilleur prix et dans les meilleurs délais ?

(c) Le Monde, 1999.

L'explication du rôle des loueurs d'avions s'appuie sur une ouverture classique. Le premier paragraphe offre une vraie réponse à une interrogation rhétorique. Cette ouverture anecdotique sert de ressort à une explication

économique plus ardue et plus générale, qui aboutit à un jugement critique souligné par l'intertitre, et à une question de fond en fin d'article, en écho de la première. On note dans ce texte une fin marquée par deux comparaisons explicites en parallèle (*loueurs* comme *spéculateur boursier* et *compagnies* comme *chaînes hôtelières*). Le titre met en facteur une explication (les cinq premiers paragraphes) et une réflexion (les deux derniers). La relation au titre de la deuxième partie réflexive est marquée par une transition *Quant au loueur lui-même*.

Au contraire, dans les textes énigmatiques, la régularité des mises en rapport entre début et fin est brouillée. Tous ces textes utilisent le principe de la temporisation et présentent une fausse fin ou une fin décalée. Les deux premiers textes amusants présentent une trop grande symétrie gauche droite, une trame fortement marquée au niveau local, qui fait oublier le fil du récit. Chaque paragraphe dans la *tertulia* est répétitivement borné par une date, à laquelle répond une interrogation en leitmotiv. Le titre appelle une explication logique, mais le masquage des débuts de paragraphe par la litanie des dates crée une parodie de récit chronologique. Deux pistes explicatives fournissent certes une alternative, pour la solution de l'énigme, mais l'alternative enfouie dans le premier paragraphe est étouffée par des réponses négatives. Reste une objection, partiellement négative elle aussi, associée à une forme exclamative *tiens tiens !*. Le texte annoté est évidé.

Oui mais... que faisait le petit dans le grand?

§1 Juin 1966. Le père Daniel se promène, comme à son habitude, sur ce qui va bientôt devenir la route BR-216, à Santa Maria (Etat du Rio Grande do Sul, Brésil). [...] Sur le chemin du retour une seule question le hante : *Que faisait le petit dans le grand?* Daniel Cargnin insiste : *Je ne suis pas un scientifique. Je ne fais que chercher, et si je trouve, alors, je protège...* A première vue, il ne semble y avoir que deux possibilités. Soit le rynchosaure a avalé le petit animal peu de temps avant d'être enseveli à tout jamais. Impossible, le rynchosaure est un animal rigoureusement herbivore ! Soit le petit était un fœtus sur le point de naître. Impossible, le rynchosaure est ovipare ! **Mais il est vrai que l'on n'a jamais trouvé d'œufs de rynchosaure... tiens, tiens !**

§3 Janvier 1968. Le père Daniel arrive à neuf heures du matin à l'endroit exact de sa découverte et commence à peigner le terrain grain par grain. [...] **J'ai réussi !** *Mais que faisait le petit dans le grand ?*

§4 Octobre 1975. Daniel Cargnin reçoit un tirage à part du très prestigieux *Journal of Paleontology*, accompagné d'une affectueuse dédicace manuscrite des auteurs, J.F. Bonaparte et M.C. Barberena. [...] **Quel honneur inattendu ! Lui qui ne recherche pas, mais trouve et protège ! Mais s'il vous plaît, que quelqu'un vienne à mon aide...**

§5 Septembre 1998. **Je suis** en face du père Daniel. Les médecins lui ont interdit toute émotion, il a mal aux genoux et sa vue faiblit. Rien de tout cela ne l'a empêché de me conduire, avec une précision algébrique et pendant deux jours, sur le gisement triasique de ses amours. Nous prenons ensemble un petit verre d'excellent "cachaça" lorsqu'il me raconte quelque chose de très important pour lui. Cela s'est passé il y a quelques années, sur un chemin creux où il se promenait tranquillement. Soudain son odorat fut assailli, offensé, par la puanteur émanant du cadavre d'une vache. Il pressait le pas en se pinçant le nez lorsqu'un fort battement d'aile noire surgit, justement, de l'intérieur de la charogne et lui fit se dresser les cheveux sur la tête. **Ce fut alors**, un millième de seconde plus tard, que son visage s'éclaira d'une lumière qui n'était pas celle de l'après-midi.

La fausse conclusion est marquée par des phrases exclamatives en presque fin des §3 et 4. Elles font écho ici à l'exclamation finale du §1. Cependant, l'adjonction des relances (interrogative ou à point de suspension) contredit le signal de fin. Dans le dernier paragraphe, un nouveau personnage apparaît, l'auteur, et l'on est alors conduit à ré-interpréter toutes les fins de paragraphes du récit précédent comme indices de *monologue* qui forment le véritable trait oppositif avec le dernier paragraphe de *dialogue*. L'affirmation et la négation s'entrecroisent dans ce paragraphe, mais la dernière phrase est partiellement négative, en écho de celle du paragraphe d'ouverture. On change donc de critère et on abandonne un critère saillant (l'exclamation) contre un critère minime (une négation partielle) pour établir une relation de concordance globale. La mise en scène, appuyée par le passé simple, permet de décrypter la comparaison finale, puisque ici comparaison vaut raison.

Dans le conte, la compréhension est moins immédiate, car dans notre folklore les renards ne sont pas des animaux magiques, tout au plus rusés. D'autre part, la traduction fait perdre une partie des indices. Les ruptures de ton

et de politesse sont moins marqués en français, ils sont moins exploités aussi dans le discours. Que le samouraï soit décrit au début de manière désinvolte comme un matamore prépare la chute, la répartie du paysan, qui n'est pas un paysan, parce qu'il s'adresse au samouraï comme à un enfant. Cependant, le masquage de progression par instauration d'une symétrie gauche droite exagérée est un procédé exactement comparable à celui de l'énigme des dinosaures. Ce sont ici des onomatopées et des ponctuations de récit (*voyez, bon, là-dessus*) qui brouillent la perception.

Dans ce conte, l'exclamation ponctue la majeure partie du récit. L'interrogation, qui normalement précède, est reportée en fin de discours. La tirade de fin cumule la négation et l'interrogation. La cohérence globale est marquée discrètement par des formes temporelles, *il y a très très longtemps, un beau jour* et *à ce moment-là*. Le conte fini, il reste une contradiction à résoudre. En effet, le titre désigne un renard ordinaire, un « renard des champs », or les contes parlent de créatures surnaturelles et le début du conte met effectivement en scène un « renard mage ». L'ordre d'apparition des personnages, associé aux marques temporelles, est un indice. Les paysans sont présentés juste après le renard, de façon discrète, quoique redondante comme *paysans de la campagne*, avec statut de personnage secondaire. Ils disparaissent. Mais au final les habitants de la campagne n'ont-ils pas le beau rôle ?

Les données qui nous manquent pour comprendre d'emblée la troisième histoire sont les connaissances historiques, la raison de la mise à feu des pailles par exemple (pour désinfection). Un historien pourrait soupçonner dès cet instant la peste. La syntaxe ancienne (l'antéposition des pronoms) contribue aussi au dépaysement sur le plan linguistique. Pourtant, il y a aussi un désir de l'auteur de temporiser, faire attendre la conclusion de la deuxième visite en donnant un luxe de détails sur le chemin à suivre pour rejoindre la lingère, par exemple. Les marques saillantes de bornage du début et de la fin sont des

déclarations de récit, associées à la concession *veux dire une aventure, qui pour être de peu de conséquence...* ; *voulu dire cette aventure, quoique ...* La concession finale est double *quoique qu'elle soit de personne de peu, mais elle était si jolie*. Les « mais » ponctuent en effet les temps forts du récit. Le bornage n'est pas réellement en décalage ici, l'impression n'est pas celle d'un jeu. De même que, dans l'exemple 4, le premier paragraphe est un prétexte pour introduire une autre explication, de même, dans l'histoire de Bassompierre, le premier récit de rencontre apparaît comme un prétexte introduisant cette fois un mystère sur le sort de la femme. On comprend a posteriori qu'il y a contradiction, un début léger « une aventure ... extrêmement agréable », annonce fallacieuse puisque l'aventure se révèle tragique, source d'inquiétude et source de regret.

L'insistance sur la conduite facile de l'aventure dans le début du texte apparaît comme une fausse piste, elle est contredite dans la seconde partie par le récit des conditions posées par la lingère, le personnage actif. La raison n'en serait-elle pas la contradiction des qualifications, de ce qui est *répété* sur l'isotopie du moindre (une aventure « de peu de conséquence » « un peu ému » « personne de peu ») et ce qui l'infirme, l'isotopie *déclinée* du « plus » : « elle me plut bien fort » « pouvant dire n'avoir jamais vu femme plus jolie ni qui m'eût donné plus de plaisir » « [j']eusse désiré pour beaucoup de la pouvoir revoir » ?

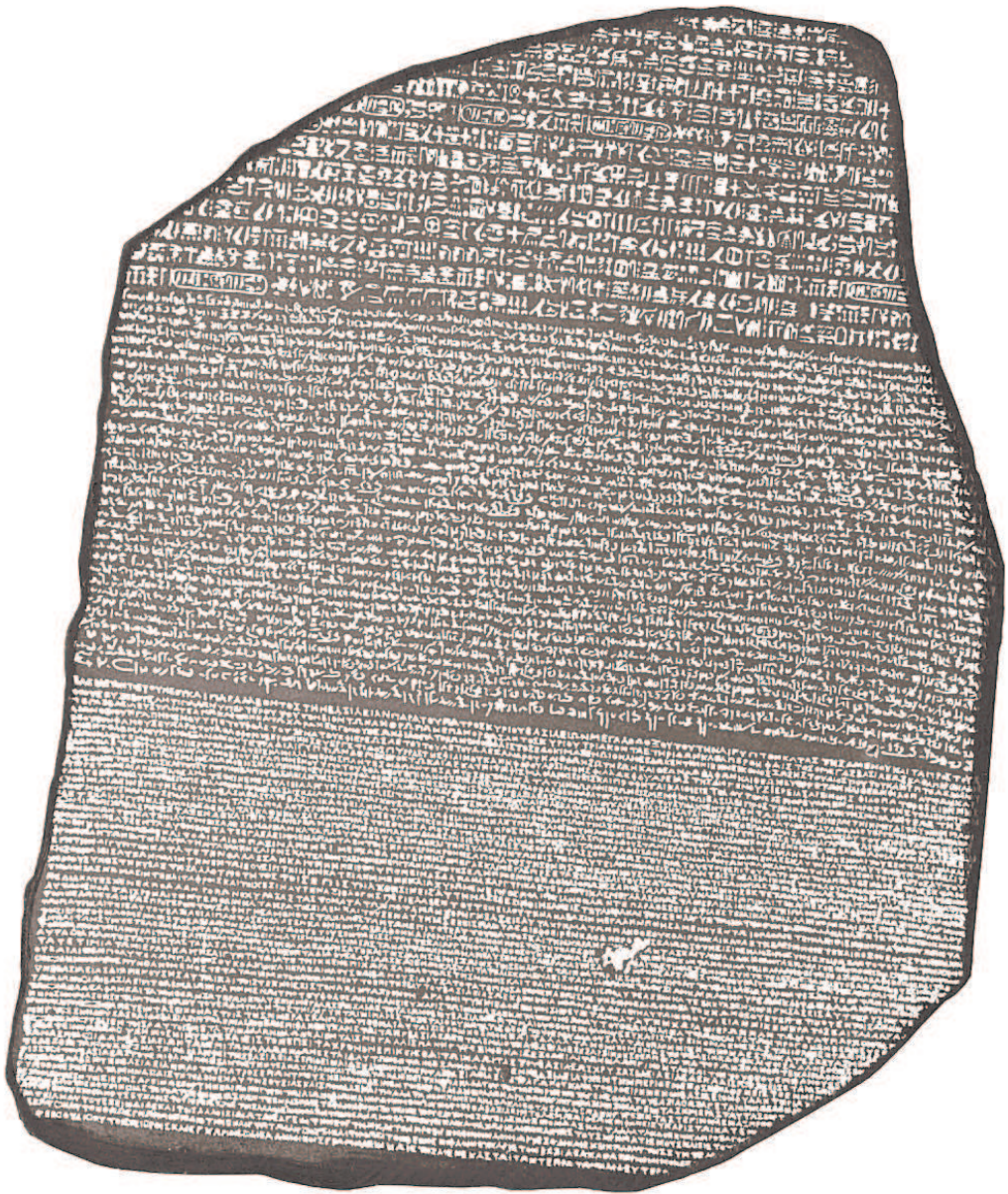
Dans les textes énigmatiques, le lecteur doit pouvoir passer d'une analyse locale à une analyse globale. Les marques en écho sont discrètes (des négations partielles, des marques temporelles, des concessions partielles). Elles sont en fonction déterminante, localement subordonnées, mais elles permettent de passer des épisodes à la disposition générale. Elles changent l'interprétation, comme des petites unités en retenue dans une addition peuvent faire changer l'ordre de grandeur du résultat.

Dans les contes où la servante devient reine, la logique du récit est fondée sur le retournement explicite des rôles (Propp, 1958). Dans l'énigme au contraire, il n'y a pas de retournement. Chacun reste à sa place. Les petits restent petits. Le sens naît plus subtilement de l'allusion, du changement de polarité dans les relations, du changement de repère. Et si l'intelligence triomphe, c'est que le lecteur y a mis aussi du sien.

Références bibliographiques

- MONTANDON, A. (1993) *Les formes brèves* Paris, Hachette.
PROPP, V. (1958) *Morphology of the Folktale*, La Hague, Mouton.
RASTIER, F. (2001). Arts et sciences du texte. Paris, Puf.
VANDENDORPE, C. (1998) La lecture de l'énigme *Alsic* Vol.1 n°2

http://alsic.u-strasbg.fr/Num2/vanden/alsic_n02-rec2.htm



Explorer et Appréhender le web

Fabien Pfaënder, Mathieu Jacomy
Université de Technologie de Compiègne
Boîte Postale 60319
60603 Compiègne cedex, FRANCE
+33-3-44-23-52-48
{fpaende,jacomyma}@utc.fr

Introduction

Les systèmes complexes sont des systèmes dont nous n'arrivons pas à nous saisir. La complexité qui les caractérise naît de l'impossibilité à découvrir l'organisation ou la structure de ces systèmes et d'en faire un objet de pensée. Ils se composent en effet d'un grand nombre de données hétérogènes, dynamiques ou non, liées entre elles sans que l'on ne sache comment. C'est de la manière dont ces données sont liées que l'on peut faire naître des connaissances. Comprendre les systèmes complexes et en extraire l'essence est d'autant plus important que ces systèmes sont de plus en plus nombreux aujourd'hui. Deux raisons principales à cela : d'une part, l'utilisation du numérique permet d'encoder toutes les données quelles qu'elles soient et de les manipuler formellement, indépendamment de leur signification. D'autre part, la démarche scientifique pousse à considérer le maximum de données afin de "coller" au mieux à la réalité du phénomène à observer. Toutes les données contextuelles possibles doivent donc être prises en compte pour expliciter et modéliser le phénomène. Tout le problème consiste à savoir quelles données sont pertinentes et interviennent dans le phénomène et lesquelles ne le sont pas. Aussi, et pour être sûr de ne pas en omettre, il est plus prudent d'en intégrer un maximum au système qui du même coup voit sa complexité augmenter. Le problème de la circonscription du phénomène n'est pas nouveau et les mésopotamiens se posaient déjà la question [Bottero 1987] de la clôture d'un phénomène en considérant un maximum de possibilités dans leur divination déductive, y compris les plus improbables. Ils prévoyaient par exemple le cas où un mouton aurait 2 foies, mais aussi 3, 4 et ainsi de suite jusqu'à 7 et les significations associées. De la même façon, la théorie du chaos pose pour hypothèse qu'un mouvement d'aile de papillon d'un côté de la planète aura des conséquences désastreuses de l'autre côté [Lorenz 1972] et trouve aujourd'hui une signification toute particulière. On voit bien ici que la tendance consiste à prendre en compte un maximum de données apparentées dans un modèle très large et ouvert. Cela prend de nos jours une ampleur sans précédent avec le besoin impérieux de mieux comprendre le monde qui nous entoure, indispensable pour prévoir notre futur proche.

On trouve donc des systèmes complexes dans tous les domaines. Il en est cependant un, transdisciplinaire et immédiatement accessible, dont les enjeux industriels et scientifiques sont particulièrement importants : le web. Ce dernier est en effet un système complexe très dynamique dont l'explicitation permettrait de tirer efficacement des connaissances pertinentes tant du point de vue documentaire que de l'organisation de ces documents. Le web est une construction qui met en présence des documents hypertextuels liés entre eux. Les documents forment des pages qui, suivant une même unité sémiotique, forment une entité que l'on appelle site Internet (il existe d'autres définitions du site web et cette notion est source de nombreux débats). La distribution de ces sites et documents liés entre eux forme un réseau complexe dont la topologie (distribution des liens) et la distribution du sens suivent des règles qui nous sont, dans une large part, inconnues. La diversité des formats sémiotiques des pages et la liberté de création des liens contribuent à rendre le système hétérogène,

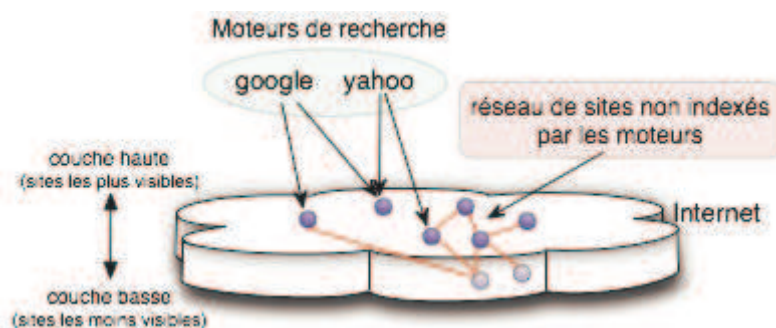


Illustration 1: Organisation du web

augmentant d'autant la difficulté de mettre à jour son organisation globale. Des structures doivent pourtant exister au sein de ce système complexe et les mettre à jour et les manifester est indispensable pour arriver à appréhender le web dans son ensemble et en faire naître de la connaissance. Bien sûr il existe déjà des outils comme les moteurs de recherche (Google™ et Yahoo™ en sont les représentants les plus célèbres) qui permettent de plonger dans la structure de manière ponctuelle mais ils ne permettent pas pour autant d'en rendre compte. Ils sont comme des satellites donnant un accès immédiat aux documents les plus visibles issus d'une requête avec une pertinence relative. Les navigateurs Internet (Firefox, Internet Explorer, etc.) permettent eux aussi un accès aux documents et une très faible spatialisation en utilisant les boutons "avance" et "recul" et plusieurs fenêtres ou onglets. Cela reste très limité et ne permet pas de se construire une représentation [Ghitalla 2003] efficace de la structure de telle ou telle localité d'Internet. Tout au plus cela autorise la création d'un micro-espace mais il est trop immersif pour être efficace. Car c'est là le principal enjeu ; proposer un outil qui explore la structure hypertextuelle et tire l'essence avant d'en rendre compte. Le résultat ne sera alors plus une liste de résultats décorrélés incomplète ou un empilement de fenêtres mais une présentation synoptique, un guide, qui rend compte de la structure dévoilée d'Internet et dont le résultat est par conséquent plus pertinent tant du point de vue objectif (les documents sont

pertinents) que subjectif (le navigant a prise sur ses résultats et leur organisation).

Afin de construire cet outil, nous proposons une méthode d'exploration et de préhension de l'organisation et de la structure du web qui peut, sous certaines conditions, être mobilisée pour explorer n'importe quel système complexe. Dans un premier temps, nous présenterons la question de la prise, de la trace et de l'appréhension, base théorique qui sous-tend l'ensemble de la démarche. Dans un second temps, nous présenterons les trois étapes de la méthode que sont la récolte de données, leur exploration et enfin leur synthèse pour transmettre les connaissances acquises sur le système.

Appréhender

Que signifie comprendre un système complexe comme Internet ? Mettre à jour sa structure est un pas mais ce n'est pas suffisant. Cela permet d'outiller une exploration vers des parties du système mais cela ne signifie pas pour autant que l'on est en mesure d'en apprécier l'ensemble. La connaissance du système n'est donc que partielle. Comprendre un système complexe signifie pouvoir s'en saisir comme d'un objet, pouvoir l'appréhender. Le système une fois saisi offre un espace de perception dont on connaît les lois [Lenay 2005]. On sait s'y déplacer et l'on sait comment les éléments qui le composent sont agencés. Il est alors aisé de l'utiliser et d'en tirer les informations que l'on souhaite. Cependant atteindre ce but est éminemment difficile, et ce dans la mesure où le système complexe est ce qu'il est justement parce qu'il n'offre pas de prise directe. Il faut donc trouver la façon d'en faire un objet à percevoir.

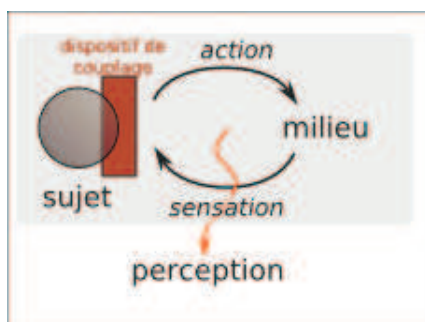


Illustration 2: Boucle sensorimotrice

Pour cela, nous nous plaçons dans le contexte d'une théorie sensorimotrice de la perception dans laquelle l'action occupe une place prépondérante. La perception n'est pas passive. Il ne suffit pas simplement d'ouvrir les yeux pour percevoir ce qui se trouve en face de nous. Au contraire, percevoir est une activité basée sur un aller-retour entre un sujet et le monde qui l'entoure. Le sujet agit dans le monde sur le milieu et il obtient en retour des sensations. Lorsque que les actions conduisent toujours aux mêmes sensations, la boucle se stabilise en un invariant sensorimoteur (une loi de contingence sensorimotrice). Cette régularité dans la boucle est la condition à laquelle il est possible de percevoir [O'Regan 2001]. De ce fait, percevoir le web revient à trouver des régularités dans ce milieu par un jeu d'actions et de sensations. Dans la mesure où il n'est pas possible d'agir directement sur le milieu, le sujet percevant est couplé avec un dispositif qui

donne un pouvoir d'action spécifique. Dans ce cas, les actions sur le milieu sont effectuées via le dispositif qui devient une prothèse. Dans le cas d'un système complexe comme le web, le dispositif proposé est une interface interactive qui propose un jeu d'actions desquelles naîtra une perception du système.

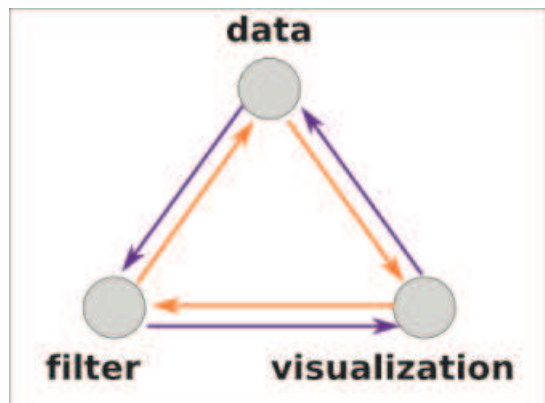


Illustration 3: Deux objectifs dans la construction des visualisations

Afin d'être la plus efficace possible, l'interface fait appel à la visualisation d'information (infoviz en anglais pour information visualization) définie comme l'utilisation de représentations visuelles, interactives, et sur ordinateur, de données abstraites pour amplifier la cognition [Schneiderman 1999]. Le caractère synoptique de la perception visuelle couplé aux possibilités du numérique en terme de manipulation et de présentation dynamique font de ce couplage un outil capable de faire naître des connaissances qu'il serait très difficile, sinon impossible, d'obtenir [Ware 2004]. Les visualisations peuvent prendre différentes formes : diagrammes, tableaux, cartes, graphes, etc., et peuvent être bidimensionnelles ou tridimensionnelles suivant le but que l'on recherche. Elles sont construites à partir de données de base, sur lesquelles on applique des filtres ou des traitements, que l'on visualise pour en prendre connaissance. Dans ce cas, il s'agit de transmettre une connaissance particulière sur le système et la visualisation doit être contrainte perceptivement pour qu'il ne soit pas possible de tirer des conclusions erronées. Un diagramme en deux dimensions est un exemple de visualisation contrainte. S'il est bien fait, il n'est pas possible d'en comprendre autre chose que ce qu'il présente. Ce n'est pas le cas des graphes ou des visualisations en 3 dimensions qui peuvent être mal interprétés. Une forme peut par exemple apparaître avec un angle de vue particulier mais il se peut qu'elle ne soit que le fruit d'un jeu de perspective. Un utilisateur incapable de changer le point de vue et qui ne ferait pas attention pourra d par ce biais se construire de fausses connaissances et c'est ce qu'il faut à tout prix éviter. L'interactivité est donc réduite de même que la 3 dimension et les visualisations ouvertes car le concepteur n'est pas capable de penser à l'avance toutes les connaissances qui pourraient naître de telle ou telle manipulation et cela perdrait alors son intérêt. Pour réduire au maximum les chances de mauvaises interprétations, ces visualisations sont construites suivant les règles de base du design graphique [Bertin 1967][Tufté 1993].

Il existe un second type de visualisations très important pour explorer les systèmes complexes : les visualisations comme espace d'actions. Dans ce cas, la visualisation est utilisée pour changer le filtre ou le traitement ou bien encore modifier leurs paramètres ce qui a pour effet de changer les données visualisées. Les actions sur ces visualisations donnent lieu à des sensations différentes et si

des régularités sont détectées alors naît une nouvelle connaissance, autrement dit une nouvelle hypothèse sur l'organisation . Et cette dernière provient directement de la manipulation. Ces visualisations sont donc ouvertes, interactives et exploratoires pour permettre aux utilisateurs de se construire des espaces d'actions et donc de perception. Ce type de visualisations est construit suivant les règles de l'infoviz [Schneiderman 1999][Munzner 2000] et elles sont principalement composées de graphes mais pas seulement.

L'interface finale est un mélange de ces deux types de visualisations et est elle-même une visualisation interactive synoptique de l'ensemble du processus. L'utilisateur est alors capable d'appréhender facilement l'enchaînement des traitements qu'il met en oeuvre. Les traces de ses actions sont directement accessibles et il est plongé au coeur de son processus, en immersion, ce qui limite les risques d'erreurs liés à la méconnaissance des actions précédentes, fréquentes lorsque l'on enchaîne beaucoup de traitements. Le processus global n'est pas fait de traitements automatiques et l'utilisateur intervient à tous les niveaux en manipulant l'interface, les traitements et les visualisations. Des traces ont été ajoutées à cette trace globale pour permettre à l'utilisateur mais aussi aux chercheurs s'intéressant à la découverte de connaissances de rejouer le film de l'exploration afin de comprendre comment et pourquoi les connaissances sont nées. La trace globale est bien entendu archivée ainsi que toutes les actions sur les filtres (paramètres) mais ce n'est pas tout. Sont aussi conservés l'ensemble des déplacements et des actions sur les visualisations, de sorte qu'un observateur puisse comprendre comment l'utilisateur a pu faire naître des connaissances qui proviennent des manipulations sur les espaces d'actions. L'interface finale est donc un outil utilisé à la fois par les explorateurs experts et par les chercheurs pour comprendre et pouvoir déconstruire à des fins de capitalisation, le processus global d'appréhension du système complexe.

Fort des ses ancrages théoriques, nous avons construit une méthode d'exploration du web en trois temps : rendre disponible les données, les explorer, et enfin manifester leur essence aux utilisateurs finaux.

Cliché du web

La première étape en amont de l'exploration consiste à prendre un cliché du web qui sera ensuite rendu disponible via une base de donnée. Ce cliché consiste à indexer l'ensemble des pages hypertextuelles du web, ce qui est bien entendu impossible compte tenu du nombre de pages qu'il contient. Google indexe à ce jour environ 8 milliards de pages et Yahoo plus de 20 milliards (information disponible sur leurs sites) et l'on estime qu'ils ne couvrent que 10% du nombre total de pages. La solution est alors d'indexer un sous-graphe du web. Pour ce faire on utilise la démarche suivante : on choisit des pages comme points

d'entrées, on envoie des robots à partir de ces pages qui vont suivre tous les liens hypertextes et indexer tous les contenus dans une base de donnée [Drugeon 2005]. Le problème est encore une fois le même que les mésopotamiens. A quelle profondeur en terme de nombre de clics faut-il aller, où doit-on s'arrêter ? Si l'on estime que le diamètre du web est de 19 clics [Albert 1999] alors une profondeur 19 reviendrait à indexer tout le web. D'autres problèmes viennent se superposer à cette question de taille. Les pages sont organisées en une unité sémantique et sémiotique que l'on appelle site Internet. Une solution pour réduire le nombre de documents à indexer est alors de ne considérer que les sites et non plus les pages. On peut conserver les pages mais la représentation de la connectivité est limitée aux sites. Du coup le stockage des données est plus réduit. On peut également choisir de ne considérer que les liens et un ensemble de mots clés pour éviter de stocker toute les pages en entier. De plus le web étant très dynamique, les changements dans les pages, d'un crawl à un autre, sont consignés et indexés aussi. Dans une mesure générale, les parcours des robots (crawl) ont pour but d'indexer tout le web sur un sujet particulier qui sera ensuite exploré. Pour cela on demande à des experts de ce domaine de donner leurs points d'entrées (des pages pertinentes). Ensuite les robots partent de ces pages et suivent les liens. Si l'on s'aperçoit que le contenu sémantique est trop éloigné du sujet de départ, le robot peut ne plus poursuivre (principe du focused crawling [Chakrabarti 1999]) dans cette voie et se recentrer sur d'autres pages.

Pour être certains d'avoir capté toutes les ressources intéressantes, le crawl est généralement lancé plusieurs fois et l'on détecte ainsi si des régularités apparaissent. Le processus de détection des régularités, décrit dans la section précédente, permet de stabiliser le sous-graphe du web extrait. Toutes les informations relatives au processus d'indexation des pages sont bien entendues conservées et synthétisées/manifestées en direction de l'utilisateur pour qu'il sache à tout moment sur quelles bases il travaille. Cette étape est encore largement expérimentale et basée sur le travail d'un utilisateur qui seul décide quand le cliché est satisfaisant, fort des indices statistiques et des visualisations dont il dispose. Quand le cliché le satisfait alors le travail d'exploration peut commencer.

Exploration

Les sous-graphes du web obtenus à l'étape précédente doivent maintenant faire l'objet d'une analyse pour en déduire l'organisation. Dans la mesure où cette étape résulte à chaque instant des choix de l'utilisateur, c'est bien une exploration qui procède d'étapes qui vont d'hypothèses en mises à l'épreuve qui infirment ou confirment une tendance. Chacune de ces étapes permet de passer à la suivante par une rupture en transformant la « matière » en « forme », ainsi que l'a décrit [Latour 1996] pour l'activité scientifique en général. Ici la matière est

constituée des données de l'étape et leur forme est la visualisation qui leur est associée, de manière plus ou moins forte suivant le type de visualisation choisie.

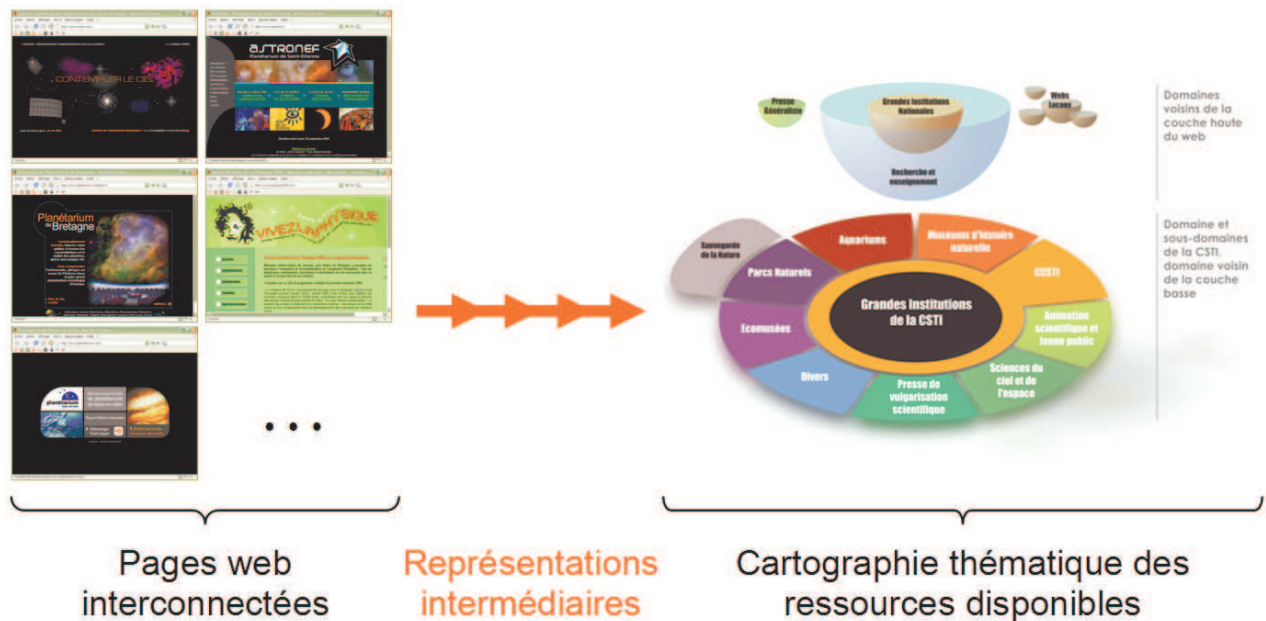


Illustration 4: Du web immersif au web synoptique

Chaque rupture est un nouveau traitement qui représente une nouvelle hypothèse et qui occasionne de nouvelles données qui prennent formes dans une nouvelle visualisation.

la visualisation en graphes

Voici un scénario d'analyse simplifié du web inspiré des méthodologies que nous avons élaborées. Le sujet choisi est la solidarité internationale, le site de départ : <http://www.ritimo.org> et la profondeur de 3 clics. Une fois indexé, ce sous-graphe d'Internet est mis en forme grâce à un logiciel d'affichage de graphes, dont la principale fonction est d'organiser spatialement le graphe sur l'écran de l'ordinateur pour le rendre lisible. Cette fonction appelée « force vector » applique des lois physiques à tous les nœuds du graphe : les nœuds se repoussent mais les liens les attirent. Ainsi manifesté dans un espace pseudo-physique, le sous-graphe du web devient manipulable de façon assez semblable à un objet matériel. La suppression d'un nœud entraîne la réorganisation de tout le graphe (on peut se représenter un tel graphe actif comme une sorte de sommier à ressorts tendu dans le vide).

Ce graphe actif est manipulé pour en faire émerger des motifs significants. Cette étape est complexe, et sera détaillée plus loin. les possibilités de manipulation du graphe permettent d'explorer les données. Nous éliminons les nœuds qui ne nous intéressent pas : tous les noeuds qui ne sont pas assez connectés. Il nous reste donc les nœuds les plus connectés, qui font un « paquet

» sur l'écran. Nous cherchons à identifier comment les nœuds restants sont connectés entre eux. Nous regroupons les nœuds sous différentes thématiques de façon à partitionner le graphe. Cette étape se fait par un marquage des nœuds par un code couleurs. Le classement en thématique est fait par l'utilisateur qui navigue sur les pages et fait lui même l'analyse sémantique. Il est aidé dans sa tâche par des outils sémantiques statistiques. Ainsi, petit à petit les quatre catégories suivantes émergent : « recherche et enseignement », « développement durable », « commerce équitable » et « divers ». Le graphe actif a été « dompté » et sa complexité réduite à une classification des sites les plus représentatifs. Les résultats de la classification sont ensuite transcrits par des tableaux, des copies d'écran, des textes... Une vue synoptique peut être produite pour rendre compte de l'analyse du domaine Internet de la solidarité internationale.

Le cœur du processus d'analyse et d'interprétation du web se trouve donc au niveau de l'exploration du graphe. Il y a plusieurs raisons à cela et la première est que le graphe actif est une présentation à la fois synoptique et analytique du web. Il est possible de distinguer des formes générales mais il faut pour cela manipuler le graphe et repérer les régularités. La seconde raison est que par nature, le graphe permet une double mise en présence du web par l'espace graphique du graphe et par la possibilité d'obtenir les pages Internet via le navigateur classique, auxquelles l'accès est facilité par l'équivalence page-nœud. Le graphe actif est susceptible d'être manipulé graphiquement pour faire apparaître diverses caractéristiques des pages. Les scores résultant d'algorithmes divers peuvent être manifestés par l'emplacement des nœuds, leur couleur, leur taille ou d'autres dispositifs plus sophistiqués.

Au final, la polyvalence du graphe le rend particulièrement efficace pour permettre une activité d'expertise basée sur la mise en co-présence de l'expert et de son objet.

Processus Général

La visualisation en graphe est un donc un outil efficace en terme de découverte de connaissances et tout particulièrement pour la recherche de motifs. De manière plus précise, le graphe est utilisé pour rendre perceptible des corrélations entre la topologie, la sémantique et la sémiotique des pages ou sites web. Dans la mesure où nous avons affaire à un système dont nous ne connaissons pas l'organisation entre les données, la tâche principale s'oriente vers la découverte de descripteurs permettant d'expliquer les relations entre ces données. Les graphes sont des objets ouverts et manipulables sur lesquels les actions de changement de points de vue ou de transformation de la présentation

peuvent faire ressortir des régularités. Autrement dit la boucle sensorimotrice sur la visualisation en graphe fait naître une perception qui est précisément une perception de la corrélation inscrite au départ. Cette perception est ensuite traduite en un descripteur du système qui représente une hypothèse sur l'organisation. Reste à la vérifier et c'est là le rôle de la première catégorie de visualisations qui visent à présenter le résultat d'un traitement de manière synthétique.

Dans le processus général, nous avons donc des données à notre disposition et des hypothèses de départ sur leur organisation. C'est le cas en topologie par exemple où [Kleinberg 1998] a proposé de considérer que les sites peuvent être classés en fonction de leur nombre de liens entrants et sortants (un portail aura beaucoup de liens sortants et une autorité beaucoup de liens entrants) et que cela fait sens dans l'organisation du web. Fort de ces hypothèses de départ, l'exploration consiste à mettre en place une double stratégie. Il faut dans un premier temps se donner les moyens de vérifier la pertinence de ces descripteurs ou hypothèses. C'est le rôle des visualisations contraintes. Elles permettent de vérifier des traitements construits sur la base du descripteur de la manière la plus juste qui soit. Si l'on souhaite vérifier que la densité de connectivité sur le web suit bien une loi de puissance [Barabasi 2002] par exemple, une visualisation en graphe ne sera d'aucune aide, contrairement à une visualisation en histogramme. La liberté de perception face à ce dernier type de présentation est très faible. Ensuite, si les descripteurs existants ne suffisent pas (et le système n'est plus complexe si ce n'est pas le cas) alors il faut trouver de nouveaux descripteurs. On fait ici appel au second type de visualisations. La dimension exploratoire décrite ci-dessus permet de découvrir de nouvelles prises, de nouveaux descripteurs qui à leur tour font faire l'objet d'un test en visualisation contrainte. Il faut préciser ici que les tout premiers moments de l'analyse d'un sous-graphe du web sont purement exploratoires et non pas analytiques. Il est nécessaire de se familiariser avec le graphe, de se forger des points de repère pour s'orienter dans le travail, d'aller voir les sites principaux, avoir un premier rapport avec l'objet avant de tenter de l'expliquer. Tout participe de la saisie de l'objet et doit la faciliter.

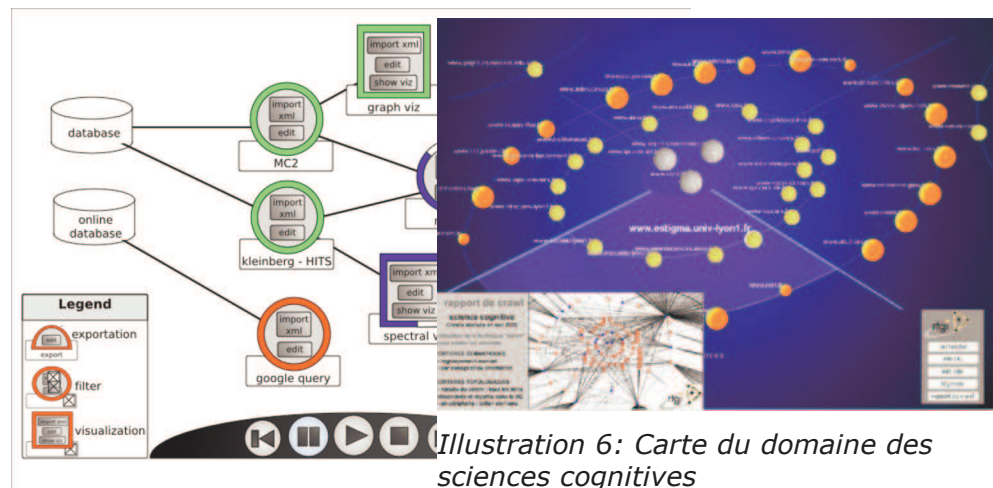


Illustration 6: Carte du domaine des sciences cognitives

La boucle entre découverte de descripteurs et leur validation s'arrête lorsque les descripteurs trouvés permettent d'expliquer de manière satisfaisante l'organisation du sous-graphe du web que nous avons indexé. Il s'agit là encore d'une application de la théorie sensorimotrice de la perception à un échelle plus large. Les actions sont les différents descripteurs mis en oeuvre et les sensations, l'allure du jeu de données ainsi constitué. La perception résultante est une appréhension de l'ensemble du processus autant que du sous-graphe indexé lui-même et donc d'une partie du web. Afin de favoriser cette exploration de second ordre, l'interface de manipulation des traitements est elle-même libre et favorise la perception de l'ensemble de la chaîne. La trace est disponible tout au long du processus et elle est en plus de cela visible et donc perceptible. Toutes les visualisations sont, de plus, présentes en même temps. La conséquence de cela est que l'on dispose à tout moment de différentes vues sur un phénomène unique. Ces présentations vécues ensemble comme autant de points de vues sur une même organisation réelle du web enrichissent la préhension en la facilitant.

Conclusion

Une fois le sous-graphe appréhendé, son organisation est mise à jour et l'on dispose d'un corpus de documents dont on connaît la structure. Vient alors la dernière étape qui consiste à rendre compte de cette structure. Dans le cas du web, cela signifie être en mesure d'expliquer le sous-graphe étudié. Ici encore tout dépend du besoin. Certains souhaiterons disposer d'indicateurs statistiques sur le sous-graphe et d'autres voudront pouvoir y naviguer mais de manière efficace et sans se perdre. Dans la mesure où l'explorateur expert sait quels traitements il a effectué pour parvenir à son résultat, il faut pour qu'elle soit comprise que son explication reflète ces traitements. C'est le cas avec la carte présentée dans l'illustration 6 page suivante. L'exploration précédant la carte a été effectuée à partir du descripteur de Kleinberg avec les portails et les

autorités, que l'on retrouve sur la carte car ce descripteur sert de métrique pour classer les sites du centre à la périphérie. De la même façon les sites ont été classés sémantiquement en catégories et l'on retrouve ces catégories comme des quartiers sur la carte. Cette visualisation est également interactive. Elle permet d'afficher différentes informations quant à la nature des sites et de les charger dans le navigateur. On se trouve donc face à une boucle où l'on est parti du web, on l'a isolé, exploré et analysé pour finalement y retourner mais cette fois-ci avec un guide. Ce dernier est d'autant plus pertinent qu'il présente la structure de la portion du web étudiée.

Au final, pour celui qui expertise un sous-graphe du web, la problématique n'est pas d'appliquer des traitements automatisés mais de s'engager dans les « plis » du web et de constituer des « prises » qui puissent être transmises. Il s'agit d'explorer le web en l'appréhendant par divers outils, et d'identifier des traces intéressantes. Ces traces sont tout d'abord diffuses, et par les différentes étapes que l'expert effectue, il cherche à reporter ces traces de représentation en représentation jusqu'à produire des vues synoptiques du domaine et des « zooms » sur des saillances pertinentes. Au cours de ce processus, le web perd en multiplicité, en flexibilité, en complexité, mais gagne en universalité et en lisibilité.

Cette méthode permet donc à partir du web d'en tirer l'essence via des dispositifs adaptés qui donnent une large part aux visualisations. L'usage à bon escient de ces dernières permet de se saisir de l'objet étudié et d'en faire un objet rationnel qui n'échappe pas à notre prise. Les visualisations sont à la fois des outils de capitalisation et de production de traces qui permettent aux utilisateurs de s'en saisir et de suivre des fils, des intuitions, sans se perdre. Une première version de l'interface existe déjà dans le logiciel TARENTE [Ghitalla 2004] et nous sommes en train de développer une seconde version baptisé FEW³ (Framework for Exploring World Wide Web) en partenariat avec l'INA.

Bibliographie

Albert, B., Barabasi, A.-L., Jeong, H., *Diameter of the World Wide Web.* Nature, 401(6749):130--131, September 1999.

Barabasi, A. L., *Linked : The New Science of Networks,* Perseus Publishing, 2002.

Bottero, J., *Mésopotamie, L'écriture, La raison et les dieux,* Gallimard, 1987.

Bertin, J., *Sémiologie Graphique. Les diagrammes, les réseaux, les cartes.*

- Paris-La Haye : Mouton, Gauthier-villars, première édition, 1967.
- Chakrabarti, S., van der Berg, M., & Dom, B.,** *Focused crawling: a new approach to topic-specific Web resource discovery.* In Proceedings of 8th International World Wide Web Conference (WWW8), 1999
- Drugeon, T.,** A technical approach for the french web legal deposit. IWAW 2005.
- Ghitalla, F., Diemert, E., Maussang, C. , Pfaender, F.,** *TARENTE : an Experimental Tool for Extracting and Exploring Web Aggregates,* ICTTA'04, Damas, Syrie, 2004.
- Ghitalla, F., Lenay, C.,** *Les territoires de l'information. Navigation et construction des espaces de compréhension sur le web,* dans *La navigation,* Les cahiers du numérique, Hermès Edition, Paris, 2003
- Kleinberg, J.,** *Authoritative Sources in a Hyperlinked Environment,* in Proc. of the ACM-SIAM Symposium on Discret Algorithms, ACM Press, 1998.
- Latour, B.,** *Le pédofil de Boavista, montage photo-philosophique,* in *Petites leçons de sociologie des sciences,* Coll. "Point", Seuil, p171-225, 1996.
- Lenay, C.,** *Constitution de l'espace et immersion,* Arob@se, www.univ-rouen.fr/arobase, volume 1, pp. 85-93, 2005.
- Lorenz,** *Predictability: Does the Flap of a Butterfly's Wings in Brazil set off a Tornado in Texas?,* meeting of the American Association for the Advancement of Science in Washington, D.C., 1972.
- Munzner, T.,** *Interactive Visualization of Large Graphs And Networks,* Ph.D. Dissertation, Stanford University, June 2000
- O'Reagan, K., Noë, A.,** *A Sensorimotor account of vision and Visual Consciousness,* Behavioral and Brain Sciences, 24 (5), 2001.
- Shneiderman, B., Card, S.-K., MacKinlay, J.-D.,** *Readings in Information Visualization, Using Vision to Think,* Morgan-Kaufmann Publishers, New-York, 1999.
- Tufte.** *Visual Display of Quantitive Information.* Graphic Press, Cheshire, connecticut, seconde edition, 1993.
- Ware, C.,** *Information visualization: Perception for design.,* Morgan-Kaufmann Publishers, San Francisco, 1999.

Utiliser des traces de la dimension globale d'un corpus pour l'accès au contenu des documents

Pierre Beust, Thibault Roy

GREYC CNRS UMR 6072 – ISLanD & Pôle ModesCoS de la MRSH
Université de Caen - Basse Normandie
14032 Caen Cedex, France
{pierre.beust, thibault.roy}@info.unicaen.fr
<http://users.info.unicaen.fr/~beust/>

Résumé : Dans nos travaux nous cherchons à développer des outils adaptés à certaines tâches informatisées où les spécificités socio-linguistiques des utilisateurs (par exemple leurs centres d'intérêt, leurs habitudes terminologiques, leurs corpus de travail) et leurs interprétations sont au centre de l'interaction homme-machine. La question du sens (sa construction et sa nature) est bien sûr très liée aux rapports entre des documents (majoritairement textuels) et des sujets interprétants travaillant avec ces documents et en produisant par ailleurs. Nous nous intéressons à la dimension intertextuelle du sens des documents dans le but d'améliorer les méthodes d'accès au contenu. Notamment nous cherchons à expérimenter d'un point de vue opératoire le principe herméneutique de la détermination du local par le global. Notre approche consiste à rechercher et identifier des traces de la dimension globale d'un corpus pour les utiliser dans des analyses locales.

Mots clés : TAL, Système individu-centré, Interface de lecture rapide

1. Introduction

La question des traces en informatique est assez transversale. On la retrouve dans un bon nombre de problèmes et dans les solutions apportées à ces problèmes. Par exemple, l'importance donnée aux fichiers de log dans le fonctionnement des systèmes d'exploitation évolués en est un exemple, de même que le développement des cookies dans les connections Internet (traces laissées par des serveurs sur des machines clients) ou encore les fichiers temporaires générés par des applications. Toutes les activités de l'utilisateur sont donc retracées dans et par les systèmes informatiques, dans certains cas dans un but d'authentification des utilisateurs ou dans d'autre cas dans un but de sauvegardes de données.

Dans ces problèmes, principalement d'ordre technique, générer des traces est la réponse au problème. Dans d'autres types de problèmes la question n'est pas de générer des traces mais plutôt d'en identifier et de chercher à les exploiter le mieux possible. C'est assez souvent le cas dans certaines approches du Traitement Automatique des Langues. Par exemple dans les approches d'analyse de corpus dites endogènes, on cherche à exploiter du matériau linguistique attesté pour en déduire des ressources qui vont en retour permettre d'alimenter

des analyses plus fines. C'est par exemple le cas des travaux de (Déjean 1998) qui cherchaient à extraire à partir de textes en langues inconnues des suffixes ou affixes redondants pour améliorer les analyses syntaxiques multilingues. De même que les approches endogènes, les méthodes d'analyses statistiques de gros volumes de données textuelles (par exemple les approches de type Zipf) cherchent à identifier des traces significatives qui ne sont souvent pas à la portée de l'analyse humaine.

Notre problématique dans cet article se situe dans ce cadre. Il s'agit d'identifier des traces que l'on peut caractériser de manière opératoire pour améliorer en retour les méthodes d'accès au contenu des textes électroniques. Nous relatons ici une recherche en cours sur la cartographie thématique personnalisée dans laquelle on cherche à exploiter la dimension intertextuelle du sens des documents.

Dans une première partie nous préciserons au regard de quel cadre épistémologique nous analysons la question de la construction du sens. Nous détaillerons ensuite la question de la détermination du global sur le local. Nous présenterons alors nos travaux et nos expériences en cours dans le domaine de la cartographie thématique de corpus.

2. La question de la construction du sens

Gérard Sabah¹ précise ce que peut être le sens du point du résultat visé par telle ou telle sémantique :

- Préciser des conditions de vérité (sémantique vériconditionnelle)
- Décrire une expression comme l'ensemble des propriétés théoriques que possèdent les concepts correspondants (sémantique intensionnelle)
- Décrire une expression comme l'ensemble des objets ou des situations du monde de référence que cette expression peut désigner (sémantique dénotationnelle ou référentielle)
- Chercher à décomposer le contenu des mots en éléments de sens plus primitifs pour étudier les possibilités de combinaison de ces éléments (sémantique componentielle)
- Décrire une expression comme l'ensemble des actions à effectuer pour trouver l'objet désigné (sémantique procédurale)
- Mettre en évidence les marqueurs et les constructions utilisées pour qu'un énoncé puisse servir comme un argument en faveur d'un autre énoncé (sémantique argumentative)

Aucune de ces sémantiques n'a complètement tort ou complètement raison. Par exemple, il existe bon nombre d'énoncés en langue naturelle dont le sens a un certain rapport avec la vérité ou encore la référence, mais cela ne veut pas dire que *tout* énoncé est interprétable en terme de vérité ou de référence.

¹ <http://www.limsi.fr/Individu/gs/textes/ATALA-14.12.96/LePointSurLeSens.html>

Selon cette catégorisation, nous situons notre approche de la notion de sens dans le cadre d'une sémantique componentielle et plus précisément dans un héritage scientifique de la Sémantique Interprétative (SI) de François Rastier (Rastier 1987), elle-même en filiation avec les travaux en sémantique structurale dont l'origine remonte aux travaux de Saussure. De la SI on retient principalement deux aspects. Premièrement, on en tire (parfois avec quelques adaptations au cadre d'une instrumentation informatique²) un « appareillage » théorique fin pour la description d'effets de sens. Par exemple, les notions de sèmes, d'isotopies (i.e. récurrences d'un même sème dans un texte), ... Deuxièmement, on s'approprie de la SI un positionnement épistémologique relativement à la question du sens. Celui-ci consiste à préférer la tradition rhétorique et herméneutique à la tradition logico-grammaticale. Ainsi on défend, d'une part, le principe selon lequel le global détermine le local (ce qui marque une rupture avec le principe de compositionnalité) et, d'autre part, que le sens ne peut pas être intégralement objectivé (Rastier 1998).

Si le sens ne peut pas être objectivé il peut encore moins l'être de manière formelle (comme cela est souvent le cas dans la tradition logico-grammaticale). On rejoint ici l'avis de (Nicolle 2005) pour qui le sens n'est jamais capturé par ses représentations. Toute représentation du sens est forcément incomplète et il n'y a donc pas de langage formel qui puisse reproduire fidèlement le sens d'un énoncé en langue naturelle. Anne Nicolle en tire la conséquence que la langue est un langage terminal. La compréhension, l'interprétation langagière, est en cela bien distincte de l'interprétation logique se résumant à la traduction dans un autre langage car pour nous, il n'y a pas d'autre langage.

Ainsi, à la manière de Courcil qui définit les principes de non consignation et de non préméditation de la chaîne parlée (Courcil 2000), nous défendons un principe de non transformation du sens en langue naturelle dans la mesure où il n'y a pas de pensée construite possible qui ne soit pas déjà sous forme langagière. Le sens est donc une réalité concrète intralinguistique et subjective. Dans le cadre de la SI, l'interprétation est considérée comme une perception sémantique, perception forcément individuelle, dont toute tentative d'objectivation est une sommation incomplète de points de vue. Ainsi le sens d'un texte est une interprétation à un moment donné et dans une tâche donnée d'un sujet interprétant, ce qui est à notre avis un argument fort pour une instrumentation de la sémantique des langues individu-centrée.

Dans la logique des partis pris épistémologiques qui sont les nôtres, nous préférons de loin l'idée d'une instrumentation du sens à celle de la construction du sens. D'une manière imagée, il nous semble qu'un traitement sémantique informatisé a plus de points communs avec un outil tel un microscope par exemple, c'est-à-dire quelque chose qui nous montre ce qui est déjà là mais qu'on ne voyait pas de cette façon, qu'avec un outil tel un transformateur

² C'est par exemple le cas concernant la notion de sème que nous avons redéfini dans (Beust 1998).

électrique qui produirait à partir de quelque chose une autre chose qui n'existait pas avant. Il s'agit donc de considérer que le sens tel que le produit une interprétation humaine n'est pas à la portée d'un seul traitement informatique. Cela ne veut pas dire pour autant que les machines ne puissent pas interpréter des textes. Elles le font à leur manière comme les sujets humains le font aussi à leur manière. Nous opposons ici l'idée d'une Interprétation Calculatoire (IC) à celle d'une Interprétation Humaine (IH). Ces deux formes d'interprétation ne sont pas en concurrence car l'IC n'a en aucun cas le but de supplanter l'IH. Au contraire, nous les pensons comme complémentaires dans le sens où une interprétation calculatoire a pour objectif de produire dans l'interaction des signes (notamment au moyen de techniques de visualisation adéquates) qui vont participer aux interprétations humaines du ou des utilisateurs.

La recherche de ce qui est à la limite des deux formes d'interprétation et qui pourrait se prêter à une modélisation dans le cadre d'une interprétation calculatoire nous semble mettre en avant la problématique des traces de deux manières :

- Comment une interprétation calculatoire peut-elle produire des traces utiles et efficaces à l'interprétation humaine ? Nous proposons pour cela différentes méthodes de visualisation telles la coloration de textes ou encore la cartographie de corpus.
- Comment identifier et exploiter des traces des relations intertextuelles entrant en compte dans les interprétations et donc dans l'accès au contenu d'un document. Nous cherchons ici notamment à opérationnaliser des traces de la détermination du local par le global.

3. Le rapport local/global dans la contextualisation du sens

Le principe de détermination du local par le global propre à l'approche herméneutique est un principe de contextualisation du sens (principe alternatif à la compositionnalité). La question de la contextualisation du sens de telle ou telle unité linguistique du texte (mot, syntagme, paragraphe par exemple) est d'une grande importance dans la perspective de la sémantique interprétative car elle est la base de l'établissement de parcours interprétatifs, c'est-à-dire des suites d'opérations permettant d'assigner un ou plusieurs contenus à des expressions (ce qui explique comment les langues peuvent s'acquérir réflexivement par leur pratique). Dans la contextualisation entrent en compte, selon nous, 3 notions : le co-texte, le contexte extralinguistique et l'intertexte :

- On entend par co-texte d'une unité linguistique son « entourage » dans le texte, c'est-à-dire un passage de texte : une zone de localité sémantique pertinente autour d'une unité. Cette zone est appelée *Période* (Rastier & al. 1994, p. 116) et elle est délimitée par l'étendue des relations d'isotopies, de prédication et d'anaphore.

- Le contexte extralinguistique regroupe les conditions pragmatiques liées à l'interprétation du texte. Dans le cadre de nos expérimentations en recherche d'information, on limitera ce contexte à l'utilisateur et sa tâche (voir au groupe d'utilisateurs et leur tâche).
- L'intertexte rassemble tous les documents que l'utilisateur estime liés à un texte du point de vue de son interprétation. Tout texte mis en relation avec d'autres textes en reçoit des déterminations sémantiques et modifie potentiellement le sens de chacun des autres textes (c'est le principe d'architextualité chez Rastier³). Un document peut appartenir à plusieurs intertextes comme un texte peut s'interpréter dans plusieurs intertextes en fonction des relations sémantiques établies, c'est-à-dire des objectifs interprétatifs. On peut penser par exemple qu'un même article de presse n'indique pas le même point de vue quand il y est fait référence dans une revue de presse des plus sérieuse ou dans les colonnes d'un journal satirique.

Ainsi contextualiser, c'est établir au sein du co-texte des parcours interprétatifs qui tiennent compte du contexte extralinguistique et de l'intertexte. Ce que l'on peut analyser avec les moyens d'une sémantique componentielle comme mettre en évidence dans certains passages du texte des sèmes particulièrement importants. Ainsi caractériser une trace de la détermination du local par le global peut être d'identifier localement des sèmes pertinents issus du global (le contexte ou l'intertexte). Le contexte et l'intertexte restent deux notions aux contours assez flous : le contexte est vu dans le cadre de nos expérimentations via les ressources personnelles construites par le ou les utilisateurs, et l'intertexte via le corpus d'étude⁴ sur lequel les outils logiciels développés permettent de travailler. Le corpus « matérialisant » l'intertexte en tire un rôle central par rapport à la construction du sens en contexte, à l'égal des ressources exploitées, qui lui confère bien plus d'importance qu'un simple réservoir de formes attestées.

Identifier et caractériser l'importance, ou pas, d'un sème dans la contextualisation du contenu d'une unité est le résultat d'opérations interprétatives d'afférence. L'afférence consiste en la production d'un sème qui vient créer ou renforcer une isotopie. Selon Olivier Mezaïlle⁵, il y a une telle production d'un sème isotopant dès lors qu'est établie une relation sémantique d'assimilation ou de dissimilation. L'assimilation est une afférence co-textuelle (c'est-à-dire que le sème afférent est vu comme inhérent dans d'autres unité du co-texte) d'un sème générique par présomption d'une récurrence sémique. Par

³ cf. http://www.revue-texto.net/Reperes/Glossaires/Rastier_Glossaire.html

⁴ Le corpus n'est pas pour autant une simplification de l'intertexte, c'en est selon (Rastier 1998, note de bas de page n°17) une objectivation : « Le corpus est la seule objectivation possible (philologique) de l'intertexte, qui sinon demeure une notion des plus vagues »

⁵ cf. « Quels mécanismes pour (r)établir la cohésion sémantique textuelle ? Sur la prééminence des processus d'assimilation et de dissimilation dans l'interprétation des énoncés contradictoires et métaphoriques » disponible en ligne à l'adresse : <http://www.chez.com/mezaïlle/contraphore.htm>

exemple, dans l'énoncé *Voici des choux, des concombres et des scoubidou* l'assimilation consiste en une afférence d'un sème à *scoubidou* où le but est de renforcer une répétition déjà initiée dans le co-texte. En l'occurrence, il peut s'agir d'enrichir le contenu sémique de *scoubidou* avec le sème afférent générique⁶ /légume/ pour rendre le thème de l'énoncé uniforme⁷. L'opération interprétative inverse de l'assimilation est la dissimilation. Alors que l'assimilation diminue, par afférence, les contrastes forts, la dissimilation, quant à elle, augmente les contrastes faibles. La dynamique des sèmes en cause dans la dissimilation n'est plus une afférence de sème générique, comme c'est le cas pour l'assimilation, mais une afférence de sèmes spécifiques pour différencier, dans un co-texte, les contenus sémantiquement proches. Par exemple, dans l'énumération *routes et autoroutes*, le contenu de *route* doit décrire une signification spécifique qui exclut la signification de *autoroute* et non une signification générique qui inclut cette signification⁸.

L'assimilation et la dissimilation sont des formes d'afférences co-textuelles mais la sémantique interprétative décrit également une autre forme d'afférence : l'afférence socialement normée. Dans de tels cas d'afférence, il y a bien enrichissement contextuel du contenu d'une lexie dans un énoncé par un (ou plusieurs) sème(s) (c'est bien pour cela qu'il s'agit toujours d'une afférence) mais cette fois, le ou les sèmes afférents ne sont pas inhérents dans d'autres contenus d'unités linguistiques du co-texte. L'afférence est alors le fait d'une norme sociale partagée au sein d'une communauté linguistique. C'est, par exemple, le cas du sème /tristesse/ afférent au contenu de *noir* dans *il broie du noir* ou encore le cas du sème /bonheur/ dans *rose* dans *la vie en rose*. Là où l'afférence co-textuelle (par assimilation ou dissimilation) est le résultat d'un parcours interprétatif local, l'afférence socialement normée résulte de parcours interprétatifs beaucoup plus globaux (à titre d'exemple d'afférences globales sur l'ensemble d'un corpus, on peut citer l'étude de (Rastier, 1987) sur le roman de Stendhal *Le rouge et le noir* où il y a une afférence jusque dans le titre du roman des sèmes /armée/ et /Église/).

Le concept d'afférence est un outil théorique très fin pour la description de la dynamique des sèmes dans les corpus et pour la caractérisation des effets du global sur le local. Le problème de la modélisation de l'afférence (surtout en ce qui concerne l'afférence socialement normée) c'est qu'il faudrait avoir des

⁶ L'assimilation concerne bien des sèmes génériques car c'est le sème /légume/ qui est afférent et pas le sème /vert/ inhérent à *choux* et *concombre* mais spécifique, ce qui a pour conséquence qu'il n'y ait pas de défaut d'assimilation dans *Voici des choux, des concombres et des carottes*

⁷ On retrouve ce type d'assimilation dans certaines formes d'humour lorsque le sème afférent est particulièrement inattendu dans le contenu sémique de la lexie en cause. Le titre du livre de G. Lakoff *Women, Fire and Dangerous Things : What Categories Reveal about the Mind* (1987) en témoigne en provoquant l'afférence du sème /dangereux/ au lexème *Women*.

⁸ Par exemple, dans *routes et autoroutes*, le contenu de *route* doit décrire une signification spécifique qui exclut la signification de *autoroute* et non une signification générique qui inclut cette signification. La dissimilation est encore plus flagrante dans l'exemple suivant où il y a répétition de la même lexie dans *Il y a musique et musique*. Ici, la dissimilation consiste, à distinguer par des sèmes spécifiques les deux significations de *musique*. Ainsi, on peut afférer à la première le sème spécifique /agréable/ et à la seconde /désagréable/.

ressources très larges pour pouvoir l'expérimenter de façon opératoire dans un traitement automatique. Comme on l'a dit plus haut, un texte n'est pas lié à un unique intertexte. Du point de vue du sujet interprétant et de son histoire propre, tous les intertextes forment un univers de textes construit individuellement de manière le plus souvent inconsciente. T. Thlivitit l'appelle l'*anagnose* (Thlivitit 1998, p. 41). Dans la perspective centrée utilisateur qui est la nôtre (de même d'ailleurs que T. Thlivitit) l'anagnose serait idéalement ce qu'il faudrait formaliser pour représenter au mieux l'utilisateur et modéliser avec satisfaction ses afférences. Seulement, il n'est pas du tout évident, au contraire, que cette anagnose rassemblant l'histoire d'un individu (l'histoire de ses interprétations comme son histoire propre), sa culture, sa société, les données de son époque soit formalisable. On est ici face aux mêmes problèmes que celui de la constitution des ontologies généralistes. Dès lors, les différentes ressources personnelles que nous serons amenés à construire et à manipuler ne toucheront que des infimes parties de cette anagnose. Ce n'est pas pour autant un problème car il y a moyen de personnaliser efficacement (comme l'a montré V. Perlerin⁹) des tâches de recherche d'information et d'accès au contenus de documents avec des ressources légères d'emblée non exhaustives.

4. La cartographie thématique

La plate-forme ProxiDocs¹⁰ (Roy 2005) permet de construire différentes cartes. On entend ici par *cartes* des représentations graphiques (en général planaires mais en généralisant on en produit aussi en 3 dimensions) visualisant un espace topologique et métrique construit sur une notion de distance (on exploite ici la distance euclidienne mais d'autres existent et pourraient convenir ; celle du Chi2 par exemple). Ces cartes ont pour but de fournir à l'utilisateur des représentations visuelles globales d'un corpus de textes dans son ensemble.

Ces cartes sont produites à partir de ressources terminologiques construites par l'utilisateur qui vont représenter les thèmes ou les domaines que ce dernier souhaite faire intervenir dans ses analyses. Deux types de représentations sont possibles selon les besoins de l'utilisateur : la représentation en « sacs de mots » où chaque thème est représenté par un ensemble de lexies s'y rapportant selon le point de vue de l'utilisateur ; et la représentation selon le modèle Lucia de sémantique différentielle (Perlerin 2004) où chaque domaine est représenté par un ensemble (appelé un dispositif) de catégories de lexies dont la signification est représentée par des différences de sèmes. Afin de construire de telles

⁹ (Perlerin 2004)

¹⁰ <http://www.info.unicaen.fr/~troy/proxidocs>

ressources terminologiques, nous proposons un ensemble de logiciels d'étude¹¹ complémentaires apportant une aide à l'utilisateur.

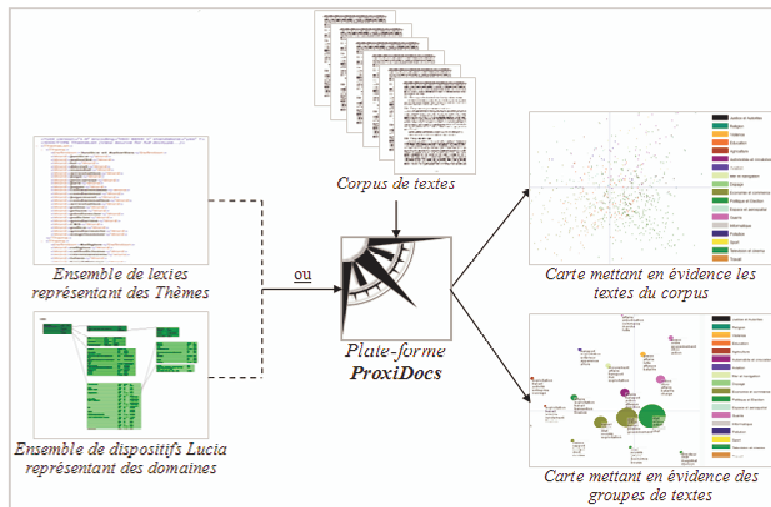


Figure 1 : Schéma d'utilisation de la plate-forme ProxiDocs.

ProxiDocs produit des cartes visualisant différentes analyses globales du corpus à partir des ressources de l'utilisateur (la figure 1 illustre les possibilités de la plate-forme exploitées dans cet article) :

- des cartes en 2 ou 3 dimensions représentant chaque texte du corpus analysé par un point. Chacun de ces points est un lien hypertexte vers le texte où les lexies des thèmes ou des dispositifs sont mises en valeur à l'aide d'une technique de coloriage. La couleur d'un point correspond au thème majoritaire repéré dans l'article représenté. De telles cartes sont utiles afin d'étudier les liens et les différences de thématiques abordées entre les documents du corpus
- des cartes en 2 ou 3 dimensions mettant en évidence des groupes de textes abordant des thèmes proches. Chaque groupe est représenté sur la carte par un disque ou une sphère de diamètre proportionnelle au nombre de textes qu'il contient. La couleur attribuée au disque ou à la sphère correspond au thème majoritaire repéré dans les textes du groupe. Les disques représentant les groupes portent un jeu d'étiquettes (au plus 5 étiquettes) indiquant les lexies localement les plus fréquentes dans l'ensemble des documents du groupe. Chacun de ces groupes est un lien hypertexte vers un rapport sur le contenu du groupe représenté. Ce rapport indique le texte le plus « représentatif » du groupe (celui situé le plus près du centre de gravité du groupe), présente la répartition des thèmes ou des catégories abordés, met en évidence les lexies répétées et propose un classement des isotopies les plus « importantes » dans les documents du groupe dans le cas d'une cartographie réalisée à partir de dispositifs

¹¹ Ces logiciels d'étude sont tous disponibles sur le site de l'équipe ISLanD du laboratoire GREYC de l'université de Caen : <http://www.greyc.unicaen.fr/island/logiciel/>

- Lucia. Ces cartes permettent d'avoir un regard sur les principaux sujets abordés ainsi que sur la répartition des thèmes dans les textes du corpus
- des cartes en 2 dimensions animées mettant en évidence l'évolution des thèmes abordés dans les textes lorsque les documents qui constituent le corpus sont datés (c'est par exemple le cas d'un corpus de dépêches d'agence de presse). Ce type de cartes permet de mettre en évidence les différentes thématiques abordées sur certaines périodes ainsi que leur enchaînement.

5. Expériences sur corpus réalisées avec ProxiDocs

Dans le but de caractériser la part de la dimension globale intertextuelle dans l'accès aux contenus de documents, nous avons mené quatre expériences différentes avec la plate-forme ProxiDocs :

- la première sur corpus généraliste avec des ressources elles aussi assez généralistes sous forme de listes de thèmes créées par nos soins.
- la deuxième sur corpus très spécialisé avec des ressources très spécifiques développées sous formes de dispositifs (représentations terminologiques différentielles) également créés par nous même.
- la troisième sur un flux documentaire avec des ressources créées par nous même et visant les différentes thématiques abordées à différents moments du flux.
- la quatrième sur un corpus et des ressources lexicales utilisées dans une recherche en cours sur la terminologie médicale par une chercheuse de l'université de Rouen.

La première expérience réalisée consiste à cartographier un corpus thématiquement hétérogène constitué d'environ 800 articles issus du journal *Le Monde* de 1987 à 1989. Cette expérience prend place dans le domaine de la veille d'informations : elle a pour objectif de découvrir les principaux sujets abordés dans cet ensemble d'articles. Les cartes obtenues à l'issue de cette expérience (présentées en figure 2) ont été réalisées avec un ensemble de 18 thèmes généralistes que nous avons construits, tels la justice, la télévision, l'éducation, etc. La carte des articles met en évidence un nombre très important d'articles de thèmes majoritaires Politique et élection (quadrant inférieur droit de la carte) et Économie et commerce (quadrant inférieur gauche). Ces observations sont confirmées par la carte des groupes d'articles, la couleur, la taille et la disposition des groupes sur cette carte donnent une information sur les thèmes abordés dans les textes du corpus ainsi que sur leur répartition. En visualisant les rapports des groupes, l'utilisateur peut avoir une idée plus précise des thèmes abordés dans les articles de chaque groupe : il est ainsi facilement observable que le groupe de thème majoritaire Politique et élection contient principalement des articles traitant des futures élections européennes. Les différentes cartes construites durant cette expérience nous ont alors permis de mettre rapidement

en évidence les grandes tendances du corpus (principaux sujets abordés dans les textes et groupes de textes abordant des thèmes proches), ce qui est un premier résultat satisfaisant pour une interface de lecture rapide particulièrement utile dans une tâche de veille d'informations.

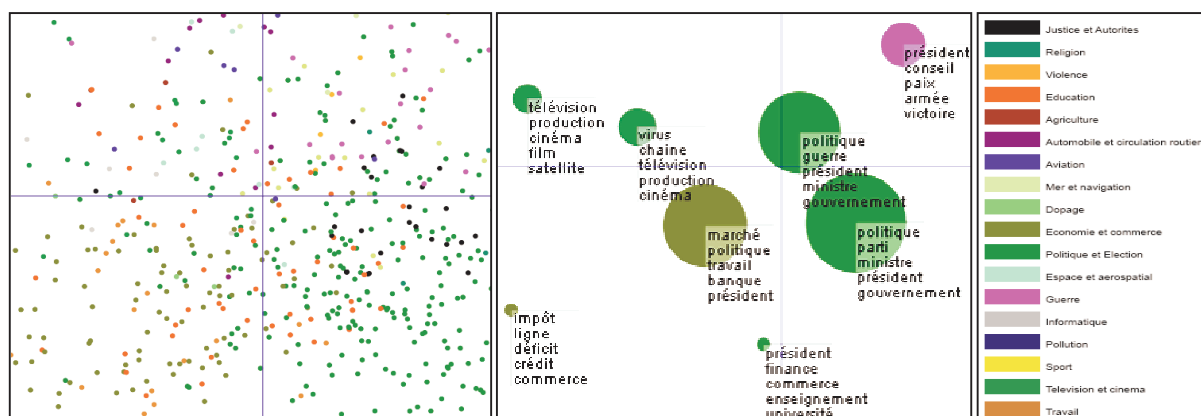


Figure 2 : Cartes thématiques obtenues à partir d'un corpus d'articles de journal et de thèmes généralistes. La carte mettant en évidence les articles du corpus est située en haut et à gauche de la figure, la carte mettant en évidence des groupes de textes est située en haut et à droite, la légende de couleur de la carte est indiquée sur la partie inférieure.

La deuxième expérience présentée ici consiste à observer un fait de langue : la métaphore conceptuelle au sens de Lakoff et Johnson (1980). Cette observation est faite sur un corpus d'environ 300 articles boursiers issues du journal *Le Monde* entre 1987 et 1989. Les analyses ont porté sur trois métaphores conceptuelles : la *météorologie boursière*, la *guerre économique* et la *santé financière*, un nombre important de ces trois métaphores ayant été observé dans notre corpus d'étude¹². Les cartes obtenues à l'issue de cette expérience (présentées en figure 3) ont été réalisées avec un ensemble de 3 dispositifs Lucia représentant les domaines de la météo, la santé et la guerre. Les cartes obtenues nous ont notamment montré une proximité entre des documents contenant des emplois métaphoriques d'un même lexique. Il a aussi été possible d'observer un lien entre le type d'articles (bilans, dépêches, etc.) et les métaphores conceptuelles qui y apparaissaient. De cette manière, nous avons pu déterminer que les métaphores de la *guerre économique* se situaient plutôt dans des dépêches détaillant des événements boursiers ponctuels alors que les métaphores de la *météorologie boursière* et de la *santé boursière* se retrouvaient de manière simultanées dans les bilans boursiers hebdomadaires et mensuels.

¹² Des extraits de notre corpus d'étude illustrant respectivement ces trois métaphores conceptuelles :

- « Une véritable **tempête** de hausses, alimentée par une marée de capitaux, étrangers pour partie, en quête de placement. » Le Monde 03/08/87
- « Le dénouement dans la **bataille** autour de la première banque commerciale privée du pays a eu peu d'effet sur les cours. » Le Monde 27/02/89
- « La pente fut longue à remonter, et il fallut bien douze mois pour **panser** les **plaies** du sinistre et à commencer à croire à de nouveaux records d'altitude pour le CAC. » Le Monde 01/08/89

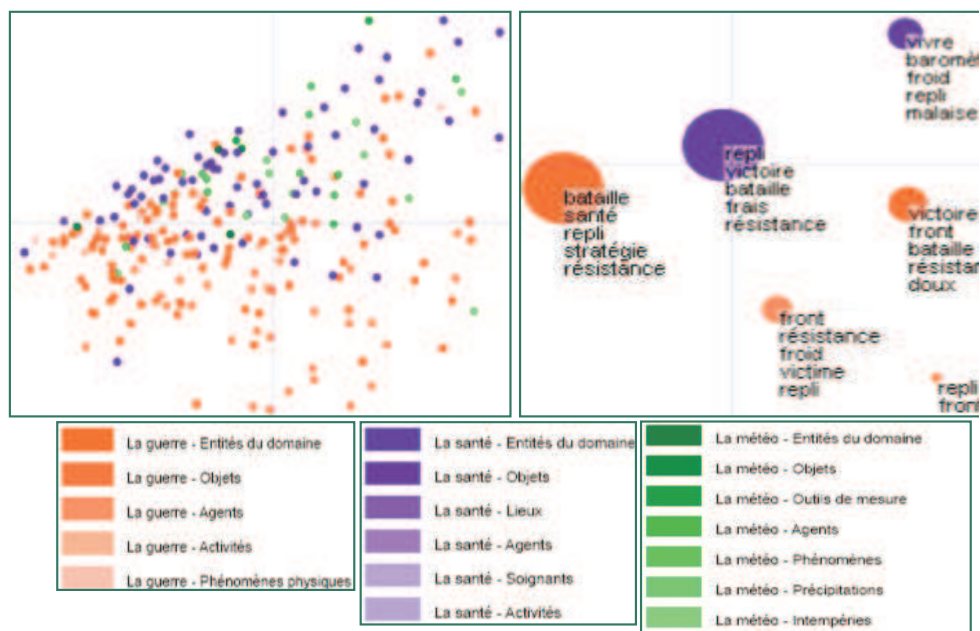


Figure 3 : Cartes thématiques obtenues à partir de notre corpus constitué d'articles boursiers et de dispositifs Lucia représentant les domaines des métaphores conceptuelles étudiées.

La troisième expérience consiste à cartographier un forum de discussion spécialisé portant sur l'apprentissage d'un langage de programmation. Le forum étudié est issu de la plate-forme INES¹³. Il permet à des étudiants de DEUST Technicien des Systèmes d'Information et de Communication d'échanger des messages en rapport avec leur formation. Ce forum est constitué d'environ 200 messages échangés entre le 18/02/2003 et le 27/04/2005 par 27 intervenants différents (enseignants et étudiants). Des cartes ont été construites à partir d'un ensemble de 5 thèmes spécialisés que nous avons construits, ces thèmes portant sur l'enseignement, son déroulement, la recherche d'informations, etc. Elles ont mis en évidence les sujets principalement abordés dans les messages. Il est ainsi possible d'observer que la thématique liée au déroulement des enseignements est très majoritaire dans le forum, alors que la thématique liée au contenu même des enseignements est très peu abordée. Ces résultats globaux peuvent alors être des signes pour les enseignants de certaines attentes de leurs étudiants.

La quatrième et dernière expérience présentée ici consiste à cartographier un corpus d'articles scientifiques médicaux à partir de ressources lexicales portant sur ce domaine. Cette expérience est en cours de réalisation et s'inscrit dans une recherche sur la terminologie médicale menée par une chercheuse en TAL de l'université de Rouen (Aurélie Néveol¹⁴). La carte présentée en figure 4 illustre les premiers résultats obtenus. Elle met ainsi en évidence des sous-ensembles d'articles et caractérise chacun de ces sous-ensembles par ses « métatermes¹⁵ » les plus fréquents. Lors du passage de la souris sur l'un de ces métatermes, les métatermes identiques caractérisant les autres groupes se mettent en relief

¹³ <http://www.dep.u-picardie.fr/>

¹⁴ aneveol@insa-rouen.fr

¹⁵ Les métatermes sont des regroupements de mots-clés réalisés en fonction de spécialités médicales, par exemple le métaterme « ophtalmologie » regroupe les termes « œil », « myopie », « hypermétropie », etc.

permettant ainsi de mettre en évidence les métatermes partagés ou non entre sous-ensembles de documents. Cette mise en relief peut alors être particulièrement utile dans une tâche d'indexation de ces groupes d'articles. Ici l'indexation peut être guidée principalement par la dimension intertextuelle des documents plutôt qu'uniquement par leur contenu.

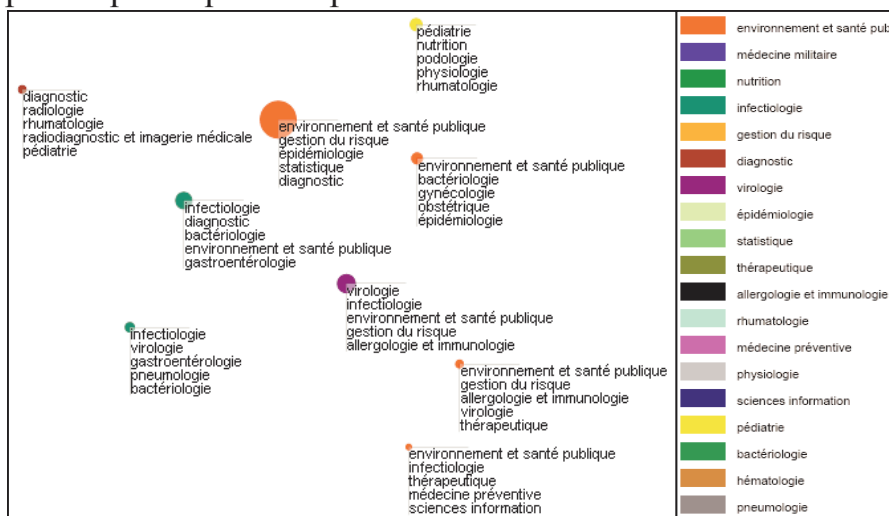


Figure 4: Carte thématique obtenue à partir d'un corpus d'articles médicaux et de ressources terminologiques spécialisées dans ce même domaine.

6. Conclusion

Les expériences présentées ci-dessus constituent des premières tentatives d'exploitation de traces de la dimension globale d'un corpus ou d'un flux documentaire pour les méthodes d'accès au contenu des documents. D'autres tentatives sont en cours de réalisation. Notamment, nous cherchons à exploiter l'importance relative d'une isotopie par rapport à une autre au sein d'un document. L'idée étant de donner (au moyen d'une pondération relative) plus d'importance aux isotopies du document qui sont directement liées à des caractéristiques globales de son intertexte, soit au niveau du groupe de documents dans lequel il figure, soit au niveau du corpus dans son ensemble.

Ces quelques recherches de traces de la dimension globale sont encore au stade de la modélisation et de l'implémentation au sein de l'outil ProxiDocs. Les enjeux qui font suite à l'implémentation consistent à en mettre en place d'autres ayant pour but de les évaluer. On pourra alors à savoir qu'est-ce qu'une « bonne » trace de l'intertextualité et de façon duale qu'est-ce qu'une « fausse » trace si tenté que cela existe. On pourra également chercher à caractériser en quoi les méthodes statistiques utilisées permettent de fournir des traces pertinentes de ce qu'elles représentent.

Ce type de questionnement scientifique montre en tout cas que l'amélioration des compétences sémiotiques et interactionnelles des machines constitue un domaine de recherche à part entière au croisement de plusieurs disciplines et pas uniquement un domaine d'ingénierie. En cela, à l'encontre de Rastier (Williams

& al. 2005, p. 41), nous militons pour une scientificité propre des traitements sémantiques et plus largement des traitements automatiques des langues.

Références

(Beust 1998) Beust P., *Contribution à un modèle interactionniste du sens*. Thèse d'Informatique de l'université de Caen, France, 1988.

(Courcil 2000) Coursil, J., *La fonction muette du langage*, Ibis Rouge Éditions, presses Universitaires Créoles, Guadeloupe, 2000.

(Déjean 1998) Déjean H., *Concepts et algorithmes pour la découverte des structures formelles des langues*, Thèse d'Informatique de l'université de Caen, France, 1988.

(Nicolle 2005) Nicolle N., *Comparaison entre les comportements réflexifs du langage humain et la réflexivité des langages informatiques*, 12^e journées de Rochebrune.

(Perlerin 2004) Perlerin V., *Sémantique légère pour le document*, Thèse d'Informatique de l'université de Caen, France, 2004.

(Rastier & al. 1994) Rastier F. avec la collaboration de M. Cavazza et A. Abeillé, *Sémantique pour l'analyse, de la linguistique à l'informatique*, Paris-Milan-Barcelone : Masson, 1994.

(Rastier 1987) Rastier F., *Sémantique interprétative*. Presses Universitaires de France, Paris, 1987.

(Rastier 1998) Rastier F., Le problème épistémologique du contexte et le problème de l'interprétation dans les sciences du langage, *Langages*, n°129, pp.97-111.

(Roy 2005) Roy T., *Une plate-forme logicielle dédiée à la cartographie thématique de corpus*, RECITAL 2005, Actes de la conférence TALN 2005, Dourdan (France), juin 2005

(Thlivitit 1998) Thlivitit T., *Sémantique Interprétative Intertextuelle*. Thèse d'Informatique de l'université de Rennes I, 1998.

(Williams & al. 2005) sous la direction de Williams J., *La linguistique de corpus*, Presses universitaires de Rennes, avril 2005, ISBN 2-7535-0046-0.

Rapports entre traces et signes dans l'activité de langage : Exemples de décryptages linguistiques assistés par ordinateur

Pierre-Sylvain Luquet - Emmanuel Giguet
{Pierre-Sylvain.Luquet, Emmanuel.Giguet}@info.unicaen.fr
GREYC - CNRS UMR 6072
Université de Caen, Bd. Maréchal Juin
14032 Caen Cedex, France
Tél.: +33 (0)2 31 56 73 {98 – 97}

Résumé

Dans l'activité de langage la première tâche est de distinguer dans l'environnement sonore ce qui relève de la parole du reste des sons. Il s'agit donc d'analyser le son, c'est-à-dire de relever les traces permettant d'isoler les sons de la langue. Les sons ainsi identifiés prennent alors le statut de phonèmes. Ce ne sont plus simplement des sons mais des signes qui, organisés en chaînes, permettent l'accès au lexique. Cette activité de décodage des traces de notre environnement fait partie du processus d'interprétation, associé à tout dispositif de perception. Dans le cas de la lecture, il faut identifier les caractères du reste des éléments perçus pour accéder à la chaîne faisant sens.

Le repérage de traces dans un flux d'informations permettant d'en isoler les signes est générique et systématique. Dans cet article, nous transposons ce problème au traitement automatique des langues. Nous nous intéressons à deux cas complémentaires : l'un concernant le traitement de la parole, l'autre concernant le traitement de corpus écrits. Dans le premier cas il s'agit de retrouver dans le flux audio la présence d'une catégorie acoustique, en l'occurrence la nasalité. Dans le second cas, nous allons chercher à retrouver dans un corpus multilingue des équivalences sémantiques entre segments de différentes langues. Dans les deux cas, il s'agira d'identifier les traces des signes recherchés dans le flux analysé.

Mots clefs

Traces, signes, phonétique, phonologie, lexique multilingue, motifs récurrents, isotopie.

1. Modèle de langage différentiel

Pour Saussure, le langage repose sur un principe d'organisation. Autrement dit, il ne s'agit pas d'un système de signes à proprement parler, mais d'un système de rapport entre valeurs. En outre, les rapports entre valeurs se limitent à des

rapports entre valeur de même type. C'est pour cette raison qu'il nommera ces valeurs « valeurs pures » (Saussure 86). Ainsi, dès l'origine de la description que Saussure fait du langage, il nous plonge dans un système. Dans ce système, ce sont les positions de ses constituants et les relations qu'ils entretiennent entre eux qui créeront le sens, et ce quelque soit l'échelle, le grain, à laquelle se place l'analyse. Que nous nous intéressions à la phonologie, au lexique, à la syntaxe, à la sémantique rien ne porte de sens de façon autonome, celui-ci n'est construit que par *différence* avec d'autres entités de langage avec lesquelles il est en relation. Ainsi, le sens, le langage, n'existe que dans le mouvement et dans l'instant où il est élaboré par son énonciation, c'est-à-dire à l'instant où les relations entre constituants sont lues, interprétées et éventuellement actualisées par la perception même de leur mise en oeuvre dans un contexte nouveau¹.

Cette question de la différence porte en elle-même la question du rapport entre la trace et le signe. Le sujet parlant, et d'une manière générale la communauté parlante a pour activité de transformer les traces en signes, la différence est le moyen. Cette activité peut être modélisée comme une fonction de transformation d'un domaine de départ **D** qui est celui du monde réel² vers **D'** domaine d'arrivée qui est la sphère du langage. Cette fonction repose sur une double capacité :

1. l'assignation (d'une trace à un signe³),
2. la distanciation (entre les signes).

Cette seconde capacité de la fonction de transformation est la garantie que le système de la langue ne s'effondre pas sur lui-même malgré sa nature réflexive (Mauger 2005). C'est un mécanisme de contrôle qui nous permet de ne pas toujours assigner un nouveau signe à un même objet. Autrement dit, s'il n'y a pas de nouvelle différence alors nous ne réalisons pas de nouvelle assignation.

2. TAL, traces et signes

Les logiciels d'étude et de recherche en Traitement Automatique des Langues manipulent des représentations d'objets de la langue dans le but d'en étudier les propriétés afin de valider une hypothèse. L'hypothèse repose sur des indices que nous avons détectés ou que nous supposons être dans le matériau linguistique étudié, et ces indices constituent la trace à étudier. Si l'hypothèse est validée par l'expérience alors les traces ont bien le rôle assigné par le chercheur et ne sont plus anonyme mais constitue un signe.

L'activité de modélisation consiste donc à identifier manuellement ou automatiquement un faisceau d'indices constituant la trace du signe recherché, puis à implémenter une fonction réalisant l'exploitation de cette trace. L'ensemble de départ **D** est constitué du matériau étudié ; l'ensemble d'arrivée

¹ Y compris en situation de soliloque.

² Nous désignons ici le monde que nous percevons directement ou indirectement et ne posons pas la question de la matérialité des choses.

³ Signe qui néanmoins sera le plus souvent protéiforme

D' est une représentation du matériau. La fonction réalise un calcul sur la base d'une hypothèse posant la relation entre le faisceau d'indices et le signe.

3. Nature de la transformation

Nous présentons dans cet article deux exemples de fonction de transfert de la trace vers le signe de nature tout-à-fait différente. La première, le décodage acoustique, a pour but de passer d'une trace physique à un signe représentant cette trace par apprentissage supervisé de la différence entre signaux acoustiques. La seconde, la constitution automatique de lexique multilingue, exploite une autre forme que peut prendre la trace : la répétition de séquences dans une collection d'objets.

Nous verrons dans les deux cas comment, par le biais d'un modèle informatique adéquat nous appréhendons ces traces et parvenons à les exploiter dans la construction de nos systèmes.

3.1. Reconnaissance de la trace

Dans le cas du décodeur acoustique, la trace à identifier dans le signal est décrite en termes phonologiques et phonétiques. Les membres de l'école de Prague ont réalisé une classification des différences des sons de la langue. Néanmoins, si nous savons différencier une voyelle orale de son équivalent nasal (par exemple l'opposition /a/ vs /a~/) la mesure et la description analytique de cette différence en termes physiques est difficile voire impossible si la population étudiée est trop importante, les variabilités entre individus étant trop importantes. Autrement dit, nous ne disposons pas d'un modèle formel de la trace et de ses indices mais d'une description différentielle de plus haut niveau.

Cette description nous conduit à poser l'hypothèse suivante : il existe dans le signal un faisceau d'indices physico-acoustique permettant l'apprentissage d'un modèle statique de la trace opposant les échantillons de signal appartenant ou n'appartenant pas à une catégorie phonétique donnée.

A partir de cette hypothèse nous cherchons à acquérir la trace par apprentissage différentiel supervisé. La partition du corpus d'apprentissage en sous-ensembles opposant des échantillons de signal opposés par leur catégorie constitue le faisceau d'indices sur lequel le classifieur doit se focaliser.

3.2. Construction de la trace

Dans le cas de la construction automatique de lexiques multilingues à partir de bitextes de la communauté européenne, la recherche de la trace s'est inscrit dans le cadre d'une démarche expérimentale, par la formulation d'une hypothèse initiale et son affinement par confrontation avec le corpus. Le modèle de la trace se présente sous la forme d'un faisceau d'indices morpho-positionnels.

Nous disposons d'une grande quantité de textes traduits et alignés et cherchons à localiser des relations de traductions entre segments sous-phrastiques (mots,

termes, collocations, ...). Nous faisons l'hypothèse que deux segments sont en relation de traduction, et correspondent donc à un seul signe, s'ils laissent une trace équivalente dans les sous-corpus monolingues.

Le modèle de la trace que nous exploitons actuellement est le suivant :

- un terme répété plusieurs fois dans un document est généralement traduit par un même équivalent dans la traduction, à des positions équivalentes et aux variantes morphologiques près ;
- un terme présent dans un document n'a pas obligatoirement une traduction unique dans le corpus. Une exploitation de ce modèle par la machine va nous permettre de repérer et de construire les traces formant les hypothèses de bijection sémantique. Ces hypothèses doivent ensuite être triées, automatiquement pour les cas simples, avec supervision humaine pour les cas plus ambigus.

4. Expériences

4.1. Décodage acoustique

Les décodeurs acoustiques, tels que ceux développés au LIMSI (Lamel & Gauvain, 1993), atteignent des taux de reconnaissance proche des 75%. Cependant, les limitations restent nombreuses et la critique la plus largement formulée est la quasi absence de connaissances sur le langage dans les modèles sous-jacents (Plaut & Kello, 1999). Les travaux actuels s'articulent autour de deux axes. Le premier s'intéresse à l'amélioration des techniques de description du signal (Chetouani *et al.*, 2002). Le second est orienté vers la production : acquisition de connaissances concernant les gestes articulatoires des locuteurs (Vaxelaire *et al.*, 2002), leurs influences sur le signal (Montagu, 2002), et les processus cognitifs mis en jeu (Hawkins, 2003). Ces connaissances font l'objet de différentes études visant à leur intégration dans les systèmes de reconnaissance automatique de la parole (Wrench & Richmond, 2000). Nous décrivons dans ces lignes une approche sensiblement différente qui s'appuie sur la *notion de différence*, charpente du langage, et son instanciation phonologique. Saussure affirme que « *dans la langue, il n'y a que des différences [. . .] sans termes positifs* » (Saussure, 1986). Coursil reprend à son compte cette notion dans (Coursil, 1992)₁ et affirme à son tour « *Pour tout phonème x, il existe un phonème y tel que y = x à une et une seule différence catégorique près* ». C'est à partir de cette dernière affirmation qu'il construit la *topique des phonèmes du français contemporain*, topique sur laquelle nous appuyons notre méthode de recherche de la trace dans le signal.

Phonologie et topique des phonèmes du français contemporain. La topique₂ des phonèmes du français contemporain décrite dans (Coursil *et al.*, 2000) et (Coursil, 1992) peut être vue comme un graphe dans lequel chaque noeud est un phonème et où seuls les phonèmes distincts par une et une seule caractéristique sont reliés. Chaque lien porte une étiquette, le label de cette étiquette est le nom

d'un trait, et plus précisément le nom du trait dont la valeur diffère entre deux phonèmes liés.

La topique forme donc un espace à 7 dimensions, chaque dimension étant une catégorie phonologique : laxité, hauteur, compacité, voisement, stridence, nasalisation et bémolisation³. Les phonèmes sont positionnés dans cet espace, et leurs coordonnées sont définies par les valeurs associées à chacune de leur catégorie. Il devient alors possible de mesurer la distance entre deux phonèmes dans cet espace. Cette distance est la différence garantissant la stabilité du système et dont nous devons retrouver la trace dans le signal.

Dans la série d'expériences proposées ci-après, nous nous concentrons sur le trait de nasalité. Du point de vue de la mécanique articulatoire, la nasalité est décrite comme une connexion du conduit vocal avec le conduit nasal par le biais de l'abaissement du vélum. Pour Fant dans (Jakobson, 1980) les consonnes nasales sont « caractérisées par un spectre où F2 est faible ou bien absent »; Delattre affine la description en précisant que pour les voyelles nasales, comparées aux orales, F1 perd une bonne part de son intensité au profit de F2. Plus récemment, (Feng & Kotenkoff, 2004) ont mené des observations basées sur une technique d'enregistrement du locuteur en différenciant les prises de son en provenance du conduit vocal et du conduit nasal. Ils ont constaté que l'abaissement du vélum a deux effets distincts : pour le conduit vocal le rétrécissement engendre le rapprochement des formants F3 et F4, et pour le conduit nasal sa connexion entraîne un rayonnement au niveau des narines caractérisé par une concentration dans les basses fréquences et aux alentours de 3000 Hz.

4.1.1. Corpus et paramètres

Corpus. Nous faisons ici l'hypothèse que le corpus présentant le moins de difficultés pour réaliser la partition oralisé vs. nasalisé est constitué de paires minimales oral vs. nasal. Pour réduire encore la difficulté, nous nous sommes concentrés sur les phonèmes dont la production pouvait être tenue. Aussi nous avons construits trois premiers corpus (voire Tableau 1⁴) réalisant une gradation de la difficulté de focalisation du classifieur sur la trace à isoler. Le premier corpus, **C1** est constitué d'une seule paire minimale (/o~/ & /o/), dont la seule variation est le trait de nasalisation⁹. Les corpus **C2** et **C3** sont plus complexes : sur les 7 caractéristiques mises en jeu 4 varient. Pour les orales (/o/ & /a/) les variations portent sur la laxité, la compacité et la bémolisation ; la hauteur intervient en plus pour les nasales (/o~/ & /a~/)⁵.

⁴ Code des catégories : L = Laxité, H = Hauteur, C = Compacité, V = Voisement, S = Stridence, N = Nasalisation et B = Bémolisation

⁵ Les quantités du Tableau 1 (nombre de phonèmes et nombre de fenêtres) données ici correspondent pour chaque corpus aux versions de test. Les versions d'apprentissage sont du même ordre de grandeur

| | Phonèmes | Nb. Phonèmes | Nb. fenêtres | Maintenus | Variations |
|----|-----------------|--------------|--------------|-----------|---------------|
| C1 | /o~/-/o/ | 22 | 2250 | oui | N |
| C2 | /o~, a~/-/o, a/ | 20 | 2000 | oui | N, L, C, B, H |
| C3 | /o~, a~/-/o, a/ | 24 | 470 | non | N, L, C, B, H |

Tableau 1

Outils et paramètres. L'outil principalement utilisé dans cette expérience est le logiciel d'analyse acoustique PRAAT. Nous l'utilisons aussi bien pour les phases de préparatoires que pour l'apprentissage et les tests de nos classifieurs. Chaque signal de parole à analyser est segmenté en tranches de 30ms avec un décalage entre tranche de 10ms, à chaque tranche est appliqué une fenêtre de Hamming. Chaque tranche de signal fenêtré constitue un vecteur labellisé ; le label étant la caractéristique acoustique du signal. Dans le cas présent nos vecteurs sont labellisés "nasal" ou "oral" suivant le phonème à partir duquel ils ont été générés.

Nos classifieurs sont générés par l'application PRAAT et nous utilisons des réseaux de neurones type perceptron à une couche cachée. Lors des phases d'apprentissage la méthode d'évaluation d'erreur (et d'arrêt du cycle d'apprentissage) est calculée suivant la méthode *minimum squared error* avec une tolérance de $1e^{-7}$.

4.1.2. Résultats

Nous présentons les résultats d'une série de trois expériences. La première met en oeuvre une variation phonétique restreinte puisque le corpus est limité à une paire minimale de phonèmes (/o/ et /a/). La deuxième expérience cherche à valider les premiers résultats en augmentant le nombre de caractéristiques qui varie entre phonèmes (nous opposons /a, o/ à /a~, o~/). Enfin, la troisième expérience reproduit la seconde mais sur des phonèmes non tenus.

Variation restreinte. Les résultats donnés ici (Tableau 2) concernent le corpus C1 et ont été obtenus au terme d'un apprentissage de 400 cycles. Les phonèmes sont tenus et la seule variation phonologique mise en jeu est la nasalisation. Nous obtenons un taux d'erreurs faible (2,8%), mais nous voyons surtout que le nombre de fenêtres incorrectement classifiées de façon continue est très faible en regard du nombre de fenêtres par phonème. Le risque de mal classifier un phonème est donc très faible.

| | | | |
|------------------------------|------|----------------------|------|
| fenêtres | 2252 | fenêtres par phonème | 102 |
| erreurs | 63 | groupes d'erreur | 17 |
| taux | 2,8% | erreurs par groupe | 3,71 |
| erreurs consécutives maximum | | | 8 |

Tableau 2

Augmentation de la dissemblance. Les résultats donnés dans le Tableau 3 concernent le corpus C2. Ils sont obtenus après 600 cycles d'apprentissage. Comme précédemment les phonèmes sont maintenus mais plusieurs variations

phonologiques sont ici mises en jeu. En effet l'opposition /a, o/ - /a[~],o[~]/ met en jeu les caractéristiques de laxité, hauteur, compacité et bémolisation. La focalisation du classifieur sur la caractéristique de nasalisation est donc rendue plus complexe en raison du bruit apporté par les autres variations. Le taux d'erreurs est de 2,3% et le nombre de fenêtres contiguës mal classifiées se limite à 9. Le risque de mal classifier un phonème reste donc très faible.

| | | | |
|------------------------------|------|----------------------|-----|
| fenêtres | 1996 | fenêtres par phonème | 100 |
| erreurs | 47 | groupes d'erreur | 16 |
| taux | 2,3% | erreurs par groupe | 2,9 |
| erreurs consécutives maximum | | | 9 |

Tableau 3

Phonèmes non tenus. L'expérience menée ici est similaire à l'expérience précédente (corpus C2), mais concerne des phonèmes non maintenus. En revanche nous opposons toujours les couples de phonèmes /a, o/ et /a[~],o[~]/, et conservons donc toujours les mêmes dissemblances et la même difficulté de focalisation pour le classifieur. Les résultats obtenus (Tableau 4) restent très intéressants. Au terme d'un apprentissage de 600 cycles, nous observons un taux d'erreur de 15,8%. Le nombre d'erreurs est de 7 fenêtres mal classifiées. Néanmoins ce nombre reste acceptable, dans le cas d'une stratégie de classification *winner-takes-all* dans la mesure où un phonème compte en moyenne 20 fenêtres.

| | | | |
|------------------------------|-------|----------------------|-----|
| fenêtres | 469 | fenêtres par phonème | 20 |
| erreurs | 74 | groupes d'erreur | 30 |
| taux | 15,8% | erreurs par groupe | 2,5 |
| erreurs consécutives maximum | | | 7 |

Tableau 4

4.2. Génération automatique de lexique multilingue

Ce second exemple traite de la génération automatique de bases lexicales multilingues. L'objectif est la localisation et l'extraction automatique de relation de traduction entre segments sous-phrastiques, à partir d'un corpus de textes traduits.

Dans la section 3.2, nous avons présenté la trace actuellement étudiée. La mise en œuvre du critère morphologique se traduit par la recherche de motifs récurrents : des segments subphrastiques répétés apparaissant avec une fréquence égale ou très proche dans les deux traductions. Le critère positionnel, quant à lui, pondère le critère morphologique en exigeant que les segments répétés apparaissent à des positions comparables dans les deux documents. Ce critère positionnel est mis en œuvre à l'aide d'une technique d'alignement des

textes au grain phrastique. Plusieurs algorithmes tels que celui de (Gale & Church 1991) existent pour obtenir un tel alignement.

Dans notre dernier modèle de la trace, il est à noter que le critère morphologique a subi un affinement. Précédemment, le modèle requérait des segments répétés strictement à l'identique au sein d'un même texte, nous avons relâché cette contraintes en imposant des segments répétés, aux variantes morphologiques près. Ces variantes morphologiques sont calculées de manière endogène, à l'aide d'un algorithme dérivé des travaux de (Déjean 1998).

4.2.1. Méthode de résolution

La résolution se déroule en deux étapes basées sur deux hypothèses sous-jacentes. La première porte sur les documents et la seconde sur l'ensemble du corpus.

- **Hypothèse 1** : soit un bitexte composé de T_1 et T_2 . Si une séquence S_1 est répétée dans T_1 dans un certain nombre de phrases P_{1i} , nous supposons qu'une séquence S_2 correspondant à la traduction de S_1 apparaissent dans les phrases P_{2j} de T_2 où P_{2j} sont les phrases alignées avec P_{1i} .
- **Hypothèse 2** : soit un corpus de bitextes composé des langues L_1 et L_2 . Il n'y a pas de garantie pour qu'une séquence S_1 répétée dans plusieurs textes de L_1 ait une unique traduction dans les textes correspondant de la langue L_2 .

4.2.2. Résolution : 1 – isoler la trace.

La première étape se déroule à l'échelle des documents. Elle consiste à relever un certain nombre d'indices dans le flux de document. Ils forment les traces de notre processus.

La détermination des séquences répétées se déroule suivant un algorithme glouton basé sur l'extraction de termes décrite dans (Vergne, 2005). L'algorithme est paramétré avec les longueurs minimale et maximale d'une séquence, ainsi que la fréquence minimale d'apparition d'une séquence pour la garder. Nous ne retenons pas les sous séquences lorsqu'elles sont aussi fréquentes que la séquence dont elles sont issues. Pour chaque séquence retenue, nous mémorisons ses positions dans le document.

Notons que nous prenons en compte la flexion des termes de nos séquences sans pour autant utiliser de ressources externes au document de type lexique ou parseur. La méthode endogène utilisée s'inspire des travaux de (Déjean, 1998) et permet de calculer aussi bien les préfixes que les suffixes.

Chaque séquence est ensuite vectorisée dans un espace orthonormé de dimension n , où n est le nombre de segment du texte d'une langue donnée. La valeur de chaque composante du vecteur est le nombre de répétition de la séquence dans le segment correspondant.

Soit une séquence apparaissant dans les segments 45, 46, 48, 51, 51 d'un texte de n segments, le vecteur correspondant sera alors le suivant :

| | | | | | | | | | | | |
|---|---|-----|----|----|----|----|----|----|----|-----|-----|
| 1 | 2 | ... | 45 | 46 | 47 | 48 | 49 | 50 | 51 | ... | n |
| 0 | 0 | | 1 | 2 | 0 | 1 | 0 | 0 | 2 | | 0 |

Nous procédons ensuite à l'alignement des séquences entre les deux langues des bitextes et proposons des relations de traduction entre séquences. Cet alignement est calculé à partir du cosinus des vecteurs associés à chaque séquence. Une valeur de seuil sur le scalaire obtenu nous permet de réaliser un premier filtrage sur les mises en relation entre séquences.

4.2.3. Résolution : 2 – confirmer l'hypothèse.

A l'issue de première phase de résolution à l'échelle des documents, nous disposons d'un certain nombre d'hypothèses de relation de traduction entre des segments d'une langue vers une autre. Le passage à l'échelle du corpus va nous permettre de confirmer ou d'infirmer les hypothèses précédemment posées. La confirmation d'une hypothèse est ici vue comme le passage du statut de trace à celui de signe. Autrement dit, une séquence est considérée comme équivalente à une autre dans le système sémiotique mis en place.

Pour qu'une hypothèse résiste au filtrage au niveau du corpus elle doit respecter l'un des deux critères suivant :

1. avoir été extraite dans deux bitextes
2. si elle n'a été trouvée que dans un bitexte alors les deux séquences doivent être composées des mêmes chaînes de caractères ou si sa fréquence est supérieure à un seuil déterminé empiriquement.

Les relations de traductions sont ensuite classées par rang suivant le produit de la fréquence de leur apparition et du nombre de documents desquels elles ont été extraites.

4.2.4. Résultats

Nous proposons en annexe les résultats produits à partir de 20 bitextes français – anglais issus du corpus « Acquis Communautaire ». Parmi les propositions correctes, nous avons trouvé des séquences dépendantes du domaine considéré :

- Termes réglementaires de la CEE :
EEC initial verification ↔ *vérification primitive CEE*,
Regulation (EEC) No ↔ *règlement (CEE) n°*.
- Autres lexiques :
rear-view mirrors ↔ *rétroviseurs*,
poultry ↔ *volaille*.
- des séquences invariantes :
km/h ↔ *km/h*, *kg* ↔ *kg*, *mortem* ↔ *mortem*.

Nous avons également des hypothèses de traduction correcte concernant des séquences de taille variée :

territory ↔ *territoire*, *fresh poultrymeat* ↔ *viandes fraîches de volaille*,
Member States ↔ *États membres*, *Whereas* ↔ *Considérant que*,
Having regard to the Opinion of the ↔ *vu l'avis*.

La plupart des alignements faux ont été extraits d'un seul bitexte ce qui nous incite à affiner nos seuils de confirmation d'hypothèse :

on ↔ *la commercialisation des*.

5. Conclusion

Au travers de deux tâches d'analyse automatique des langues, l'une relative au traitement de l'oral et plus précisément à la classification de signal, et l'autre relative au traitement de l'écrit et portant sur l'alignement automatique, nous avons montré l'omniprésence de l'interaction entre trace et signe dans notre activité de modélisation linguistique.

Nous avons particulièrement insisté sur la similarité des méthodes sous-jacentes à la conception de ces logiciels d'étude et de recherche, laissant percevoir l'existence d'une méthode de résolution abstraite et générique. Cette méthode abstraite de reconnaissance de forme s'appuie sur la caractérisation préalable d'un faisceau d'indices constituant la trace du signe recherché, et sur l'implémentation d'une fonction réalisant l'exploitation de cette trace afin de mettre à jour les signes visés.

La caractérisation des hypothèses quel que soit la tâche repose sur la pré-existence d'un modèle linguistique à valider. Dans le cas du traitement de l'écrit, la description de la trace repose sur une description formelle de nature symbolique. Dans le cas du traitement de l'oral, cette description est acquise automatiquement par un processus d'apprentissage différentiel.

Références

- CHETOUANI M., GAS B. & ZARADER J. (2002). Coopération entre codeurs neuro-prédictifs pour l'extraction de caractéristiques en reconnaissance de phonèmes. In *Reconnaissances des formes et intelligence artificielle*.
- COURSIL J., HILAIRE J.-C., MONTLOUIS-CALIXTE D. & REMI C. (2000). *Projet anadia*.
- COURSIL J. (1992). *Essai d'intelligence artificielle et de linguistique générale*. PhD thesis, Université de Caen.
- DÉJEAN H. (1998). Morphemes as Necessary Concept for Structures Discovery from Untagged Corpora. In *Workshop on Paradigms and Grounding in Natural Language Learning*, p. 295-299.
- FENG & KOTENKOFF (2004). Vers un nouveau modèle acoustique des nasales basé sur l'enregistrement bouche - nez séparé. In *Journées d'Etude sur la Parole*.
- GALE W.A. & CHURCH K. W. (1991). A Program for Aligning Sentences in Bilingual Corpora. In *Proc. 29th Annual Meeting of the Association for Computational Linguistics*, p. 177-184.
- HAWKINS S. (2003). Roles and representations of systematic fine phonetic detail in speech understanding. *Journal of Phonetics*.
- JAKOBSON R. (1980). *La charpente phonique du langage*. Paris : Editions de Minuit.

- LAMEL L. & GAUVAIN J. (1993). High performance speaker-independent phone recognition using cdhmm. In *European Conference on Speech Communication and Technology*.
- MAUGER S. & LUQUET P.-S. (2005). Réflexivité dans un processus d'interaction homme-machine Exemple de résolution de charades assistée par ordinateur. In *Journées de Rochebrune 2005*.
- MONTAGU J. (2002). L'articulation labiale des voyelles nasales postérieures du français : comparaison entre locuteurs français et anglo-américains. In *Journées d'Étude sur la Parole*.
- PLAUT D. C. & KELLO C. T. (1999). *The Emergence of Language*, chapter The Emergence of Phonology from the Interplay of Speech Comprehension and Production : A Distributed Connectionist. Lawrence Erlbaum Assoc : Mahwah.
- SAUSSURE F. (1986). *Cours de linguistique générale*. Paris : Mauro Payot.
- VAXELAIRE B., FERBACH-HECKER V. & SOCK R. (2002). La perception auditive de gestes vocaliques anticipatoires. In *Journées d'Étude sur la Parole*.
- VERGNE J. (2005). Une méthode indépendante des langues pour indexer les documents de l'Internet par extraction de termes de structure contrôlée. In *Conférence Internationale sur le Document Electronique*.
- WRENCH A. A. & RICHMOND K. (2000). Continuous speech recognition using articulatory data. In *International Conference on Spoken Language Processing*.

| Source | ndoc | freq | cos | cible |
|--|------|-------|------------------------------|-------|
| And | 12 | [336] | - et | |
| Member | 10 | [206] | - membre~ | |
| Member State~ | 10 | [201] | - membre~ | |
| Member States | 13 | [143] | - États membres | |
| The | 4 | [392] | - d~ | |
| Of | 5 | [313] | - de~ | |
| EEC | 9 | [118] | - CEE | |
| Directive | 11 | [95] | - directive | |
| State | 4 | [71] | - membre | |
| provisions | 8 | [35] | - dispositions | |
| Whereas | 10 | [28] | - considérant que | |
| Member State | 4 | [63] | - membre | |
| 2~ | 5 | [43] | - 2 | |
| the Council | 10 | [19] | - Conseil | |
| approv~ | 3 | [62] | - modèle | |
| requirements | 6 | [30] | - prescriptions | |
| mark~ | 6 | [30] | - marque~ | |
| 5 | 8 | [22] | - 5 | |
| EEC pattern approval | 4 | [35] | - CEE de modèle | |
| verification | 4 | [34] | - vérification | |
| Council Directive | 9 | [15] | - Conseil | |
| EEC initial verification | 5 | [27] | - vérification primitive CEE | |
| Having regard to the Opinion of the | 8 | [16] | - vu l'avis | |
| THE | 8 | [16] | - DES | |
| - | 4 | [32] | - - | |
| II | 6 | [20] | - II | |
| ; | 5 | [20] | - ; | |
| Regulation | 4 | [24] | - règlement | |
| of Annex I | 4 | [18] | - l'annexe I | |
| 5~ | 5 | [13] | - 5 | |
| 10 | 5 | [13] | - 10 | |
| maximum | 3 | [21] | - maximal~ | |
| instruments | 4 | [14] | - instruments | |
| instrument~ | 2 | [27] | - instrument~ | |
| I~ | 3 | [18] | - I | |
| separate | 2 | [26] | - particulier~ | |
| method~ | 3 | [17] | - méthode~ | |
| may | 5 | [10] | - peuvent | |
| seed~ | 1 | [48] | - semences | |

| | | | |
|--------------------------|---|------|--------------------------------|
| se~ | 1 | [46] | - semences |
| methods | 3 | [15] | - méthodes |
| seed | 1 | [45] | - semences |
| pattern~ | 1 | [44] | - modèle~ |
| by | 4 | [11] | - par |
| Of Annex | 3 | [14] | - point~ |
| 9 | 4 | [10] | - 9 |
| 14 | 3 | [13] | - 14 |
| provision~ | 3 | [13] | - dispositions |
| Tests | 2 | [19] | - essais |
| Committee | 2 | [18] | - Comité |
| poultrymeat | 1 | [36] | - de volaille |
| and methods of | 4 | [9] | - de contrôle |
| Regulation (EEC) No | 3 | [12] | - règlement (CEE) n° |
| OF | 6 | [6] | - DES |
| particular | 3 | [11] | - notamment |
| 19~ | 3 | [11] | - 19~ |
| certain | 3 | [11] | - certain~ |
| marks | 3 | [11] | - marques |
| mark | 4 | [8] | - la marque |
| directive | 2 | [16] | - directive particulière |
| trade | 2 | [16] | - échanges |
| pattern approval | 1 | [31] | - de modèle |
| pattern approval~ | 1 | [31] | - de modèle |
| 4~ | 5 | [6] | - 4 |
| approximat~ | 3 | [10] | - rapprochement |
| other | 3 | [10] | - autres que |
| for liquid~ | 2 | [15] | - de liquides |
| measure~ | 3 | [10] | - mesur~ |
| measuring instruments | 3 | [10] | - instruments de mesure |
| separate directive | 2 | [15] | - directive particulière |
| EEC pattern | 1 | [29] | - CEE de modèle |
| equipment | 2 | [14] | - dispositif~ |
| EEC pattern approval~ | 1 | [28] | - CEE de modèle |
| competent | 2 | [13] | - compétent~ |
| This Directive | 2 | [13] | - la présente directive |
| 26 July 1971 | 4 | [6] | - du 26 juillet 1971 |
| on | 1 | [23] | - la commercialisation des |
| fresh poultrymeat | 1 | [23] | - viandes fraîches de volaille |

| | | | |
|---------------|---|------|---------------------|
| EEC) No | 2 | [11] | - CEE |
| inform~ | 2 | [11] | - inform~ |
| the competent | 2 | [11] | - compétent~ |
| into force | 3 | [7] | - en vigueur |
| mass~ | 1 | [19] | - à l'hectolitre |
| type-approv~ | 1 | [19] | - CEE |
| than | 2 | [9] | - autres que |
| weight | 1 | [18] | - poids |
| amendments to | 2 | [9] | - les modifications |

Un modèle des traces pour l'interaction entre des agents

Anne Nicolle
GREYC UMR 6072 CNRS (équipe ISLanD)
Université de Caen

Résumé :

Les agents sont des processus interactifs de durée indéfinie (êtres vivants, robots, agents logiciels) qui ont des capteurs pour récupérer des informations dans leur environnement et des effecteurs pour y produire des événements, des flux ou des traces. Ces agents ont une apparence variable qui signale en partie de leur état interne, que les autres agents peuvent interpréter comme des traces. Les événements captés et produits par des processus sont soit ponctuels, soit organisés en flux. Ils ont des qualités (épaisseur, force, couleur, intensité, signature...) qui les distinguent les uns des autres. Ils peuvent laisser des traces persistantes qui pourront être interprétées en dehors du contexte où ils ont été produits. Leur mémoire peut garder une trace des interactions.

Cet article propose un modèle computationnel qui organise les notions d'événement, de flux, de trace, d'apparence. Il rend compte de leur utilisation dans la dynamique des interactions entre agents.

Mots-clés : agents, processus, événements, interaction, traces

1 Introduction

Les humains ont construit de nombreux types de machines qui instrumentent leurs interactions. Avec l'évolution technologique, les activités des machines et des robots sont gérées par l'informatique, là où des procédés mécaniques ont longtemps été suffisants. Cette évolution amène à décrire tous les processus artefactuels par des programmes informatiques, qui contrôlent le fonctionnement des machines et leurs interactions. Pour les étudier de manière générique, on appelle processus interactifs ou agents les êtres vivants, les machines et les robots, ainsi que les programmes interactifs qui tournent sur les ordinateurs. L'interaction entre les agents se produit quand les actions des uns affectent les autres, que ce soit intentionnellement ou non.

Avec la possibilité de faire exécuter les programmes en monde ouvert, et donc de créer des dynamiques formelles en interaction, apparaît un enjeu scientifique : rendre compte de la formalisation des processus par du logiciel exécuté par des processeurs alors que les mathématiques de la cybernétique créent seulement des structures et des fonctions dont l'interprétation est laissée aux humains. La condition de cette modélisation dynamique est la réhabilitation de la notion de cause finale, qui permet de passer d'un modèle descriptif à un modèle causal.

L'implantation informatique du modèle causal permet de vérifier qu'elles produisent bien un analogue formel des phénomènes étudiés [Nicolle, 2004a].

Peter Wegner argumente depuis plusieurs années pour le passage des théories de l'informatique du paradigme du calcul au paradigme de l'interaction [Wegner, 1997]. Il a défini des modèles théoriques pour ce paradigme [Wegner, 1998], rendant compte de la persistance de la mémoire des machines et du traitement des flux provenant d'un monde ouvert. Dans ce contexte des théories de l'interaction, des éléments pour une théorie des processus interactifs de durée indéfinie ont été proposés [Nicolle, 2004a]. La question du temps [Nicolle, 2004b] comme temps réel et comme construction de la mémoire peut être mise en relation avec la question des traces. Pour l'interaction synchrone au présent, les événements produits et captés par les agents sont seuls nécessaires. Pour les interactions asynchrones, ou tenant compte du passé, ou préparant le futur, les traces produites ou captées par les agents sont le support et l'enjeu de leur comportement. Les flux jouent un rôle intermédiaire.

2 Une théorie des processus interactifs

Elle prend en compte trois entités structurelles :

- les objets, qui contiennent des fonctions et des données,
- les processus, qui créent, utilisent et détruisent des objets,
- les systèmes, qui regroupent des processus en interaction.

Les systèmes peuvent aussi être vus de l'extérieur comme des processus. L'encapsulation du fonctionnement interne du système sépare l'interface publique des interactions du fonctionnement interne : c'est un analogue de la membrane des cellules ou de la peau des animaux. Avec la notion d'objets actifs, les processus peuvent être vus comme des objets ayant un comportement autonome.

Les processus interagissent en s'envoyant des événements. En informatique, un événement d'abord est un signal physique (frappe d'une touche, clic souris...) capté par un objet (fenêtre, bouton, case...) qui est abonné à ce type d'événement. L'objet a alors une méthode pour traiter cet événement : il peut changer d'état, l'envoyer à un autre objet, créer un nouvel objet... Mais les objets logiciels peuvent aussi s'envoyer des événements qui ne proviennent pas d'un signal physique (top d'horloge). Un événement est ponctuel mais il peut transporter des informations, comme son origine ou son intensité. Les événements de même type captés par le même objet peuvent former des flux. Un flux n'est pas un ensemble puisque l'arrivée des événements et leurs qualités ne sont pas connus à l'avance. On les note $F = e \bullet F$, où e représente l'événement actuel, \bullet la concaténation et F le flux. On ne peut en avoir qu'une connaissance statistique (domaine, loi de probabilité). Les flux (bruits, lumière) sont duratifs mais ils doivent être entretenus par de nouveaux événements pour maintenir leur existence dans le temps. Les flux peuvent être des flux de matière (odeurs, rivières), des ondes (sons, lumière), des flux d'objets (voitures sur une route, gens dans une foule). Ils peuvent être discrets ou continus. Les flux continus peuvent être étudiés en positionnant une fenêtre qui les discrétise. Les flux discrets peuvent être tamponnés, un nombre fini d'éléments du flux est alors mémorisé tant que les suivants ne les remplacent pas.

Les événements et les flux peuvent produire des traces. Comme exemple de traces, on peut donner les traces de pas, l'écriture, les partitions musicales ou les dessins. Comme exemple en informatique, dans un traitement de textes, les touches tapés par l'utilisateur forment un flux de caractères interprétés par le logiciel pour former un texte et le mettre en forme. Le résultat est une trace du processus d'écriture sur l'écran qui sera peut-être conservé dans un fichier si l'utilisateur produit un événement déclenchant l'enregistrement. Une trace externe est produite par un flux ou par un mouvement, elle nécessite une matière pour se réaliser, mais c'est quelquefois sa forme qui la caractérise, on peut alors la reconnaître sur une photo. Les traces sont permanentes à l'échelle d'observation où elles sont considérées comme des traces. Les traces physiques disparaissent progressivement : évaporation, usure... Suivant l'échelle d'observation, on peut ignorer ce phénomène ou considérer la trace comme un processus dont on observe l'apparence variable. Nous emploierons quelquefois flux comme hyperonyme de flux et de traces quand il n'y a pas de raison pour les distinguer. Un processus perçoit un autre processus via les événements qu'il lui envoie et via son apparence, qui assure son identité et signe ses manifestations. Comme exemple d'apparence, on peut donner la position du corps et les mimiques des êtres vivants, les vêtements et le maquillage des humains, les interfaces graphiques sur les écrans des ordinateurs. L'apparence d'un processus est ce qui est observable d'un processus par les autres processus du fait de leur appareil cognitif, sans envoi de message explicite. L'apparence est observée comme un flux, elle peut laisser des traces comme la signature, les enregistrements audio, photo ou vidéo.

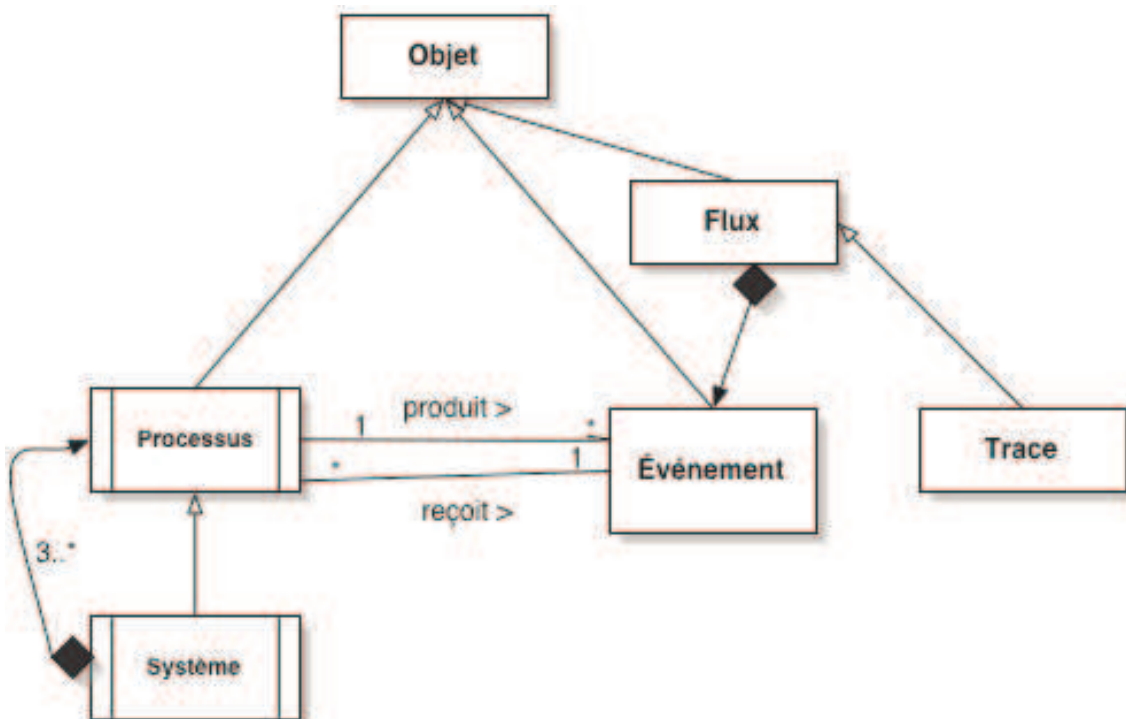


Figure 1 - Relations entre processus, systèmes, flux et traces

On obtient donc le diagramme UML¹ ci-dessus, qui rend compte des relations entre les trois entités structurelles citées et les trois entités interactionnelles que sont les événements, les flux et les traces.

On appelle environnement d'un processus les autres processus avec lesquels ce processus peut interagir et l'espace comme lieu de transmission des événements, des flux et comme lieu d'inscription des traces. L'espace et le temps n'apparaissent pas dans la figure 1 parce qu'ils ne sont pas représentés. Ils interviennent par l'intermédiaire des attributs des objets comme leur position, comme des événements (top d'horloge, rencontre) ou comme des flux (signaux GPS de position d'un mobile).

3 Modèle des interactions

Les processus interactifs peuvent capter et produire des événements, des flux et des traces persistantes ayant des attributs observables (force, intensité, couleur...). Toute interaction passe par un signal physique produit par un ou des agents, transmis à travers un canal de communication et reçu par un ou des agents. Les agents ne captent pas toutes les traces présentes dans leur environnement, pas plus qu'ils ne perçoivent comme événements tous les signaux qui sont émis. L'interaction entre des agents vivants suppose qu'ils aient les moyens biologiques de produire et de capter certains signaux. Capturer nécessite un appareil cognitif approprié et une attention tournée vers ce type de signal. À l'état de sommeil, on capte moins de signaux et on ne fait pas attention aux traces. La disposition d'écoute est donc nécessaire pour que l'interaction ait lieu, ainsi que la connaissance du sens des signes pour les interpréter.

Quel est l'équivalent de ces dispositions pour les agents logiciels ? Les agents purement mentaux, comme les agents BDI interprètent les interactions dans un monde représenté, donc sans distinction entre les dimensions contraignant les effets de leurs actions sur le monde. Par contre, les agents situés, en prise sur un monde ouvert, doivent tenir compte de la manière dont le monde est structuré : monde physique, monde social, monde propre car les dimensions physique, biologique et sociale forment la substance et l'enjeu des interactions. Un exemple permet de montrer comment elles contraignent la réussite des interactions : par la parole, je peux ouvrir la séance mais je ne peux pas ouvrir la porte (sauf dans les contes "Sésame, ouvre toi !"). Il faut que quelqu'un ou quelque chose interprète mes paroles et réalise l'action physique nécessaire à l'ouverture d'une porte.

- La dimension physique est constituée par le temps présent (ou temps réel), l'espace, les objets matériels et leurs transformations, les substances qui composent les objets, les champs de force auxquels les objets sont soumis, les signaux. En particulier, toute trace a une dimension physique qui donne forme à une substance.

- La dimension sociale est constituée par les relations entre les agents, et par leur cristallisation dans des institutions :

¹ www.uml.org

- les systèmes de signes dont les langues naturelles,
- les organisations, les rôles, les normes,
- les valeurs partagées : "liberté, égalité, fraternité"... la culture et la connaissance,
- les projets, stratégies et tactiques de groupe
- le temps social (horloges, calendriers).

Les traces ont une dimension sociale quand elles nécessitent la connaissance de conventions pour être interprétées : l'écriture en est un exemple, mais aussi les arts, les cartes et les plans.

- La dimension du soi, ou monde propre est une dimension biologique et psychologique pour les agents vivants. Elle est d'abord une dimension biologique qui distingue l'agent de son environnement et qui est en jeu dans le maintien de sa viabilité. Elle se constitue ensuite comme dimension psychologique par les relations de l'agent au monde : aux sensations, aux émotions et aux actions qui construisent sa mémoire, ses perceptions et ses décisions. Elle est aussi une dimension mentale pour les agents communiquant par la parole, qu'ils soient humains ou artificiels. Pour les agents artificiels, cette dimension du soi est assurée par la réflexivité, quand elle existe.

Ces trois dimensions du monde se co-construisent. Aucune n'est détachable des autres, mais cette distinction est utile car elle permet d'expliquer la dynamique des interactions et le statut des traces. Les objets techniques sont un bon exemple de la non séparabilité des dimensions du monde, objets matériels assurément, ils peuvent être détournés de l'usage prévu². Ce sont aussi des objets sociaux, qu'on invente, qu'on construit, qu'on achète, et leur propriétaire peut les considérer comme faisant partie de son monde propre (ma maison, ma voiture, mes skis...). Ce sont des traces de l'activité humaine et quelquefois aussi des traces qui prolongent l'apparence de leur possesseur.

3.1 Les agents

Les agents ont une connaissance du monde physique et du monde du soi par leurs sensations et par les effets observables de leurs actions. Leurs interactions construisent le monde social et reconstruisent le monde dans les deux autres dimensions comme un monde partagé entre les agents. Les trois dimensions distinguées dans le monde s'appliquent aussi à l'analyse des agents :

- La dimension physique d'un agent interagit avec le monde physique par ses capteurs et ses effecteurs qui délimitent ses interventions ;
- La dimension sociale d'un agent gère ses relations aux autres agents : les conflits, la coopération, l'interaction entre les agents pour construire un monde partagé, les rôles pris dans les organisations pour maintenir la viabilité du groupe et de l'espèce, la constitution et le maintien des institutions humaines ;
- La dimension du soi des agents naturels correspondant dans un premier temps aux besoins que l'agent doit satisfaire pour assurer sa viabilité, comme

² Le quai de Tatihou (www.tatihou.com) est jonché de coquilles de moules cassées car les mouettes l'utilisent à leurs propres fins.

individu et comme espèce, et aux émotions (peur, colère, joie, peine...) qu'il éprouve. Ensuite, l'agent forme des perceptions, des désirs et des intentions relativement à la gestion de ses sensations et de ses émotions. La perception filtre les sensations et les interprète en fonction des besoins, désirs et intentions de l'agent. Le jugement et la décision amènent à former des actions. Ce niveau assure la gestion du présent. Pour les agents artificiels, cette dimension est implantée par une boucle de comportement primaire :



Cette boucle se dédouble au niveau réflexif pour la gestion du passé et du futur.

Un agent peut jouer trois rôles dans une interaction primitive : producteur, destinataire et observateur. On retrouve ces trois rôles dans les interactions langagières : locuteur, allocutaire et observateur. Un agent qui produit un signal le reçoit aussi dans la plupart des cas. L'observateur est un agent qui capte un signal qu'il n'a pas produit et qui n'intervient pas à la suite de ce signal. Donc dans tous les rôles, le signal est capté et il n'y a donc pas de symétrie entre les rôles de producteur et de destinataire. Beaucoup de modèles des interactions en sciences humaines sont faits du point de vue d'un observateur extérieur alors que la conception des agents logiciels nécessite de modéliser les interactions du point de vue des trois rôles que l'agent peut prendre au cours du temps.

3.2 *L'interaction*

On distingue deux formes d'interaction suivant qu'elles mobilisent ou non des conventions partagées.

- La première forme d'interaction, non intentionnelle, se produit via les modifications de l'environnement physique, comme des traces de pas, des bruits, des odeurs, des débuts de construction. S'il y a interaction, c'est parce que l'agent interprète les modifications non intentionnelles de l'environnement ou le contact physique comme des signaux qui vont contraindre ses actions ultérieures. Les traces prennent comme support la matière à laquelle elles donnent une forme, les flux prennent comme support les ondes lumineuses ou sonores alors que contact physique prend l'agent lui-même comme support.
- La seconde forme d'interaction utilise des messages, c'est-à-dire des modifications intentionnelles de l'environnement suivant des conventions partagées entre l'émetteur et le récepteur du message (cris, gestes, paroles, dessins). Ces messages ont un mode d'existence physique pour pouvoir être transmis, mais ils ne s'y réduisent pas car on ne peut pas les interpréter sans en connaître le code³. Les oiseaux par exemple communiquent ainsi par leurs chants, ils se comprennent dans la même population mais les populations différentes de la même espèce ne se comprennent pas. Ces signaux sont des messages en vertu de conventions sociales, qui fixent les règles pour coder et

³ Evelyne Andreewsky disait un jour à Rochebrune : "Si le sens était dans les mots, on comprendrait toutes les langues".

décoder les messages, mais aussi les routines d'interaction qui déterminent qui peut parler à qui et comment, les conventions de début et de fin du dialogue, la gestion des tours de parole. Les messages compris ont une existence mentale qui détermine les actions accomplies par l'agent suite à l'interprétation de ce message. Il faut donc que l'agent se soit approprié des conventions, ce qui est une sorte de trace mentale du social. Cette forme d'interaction est le lieu de l'attribution et de l'acceptation des rôles sociaux via les rites qui les reproduisent au cours du temps. Ces rites sont une forme procédurale des traces : avant d'être consignées par écrit, ils existent dans la mémoire des acteurs et dans les cérémonies qui les font revivre périodiquement.

Ces deux formes d'interaction peuvent permettre la coopération et la compétition entre les agents. Dans les deux cas, une coordination peut se mettre en place. Le §4 traitera de ces questions après un approfondissement de cette deuxième forme d'interaction qui utilise des conventions sociales qu'on appelle la communication.

3.3 *La communication*

Les humains interagissent entre eux selon plusieurs modes. Ils peuvent entrer dans des activités partagées, comme jouer au foot. Ils peuvent communiquer par des signaux (routiers, de navigation... dont le sens n'est pas négociable). Ils peuvent utiliser des codes comme les notations mathématiques et les langages de programmation avec une sémantique non ambiguë. Mais ils utilisent principalement pour communiquer les langues naturelles qui ne fonctionnent pas selon le modèle du code mais selon le modèle du transfert [Coursil, 2000]. Ce niveau symbolique permet de parler des choses en leur absence, et de parler de choses imaginaires comme des mythes, des histoires, de faire des projets. C'est le seul mode de communication avec lequel on peut paraphraser tous les autres. Les agents logiciels sont dans ce monde symbolique par construction et par leur usage.

Pour que les messages fassent sens, il faut partager un monde accessible par la perception et l'action, il faut partager un code ou une langue pour engrammer des messages, et il faut partager un protocole de communication. Ce protocole fait qu'un agent accepte un message d'un autre agent comme s'adressant à lui. Trois niveaux de modélisation sont donc nécessaires :

1) un niveau physique pour l'engrammation du message dans une substance et sa diffusion par un canal de communication, pour fixer le lieu et le temps du message. L'électricité, le téléphone, l'informatique ont étendu les possibilités techniques d'engrammation et de diffusion des messages.

2) un niveau biologique/cognitif qui détermine la possibilité de produire et de recevoir des messages d'un certain type (parole, écriture, signaux télégraphiques...). Pour les agents logiciels, la numérisation des signaux est requise.

3) un niveau social qui fixe les conventions de structure et de contenu. Il comporte la gestion des communications, la syntaxe et la sémantique des messages [Steels, 1997].

On peut distinguer deux formes de communication : la communication synchrone qui utilise des flux et la communication asynchrone qui utilise des traces.

- Un message est synchrone si le locuteur et l'allocataire sont en présence : gestes ou paroles saisis au moment de leur émission. Il peut ou non y avoir co-présence des interlocuteurs dans l'espace. S'il y a co-présence il est possible de faire référence au monde partagé par la perception et l'action. En cas d'échange téléphonique ou vidéo-téléphonique, seule une partie du monde est accessible à distance (bruits dans la pièce, par exemple, ou champ de la caméra).

- Un message est asynchrone s'il est émis hors de la présence de l'allocataire, et si la réponse n'est pas attendue au moment même par le locuteur. Les messages asynchrones nécessitent l'utilisation de traces observables hors du lieu et du moment de leur production : dessins, textes écrits, enregistrements audio ou vidéo. Les messages asynchrones peuvent être adressés à des agents en particulier, qui les reçoivent au bout d'un temps indéterminé (lettres, mails). Dans ce cas, le monde est partagé par la représentation que les acteurs s'en font. Ils peuvent être placés dans un endroit accessible par plusieurs agents pour ceux qui passent (annonces sur un tableau, signaux routiers). Il y a un espace partagé à distance dans le temps. Certaines choses peuvent avoir changé entre l'émission et la réception, ce qui perturbera la bonne compréhension des messages.

3.4 La coopération

Les agents interagissent avec leur environnement physique et avec les autres agents de leur environnement social. Cette interaction est la mise en œuvre de leur rapport au monde, interprété comme coopération ou comme compétition.

Les interactions entre les êtres vivants et leur environnement ont pour objectif principal de maintenir leur viabilité en assurant la subsistance et l'élimination des déchets. Les interactions entre les agents artificiels ont pour objectif de réaliser au mieux les fonctionnalités attendues de ces agents du fait de leur création. Dans les deux cas, les interactions ont lieu soit parce qu'il y a conflit, soit parce qu'il y a coopération pour réaliser une tâche commune.

La coopération est nécessaire quand un agent ne peut pas atteindre ses buts sans l'aide des autres agents. Cette situation est fréquente même chez des espèces primitives, édification d'une fourmilière par exemple. Un agent peut avoir besoin d'un autre agent parce que cet agent a des compétences qu'il n'a pas, ou parce qu'il faut être plusieurs pour réaliser la tâche (pousser un objet trop lourd). La coopération n'est pas forcément consciente, elle peut résulter d'un comportement automatique, comme la construction des ruches ou des termitières. Dans ce cas, ce sont les traces des activités des agents qui orientent l'activité actuelle. Les buts nécessitant la coopération peuvent être des buts sociaux qui assurent la survie du groupe ou de l'espèce. Lorsque ce sont des buts individuels, un agent qui en aide un autre peut attendre une aide en retour ou se faire payer son travail. L'argent est donc une trace de l'interaction socialement constituée.

La compétition entre agents peut avoir plusieurs sources. Les buts des agents peuvent être incompatibles : dans une situation de jeu, chacun cherche à être

gagnant mais la réalisation de ce but pour un des joueurs rend impossible la réalisation du but des autres joueurs. Les ressources dont les agents ont besoin peuvent être rares et l'utilisation d'une ressource par un des agents peut empêcher un autre agent de réaliser son but. Par exemple, si l'imprimante est utilisée je dois attendre pour imprimer mon article et je ne pourrai pas le poster avant le départ du courrier. La compétition crée des situations de conflit qui peuvent être résolues par la lutte ou par la coordination.

Dans les deux cas, coopération ou compétition, une coordination peut être nécessaire pour améliorer le fonctionnement global du système. Lorsque plusieurs agents travaillent sur le même lieu, utilisent les mêmes ressources, où résolvent des sous problèmes qui ne sont pas complètement indépendants (conception d'un objet complexe par exemple), ils doivent accomplir, en plus des tâches liées directement au problème traité et des tâches de coordination. Ces tâches ne sont pas directement productives mais elles améliorent les tâches productives. Les tâches de coordination les plus simples nécessitent des procédures comme la gestion des files d'attente. Les tâches plus complexes sont organisées via des traces : plans, cartes, rapports. Ces traces sont le témoin de représentations partagées nécessaires pour l'organisation de la coopération. Comme pour les rites, le fonctionnement des files d'attente est une trace des comportements de coordination manifestée dans un processus.

4 Les traces

Les premières traces laissées par les êtres vivants sont des conséquences involontaires de leur activité, mais l'observation des interactions qui en découlent les amènent à construire des traces intentionnelles. Il y a plusieurs sortes de traces, selon qu'elles proviennent d'une activité biologique, comme dans les sociétés animales, ou selon qu'elles proviennent d'une activité sociale au sens plein du terme, résultant de l'institution de représentations partagées [Morand, 2004]. Cette distinction ne se recoupe pas avec la distinction intentionnel – non intentionnel car il existe des stratégies de marquage chez les animaux, par exemple pour envoyer leurs prédateurs dans une mauvaise direction. On obtient donc 4 cas :

| intensionnel | conventionnel | exemples |
|---------------------|----------------------|-----------------------------|
| non | non | Traces de pas, odeur |
| non | oui | |
| oui | non | Barrage |
| oui | oui | Messages, textes, documents |

4.1 Description des traces

Dans les populations animales, la coordination a lieu de manière déterminée par la structure sociale, mais elle est aussi le lieu où la structure sociale se met en place. Les animaux d'un même clan coopèrent pour réaliser des buts collectifs, trouver et rapporter la nourriture, construire des abris, élever les petits. Ils ne coopèrent pas avec les membres des autres clans mais ils les combattent. La coordination a lieu au moment où elle est nécessaire, à travers la perception

conjointe des situations et des actions des autres. Les traces y tiennent une place importante : traces de pas, odeurs, phéromones, mais aussi traces de l'activité comme des déplacements d'objets, des branches cassées...

Dans les sociétés humaines, cette forme de coopération a aussi lieu, mais l'existence du langage permet de prévoir et d'organiser cette coopération alors que la situation n'est pas présente (alliances, planification) [Cyrulnik, 1995]. Le langage permet aussi d'améliorer la coopération après coup :

- par remémoration des situations difficiles ou des situations d'échec pour y remédier ;
- en racontant les situations glorieuses pour transmettre des modes de coopération efficaces.

Le langage écrit est une forme de trace instituée socialement qui traverse le temps, comme les dessins et plus récemment les photos, le cinéma, les enregistrements audio et vidéo. Toutes les traces socialement constituées utilisent des conventions qui ne peuvent être formées que dans le langage.

4.2 Structure des flux et des traces

Les flux naturels peuvent être discrets ou continus. La saisie d'un flux continu est une action qui discrétise le flux relativement à l'appareil cognitif et à la pulsation du processus. Un flux est donc composé d'objets qui arrivent l'un après l'autre, soit par nature, soit parce qu'on les observe à travers une fenêtre. Lorsqu'un flux est continu, la continuité du flux doit être reconstruite par les processus, elle n'est pas saisie d'emblée. Lorsqu'un processus produit un flux, il peut être continu, car le mouvement est continu, contrairement à la perception.

Certains flux diffusent dans toutes les directions (son, lumière, odeur), d'autres sont linéaires, (ou plutôt quasi-linéaires puisqu'ils peuvent avoir une section). Les traces sont produites par des flux, mais ne gardent pas toujours la mémoire de l'ordre dans lequel chaque élément de la trace a été produit. Les traces peuvent se placer dans l'espace correspondant à l'environnement en fonction de la structure de l'espace et de la structure du flux ou du mouvement qui les créent. Un trait sur le sol est une trace linéaire, un dessin une trace à deux dimensions, une sculpture ou une maison, une trace à trois dimensions.

4.3 L'appareil cognitif des agents logiciels

On supposera que les agents logiciels sont équipés d'un appareil cognitif qui leur permet de reconnaître, de produire et de traiter par des opérations certains types d'objets. Si un processus produit dans l'environnement un objet d'un certain type, par exemple un cube bleu, un autre processus peut reconnaître cette forme et cette couleur s'il est équipé pour cela. Sinon, il ne perçoit pas ces valeurs, ou il les perçoit autrement. Un objet est donc vu abstraitement comme un couple (codage binaire, codage binaire) qui est vu par certains processus comme un objet, c'est-à-dire comme un couple (type, valeur) signifiant pour lui. La théorie des processus interactifs nécessite donc de remplacer la théorie de l'information par une théorie des objets qui circulent entre les processus. Les flux, les traces et la mémoire sont composés d'objets ou sont décomposés en

objets. La communication suppose en plus de l'identification des objets la possibilité de les nommer.

L'appareil cognitif est caractéristique de l'espèce. Il peut être caractérisé comme structure algébrique et comme dynamique. Pour les processus informatiques, la structure algébrique est un système de types (classes, interfaces, objets primitifs). La dynamique est décrite en tant que dynamique informatique et pas en tant que système d'équations. En effet, poser les équations qui régissent un mouvement ne suffit pas à en rendre compte si le système d'équations n'est pas soluble, ce qui est presque toujours le cas (Problème des trois corps en astronomie). Écrire un programme construisant la dynamique d'un processus, l'enchaînement de ses actions et de ses interactions, permet de l'étudier d'une part dans son exécution et d'autre part dans son écriture. Il n'y a pas d'équivalent entre cette écriture, qui est une création sociale humaine, et le fonctionnement autonome de la vie des agents naturels. Il est donc particulièrement intéressant de construire des processus ayant des comportements analogues à des êtres vivants ou à des êtres humains. Pour communiquer entre eux et avec les humains, les agents logiciels doivent produire, capter et traiter les valeurs irréductibles à leur code, c'est à dire relevant de la relation signifiant/signifié.

4.4 Du modèle des processus au modèle des traces

Les traces sont une forme de mémoire externe pour les processus. Nous ne savons que très indirectement comment est organisée la mémoire des processus vivants. La mémoire des processus artefactuels est construite et donc son organisation est explicite [Nicolle & al., 2001]. Pour faire une théorie des processus, on peut considérer l'espace des traces et l'espace mémoire des processus comme ayant une même organisation : même structure et même fonctionnement. Comme la mesure ne peut pas être présupposée, il faut le voir seulement comme un espace topologique et pas comme un espace cartésien.

Les processus à flux ou à trace produisent, captent et mémorisent des objets qui forment des flux et des traces (figure 1). Leur compétence interactionnelle dépend de leur appareil cognitif : propriétés des capteurs, des effecteurs et de la mémoire. On ne veut pas ramener les valeurs des objets à leurs propriétés physiques, puisqu'il s'agit d'une théorie des processus installés dans les trois dimensions du monde. Du reste, les valeurs au sens d'un processus abstrait sont des valeurs numérisables, donc des formes et pas des substances.

La théorie des machines de Turing n'est pas utilisable pour les processus interactifs, car il n'est pas possible de modifier leur ruban pendant le calcul. Peter Wegner a proposé un autre formalisme, les machines de Turing persistantes [Wegner, 1998], pour rendre compte de l'interaction au niveau séquentiel. Mais pour qu'une théorie soit utile, il faut pouvoir l'appliquer à des processus de taille réaliste. Or la mémoire et les événements y sont analysés comme une suite de codes binaires, comme en théorie de l'information. C'est un grain d'analyse trop fin, qui amène à des preuves extrêmement compliquées parce qu'elles ne s'appuient pas sur une structure de données riche. D'autre part, la dualité entre signifiant et signifié dans les systèmes de signes n'est pas prise en compte. En effet, la tradition de la formalisation logique assimile le signifiant

et le signifié des signes, ce qui amène l'indécidabilité des systèmes formels. On fait donc l'hypothèse que le bon niveau de description est le niveau des « objets » au sens de la programmation par objets, et l'on cherche donc dans le même mouvement à faire une théorie des objets, dans leurs relations aux processus. Les processus peuvent manipuler et transmettre des objets dont la granularité et les attributs doivent être étudiés. Ces objets ne sont pas réductibles à leur codage, ils ont une histoire qui les charge de conventions partagées : ils font valeur.

5 Conclusion

Les traces des activités passées des agents organisent leurs activités ultérieures. Les premières traces sont des conséquences involontaires de l'activité, mais l'observation des interactions qui en découlent amènent à construire des traces intentionnelles.

Dans les sociétés humaines, beaucoup de types de traces sont institués : l'écriture, les dessins, l'argent, la signalisation. Elles forment des systèmes symboliques [Nicolle, 2001]. Ces traces sont le témoin d'une activité passée, mais elles peuvent être construites en vue d'une activité future : plans d'architecture, cartes routières, signaux routiers.

La conception d'agents artificiels pose de manière plus exigeante la question de ce qui fait trace, car toutes leurs interactions passent par des types d'événements et de traces conventionnels. Par contre, on peut observer le fonctionnement de leur mémoire, ce qui n'est pas possible pour les êtres vivants. L'articulation des modèles de la mémoire externe et de la mémoire interne est alors possible.

6 Références

- Jacques Coursil, 2000, *La fonction muette du langage*, Ibis Rouge Éditions, presses Universitaires Créoles, Guadeloupe
- Boris Cyrulnik, 1995, *La naissance du sens*, Hachette
- Bernard Morand, 2004, *Logique de la conception. Figures de sémiotique générale d'après Charles S. Peirce*, Éditions L'Harmattan, Collection L'ouverture philosophique, Paris
- Anne Nicolle, 2001, *La question du symbolique en informatique*, in *La cognition entre individu et société*, ARCo'2001, pp. 345-358, Hermès, Paris
- Anne Nicolle, Pierre Beust, Vincent Perlerin, 2001, *Un analogue de la mémoire pour un agent logiciel interactif*, *Revue In Cognito*, n° 21, pp. 37-66
- Anne Nicolle, 2004a, *Prolégomènes à une théorie des processus interactifs de durée indéfinie*, publication du séminaire *Du Sujet : Théorie et Praxis*, pp. 29-45, MSH Paris, et enregistrement vidéo sur : <http://lutecium.org/stp/anne.html>
- Anne Nicolle, 2004b, *Temps et processus*, Journées de Rochebrune « Le temps dans les systèmes complexes naturels et artificiels », Megève, France, ENST 2004 S 001, pp. 229, 239

- Luc Steels, 1997, The origins of syntax in visually grounded robotic agents. In: Polack, M. (ed.) Proceedings of IJCAI 97. AAAI Press, Menlo Park Ca
- Peter Wegner, 1997, Why Interaction is More Powerful than Algorithms, *Communications of the ACM* 40 (5)
- Peter Wegner, 1998, Interactive Foundations of Computing. *Theoretical Computer Science* 192, Feb. 98

Construction et déconstruction des *Connaissances* par les informaticiens

Francis Rousseaux¹ — Thomas Bouaziz²

¹Laboratoire CRéSTIC (EA 3804)
Université de Reims URCA
francis.rousseau@univ-reims.fr

²Association ExperiensS
www.ExperiensS.org
thomas.bouaziz@experienss.org

RÉSUMÉ. Lorsqu'en 1982 Allen Newell a inventé une nouvelle acception des connaissances pour permettre aux informaticiens d'unifier les deux inspirations fondatrices de l'intelligence artificielle, il était peut-être loin d'imaginer l'immense succès qu'allaient rencontrer ses propositions. La manière de penser les documents numériques et les outils interactifs d'accès à leurs contenus allaient s'en trouver profondément marquée, ouvrant à nombre d'applications innovantes. Pourtant, une analyse de systèmes d'aide à l'interprétation de documents réalisés depuis les quinze dernières années sur la base des propositions de Newell révèle une surprise de taille : pour concevoir une grande partie de ces systèmes, les ingénieurs se sont en réalité efforcés de déconstruire ces propositions. Typiquement, il apparaît qu'à la notion de connaissances est préférée celle de collections. Mais cette dernière ne pouvait peut-être émerger qu'en mobilisant la première pour la tordre et la détourner.

ABSTRACT. Allen Newell wrote his paper « The Knowledge Level » in 1982, inventing a new meaning for knowledge, useful for computer scientists. This proposition was very innovative, allowing systems designers to think differently the notions of digital document and browsing among content. But by analysing several recent systems, we have found that most of it tries to build down that Knowledge approach and to replace it by a Collection approach.

MOTS-CLÉS : connaissances, collections, fouille de données numériques, apprentissage machine, musique, cartographies.

KEYWORDS: knowledge, collections, digital data mining, machine learning, music, maps.

1. Introduction

Nous sommes marqués par une tradition qui voudrait que les documents aient un contenu de connaissances. Cette figure du contenant excède pourtant manifestement le phénomène qu'elle prétend décrire.

Pourquoi ne pas résider au plus près du phénomène éprouvé, quitte à retarder l'acceptation d'explications trop *ad hoc* pour rester stimulantes ? Voici ce que nous vivons *en première main* : lorsqu'elle réussit, la "prise de connaissance" d'un document ébranle et anime la pensée, ouvre un horizon d'investigation et d'altérité, phénomène se traduisant par le désir de sa continuation, par une persévérance dans la confrontation au document (ou par glissements, à d'autres documents), et conduisant typiquement à la production de nouveaux documents. Lorsqu'ils se laissent recevoir sous la figure de l'attraction (continuation/répétition) et non de la répulsion, les documents donnent à concevoir et à produire d'autres documents.

Aussi, le propre de la culture est de ne jamais commencer et de ne pas finir dans l'instant, la vraie question étant celle de la possibilité de sa *reprise* et de sa *métamorphose* [Merleau-Ponty, 1969]. Si ce que je dis n'est déjà plus ce que je pense du fait même que je le dise, comment la pensée pourrait-elle consister en une manipulation de connaissances déjà là, ou en un simple ré-agencement d'unités élémentaires stockées dans des réservoirs appelés documents ?

Les informaticiens sont souvent des défenseurs ardents et innocents de cette tradition qui défend la thèse du contenu de connaissances des documents. Non pas qu'ils aient été recrutés par quelque promoteur militant de cette thèse, ni même qu'ils soient particulièrement intéressés au débat. Mais l'histoire de l'informatique, qui s'origine avec celle de l'Intelligence artificielle (IA), les conduit naturellement à prendre position implicite sur un terrain somme toute stratégique : en effet, à l'heure du document numérique, ce sont bien souvent les informaticiens qui sont amenés à concevoir et réaliser les systèmes d'accès et de navigation qui recouvrent désormais le monde de leurs réseaux et de leurs services [Enjalbert, 2004].

Dans cet article, nous proposons d'enquêter sur l'invention des Connaissances¹ en informatique, qui constitue selon nous l'origine du *regard sous influence* que portent beaucoup d'informaticiens sur la question du numérique.

2. L'invention des Connaissances en informatique

Si l'informatique est une discipline jeune, elle n'en a pas moins une histoire riche et tourmentée, qui commence avec la Guerre froide sous la forme d'un vaste et ambitieux projet transdisciplinaire dont le nom seul dit assez les finalités, si tant est

1. Pour bien indiquer que la notion de *Connaissances* des informaticiens ne coïncide pas de manière évidente avec la notion de connaissance du sens commun, bien que le vocable choisi par les premiers vise à provoquer cette coïncidence, nous emploierons un C majuscule pour distinguer l'acception intrusive du vocable. Sur son pluriel systématique, nous aurons l'occasion de nous expliquer dans la suite.

qu'on se souvienne du sens de renseignement que revêt *intelligence* en anglais : *Artificial Intelligence*.

Le domaine de recherche est marqué par les propositions fondatrices d'Alan Turing, dynamisé par les prévisions dithyrambiques d'un Herbert Simon, pour ne rien dire des autres contributions déterminantes. Cela, tout le monde le sait. Mais on oublie souvent le rôle fondamental joué par Allen Newell, qui a littéralement inventé une acception nouvelle d'une notion très importante en métaphysique, en construisant les *Connaissances* au sens des informaticiens. Il s'agissait de bâtir une notion opérante et propice aux concepteurs et programmeurs de systèmes informatisés, tout en essayant de les convaincre qu'ils tenaient là le Graal que la métaphysique, depuis la nuit des temps, s'échinait en vain à définir. Nombreux et judicieux sont les informaticiens qui ont cherché à innover sur la base des propositions de Newell, et plus nombreux encore sont ceux qui ont subi ses idées sans même pouvoir situer leurs origines.

2.1. La situation de l'Intelligence artificielle en 1982

Quelle était la situation de l'IA en 1982, lorsque Newell entreprit de rédiger son fameux article "*The Knowledge Level*" [Newell, 1982] ? Pour le dire poétiquement, le programme de recherche IA était au bord de sombrer dans son gouffre originaire, qui ne cessait alors de se creuser entre les deux bords de sa circonscription utopique, ironiquement instituée par Alan Turing à travers les deux figures de sa Machine et de son Test [Turing, 1995].

La Machine de Turing est une machine virtuelle logique, qui sera plus tard architecturée par von Neumann [von Neumann, 1996] et bientôt réalisée matériellement dans le silicium des ordinateurs, qui permet d'opérationnaliser et de simuler certains phénomènes temporels et/ou causaux par assimilation de la raison nécessaire (*Modus Ponens* ou principe de déduction) avec la causalité, puis effectuation automatique des inférences logiques transformées en calcul [Turing, 1939]. Du côté de la Machine de Turing, l'IA consiste en un corpus de techniques de programmation spécifiées pour aborder des questions de Résolution de problèmes², celles-là même que Newell adresse avec son *General Problem Solver* (GPS).

Le Test de Turing, lui, rattache au dialogue intersubjectif le mystère de sa *continuation* : un interlocuteur est considéré comme intelligent dès lors qu'il fait *rebondir* le dialogue. Tout acteur intelligent doit demeurer ainsi dans le champ du dialogue constructif et tenir son interlocuteur en haleine en proposant des tirades stimulantes. Quant à l'initiative de la rupture dialogique, chacun entend bien en disposer dans son propre intérêt et en imposer la maîtrise. C'est ainsi qu'un Humain peut en venir à personnifier un interlocuteur artificiel, dès lors que ce dernier est considéré comme intelligent, capable de prolonger l'échange dans la durée et de

² De manière générale, nous ferons débiter par une lettre majuscule tous les termes qui, dans cet article, font référence à des notions répertoriées de l'informatique.

rester à la merci du locuteur humain³. Aussi du côté du Test de Turing, l'IA est-elle une investigation phénoménologique sur la question du sujet dialoguant.

Force est de constater que l'IA s'est développée sans parvenir jamais à combler le gouffre béant entre ces deux pôles également fondateurs. En 1982, le hiatus est critique : impossible de rapprocher les deux bords, malgré les promesses présomptueuses⁴ et les investissements massifs. Le projet de recherche avance pourtant sur son versant Machine, mais l'IA s'enferme dans la programmation des ordinateurs et les tentatives de satisfaire des exigences d'environnement technique et d'interfaces spécifiques, sous couvert des paradigmes dominants de la Théorie de l'information, se trouvant peu à peu réduite à une techno-science. Inversement, le projet tourne court sur son versant Test, inspiré par la philosophie analytique, et demeure à l'état d'une gnose consistant à s'interroger en boucle ouverte sur les similitudes entre une Machine intelligente et un système nerveux humain. En conséquence de cette schizophrénie, les systèmes d'IA ne sortent pas significativement des Laboratoires de recherche, et cet échec retentissant devient difficile à cacher : il fallait se résoudre à jeter l'éponge ou bien ouvrir de nouvelles voies en tentant « le tout pour le tout ».

Newell veut agir pour défendre les chances de l'approche IA, refusant la réduction de l'ambitieux programme de recherche de Turing à une ingénierie au service de la Théorie de l'information. De par sa culture, en tant qu'ingénieur passionné par le versant ingénierie de l'IA, c'est du côté de la Machine que Newell intervient, en proposant de considérer les ordinateurs comme des systèmes en couches (*Symbol Level*), auxquels il ajoute une couche supérieure (le *Knowledge Level*), construite pour atteindre le versant Test de l'IA et solder définitivement la crise menaçante, par liquidation pure et simple du gouffre empli de questions scarifiant la discipline.

2.2. La proposition d'Allen Newell : The Knowledge Level

Newell invente les *Connaissances* pour résoudre la question problématique de l'organisation Humain-Machine en Intelligence artificielle : à la question controversée "Qui est intelligent, de l'Humain ou de la Machine ?", il répond "Rendons-les intelligents ensemble, comme couple/groupe/organisation hybride multiagent; les Connaissances seront le point d'articulation du couplage, l'interdépendance du couple".

Les Connaissances désignent ainsi en informatique la condition de possibilité de l'hypothèse du Niveau des connaissances de Newell. Une coopération interactive Humain-Machine est ouverte sur la base d'un principe de rationalité (j'aime exprimer ce principe à la manière de Montaigne : dis-moi ce que veux, ce que peux, ce que sais, je te dirai ce que fais). En tant qu'elles sont manipulables par l'Humain, qui peut

3. Souvenons-nous du système ELISA [Weizenbaum 1975] : le système jouait le rôle du psychanalyste, et non pas de l'analysé.

4. Les systèmes intelligents devaient remplacer l'Humain jusque dans les conduites/activités réputées intelligentes qu'il déploie, étendant la taylorisation à des sphères qui lui échappaient jusqu'alors.

désormais penser sa pensée comme outil rationnel et finalisé, les Connaissances épuisent/réduisent/décrivent le phénomène de la pensée. En tant qu'elles sont représentables et engrammables dans des systèmes informatiques, elles informent les ordinateurs des situations et des degrés de liberté des actions humaines, qui pourront mobiliser l'opération d'instanciation⁵ ainsi que les inférences logiques pour s' enrôler dans des raisonnements rationnels divers. Les Connaissances offrent finalement un support pour déchiffrer, modéliser, et manipuler l'énigme de la pensée et des actions humaines.

Grâce à Newell, l'IA quitte à la fois le champ limité de la programmation des ordinateurs et la poétique turingienne pour conquérir les organisations, considérées comme des communautés d'agents humains en interaction, dans lesquelles il s'agit désormais d'insérer avec productivité des agents artificiels rationnels. Car même si Newell s'adresse en priorité au couple Humain-Machine, ce seront bientôt des sociétés multiagent bien plus larges⁶ qui vont être modélisées au Knowledge Level.

Newell eut en effet le talent de concevoir le geste salvateur jusqu'au bout, en nommant le lieu du monstre dual⁷ qu'il venait de créer et en le baptisant de façon à l'amadouer et l'appriivoiser. Le lieu du monstre était d'ailleurs déjà pointé par Turing : l'interaction Humain-Machine, comme lieu dépassant à la fois les deux protagonistes, unifiant mystérieusement la Machine au Test. Sa condition de possibilité ? Partager les Connaissances. L'Humain permet à la Machine d'acquérir ses Connaissances, quand la Machine offre à l'Humain d'apparaître comme un interlocuteur digne de lui. À peine opérationnalisée, la notion de Connaissances est ainsi associée à la connaissance du sens commun, pour faire d'un monstre une figure reconnue et opérante (c'est la tératologie — la science des monstres, qui nous enseigne ce subterfuge).

2.3. Les conséquences du geste de Newell

Les Connaissances de Newell sont logiques, plus précisément téléologiques, *hors du temps et du désir humains*. Ses Connaissances peuvent être régionales/domaniales, orientées métier, mais elles ne sont pas situées, elles sont littéralement inhabitées, comme le montre le type d'aporie (logique) auquel Newell admet se heurter (*The Lady and the Tiger*). Newell est chercheur en Résolution de problèmes, il a réalisé son GPS : pour lui, la vie est un vaste problème, et vivre est un vaste mécanisme de résolution de ce problème. Les Connaissances de Newell prétendent réduire la connaissance banale de l'imaginaire humain, *narrative et discursive*, déployée dans la

5. L'instanciation est le grand impensé de l'informatique, qui assimile violemment le *singulier* au *particulier* (Rousseaux, 2004) par le truchement de la fameuse subsomption, alors même qu'Aristote [Aristote, éd.1993] affirmait déjà que la Science ne peut parler que du *général* et reste impuissante à rien dire du *singulier*.

6. L'approche a en effet le mérite de proposer nativement des représentations et des dispositifs d'extraction/acquisition des Connaissances. Dans les faits, elle conduira d'ailleurs souvent des institutions (qui prendront la proposition de Newell au pied de la lettre et s'y conformeront avec un conformisme qui frisera parfois une taylorisation d'un genre nouveau et inédit) à prescrire leurs organisations et leurs savoirs (1982 n'arrive jamais que deux ans avant la funeste prévision de Georges Orwell).

7. Comme naguère Janus, les Connaissances ont deux faces, l'une tournée vers le *Symbol Level* des Machines, l'autre vers les actions humaines, toujours rationnelles et finalisées selon Newell.

durée, pour la recouvrir définitivement d'une chape de rationalité immédiatement finalisée.

À l'occasion, l'exigence de rationalité des Connaissances stigmatisera l'irrationnel de la connaissance narrative : quel est votre *problème* ? Si vous n'en avez pas, si vous ne parvenez pas à en exhiber une expression formelle canonique, alors vous n'avez *pas de problème*, vous allez bien ... Les Connaissances — *Knowledge* (symbolisé par la lettre K en anglais, comme la nouvelle de Kafka), auront parfois des allures kafkaïennes.

Prise au pied de la lettre, la proposition de Newell appauvrit certes notre pensée, mais elle a aussi sa productivité et sa capacité d'innovation propres. De plus, rien n'empêche de chercher à la déconstruire, en mettant en avant des inventions moins durement réductrices comme la notion de *collection* [Vignaux, 2004], davantage située, se déployant ici et maintenant, en singularité, dans un champs d'attraction dynamique. Le parcours peut devenir chorégraphique/scénographique, et renoncer au caractère topologique, toujours *a priori* cartographiable, de son inscription. Mais de cela nous parlerons plus tard.

Car la question de l'interaction collaborative est certes celle de son établissement équitable et fructueux, mais aussi celle de sa durée, c'est-à-dire de son déploiement dans le temps⁸.

2.4. L'invention de Newell sur le terrain : retours d'expériences

L'analyse d'exemples de systèmes informatisés que nous connaissons bien, pour avoir participé de près ou de loin à leur conception/réalisation, nous autorise une étude critique de la productivité des Connaissances inventés par Allen Newell. On appuiera cette analyse sur les systèmes de navigation dans des recueils numériques de morceaux de musique LE MUSICOLOGUE et CUIDADO d'une part, et les systèmes de contrôle de situations CHEOPS et VIRTUALIS d'autre part.

On trouvera une description plus détaillée de ces systèmes en Annexe du présent article.

3. Naviguer au travers de documents numériques musicaux

La mise en place d'un système de navigation au travers de documents musicaux numérisés pose toujours de difficiles problèmes préalables d'acquisition et de restitution, mais aussi de représentation et d'interface Humain-Machine. Lorsque ces problèmes sont enfin surmontés, c'est seulement alors qu'apparaît la difficulté fondamentale : mobiliser les outils basés sur le *Knowledge Level* de Newell pour néanmoins subvertir ses propositions initiales, affadir l'instanciation et la

8. Newell aurait sans doute prétendu que son modèle permettait de coiffer la question de la durée en la réduisant à la variation dynamique et la mise à jour des Connaissances. Mais la dynamique n'est vraiment efficace que lorsqu'elle parvient à ne voir dans la statique qu'un cas particulier du mouvement, et pas le contraire. Avec Newell, le "calcul différentiel" sur les Connaissances reste à inventer.

classification *a priori*, pour aller vers la similarité en situation et la collection déployée *en singularité*.

3.1. Présentation du système de navigation musicale LE MUSICOLOGUE

Le système LE MUSICOLOGUE (cf. Annexe A.1) a été conçu et réalisé par une petite équipe d'informaticiens et de musicologues entre 1987 et 1990. Parmi les ambitions du système, on a voulu permettre à un élève venant de s'exercer à la dictée musicale sur une certaine pièce de s'en voir proposer d'autres par le système [Rousseaux & Saoudi, 1991], selon une cohérence optimale du cursus, dans le sens d'une collection d'exercices travaillés adaptés la progression de l'élève.

3.2. Le cursus comme collection de pièces travaillées

LE MUSICOLOGUE aide l'élève à constituer sa collection de pièces travaillées. Collectionner est plus originaire que catégoriser. C'est dans le fil du temps, du *lewensvelt*. C'est particulièrement vrai dans le cas d'un travail sur des pièces musicales, dont l'empreinte du succès est la continuation d'une activité qui ne cesse ni ne répète son objet, mais se prolonge sur des objets dont la succession fait parcours de collection [Rousseaux 2004], un peu comme lorsqu'on constitue une collection d'œuvres d'art (même si l'appropriation des objets temporels n'est pas comparable à l'appropriation des objets spatiaux). Mais si la trace dans le monde d'une activité n'est autre que sa continuation, comment installer un dialogue Humain-Machine, et sur quel type de Connaissances médianes l'instaurer ?

Dans le cas de l'environnement d'insertion de LE MUSICOLOGUE, l'élève laisse des traces de son activité d'exercice autres que la sélection préliminaire de la pièce qui l'occupe : l'évaluation de son travail, ainsi que l'évaluation de son niveau dans le cursus, ont été soigneusement conçues pour donner prise à un système apprenti qui pourra stimuler l'intérêt de l'élève en lui proposant des pièces *intéressantes* à travailler, parmi lesquelles l'élève aura tout loisir de faire son choix motivé. Mais qu'en serait-il de l'activité d'écoute musicale pure, sans prise de note ni autre trace que le seul désir de sa continuation ? Pourrait-on envisager un système qui se propose pour aider l'auditeur à constituer un parcours/collection, alors même qu'aucun but extérieur à l'activité en situation ne peut être assigné au système ? C'est l'objectif du *Music Browser* développé par Sony-CSL dans le cadre du projet européen CUIDADO, co-ordonné par l'Ircam entre les années 2000 et 2003 [Vinet, Herrera et Pachet, 2002].

3.3. Étude du système de navigation CUIDADO

La navigation musicale au sein de vastes corpus de titres numérisés est très influencée par la notion de *genre*, elle-même héritée de la nécessité de choisir physiquement les CDs qu'on désire se procurer parmi les bacs et les rayonnages des

grands magasins spécialisés. La fin du support CD signe la fin de l'hégémonie de cette activité d'achat et laisse place à une ribambelle d'activités concurrentes et prétendant infléchir l'indexation, d'où l'avènement de régimes d'indexations pléthoriques et concurrents. C'est pourquoi le *Music Browser* de CUIDADO (cf. Annexe A.2) propose, concurremment à une indexation par métadonnées éditoriales, des possibilités de fouille *culturelle* et *acoustique*, renonçant d'ailleurs à imposer des catégorisations exclusives basées sur ces types d'index, mais encourageant l'utilisateur à glisser par une recherche de similarités aussi transversale⁹ et interactive que ses caprices le lui inspirent [Pachet 2000 et 2003]. C'est encore une fois l'esprit de collection qui est à l'œuvre, et le système offre au collectionneur/auditeur des opportunités qui se conjuguent sur des plans différents mais toujours simultanément activables, liberté lui étant laissée de choisir celle sur laquelle il va localement exercer son contrôle.

3.4. Conclusion partielle et premiers enseignements

Les différences entre les systèmes LE MUSICOLOGUE et CUIDADO ne sont pas tant techniques qu'épistémologiques. D'un système à l'autre, on passe (LE MUSICOLOGUE) d'un monde de catégories formelles où la Machine tend à piloter la boucle d'événements interactifs à (CUIDADO) une situation où il est question de collections singulières et où l'Humain tend à rester maître et ultime responsable de cette boucle d'événements, ainsi que des résultats auxquels le système Humain-Machine aboutit. Encore ne parle-t-on plus de résultats dans le cas de CUIDADO mais de *parcours*, la connaissance étant située en ce sens qu'elle est toujours engagée dans une narration qui s'entretient sans nécessité de buts exogènes ou endogènes au système, par glissements en similarité.

4. Naviguer au travers de documents carto/scénographiques

Dans cette deuxième phase d'analyse de systèmes, on s'intéresse à des réalisations qui traitent cette fois de documents de nature cartographique ou scénographique. Exactement dans le même mouvement qui va du système LE MUSICOLOGUE (1990) à CUIDADO (2000) en affranchissant la conception du système informatisé d'exigences téléologiques explicites ou implicites, nous présentons les systèmes CHEOPS (1995) et VIRTUALIS (2005) en mettant en évidence un mouvement de même nature.

9. C'est bien une combinaison des descriptions qui est construite par l'Humain dans *Music Browser*, de type *et/ou* plutôt que de type *ou*, qui permet de faire glisser ses désirs, en appui sur des similarités toujours partielles mais néanmoins stimulantes, les aspects quantitatifs s'appréhendant d'emblée qualitativement, et la Machine étant en retrait bienveillant par rapport à l'utilisateur humain.

4.1. Étude du système de suivi de situations critiques CHEOPS

CHEOPS [Rousseaux, 1995] est un système d'aide à la décision en situations de crise géopolitique, conçu et réalisé entre 1990 et 1995, et intégrant des composants informatiques pour réaliser la possibilité de débats contradictoires de niveau stratégique entre un Humain et un agent artificiel argumenteur [Rosenthal-Sabroux et Rousseaux, 1996; Zacklad et Rousseaux, 1996]. Cet agent tente toujours de présenter d'autres facettes de la situation tactico-stratégique au décideur, pour éviter à ce dernier d'oublier des contradictions fructueuses possibles (et surtout celles qui pourraient faire triompher un éventuel contradicteur ou adversaire).

Dans CHEOPS (cf. Annexe A.3), les documents numériques fondamentaux étaient des cartes géographiques (*raster* et/ou vecteur) augmentées de symboles décrivant une situation courante, et dont on supposait l'interprétation quasi-objective au niveau tactique, pour rechercher des interprétations raisonnables/justifiables au niveau stratégique (le rapport tactique/stratégique ressemble au rapport local/global de nos descripteurs solfégiques de la musique).

Mais CHEOPS permet aussi bien de légitimer les décisions que de les former, les intentions stratégiques du décideur infléchissant l'interprétation de la situation en concurrence avec les analyses tactiques qui en émergent. Quant aux explications qui sont supposées remonter du terrain jusqu'aux États-majors, elles sont concurrencées par des justifications d'action intentionnelles qui prescrivent l'interprétation du terrain.

Avec CHEOPS, il devient clair que le statut des Connaissances introduites par les informaticiens dans le paysage Humain-Machine peut avoir des conséquences sur les organisations, voir même des effets directs sur leur institution. D'abord "à extraire" en amont du cycle de conception des systèmes intelligents, les Connaissances sont vite devenues l'enjeu de l'Acquisition des Connaissances comme "principe constructif de modélisation" [Clancey 1993]. Mais finalement, ce principe modéré est vite devenu principe de management et d'organisation (par exemple en vue d'unifications normatives des vues au sein d'une même organisation), et donc principe d'institution organique. C'est ainsi que l'innovation de Newell tend à ébranler le champ d'insertion des outils qu'elle conduit à produire, c'est-à-dire l'organisation elle-même, sommée de laisser place à des agents rationnels artificiels, et finalement elle-même finalisée et normalisée en conséquence.

Au fond, CHEOPS n'est pertinent que si l'on considère qu'il est possible de mettre à distance une situation pour procéder à des décisions, vues comme des réorganisations partielles de certaines composantes de cette situation. La situation serait en quelque sorte *congelée* avant cette opération, et ne serait pas affectée par la visée délibératrice. Or une grande partie de notre investigation philosophique, alors que nous animions entre 1997 et 2000 le groupe de recherche « Reconstitution de la décision politico-stratégique » du Collège international de philosophie, a consisté à montrer les considérables limites de cette vision.

C'est suite à cette enquête philosophique que l'idée du système d'aide à la scénographie VIRTUALIS est née, comme dispositif Humain-Machine destiné à explorer d'autres postures de décision.

4.2. Le système de génération de collections d'interactions VIRTUALIS

VIRTUALIS (cf. Annexe A.4) est un système établi sur l'idée qu'un spectacle vivant peut être considéré comme une collection/procession d'interactions sous contrainte, et que la mise en œuvre de procédés de rétention de certains échanges interactifs peut ouvrir l'œuvre (au sens de l'œuvre ouverte d'Umberto Eco), en permettant de densifier l'espace des interactions [Rousseaux et Bonardi, 2004].

Par exemple, Alain Bonardi, principal concepteur de VIRTUALIS [Bonardi et Rousseaux, 2001; 2004], a mis en œuvre le système dans la création d'une pièce de Geneviève de Gaulle mettant en scène une narratrice et une danseuse Nô, mais aussi un immense écran de fond de scène sur lequel étaient dessinés des mobiles qui s'animaient en lien avec les émotions particulières de la voix de la narratrice. Ainsi, l'influence immédiate de la voix se trouvait concurremment médiatisée par l'écran et acquérait ainsi une portée et une rémanence supplémentaires dans l'ordre de l'épaisseur temporelle.

Ce qui nous intéressait dans ce travail était de parvenir à affadir l'instanciation¹⁰ [Rousseaux et Bonardi, 2004], le fameux impensé de l'informatique, en proposant qu'une mise en scène ne soit plus contrôlée et spécifiée par des variations d'instanciation dans des ontologies de personnages et d'actions, mais davantage par un glissement de situation contrôlé par la situation elle-même.

Par une approche Fouille de données interactive, on représente l'exemple comme une spécialisation de l'ensemble des cas, et on cherche d'autres spécialisations voisines, mais sans disposer par avance d'une ontologie. L'utilisateur accepte de la façonner à sa main avec l'aide interactive de la machine, de manière *ad hoc*.

Le système des interactions multimodales dans *La traversée de la nuit* repose sur un système homme-machine « autarcique » : une comédienne, Valérie Le Louédec, disant l'intégralité du texte, une danseuse, Magali Bruneau, accomplissant un certain nombre de gestes inspirés du théâtre Nô et un ordinateur multimédia, acteur artificiel. L'ordinateur se manifeste sous forme d'images projetées sur un écran de fond de scène de très vastes dimensions (la comédienne et la danseuse en voient toujours au moins une partie sans se retourner), provoquant la réaction des deux comédiennes,

10. *Instanciation* est un anglicisme couramment utilisé par les informaticiens, qui renvoie au mot *instance* signifiant *exemple, cas*; l'instanciation généralise en quelque sorte l'opération, utilisée par les mathématiciens, d'affectation d'une valeur numérique à une variable : pour parler du réel, les informaticiens *instancient* des classes abstraites, décrétant ainsi que telle ou telle entité est un cas particulier d'une classe, elle-même reliée à d'autres classes par des hiérarchies de généralité et/ou des propriétés formelles, l'ensemble du dispositif constituant ce qu'on appelle une Ontologie (les ontologies prétendent ainsi décrire des pans de connaissances mondaines très utilisées en intelligence artificielle), parfois une *conception à objets* (une conception à objets est constituée de graphes d'héritage conçus pour donner lieu à des programmes informatiques par simple instanciation de paramètres clés).

notamment de la danseuse adaptant la réalisation de sa gestuelle aux mouvements et qualités de l'image. Or, les deux actrices sur scène constituent les deux versants — conscient et inconscient — du même personnage, selon les traditions du *shite* et du *waki* du théâtre Nô. Entraînée dans ses déplacements par la danseuse, la comédienne adapte elle aussi sa déclamation, sans compter les moments où elle regarde aussi l'écran. Pour boucler la boucle, l'ordinateur capte les états émotionnels de la voix de la comédienne.

4.3. Nouvelles conclusions et enseignements supplémentaires

Avec VIRTUALIS, Alain Bonardi et moi découvrons que la technologie permet de se passer des concepts mêmes qui lui ont permis d'exister, ou plus exactement de les déconstruire, au sens de Derrida. Ainsi, après même qu'elles aient subrepticement prescrit des organisations dans leurs méthodes de travail et de conception, les Connaissances moléculaires et cartographiées de Newell pouvaient être déconstruites. La notion laisse place à autant de scénographies/chorégraphies qu'il y a d'expériences à l'œuvre, dans un mouvement qui rappelle celui de la *concrétisation* au sens de Simondon [Simondon 1989], l'innovation consistant à déconstruire ces Connaissances pour en liquider l'artifice, tout en continuant à mobiliser les outils que la notion dogmatique a permis de constituer.

Après l'Information et le *Rapport fonds/forme*, c'est au tour des notions de *Connaissances* et de *Contenu des documents numériques* d'être mises en tension.

5. Déconstruction des *Connaissances* et invention de la *Collection*

En revisitant ces expériences de conception de systèmes intelligents d'aide à l'interprétation de documents numériques, nous avons constaté que la proposition d'une description *a priori* de la connaissance, hors situation vécue, nous a souvent laissé insatisfaits, et que nous lui avons préféré celle de glissement des situations, fondant une interaction Humain-Machine dans la durée, au travers de formes narratives, et permettant la constitution de collections inspirées, à l'intentionnalité mouvante et vivante.

Les documents numériques ne *contiennent* pas de connaissances. Leur « prise de connaissance » est bien plutôt un processus élaboratif de collection, qui vise concurremment sa terminaison et sa continuation, les deux visées étant nécessaires ensemble parce qu'elles entretiennent leurs possibilités réciproques.

6. Bibliographie

Aristote, (éd. 1993) *Métaphysique*. Garnier Flammarion.

Bonardi A. et Rousseaux F. (2004) « A la croisée de l'informatique et du théâtre : similarité en intension vs en extension ». *Conférence Internationale ARCo*, Compiègne, décembre 2004.

- Bonardi A. et Rousseaux F. (2001) « Interagir avec un contenu opératique : le projet d'opéra virtuel interactif Virtualis ». *Revue d'Interaction Homme-Machine*, numéro spécial *Interaction et Documents*, Volume 2, n°1, juin 2001.
- Clancey, W. (1993) « Notes on Epistemology of a Rule-based Expert System ». *Artificial Intelligence* 59: 191-204, Special issue Artificial Intelligence in Perspective.
- Enjalbert P. et Gaio M. (2004) « Actualité d'une approche sémantique du document numérique », *Conférence Internationale sur le Document Electronique*, La Rochelle, 22-28 juin 2004.
- Merleau-Ponty, M. (1969) *La prose du monde*. TEL Gallimard.
- Newell, A. (1982) « The knowledge Level ». *Artificial Intelligence*. Vol. 18, p.87-127.
- Pachet, F. (2000) *A Taxonomy of Musical Genres*. RIAO, Paris.
- Pachet, F. (2003) « Nom de fichiers : Le nom ». *Actes du séminaire STP de la MSH Paris*.
- Rosenthal-Sabroux C. et Rousseaux F. (1996) *Systèmes informatisés d'aide au traitement des crises*. Article "acquisition des connaissances, aide à la décision et gestion de crises", "Gestion de crises et médecine des catastrophes", dirigé par Renaud Vié Le Sage, Elsevier, novembre 1996.
- Rousseaux F. et Bonardi A. (2004) « *Music-ripping* : des pratiques qui provoquent la musicologie ». *MUSICAE SCIENTIAE*, numéro spécial 2003-2004.
- Rousseaux F., Bonardi A. et Zeppenfeld C. (2004) « Une mise en scène inspirée de la fouille interactive de données ». *Conférence Internationale sur le Document Electronique*, page 205-112, La Rochelle, 22-28 juin 2004.
- Rousseaux, F. (1989) LE MUSICOLOGUE, A Learning Apprentice System for Music Education. *Artificial Intelligence & Music*, IJCAI, Detroit, août 1989
- Rousseaux, F. (1995) Contribution à une méthodologie d'acquisition des connaissances pour l'ingénierie des SIC : l'exemple de Cheops pour l'aide à la gestion de crises collectives à caractère géographique. Mémoire d'habilitation à diriger des recherches, rapport de recherche LAFORIA.
- Rousseaux F. et Saoudi K. (1991) « Informatique musicale, pédagogie et communication ». *Revue de pédagogie musicale MARSYAS de l'Institut de Pédagogie Musicale*, dossier n°20 "Informatique et pédagogie", décembre 1991.
- Rousseaux, F. (2004) « Collecting or classifying? — Intimacy clarified: a collector's ritual ». Special Interest Group on "Philosophy and Informatics", *Professional Knowledge Management International Conference*, Ulm, octobre 2004.
- Simondon, G. (1989) *Du mode d'existence des objets techniques*. Aubier.
- Turing, A. (1995) « Computing Machinery and Intelligence ». *Mind* LIX, n°236, 1950; réédité dans les *Collected Works of A-M. Turing*, vol 3. *Mechanical Intelligence*; traduction française dans A-M. Turing, J.-Y. Girard, « La machine de Turing », Seuil, Paris.
- Turing, A. (1939) « Systems of Logic based on Ordinals ». *Proceedings of the London Mathematical Society* n°45.
- Vignaux, G. (2004) La notion de collection : genèse, développements, valorisations. *Rapport interne CNRS du RTP 33 "Document et contenu : création, indexation, navigation"*.
- Vinet H., Herrera P. et Pachet F. (2002) « The Cuidado Project ». *Actes des conférences ISMIR*, Paris.
- Von Neumann, J. (1996) *Theory of Self-Reproducing Automata*. London and Urbana, University of Illinois Press; traduction française : *Théorie générale et logique des automates*, Paris, Champ Vallon.
- Weizenbaum, J. (1975) *Raisonnement de l'homme, intelligence de la machine*. Edition de l'informatique.
- Zacklad M. et Rousseaux F. (1996) « Modelling Co-Operation in the Design of Knowledge Production Systems : The MADEINCOOP Method - An example in the field of C3I systems ». *Computer Supported Co-operative Work : The Journal of Collaborative Computing* 1-22, Kluwer Academic Publishers, the Netherlands.



Du roman policier à la criminologie: traces et abduction

Danièle Bourcier

CNRS - Centre Marc Bloch
Schiffbauerdamm, 19
D-10117 BERLIN
bda@cmb.hu-berlin.de

« Une fois que vous avez exclu l'impossible, ce qui reste,
aussi improbable que cela soit, doit être la vérité »
Sir Arthur Conan Doyle, *Les aventures de Sherlock Holmes*

Résumé

L'énigme policière sera considérée comme le prototype de l'énigme en général. L'intérêt de ce type d'énigme réside en ce qu'elle a fait l'objet d'une littérature de fiction (le roman policier) mais aussi de développements scientifiques et techniques remarquables en criminologie. Nous avons tenté à partir de ces deux approches de revoir la pertinence de la fameuse méthode de Sherlock Holmes. Nous y relèverons cependant quelques insuffisances sur la nature du raisonnement qui serait mis en œuvre au cours de la résolution d'énigmes, lui préférant l'hypothèse de l'abduction à celle de la démonstration. L'abduction sera vue comme un raisonnement dialectique, celui de la meilleure explication possible développée au cours d'un dialogue sur la recherche d'une hypothèse. Ce qui correspond bien au cadre du procès criminel ultérieur, soumis au principe du contradictoire et du doute qui, comme chacun sait, profite toujours à l'accusé.

Mots clefs

raisonnement, abduction, criminalistique, criminologie, roman policier, intelligence artificielle, sérendipité

Présentation

A la fin des années 1980, le quartier de l'East End à Londres était très dangereux. Des criminels en tous genres exerçaient leur pratique dans des rues mal éclairées et les taudis de White Chapel. Un matin de novembre 1888, un homme frappa à la porte d'une prostituée de 24 ans. Par la fenêtre cassée, il découvrit l'horreur : le corps de la jeune femme était démembré et découpé en morceaux. La police appelée sur place constata que le foie avait été mis entre les pieds de la victime et qu'une de ses mains avait été placée dans son estomac... Plusieurs meurtres de femmes ont eu lieu à la même époque dans des circonstances similaires et avec la même sauvagerie, ce qui laisse penser que ces actes étaient commis par un même personnage mystérieux surnommé Jack

l'Eventreur. En dépit de récits de témoins et d'une quantité importante d'indices, aucun suspect plausible n'a été identifié jusqu'à aujourd'hui.

Comment interpréter les traces d'une énigme ? Comment les affaires criminelles sont-elles résolues ? Certaines n'ont jamais été instruites (mort de Napoléon ou suicide de Van Gogh), ou révisées (procès de Sacco et Vanzetti), ni résolues comme celle de Jack l'Eventreur¹. Mais Oedipe qui était son propre juge instructeur, l'a en partie résolue en reconstituant le mythe familial.

La résolution d'une énigme et celle d'une affaire criminelle posent des questions similaires : il faut repérer des indices, analyser des faits, interpréter des *traces*, puis, à l'aide de raisonnements, de techniques ou de méthodes scientifiques, les rapprocher et construire une solution plausible. Cette solution est souvent la seule acceptable en l'état des connaissances mais elle peut toujours être réfutée si l'on trouve de nouveaux faits qui la contredisent. L'énigme policière est bien le prototype de toute énigme.

Deux aspects doivent donc être distingués dans la résolution d'une énigme : l'aspect cognitif, propre aux qualités individuelles du détective comme l'intuition, le sens de l'observation, et surtout le raisonnement et l'apprentissage; l'aspect vérificatif qui sollicite des outils et des expertises fondées scientifiquement qui ont aidé à construire l'hypothèse la plus acceptable.

En ce qui concerne l'aspect *cognitif*, le raisonnement est fondamental. En effet pour vérifier la plausibilité de la solution, des allers et retours sont nécessaires entre l'observation des faits et la formulation des hypothèses. Ces allers et retours se font suivant des types de raisonnement particuliers. Pour Sherlock Holmes il s'agissait d'un raisonnement synthétique. Plus généralement la construction d'hypothèses est le fruit d'un raisonnement spécifique : l'*abduction*. Mais une bonne abduction se teste par rapport à des observations pertinentes et un réel vérifiable.

Le deuxième aspect, dit *vérificatif*, est l'objectif principal de toute démarche scientifique. Mais elle est devenue fondamentale dans l'instruction d'un dossier judiciaire. La criminalistique permet de constituer des preuves ou des indices qui aideront le policier puis le juge à construire- comme Oedipe- des scénarios plausibles, à charge ou à décharge. Cette science appelée *forensic* en anglais² mobilise de plus en plus de connaissances scientifiques ou techniques.

Cette distinction ne serait pas reconnue par Sherlock Holmes qui ne sépare pas les aspects scientifiques des aspects intuitifs. Nous sommes d'accord avec lui pour accepter que toute résolution d'énigme articule des signes (une magie) et des faits (une science). Cette distinction ne sert qu'à la clarté de l'exposé.

Par rapport au thème de la trace et de l'énigme qui a été lancé pour Rochebrune 2006, j'aimerais introduire une réflexion sur les rapports entre

¹ Les énigmes non résolues peuvent ressortir à tout moment : en 1991 un ferrailleur de Liverpool prétendit qu'on lui avait donné le journal intime de Jack l'éventreur.

² Forensic vient de *forum* qui signifie le marché : dans la société romaine, la justice était rendue sur la place publique.

compétences cognitives et science de la preuve. Je prendrai appui sur la *méthode de Sherlock Holmes*, exposée et révisée dans maints ouvrages et sites dédiés au célèbre détective³. La partie « raisonnement » de cette méthode cependant, en l'état actuel des travaux en sciences cognitives, mérite une mise à jour critique.

On montrera en particulier comment l'*abduction* relie les découvertes et les énigmes et que ce raisonnement est aussi celui qui est à l'œuvre dans la sérendipité.

1- Une méthode cognitive de résolution des énigmes

Il a regardé partout. Il a inspecté le plâtre des murs et les interstices du parquet à la loupe. Les deux détectives de Scotland Yard qui avaient appelé Holmes sont un peu moqueurs. Ils vont en être pour leurs frais. "*Je vais vous dire quelque chose qui pourra vous aider*, leur lance finalement Sherlock Holmes. (...) *Il y a eu meurtre, et l'assassin était un homme. Il mesure plus d'un mètre quatre-vingts, se trouve dans la force de l'âge, il a des petits pieds pour sa taille, porte de solides bottines à bout carré et fume des cigares de Trichinopoly. (...) Selon toute probabilité, le meurtrier a un visage rubicond, et les ongles de sa main droite sont particulièrement longs. Ce ne sont que quelques indications, mais elles pourraient vous être utiles.*" Là où les autres n'ont rien vu, il a su regarder. Les indices parlent d'eux-mêmes. C'est évident...

L'intuition, le sens de l'observation et le raisonnement font partie des canons du bon détective.

Le roman policier est la plupart du temps construit autour de cas réels. Conan Doyle reprend par exemple des crimes qui ont défrayé les chroniques à la fin du XIX^e à Londres : ils constituent un matériau qui continue d'inspirer la criminalistique⁴. Fombonne (1996) explique ainsi que « *Si Sherlock Holmes savait observer, prélever et faire parler l'indice, c'est parce qu'il appartient d'une part à l'exceptionnel par la nature de ses enquêtes, mais d'autre part au banal par l'extrême rusticité des moyens techniques qu'il emploie*". Il conclut alors que "*seul le raisonnement du personnage permet d'établir une vérité dont la portée prospective de l'indice, pourtant fulgurante, n'est que le corollaire*"...

L'intuition travaillée

Pasteur avait signalé que le hasard en science ne favorise que « les esprits préparés ». Il en est de même pour l'intuition. L'intuition est, chez Holmes, préparée, construite. L'un des traits les plus marquants du célèbre détective auprès du grand public est sans aucun doute la méthode qu'il emploie dans ses

³ Cet article doit beaucoup au site des « holmesiens » français : www.sshf.com. Je les en remercie.

⁴ "*A police expert or examining magistrates would not find it a waste of time to read Conan Doyle's novels*" disait notamment le docteur Locard qui, expert en criminalistique, a été à l'origine du principe de Locard suivant lequel lorsque deux objets sont en contact, de petits morceaux de matières passent de l'un à l'autre (cité dans B. Kaye, *Science and the Detective*, Weinheim, VCH, 1995 p. 71).

investigations. Chacun a en tête le souvenir d'un Sherlock Holmes furetant ici et là, loupe à la main, à la recherche du plus petit indice. Cependant, c'est sans doute la seconde phase de sa méthode qui força l'admiration et donna à Sherlock Holmes la réputation qui ne l'a pas quittée depuis. Prenons un exemple d'exploitation de ces données, tiré de *La Ligue des rouquins* : *"En dehors des faits évidents que M. Wilson a quelque temps pratiqué le travail manuel, qu'il prise, qu'il est franc-maçon, qu'il est allé en Chine, et qu'il a beaucoup écrit ces derniers temps, je ne puis rien déduire d'autre !"*

Holmes continue cette succession de détails biographiques et personnels en expliquant le cheminement de sa pensée : *"Votre main droite est presque deux fois plus large que la gauche. Vous avez travaillé avec, et ses muscles ont pris de l'extension. [...] En contradiction avec le règlement de votre ordre, vous portez, en guise d'épingle de cravate, un arc et un compas. [...] Que peut indiquer d'autre cette manchette droite si lustrée ? et cette tâche claire près du coude gauche, à l'endroit où vous posez votre bras sur votre bureau ? [...] Légèrement au-dessus de votre poignet droit, il y a un tatouage : le tatouage d'un poisson, qui n'a pu être fait qu'en Chine."*

Bien que ses qualités ne soient pas toujours appréciées à leur juste valeur (*"Eh bien ! c'est formidable. Au début, j'ai cru que vous étiez intelligent, mais je m'aperçois que ça n'était pas si malin, au fond !"* s'écrie M. Wilson), Holmes expose un vrai raisonnement, logique, rationnel et qui ne laisse aucune place, dit-il, à l'intuition.

L'intuition est si bien travaillée qu'elle se mêle au processus de découverte. Pourtant dans ses enquêtes, ce tour de force n'est qu'une étape dans le processus qui mène à la résolution d'un *"mystère"* : la place faite à l'observation est fondamentale pour la suite du raisonnement.

Un sens aigu de l'observation

Sherlock Holmes ne fait pas de psychologie : il observe systématiquement les traces que laissent les personnages qu'ils croisent : *"Il est difficile de se servir quotidiennement d'un objet sans que la personnalité de son possesseur y laisse des indices."* (*Le Signe des quatre*).

Ce qui retient particulièrement son attention, ce n'était pas l'aspect physique de ses clients mais plutôt leurs vêtements ou leurs accessoires. *"Mon premier regard, s'il s'agit d'une femme, est pour les manches. S'il s'agit d'un homme, pour les genoux du pantalon"* (*Une Affaire d'identité*). Toutefois, plus tard, il s'intéressera aux détails corporels : *"Il faut toujours commencer par regarder les mains, Watson. Ensuite les poignets de la chemise, les genoux du pantalon et les souliers"* (*L'Homme qui grimpa*).

Le meilleur exemple de cette règle figure dans l'aventure de L'Escarboucle bleue, où Holmes infère une série d'indications à partir du chapeau de quelqu'un qu'il n'a jamais vu, et qu'il ne connaît pas. Voici les indices tirés de l'observation du chapeau en question:

"D'après la surface frontale, son propriétaire est un intellectuel. Et alors qu'il était très à l'aise ces trois dernières années, il a mangé récemment de la vache enragée. On peut dire aussi de lui qu'il avait le don de prévoyance : moins à présent toutefois que jadis ; perte imputable, sans doute, à une certaine dégénérescence morale qui, jointe à des revers de fortune, paraît due elle-même à une influence détestable : la boisson, évidemment. Et ceci serait la justification du fait non moins évident que sa femme ne l'aime plus. [...] Cependant cet homme n'a pas perdu tout respect de soi-même. Nous nous trouvons en présence d'un gentleman qui mène une existence sédentaire, qui sort peu, qui n'est plus en bonne forme physique, qui est entre deux âges, et dont les cheveux grisonnent... Pour les cheveux, il les a fait couper ces jours-ci, et il les discipline d'habitude avec du cosmétique [...] etc."

De la même façon, les Princes de Sérendip ont induit que la mule était borgne de l'œil droit parce qu'ils avaient observé sur les bords de la route que l'herbe avait été broutée seulement du côté gauche (Van Andel, Bourcier, 2001).

La manière dont on cherche

L'intuition, l'observation impliquent aussi une méthode : *"Comment avez-vous vu cela, Holmes ? Parce que je le cherchais."* (Les Hommes dansants). L'art de l'observation n'est pas une question d'intelligence. Il s'agit d'un entraînement constant et d'un travail de recherche approfondi.

Holmes *"ne voit pas plus que les autres, mais il est entraîné à remarquer ce qu'il voit."* (Le Soldat blanchi). L'observateur occasionnel regarde, mais ne sait pas s'arrêter sur les détails *importants* parce qu'il n'a pas *appris* à le faire. Holmes sait tout ce qu'il faut savoir sur les tissus, les boutons, les lunettes, les chaussures, les cicatrices, les traces et certaines parties du corps. Il écrira d'ailleurs plusieurs études sur ces « détails » : De la distinction entre les cendres de divers tabacs, La Détection des empreintes de pas, Les Tatouages, Les Différentes formes d'oreilles humaines, etc.

Lorsqu'il est sur les lieux du crime Holmes sait où chercher car il sait ce dont le criminel a eu besoin et ce qu'il a laissé derrière lui. *"Vous connaissez ma méthode, Watson. Je me mets à la place de l'homme, et ayant d'abord évalué l'ampleur de son intelligence, je m'efforce d'imaginer comment j'aurais moi-même agi dans des circonstances analogues."* (Le Rituel des Musgrave).

Sherlock Holmes met progressivement en application pratique l'ensemble de ses méthodes de constat, recherche et identification du criminel, méthodologie affinée au fil de l'expérience :

- aborder l'affaire avec l'esprit libre sans tenir compte d'avis extérieurs

Dans l'enquête sur la mort de Sir Eustace Brackenstall, du Kent (Le Manoir de l'Abbaye, 1897), Holmes revient sur ses premières observations, sentant par instinct que tout est faux sur la scène du crime : *"Si je n'avais pas considéré les choses comme sûres et certaines, si j'avais procédé à mes examens avec le soin que j'aurais déployé si nous avions abordé l'affaire l'esprit libre, sans histoires toutes faites pour me brouiller la cervelle, n'aurais-je pas alors découvert une piste sur laquelle nous aurions pu galoper ?"...*

- examiner méthodiquement le lieu du crime, par balayage d'un « cercle virtuel d'investigation »

Toute observation devra converger tôt ou tard vers le centre : la victime.

En juillet 1895, lors de l'assassinat du Capitaine Peter Carey, le corps transpercé par un harpon, Holmes explique au jeune inspecteur Hopkins que *"du moment que le criminel se déplace sur deux pieds, il y a toujours un foulage, une dentelure, une éraflure, une modification minime de l'état du sol que le chercheur scientifique peut détecter."*

- recueillir les témoignages de toute personne, quels que soient son statut et son apparence

"la première qualité d'un enquêteur criminel étant de pouvoir percer un déguisement."

Une filature ou une surveillance peut être nécessaire dans le cadre de la recherche et de l'identification du criminel, premier pas dans la présomption de culpabilité, ou dans la preuve d'une cachette, d'une liaison, d'un domicile, d'une complicité, voire d'une totale innocence (cf. "Le chien des Baskerville).

2- Du recueil des indices à la vérification des faits : la criminalistique

Pour soutenir l'intuition du détective ou du policier, il existe des méthodes et des technologies tendant à constater le crime, à en rechercher l'auteur. On a vu que dans les pays Anglo-Saxons, les "Forensic Sciences" (Sciences criminelles) sont définies comme l'ensemble des procédés ou techniques mis en œuvre à la suite de la survenance d'un fait judiciaire et visant par la recherche, le rassemblement et l'exploitation scientifique des indices, à en comprendre les mécanismes en vue d'en identifier les personnes en présence (auteurs ou victimes) ce qui permet la constitution de preuves exploitables au cours du procès pénal.

Les débuts de la police scientifique : le détective comme initiateur

Au vu de l'ensemble de ses méthodes, Sherlock Holmes se pose en véritable initiateur en matière de police technique et scientifique.

L'apprentissage du constat constitue le chapitre central de la police technique. Le but est de recueillir sur place les données du problème, en obéissant à des règles logiques, et en s'adaptant à la fugacité des preuves, les traces, les témoignages; et à des règles techniques, celles-ci se déroulant dans un ordre rationnel immuable : photographie des lieux et des sujets, relevé des traces, description écrite des lieux, établissement d'un plan, récolement des témoignages.

Le Canon représente la seule source d'informations sur les différentes méthodologies de Sherlock Holmes dans le constat, la recherche et l'identification du criminel⁵. Mais elles sont la plupart du temps noyées dans une description littéraire, au grand mécontentement du détective qui se veut « scientifique ». Holmes proteste lors d'une conversation avec son biographe, rapportée dans le problème du "Manoir de l'Abbaye" (1897 – publié en 1904) : *"Vous avez la détestable habitude de considérer toute chose du point de vue du conteur et non du point de vue du chercheur scientifique. Par-là, vous avez démoli ce qui aurait pu être une suite instructive et même classique de démonstrations* (je souligne). *Vous négligez la finesse et la délicatesse de mes déductions* (je souligne) *pour insister sur des détails dont le caractère sensationnel excite peut-être la curiosité du lecteur mais ne l'éduque sûrement pas ! ... Je me propose de consacrer les années de ma vie déclinante à réunir en un seul volume tout l'art du détective...*". Mais aucun exemplaire de ce volume n'a encore été publié ...

Au moment où le détective officiait (vers 1880), la science au service des affaires criminelles commençait : les officiers de la police procédaient certes de façon empirique mais avaient une profonde connaissance du milieu.

Avant de s'engager dans une carrière liée à la criminalistique, il est nécessaire de faire un apprentissage minutieux et rigoureux sur tous les domaines directement ou indirectement concernés... Il en était de même pour Sherlock Holmes entre l'été 1874 où il quitte l'université et 1877, qui passe son temps à *"polir les outils de son métier"*, en étudiant toutes *"ces branches de la Science"* dans lesquelles il doit devenir expert avant de pouvoir s'établir dans sa profession. Il assiste ainsi aux cours de l'hôpital St Bartholomew à West Smithfield, près de la cathédrale St Paul ; plus exactement aux cours de démonstrations d'anatomie pathologique et de chimie pratique, études auxquelles s'ajoutent ses propres expériences scientifiques...

Le domaine scientifique revêt donc pour Holmes, dès le début de son apprentissage, un caractère tout à fait primordial dans la recherche du criminel.

⁵ Le "Canon" : l'ensemble des aventures de Sherlock Holmes relatées par Watson, soit quatre romans et cinquante-six nouvelles, publiées par Sir Arthur Conan Doyle entre 1887 et 1927.

Cette attitude est révolutionnaire dans le contexte de l'époque. Dans le même temps, il a acquis une solide connaissance de l'histoire du crime, en se fondant sur l'observation que *"tout arrive de manière cyclique"*. Tout au long de sa carrière, il continuera à accumuler toutes les informations et savoirs liés à telle ou telle enquête criminelle comme le livre de J.G. Wood, "Out of Doors", où ce dernier parle de la *Cyanea Capillata*, la "Crinière du lion". Le Canon donne des indications précises sur l'utilisation des sciences par Sherlock Holmes dans la recherche de la vérité : dans l'Étude en Rouge (1881), à la suite d'un travail intensif au laboratoire de chimie de Bart's, S. Holmes trouve un réactif pouvant être précipité par et uniquement par l'hémoglobine, même sous une forme hautement diluée. *"La découverte médico-légale la plus utile qu'on ait faite depuis des années"*.⁶ Le Docteur Watson rapporte dans l'aventure de Shoscombe Old Place (mai 1902) l'utilisation par Holmes d'un microscope pour identifier des restes de colle sur une casquette trouvée près d'un corps, et appartenant au meurtrier présumé. Le détective explique que depuis qu'il a coincé un faux monnayeur par la limaille de zinc et de cuivre qui se trouvait dans la couture de sa manchette, on commençait à mesurer l'importance du microscope...

Les méthodes techniques et scientifiques de Sherlock Holmes sont le reflet de cette époque où l'on peut observer une avancée certaine dans tous les domaines de la future "Criminalistique".

Sherlock Holmes a fait avancer l'ensemble des méthodes techniques et scientifiques, permettant le constat, la recherche et l'identification du criminel. Edmond Locard (1877-1966), grand criminologue français, fondateur de l'Académie Internationale de Criminalistique et « holmésien » convaincu, conseillait aux policiers en formation de lire les Aventures de Sherlock Holmes ; et aujourd'hui, les Techniciens en Identification Criminelle appliquent à la lettre la méthode du détective britannique, qui *"était fondée sur l'observation des riens"*...

La criminalistique au service de l'intuition

La police scientifique fait appel aux sciences et aux méthodes les plus variées (Diaz, 2005). On citera les principales dont certaines ont été utilisées dans les enquêtes de Sherlock Holmes pour vérifier ses géniales intuitions.

- **la balistique** - En 1910, les forces de police commencent à recourir aux indices balistiques dans leurs enquêtes sur des crimes où sont utilisées les armes à feu.

⁶ June Thomson, dans son livre "Watson et Holmes", explique que *"Christine L. Hubert a suggéré que la méthode employée utilisait l'hydroxyde de sodium ainsi qu'une solution saturée de sulfate d'ammonium. Lorsque ces éléments sont ajoutés à de l'eau distillée ne contenant qu'une goutte de sang, une poussière brunâtre est précipitée, dénotant la présence d'hémoglobine"*.

Dans l'affaire de la Maison vide (avril 1894), les charges qui pèsent sur le Colonel Moran au sujet du meurtre de Ronald Adair et de la tentative d'assassinat sur la personne du détective sont abandonnées, faute de preuves irréfutables sur le lien entre les balles explosives de revolver récupérées aux domiciles de Adair et du 221 B Baker Street, et du fusil à air de l'ingénieur allemand von Herder trouvé en la possession de Moran.

- **les documents** - Trois catégories d'études doivent être distinguées : la comparaison d'écriture (trouver les différents points de correspondance entre deux documents) ; l'examen graphologique (déterminer la personnalité du scripteur) ; et la chromatographie (comparaison des encres).

Au sujet des deux premières catégories, dans le problème intitulé par le Docteur Watson "Une affaire d'identité" (avril 1889), le détective compare par exemple deux séries de documents tapés à la machine, une émanant de Hosmer Angel, disparu le jour de son mariage, et l'autre de James Windibank, le beau-père de la mariée. Partant du principe que les machines à écrire présentent des singularités, autant que l'écriture humaine, Holmes prouve ainsi que Windibank et Angel sont une seule et même personne...

La troisième catégorie peut être illustrée par l'affaire des deux femmes (1886) où un document « authentique » et objet d'un chantage, daté de 1848, porte une signature faite à l'encre bleu-noir, encre inventée en 1856...

De même, quand on trouva le « journal intime » de Jack l'éventreur en 1911, on pratiqua de nouveaux tests pour analyser l'écriture manuscrite et les matériaux. Actuellement on ne sait toujours pas si le journal est authentique ou non (Owen, 2000).

- **les empreintes digitales** - Dans l'affaire de l'entrepreneur de Norwood (août 1894), la police et le détective privé prennent en compte une empreinte de pouce ensanglanté laissée sur un des murs du couloir d'entrée par l'assassin de Jonas Oldacre ; mais avec deux interprétations contraires, Holmes va prouver avec subtilité l'innocence de son client. Les empreintes sont désormais scannées systématiquement à chaque entrée sur le territoire américain dans la lutte anti-terrorisme menée par ce pays depuis le « 11 septembre ». ⁷

- **l'identification des véhicules** - Dans l'affaire de l'école du Prieuré (mai 1900), grâce à une parfaite connaissance des différentes marques de pneus de bicyclette, Holmes arrive à retracer les événements à partir de la disparition du jeune Lord Saltire et du professeur d'allemand ...Heidegger (?).

- **l'entomologie légale** - Les méthodes classiques pour fixer la date d'un décès s'appuient sur "les changements post-mortem" (rigidité cadavérique, refroidissement, évolution de la putréfaction, lividité). On peut désormais déterminer la datation de la mort à partir des insectes identifiés sur le corps dont on sait qu'ils se succèdent dans un ordre invariable.

⁷ Ce qui signifie désormais une criminalisation *totale* de toutes les informations recueillies lors du déplacement des personnes....

- **l'anthropométrie** - Le procédé donne les caractéristiques d'un individu (taille, longueur, largeur maximales de la tête, longueur du médius gauche, et du pied gauche, couleur de l'iris gauche). Alphonse Bertillon a appliqué avec succès en 1880 une méthode en matière de police permettant l'identification des malfaiteurs récidivistes. Cette méthode intéressera Sherlock Holmes.

Soumise à de graves erreurs de mesure, cette technique va être supplantée par la dactyloscopie c'est-à-dire l'identification des individus à partir de leurs empreintes digitales. La méthode, ancienne, – la pression du pouce sur une tablette d'argile mou était déjà utilisée comme signature 3000 ans avant J.C. – est renouvelée en 1892 par Sir Francis Galton, un savant anglais, cousin de Charles Darwin. Karl Landsteiner (Prix Nobel 1930) découvre en tant que biologiste les groupes sanguins et le facteur Rhésus, ce qui permet d'exclure avec certitude tel individu dans une enquête criminelle où des traces de sang ont été révélées.

D'autres branches de la police scientifique se sont développées mais elles sont trop récentes pour que Conan Doyle ait pu s'y référer : **la biologie légale** avec la découverte de l'A.D.N. et sa mise en pratique à partir de 1986 dans les enquêtes sur les crimes sexuels ; les activités liées à l'**identification humaine** (reconstituer un visage à partir du crâne, vieillissement...) ; **l'anthropologie légale** (étude du squelette) ; **l'odontologie légale** (étude des dents).

Il faut ajouter **l'informatique** et **l'intelligence artificielle** qui peuvent croiser toutes les données contenues dans des fichiers, et utiliser des modèles de décision et des profils de comportement pour identifier des classes de «suspects». Des logiciels s'inspirent maintenant de la méthode cognitive du détective pour simuler son raisonnement (Bourcier, 1995). HOLMES... en Grande-Bretagne est un logiciel qui suggère une liste de suspects potentiels et « Big Floyd » aux USA aide les enquêteurs à obtenir des profils psychologiques de tueurs en série d'après les indices disponibles.

3- L'abduction : une heuristique pour dénouer les énigmes ?

Mais il est une partie de l'activité de résolution d'énigmes sur laquelle j'aimerais revenir. En effet, dans la méthode de Sherlock Holmes, telle qu'elle est décrite, il est question du raisonnement holmesien. Cependant la description « logique » des raisonnements est la partie la moins développée du Canon. Par exemple, l'abduction qui a fait l'objet d'une recrudescence d'analyses depuis une dizaine d'années en sciences cognitives n'est pratiquement jamais citée alors qu'elle semble représenter le mieux la démarche cognitive et argumentative de Holmes.

On peut brièvement reprendre la description des trois raisonnements canoniques à partir d'Aristote et Pierce. La déduction est généralement décrite par son prototype, le syllogisme. L'induction est considérée comme un raisonnement statistique. Le troisième, l'abduction – ou, suivant la terminologie d'Aristote, la

réduction- est considérée comme un raisonnement de plausibilité, « un mélange d'induction et d'hypothèse » (Pierce) dont la formalisation doit le moins à la logique. Pour résumer, l'abduction crée l'hypothèse, l'induction la vérifie, la déduction l'applique.

Reprenons l'exemple canonique de l'abduction :

Tous les haricots de ce sac sont blancs

Ces haricots sont blancs

Ces haricots proviennent de ce sac

L'abduction est réservée à la méthode de raisonnement qui conduit à trouver de nouveaux résultats. Elle permet dans l'exemple ci-dessus de passer de l'observation triviale « les haricots blancs sont sur la table » à une nouvelle inférence « ces haricots viennent de ce sac ».

Rapprochons ce raisonnement de celui du policier (Walton, 2004):

La fenêtre a été forcée avec un couteau

Un morceau de la lame a été trouvé sur le bord de la fenêtre

Le meurtrier devait posséder un couteau.

Ce raisonnement est aussi celui des trois princes de Sérendip qui vont errer de par le monde et se livrer à toutes sortes de découvertes à partir de traces diverses. La *sérendipité* est devenue le terme qui désigne « la capacité à découvrir, à inventer, à créer ou à imaginer un résultat- une explication, une théorie, une loi- sans l'avoir délibérément cherché » (van Andel, Bourcier, 2001). Leur sagacité les entraîne loin, y compris en prison. Ils savent relever les indices de ce qui, par abduction, peut devenir une hypothèse qu'il restera à vérifier, par induction, ou mieux, si la preuve est forte, par déduction.

Les deux raisonnements de la méthode du célèbre détective

Revenons à Sherlock Holmes. La description du raisonnement qu'il mène fait référence soit à la déduction, soit au raisonnement synthétique : « *Vous avez préféré rendre compte d'affaires qui en elles-mêmes auraient été du dernier banal si elles n'avaient permis à ces facultés de déduction et de synthèse logique dont j'ai fait ma spécialité de se déployer hardiment.* »

Une partie de sa méthode est décrite comme le prototype de la déduction. Elle a longtemps été le terme qui englobait toutes les facultés mentales de Sherlock Holmes. Il parle d'ailleurs souvent de « démonstrations ». La déduction (ou induction) ne semble pas pourtant être le raisonnement spécifique de la découverte, ni de la résolution d'une énigme. On connaît l'argument des Sophistes suivant lequel la déduction partant de l'hypothèse connue ne peut permettre de découvrir des hypothèses inconnues. Holmes ne déduit ni n'induit. La déduction est un processus d'élimination, de soustraction qui a longtemps été d'un grand secours aux enquêteurs antérieurs au détective. Mais la déduction est

une opération qui part du général pour arriver au particulier. Elle peut être tout au plus un mode d'exposition *a posteriori*. L'induction va, elle, du particulier au général : lorsque Holmes conclut que le propriétaire d'un fiacre est un médecin de médecine générale quand il remarque son panier d'osier contenant divers instruments médicaux, il part effectivement d'un fait particulier pour arriver à une généralité. Mais aucun raisonnement inductif ne peut vérifier une hypothèse et en tout état de cause il ne pourrait suffire à en lancer de nouvelles.

Un deuxième raisonnement plus complexe est évoqué dans la méthode de Sherlock Holmes. Il est clairement cité dans l'aventure des Hêtres-Rouges : la synthèse logique. De quoi s'agit-il ?

La synthèse consiste à construire une hypothèse à partir de plusieurs éléments déjà existants. Une fois que Sherlock Holmes a effectué toutes ses observations et "*déductions*", il doit trouver le fil qui les relie intégralement. Il lui faut une histoire, ou une explication. "*Aucune combinaison d'événements n'échappe à l'explication humaine. Une sorte d'exercice mental, sans aucune garantie de vérité, m'indique une ligne possible qui correspond aux faits. C'est, je le confesse, un travail de pure imagination ; mais combien de fois l'imagination ne s'est-elle pas révélée mère de la vérité ?*" (La Vallée de la peur). Ce raisonnement est de toute évidence une abduction, souvent appelée « raisonnement de la meilleure explication possible ».

Son imagination l'amenait à un dialogue intérieur, méthode que l'on retrouve dans les travaux sur l'abduction judiciaire : "*Sherlock Holmes, quand il avait un problème à résoudre, pouvait demeurer des jours entiers, et même une semaine sans se reposer : il tournait et retournait les faits dans sa tête, les examinait sous tous les angles jusqu'à ce qu'il eût bien approfondi le mystère, à moins qu'il ne trouvât insuffisants ses renseignements*" (L'Homme à la lèvre tordue).

Dans sa démarche cognitive, il pouvait avoir jusqu'à "*sept explications distinctes ; chacune se rapportant aux faits tels que nous les [connaissions]*" (Les Hêtres-Rouges). Une seule de ces explications s'avèrera être la solution de l'énigme. D'où la célèbre maxime qu'il cite aussi comme une règle : "*Lorsque vous avez éliminé l'impossible, ce qui reste, aussi improbable que cela paraisse, doit être la vérité.*" (Le Signe des quatre).

L'abduction : le vrai raisonnement de Sherlock Holmes ?

Quels sont les travaux sur le raisonnement qui nous font penser que Holmes ferait appel au raisonnement abductif plutôt qu'à d'autres types de raisonnement ? L'abduction est considérée désormais plus comme un argument dialogique que logique (Blair, 1999). Examinons plusieurs recherches sur l'abduction. Pour Josephson (1994), la partition canonique en trois raisonnements (voir plus haut) peut être réduite à deux : il y aurait le raisonnement déductif et le raisonnement abductif qui comprendrait le

raisonnement inductif. La distinction reposerait sur le fait que la déduction est prédictive et l'abduction non. En effet, on ne peut fermer la recherche d'hypothèses par une déduction notamment en matière de procès pénal. Le résultat du procès n'est pas « prédictible », sinon il n'y aurait pas besoin d'entendre la défense. Pour Hintikka (1998), l'intérêt de dire qu'il s'agit d'une abduction porte sur l'ouverture possible à de nouveaux résultats. L'abduction est ampliative et la déduction tautologique. Enfin Wigmore(1913) considère que l'abduction équivaut à « analyser l'inférence pour une meilleure explication ». Cette analyse se fait dans le cadre d'un procès dans lequel les deux parties présentent leurs arguments et preuves selon des règles de procédure précises. La première étape consiste à établir l'existence de faits. Une question sur le pourquoi et le comment est alors demandée à un participant. Le troisième moment consiste à permettre à l'autre de répondre en offrant une explication. A travers ce dialogue, des alternatives sont exposées. On évalue alors ces explications et la meilleure est sélectionnée. Sherlock Holmes procède en réalité à un monologue intérieur où les arguments pour ou contre ne sont pas toujours explicites ou débattus. Etendue au procès, cette démarche pourrait contrevenir aux règles du contradictoire. On a pu dire cependant que si Holmes se contente de ce monologue, c'est qu'il connaît l'univers où il évolue de façon précise. L'objet qu'on découvre est celui qu'on invente (*venio in*). Inventer⁸ c'est découvrir l'inconnu. Inventer, c'est être sensible à ce qui « existe déjà », c'est *dévoiler* l'objet qui, présent, demeure invisible. Ainsi dit-on d'un trésor, qu'il a été inventé par la personne qui l'a découvert. L'adéquation entre les propositions de Sherlock Holmes et la réalité - son invention- doit tout à la symbiose entre le détective et le *déjà vu* de son monde de référence : *"Sherlock Holmes réussit toujours ses raisonnements avec la plus grande classe car ces derniers correspondent parfaitement à la réalité de son monde. Conan Doyle a simplement créé un univers à la mesure du héros, un univers dont le détective est le centre et qui reflète son égocentrisme. Le monde est construit pour et autour du personnage qui peut ainsi y prévoir (je souligne) l' "avenir" en choisissant les hypothèses les plus probables car il en connaît le fonctionnement. Et contrairement à Watson, pas d'échecs possibles pour le héros. Ainsi, toutes les abductions faites sur l'espace sont justes car cet espace est organisé en fonction des besoins de Holmes."* (Ravenel, 1987)

L'abduction peut-elle être assimilée à une synthèse ? La synthèse de l'ensemble des données recueillies lors du constat permet par un "raisonnement analytique" rigoureux de remonter des effets aux causes (abduction); sachant qu'il *"faut toujours examiner deux hypothèses : l'une pour, l'autre contre"* ; règle n°1

⁸ Voir Code civil, art. 716 : « Le trésor est toute chose cachée ou enfouie sur laquelle personne ne peut justifier sa propriété, et qui est découverte par le pur effet du hasard [...] L'inventeur d'un trésor s'entend de celui qui, par le seul effet du hasard, met le trésor à découvert ».

dans toute enquête criminelle." (Peter le Noir - juillet 1895 - publié en 1904). S. Holmes explique lui-même dans la rédaction de l'affaire du Soldat Bianchi que le processus de "synthèse-analyse-déduction" est fondé sur l'hypothèse que lorsque l'on a éliminé tout ce qui est impossible, il ne reste plus que la vérité, quelque improbable qu'elle paraisse. Il arrive que plusieurs explications s'offrent encore à l'esprit ; dans ce cas on les met successivement à l'épreuve jusqu'à ce que l'une ou l'autre s'impose irrésistiblement. Mais l'abduction semble plutôt relever de la construction d'un dialogue que de la démonstration du résultat.

L'intérêt du raisonnement abductif comme procédure dialogique

Telle qu'elle est étudiée actuellement notamment dans le procès judiciaire⁹, l'abduction tend à se distinguer de plus en plus de la logique et à se référer au *raisonnement dialectique*. Du raisonnement dialectique, elle tirerait sa force par la description systématique de mouvements argumentatifs mis en jeu contradictoirement.

Ensuite, elle servirait autant à établir une hypothèse qu'à collecter de l'information, proposer et évaluer une explication, résoudre un problème pratique ou un conflit d'opinion, structurer une plaidoirie ou une motivation, clôturer un dialogue. Autant de fonctions qui relèvent plus de la théorie de l'argumentation que de la logique de la démonstration.

Enfin d'une façon générale, l'intérêt de dire qu'il y a abduction réside en ce que les parties admettent que de nouveaux résultats ou de meilleures explications sont toujours possibles.

Il y aurait même un danger à considérer que l'on peut passer « facilement » d'une abduction à une déduction comme tendrait à le montrer le schéma de Pierce (voir plus haut). En effet, le mis en examen est présumé *innocent* et le doute du jury lui profite toujours. Un des principes fondamentaux de la procédure judiciaire est le principe du contradictoire : il présuppose que tout élément nouveau peut à tout moment renverser l'hypothèse de départ dans un débat assurant l'égalité des parties.

Dénouement

Une énigme se doit d'être dénouée. Mais le nœud est superposable à l'inconscient disait Lacan. Le dénouement, comme le transfert en psychanalyse, est donc à la fois résistance et interprétation. Il aurait été sans doute pertinent de relier les mécanismes de l'énigme oedipienne avec l'inconscient de Sherlock

⁹ Douglas Walton (2004) pour le raisonnement judiciaire : pp. 22, 31, 169, 229.

Holmes. On s'en gardera bien. Ce que l'on peut supposer cependant, au plus près de ce que peuvent nous apporter sciences cognitives et criminalistiques, c'est que ce qui structure le dénouement, c'est une fiction (littéraire) ou une présomption (judiciaire). Dans la découverte d'un sens, tant que les faits ne peuvent être insérés dans une déduction logique – ce qui n'existe que dans le monde idéal des mathématiques –, on ne se situe pas dans le discours de la vérité mais dans celui de la plausibilité ou de l'imagination. Dans une affaire criminelle ou un roman policier, c'est la cohérence de la narration qui rend l'enchaînement des faits et l'abduction acceptables.

En général, un faisceau d'indices ou de preuves peut suffire à *convaincre* le juge ou le jury, ou sinon, à introduire un doute qui profitera à l'accusé. Remarquons cependant une asymétrie : le témoin doit jurer de dire la vérité, toute la *vérité*, mais le juge produit un jugement qui sera *tenu pour vrai*¹⁰ puisque, suivant l'adage latin, *Res iudicata pro veritate habetur*.

Bibliographie

- P. van Andel & Danièle Bourcier (2001), "Serendipity and Abduction in proofs, presumptions and emerging laws", *Cardozo Law Review*, vol. 22, N° 5-6, p. 1605.
- J.A. Blair (1999), "Walton's argumentation schemes for presumptive reasoning : a critique and development" In *Proceedings of the 4 th I.C. I.S. for the Study of argumentation*, ed. F. H. van Eemeren, R. Grootendorst, J. A. Blair, Ch. A. Willard, Amsterdam, SicSat, pp. 56-61.
- Danièle Bourcier (1995), *La décision artificielle*, Paris, PUF, Les voies du droit.
- Danièle Bourcier & Monique de Bonis (1999), *Les paradoxes de l'expertise, Savoir ou juger*, Plessis Robinson, Les empêcheurs de penser en rond, Synthélabo.
- Arthur Conan Doyle, *L'ensemble du Canon*, Paris, édit. Le Livre de Poche.
- Arthur Conan Doyle & J. Dickson Carr, *Les exploits de Sherlock Holmes*, Paris, Le Livre de Poche.
- Charles Diaz (2005), *La police technique et scientifique*, Paris, PUF, Que Sais-je ?
- Jacques Fombonne (1996), *La Criminalistique*, Paris, PUF, Que Sais-je ?
- J.R. Josephson & S.G. Josephson (1994), *Abductive inference: Computation, Philosophy, Technology*, New York, Cambridge University Press, p. 28.
- J. Hintikka (1998), "What is abduction ?" *Transactions of the Charles S. Peirce Society*, 34, p. 503.
- David Owen (2000), *Crime et science, Les crimes à l'épreuve de la science*, Paris, Tana éditions.
- Loïc Ravenel (1987), *Les aventures géographiques de Sherlock Holmes*, Paris, Editions Larousse p. 226 à propos d'une article d'Umberto Eco, « Conjectures d'Aristote à Sherlock Holmes », *La Magazine Littéraire*, avril 1987.
- Sherlock Holmes et la France*, ouvrage collectif dirigé par la Société Sherlock Holmes de France, collection Bilipo - pp. 54 à 56.
- June Thomson (1996), *Watson et Holmes*, Paris, Edit. Librairie des Champs Elysées, coll. Le Masque.
- Douglas Walton (2004), *Abductive reasoning*, The university of Alabama Press.
- J. H. Wilmore (1913), *The principles of judicial proof*, Boston, Little & Brown, p. 25.

¹⁰ On a vu comment "se tient" parfois la vérité dans l'inacceptable premier procès d'Outreau, 2004.

Traces légales

Roger F. Cozien
234 rue Championnet 75018 Paris
email : roger_cozien@yahoo.fr

Résumé

Le concept ancien de «scène de crime» est central dans l'exercice de la police judiciaire. C'est le départ d'une série de gestes techniques et légaux visant à trouver les auteurs de l'infraction pénale. Sur cette scène de crime la recherche des traces laissées par les protagonistes du drame est devenue un acte légal essentiel et incontournable. Avec le temps, et l'avènement de nouveaux moyens de commission d'actes délictueux, les techniques de recherche, de prélèvement et d'exploitation de ces traces ont considérablement évoluées. On regroupe ces techniques sous le terme «criminalistique» ou «forensic» pour sa traduction anglaise.

Mots clés : police judiciaire, criminalistique, forensic, preuve scientifique, scène de crime, trace.

1 Introduction

En matière pénale et surtout de police judiciaire, il existe un principe fondamental et inamovible. Une infraction, pour être poursuivie doit impérativement reposer sur trois éléments consubstantiels : un élément légal, un élément moral et un élément matériel. Il s'agit là d'une union si forte que certains auteurs parlent de **trinité**.

L'élément légal signifie que l'action poursuivie doit être explicitement illégale et donc prévue et réprimée par un texte d'un des nombreux codes du droit français (*principe de légalité*) : tuer, voler, escroquer, faire un excès de vitesse, etc.

L'élément moral est plus subtil à expliquer. Son aspect le plus simple repose sur l'intentionnalité du ou des auteurs. Il faut donc expliciter les motivations du ou des auteurs, la conduite délictuelle volontaire, la prise de risque, la mise en danger d'autrui, ... démontrer donc que le ou les auteurs n'ont pas succombé à une force si puissante qu'ils ne pouvaient s'en soustraire et qui les aurait poussé à commettre l'acte illégal.

Enfin l'acte matériel impose que la constatation de l'action illégale, puis sa poursuite, s'appuie sur des éléments matériels qui relient le ou les auteurs à l'action illégale : un cadavre, l'arme du crime, une mesure de vitesse, des relevés de comptes bancaires, une trace de sang ou de sperme, de la poudre, des logs informatiques de connexion, etc.

L'élément matériel à une double portée : premièrement démontrer la matérialité des faits, mais également, dans certains cas, éclairer l'élément moral. Les éléments matériels peuvent par exemple permettre d'établir ou non l'intentionnalité de l'auteur, la bande organisée, ou encore la préméditation.

La modification, pour ne pas dire la complexification, du droit pénal, avec en particulier le renforcement de la nécessité de préserver les droits de la défense, l'avènement de nouvelles formes de crimes et délits, l'intensification de la violence des atteintes aux personnes et aux biens, ont imposé à la justice et aux forces de l'ordre une extrême rationalisation des moyens de prélèvement et d'analyse des traces laissées par les protagonistes du délit ou du crime.

Ainsi, la science s'est insinuée dans tous les domaines de la police judiciaire, amonts et aval, pour donner naissance à la criminalistique moderne. Chaque trace devient une énigme en elle-même et appelle donc la mobilisation de processus de résolution propres à la police judiciaire et à la criminalistique. Chaque trace, chaque énigme résolue, participe à la construction d'un sens particulier : la construction de cette trinité infractionnelle et plus spécifiquement en dévoilant l'élément moral et l'élément matériel.

2 La police judiciaire

En France comme dans tous les pays démocratiques de droit, la recherche des auteurs d'infractions pénales est dévolue à des services spécialisés et sous le contrôle permanent de la justice. Ces services sont dits de « police judiciaire » et relèvent principalement, dans notre pays, de la Gendarmerie Nationale et de la Police Nationale.

La police judiciaire (quelque soit le corps d'appartenance) est un auxiliaire indispensable à l'action de la justice répressive. Son rôle commence bien avant la juridiction de jugement. De façon générale, la police judiciaire est chargée de constater les infractions à la loi pénale, d'en rassembler les preuves et d'en rechercher les auteurs (*art. 14 de CPP*). Ces actes s'accomplissent lors de la phase policière, lors de ce que le Code de Procédure Pénale désigne sous le nom global d'*enquête*. Avant l'ouverture d'une information judiciaire, elle éclaire par ses investigations le Ministère public sur l'opportunité des poursuites. Après l'ouverture d'une information, elle seconde le magistrat chargé de l'instruction en agissant sous commission rogatoire.

Le CPP fait la distinction entre trois grands types d'enquête, plus un quatrième plus marginal : l'enquête préliminaire (*art. 75 à 78 de CPP*) et l'enquête sur infraction flagrante (*art. 53 à 74 du CPP*) sont les deux enquêtes de la phase policière. Nous trouvons également l'enquête sous commission rogatoire (*art. 81 et 151 à 155 du CPP*). Enfin, nous trouvons l'enquête suite à la découverte d'un cadavre. Dans tous les cas d'enquête, la police judiciaire procède à des investigations matérielles. Un des actes les plus fondamentaux consiste à veiller à la conservation de tout ce qui peut servir à la manifestation de la vérité. Il s'agit dans les cas simples d'objets usuels : armes, vêtements, argent, ... Cependant, dans les cas plus complexes, il ne s'agit pas d'objets mais de traces laissées par des objets, ainsi que les protagonistes du crime ou délit. Au cours du temps, il a fallu mettre au point et perfectionner sans cesse des techniques de prélèvement, de préservation et d'exploitation de ces traces.

L'avènement de nouvelles techniques viennent renforcer le champs d'action de la criminalistique. C'est en particulier vrai pour les traces microscopiques et les traces biologiques. Mais également, dans les domaines de l'électronique et de l'informatique, ou encore dans celui de l'identification des victimes. Tous les domaines de la science et des technologies sont mis à contribution. L'article 60 du CPP et la loi 99-515 du 23 juin 1999 permettent à la police judiciaire de recourir à toutes personnes qualifiées. En interne, Police et Gendarmerie se sont dotées de leurs propres laboratoires de police scientifique. Cependant, toute personne ou structure qualifiée peut concourir, par réquisition, à la manifestation de la vérité.

3 De la preuve

La recherches des éléments matériels, des traces et indices participe à l'idée générale de «rassemblement des preuves». La manière de se procurer ces preuves n'est pas entièrement libre. Elles doivent être obtenues selon des procédures que seule la loi peut réglementer. Ceci afin d'assurer l'efficacité de la preuve afin qu'elle soit incontestable, et d'autre part, afin d'éviter les abus qui pourraient résulter d'investigations sans limites.

Les contraintes imposées dans la collecte des preuves sont variables. La réglementation devient d'autant plus stricte et minutieuse que le procès pénal avance. Au début des investigations il est possible de se contenter de vraisemblances, il faut, au sortir des investigations préliminaires, établir les faits d'une manière aussi précise et indiscutable que possible. Si pendant longtemps dans l'histoire de la police judiciaire, les aveux, les témoignages, mais encore les couteaux ensanglantés ont suffi à forger la conviction des enquêteurs et des juges, l'évolution du droit impose un recours massif aux techniques d'investigations scientifiques.

Si on laisse de côté les aveux et les témoignages, la recherche de preuves passe nécessairement par les constatations matérielles. Elles ont pour but de relever directement les diverses circonstances qui éclaireront sur le commission de l'infraction ou sur l'identité des auteurs : traces, emplacement, objets, etc. C'est ce que l'on nomme des «pièces à conviction». Ces pièces à conviction, et c'est particulièrement vrai pour les traces, induisent des «présomptions». Il s'agit de faits ou indices qui établissent plutôt une probabilité qu'une preuve véritable. Ils sont pour cette raison qualifiés de «preuves conjecturales». La trace, que se soit une tâche de sang, de sperme, une empreinte digitale, un fichier informatique, est typiquement de cette nature.

Cette nature conjecturale ne va pas sans poser de problème lors du procès. En effet, c'est le juge qui en apprécie librement, en suivant son intime conviction, leur force probante. Il peut donc en toute légalité les faire prévaloir sur un mode de preuve plus immédiat, comme le témoignage ou l'écrit. Et c'est là que réside toute l'ambiguïté : la loi considère les indices comme des moyens de preuve secondaires. Ainsi, un procès verbal, qui fait foi jusqu'à preuve du contraire peut être combattu par des écrits ou des témoignages, mais non par des indices (*art. 431 et 537 du CPP*).

On comprend donc que la notion de preuve est très relative et conjecturale. Parallèlement, l'imprécision, la fragilité et la relativité du témoignage humain, ont largement été démontrées dans l'histoire judiciaire. Ainsi, au cours du temps, les preuves indiciales ont gagné une place

nouvelle dans l'intime conviction des juges. En contre partie, la preuve indiciale a dû conquérir ce nouveau statut en présentant la rigueur de l'approche scientifique. Ainsi, la criminalistique est-elle née de cette confrontation entre les exigences formelle du droit positif et les faiblesses humaines. Dernièrement, les dispositions relatives à la préservation de la présomption d'innocence renforcent l'importance de l'expression de la preuve matérielle par la voie scientifique et par le droit du prévenu à rester silencieux durant sa garde à vue.

4 De la trace à la preuve scientifique

Nous allons maintenant parler de *criminalistique*, qui est habituellement considérée comme la science des causes du crime. Avec le temps, et la science progressant, la criminalistique couvre l'ensemble du phénomène criminel. Nous n'allons pas expliquer ici l'ensemble des domaines de la criminalistique moderne tant ils sont nombreux : balistique, empreintes digitales, écritures, documents, traces biologiques, fibres et textiles, incendies et explosifs, photographie, spectrométrie, médecine légale, anthropométrie, informatique, entomologie, etc.

4.1 La scène de crime

La criminalistique (ou police scientifique) peut être mobilisée à de nombreux moments, plusieurs fois durant la même enquête, sur des événements importants comme marginaux. Cependant, nous pouvons en règle générale poser le départ du processus criminalistique sur la «*scène de crime*». Contrairement à ce que l'on peut voir dans les séries télé, cette notion va bien au delà de sa dimension purement géographique : les rubans marqué POLICE autour du cadavre ! De fait, l'acception anglo-saxonne est en train de devenir la norme : la scène de crime est l'ensemble des éléments connectés à une infraction, quelles qu'en soit la forme. Nous pouvons donc y inclure les lieux à proprement parlé, les victimes, les témoins, les suspects, les objets, et bien sûr, les traces en tout genre.

D'autres endroits peuvent y être inclus relativement à leur fonctionnalité : endroits de rencontres, chambres d'hôtel, des véhicules, etc. Durant la phase policière, voire durant l'instruction, la criminalistique pourra être employée dans chacun de ces lieux ou sur chacun de ces objets ou ... individus. Il faut noter que le terme de «*scène de crime*» est circonstanciel. En effet, la police scientifique peut sans distinction être aussi employée sur les accidents. On se rappelle les tragiques accidents du Concorde ou du tunnel du Mont Blanc. Ainsi, de façon très générale, la criminalistique participe pleinement à la «*manifestation de la vérité*».

L'importance de la scène de crime est primordiale. On parle donc de gestion et de gestionnaire de scène de crime. Le but d'une telle gestion est de rechercher de façon dynamique et d'exploiter au mieux tous les éléments d'information qu'elle peut receler, au premier rang desquels les traces en tout genre. Nous pouvons classer ces traces, et plus globalement les indices trouvés sur la scène de crime, en trois catégories : les indices matériels visibles. Les traces latentes invisibles qui ne peuvent être révélées que par des procédés spécifiques. Enfin, nous trouvons les indices immatériels.

Les indices ou traces latentes sont les plus nombreuses. Si l'auteur ne peut que difficilement les faire disparaître (à ne pas confondre avec la dissimulation), elles sont particulièrement fra-

giles. Ainsi, les techniques forensiques ont également dû adopter une démarche qualitative en vue de la préservation de ces traces : protection draconienne des lieux, une recherche exhaustive des traces et indices, utilisation de procédés scientifiques spécifiques et éprouvés pour la révélation des traces latentes. Dans la pratique anglaise ces trois éléments se trouvent appliqués bien plus drastiquement qu'en France. Il y a désignation d'un gestionnaire de scène de crime chargé de la protection des indices latents et garant des intervenants.

En tout état de cause, l'évolution «des» droits européens, mais surtout des sciences et techniques, tend à faire converger les méthodes criminalistiques. Il existe un domaine qui est particulièrement «convergeant» tant en méthode qu'en droit, celui de l'exploitation des traces contenant de l'ADN.

4.2 Traces génétiques

Indéniablement, les dernières années ont vu le considérable renforcement de l'importance des empreintes génétiques dans la panoplie de la criminalistique. Avant 1985, et les travaux d'Alec Jeffreys sur l'ADN, l'identification des individus dans le domaine biologique se faisait par les groupes sanguins. Cette pratique était très peu discriminante et permettait surtout des exclusions.

Il existe de nombreuses méthodes d'analyse du matériel génétique, et nous trouvons des différences entre les pays qui sont à même de les réaliser. Cependant, l'ensemble des recherches dans ce domaine consiste à rendre exploitable des prélèvements de plus en plus petits, de plus en plus potentiellement souillés et, sur tout type de support.

La question récurrente associée à ce genre de méthode concerne le risque d'erreur. La réponse est simple si les profils génétiques sont différents : l'exclusion sera certaine. À contrario, si les profils sont identiques, l'identité ne sera jamais totalement certaine. Cependant, l'utilisation de plusieurs méthodes de «révélation» et l'étude des fréquences des séquences d'ADN dans la population permettent de dire que la comparaison des empreintes génétiques dépasse de très loin la fiabilité de toutes les méthodes de discrimination et d'identification biologique. Il est surtout extrêmement surprenant de savoir que ces résultats sont possibles à partir de traces extrêmement infimes et variées.

Les limites de cette technique concernent premièrement, la quantité et la qualité du matériel. C'est certainement la limite la plus contraignante. Des expériences faites sur des mégots de cigarettes permettent d'estimer à 50% les chances d'établir un profil génétique, avec au moins un système d'amplification. C'est pour cela que toutes les étapes de détection, de prélèvement et de transport, avant le laboratoire d'analyse, doivent faire l'objet de soins extrêmes. La deuxième limite importante réside dans la contamination de l'ADN prélevé dans les traces laissées sur la scène de crime. Cette ADN présente sur la scène de crime, se nomme «ADN matrice». Toute autre ADN matrice polluant les prélèvements sera également amplifiée lors de l'analyse. Il peut, par exemple, s'agir de l'ADN des enquêteurs.

Les indications d'utilisation des empreintes génétiques sont nombreuses. Citons en premier lieu les tests de paternité. Mais également, des cas bien plus sordides comme les agressions

sexuelles. Certes, les traces laissées par le ou les agresseurs sont exploitées, mais le transfert des cellules de la victimes sur les agresseurs apporte également des éléments potentiellement discriminants. En cas de viol en réunion, les empreintes génétiques peuvent apporter des informations sur le nombre d'agresseurs. Soulignons également, l'apport important de cette technique dans l'identification des cadavres, y compris lors des catastrophes majeures.

4.3 Le cadre juridique des empreintes génétiques

Se sont les deux lois de bio-éthique qui fixent le contexte juridique dans l'utilisation légale des technologies liées à l'ADN, en particulier les conditions de mise en œuvre de l'identification d'une personne. Il existe également un décret qui organise les conditions d'agrément des personnes habilitées à réaliser légalement les tests génétiques judiciaires. Nous trouvons donc la loi n° 94-653 du 29 juillet 1994 relative au respect du corps humain. La loi n° 94-654 de la même date, relative au don et à l'utilisation des éléments et produits du corps humain. Enfin, le décret n° 97-109 du 6 février 1997 relatif aux conditions d'agrément de personnes habilitées à procéder à des identifications par empreintes dans le cadre d'une procédure pénale.

Créé en 1998 par la loi du 17 juin relative à la prévention et à la répression des infractions sexuelles, le «*Fichier National Automatisé des Empreintes Génétiques*» (FNAEG) est commun à la Police Nationale et à la Gendarmerie Nationale. Cette création consistait à combler le retard de la France dans le domaine de la préservation des échantillons génétiques. Ce fichier est destiné à centraliser les prélèvements des traces génétiques des personnes condamnées pour crimes et délits sexuels. Il est également possible de comparer les empreintes de suspects aux données du fichier, mais sans qu'elles soient conservées. Le stockage des échantillons à proprement parlé est doublé d'un système informatique de gestion. Le système utilisé a été emprunté à la police fédérale nord américaine et est opérationnel depuis 2003. Le système policier français bivalent fait que la Police Nationale est en charge de la partie informatique, tandis que la Gendarmerie Nationale est chargée du stockage et de la préservation des échantillons.

Dans le cadre de la loi relative à la sécurité intérieure, le champ du FNAEG devrait bénéficier d'une extension de son champ d'application propre à la délinquance de masse. Cette loi a étendu le FNAEG à la quasi-totalité des crimes et délits d'atteintes aux personnes et aux biens. Ainsi, les profils génétiques des auteurs présumés de violences volontaires correctionnelles, de vols simples ou aggravés, de recel et de blanchiment, ainsi que les auteurs présumés d'infractions à la législation sur les armes, pourront faire l'objet d'une insertion dans le FNAEG, comme leurs empreintes digitales le sont dans le FAED.

Cette loi a introduit l'examen des profils génétiques dans le cadre des procédures de recherche des causes de la mort ou de recherche des causes d'une disparition inquiétante. Cette disposition permet à l'office central chargé des disparitions inquiétantes de personnes de bénéficier d'un outil d'aide à l'enquête complétant l'apport du fichier des empreintes digitale. Il existe d'autres dispositions importantes qui, tout en réaffirmant le contrôle de l'autorité judiciaire sur le FNAEG, vont faire du FNAEG au XXI^{ème} siècle, ce qu'aura été le fichier des empreintes digitales au début du siècle précédent. La trace continue donc d'être, et pour longtemps, un des éléments clé de l'enquête policière.

Pour être complet sur le FNAEG, retenons que sa base de données est alimentée par les traces ou échantillons biologiques recueillis dans le cadre d'une enquête concernant l'une des infractions mentionnées à l'article 706-55 du Code de procédure pénale, à savoir :

- Les traces biologiques issues de personnes inconnues ;
- Les échantillons biologiques prélevés sur des personnes à l'encontre desquelles il existe des indices graves ou concordants rendant vraisemblable qu'elles aient commis l'une des infractions susmentionnées ;
- Les échantillons biologiques prélevés sur des cadavres non identifiés ;
- Les traces biologiques issues de personnes inconnues recueillies pour des recherches des causes de la mort ou pour recherche des causes d'une disparition inquiétante ou suspecte ;
- Les échantillons biologiques prélevés sur des ascendants et descendants d'une personne disparue ;
- Les échantillons biologiques prélevés sur des personnes définitivement condamnées.

5 L'exemple Canadien

L'importance des empreintes génétiques est parfaitement connue dans de nombreux pays. Parmi les pays occidentaux industrialisés, la France est loin d'avoir été la première à rationaliser son emploi, en particulier par la mise en place d'un fichier central. Dès 1995 la Grande Bretagne avait son propre système. Le Canada est un pays en pointe pour les sciences forensiques. Les autorités canadiennes ont largement contribué à développer l'utilisation des empreintes génétiques. Les techniques de prélèvement, d'amplification et de comparaison sont les mêmes qu'en Europe. Dans la deuxième partie des années 90, le Canada s'est engagé dans une large réflexion sur l'utilisation de la génétique à des fins légales. En particulier, il a été mis en place une banque des échantillons génétiques qui sera alimenté par une liste d'infractions bien plus longue qu'en France.

Ce qui est intéressant, c'est que le Canada avait engagé une réflexion assez large sur l'utilisation de cette technologie à des fins légales. Il existe au Canada un Commissaire à la protection de la vie privée (CPVP). Cette autorité, tout en soulignant l'intérêt des empreintes génétiques pour la police judiciaire, s'est opposée par principe à la constitution d'une banque de données de caractère génétique de la population masculine. Elle met un avant le risque de dérive dans l'utilisation de cette base. Pour ma part, je n'ai pas trouvé particulièrement pertinents les arguments avancés, qui ont principalement consisté à agiter une fois de plus l'épouvantail des atteintes aux libertés civiles. Le commissaire propose également une série de six restrictions au cas où l'on créerait finalement une telle banque.

À l'opposé, l'Association du Barreau Canadien (ABC) a accueilli favorablement la création d'une telle banque de données pourvu que les libertés individuelles, dont la vie privée, soient respectées. Elle a cependant proposé certaines restrictions dans l'usage de la banque : utilisation à des fins limitées, empreintes génétiques uniquement associées à des crimes et délits graves et violents. Dans le cas des jeunes contrevenants, que les prélèvements soient préservés pour une période limitée. L'ABC se prononce contre le prélèvement obligatoire d'échantillon lors de l'arrestation. Que des personnels de la Police soient spécialement formés au prélèvement. Enfin, tout comme le CPVP, l'ABC recommande que seuls les résultats d'analyse soient conservés dans une base de données, et non pas les échantillons.

Enfin, l'Association Canadienne des Policiers (ACP) estime que les seules restrictions à appliquer à une telle banque sont celles appliquées à celle des empreintes digitales aux termes de la loi sur «*l'identification des criminels*». Ainsi, l'ACP préconise que le matériel génétique soit utilisé dans la pleine expression de ses potentialités. Toutefois, l'ACP indique qu'une telle banque de données ne doit servir que la justice et l'identification des criminelles.

Après toutes ces discussions et études, une banque d'échantillons génétiques a finalement été mise en place au Canada. Dans toutes les provinces et les territoires du Canada, les agents de police ont appris à prélever des échantillons d'ADN, qu'ils envoient ensuite à la Banque nationale de données génétiques pour analyse. À l'aide d'une trousse spéciale, ils prélèvent quatre gouttes de sang de personnes reconnues coupables de certains crimes. Tout au long de l'analyse, l'échantillon n'est identifié que par un simple code-barres, afin que le personnel de la Banque ignore l'identité du donneur. Des empreintes génétiques sont également prélevées sur les lieux du crime : à partir de taches de sang, de cheveux et même de cellules de la peau laissées sur l'arme du crime. S'il se trouve que l'empreinte génétique relevée sur les lieux du crime correspond à celle d'un criminel connu, il y a communication avec le laboratoire pour connaître l'identité de la personne en question. Un second échantillon de cette personne sera alors prélevé, lequel servira au procès. Les informations fournies par la Banque sont employées strictement à des fins d'application de la loi. Tous les autres usages, y compris la recherche médicale, sont prohibés par la loi et sujets à sanction.

Les enquêteurs canadiens ont également remarqués que depuis quelques années, les juges s'appuient de plus en plus sur les preuves médico-légales plutôt que sur les témoignages. Cependant, des juristes canadiens font des rappels de prudence, comme Laura Stevens (juriste et criminaliste) : «l'empreinte génétique prouve simplement que les cellules d'une personne se trouvaient à un certain endroit, mais en aucun cas que cette personne a commis le crime. L'empreinte génétique peut toutefois constituer de loin la principale preuve à conviction. Cette seule preuve peut occulter le reste de la cause».

De la même façon Brad Odsen explique que «L'analyse de l'ADN est certes un outil médico-légal très puissant et fiable. Mais n'oublions pas que pendant 40 ou 50 ans, les cheveux et les fibres constituaient les meilleures preuves à présenter lors d'un procès. De fait, les cheveux et les fibres étaient d'importantes pièces à conviction dans la cause de David Milgaard (un citoyen de la Saskatchewan déclaré coupable d'un meurtre en 1969, mais innocenté en 1997 grâce à des tests d'ADN). Il y a quelques années, tout le monde croyait faire ce qu'il fallait en utilisant des cheveux et des fibres comme pièces à conviction. Je me demande ce que les scientifiques penseront des preuves génétiques dans 40 ou 50 ans». Il poursuit en disant «il faut rester vigilant et sceptique. Ce n'est pas parce que le gouvernement affirme que les empreintes génétiques sont là pour notre bien que nous devons les accepter. Ce n'est pas parce que la loi dit quelque chose que cette chose va automatiquement se produire. Les gens qui instaurent le système à l'origine peuvent avoir les meilleures intentions du monde, mais les choses changent. Les gens changent aussi. L'information, elle, reste. Quel usage en fera-t-on ?»

Enfin, citons Randy Topp (enquêteur) : «De par ma fonction qui est de rechercher la vérité, j'attache une grande importance à la preuve par l'ADN. L'empreinte génétique a permis

de résoudre de nombreuses enquêtes criminelles et de gagner d'innombrables heures lors des enquêtes et des procès. Mais il existe d'autres types de preuves. Il n'y a pas que l'ADN. Il ne faut pas l'oublier et alors les enquêteurs auront un meilleur dossier à présenter au tribunal».

6 Conclusion

Dans toute enquête de police judiciaire, hier comme aujourd'hui, la *trace* est au centre des investigations. Ormis les cas simples où le criminel est, par exemple, filmé en train de tuer sa victime, il n'y a que les traces de l'acte délictueux/criminel qui seront à la portée des enquêteurs et de la justice. Ces traces sont de natures et de quantités extrêmement variées. De la même façon, les recueillir fera appel à des techniques plus ou moins perfectionnées. Quoiqu'il en soit, nous nous sommes rendus compte du statut particulier de la trace lors d'investigations judiciaires. La trace n'est pas en elle-même que très peu souvent porteuse de sens. En effet, on s'attend bien souvent à trouver un type de trace précis sur une scène de crime. La trace est plutôt un vecteur de révélation de sens. Pour ce faire, la criminalistique est un moyen de révélation qui va permettre à la trace de produire tout son sens, dans le contexte de l'enquête.

Les traces sont pour les enquêteurs des clichés, des polaroids de l'action délictueuse ou criminelle. Ils parsèment la scène de crime, et ouvrent une étroite et fugace fenêtre sur les événements qui la constituent. Mais ils souffrent de graves carences car ils ne comportent en général pas d'éléments de datation, ni n'exhibent explicitement les liens qui les relient entre eux. Certes, la criminalistique en s'appuyant sur la science peut aider à la révélation de ces éléments. Mais cela n'est en général pas suffisant. Il faut aux enquêteurs produire du sens autour des traces. Il n'y a pas de règle en la matière, l'expérience, le contexte de l'affaire, le droit, etc. participent tous au processus. Depuis quelques années une nouvelle discipline nommée «*analyse criminelle*», épaulée par des logiciels spécialisés, tente d'aider les enquêteurs.

Avec les possibilités que donne la criminalistique moderne, nous devons considérer que le statut de la trace dépasse largement la seule caractérisation de l'élément matériel. La quantité d'information apportée par un trace est maintenant virtuellement bien plus grande que sa seule présence pourrait apporter. Nous nous sommes attardés sur les traces génétiques, mais il y aurait long à dire sur la présence de certaines cohortes d'insectes sur un cadavre permettant de dater précisément l'heure de la mort, voire de dire si le corps a été déplacé. La scène de crime est potentiellement parsemée de traces qui vont aussi renseigner les enquêteurs sur les éléments moraux du délit ou du crime, en particulier l'intentionnalité. C'est une conquête considérable dans le fonctionnement de l'enquête judiciaire et cela ajoute au déséquilibre entre le témoignage et la preuve indiciale. La démarche de la criminalistique, son importance croissante, valide l'hypothèse simple qu'on ne peut pas ne pas interagir avec notre environnement. Plus encore, le niveau d'interaction est directement proportionnel au niveau de violence (qui n'est pas que physique) déployée pour la commission de l'infraction. Les traces ne sont que le reflet de ces interactions, elles deviennent légales parce que c'est la science qui en révèle toute l'information. Peut être parce qu'en la matière, la science est aussi aveugle que se veut la justice.

Signification et Contextes

Delphine Battistelli⁽¹⁾, Jean-Luc Minel⁽¹⁾ et Sylviane R. Schwer⁽²⁾

(1) Laboratoire LaLICC (UMR 8139 CNRS et Université Paris 4)

(2) LIPN (UMR 7030 CNRS et Université Paris 13)

*La Nature est un temple où de vivants piliers
Laissent parfois sortir de confuses paroles :
L'homme y passe à travers des forêts de symboles
Qui l'observent avec des regards familiers.
(Correspondances, Baudelaire)*

Résumé

Nous montrons à travers des exemples en mathématiques, informatique et linguistique, comment la notion de contexte intervient pour la désambiguïsation d'un signe donné. Nous nous interrogeons alors sur les problèmes de modélisation du contexte et des langages nécessaires.

En mathématiques, nous prenons l'exemple des différentes structures d'ordre induit par la relation "usuelle" \leq dans les ensembles de nombres entiers naturels \mathbb{N} , de nombres entiers relatifs \mathbb{Z} , des nombres rationnels \mathbb{Q} et des nombres réels \mathbb{R} . En informatique, nous regardons la notion de typage et l'interprétation d'une opération multiplicative. En linguistique, nous analysons la signification que l'on peut donner dans une phrase à un terme calendaire comme "le lendemain" .

1 Introduction

Pour parler de la « signification en contexte », on renvoie le plus souvent aux réflexions et travaux menés dans le cadre de l'étude des langues, à savoir la linguistique. On désigne alors par ce terme le résultat tout autant que la procédure qui permettent d'assigner à une unité linguistique donnée une signification dans le contexte dans lequel elle se trouve ; dans ce cadre, une des difficultés réside justement dans le fait d'être amené à préciser l'unité linguistique ainsi que le contexte.

Une certaine vision des différentes approches linguistiques de cette question peut alors amener à considérer ces dernières selon trois points de vue, correspondant aux trois façons de se prononcer sur l'unité sémantique de base de la langue : les approches *atomistes*, qui prennent comme unité de base la *lexie*, supposée avoir isolément une signification extralinguistique ; les approches *molécularistes*, qui prennent la *phrase* (ou la proposition grammaticale) comme véhicule premier de la signification ; (iii) les approches *holistes*, qui prennent pour unité sémantique de base le *discours*. Une quatrième approche, qui dépasse le cadre même de la linguistique, est la *pragmatique*, qui considère qu'un discours n'est en fait pas interprétable hors du *cadre culturel* dans lequel il occure.

L'atomisme est le cadre privilégié des terminologues, qui collaborent, comme l'a souligné François Rastier [18], à l'entreprise normative, qui a tât fait de transformer des usages en normes, et des normes en règles normalisant les discours de métiers, voire des sociétés. Les travaux effectués dans le cadre de l'approche d'exploration contextuelle [10] relèvent à la fois du molécularisme (association d'une valeur aspectuo-temporelle à une proposition) et du holisme (repérage dans les textes d'unités textuelles relevant de telle ou telle catégorie sémantico-discursive,

comme par exemple une *définition*). Les cadres du discours [4] relèvent quant à eux de l'holisme. Les travaux du cogniticien Dan Sperber et du linguiste Deirdre Wilson ont introduit la pragmatique au niveau atomiste [19] et certains travaux à l'origine de la SDRT [1] relèvent également de la pragmatique.

Dans cette conceptualisation strictement linguistique, il reste cependant à remarquer que ni l'unité sémantique de base ni le contexte n'ont de statut véritablement clairement défini, l'un comme l'autre pouvant renvoyer à différentes unités discursives (ou textuelles) : le mot (la lexie), la proposition grammaticale, la phrase, le paragraphe, le texte entier.

S'agissant plus particulièrement du contexte, on s'aperçoit que l'on peut également aborder ce dernier sous divers angles, qui vont du linguistique *stricto sensu* (c'est l'angle implicite sous lequel était abordé la notion de contexte dans ce qui a été dit précédemment à propos des approches linguistiques de la signification), à l'extralinguistique en passant par ce que l'on va pouvoir appeler le « contexte historique », à savoir l'époque et le lieu d'interprétation d'un signe donné ; ainsi par exemple, les signes mathématiques ont une histoire et une interprétation contextuelle. Aussi, dans cet article, nous tenterons de revenir sur ces différentes façons d'appréhender le contexte pour l'interprétation d'un signe, linguistique ou non.

La suite de cet article est divisée en deux parties principales. Dans la première partie, nous prendrons pour exemples les deux signes mathématiques correspondant aux deux relations fondamentales des mathématiques : = et <, avec une étude plus détaillée du signe <. Puis, nous rappellerons quelques aspects polysémiques en informatique. Dans une seconde partie, nous aborderons des problèmes purement linguistiques en abordant diverses « expériences » de modélisation linguistico-informatique : après avoir succinctement montré qu'analyser la signification d'un terme comme *le lendemain* en tant que purement calendaire conduit nécessairement à des erreurs d'interprétation, nous montrerons que l'étude de certains phénomènes linguistiques implique une approche holiste et nous présenterons un modèle formel de représentation prenant en charge ce problème.

2 du côté des langages formels

2.1 Les signes des deux relations fondamentales des mathématiques : = et <

Classer et ordonner sont les deux opérations les plus fondamentales de l'"homo rationalis". Toute classification ou taxonomie est subjective. Nommer, classer et hiérarchiser sont le propre de l'homme. Plusieurs mythologies du commencement racontent le geste d'introduction de l'ordre dans quelque chose de donné antérieurement, indépendant du Temps, un désordre initial (Tohu-Bohu, *κωρα*, un chaos, un réceptacle dans lequel tout est mélangé), et puis l'*On* vient et l'*On* y met bon ordre. Chez Platon, il y a avant le Temps un mouvement désordonné sans limite ni mesure, et le Demiurge y met de l'ordre en séparant et en distinguant. Dans la Bible, la création du monde se fait *ex nihilo*, en nommant les entités et en leur attribuant des fonctions diverses. La classification consiste donc à mettre chaque objet, chaque pensée, chaque événement, etc. dans une boîte étiquetée, qu'Aristote appelle catégorie, mais qu'on peut appeler classe, concept, etc. Le procédé repose généralement sur les critères de nature ou de fonction. Il s'agit d'ôter les caractères individuels pour ne traiter que du général, essentiellement parce qu'on a trop d'éléments à gérer. La hiérarchisation provient aussi du fait que l'on ne peut pas tout traiter à la fois. Mettre dans la même boîte est le fait mathématique d'une relation d'équivalence,

dont la plus fine (ou l'archétype) est l'égalité, notée $=$. Hiérarchiser est le fait mathématique de la relation d'ordre, dont l'archétype est la relation de *précédence*, notée $<$.

Nous allons rappeler l'origine de ces deux signes, mais compte-tenu que (i) $=$ reste un cas particulier du signe d'équivalence \cong , (ii) Violaine Prince [17] a proposé une réflexion approfondie sur l'ambiguïté de la signification de l'égalité et de son signe, nous nous intéressons ici tout particulièrement au signe $<$.

Dans les livres imprimés, jusqu'en 1557, l'égalité était généralement exprimée par des mots tels que *aequales*, *aequantur*, *esgale*, *faciunt*, *ghelijck*, ou *gleich*, mais parfois aussi par l'abréviation *aeq* (Cajori vol. 1, page 297). Le symbole ($=$) a été utilisé pour la première fois par Robert Recorde (1510-1558) en 1557 dans le *Whetstone of Witte*. Il utilisa une paire de parallèles plus allongées que le symbole actuel « parce qu'il n'y a pas deux choses plus égales ». Ce symbole ne réapparut pas avant 1618, dans un appendice anonyme d'une traduction anglaise de la *Descriptio* de Napier puis en 1631, sous les plumes de Thomas Harriot et William Oughtred [3].

La première apparition des symboles $<$ et $>$ date de 1631 dans *Artis Analyticae Praxis ad Aequationes Algebraicas Resolvendas* (Les arts de l'analyse appliqués à la résolution des équations algébriques, œuvre posthume de Thomas Harriot (1560–1621) : « Signum majoritatis ut $a > b$ significet a majorem quam b » et « Signum minoritatis ut $a < b$ significet a minorem quam b ». La légende raconte qu'alors qu'Harriot explorait l'Amérique du Nord, il vit un habitant portant un tatouage sur son bras : un symbole ressemblant à l'entrelacement des deux symboles $<$ et $>$ qu'il aurait alors désolidarisé pour l'usage relationnel plus petit et plus grand entre les nombres (entiers naturels).

2.1.1 Le succès du signe $<$

Notons que le signe $<$ est devenu l'archétype de la relation d'ordre [2, p. 7]. Ainsi, Hans Kamp [13] utilise ce signe pour signifier la précédence événementielle ; Elzbieta Hajnicz [12] utilise systématiquement ce symbole pour désigner l'ordre de toute structure temporelle ponctuelle.

2.1.2 Propriétés des structures d'ordre linéaires numériques

Nous nous proposons dans cette partie de revenir à la définition numérique initiale du signe $<$, employé à l'époque pour les nombres entiers positifs, pour les ensembles de nombres entiers naturels \mathbb{N} , de nombres entiers relatifs \mathbb{Z} , des nombres rationnels \mathbb{Q} et des nombres réels \mathbb{R} . Cette définition est la suivante : $x < y$ si et seulement si $y - x$ est positif.

$<$ induit une structure d'ordre linéaire discrète avec plus petit élément sur \mathbb{N} , ce qui signifie que tous les nombres entiers naturels admettent un successeur et tous, sauf zéro, un prédécesseur. Toute régression sur \mathbb{N} est finie, c'est-à-dire que quelque soit le nombre n choisi, on atteint 0 en un nombre fini d'étapes, ce qui est le moteur des preuves de terminaison de programmes.

$<$ induit une structure d'ordre linéaire discrète sans élément extrémal sur \mathbb{Z} . Tout élément possède un unique successeur et un unique prédécesseur. Les régressions sur \mathbb{Z} sont infinies. Un résultat fondamental de la théorie des ordres est le théorème suivant :

Etant donné un ordre linéaire non vide discret : (i) s'il admet un plus grand et un plus petit élément, alors il est en correspondance naturelle avec l'ordre naturel sur $\{1, 2, \dots, n\}$ où n est le cardinal de l'ensemble (nécessairement fini) ; (ii) s'il admet un plus petit élément mais pas de plus grand élément, alors il est en correspondance naturelle avec l'ordre naturel des entiers naturels \mathbb{N} ; (iii) s'il admet un plus grand élément mais pas de plus petit élément, alors il est en

correspondance naturelle avec l'ordre naturel des entiers négatifs ; (iv) s'il admet ni plus grand ni plus petit élément, alors il est en correspondance naturelle avec l'ordre naturel de l'ensemble de tous les entiers relatifs \mathbb{Z} .

\mathbb{N} admet de plus la propriété selon laquelle entre deux éléments a et b , il n'existe qu'un nombre fini d'éléments intermédiaires. Ce n'est pas vrai de tous les ordres discrets : prenons par exemple l'ensemble \mathbb{A} constitué de l'ensemble \mathbb{N} suivi d'une copie de $\overline{\mathbb{N}}$, entre 1 et $\overline{1}$, il existe une infinité d'éléments, bien que tout élément de \mathbb{A} possède un successeur et tout élément sauf 0 possède un prédécesseur. Cet ensemble est bien en correspondance naturelle avec \mathbb{N} en faisant correspondre à \mathbb{N} la suite des nombres pairs et à sa copie la suite des nombres impairs.

Un ensemble \mathbb{E} est dense pour la relation d'ordre \mathcal{R} si entre deux points x et y distincts quelconques de \mathbb{E} comparables, il y a un troisième point (propriété de dichotomie). L'ensemble des rationnels \mathbb{Q} et l'ensemble des réels \mathbb{R} sont des ordres denses pour l'ordre naturel. \mathbb{Q} reste un ensemble dénombrable, c'est-à-dire qu'on peut lui trouver un ordre discret, ce qui n'est pas le cas de \mathbb{R} . Le résultat suivant est également important :

Etant donné un ordre linéaire non vide dense et énumérable (\mathbb{E}, \leq) : (i) il existe une partie S de \mathbb{Q} , ensemble des nombres rationnels, tel que (\mathbb{E}, \leq) est en correspondance naturelle avec \mathbb{Q} , ordonné naturellement ; (ii) si (\mathbb{E}, \leq) admet ni plus grand ni plus petit élément, alors il est en correspondance naturelle avec l'ordre naturel de l'ensemble de tous les nombres rationnels \mathbb{Q} .

\mathbb{R} n'est pas dénombrable mais il est topologiquement complet pour $<$, c'est-à-dire que toute suite croissante majorée admet une limite supérieure.

Ainsi, le symbole $<$, devenu le signe générique des structures d'ordre (ponctuelles), désigne des ordres de natures très diverses sur les différents ensembles numériques. Cela explique peut-être une certaine confusion chez certains linguistes réclamant pour le temps un ordre à la fois dense (continu) et discret ou affirmant que les intervalles de \mathbb{R} sont formés de points contigus, niant *de facto* la densité de cet ensemble, puisque deux éléments sont contigus, s'il n'existe pas d'élément intermédiaire de même nature. Ainsi pour Co Vet [21, p. 38], « Nous le [Temps] le concevons comme un ensemble infini de moments ordonnés linéairement selon la relation \leq (antérieur ou simultané à) et comme étant de nature discrète et dense ». De même, pour Zlatka Guentcheva [11, p. 26–27] « L'intervalle [de Temps] est un ensemble orienté d'instant contigus », et d'expliquer : « L'intervalle est donc un continu construit d'instant contigus au sens où il est toujours possible d'insérer un instant entre deux instant ».

2.2 Dans les langages informatiques

Le principe consistant à utiliser le même symbole pour signifier des opérations différentes est celui de *polymorphisme*. Bien connu des langages orientés objets, il est utilisé fréquemment dans la plupart des langages de programmation ; par exemple en APL, la signification de chaque opérateur dépend du nombre de ses arguments (en général un ou deux). On peut alors dire que pour APL le contexte de l'opérateur est le nombre de ses arguments. Pour d'autres langages, ce sera le type de ses arguments. Mais déjà l'écolier a rencontré ce phénomène dans ses leçons de mathématiques puisque le signe $-$ peut signifier *le signe d'un entier relatif* (-2 est le signe négatif de l'entier relatif -2) ou *être l'opposé de* s'il a un unique argument ($-(-2)$ est l'opposé de -2) et *soustraction* s'il possède deux arguments ($5 - 2$ est la soustraction de 2 à 5).

3 Du côté de l'analyse linguistique

3.1 interprétation du terme lendemain

Les équipes impliquées dans la constitution de systèmes d'annotation temporelle des textes ont consacré l'essentiel de leurs efforts aux principes de calculs des liens entre les expressions temporelles dites calendaires et les calendriers (le calendrier grégorien par exemple). Ces travaux ont conduit à des systèmes performants. En revanche, peu d'efforts ont été consacrés à décrire les processus de calcul encodés, de façon à rendre compte de la façon dont l'humain utilise et exprime à travers la langue les informations temporelles à même d'être ancrées dans un système calendaire.

En effet, quelle signification doit on donner dans une phrase à des termes calendaires comme *le lendemain* dans des énoncés du type :

(1) Lundi, nous avons rejoint nos amis vers 22h au cabaret et nous étions couchés à deux heures du matin ; le *lendemain*, nous quitions la ville et visitons les Dupont.

(2) Lundi, Jean a visité les Durant, le *lendemain* les Dupont.

Les analyses purement calendaires conduisent nécessairement à des erreurs d'interprétation. En effet, les systèmes actuels, fondés sur les calendriers grégoriens ou spécifiques à des métiers calculent de la façon suivante les informations calendaires de (1) : deux heures du matin correspond à une journée du calendrier grégorien, en l'occurrence un mardi, puisque 2h du matin est postérieur à 22h, donc détermine le jour suivant. Le lendemain est donc calculé en mercredi. Le calculs pour (2) correspond au mardi. Supposons que les Lundi de (1) et (2) correspondent au même jour, tout auditeur pensera qu'il est possible que nous ayons rencontré Jean chez les Dupond, ce que le calcul nie. En effet, les systèmes actuels ne prennent pas en compte la relation forte qui lie un jour humain avec les alternances des périodes de veilles et de long sommeil et ne se conforme qu'au caractère normatif des calendriers. Partant de ce genre de constatation, nous travaillons sur un système de représentation qui, outre le fait qu'il permette de mettre en oeuvre ces calculs, soit plus proche de notre " compréhension " d'un texte qui implique différents modes de parcours (vertical, horizontal ou mixte) au sein d'un système calendaire comportant différentes références linguistiques fondées notions de C-chronologie (dans le cas d'une référence purement calendaire) et de E-Chronologie (dans le cas de référence événementielle). Cette dernière notion permet de créer de nouveaux calendriers fondés sur des événements cycliques comme les expositions, les types d'élection, . . .

3.2 Un exemple d'exploration contextuelle

L'approche développée tout d'abord dans le projet SERAPHIN [14] puis étendue dans le projet Filtext [15] est fondée sur la méthode d'exploration contextuelle [9, 10]. Cette méthode vise à identifier les connaissances linguistiques en les restituant dans leurs contextes et en les organisant en tâches spécialisées. Pour appliquer la méthode d'exploration contextuelle, le linguiste doit accumuler des marqueurs linguistiques qui peuvent être des unités lexicales simples ou complexes. Ces marqueurs linguistiques, lorsqu'on les identifie dans un texte, constituent des indices de certaines notions discursives. Certains indices, ceux dont la valeur sémantique est à désambiguïser, sont appelés indices déclencheurs. Le linguiste doit ensuite spécifier, sous forme de règles d'exploration contextuelle, les dépendances contextuelles entre ces indices. La partie « Action » d'une règle spécifie le segment textuel et précise l'annotation sémantique qui doit être attribuée à celui-ci. Chaque règle d'exploration contextuelle permet d'attribuer une étiquette aux

phrases qui contiennent les indicateurs et les indices pertinents et contribue ainsi à produire une structure hiérarchisée « décorée » par des informations sémantiques. Une dizaine d'annotations (« Annonce Thématique », « Conclusion », « Définition », etc.) ont ainsi été définies. Ces annotations sont celles utilisées dans le système de résumé Filtext mais on peut aussi faire appel à des annotations dans le domaine de la temporalité. Les règles définies permettent d'attribuer à des formes verbales conjuguées certaines valeurs sémantiques dites aspecto-temporelles. Rappelons que le système mis en place est fondé sur le principe selon lequel certains marqueurs linguistiques fournis par le contexte constituent des indices (complémentaires). Les marqueurs aspecto-temporels uniquement grammaticaux ne sont pas suffisants pour déterminer les valeurs sémantiques associées aux occurrences d'un morphème temporel : il faut tenir compte d'une part, du contenu lexical du verbe et d'autre part, de la coprésence de certains indices pertinents dans le contexte. Aussi, pour déterminer la valeur sémantique d'un marqueur grammatical aspecto-temporel, plusieurs composantes interviennent : (i) identification du marqueur morpho-syntaxique ; (ii) catégorisation sémantique des éléments lexicaux (verbes en particulier) ; (iii) identification d'indices repérés dans le contexte. Le système décide de la signification qu'il convient d'attribuer à une occurrence de forme en fonction de ces trois composantes. Les travaux de H.G. Oh [16] sur le passé composé ont permis de dégager un certain nombre de valeurs associées à ce temps. Nous donnons ci-après un exemple de règle permettant de calculer la valeur aspectuelle d'évènement accompli non-achevé du passé-composé.

[R111] Si l'on relève une occurrence X associée au morphème du passé composé
 & si X désigne un archétype dynamique de transitivité
 & si présence d'un substantif suivi de X
 & si présence de la préposition suivie d'une indication de durée
 ou présence d'un élément de la liste L13¹ dans la même proposition
 Alors X, sauf indication contraire, a la valeur de processus accompli non achevé qui donne naissance à un évènement

Exemples d'analyse :

Jean a peint le mur *pendant trois heures*.

L'ouverture du point de passage a permis peu à peu l'instauration d'un trafic à peu près régulier.

3.3 Modélisation du contexte linguistique

Comme le montre la section précédente, l'étude de certains phénomènes linguistiques implique la nécessité de prendre en compte la dimension textuelle, c'est-à-dire qu'il faut faire appel au contexte extérieur à la phrase pour les expliquer [4]. Les connecteurs qui marquent la progression thématique, les mécanismes de prise en charge énonciative, les différentes formes de reprise référentielle (pronoms, déictiques, anaphores, cataphores), les chaînes de liage, parmi d'autres, exigent le traitement de structures du texte qui vont souvent au-delà de la phrase. Pour pouvoir exploiter dans les logiciels de TALN les recherches qui sont menées dans cette direction, il est nécessaire de disposer d'outils formels qui permettent de travailler de manière systématique et consistante dans une dimension textuelle.

Nous décrivons brièvement dans cette partie sur quelle représentation nous nous appuyons pour modéliser un texte [6], et comment nous pouvons exploiter cette représentation pour assister un lecteur dans sa tâche de compréhension du texte.

¹La liste L13 contient des expressions temporelles définissant le processus, comme *peu à peu*, *progressivement*, *petit à petit*.

La modélisation du texte comporte deux aspects : l'un conceptuel : type d'information à représenter, unités à considérer, organisation entre les unités, mécanismes de manipulation de la structure, etc. ; et l'autre logiciel : comment représenter et manipuler l'information dans le cadre d'un développement informatique. Notre modélisation d'un texte s'inspire à la fois des propositions du modèle ContextO [7] et de celles du modèle TEI Light [20]. Afin de modéliser la structure logique d'un texte de la manière la plus générale possible en incluant les éléments textuels qui peuvent être le mieux exploités par notre plate-forme, notre vision d'un texte est celle d'une hiérarchie d'unités textuelles typées. La représentation d'un texte se divise en deux parties : le corps, où les unités textuelles se structurent d'une façon hiérarchique et la tête, où s'expriment les relations non hiérarchiques entre ces mêmes unités.

3.3.1 Le corps du texte : les éléments hiérarchiques

L'unité textuelle de base (dorénavant UT) comporte différents attributs selon qu'elle s'agisse d'une feuille ou d'un nœud dans la hiérarchie textuelle. De plus, les UT possèdent des attributs fixes et d'autres dynamiques, pouvant avoir, optionnellement, un titre. Une unité textuelle "feuille" correspond en général à une unité appartenant à la granularité la plus fine dans la hiérarchie définie ; c'est celle qui porte l'information lexicale, dénotée par un attribut fixe nommé "chaîne". Une unité textuelle "nœud" a pour rôle de révéler la hiérarchie textuelle, autorisant l'emboîtement des différentes unités. En conséquence, elle ne porte pas directement d'information lexicale. Il est néanmoins possible d'inférer l'information lexicale véhiculée par une UT nœud en parcourant la hiérarchie des UT.

3.3.2 La tête du texte : les éléments non hiérarchiques

Afin de représenter des constituants textuels ne pouvant pas être représentés dans le corps du texte, il est nécessaire de disposer de constructeurs de nouvelles unités textuelles. Les constructeurs proposés [5] sont : ensemble, séquence, référence et graphe, qui disposent des propriétés prouvées, facilement manipulables.

Les nouvelles unités textuelles construites observent ces trois règles :

- elles sont construites exclusivement à partir d'UT de base existantes dans le corps (i.e. elles ne peuvent pas créer de nouvelles UT de base) ;
- elles peuvent avoir des annotations leur appartenant, ces annotations ne pouvant pas être héritées par les UT de base ;
- elles peuvent se combiner afin de construire des éléments plus complexes.

3.4 Navigation textuelle

La modélisation présentée ci-dessus va être exploitée pour représenter visuellement un texte sous différentes formes techno-sémiotiques et va offrir la possibilité de proposer à un lecteur des opérations de navigation. Ces opérations, qui modélisent des connaissances de lecture d'un lecteur expert, pour un point de vue sur le texte (par exemple un point de vue de lecture fondé sur des expressions temporelles) ont pour objectif d'assister le lecteur dans son parcours de lecture. Ainsi, le lecteur associe les propositions de navigation offertes par le système de navigation avec sa propre interprétation du texte.

3.4.1 Connaissances de navigation

La navigation est conceptualisée comme une opération qui relie une UT (unité textuelle) source avec une UT cible. Une opération de navigation comprend une *source*, une *cible*, une ou plusieurs *conditions*, et un *empan*. L'exécution d'une opération est ainsi soumise à des conditions qui contraignent les attributs des UT considérées. L'*empan* de texte peut être spécifié de manière à restreindre l'espace de recherche des unités textuelles cibles. Chaque opération est typée avec une valeur qui appartient à l'ensemble $\{Premier, Dernier, Suivant[i], Précédent[i]\}$. Ces valeurs spécifient d'une part l'orientation, c'est-à-dire dans quel sens (avant ou après l'unité textuelle source) doit être effectué la recherche de l'unité textuelle cible, et d'autre part le référentiel, *absolu* (*Premier*, *Dernier*), ou *relatif* (*Suivant[i]*, *Précédent[i]*), par rapport à la source. Dans le cas d'un référencement relatif, l'index *i* permet de spécifier le rang de la cible recherchée. Par exemple, l'expression « Type= Suivant[3] » s'interprète comme la recherche de la troisième UT située après l'UT source et qui satisfait aux conditions spécifiées.

Les conditions expriment des contraintes sur les valeurs des attributs des UT source et/ou cible. Ces conditions simples peuvent être combinées entre elles avec les opérateurs logiques (ET, OU, NON). Afin d'illustrer l'intérêt de ces connaissances de navigation, reprenons l'exemple (1) :

(1) Lundi nous avons rejoint nos amis vers 22h au cabaret et nous étions couchés a deux heures du matin ; le lendemain, nous quitions la ville et visitons les Dupont.

Une opération de navigation qui proposera au lecteur, positionné sur l'ÔUT « lendemain », de se positionner successivement sur les expressions temporelles précédentes dans le texte, « deux heures du matin », « Lundi », fournira à celui-ci les moyens d'interpréter correctement ces expressions.

Références

- [1] Nicholas Asher, Alex Lascarides, The Semantics and Pragmatics of Presupposition *Journal of Semantics*, 1998, **15**, p. 239–299.
- [2] Jean-Michel Autebert, *Théorie des langages et des automates*. Paris Masson, 1994.
- [3] Florian Cajori, *A history of mathematical Notation*. Dover Publications - Chicago 1928/1993
- [4] M. Charolles, L'encadrement du discours : univers, champs, domaines et espaces. Cahier de Recherche Linguistique, Landisco, Université Nancy 2, (1997), 1–73.
- [5] [Couto, J., Minel, J.-L., Outils dynamiques de fouilles textuelles, *Actes de RIAO 2004*, Avignon, (2004) 420–430.
- [6] Couto J., L. Lundquist, J.L. Minel. Navigation interactive pour l'apprentissage en linguistique textuelle, *EIAH 2005*, Montpellier, (2005) 45–56.
- [7] G. Crispino, *Une plate forme informatique de l'Exploration contextuelle : modélisation, architecture et réalisation (ContextO)*. Application au filtrage sémantique de textes. Thèse de Doctorat, université Paris-Sorbonne.
- [8] Jean-Pierre Desclés, Jean-Luc Minel, Interpréter par exploration contextuelle, in *Interpréter en Contexte*, sous la direction de Francis Corblin, Paris, Éditions Hermès, p. 305-328, 2005.

- [9] J.-P. Desclés, C. Jouis, H.-G. Ho, D. Maire-Reppert, Exploration Contextuelle et sémantique : un système expert qui trouve les valeurs sémantiques des temps de l'indicatif dans un texte, *Knowledge modeling and expertise transfer*, D. Herin-Aime, R. Dieng, J.-P. Re-gourd, J.P. Angoujard (éds), Amsterdam, p. 371-400, 1991.
- [10] J.-P. Desclés, J.-L. Minel, Interpréter par exploration contextuelle, in *Interpréter en Contexte*, sous la direction de Francis Corblin, Paris, Éditions Hermès, p. 305-328, 2005.
- [11] Zlatka Guentcheva, *Temps et aspect : l'exemple du bulgare contemporain*, Éditions du CNRS, 1990.
- [12] Elzbieta Hajnicz, *Time Structures*, LNIA 1087, Springer Verlag, 1996.
- [13] Hans Kamp, Events, Instants and Temporal Reference, in *Semantics from Different points of view*, Springer Verlag, Berlin (1979) 376–417.
- [14] D. Leroux, J.-L. Minel, J.-L. Berri, SERAPHIN project, *Proc of the First European Conference of Cognitive Science in Industry*, Luxembourg, p. 275-283, 1994.
- [15] J.-L. Minel, E. Cartier, G. Crispino, J.-P. Desclés, S. ben Hazez, A. Jackiewicz, Résumé automatique par filtrage. Sémantique d'informations dans des textes, Présentation de la plate-forme FilText, *Technique et Science Informatiques*, 3, Paris, 2001.
- [16] H.-G Oh, Les temps de l'indicatif du français en vue d'un traitement informatique : passé composé, Thèse de doctorat, Université Paris-Sorbonne, 1991.
- [17] Violaine Prince, L'égalité à soi n'est pas l'identité. 12^{ème} journées de Rochebrune : rencontres interdisciplinaires sur les systèmes complexes naturels et artificiels, *Réflexivité et auto-référence dans les systèmes complexes*. ENST 2005 S 001, Paris, 207–218.
- [18] François Rastier, Le terme : entre ontologie et linguistique, *La banque des mots*, 1995, 7, p. 35–65. Accessible par http://www.revue-texto.net/Inedits/Rastier/Rastier_Terme.html
- [19] Dan Sperber, Deirdre Wilson, *Relevance : Communication and Cognition*, Oxford : Basil Blackwell (1986).
- [20] TEI 2005, Text Encoding Initiative Website : <http://www.tei-c.org>
- [21] Co Vet, *Temps, aspects et adverbes de temps en français contemporain*. DROZ, Genève, 1980

Sur les traces d'un temps fractal

Jérôme Cardot* et Sylviane R. Schwer†

Résumé

La plupart des modèles du temps, que ce soit en informatique, mathématique, philosophie ou linguistique se fondent sur les ensembles de nombres entiers naturels \mathbb{N} , de nombres entiers relatifs \mathbb{Z} , des nombres rationnels \mathbb{Q} et des nombres réels \mathbb{R} , en correspondance avec la droite géométrique réelle selon la structure d'ordre linéaire que l'on souhaite : discrète avec origine pour \mathbb{N} , discrète sans origine pour \mathbb{Z} , dense pour \mathbb{Q} , continu pour \mathbb{R} . À ceci près que le terme de continu est pour le moins ambigu dans la littérature non mathématique, se confondant plus souvent avec la notion de densité qu'avec celle de complétude. Ce n'est pas si grave que cela car en logique du premier ordre, il est impossible de différencier \mathbb{Q} et \mathbb{R} . En revanche, imposer \mathbb{R} comme représentation ultime du temps n'est pas une fatalité. En effet, l'utilisation du modèle des S -langages de Schwer pour représenter les relations temporelles qualitatives entre entités, avec l'usage des opérations standard de la théorie des langages formels que sont l'insertion et la substitution, nous ont conduits à penser à une structure fractale du temps. Cette intuition est maintenant confortée grâce au codage de Cardot des S -mots dans un ensemble de permutations, qui fait apparaître une répartition des S -mots sous la forme de triangle de Sierpiński, image fractale bien connue. Ces pérégrinations nous conduisent à poser la question d'une modélisation du temps par des ensembles de nombres non standard, et une redéfinition de l'instant, non comme ponctuel mais comme halo, rejoignant la vision leibnizienne des infinitésimaux petits.

Mots clés : S -langages, temps, fractales.

1 Introduction

La plupart des modèles du temps, que ce soit en informatique, en philosophie ou en linguistique utilisent sur les ensembles de nombres naturels, relatifs, ration-

*Université de Bretagne Occidentale – Brest

†LIPN UMR 7030 CNRS-Paris 13

nels ou réels, tels que les mathématiques modernes les définissent. L'ensemble des réels \mathbb{R} est un ensemble de points qui possède des propriétés topologiques qui permettent les opérations de calculs, notamment les calculs différentiel et intégral. La construction de \mathbb{R} est motivée par une propriété de continuité, qu'elle soit fondée sur une définition propre de passage à la limite ou bien sur une l'idée plus simple de coupure, ne correspond pas à l'idée intuitive de continuité.

En effet, il y a deux propriétés comprises sous le terme de continuité : d'une part la propriété de divisibilité et d'autre part la propriété de passage d'une entité à une autre de même nature sans saut. La division à l'infini, ou plus précisément *ad libidum* divise l'ensemble en parties homogènes à lui-même, ce qui réfère à *l'infini en puissance* aristotélien. Le recollement de deux parties de même nature, en prolongement l'une de l'autre en un tout homogène par une couture assurant le passage sans "trou" est aussi la limite aristotélienne. Limite qui n'est pas de même nature que les entités qu'elle délimite tout en les reliant, assurant le passage de l'une à l'autre sans aspérité. C'est pourquoi nous pensons que si la droite géométrique est la figure métaphorique pour le temps, l'ensemble des réels n'est pas le modèle idoine pour représenter le temps.

En effet, si le temps n'est pas une donnée directement accessible, ses traces sensibles sont les relations qu'il instaure entre les entités soumises à son joug : relations de précédence et de simultanéité. L'homme s'est également doté de moyens permettant à la fois de situer ces objets de façon universelle et de mesurer leur durée, cela par le biais des chronologies, des calendriers et des horloges. Ces outils, conceptuellement, sont des structures d'ordre linéaires et discrètes, en correspondance naturelle avec l'ensemble des entiers naturels ou entiers, selon qu'on leur attribue une origine, premier élément ou non. Les systèmes horo-calendaires sont des arborescences de structures d'ordre linéaires et discrètes qui permettent de passer d'une échelle à une autre, et en particulier, potentiellement, de diviser l'unité de mesure en unités plus petites (an/jour) *ad libidum*. Ce procédé rappelle les constructions fractales. Ce qui nous ramène à la première propriété de continuité. D'autre part, le passage d'une occurrence d'une unité à la suivante (par exemple "du jour au lendemain") constitue bien une couture qui à la fois permet de délimiter les deux occurrences tout en assurant la continuité de ce passage.

Enfin, sur la ligne métaphorique du temps, considéré comme un *continuum*, seuls émergent les événements "remarquables" qui, s'ils font rupture dans la vie, en assurent également la cohérence¹.

Ces remarques justifient le choix d'une approche d'une autre nature. C'est pourquoi S. R. Schwer a proposé celle des S-langages, fondée sur la théorie des langages formels. En effet, l'idée n'est pas de représenter le temps mais uniquement les traces temporelles des événements qui font sens, dans un contexte déterminé. Dans la section suivante, nous rappelons quelques traits des S-langages.

¹La cohérence est dans le fait que la ligne de vie peut devenir une ligne dite brisée, c'est-à-dire droite par morceaux, qui n'en demeure pas moins continue, grâce aux points de brisure.

Puis nous nous intéresserons à l'organisation sous forme de treillis de ces S-langages, et plus précisément des S-langages formés de S-permutations. Par un codage spécifique, dû à Jérôme Cardot, ces S-permutations se plongent dans un treillis de permutation, selon une répartition fractale. Cela nous amène à poser la question d'une structure fractale possible du temps.

2 Retour sur les S-langages

La théorie des S-langages, développée par S. R. Schwer² (Schwer 2001) est une généralisation de celle des langages formels, c'est-à-dire des suites finies de lettres, dont la lecture se fait en épelant chacune des lettres dans leur ordre d'écriture (ici de gauche à droite) dans la séquence. Ainsi, tout mot peut être considéré comme une trace du temps physique, ponctués de *moments remarquables* : le premier moment correspond à la signification donnée par l'occurrence de la première lettre de la séquence, le second moment correspond à la signification donnée par l'occurrence de la deuxième lettre de la séquence, ... Deux moments remarquables consécutifs déterminent une période de temps ou durée. Par exemple, le mot *abab* de la Figure 1 décrit le fait que l'intervalle noir, dénoté *aa*, contient l'intervalle blanc, dénoté *bb*. Une autre interprétation est la suivante : *aa* et *bb* dénotent deux entités temporelles duratives A et B ; le mot *abab* dit alors que A a commencé en premier et s'est terminé en premier, mais après le début de B. Cette représentation est fondée sur le concept d'*identité d'entités*, que l'on trouve par exemple en conception orientée objet et s'oppose à la notion leibnizienne des indiscernables³. Cette entité *P* est représentée par autant d'occurrences de son identité qu'il lui est survenu quelque chose. Si ce quelque chose dure, il est considéré comme intervenant deux fois : la première fois en son début et la seconde en sa fin. Ainsi sur la Figure 2, l'entité A est décrite par trois événements successifs duratifs, l'entité B par deux événements duratifs séparés par un événement ponctuel et l'entité C par un seul événement duratif. Les durées sont toujours qualitatives, mais relatives les unes aux autres. Mais la relation ternaire décrivant l'ordre chronologique global n'est pas exprimable sous forme de mots, car certaines extrémités sont simultanées. En effet, décrire les positions temporelles relatives nécessite de pouvoir représenter à la fois la simultanéité et la précédence. C'est donc non sur l'alphabet des identités mais sur l'alphabet constitué des parties de l'ensemble des identités - appelé S-alphabet, dont les éléments sont les S-lettres - qu'il s'agit de travailler, l'ordre des occurrences des S-lettres dans le S-mot donnant *de facto* la précédence.

²Qui a présenté à Rochebrune un article sur modalité temporelle et S-mots (Schwer 2004)

³Qui dit que si deux entités ne sont pas discernables, alors ce sont les mêmes, ce qui dans notre théorie se traduit par : si deux entités ont mêmes extensions temporelles, elles sont identiques. En revanche, la théorie des S-langages emprunte à Leibniz sa représentation des moments comme classes d'équivalence de toutes les occurrences d'événements se produisant simultanément, cela même grâce à la notion d'identité des entités.

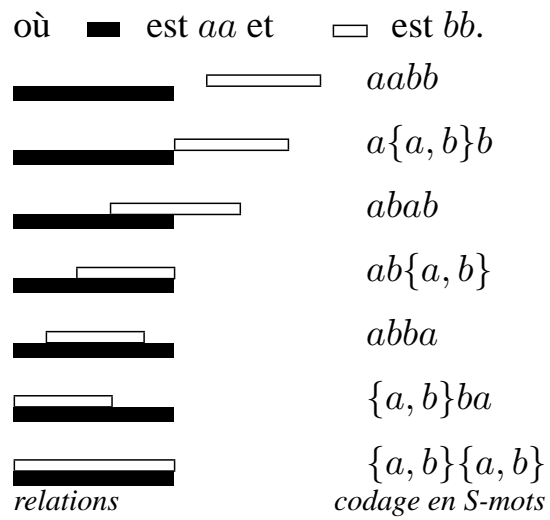


FIG. 1 – Relations entre deux intervalles

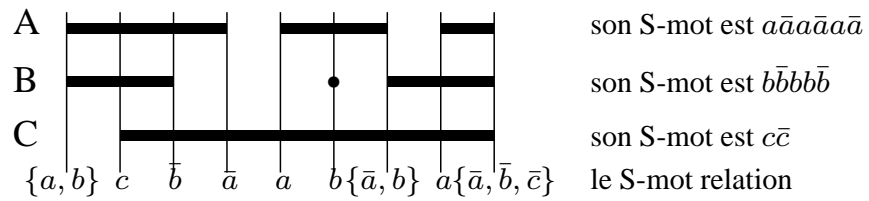
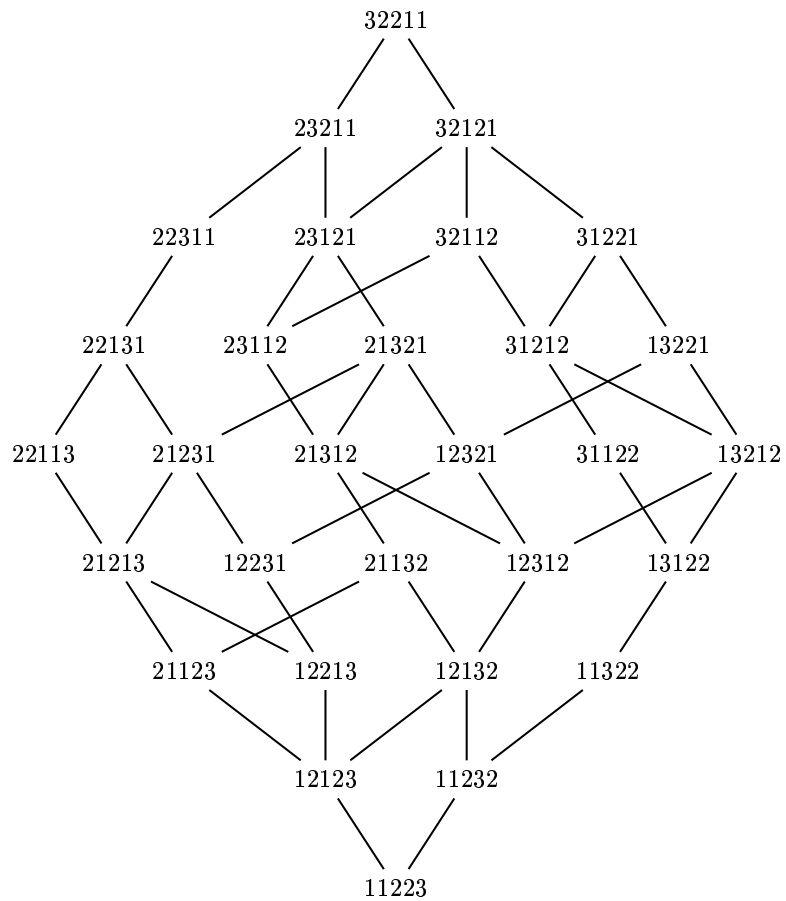


FIG. 2 – Une relation ternaire entre trois entités A, B and C.



3 Treillis des S-permutations

Les éléments développés ici, qui font partie d'un travail mené par les auteurs en collaboration avec Claude Barbut, ont pour but d'étudier certaines propriétés des S -mots ne possédant qu'une seule occurrence de chaque lettre, et plus précisément de comparer ce treillis des S -mots, appelé S-permutations avec le treillis des permutations ou les treillis d'arrangements avec répétition, que la notion de S-mots généralise (Schwer 2001) du même ordre – c'est-à-dire utilisant les mêmes lettres. En effet, il ressort de (Autebert and Schwer 2003) que ces treillis ne sont pas distributifs, dès qu'il y a plus de deux lettres en jeu. Or les treillis de permutations et d'arrangements avec répétition ne sont pas distributifs mais seulement semi-distributifs. Les treillis de S-permutations sont-ils aussi semi-distributifs ? Nous n'avons à ce jour, pas de réponse. Différentes méthodes ont été développées, parmi lesquelles différents codages, permettant le plongement des treillis des S-permutations dans d'autres treillis. Nous présentons ici le codage de Cardot, qui, s'il n'a pas (encore) permis de répondre à la question, donne un nouveau regard sur les S-mots correspondants.

Nous introduirons d'abord les structures algébriques mises en œuvre, et montrerons comment le plongement envisagé conduit à un codage numérique et à une représentation des S -mots comme des chemins sur le triangle de Sierpiński

3.1 Caractérisation des treillis

Illustrons la notion de treillis, dans le cas des ensembles finis.⁴

Une *relation d'ordre strict*, souvent notée $<$, sur un ensemble fini T peut être représentée par un graphe sans cycle sur T (pour traduire l'interdiction d'avoir à la fois $a < b$ et $b < a$).

- i. $a < b$ si, et seulement si, il y a un chemin de a à b dans le graphe ;
- ii. si on ne représente pas les « *raccourcis* » du graphe, le graphe est appelé *diagramme de Hasse* de la relation d'ordre ;
- iii. à la relation d'ordre strict, on associe une relation d'ordre large, \leq :

$$a \leq b \text{ si, et seulement si, } a < b \text{ ou } a = b$$

ce qui revient à considérer aussi les chemins de longueur nulle du graphe.

Si de plus les conditions suivantes sont vérifiées, on dira que T est un *treillis* :

- iv. il y a un élément de T plus petit que tous les autres ;
- v. il y a un élément de T plus grand que tous les autres ;

⁴Pour une définition plus formelle des treillis, voir par exemple (Szász 1971; Davey and Priestley 2002).

- vi. si a et b sont dans T , il y a un élément de T , que l'on note $a \vee b$, qui est le plus petit des éléments de T plus grands que a et que b ; autrement dit, si un chemin du diagramme, commençant en a , croise un chemin commençant en b , on peut aussi atteindre le premier croisement des deux chemins, depuis a ou depuis b , en passant par $a \vee b$;
- vii. si a et b sont dans T , il y a un élément de T , que l'on note $a \wedge b$, qui est le plus grand des éléments de T plus petits que a et que b ; autrement dit, si un chemin du diagramme, finissant en a , croise un chemin finissant en b , on peut aller du dernier croisement à a ou b en passant par $a \wedge b$.

La figure 3 présente deux exemples classiques de treillis : le treillis des parties d'un ensemble, ordonnées par la relation d'inclusion, pour lequel \wedge et \vee correspondent à l'intersection et la réunion, et le treillis des diviseurs d'un nombre, ordonnés par la divisibilité, où $a \vee b$ est le plus petit multiple commun de a et b , et $a \wedge b$ leur plus grand diviseur commun.

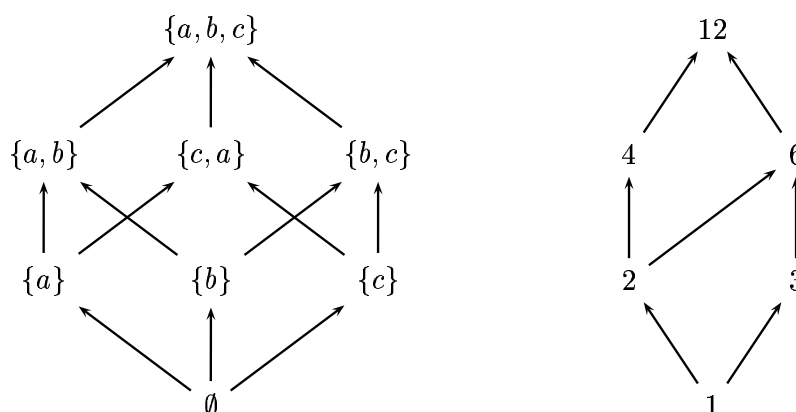


FIG. 3 – Deux exemples de treillis

3.2 Codage par un mot de longueur double

Notre objectif est de représenter les S -permutations sur un alphabet X par des mots sur ce même alphabet ou sur un alphabet qui lui est proche, de façon à profiter de résultat connus sur les permutations. Le codage se fait sur des S -mots quelconques, comme nous l'exposons ici.

Chaque S -lettre d'un S -mot σ est lue deux fois : une fois dans l'ordre alphabétique, et une fois dans l'ordre alphabétique inverse. Cette deuxième lecture correspond, pour aider la lecture, à un surlignage des lettres.

Par exemple, la S -permutation $\sigma = a\{a, b, c\}\{a, c\}$ est codée $\sigma^\diamond = (a\bar{a}bcd\bar{d}\bar{c}\bar{b}ac\bar{c}\bar{a})$

Si on ne lit que les lettres simples (non barrées), on obtient $\sigma^- (abcdac)$, la plus grande permutation inférieure à σ , et on obtient $\sigma^+ (\bar{a}\bar{d}\bar{c}\bar{b}\bar{c}\bar{a})$, la plus petite permutation supérieure à σ , en ne lisant que les lettres barrées.

L'intervalle $[\sigma^-, \sigma^+]$ est alors le meilleur encadrement possible de σ par des permutations.

Parmi les ordres possibles sur les lettres de σ^\diamond , ceux qui correspondent aux S -mots vérifient les conditions suivantes :

- i. on n'échange pas des occurrences (simples ou barrées) de la même lettre ; autrement dit on ne sort pas de l'intervalle $[(a\bar{a}b\bar{b}c\bar{c}d\bar{d}), (d\bar{d}c\bar{c}b\bar{b}a\bar{a})]$
- ii. entre ℓ et $\bar{\ell}$, soit il n'y a rien, soit il y a α suivie de $\bar{\alpha}$ (avec $\alpha > \ell$). Les lettres peuvent donc être embrassées ($ab\bar{b}\bar{a}$) mais pas alternées ($ab\bar{a}\bar{b}$) ; cette condition correspond aux langages de Dyck (langage de parenthèses, cf. (Autebert 1994)).
- iii. la contrainte selon laquelle les parenthèses intérieures sont d'un niveau plus élevé que les parenthèses extérieures correspond à imposer un ordre aux règles de simplification :
 - d'abord $z\bar{z} \longrightarrow \varepsilon$
 - en dernier $a\bar{a} \longrightarrow \varepsilon$
- iv. on peut aussi exprimer cette contrainte de la façon suivante : une chaîne de lettres non barrées est nécessairement croissante, une chaîne de lettres barrées est nécessairement décroissante.
- v. en outre, la lecture des lettres barrées doit dépiler toutes les lettres non barrées lues ; ainsi le mot $ab\bar{b}c\bar{c}\bar{a}$ ne code-t-il aucun S -mot ; c'est une contrainte supplémentaire par rapport aux langages de Dyck.

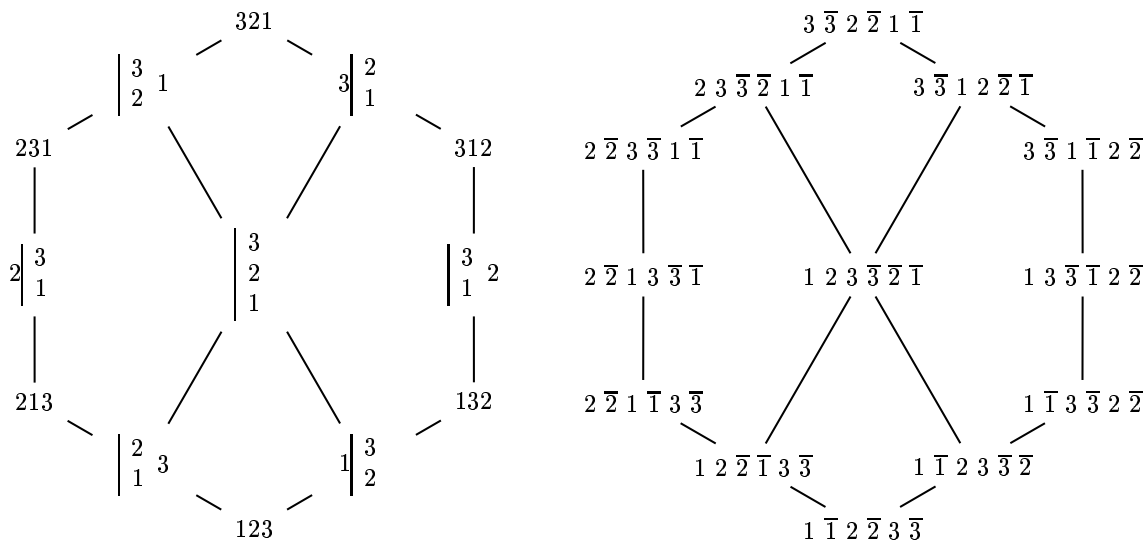


FIG. 4 – Treillis des S -mots sur 3 lettres, et leur linéarisation

3.3 Représentation graphique

La représentation pour les mots codés est choisie de telle sorte que :

- i. chaque lettre est représentée dans le plan par un vecteur différent ; comme le code linéarisé d'un S -mot comprend toutes les lettres de l'alphabet $X \cup \bar{X}$, tous les S -mots sont alors représentés par des chemins de même origine et même extrémité ;
- ii. une symétrie permet de passer d'une lettre barrée à la lettre non barrée correspondante ;
- iii. des lettres différentes sont représentées par des vecteurs différents.

Ces contraintes sont vérifiées en associant à chacune des lettres un vecteur de longueur une puissance de 2 (ce qui assure que des combinaisons formées de lettres différentes se représentent par des vecteurs différents), une direction étant consacrée aux lettres barrées, une autre aux lettres non barrées.

La figure 5 représente ainsi les chemins associés aux mots de 8 lettres associés aux S -mots $abcd$, $a \begin{array}{c} c \\ b \end{array} d$ et $d \begin{array}{c} c \\ b \\ a \end{array}$



FIG. 5 – Représentation linéarisée : $(a\bar{a}b\bar{b}c\bar{c}d\bar{d})$, $(a\bar{a}bc\bar{c}b\bar{d}d\bar{d})$, $(d\bar{d}abc\bar{c}b\bar{a})$.

Il est alors utile de caractériser l'ensemble des chemins correspondant aux mots linéarisant les S -mots sur un alphabet. On obtient ainsi un automate reconnaissant les S -mots linéarisés. C'est ce que représente la figure 6, où les points marquent les bifurcations possibles entre un segment horizontal (qui représente des lettres non barrées) et un segment vertical (qui représente les lettres barrées).

Rappelons (contrainte v . du codage) qu'une série de lettres barrées doit comprendre toutes les lettres non barrées de la série qui précède, et sera donc représentée par un segment de même longueur : après un segment horizontal, le chemin représentant un S -mot remonte verticalement jusqu'à la diagonale, et chaque S -lettre est représentée par un triangle isocèle rectangle.

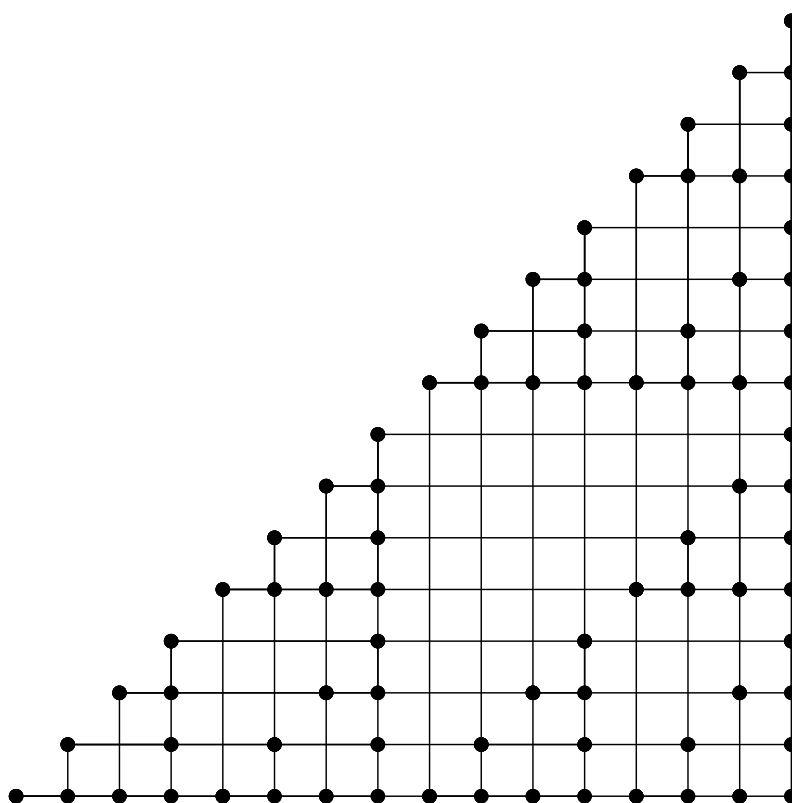


FIG. 6 – Automate pour les S -mots linéarisés d'ordre 4

La figure 6 présente les traces d'un aspect fractal de la construction réalisée : les points de bifurcation sont en effet disposés suivant une figure fractale bien connue, le triangle de Sierpiński.

4 Conclusion

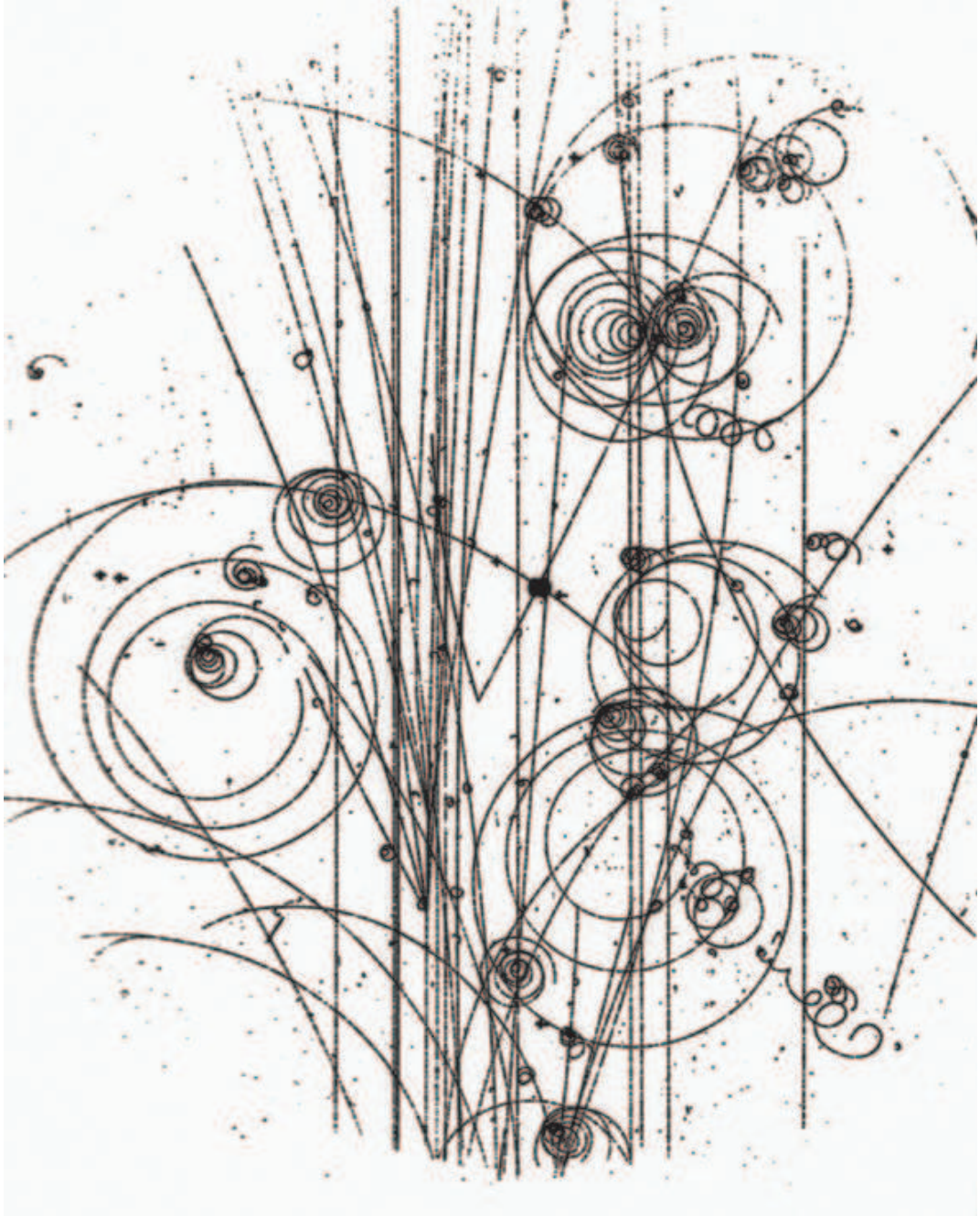
Ainsi, nous avons suggéré que la représentation du temps sous la forme de S -mot conduisait à une structure temporelle dense et satisfaisant des critères de

non-standardisation des ensembles de nombres candidats \mathcal{Q} ou \mathbb{R} . Certaines réflexions du mathématicien et philosophe de la nature Whitehead (1861–1947) vont déjà dans cette direction, bien qu’il affirme refuser, comme Weierstrass, l’existence des infinitésimaux. Mais l’analyse non standard introduite par Robinson était encore à venir (1961) !

Plus déroutant encore est la transformation de Cardot, qui plonge le treillis des S-mots d’un ordre donné dans une structure fractale. Quel sens donner à cela ?

Références

- Autebert, J.-M. (1994). *Théorie des langages et des automates*. Paris : Masson.
- Autebert, J.-M. and S. R. Schwer (2003). On generalized Delannoy paths. *SIAM Journal on Discrete Mathematics* 16(2), 208–223.
- Davey, B. A. and H. A. Priestley (2002). *Lattices and Order*. Cambridge University Press.
- Schwer, S. R. (2001). S -arrangements avec répétitions. *Comptes rendus de l’Académie des Sciences de Paris* (334), 261–266.
- Schwer, S. R. (2004). Relations temporelles qualitatives et langages formels. In D. Badarioti (Ed.), *Proceedings of the Spatial and Temporal Reasoning Workshop the 18th IJCAI*, pp. 329–340. Paris : ENST.
- Szász, G. (1971). *Théorie des treillis*. Paris : Dunod.



Acquisition de la terminologie et définition des tâches à effectuer, deux principes indissociables

Mathieu Roche

Équipe TAL, LIRMM, UMR 5506, Université Montpellier 2
mathieu.roche@lirmm.fr

Résumé

L'acquisition de la terminologie à partir de textes spécialisés est une étape essentielle. Elle est souvent indispensable pour extraire des connaissances permettant une meilleure compréhension de ces textes. Ainsi, les termes extraits peuvent représenter des traces utiles pour l'expert du domaine étudié. Afin de définir la pertinence de la terminologie extraite, cet article montre la nécessité de rigoureusement prendre en compte les tâches et les sous-tâches à effectuer avec les termes. En effet, d'une tâche à l'autre, les terminologies obtenues à partir d'un même corpus peuvent se révéler extrêmement différentes.

Mots clés

Fouille de textes, TAL, Terminologie

1. Introduction

Dans cet article, nous nous intéressons à l'étude des groupes de mots qui sont appelés des **collocations**. Nous nous intéressons plus particulièrement à l'extraction des collocations à partir de corpus (ensembles de textes homogènes). Comme nous le montrerons dans cet article, ces collocations peuvent représenter des traces qui seront utiles pour l'extraction de connaissances dans les textes. Ces connaissances extraites sont alors susceptibles d'apporter une meilleure compréhension des textes étudiés. Les collocations pertinentes sont des traces linguistiques de concepts. Cette définition a été introduite par Yves Kodratoff dans (Kodratoff, 2004). Dans cet article, nous montrerons comment les collocations peuvent ou non représenter de véritables traces linguistiques de concepts. Ainsi, nous montrerons que la pertinence d'une collocation devra être définie selon une tâche déterminée par l'expert.

Plus formellement, (Clas, 1994) donne deux propriétés définissant une collocation. Dans un premier temps, une collocation est définie comme un groupe de mots ayant un sens global qui est déductible des unités (mots) composant le groupe. Par exemple, "jour faste" est considéré comme une collocation car le sens global de ce groupe de mots peut être déduit des deux mots "jour" et

"faste". En nous appuyant sur cette définition, l'expression "tirer son chapeau" n'est pas une collocation car son sens ne peut pas être déduit de chacun des mots. De telles formes sont appelées des **combinaisons figées**¹. Une deuxième propriété est ajoutée par (Clas, 1994) pour définir une collocation. Le sens des mots qui composent la collocation doit être limité. Par exemple "acheter un chapeau" n'est pas une collocation car le sens de "acheter" et de "chapeau" n'est pas limité. En effet, de multiples objets, voire des personnes, peuvent être achetés. De tels groupes de mots sont appelés des **combinaisons libres**.

Dans nos travaux, nous considérons les combinaisons figées et libres comme des collocations contrairement à la définition de (Clas, 1994). Prenons l'exemple "first lady" issu d'un ensemble de textes étudiés au cours du challenge TREC'2004 (Text REtrieval Conference)². Une telle collocation devrait être considérée comme une combinaison figée désignant la femme du chef de l'exécutif. Dans le cas des textes décrivant la campagne sénatoriale d'Hillary Clinton en 2000, cette combinaison est utile et pertinente pour caractériser Hillary Clinton. Cet aspect pragmatique, nous a motivé à considérer les combinaisons figées mais également libres comme des collocations.

Dans cette étude qui a pour but de définir de quelle manière définir la pertinence d'une collocation, nous nous appuyons sur deux corpus écrits en français :

- Corpus composé de Curriculum Vitæ fournis par la société VedioBis (120000 mots après divers prétraitements décrits dans (Roche, 2004)). Une des particularités de ce corpus tient au fait qu'il est composé de phrases très courtes avec de nombreuses énumérations.
- Corpus issu du domaine des Ressources Humaines (société PerformanSe) correspondant à des commentaires de tests de psychologie de 378 individus (600000 mots). Les textes sont écrits par un seul auteur qui emploie un vocabulaire spécifique avec l'utilisation de tournures souvent littéraires.

Lorsque l'on décrit les systèmes d'extraction de la terminologie, il est coutume de préciser que l'on souhaite extraire les collocations pertinentes. Cependant, une des difficultés majeures est de déterminer ce que nous entendons par "collocation pertinente". Ainsi, la section 2 définit la pertinence d'une collocation. Dans la suite de cet article, les collocations pertinentes seront appelées des "**termes**". Les sections 3 et 4 présenteront respectivement le type de collocations extraites pour les tâches de normalisation des textes et de construction

¹Les combinaisons figées non ambiguës peuvent alors être lexicalisées pour former des expressions lexicalisées.

²<http://trec.nist.gov/>. L'exemple donné est issu de la tâche Novelty du Challenge TREC'04. Cette tâche consiste à rechercher des phrases pertinentes et nouvelles à partir de textes journalistiques.

d'une classification conceptuelle. Enfin, nous présenterons quelques perspectives à ce travail.

2. Qu'est ce qu'une collocation pertinente ?

Pour définir la pertinence d'une collocation ou plus généralement d'une trace présente dans les textes, il est nécessaire de déterminer la tâche que l'on souhaite effectuer avec celle-ci. Nous montrerons par exemple dans les sections suivantes qu'une collocation peut être pertinente pour une tâche de normalisation des textes mais inutile pour construire une classification conceptuelle.

Il est également important de préciser que nous pouvons ajouter un niveau supplémentaire relatif aux sous-tâches pour une tâche principale. À titre d'exemple, une classification conceptuelle peut être construite pour différents objectifs :

- Découvrir des règles d'association entre concepts présents dans les textes (Kodratoff *et al.*, 2003, Azé & Roche, 2003) ou entre les instances de concepts (Janetzko *et al.*, 2004).
- Extraire des informations en utilisant des patrons d'extraction grâce aux concepts (Freitag, 1998), etc.

Ces deux objectifs propres à la tâche principale d'acquisition d'une classification conceptuelle caractérisent plusieurs sous-tâches. La pertinence d'une collocation peut en être affectée. Par exemple, une collocation peut être adaptée pour une classification conceptuelle pour rechercher des informations dans les textes à base de patrons d'extraction. Pourtant, cette même collocation peut être jugée comme non pertinente pour découvrir des règles d'association entre concepts. Nous reviendrons en détail sur cet exemple dans la section 4. Ce fait est donc à prendre en considération dans la définition des collocations pertinentes. Dorénavant, nous considérons que si des sous-tâches propres à une tâche principale sont identifiées alors l'expert devra définir la pertinence selon les sous-tâches à réaliser.

Plus généralement, nous montrons dans cet article qu'une suite de mots présente dans les textes qui respecte une structure grammaticale déterminée est une **trace pertinente** si elle est utile pour une tâche ou une sous-tâche à réaliser par l'expert. Dans les sections suivantes, nous discutons de l'utilisation des mesures statistiques pour deux tâches principales : la normalisation des textes et la classification conceptuelle. Ces mesures statistiques permettent de classer les collocations extraites. Nous allons expliquer de quelle manière ces mesures peuvent aider à privilégier les collocations propres à une tâche principale. En effet, nous estimons pouvoir identifier les mesures statistiques les plus adaptées pour une tâche principale particulière mais il nous semble difficile de déterminer les mesures classiques à utiliser pour les sous-tâches souvent plus fines.

3. Acquisition des termes pour la normalisation des textes

Dans la chaîne globale de fouille de textes (Kodratoff *et al.*, 2003, Mathiak & Eckstein, 2004), après l'acquisition du corpus, sa normalisation est la première tâche importante à effectuer. Ces normalisations consistent à éliminer le bruit présent dans les textes, à uniformiser le vocabulaire, etc. Nous donnons ci-dessous, quelques exemples de sous-tâches utiles pour la phase de normalisation :

- détection du bruit dans les textes : problèmes liés au nettoyage, aux fautes d'orthographe, etc.
- détection de collocations utiles pour la phase de constitution ou de mise à jour de lexiques de noms propres : noms de lieux, noms de sociétés, noms d'établissements, noms d'organisations (politiques, associatives, religieuses, etc.), couples "prénom" associé à un "nom de famille", etc.

Sur le corpus de CVs, les deux sous-tâches de normalisation qui sont le nettoyage (présence de bruit et de fautes d'orthographe) et la reconnaissance de collocations pouvant constituer un lexique de noms propres ont été considérées. Ces lexiques peuvent être utilisés de différentes façons pour la tâche de normalisation. Les collocations d'un tel lexique de noms propres peuvent être utiles pour une phase de pré-terminologie. La pré-terminologie consiste à extraire des collocations particulières qui peuvent être considérées comme des mots à part entière. Ces collocations étant repérées, un trait d'union (ou un "blanc souligné" plus spécifiquement utilisé pour ce type d'entités nommées) peut être placé entre chacun des mots composant ces collocations. Ceci permet alors de les considérer comme des mots à part entière lors des étapes suivant la phase de normalisation (étiquetage grammatical, analyse syntaxique, etc.). Bien entendu, ce type de collocation est souvent spécifique aux domaines étudiés.

Cependant, d'un corpus à l'autre et d'un lexique de noms propres à l'autre, leur utilisation peut différer. Notons ci-dessous des exemples liés à une utilisation différente des lexiques :

- *Utilisation différente d'un même lexique selon les corpus.*
Par exemple, sur un corpus d'articles de journaux (Soboroff & Harman, 2003), les collocations issues d'un lexique de prénoms associés à un nom de famille doivent être considérées comme de la pré-terminologie. Ce type d'information est en effet essentiel dans un corpus journalistique. Sur un corpus de CVs, de telles informations sont inutiles et inopportunes dans un souci d'anonymat de certaines informations textuelles. Il est donc nécessaire de les supprimer. Cet exemple montre donc qu'un même lexique peut être utilisé de manière différente selon les corpus.

- *Utilisation différente des lexiques sur un même corpus.*

Par exemple, l'utilisation de lexiques des noms de sociétés peut être utile dans une phase de pré-terminologie sur un corpus de CVs. Au contraire, sur ce même corpus de CVs, un lexique de prénoms associés à un nom de famille est utilisé afin de supprimer ces informations. Cet exemple montre donc que, sur un même corpus, les lexiques peuvent être utilisés de manière différente.

Plusieurs expérimentations sur le corpus de CVs non normalisé ont été effectuées afin de comparer deux mesures statistiques. Pour cela, nous nous sommes appuyés sur la relation Nom-Nom du corpus de CVs qui présente le nombre le plus important de collocations (voir tableau 1). Pour extraire ces collocations, une étape préalable est effectuée qui consiste à étiqueter grammaticalement les mots des textes avec l'étiqueteur de Brill (Brill, 1994). Ceci permet alors d'extraire aisément les collocations respectant des patrons précis de type Nom-Nom, Nom-Préposition-Nom, Nom-Adjectif, Adjectif-Nom, etc. Notons que l'approche d'extraction de la terminologie que nous utilisons suit un processus itératif détaillé dans (Roche, 2004) afin de construire des termes complexes (composés de nombreux mots). Par exemple, si à la première itération, le terme "système complexe" de type Nom-Adjectif est extrait, à la deuxième itération, nous pouvons extraire le terme "système complexe naturel" (exemple issu du titre des thèmes des Journées de Rochebrune).

| Collocations | Nombre total | Collocations | Nombre total |
|--------------|--------------|--------------|--------------|
| Nom-Prép-Nom | 5340 | Nom-Adjectif | 2904 |
| Nom-Nom | 9394 | Adjectif-Nom | 878 |

Table 1: Nombre total de collocations sur le corpus de CVs non normalisé.

Nous montrons, dans la suite, de quelle manière les mesures statistiques peuvent être utilisées pour identifier ce type de collocations pouvant représenter des traces utiles pour l'étape de normalisation. Afin de classer les collocations extraites, nous nous appuyons sur deux mesures classiques du domaine : l'Information Mutuelle (Church & Hanks, 1990) et le Rapport de Vraisemblance (Dunning, 1993). Ces mesures sont décrites dans (Roche, 2004). Comme cela est montré dans (Daille, 1994, Roche, 2004), le Rapport de Vraisemblance privilégie les collocations les plus fréquentes contrairement à l'Information Mutuelle qui place en tête les collocations les plus rares.

Dans le tableau 2, nous présentons l'analyse manuelle des 100 premières collocations Nom-Nom en utilisant ces deux mesures statistiques. Dans ce tableau, les expérimentations réalisées avec l'ensemble des collocations de type Nom-Nom extraites sans appliquer d'élagage sont présentées. Nous avons associé

manuellement chacune des 100 premières collocations binaires aux différentes sous-tâches propres à la normalisation. Ce tableau montre que certaines mesures statistiques sont plus ou moins adaptées pour la tâche de normalisation. Ainsi, près des deux tiers (65%) des premières collocations extraites avec l'Information Mutuelle sont utiles pour la tâche globale de normalisation (voir tableau 2). A contrario, avec le Rapport de Vraisemblance, moins d'un tiers (28%) sont utiles pour la phase de normalisation. Précisons également que l'Information Mutuelle est particulièrement efficace pour constituer ou enrichir des lexiques composés des couples "prénom nom" contrairement au Rapport de Vraisemblance. Les travaux de (Thanopoulos *et al.*, 2002) ont montré que, d'une manière générale, l'Information Mutuelle permettait d'extraire des entités nommées.

| Sous-tâches et leur description | | IM | RV | |
|---|--|------|--------|-----------------|
| Nettoyage | { Bruit Fautes d'orthographe | 5% | } 5 % | |
| | | 0% | | 1% } 1 % |
| Constitution ou enrichissement de lexiques de noms propres | { noms de lieux noms de sociétés, d'établissements, d'organisations association pré-noms/noms | 5 % | } 60 % | |
| | | 25 % | | 4 % } 27 % |
| | | 30 % | | 20 % } 3 % } |

Table 2: Pourcentage des 100 premières collocations Nom-Nom (première itération) vérifiant des sous-tâches de la phase de normalisation. Ces collocations sont classées avec deux mesures : l'Information Mutuelle (IM) et le Rapport de Vraisemblance (RV). Expérimentations à partir du corpus de CVs non normalisé sans appliquer d'élagage pour la relation Nom-Nom.

Dans cette section, les expérimentations présentées sont relatives au corpus de CVs. Notons que sur d'autres corpus, l'Information Mutuelle extrait également des collocations pertinentes pour la constitution de lexiques spécifiques qui sont utiles pour l'étape de normalisation. Cependant, ces lexiques peuvent être de natures différentes. À titre d'exemple, un lexique constitué de "termes littéraires" (expressions linguistiques) peut être construit à partir du corpus des Ressources Humaines. En effet, l'auteur de ce corpus utilise un vocabulaire caractéristique composé de nombreux termes littéraires qui sont extrêmement bien classés en utilisant l'Information Mutuelle : par exemple, "*statu quo*", "*voeux pieux*", "*carte blanche*", "*bâton rompus*", "*coudées franches*", etc. De telles collocations sont des combinaisons figées.

Après avoir mis en valeur les avantages de l'Information Mutuelle pour la tâche principale de normalisation, nous allons comparer le Rapport de Vraisemblance et l'Information Mutuelle pour une autre tâche principale : l'acquisition des termes pour la classification conceptuelle du domaine.

4. Acquisition des termes pour la classification conceptuelle

Pour construire une classification conceptuelle, les collocations évoquant des concepts du domaine sont extraites puis regroupées. Le tableau 3 présente des exemples de collocations associées à des concepts à partir des deux corpus en français étudiés. Les concepts sont définis par les experts selon les différents objectifs qu'ils souhaitent effectuer avec la classification conceptuelle. Nous détaillerons des exemples d'objectifs que les experts peuvent se fixer dans la suite de cet article.

| CURRICULUM VITÆ | | RESSOURCES HUMAINES | |
|------------------------|----------------------|------------------------|---------------|
| Collocations | Concepts | Collocations | Concepts |
| aide comptable | Activité Gestion | besoin d'information | Communication |
| gestion administrative | Activité Gestion | capacité d'écoute | Communication |
| chef de service | Activité Encadrement | contexte professionnel | Environnement |
| direction générale | Activité Encadrement | lieu de travail | Environnement |
| employé libre service | Activité Commerce | sentiment de malaise | Stress |
| assistant marketing | Activité Commerce | tension permanente | Stress |

Table 3: Extrait des classifications conceptuelles.

Afin de valider les collocations extraites, plusieurs catégories de pertinence (ou de non pertinence) ont été identifiées :

- **1^{ère} catégorie** : La collocation est pertinente pour la classification conceptuelle (exemple du corpus de CVs : "*baccalauréat littéraire*")
- **2^{ème} catégorie** : La collocation est pertinente mais très spécifique et pas nécessairement pertinente pour le domaine (exemple du corpus de CVs : "*écosystème marin*")
- **3^{ème} catégorie** : La collocation est pertinente mais très générale et pas nécessairement pertinente (exemple du corpus de CVs : "*situation actuelle*")
- **4^{ème} catégorie** : La collocation est non pertinente (exemple du corpus de CVs : "*jour quotidienne*")
- **5^{ème} catégorie** : L'expert ne peut pas juger si la collocation est pertinente (exemple du corpus de CVs : "*master franchisé*").

Comme précisé précédemment, les collocations représentant des traces linguistiques de concepts doivent être pertinentes selon une sous-tâche à réaliser. Par exemple, les collocations qui sont des instances de concepts peuvent être utilisées pour découvrir des règles d'association entre concepts présents dans les textes. Ceci permet alors de déterminer la force des associations qui peut exister entre les concepts. Les concepts peuvent également être utilisés pour construire des patrons d'extraction utiles pour la recherche d'informations.

Dans la première des sous-tâche (découvrir des règles d'association entre concepts), les concepts utilisés doivent être précis afin de déterminer des associations éventuelles. Ce travail a des similarités avec les approches de (Srikant & Agrawal, 1997) qui consistent à utiliser une taxonomie pour généraliser des règles d'association extraites. Dans nos travaux présentés dans (Kodratoff *et al.*, 2003, Azé & Roche, 2003) et dans la thèse de Jérôme Azé (Azé, 2003), les règles d'associations découvertes sont de la forme $concept_1 \dots concept_{n-1} \rightarrow concept_n$ où n est le nombre de concepts impliqués dans les règles d'association extraites. Le détail de l'algorithme est présenté dans (Azé, 2003). À titre d'exemple, nous donnons une règle d'association extraite à partir du corpus des Ressources Humaines : "Stress" \rightarrow "Environnement". Cette règle signifie que le stress s'exerce par l'intermédiaire de l'environnement. Cet exemple montre que les règles d'association permettent une meilleure compréhension des corpus étudiés. L'extraction des règles d'association s'effectue avec des concepts très précis qui intéressent l'expert. Ainsi, en reprenant les différentes catégories évoquées au début de cette section, les collocations pertinentes afin de découvrir des règles d'association entre concepts sont les collocations de la catégorie 1. Les collocations issues des catégories 2, 3 et 4 sont jugées comme non pertinentes. Enfin, les collocations de la catégorie 5 ne sont pas prises en considération car l'expert n'a pas été apte à les valider. Cette dernière catégorie correspond en fait à des collocations qui sont ambiguës ou qui n'ont pas pu être évaluées par méconnaissance partielle du domaine.

Définissons la seconde sous-tâche évoquée qui peut être associée à la même tâche principale (classification conceptuelle). Cette sous-tâche consiste à déterminer des concepts dans le but de construire des patrons d'extraction. Les patrons d'extraction prenant en compte des noms de personnes peuvent être utiles dans la recherche d'informations. Par exemple, le concept de *Nom de Personnes* permet d'appliquer un patron de type "Concept_Nom_de_Personnes suivi du verbe "s'intéresser à" afin d'extraire dans les textes tous les centres d'intérêt des personnes. Ainsi, l'utilisation des collocations spécifiques ou générales (catégories 2 et 3) permet de construire des patrons d'extraction couvrant davantage les textes et être donc plus efficace pour la recherche d'informations. Pour cette sous-tâche, les collocations pertinentes sont donc les collocations des catégories 1, 2 et 3 et les collocations non pertinentes sont celles issues de la catégorie 4. Ainsi, plus généralement, les collocations de la catégorie 4 ne peuvent être considérées comme de véritables traces.

Dans la suite, nous allons examiner les 100 premières collocations de type Nom-Nom du corpus de CVs classées avec l'Information Mutuelle et le Rapport de Vraisemblance. Ces collocations, analysées pour la tâche principale de classification conceptuelle, correspondent aux mêmes jeux de données que l'étude sur la normalisation (section 3). Les résultats sont présentés dans le tableau 4. Dans un premier temps, nous remarquons que les collocations pertinentes

| Information Mutuelle (IM) | | | |
|-------------------------------|--------------------|------------------------------------|-------------------------------------|
| Catégories | Nb de collocations | Découvrir des règles d'association | Construire des patrons d'extraction |
| 1. pertinent | 6 | <i>positif</i> : 6 | <i>positif</i> : 83 |
| 2. spécifique | 77 | <i>négatif</i> : 83 | |
| 3. général | 0 | | |
| 4. non pertinent | 6 | | <i>négatif</i> : 6 |
| 5. indécis | 11 | 11 | 11 |
| Rapport de Vraisemblance (RV) | | | |
| Catégories | Nb de collocations | Découvrir des règles d'association | Construire des patrons d'extraction |
| 1. pertinent | 52 | <i>positif</i> : 52 | <i>positif</i> : 88 |
| 2. spécifique | 30 | <i>négatif</i> : 45 | |
| 3. général | 6 | | |
| 4. non pertinent | 9 | | <i>négatif</i> : 9 |
| 5. indécis | 3 | 3 | 3 |

Table 4: Corpus des CVs (non normalisés). Évaluation des collocations pour deux sous-tâches : établir une classification pour découvrir des règles d'association entre concepts et pour construire des patrons d'extraction. Les 100 premières collocations de type Nom-Nom classées avec l'Information Mutuelle et le Rapport de Vraisemblance sans application d'élagage sont évaluées.

sont beaucoup plus nombreuses lorsqu'elles sont extraites avec le Rapport de Vraisemblance. L'Information Mutuelle extrait davantage de collocations spécifiques. Ceci corrobore les résultats présentés dans la section précédente qui montraient que l'Information Mutuelle était une mesure particulièrement bien adaptée pour extraire des collocations afin de constituer ou enrichir des lexiques spécifiques.

La seconde remarque à prendre en compte tient au fait que selon une sous-tâche à réaliser (par exemple, construire une classification conceptuelle pour découvrir des règles d'association entre concepts ou pour construire des patrons d'extraction) la qualité des collocations extraites diffère. Dans le tableau 4, les collocations utiles (resp. inutiles) pour chacune de ces sous-tâches sont appelées des exemples positifs (resp. négatifs). Une telle différence est significative avec l'Information Mutuelle. Bien que moins flagrantes, les différences restent importantes entre le nombre de collocations utiles et inutiles classées avec le Rapport de Vraisemblance pour chacune des sous-tâches (voir tableau 4).

5. Conclusion et Perspectives

Cet article montre que pour définir la pertinence d'une collocation, il est nécessaire de déterminer la tâche que l'expert souhaite effectuer avec la terminologie. Nous avons validé nos résultats à partir d'un corpus de CVs. Nous

avons montré que selon les tâches et les sous-tâches à réaliser, certaines mesures statistiques étaient plus ou moins bien adaptées pour classer les collocations. Une perspective importante à ce travail consiste à étudier la qualité de la terminologie extraite en définissant différentes autres tâches et en utilisant d'autres mesures statistiques.

Plus généralement, une trace dans les textes qui peut être définie comme un groupe de mots ayant une structure grammaticale déterminée est réellement pertinente si elle est utile pour la tâche que l'expert souhaite réaliser. La plupart des systèmes d'extraction de la terminologie (et plus généralement des systèmes de Traitement Automatique du Langage) s'appuient seulement sur les textes en les traitant avec des outils statistiques et/ou linguistiques pour obtenir un résultat pertinent. Seulement, pour obtenir des traces pertinentes en choisissant les paramètres ou les méthodologies les plus adaptés, il est indispensable de prendre en considération l'utilisation souhaitée des traces obtenues. Cette utilisation peut d'ailleurs être plus ou moins indépendante des textes étudiés.

Remerciements : Je remercie Yves Kodratoff, Violaine Prince et les relecteurs pour leurs conseils ainsi que Serge Baquedano (société PerformanSe) pour la constitution d'un des corpus étudiés dans cet article.

Références

- AZÉ J. & ROCHE M. (2003). Une application de la fouille de textes : l'extraction des règles d'association à partir d'un corpus spécialisé. *Revue RIA-ECA numéro spécial EGC03*, **17**, 283–294.
- AZÉ J. (2003). *Extraction de Connaissances dans des Données Numériques et Textuelles*. PhD thesis, Université de Paris 11.
- BRILL E. (1994). Some advances in transformation-based part of speech tagging. In *AAAI, Vol. 1*, p. 722–727.
- CHURCH K. W. & HANKS P. (1990). Word association norms, mutual information, and lexicography. In *Computational Linguistics*, volume 16, p. 22–29.
- CLAS A. (1994). Collocations et langues de spécialité. *Meta*, **39**(4), 576–580.
- DAILLE B. (1994). *Approche mixte pour l'extraction automatique de terminologie : statistiques lexicales et filtres linguistiques*. PhD thesis, Université Paris 7.
- DUNNING T. E. (1993). Accurate methods for the statistics of surprise and coincidence. *Computational Linguistics*, **19**(1), 61–74.
- FREITAG D. (1998). Toward general-purpose learning for information extraction. In *Proceedings of the Annual Meeting of the ACL*, p. 404–408: Morgan Kaufmann Publishers.
- JANETZKO D., CHERFI H., KENKE R., NAPOLI A. & TOUSSAINT Y. (2004). Knowledge-based selection of association rules for text mining. In *Proceedings of ECAI'04, IOS Press, Valencia, Spain*, p. 485–489.
- KODRATOFF Y. (2004). Induction extensionnelle : définition et application l'acquisition de concepts à partir de textes. *Revue RNTI E2, numéro spécial EGC'04*, **1**, 247–252.

- KODRATOFF Y., AZÉ J., ROCHE M. & MATTE-TAILLIEZ O. (2003). Des textes aux associations entre les concepts qu'ils contiennent. *Numéro spécial de la revue RNTI "Entreposage et Fouille de données"*, **1**, 171–182.
- MATHIAK B. & ECKSTEIN S. (2004). Five steps to text mining in biomedical literature. In *Proceedings of « Data Mining and Text Mining for Bioinformatics » workshop of ECML/PKDD Conference*, p. 44–49.
- ROCHE M. (2004). *Intégration de la construction de la terminologie de domaines spécialisés dans un processus global de fouille de textes*. PhD thesis, Université de Paris 11.
- SOBOROFF I. & HARMAN D. (2003). Overview of the trec 2003 novelty track. In *NIST Special Publication: SP 500-255 The Twelfth Text Retrieval Conference (TREC 2003)*.
- SRIKANT R. & AGRAWAL R. (1997). Mining generalized association rules. *Future Generation Computer Systems*, **13**(2–3), 161–180.
- THANOPOULOS A., FAKOTAKIS N. & KOKKIANAKIS G. (2002). Comparative Evaluation of Collocation Extraction Metrics. In *Proceedings of 3rd International Conference on Language Resources and Evaluation (LREC'02)*, volume 2, p. 620–625.

**De l'observation à l'intervention
dans les processus de gestion
ou Faut-il verbaliser les sens interdits ?**

*Nils Ferrand et Richard Raymond
Cemagref Montpellier*

Nils.Ferrand@montpellier.cemagref.fr;ric_raymond@yahoo.com



Source : Jean-Yves le Tetour,

<http://www.pbase.com/pingupingu/image/35993978>

*Je m'emporte et mes sens interdits
Impriment leur désordre en tout ce que je dis
Tite et Bérénice, Corneille*

Le sens d'une interrogation

Les processus de gestion collective de ressources réputées naturelles, de l'environnement ou de territoires partagés sont devenus l'objet d'études de plus en plus nombreuses. L'intérêt croissant porté à ces processus est concomitant à l'affirmation de nouveaux modes de gouvernance. Le décentrement des processus décisionnels interroge. En effet, il est maintenant communément admis que les représentations, individuelles ou collectives, orientent et structurent ces processus et les actions qui les accompagnent. Or, nombre d'études montrent que les acteurs impliqués dans un même processus de gestion ont des objectifs et des représentations parfois fort différents du système qu'ils gèrent ensemble et du système d'acteurs dont ils font partie. Les études entreprises visent donc à dévoiler et comprendre les modes de coordinations des acteurs. Comment une pluralité de sens (significations et finalités) est-elle réduite en un sens commun structurant l'action collective sur un système partagé ?

Cette interrogation conduit à travailler sur le sens des rapports que les acteurs entretiennent entre eux et avec les choses qu'ils gèrent et partagent mais aussi sur le sens que ces mêmes acteurs donnent à leurs actions et leurs revendications. Elle conduit enfin à interroger les sens que l'analyste propose pour ces rapports ou ces actions. En effet, les divers travaux d'analyse saisissent, dévoilent, proposent et, pour finir, produisent, eux aussi, du sens. Ces attributions de sens posent de nouvelles questions, méthodologiques et déontologiques, aux chercheurs. L'une d'elles est commune à toutes les recherches engagées pour une meilleure gestion de l'environnement ou un développement durable : faut-il verbaliser les sens interdits ?

Cette question, à l'allure paradoxale ou surréaliste, concerne la recherche impliquée dans l'analyse des formes de gestion collective des ressources ou des espaces réputés naturels. Elle devient prégnante lorsque les chercheurs s'interrogent sur les conséquences de la publicisation de leurs travaux concernant les processus de gestion collective qu'ils étudient. Elle fait également écho à ce qui se passe dans les processus de gestion collective eux-mêmes. Est-ce que tout peut être et doit être dit ? Est-il possible, en toute situation, d'explicitier et de dévoiler les règles, non conventionnelles, refoulées ou anormales, qui président aux avis et positions défendus par les acteurs impliqués dans un processus de gestion collective ? Il s'agit d'explorer des processus de gestion des ressources communes en observant leurs affichages, leurs traces et leurs restitutions pour comparer les attendus et les produits. Notre interrogation concerne la méthodologie et la déontologie des acteurs impliqués dans ces processus de gestion. Nous aborderons cette question à partir d'un triple point de vue.

Point de vue des acteurs impliqués dans ces processus de gestion collective, d'abord. Ces acteurs peuvent-ils toujours dévoiler le sens de leur action ou de

leur propos au sein des arènes de concertation et de gestion collective ? Ce dévoilement répond au souci de transparence prôné ici ou là comme une des conditions nécessaires à la bonne gouvernance des ressources et espaces partagés. Mais ne risque-t-il pas de mettre à mal la stratégie des acteurs concernés et, parfois, le processus lui-même ?

Le deuxième point de vue concerne la possibilité même qui est offerte aux chercheurs de saisir les sens et les significations qui sont contenus dans les discours et pratiques des acteurs observés. Est-il possible d'attribuer un sens univoque à une intention qui n'est pas la sienne ? Si, comme nous le craignons, la réponse à cette question est négative, quel est alors la valeur des apports et études scientifiques qui accompagnent les processus décisionnels ?

Enfin, le troisième point de vue interroge la pratique du chercheur. Faut-il et comment révéler les sens interdits lorsqu'on les détecte, qu'il s'agisse des non-dits stratégiques, des liaisons et coalitions invisibles ou des effets collatéraux de toutes natures ? Cette question bouscule le bien fondé de la restitution de la totalité des travaux aux acteurs de terrain. Elle interroge la publication de résultats concernant des processus en cours. Elle replace le scientifique dans le champ politique dont il veut parfois s'extraire pour justifier d'un certain recul nécessaire à une objectivité recherchée.

Nous aborderons ces aspects à partir de nos expériences analytiques et interventionnistes dans le domaine de la gestion de l'environnement et des territoires. Au cours de ces expériences, nous avons pu observer, enquêter et parfois susciter des situations particulières d'interactions multi-acteurs. Des représentations ont alors été extériorisées, soit dans des discours, soit dans des artefacts. Certaines ont été échangées et ont évolué jusqu'à former un corpus commun, plus ou moins consensuel. Nous avons constaté que l'explicitation de significations et de finalités peut conduire différents acteurs à re-positionner leurs intentions d'actions vis-à-vis du monde, à émettre des avis et à négocier différentes formes d'engagement. Cela conduit à orienter le *sens* vers lequel ils veulent (finalité) ou disent (justification) désormais agir. Dans le même temps, les acteurs donnent *du sens* à leur action, se livrent et se dévoilent.

Notre propos n'est pas d'apporter des réponses définitives à ces questions. Il s'agit de relever quelques difficultés et interrogations liées à la production, au dévoilement et à la publicisation de sens afin d'engager une réflexion et d'ouvrir un débat sur le rôle des scientifiques dans les processus de gestion collective. La place des scientifiques dans un mode de gouvernement du monde décentré, distribué et complexe doit être cernée et éclaircie. Dans une première partie, nous proposons donc de discuter différentes catégories de sens que l'on peut projeter sur des situations de gestion et de les comparer avec les sens qui sont affirmés par les acteurs. Puis, nous nous interrogeons sur notre capacité à observer et caractériser ces différents sens, par enquête et observation. Enfin, nous évaluerons ce que peuvent être les effets induits du dévoilement des différents sens sur les acteurs et les processus.

Finalités et significations : les sens interdits de la gestion de l'environnement

Dans son ouvrage restituant la dimension stratégique de la gestion de l'environnement ou des ressources naturelles, Laurent Mermet distingue deux types d'acteurs de l'environnement : les acteurs ayant une action directe sur le système biophysique considéré et les acteurs ayant une action indirecte sur ce système (Mermet, 1992). Ces derniers élaborent des diagnostics, des scénarios et des choix de régulation qu'ils proposent ou imposent aux acteurs directs. Cette distinction entre une gestion directe et une gestion indirecte de l'environnement est ici fondamentale. C'est elle qui soutient la dimension stratégique mise en avant par l'auteur. Elle met en exergue l'importance des représentations et des significations données à voir, proposées ou imposées de différentes manières. En effet, les acteurs agissant directement sur l'environnement sont les créateurs ou « conformeurs » de l'environnement. Leurs choix d'action, leurs orientations tactiques et stratégiques sont induits par leurs représentations qu'ils ont du monde. Or, une part de ces représentations est déterminée par les informations et régulations (politiques) diffusées et promues par d'autres acteurs. Ceux-là agissent de manière indirecte sur l'environnement. Cette assertion répond aux paradigmes constructivistes tels que les explicite Philippe Corcuff (Corcuff, 1995). Elle n'est, en effet, contradictoire ni avec les théories sociales majoritairement interactionnistes, ni avec celles qui sont majoritairement structuralistes et holistiques.

Lorsqu'ils conçoivent des modes et processus de régulation, ces acteurs indirects doivent anticiper les réponses des acteurs directs (*i.e.* des agriculteurs) face à ces scénarios. Pour ce faire, ils cherchent et projettent les représentations que ces acteurs directs se feront des régulations et de leur impact, prévoyant ainsi leur comportement (adoption, défiance, adaptation...). Or, nous aborderons cet aspect dans notre deuxième partie, de multiples contraintes interdisent l'accès aux sens des rapports que ces acteurs directs entretiennent entre eux et avec le système géré. Mais, et nous abordons un point important de notre propos, quand bien même ces projections de représentations seraient exactes¹ *a priori*, leur dévoilement dans une arène publique risque immédiatement d'induire leur contradiction. En effet, nous avons remarqué que le porter à connaissance d'une politique visant à la coordination d'usage de ressources limitées impose aux concernés d'inventer des stratégies de second-ordre. Les enseignements fournis par l'observation de jeux de rôles s'avèrent, ici, très intéressants (Darré, 2005) Chaque acteur qui dispose d'une anticipation du comportement collectif, anticipe l'impact de cette politique sur ses partenaires et recherche à son tour

¹ C'est-à-dire qu'au moment de l'analyse, il y a coïncidence entre la représentation que se font les acteurs directs et l'anticipation de cette représentation par un acteur indirect.

une stratégie pour maximiser son bénéfice personnel. Cette question a été longuement débattue en théorie des jeux. Si un tel comportement stratégique réflexif se répète dans la population, l'anticipation politique initiale, même exacte, devient fautive par le jeu de ses effets dès lors qu'elle est rendue publique ou connue d'un petit nombre d'acteurs directs. Or, de nombreux processus de gestion d'une ressource partagée et limitée reflètent ce comportement. Combien de déclarations agricoles sont orientées par la connaissance des règles d'attribution des primes ou des quotas d'utilisation d'une ressource ?

Au-delà de cette dimension stratégique, le processus de dévoilement des sens des représentations et des actions des acteurs a d'autres conséquences qui interrogent le statut même de vérité de ce qui est dévoilé. Les représentations et le sens attribués à l'environnement et au système social ne sont pas stables. Ils évoluent au gré des apprentissages individuels et collectifs nés de nouvelles expériences ou de pratiques relationnelles. Construit, le sens des choses est dynamique. Les confrontations successives aux autres et aux faits sont des mises à l'épreuve qui le modèlent peu à peu. La verbalisation réflexive d'un sens est une mise à l'épreuve, dans le champ social et intersubjectif, aussi importante que peut l'être une mise à l'épreuve des choses par l'expérience. Comme l'expérience, elle est un moteur de changement qui peut rendre une signification caduque dès lors qu'elle a été explicitée publiquement et a été l'objet de réactions.

Ce constat prend une dimension particulière lorsqu'un sens nouveau se heurte au "bon sens" établi, au sens de ceux qui parlent les deux pieds dans la glaise. Parce qu'ils s'opposent au "bon sens" communément admis, combien d'argumentations, combien de critères de décision sont écartés avant même un début de réflexion ou une éventuelle mise à l'épreuve des faits ? Pourtant, il n'y a aucune raison de croire que le "bon sens" des gens du cru est plus fiable ou plus efficace qu'un sens proposé de l'extérieur. Certains sens nouveaux, contraires au "bon sens" commun, ne doivent donc pas être verbalisés au risque d'être discrédités et rendus caduques par ce discrédit.

Ces constats reposent sur un a priori, celui que les sens interdits portent une part importante de vérité. Or rien ne permet de l'affirmer. Il existe d'autres propriétés du dévoilement de sens qui concernent des propositions qui ne sont, a priori, pas vraies mais qui le deviennent par leur dévoilement. Ceci concerne, en particulier les propositions auto-réalisatrices ou les actions auto-justificatrices. Ces propositions ou ces actions, dont le sens est affirmé et explicité, peuvent modifier le comportement du système qu'elles concernent pour le conformer au sens qu'elles portent. Il en va ainsi de certains travaux d'aménagement du territoire qui, lorsqu'ils sont entamés, n'ont pas de caractère nécessaire mais dont la nécessité est établie une fois l'aménagement réalisé. De nombreux travaux routiers en sont de bons exemples : non nécessaires lorsqu'ils sont

planifiés, ils s'avèrent indispensables lorsqu'ils sont finis parce que leur réalisation a engendré une évolution du système de circulation automobile conforme à leur fonction.

Un autre constat se répète. Il est admis que toute mesure, tout indice, tout indicateur est souvent, sinon toujours, incertain. Cette incertitude est liée au processus d'évaluation ou au phénomène évalué. Dans les arènes publiques, ces incertitudes ne sont qu'exceptionnellement abordées et citées, sinon pour appuyer des débats contradictoires. Or, la finalité de l'action collective ne peut souvent être partagée et acceptée que si le sens qui la supporte est clair et établi. L'incertitude s'accompagne souvent de doute. Elle est inconfortable et n'autorise pas, au regard des principes du développement durable, des actions radicales qu'on estime pourtant devoir prendre. Pour cette raison, réfléchies, explicitées à certains moments de la concertation, les incertitudes font partie de ces dimensions passées sous silence au moment de la publicisation de la décision prise au nom et pour la collectivité. L'exemple de la gestion des pollutions d'origine agricole illustre ce propos. Les actions pour limiter la concentration d'azote dans les eaux des rivières ou des nappes ont été engagées sur la base d'un postulat clair : les nitrates étaient le facteur limitant de l'eutrophisation des eaux et étaient, de plus, toxiques. Les recherches ont montré, depuis, que ce n'était pas le cas. Néanmoins, diverses actions restent étayées et légitimées par ces croyances, sans qu'aucune incertitude ne les accompagne.

Enfin, la formulation du sens de certaines actions répondant aux principes du développement durable est un obstacle à la réussite même de ces actions. Il en va ainsi de la transparence ou de la confiance qui devrait être de mise au sein des processus de gestion collective pour un partage équitable des ressources. Or le fait même d'énoncer qu'un processus est transparent génère la suspicion de la part de ceux à qui est destiné le message. Il en est de même de la confiance. Celle-ci se donne mais ne peut se demander. Dès qu'un acteur énonce la proposition « faites moi confiance ! », la réaction est imparable : cette proposition est suspectée de masquer d'autres enjeux moins avouables. Ce phénomène a été explicité par Jon Elster sous le nom d'*effet essentiellement secondaire* (Elster, 1986). L'effet essentiellement secondaire est l'effet secondaire (ou indirect et venant en supplément) de l'effet visé. Mais cet effet ne peut être réalisé par l'agent qui se donnerait comme but premier de le produire. Jon Elster en fait l'analyse dans son essai sur les limites de la rationalité intitulé *Le Laboureur et ses enfants*. Faire en sorte que la confiance soit établie, faire en sorte que le processus soit jugé transparent mais ne pas le dire ou le réclamer, tel pourrait être le conseil qui émergerait de ce constat.

Certaines dimensions des processus de gestion collective ne sont ou ne peuvent donc pas être exprimées ni en termes d'impact ni en termes d'objectifs. D'autres ne devraient pas être avancées au nom de la vérité. Ces sens concernent les changements dans les structures relationnelles, les régimes d'équité ou de vérité et les incertitudes relatives sur les processus biophysiques et sociaux. Leurs expressions par les acteurs eux-mêmes contribueraient à mettre à mal les objectifs qu'ils recherchent. Cependant, immédiatement associée à ces premières observations, une question survient, interrogeant la pratique même de l'analyse scientifique. Sommes-nous capables, et dans quelles conditions, de saisir les sens et les significations qui sont contenus dans les discours et pratiques des acteurs observés.

Les sens projetés par l'analyse des situations de gestion

La démarche d'analyse des situations de gestion, qu'elle se fonde sur l'observation d'un processus ou sur des enquêtes ex-post, cherche à reconstituer l'évolution des positions des acteurs au cours du temps en fonction de la situation externe, des autres positions perçues et, éventuellement, des interventions d'objets ou d'acteurs intermédiaires. L'analyste interprète donc, selon ses hypothèses, des traces laissées ou rapportées. Il propose des sens. Lorsque la recherche est impliquée, il peut restituer directement ce sens aux acteurs participant au processus, qu'il s'agisse là de son contrat ou qu'il entende ainsi contribuer à une amélioration. Cette intervention s'appuie sur une double hypothèse formelle : que l'analyste a capacité à identifier les sens, à y accéder, et qu'il est bon, utile, qu'il les dévoile. Nous reviendrons sur ce dernier point dans notre troisième partie.

En supposant, comme le considère les sciences positives, qu'il y ait une position de « surplomb » du chercheur, le rôle de l'analyste est donc de traquer les sens cachés, parfois interdits, ce que l'on peut simplement nommer : *comprendre*. Quels sont les avantages ou les principes qui pourraient expliquer qu'un observateur puisse saisir les sens cachés, non-dits, masqués d'une situation de gestion ? Pourquoi les participants comprendraient-ils *moins*² le sens de leur action ou de leurs énoncés, qu'un analyste ?

On peut proposer les raisons suivantes :

- parce que les acteurs doutent et ne veulent pas risquer d'avérer des incertitudes qui président à leurs actions et risqueraient d'entamer leur légitimation,
- parce qu'étant immergés dans les enjeux et le processus, leur mode d'appréhension des questions est déterminé par un cadre normatif restreint qui les oriente, par leurs valeurs ou leurs préférences,

² La définition d'une métrique de compréhension est, au demeurant, un exercice qui paraît bien difficile.

- parce qu'ils sont manipulés par d'autres : leur compréhension est intentionnellement détournée de certaines interprétations par l'influence de certains acteurs y trouvant leur intérêt,
- parce qu'ils ne disposent pas des clés de compréhension, des modes d'analyse préalable, des théories qui permettent de « donner un sens », par analogie ou déduction – ceux-ci sont fournis par la connaissance experte ou scientifique, ou par des savoirs profanes explicatifs avérés,
- parce que l'acceptation de certaines explications ou sens les conduirait dans des directions d'action qu'ils ne peuvent accepter pour des raisons éthiques (axiologiques ou déontologiques) ou psychologiques – on se rapproche ici d'un processus de refoulement, sans prendre un sens psychanalytique,
- parce qu'ils ne disposent pas d'une ouverture suffisante, en terme de complexité, sur les facteurs et les conséquences,
- parce qu'ils ne disposent pas des outils intellectuels et calculatoires leur permettant d'envisager les effets possibles de positions ou d'actions,
- parce que les participants n'ont pas le droit social (normes) de laisser transparaître qu'ils ont compris,
- parce que c'est un mode de manipulation ou de test de l'analyste par les acteurs.

Si l'une au moins de ces raisons est vraie, alors la recherche de sens par un acteur extérieur au processus de gestion, l'analyste, est justifiée.

Les trois modalités principales de saisissement des sens sont l'observation, l'enquête et l'expérimentation. Dans l'observation, l'analyste suit au cours du temps des comportements et interactions entre acteurs, transcrit les énoncés et les constructions intermédiaires (artefacts), puis formule ses interprétations à partir de théories ou d'hypothèses explicatives. On peut alors distinguer deux types de sens : les sens qu'un autre analyste ou les participants eux-mêmes reconnaîtraient sans doute, et les sens interprétés sur lesquels il n'y a pas consensus explicite. Nous supposons que les sens attribués par l'analyste suivent toujours le principe d'autorité de la raison et ont donc valeur de vérité « en l'état actuel de nos connaissances ». Parmi les sens ne faisant pas consensus, mais vrais « en l'état actuel de nos connaissances », certains relèvent de l'interdit. Ceux-là restent masqués car il y a un refus de formulation explicite par les acteurs en interaction, par exemple en entretien avec l'analyste.

Dans l'enquête, vers un individu ou un groupe, l'analyste construit un dialogue portant sur les sens pour explorer des interprétations. Les questions qu'il pose doivent lui permettre de comprendre. Là encore l'interprétation est de mise, surtout lorsque l'analyste a recours à des questions partiellement manipulatoires qui n'abordent qu'indirectement leur objectif réel (par exemple pour interroger sur les valeurs ou pour contrecarrer des questions). Les sens interdits émergent soit explicitement soit implicitement selon que l'interrogé est seul ou non.

L'analyste peut les reformuler et les restituer pour les vérifier. La réaction de l'interrogé peut aller de l'acceptation associée à de la gratitude pour les avoir explicités, à un déni accompagné d'un refus de poursuivre l'interaction avec l'analyste.

L'expérimentation est une combinaison d'enquêtes et d'observations. Il s'agit d'induire une situation sociale ou cognitive particulière, et d'observer alors les résultats pour en déduire des interprétations. La forme la plus directe d'exploration expérimentale des sens interdits consiste à soumettre un sens interdit comme objet dans un groupe de discussion. La règle est alors que, dans cette situation construite, le sens n'est temporairement plus interdit : par exemple aborder la séduction et les rituels amoureux. Les réactions peuvent aller ici du contournement (parler malgré tout d'autre chose) à la reformulation ou à l'abord explicite.

Malgré les possibilités données à l'analyste pour saisir des sens cachés, l'accès à certains sens de représentations et de pratiques d'acteurs engagés dans un processus social reste parfois délicat. Deux aspects limitent son investigation et détournent le sens de son analyse : la subjectivité (l'accessibilité) et l'intention (le masquage). Les sens « donnés » aux choses par un acteur ne sont accessibles à autrui que par des traces dans l'interaction, dans le discours et dans des objets intermédiaires, résultats de la transcription de sens par l'acteur. L'autre interprète avec sa subjectivité et reconstruit un sens en supposant qu'il correspond à celui de l'acteur énonciateur. Il suit en cela le principe de charité avancé par Donald Davidson concernant l'intercompréhension supportée par le langage (Davidson, 1993). Si ce principe de charité est nécessaire à l'intercompréhension, il masque de nombreuses approximations ou erreurs d'interprétation qui ne peuvent être levées que par des investigations supplémentaires. Ce type de difficulté se répète, dans le champ de l'observation scientifique, de la formulation (du sens à sa matérialisation) à l'interprétation (de la matérialisation à un sens reconstitué).

De même, au cours de nombreux entretiens et enquêtes, un doute subsiste, souvent. Il s'agit d'un interdit intentionnel. Celui-ci peut être de trois ordres. Il peut s'agir d'une intention délibérée de l'informateur de cacher à l'enquêteur le sens de ses actions. Ceci peut être dû à une volonté de garder secret une stratégie, à une gêne sur les motivations réelles des actions conduites (motivations indicibles ou non conformes aux valeurs habituellement partagées). Il peut aussi s'agir d'une volonté délibérée de l'énonciateur de tenter de formuler un sens différent de celui auquel il adhère, pour satisfaire un objectif personnel, en anticipant l'interprétation que l'autre en fera. Il peut s'agir enfin d'une volonté de l'informateur de répondre à l'enquêteur en recherchant la réponse qui convient à la question posée, même si la réponse proposée ne traduit en rien le sens de ses représentations ou de ses motivations personnelles. Ce dernier cas de figure ne peut être complètement ignoré lorsque l'informateur

fait, lui aussi, preuve d'empathie et recherche à s'inscrire dans le système de valeurs et de pensées de l'enquêteur. Ainsi, de nombreuses enquêtes menées par des agronomes ou des sociologues auprès d'agriculteurs révèlent un sens aux pratiques de ces derniers qui s'inscrit dans un registre bio-technique ou dans "l'air du temps", mais ne correspond pas aux actions entreprises par l'agriculteur. Ce sens est suggéré par l'agriculteur lui-même qui adapte son discours et donne le sens de ses propos en fonction de ce qu'il pense que son interlocuteur attend de lui.

Lors d'un énoncé particulier, chacun des protagonistes peut donc saisir et interpréter différemment l'intention de sens contenue dans les propos de l'autre. Ce qui importe, c'est de relever que, s'il n'y a pas unicité des interprétations, il y a cependant une interprétation dominante, consensuelle, simple (répondant au principe de charité). De ce fait, un espace s'ouvre pour caractériser et verbaliser d'autres sens, jouant ainsi avec les formes d'interdit précédemment évoquées : formulation, interprétation, intention. Pour lever ces difficultés, trois types de reformulations sont possibles : « voulais-tu dire que... », « j'ai compris que... », et « tu as dit cela parce que... ». Cependant, ces exercices de précision sont lourds et peu propices à l'exploration de terrain.

Notre capacité à observer et caractériser les différents sens inscrits dans un processus de gestion collective n'est donc pas exempt d'interdits ou de masques. Nous remarquons, en particulier, que le statut de véracité de ce que nous pouvons formuler et restituer aux intéressés en situation de gestion collective ne peut être entièrement stabilisé. De ce fait, il existe un risque de ne pas être crédibles, et donc inutiles. Cela s'applique aux sens donnés aux savoirs scientifiques comme aux croyances profanes consensuelles.

De la verbalisation des sens interdits à l'interdit de verbaliser les sens

Si l'analyste peut saisir le sens des actions observées, doit-il pour autant le rendre public et le divulguer à la communauté observée ? Cette question replace les productions scientifiques dans leur contexte de production. En effet, un exposé, un article, un ouvrage publié n'est pas, loin s'en faut, sans conséquence pour la communauté étudiée. Certaines de ces restitutions sont faites en direction de la communauté locale. D'autres, faites en direction de la communauté scientifique n'en sont pas moins publiques et la capacité des acteurs locaux à aller chercher un document les concernant est parfois déroutante. Quoi qu'il en soit, la verbalisation de sens interdits est très souvent entendue par tout ou partie de la communauté concernée. Cette verbalisation n'est donc pas sans conséquence sur les processus mêmes de gestion collective.

Concernant les conséquences de l'analyse, nous notons que même lorsque l'énoncé d'une interprétation est faux (c'est-à-dire que le sens proposé diffère de celui qui est intentionnellement produit par l'énonciateur ou de la réalité des faits), il prend un statut relatif de vérité par la vertu de l'analyse. Le statut donné à l'analyste est transposé à ses interprétations, ces deux statuts étant conjointement révisables. Ceci a deux conséquences importantes. D'une part, si l'analyste jouit d'une opinion favorable, ses interprétations et les sens qu'il produit et formalise sont reçus avec la même faveur. En revanche, si les sens produits sont discrédités ou invalidés par une mise à l'épreuve de la réalité ou du sens commun, c'est le crédit de l'analyste et de sa communauté qui en est affecté. Ce deuxième mouvement est très mal vécu par la communauté scientifique qui perd, là, un terrain d'étude où elle a pu s'investir depuis de longues années.

Les conséquences du dévoilement de sens interdits précédemment cités apparaissent là-encore. Face à une publicisation de leur stratégie ou du sens de l'action collective, les acteurs engagent une stratégie d'ordre supérieur. De ce fait, les rôles et les relations sociales s'en trouvent être modifiés. De même, les zones d'incertitudes, les réseaux informels, les implicites se délocalisent. En effet, sources de pouvoir et d'équilibre pour les acteurs, ces aspects de la gestion collective d'une ressource partagée sont fonctionnels comme le sont les petits arrangements entre acteurs ou les réseaux informels (Beuret, 1999, Beuret et al, 2001). Lorsque les structures ou les ressorts de pouvoirs sont dévoilés, ceux-ci se déplacent car l'incertitude, l'implicite et le secret sont fonctionnels.

D'autre part, comme l'évoque Yves Luginbühl à propos des paysages ligériens et des saules émondés, la production de sens par le chercheur conduit parfois à l'adoption de ce sens dans la société locale. Adoption que le chercheur n'avait non seulement pas voulue mais dont il se défendrait de bonne fois si on lui en prêtait l'intention (Luginbühl, 1995). Dévoiler des sens cachés ou produire de nouveaux sens conduit donc, parfois, à l'importation de normes et valeurs au sein du système social cible.

Enfin, l'explicitation des objectifs mêmes de processus et des stratégies diffuses est parfois contre-productive vis-à-vis des objectifs du développement durable. Outre le fait qu'ils conduisent presque systématiquement à l'élaboration de stratégies d'ordre supérieur, ces dévoilements conduisent également à rendre inopérantes les stratégies des acteurs locaux. Ceci est particulièrement vrai pour les stratégies de construction d'une réalité partagée ou d'une communauté d'intérêt. En effet, construites par petites touches successives, établissant des conventions qui se répondent et s'étayaient mutuellement, s'appuyant sur divers éléments symboliques dont les significations et les légitimités sont ténues, ces stratégies sont disqualifiées si tant est qu'un chercheur dévoile le caractère construit et intersubjectif de ces réalités sociales. Certains dévoilements sont donc contre-productifs au regard de l'objectif de la recherche entamée ou de l'engagement politique du chercheur.

Si l'un des buts de l'activité scientifique est de démasquer les entreprises en mystification, les mythes restent nécessaires et leur dévoilement peut s'avérer dommageable pour la préservation de l'environnement. Ces considérations replacent le scientifique et l'analyste dans le champ politique dont il ne peut, comme l'annonçait Pierre Bourdieu, jamais totalement s'abstraire. Ainsi, il nous semble que le chercheur ne peut se satisfaire de la démonstration de la diversité des représentations ou des sens attribués et saisis, à tort ou à raison, par lui. Il lui faut travailler, dévoiler ou rechercher une compréhension de la hiérarchie de ces sens, de ces représentations et des enjeux qui les accompagnent en les mettant en perspective avec un objectif. Au-delà d'une recherche pour la seule compréhension des processus cognitifs et sociaux, des différentes relations au monde, il nous semble important de retrouver une certaine finalité à la production de connaissances/explicitations de sens concernant les modes de gestion collective. Il existe des besoins d'analyse et de recherche sur les sens vers des objectifs d'action.

Au-delà des questions concernant le statut des sens et des interprétations relevés et produits par le chercheur, c'est le rôle de la verbalisation et du dévoilement de ces sens qui doit être abordé. Les conséquences et les opportunités de ces dévoilements doivent être discutées. L'insertion de ces productions scientifiques dans les processus de constructions sociales et d'institutionnalisations de conventions et leurs articulations avec les processus cognitifs à l'œuvre dans les dispositifs de gestion collective doivent être analysées. Il y a là matière à soulever une véritable question. Il nous semble opportun d'ouvrir ici ce débat et de le poursuivre dans un cadre déontologique de la recherche et de l'ingénierie de l'intervention. Un programme d'étude visant à explorer systématiquement et expérimentalement ces questions pour, premièrement, qualifier notre capacité à mesurer ces sens interdits et en établir le statut, et, deuxièmement, évaluer les effets de leur verbalisation pourrait être mis en place avantageusement.

Références :

- Beuret J.-E. (1999), "Petits arrangements entre acteurs : les voies d'une gestion concertée de l'espace rural", *Natures Sciences Sociétés*, Paris, Elsevier, vol. 7, n°1 : 21-30.
- Beuret J.-E. et Trehet C. (2001), "Pour la gestion concertée de l'espace rural : appuyer des médiations territoriales", *Le courrier de l'environnement de l'INRA*, Versailles, INRA Éditions, 43 : 25-39.
- Corcuff P (1995) *Les nouvelles sociologies*, Paris : Nathan Université, 126 p.
- Darré W. (2005), *Comportements des acteurs dans le jeu et la réalité : indépendance et correspondance*. Thèse de l'ENGREF, 401 p.

- Davidson D. (1993), *Enquêtes sur la vérité et l'interprétation*, Nîmes : éd. Jacqueline Chambon, tr. fr. Pascal Engel, 420 p.
- Elster J. (1986), *Le Laboureur et ses enfants, Deux essais sur les limites de la rationalité*, Paris : Ed. de Minuit, tr. fr. A. Gerschenfeld, 208 p.
- Luginbühl Y. (1995), "Quelques avatars de la recherche sur le paysage", *Paysage au pluriel, pour une approche ethnologique des paysages*, Cahier n°9 de la collection Ethnologie de la France, Paris : Ed. de la Maison des sciences de l'homme, 219-224.
- Mermet L. (1992), *Stratégies pour la gestion de l'environnement. La nature comme jeu de société ?*, Paris : L'Harmattan, Coll. « Environnement », 205 p.
- Maurel P, Craps M, Cernesson F, Raymond R, Valkering P, Ferrand N. (à paraître), "Concepts and methods for analysing the role of IC-tools in Social Learning processes for River Basin Management", *Environmental Modelling and Software*, Special issue on IEMSS 2004 conference theme "Complexity and Integrated Resources Management", Elsevier Science.

CORPUS, TRACE et APPRENTISSAGE



Christian Champendal, PRAG d'Anglais à l'Ecole Polytechnique de l'University de Nice à Sophia-Antipolis (UNSA). Recherche et Développement au CRDL (laboratoire d'ingénierie linguistique, Professeur H. Zingle), Faculté des Lettres et Sciences Humaines - (UNSA). champend@unice.fr / +33 493 929206

RÉSUMÉ: Un corpus est un ensemble de données écrites ou orales, "dépositaire d'expériences passées". Les systèmes stochastiques d'apprentissage à mémoire (*Memory-Based Parsers*) à partir d'une base connue, probabilisée, analysent le nouveau. La définition de hiérarchies ou de dépendances donne du relief à l'interprétation et reflète la saillance des données, contrairement à un étiquetage à plat. L'organisation naturelle des éléments du corpus en familles ou Séries Lexicales, permet un traitement des particularités de chacune; complétant l'approche statistique elle l'optimise.

ABSTRACT: A corpus is a set of written or recorded data, "depository of past experiences". Memory-Based learning systems start from what knowledge is stored in the base to parse new occurrences. Unlike tagging, which appears rather flat, hierarchies or dependences gives volume to the representation of data. The natural organisation of corpus items into families or Lexical Sets, makes way for a more idiosyncratic processing that completes and optimises the statistical approach.

MOTS-CLÉS : Apprentissage, clusters, corpus, relation prédicative, série lexicale.

KEYWORDS: clusters, corpus, learning, lexical sets, predicative relations.

En Traitement Automatique des Langues (TAL), un **corpus** est un ensemble de données écrites ou orales, "dépositaire d'expériences passées". A partir d'une base connue, probabilisée, les systèmes stochastiques d'apprentissage à mémoire (*Memory-Based Parsers*), traitent le nouveau. La définition de hiérarchies ou de dépendances donne du relief à l'interprétation et reflète la saillance des données, contrairement à un étiquetage à plat (*tagging*). L'organisation des éléments du corpus en **Séries Lexicales**, permet un traitement des particularités de chacune. Elle est naturelle, puisqu'au sein d'une famille la même relation pourra être exprimée par divers lemmes. Complétant l'approche statistique, elle l'optimise.

1. Séries Lexicales

Les résultats sont extraits d'un corpus de 135 articles en TNS, soit 500 pages, 3900 paragraphes, 16 500 lignes et 350 000 mots-occurrences (2 millions de caractères et 3,27 Moctets). Les articles ont été sélectionnés dans le CD-Rom, *ICASSP'98 International Conference on Acoustics, Speech and Signal Processing*. Le corpus d'apprentissage est constitué de **Séries Lexicales** obtenues à partir des listes d'occurrences avec Z-Text®¹. Une centaine de séries est retenue après avis d'expert car quantitativement et qualitativement saillante : leur étude assure une couverture suffisante des relations du domaine.

| | | |
|---|---|--|
| <p>[R]: (<i>occ./Lemm.</i>)</p> <p><i>r1 r2 r3</i> ... <i>rn</i></p> | <p>[Use]: 2943/12)</p> <p>(1) <i>reuse</i> (2) <i>reused</i> (1) <i>reusing</i>, (15) <i>usage</i> (482) <i>use</i> (919) <i>used</i> (1) <i>useless</i> (136) <i>user</i> (56) <i>users</i> (109) <i>uses</i> (1094) <i>using</i> (17) <i>usual</i> (49) <i>usually</i> (61) <i>useful</i> (4) <i>usefulness</i></p> | <p>[Filter]: (1441/4)</p> <p>(809) <i>filter</i> (62) <i>filtered</i> (146) <i>filtering</i> (424) <i>filters</i></p> |
|---|---|--|

Figure 1. Séries Lexicales

Cette étape qui organise lexicalement les éléments de chaque série selon leurs occurrences de formes simples, composées et complexes permet déjà de repérer les liens inter-séries les plus récurrents.

2. Concordances

L'étude des **concordances** des prédicats nominaux ou verbaux, définit une **formule prédictive** valide pour les éléments prédictifs d'une même famille.

¹ Voir bibliographie.

...
 The analysis and synthesis **filter** banks **use** a 29th order lowpass prototype.[15193]
 In contrast, the recursive design **uses** only one multiplication and one subtraction![5867]
 The echo-path model **uses** a single pole single zero digital **filter**.[6429]
 In standard radar observation models, the optimum procedure is to **use** multi-dimensional
 temporal (matched) **filtering** prior to bearing estimation;[2047]
 Non-separable filters are **designed using** the method in [3] to extract the components from
 the signal.[3523]
 "Circularly Symmetric Filter **Design using** 2D Prolate Spheroidal Sequences".[14396]

Figure 3. Concordances (Extraits)

Cette formule, adaptable au type du prédicat (nominal, verbal), sensible au type du verbe (transitif, prépositionnel, ditransitif), intègre les adjoints circonstanciels, plus périphériques mais porteurs de sens. Elle est une **forme canonique** empiriquement reconstituée grâce au corpus, de type:

$C_0 = \text{Sujet(S)}$ <Prédicat(Pred)> $C_1 = \text{Complément(Od)}$
 $C_2 = \text{Complément(Od/Oi)}$
 $\text{Circ}_n = \text{Circonstanciel(Adjct)}$

L'ordre et le nombre des adjoints circonstanciels est flexible, s'agissant de constituants mobiles et optionnels. Pour <Use> et <Filter>, par exemple, la formule est:

C_0 <Use > C_1 $\text{Circ}_1 = \text{ASA}$ $\text{Circ}_2 = \text{TO}$ $\text{Circ}_3 = \text{IN}$

The algorithm **uses** two neighbouring pixels, one left and the other top, as a pioneering block to search for the best matched blocks inside a pre-defined window.

C_0 <Filter> C_1 $\text{Circ}_1 = \text{FROM}$ $\text{Circ}_2 = \text{IN/OVER}$ $\text{Circ}_3 = \text{TO}$ $\text{Circ}_4 = \text{BY USE OF}$
 In WI, pitch-cycle waveforms **are filtered in** the evolution domain to decompose the signal into two waveform surfaces, one characterising voiced speech and a second representing unvoiced speech.

Pour chaque prédicat d'une série, les places sont instanciées par des éléments privilégiés détectés par leur récurrence lors de l'extraction de résultats. <Filter>, par exemple:

$C_0 =$ {we, algorithm, filter, weight,...}
 $C_1 =$ {data, image, signal, noise, procedure, sample, waveforms,...}
 $\text{Circ}_1 =$ from {error, source, signal}
 $\text{Circ}_2 =$ in/over {domain, region, frequency band...}
 $\text{Circ}_3 =$ to do {imaging, noise removal...}
 $\text{Circ}_4 =$ by use of {statistics, optimization techniques...}

3. Clusters

Un balayage des principales séries établit leur **connectivité**. Il s'agit

d'affinités plus ou moins **larges** selon qu'un plus ou moins grand nombre d'éléments d'une série est fréquemment associé à ceux d'une autre série. L'extraction des **collocations** (formes associées) révèle automatiquement ces liens de proximité **directe** que sont les relations prédicatives (liens complémentaires à distance -1, +1, +2, notés **C₀, C₁, C₂**). La **formule prédicative**, extraite de l'examen des concordances prédicatives d'une série, prend en compte les divers liens circonstanciels (distance sur un horizon élargi du cotexte: **C_{irc1} C_{irc2} C_{irc3}... C_{ircn}**). Cette formule aide aussi à expliciter les relations sous-jacentes des formes composées, où les relations prédicatives sont empaquetées, "compactées", que l'élément concerné soit prédicatif ou non². Cette mise en forme dégage des régularités sous forme d'un treillis de relations prioritaires qui rendent le corpus intelligible.

[**FILTER**] est une série à fort caractère nominal; ses éléments prédicatifs sont largement **directement** associés aux séries: [**DESIGN, USE, CHOSE, IMAG, ESTIMAT, PERFORM,...**]; ils sont **indirectement** liés aux séries : [**DOMAIN, OPTIM, SHOW, USE, ...**]. La série [**USE**], plutôt verbale, est liée **directement** aux éléments des séries [**CODE, DATA, DESIGN, ESTIMAT, FILTER, IMAG, MODEL, OPTIM, SYSTEM,...**]; ainsi que par des liens circonstanciels avec [**SEARCH, DETERMIN,...**]. Ces deux séries ont entre elles des affinités **larges**.

| | | |
|-----------------------------|----------------------|---|
| a1/a2/a3/a4 b1/b2/b3 | b1/b2/b3 a1/a2/a3 | (...) use/uses/used/using (... ³) filter/filters/filtering (...) (...) filter/filters/filtering (...) use/uses/used (...) |
| a1/a2/a3/a4 c1/c2/c3 | c1/c2/c3 a1/a2/a3 | (...) use/uses/used/using (...) design/designs/ designing (...) (...) design/designs/designing (...) use/uses/used (...) |
| b1/b2/b3 c1/c2/c3 | c1/c2/c3 b1/b2/b3 | (...) filter/filters/filtering (...) design/designs/ designing (...) (...) design/designs/designing (...) filter/filters/filtering (...) |
| a1/a2/a3/a4 a1/a2 | b1 b1/b2/b3 | (...) use/uses/used/using (...) filter (...) design/designs/ designing (...) |
| b1/b2/b3 c1/c2/c3 ... | c1/c2/c3 | (...) use/uses of filter/filters/filtering (...) design/designs/to design/ for designing (...) |

Figure 2. Clusters (Exemples 2 à 2 et 3 à 3)

De plus, il apparaît des **clusters** ou réseaux de relations privilégiées avec

² Ex: user interface (5), user signature/s (4), user perception (6), user specified (2), user detector (3), user interference (3), user joint detection (2), user transmissions (2), user case (2), user error /s (6).

³ Les **inserts** et **expansions** sont plus ou moins contraints, de type et de taille divers : du groupe verbal (is/are/was/were/has been /have been/is being/are being used), nominal (adaptive FIR/Wiener/Lee's Filter) ou éléments adverbiaux, par exemple.

d'autres **séries communes**, deux à deux ou n à n; ces derniers sont étudiés prioritairement car éminemment rentables.

Pour la machine, l'apprentissage nécessite de **reconnaître** cette trace constituée préalablement réorganisée, rendue intelligible, porteuse de sens. C'est une **re-connaissance** qui doit également permettre d'intégrer le nouveau, le « non-déjà re-connu ». Ceci implique certaines modifications des chaînes connues contenues dans les bases de connaissances: omissions, additions, réagencements... La mise à jour incrémentale de ces bases est effectuée rétroactivement par traitement et classification des nouvelles occurrences « déviantes » (non absolument identiques à une des instances de la base de données).

Un document ou partie de document du domaine (énoncé, paragraphe, chapitre ou texte), sera donc interprété à l'aune des connaissances extraites et organisées à partir du corpus d'apprentissage. Par ailleurs, si un corpus ne contient qu'une partie ou sous-ensemble d'un domaine, il est porteur de signes avant-coureurs, empreintes plus profondes et plus denses qui vont permettre de baliser des itinéraires privilégiés.

Bibliographie

- BOD Rens, 1998 *Beyond Grammar*, CLSI Standford California
- CHAMPENDAL Christian & PITARQUE Thierry, 2005; *Lexical Sets and Text Processing*, Proceedings of The Workshop "Semantic Web Technologies for Machine Translation", in conjunction with the MT-Summit X, Cristina Vertan, Walther v. Hahn & Vladislav Kubon (Eds.), Phuket, 12 Sept 2005, pp. 10-13
- CHAMPENDAL C., 2004; *Language teaching and corpus linguistics*, IATEFL – CALL Review, pp.58 – 61, ISSN 1028-428
- KÜBLER Sandra, 2004; *Memory-Based Parsing*, John Benjamins B.V.
- MEL'ČUK Igor A., CLAS A, POLGUÈRE A, 1995 ; *Introduction à la lexicologie explicative et combinatoire*; Ed Duculot.
- TRIBBLE Chris, 2000; *Practical uses for language corpora in ELT* in A Special Interest in Computers, P. Brett and G. Motteram Eds, IATEFL.
- ZINGLE Henri *Travaux du LILLA*, 1999 ; Numéro spécial– La modélisation des langues naturelles – Aspects théoriques et pratiques.
- ZINGLE Henri, 1998 ; ZTEXT : un outil pour l'analyse de corpus in *Travaux du LILLA n°3*, pp 69-78.

Une théorie de la trace informatique pour faciliter l'adaptation dans la confrontation logique d'utilisation/logique de conception

Alain Mille – Yannick Prié
LIRIS UMR 5205 – Université Claude Bernard Lyon 1
prenom.nom@liris.cnrs.fr

1 Introduction

Un environnement informatique *perçu* offre des possibilités nombreuses et variées d'inter-actions¹ pour accompagner une activité humaine pour des tâches plus ou moins formatées. Nous ne faisons pas d'hypothèse *a priori* sur la notion d'environnement informatique, qui peut être considéré à des niveaux de granularité très variables. Le contour (ou le périmètre) ressenti d'un environnement dépend en effet fortement du contexte d'utilisation, lequel n'a de sens que pour l'utilisateur, acteur dans son environnement. Un environnement pourra alors aller d'une simple fenêtre de configuration d'une application jusqu'à l'ensemble des outils disponibles à l'utilisateur à un instant donné.

Les composants d'un environnement informatique sont conçus sur la base d'une analyse de l'activité des futurs utilisateurs et des tâches impliquées. Ils offrent différentes possibilités d'inter-action conformément à l'analyse menée. D'autre part, beaucoup de composants sont « ouverts » aux autres et peuvent être utilisés les uns avec les autres, les uns à partir des autres², *etc.* Le périmètre de l'environnement informatique *conçu* se révèle donc difficile à stabiliser à l'avance.

Un hiatus est donc inévitable entre environnements informatiques perçus et environnements informatiques conçus. Il nous semble donc que la question de l'accompagnement de l'activité de l'utilisateur par des « facilitateurs » d'appropriation des possibilités offertes se pose donc dans la conception même des environnements informatiques.

Après avoir précisé les notions de *logiques d'utilisation* et de *logiques de conception* qui se confrontent au moment de l'utilisation, nous proposons d'utiliser les traces d'inter-action comme support de facilitation de la gestion du hiatus par l'utilisateur. Enfin, nous présentons le processus de modélisation de l'utilisation nécessaire pour décrire les inter-actions et détaillons le modèle MUSETTE que nous proposons pour y parvenir. Nous illustrons ensuite brièvement la démarche de modélisation de l'utilisation avant de nous intéresser à la notion de « facilitateur » en fournissant un exemple dans le cadre d'un Environnement Informatique d'Apprentissage Humain.

¹ Inter-action : dénote le fait que les actions de l'utilisateur et les réactions de l'environnement sont dépendantes les unes des autres dans le cadre d'une activité particulière. L'utilisateur agit avec l'environnement comme inter-médiaire avec lui-même d'abord. Il n'est pas possible de dissocier l'action de l'attente du résultat de l'action qui précisément est le lieu de la confrontation logiques d'utilisation / logiques de conception.

² Il est possible par exemple d'autoriser de réaliser un document à partir d'un tableur mais d'exploiter intensivement les possibilités d'un autre environnement comme un traitement de texte dans ce tableur (et vice-versa).

Il ne s'agit pas ici de développer les questions épistémologiques importantes que l'approche soulève et de développer les enjeux d'une telle recherche, mais de contribuer à la réflexion sur la notion de traces et de leurs usages quand il s'agit d'environnements informatiques. Ces questions seront toutefois évoquées dans la discussion qui conclut l'article.

2 Logiques d'utilisation et logiques de conception

Les environnements mis à disposition sont conçus selon des *logiques de conception* et sont exploités selon des *logiques d'utilisation*. Nous choisissons ce terme *logique* pour référer à une rationalité qui pourrait être revendiquée par les concepteurs ou les utilisateurs dans leurs choix. Nous mettons ce terme au pluriel pour indiquer qu'il y a peu de chance qu'une conception ou une utilisation réponde à une seule « logique ».

Les logiques de conception s'expriment au travers de l'environnement développé par les possibilités d'inter-actions offertes avec des objets informatiques manipulables par l'utilisateur, ces possibilités variant selon les différents états prévus de l'environnement.

Quel que soit le soin apporté au processus de conception, l'utilisateur tentera d'appliquer ses propres « logiques d'utilisation ». Les manifestations de ces logiques d'utilisation se constatent par l'observation. C'est au travers de l'exploitation de « retours d'expérience » qu'elles sont habituellement utilisées pour tenter de réduire le hiatus constaté avec les logiques de conception. Il s'agit alors de mettre en place de nouvelles conceptions intégrant les logiques d'utilisation au mieux en offrant de nouvelles possibilités d'inter-actions avec l'environnement.

L'environnement informatique constitue par lui-même un terrain de confrontation de logiques de conception s'exprimant au travers des possibilités d'inter-action offertes et de logiques d'utilisation s'exprimant par les inter-actions effectivement attendues dans le cadre d'une tâche au sein d'une activité. Cette confrontation est gérée par l'utilisateur pour avancer dans sa propre tâche. Les actions possibles et les états atteignables sont liés aux possibilités issues des logiques de conception ; composer avec ces possibilités relève des logiques d'utilisations. L'idée principale de notre démarche est de fonder un principe d'accompagnement de l'appropriation par une facilitation de l'interprétation de ce qui se passe dans le cours d'action de l'utilisateur pour reprendre l'expression de Theureau [1]. D'un point de vue pratique, il s'agit pour l'utilisateur d'exploiter un dispositif de « réflexion » de ses inter-actions « révélant » d'une certaine manière la confrontation en cours liée au hiatus entre logiques de conception et logiques d'utilisation.

En guise de dispositif de « réflexion », nous proposons de fournir à l'utilisateur une trace des inter-actions telles qu'elles sont observées par un « agent » de traçage. La trace produite doit permettre la représentation des inter-actions et être le support de facilitation de la gestion du hiatus entre logiques de conception et logiques d'utilisation. En conséquence, une telle trace doit être lisible et manipulable symboliquement par l'utilisateur ; celui-ci inter-agira

nécessairement avec elle en tant qu'objet informatique particulier (voir Figure 1 par exemple) qu'il devra donc également s'approprier.

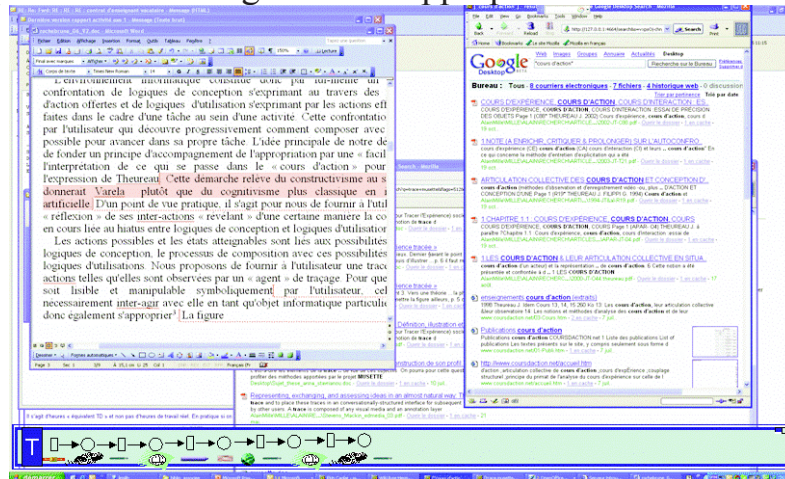


Figure 1. Une fenêtre « trace » s'ajoute dans l'environnement. L'environnement informatique perçu comprend les objets (avec leurs propriétés) permettant les inter-actions. Un objet trace (bas de l'écran) des transitions et des états se succédant sur une ligne temporelle, illustrés par les icônes des objets concernés dans l'inter-action. Les composants de la trace sont explorables et manipulables par l'utilisateur.

L'intégration d'objets traces transforme *de facto* un environnement informatique S en un nouvel environnement informatique S' . L'utilisation du système S' , intégrant la gestion de traces explicites, peut donner lieu à une modélisation de nouvelles utilisations et leur traçage/présentation, pouvant donner lieu à la mise en place d'un système S'' intégrant l'utilisation des traces de S dans S' . Il s'agit donc bien d'une dynamique de l'utilisation liée à la disponibilité de traces gérées par l'utilisateur, et la connaissance de l'utilisateur sur l'utilisation de l'environnement se construit inter-activement en étant symboliquement représentable par l'utilisateur lui-même à des fins de facilitations de ses propres tâches.

3 Modéliser l'utilisation pour pouvoir la tracer

Nous présentons d'abord le principe de modélisation de l'utilisation en introduisant les notions de *modèle d'utilisation (MU)* et de *signatures de tâches expliquées (SiTEx)* avant de détailler une manière de les mettre en œuvre avec l'approche MUSETTE.

3.1 Modèle d'utilisation et signatures de tâches expliquées

Nous appelons *modèle d'utilisation* l'ensemble des éléments qui permettent de produire une trace d'utilisation sur la base des inter-actions avec un environnement informatique. La construction d'un modèle d'utilisation nécessite qu'un modélisateur humain observe des utilisateurs dans leurs pratiques de l'environnement. L'observation « initiale » est réalisée dans le cadre d'un environnement informatique S ne comportant pas d'objets traces. C'est une

observation « ouverte »³, destinée à identifier ce que l'utilisateur prend en considération comme entités et événements lorsqu'il inter-agit avec l'environnement.

Une façon concrète de construire un modèle d'utilisation initial est de mettre en évidence ce que nous avons convenu d'appeler des « signatures de tâches expliquées » (SiTEx). Une SiTEx est un motif inter-actionnel que les utilisateurs re-connaissent dans une trace comme typique d'une tâche qu'ils peuvent alors expliquer (décrire). Les signatures peuvent avoir des « grains » variés, depuis une séquence d'opérations élémentaires jusqu'à une combinaison de possibilités d'inter-actions offertes par différents environnements, considérés comme un seul environnement. La manière d'exprimer les signatures fournit matière à symboliser les processus inter-actionnels pour mieux les manipuler ensuite. Le modèle d'utilisation initial doit permettre de produire ce que nous avons convenu d'appeler une « trace primitive ». Celle-ci constitue le premier niveau considéré comme interprétable par l'utilisateur d'une trace et exploitable par la machine via les SiTEx permettant de repérer des séquences inter-actionnelles potentiellement significatives. Nous verrons dans la section 3 comment ce principe est exploité pour construire des facilitateurs pilotés, voire co-conçus par les utilisateurs.

Le modèle d'utilisation d'un environnement⁴ S est utilisé pour mettre en place l'objet informatique *Trace* qui permet donc de « parler » de ce que l'on fait en inter-agissant avec S . Le contour de l'environnement S a donc évolué en S' (S + la trace) mais il n'est pas possible de « parler » de l'objet *Trace* dans la trace elle-même puisqu'elle est construite sur la modélisation d'utilisation de S .

La dynamique d'utilisation possède deux dimensions :

- Dynamique d'utilisation du système S , ce qui amène à passer de $MU_0(S)$ (MU initial de S) à $MU_1(S)$ (évolution de $MU_0(S)$) et d'une manière générale de raffiner chaque nouveau $MU_j(S)$ en un autre $MU_k(S)$ pour des raisons très diverses (appropriation d'éléments liés aux logiques de conception, construction de nouveaux éléments généralisant, introduction de nouvelles possibilités, *etc.*).
- Dynamique liée à l'intégration de l'objet trace dans le modèle d'utilisation. Cette fois, l'objet *trace* et ses usages étant intégrés dans le modèle d'utilisation, le MU résultant permet de décrire le système S' . Un nouvel objet trace peut alors être créé, correspondant à l'objet trace pour S qui sera décrit dans l'élaboration du $MU_0(S')$. Cette dynamique est d'une nature différente, introduisant une mise en abîme du dispositif. La représentation des traces peut alors être combinée, offrant à l'utilisateur la capacité de donner du sens à l'usage des traces elles-mêmes.

³ Par opposition à une observation « fermée » dans laquelle on ne ferait que vérifier des choses. Il est nécessaire d'observer les situations d'utilisations en tant que telles, sans vouloir les découper *a priori* selon une vision qui ne pourrait venir que de pré-conçus.

⁴ Il faut remarquer que le modèle d'utilisation d'un environnement est une façon de donner le contour *perçu* de l'environnement informatique. Le contour de l'environnement *conçu* n'est en général pas accessible.

3.2 MUNETTE : Modéliser les Utilisations et les Tâches pour Tracer l'Expérience

L'approche Musette [2] a été mise en place dans l'objectif de fournir un cadre pour l'assistance à base de trace au sein de l'équipe CEXAS⁵. La *Figure 2* illustre l'approche MUNETTE, articulante le niveau du traçage (en haut à droite) et celui de la ré-utilisation d'épisodes d'utilisation (à gauche). Le schéma se lit de la façon suivante : un utilisateur inter-agit avec un environnement (système). Un agent observateur, guidé par un modèle d'observation, génère à partir de ces inter-actions une trace primitive respectant un modèle d'utilisation. Un analyseur générique de trace en extrait alors des épisodes, en accord avec des signatures de tâches expliquées. Ces épisodes sont ré-utilisés par des agents assistants, dont l'action peut être soit spécifique avec des agents clairement distincts du système, soit intégrée grâce au système lui-même. Musette est une approche générale de modélisation, et peut être implémentée grâce à divers langages ou formalismes de représentation, à même d'en représenter les différents éléments, que nous précisons dans ce qui suit⁶.

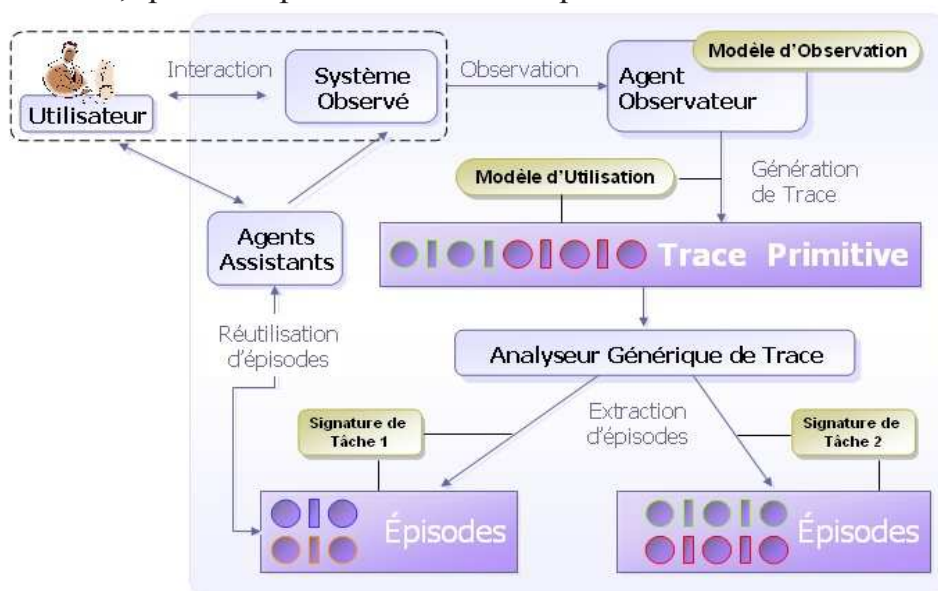


Figure 2. Approche Musette

Il convient de déterminer de quoi la trace sera constituée (modèle d'utilisation), et comment celle-ci sera générée (modèle d'observation). Les éléments d'inter-action constituant de la trace sont des *objets d'intérêt* (OI) de trois types : *entités*, *événements* et *relations*. Les entités sont des objets « présents » pour l'utilisateur dans son inter-action ; les événements des objets qui « ont lieu », qui « se passent », durant cette même inter-action ; les relations peuvent lier indifféremment des entités et/ou des événements. Un modèle d'utilisation décrit les entités, événements et relations à observer pour construire une trace primitive, à la manière d'une « ontologie de l'observation ». Associé

⁵ *Cognition, Expérience et Agents Situés*, équipe de l'axe *Données, document, et connaissances* du LIRIS.

⁶ Les contraintes éditoriales nous contraignant à aller à l'essentiel, nous invitons le lecteur intéressé à consulter les références citées, qui discutent plus précisément de nombreux points ici survolés ou passés sous silence.

au MU, un modèle d'observation décrit les règles nécessaires à la détermination des données pertinentes issues de l'environnement et à la construction effective de la trace primitive. Le modèle d'observation n'est pas spécifié dans l'approche Musette et doit être adapté à chaque environnement observé.

La trace primitive est structurée selon deux types de structures : les états, regroupant des entités, et les transitions regroupant des événements. Les premiers représentent l'état du système observé à un instant/période donné par rapport au modèle d'utilisation. Une transition regroupe un ensemble d'événements, observés entre deux états. La trace est donc *in fine* une séquence d'états et de transitions⁷.

La figure 3 présente un exemple simplifié d'un modèle d'utilisation de navigateur Web, associé à un fragment de trace primitive. Les entités du modèle d'utilisation sont les pages Web (*Page*), les Hyperliens (*Lien*), les images (*Img*) et les préférences (*Pref*). Les événements du modèle d'utilisation sont les clics de l'utilisateur (*Clic*), ses pages sauvegardées localement (*Sauv*), ses bookmarks sur une page (*Bm*) et le changement de langue (*Lang*). Dans la suite de cet exemple, nous considérerons que les pages possèdent également un attribut URL contenant leur adresse.

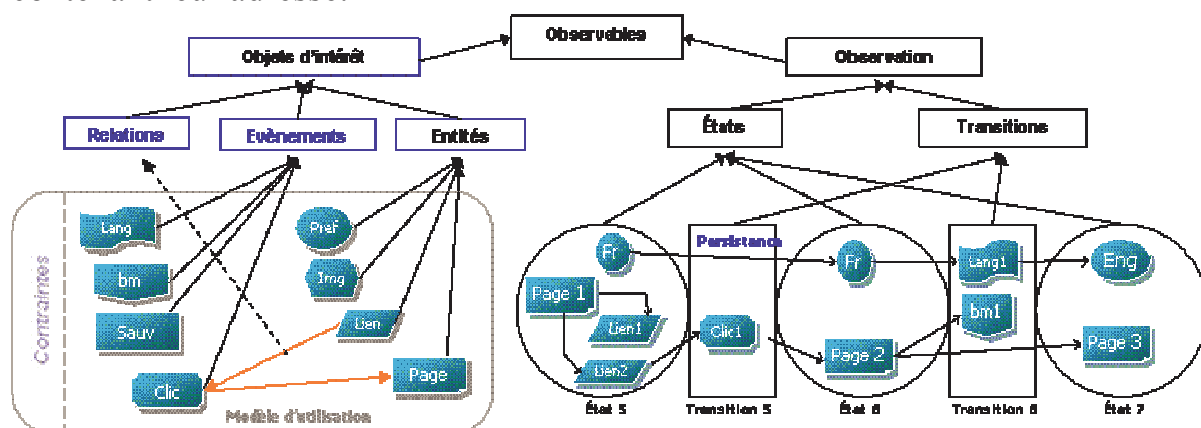


Figure 3. Un modèle d'utilisation simplifié et un fragment de trace de navigation

Pour la lisibilité de la figure, toutes les relations possibles du MU n'y ont pas été reportées. Seules apparaissent les relations d'une entité Lien vers un événement Clic, et celle d'un événement Clic vers une entité Page - signifiant simplement que le lien cliqué mène à la nouvelle page affichée. D'autres relations apparaissent uniquement dans le fragment de trace, mais sont en principe définies dans le modèle d'utilisation : d'une page à un lien (« la page contient le lien »), d'une page à un bookmark (« le bookmark s'applique à cette page »), de page à page (« cette page est rafraîchie »), de préférence à changement de langage, puis à une autre préférence (quand les préférences représentent les choix de l'ancien et du nouveau langage). On peut noter que la transition 6 comporte deux événements, indépendants l'un de l'autre dont la

⁷ Le choix des termes *État* et *Transition* est lié à la volonté de considérer le cours d'action comme un processus.

trace ne donne pas l'ordre d'apparition. La relation Persistance permet quant à elle, dans ce modèle d'utilisation particulier, de traduire le fait que c'est un même objet du point de vue du système, qui se retrouve dans deux états successifs.

Nous appelons épisode une sous-partie de la trace primitive, qui peut y être découpée à l'aide d'une signature de tâche expliquée. Une telle signature s'exprime par : (a) des motifs de graphe constitués d'objets d'intérêt et leurs relations, (b) des contraintes relatives aux positions (distances relatives) des OIs dans la trace, (c) des contraintes sur la structure interne des objets d'intérêt. Les « explications » d'une signature de tâche permettent d'apporter des précisions sur les épisodes qu'elle a détectés : un épisode sera au minimum nommé, éventuellement annoté librement – en vue d'une lecture par l'utilisateur – ou formellement – en vue d'une utilisation automatique.

Nous présentons ci-dessous deux tâches simples de navigation Web qui peuvent être identifiées par notre modèle d'utilisation simplifié, et montrons comment leurs signatures peuvent être expliquées par des annotations sur les objets d'intérêt correspondants. Premier exemple : lorsqu'un utilisateur trouve une page intéressante, il lui arrive souvent de vouloir poser un bookmark sur le site dans lequel il a trouvé la page, considérant qu'il en contient – ou contiendra – d'autres. Ce genre de tâche est aisément repérable dans la trace primitive (fig. 3) : depuis une page (Page) d'un site, l'utilisateur remonte à la page d'accueil et y pose un bookmark (Bm). Des explications sont également fournies dans cette SiTEx : l'évènement bookmark (Bm) est annoté avec une explication textuelle, potentiellement utile pour l'interprétation de l'épisode et les deux pages Web impliquées peuvent être annotées par « page interne » et « page de garde », selon une ontologie des sites Web, définie en dehors de Musette, et utilisable par un assistant dédié.

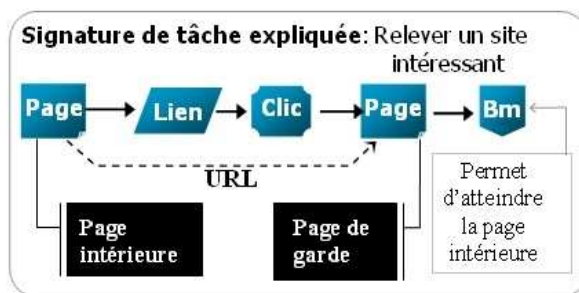


Figure 4. SiTEx et épisode (exemple 1)

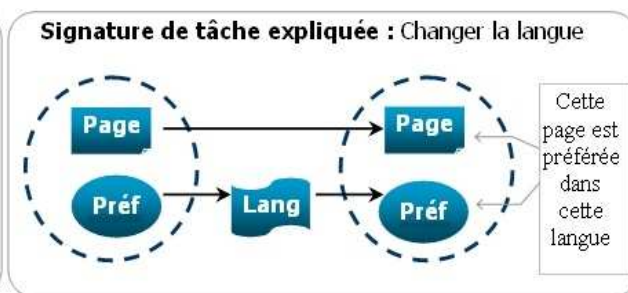


Figure 5. SiTEx et épisode (exemple 2).

Une autre tâche consiste à modifier les préférences de langue pour visualiser différentes versions de pages Web. Le changement étant représenté par un évènement Lang (lorsque la page change sans avoir cliqué sur un lien). La SiTEx de la figure 4, exprime une signature de ce type, les cercles en pointillés représentant une contrainte de co-occurrence des objets d'intérêt Page et Préf (représentant le langage courant de l'utilisation). Dans cet exemple des annotations textuelles donnent des explications quant aux objets d'intérêt concernés.

3.3 *Illustration : construction d'un modèle d'utilisation pour un environnement de veille stratégique*

L'approche Musette a été utilisée [3, 4] pour la modélisation de l'utilisation de l'outil de veille stratégique cartographique Human-Links⁸. Le modèle d'utilisation résultant contient 35 entités, 42 événements et plusieurs dizaines de relations, et sa définition a permis d'éprouver Musette ainsi que de mettre en place des éléments méthodologiques de conception de MU :

- la nécessité de définir précisément l'activité dans laquelle est utilisé l'environnement étudié (en observant les utilisateurs),
- la validation *a priori* des objets d'intérêt par leur appartenance à une SiTEx,
- la nécessité que tous les objets d'intérêt soient intelligibles par les utilisateurs (qui permettront de trancher le cas échéant),
- les compromis rendus nécessaires par le fait de travailler à plusieurs niveaux d'assistance différents (niveau de l'interface, niveau de la tâche de veille), et les manières de mêler des objets d'intérêts de ces différents niveaux, *etc.*

4 Application à la « facilitation » de tâches médiées par les environnements informatiques

4.1 *Principes d'un facilitateur*

Nous avons choisi d'appeler « facilitateur » [5] tout agent informatique exploitant le contenu de la trace d'utilisation pour fournir à l'utilisateur des éléments de compréhension de l'environnement par rapport à ses propres utilisations et la capacité de s'approprier les possibilités d'inter-action offertes dans le cadre de ses utilisations propres. Le facilitateur doit donc « suivre » au plus près le contexte d'utilisation, défini par l'utilisateur lui-même en train d'agir. Une requête de facilitation consiste pour l'utilisateur à fournir, directement ou non, un contexte d'utilisation qui fait sens pour lui et pour lequel il espère trouver des éléments complémentaires élargissant sa compréhension de la situation. Il utilise le mécanisme de signature de tâche expliquée comme support de sa requête. La signature peut être fournie de différentes façons selon le degré d'anticipation de la facilitation :

- *Nouvelle signature de tâche expliquée* : l'utilisateur réalise des inter-actions, exploite la trace pour choisir les éléments qui font sens dans sa tâche courante et utilise les éléments choisis comme constitutifs de la signature de tâche expliquée qui sera sa requête. Le facilitateur recherche dans la trace les séquences « similaires » et les propose comme épisodes candidats en les situant temporellement et par les objets qui étaient alors en relation, ce qui permet à l'utilisateur d'élargir son interprétation et d'agir en conséquence.
- *Signature de tâche expliquée cataloguée* : l'utilisateur choisit un squelette de requête dans une bibliothèque de signatures de tâches expliquées. Celle qui correspond à son utilisation courante (qui sert d'index) et les épisodes correspondants sont retrouvés par l'agent facilitateur comme dans le premier

⁸ <http://www.human-links.com/>

cas. Les écarts constatés entre les épisodes retrouvés et l'épisode en cours de construction sont mis en évidence pour faciliter l'adaptation par l'utilisateur lui-même.

- *Signature de tâche expliquée canonique* : un modélisateur de la trace a conçu (en liaison avec des utilisateurs) des signatures canoniques à respecter, ainsi qu'un agent de type assistant permettant la mise en œuvre d'une utilisation conforme aux signatures en exhibant des épisodes réels qui pourront être adaptés. Cet usage des signatures de tâche expliquée revient à faire du Raisonnement à Partir de Cas à des fins d'assistance. La partie suivante illustre l'utilisation de cette dernière approche pour la réalisation d'un agent d'assistance dans le domaine de l'apprentissage humain, en se plaçant dans le dernier cas.

4.2 Illustration d'un facilitateur : Pixed (Projet d'intégration de l'expérience pour l'enseignement à distance)

4.2.1 Présentation générale

Dans le contexte de l'apprentissage à distance, l'apprenti construit sa connaissance en inter-agissant avec les environnements d'apprentissage, tentant d'exploiter au mieux les activités éducatives disponibles. Apprendre est vu comme mobilisant (entre autres) une capacité à résoudre un problème : « Quel est le meilleur chemin à suivre pour maîtriser cette notion ? ». C'est donc une signature de tâche expliquée correspondant aux opérations d'orientation dans l'hyperespace du cours qui est à la base de l'agent d'assistance réalisé.

Le modèle de conception correspond en partie aux spécifications PAPI⁹ et intègre ce que nous avons appelé [6] un réseau notionnel annoté (RNA). Les noeuds du réseau sont des notions définies comme des buts d'apprentissage et les arcs sont constitués de relations de précédence avec un certain seuil de franchissement et de relations de suffisance possédant un attribut de seuil de maîtrise et un ratio de contribution à la notion cible. PIXED [7-10] permet aux enseignants comme aux apprentis d'annoter tout ou partie des activités éducatives (document hypertextes simples : HTML). Les annotations servent de support de communication entre enseignants, entre enseignants et apprentis et entre apprentis ; elles sont à la base des procédures d'adaptation de présentation de contenus.

4.2.2 Modèle d'utilisation

Dans cette partie, nous mettons en **gras** les éléments du modèle d'utilisation. Les objets d'intérêt sont constitués par les objets avec lesquels l'apprenant interagit durant l'apprentissage. Les **activités pédagogiques**, **documents numériques** et **Quiz** sont associés aux noeuds du **réseau notionnel annoté** (RNA). Le **cours** est construit initialement autour d'un RNA conçu par les enseignants sur des bases didactiques. Chaque apprenti navigue dans le cours par l'intermédiaire du RNA et son avancement est représenté par un sous-graphe de ce RNA qu'il peut donc enrichir en l'annotant avec des notions qui lui sont

⁹ Public and Private Information (PAPI) for learners (IEEE). <http://edutool.com/papi/>

propres et des **activités pédagogiques** qu'il aura exploitées de sa propre initiative. Les activités pédagogiques sont intégrées en tant que documents numériques avec leur description. Les activités pédagogiques sont annotées par les **notions** qu'elles expliquent, illustrent, référencent, etc., en précisant quelles sont les notions utiles, nécessaires, suffisantes, etc., pour les exploiter. Les notions étant reliées entre elles par des **relations de précedence et de suffisance** pour former un **cours** sous la forme d'un RNA particulier, ces informations sont utilisées pour construire automatiquement une première version du réseau initial.

La construction du modèle d'utilisation a été faite empiriquement, et la sémantique des relations (précedence, suffisance) est clairement orientée par des objectifs didactiques, relevant de logiques de conception. Naviguer dans un cours à l'aide d'un réseau notionnel annoté peut se révéler difficile pour l'apprenti qui par définition ne maîtrise pas les notions qui lui sont proposées à l'apprentissage. Il lui est pourtant nécessaire de se construire un chemin parmi les différentes navigations possibles pour atteindre la notion cible de l'apprentissage (au départ, cette notion est la notion cible du cours lui-même). PIXED intègre une façon originale de construire un chemin pertinent sous la forme d'un arbre présentant analogiquement (longueur et position des branches d'accès aux notions) l'effort à fournir pour atteindre la notion cible courante. Les sémantiques des relations de suffisance et de précedence permettent en effet de construire un **arbre de recommandation** pour suivre un **chemin** selon une progression notionnelle didactiquement adaptée.

4.2.3 La trace et son exploitation dans PIXED

Tout choix dans l'orientation du cours (choix d'une activité pédagogique) est tracé (y compris les inter-actions liées à l'évaluation par l'intermédiaire des Quiz). L'exploitation interne d'une activité pédagogique elle-même n'est pas tracée dans cette version de PIXED. Nous nous intéressons à renvoyer à l'apprenti une possibilité d'exploiter des traces d'apprentissages comme sources de connaissances pour assister son orientation dans la progression d'apprentissage dans un cours.

La SiTeX considérée ici est « ce que j'ai fait (objets inter-agis), ce que j'ai appris (notions “réussies”) et ce que je souhaite apprendre (notion à acquérir) » et est utilisée par l'assistant pour suggérer l'adaptation « d'un chemin » notionnel avec des activités pédagogiques alternatives qui, selon l'expérience similaire, pourraient faire sens dans la situation courante.

Quand l'apprenti navigue dans son cours en parcourant le RNA correspondant, le système trace les inter-actions d'apprentissage. Un épisode courant est constitué de la partie de trace démarrante avec la notion ciblée pour le cours (la situation finale souhaitée). Sur la base des mesures de similarités¹⁰ et des différents potentiels de réussite, le système sélectionne des épisodes similaires à la situation courante. A partir de ces épisodes, le système crée un

¹⁰ Mesures de similarité : elles sont construites à partir « d'explications » particulières liées à la signature de tâche expliquée permettant de donner une importance plus ou moins grande aux éléments. Ces explications doivent pouvoir être construites par l'utilisateur.

épisode adapté à la situation en cours, tentant de maximiser le potentiel de l'épisode courant de l'apprenti. Cet épisode permet de fournir à l'apprenti un « chemin d'expérience » (sorte de façon de continuer son épisode) qu'il pourra personnaliser et exploiter. L'apprenti peut en effet naviguer dans ce chemin d'expérience, choisir des notions relais, exploiter les activités éducatives qui ont aidé d'autres apprentis dans la même situation d'apprentissage, consulter des annotations sur ces activités éducatives (posées par d'autres apprentis, ou des enseignants). L'épisode courant est tracé pour être ajouté aux épisodes de la base dès que la notion ciblée est atteinte, tel qu'il aura été vraiment réalisé par l'apprenti.

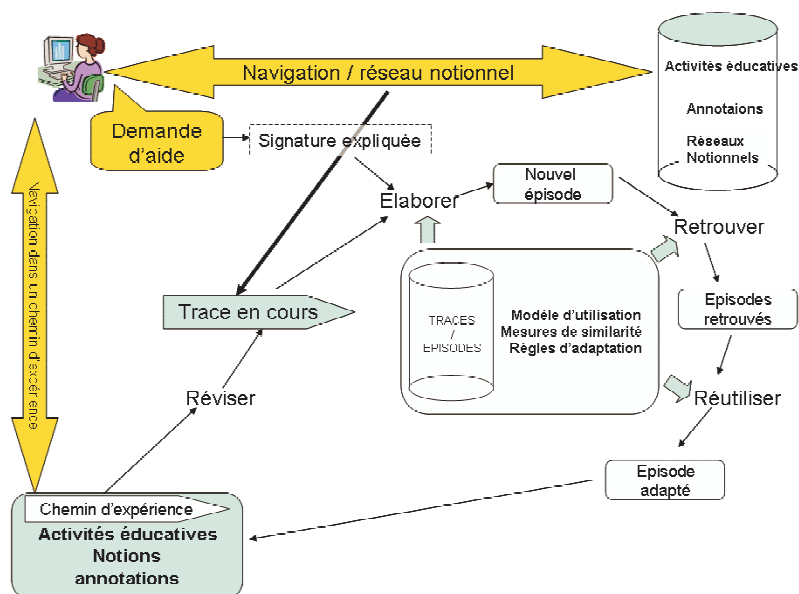


Figure 6. Le cycle Pixed : au cours de la navigation, l'apprenant demande une assistance pour le guider ; la situation d'apprentissage est représentée par une signature de tâche correspondant à « comment continuer l'apprentissage » ; l'épisode en cours correspondant à cette signature est élaboré, et les traces disponibles sont exploitées pour rechercher des épisodes passés similaires candidats ; une proposition de poursuite d'apprentissage est construite sur la base d'un épisode passé sélectionné, adapté et réapproprié par l'apprenant ; la trace en cours inscrit « naturellement » ce nouvel épisode dans la base de traçage

L'adaptation se fonde sur l'ensemble des épisodes ayant le potentiel le plus important pour proposer une fin d'épisode alternative à l'épisode courant. Une liste ordonnée des essais potentiellement les plus intéressants de ces épisodes est construite. Si plusieurs épisodes ont la même notion ciblée, l'essai au plus fort potentiel est gardé.

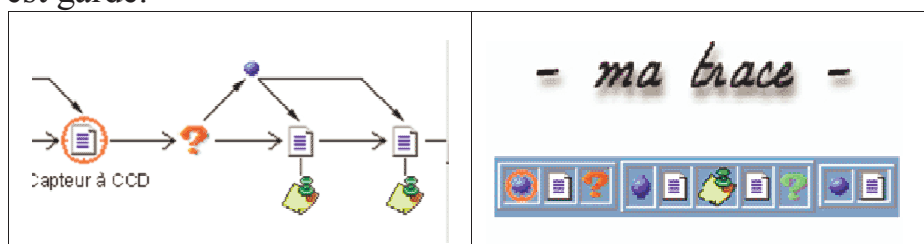


Figure 7. Chemin d'expérience proposé à l'apprenti (figure de gauche) et trace telle qu'elle s'affiche à l'apprenti pendant son travail (figure de droite). Les deux graphiques sont navigables.

5 Discussion

L'usage de traces d'inter-actions pour faciliter l'appropriation d'un environnement informatique s'applique à de nombreuses activités ou situations comme la conception [11, 12] les tâches collaboratives [13], l'apprentissage humain associé à un Environnement Informatique pour l'Apprentissage Humain (EIAH), le partage et la réutilisation d'ontologies [14], *etc.* Dans la mesure où il est possible de générer les traces à partir d'une démarche explicite de modélisation de l'utilisation, l'approche MUSETTE facilite également le développement d'ateliers pour l'analyse de situations d'utilisations d'environnements techniques instrumentés [15] et différents projets nationaux ou régionaux exploitent cette approche actuellement (acteurs des TICE, pratiques de personnes âgées, traces et EIAH).

Deux orientations complémentaires caractérisent notre recherche actuelle : expérimentation sur des terrains variés en développant des ateliers d'analyse et des facilitateurs ; formalisation de l'approche en fondant plus précisément la notion de modèle d'observation (nous faisons intervenir les notions de modèle de collecte, de source de traçage, *etc.*) et le statut des modèles d'utilisations qui se déclinent à partir d'un modèle d'utilisation initial.

De plus, nous considérons la possibilité de concevoir un « système d'exploitation » informatique incluant l'objet trace comme objet natif, disponible pour les concepteurs et les utilisateurs au même titre que l'objet fichier par exemple.

Enfin, nous nous intéressons aux résonances de cette approche réflexive de l'inter-action avec un environnement informatique avec les principes d'ontogenèse étudiés en philosophie et en psychologie.

1. Theureau, J. and G. Filippi, *Cours d'action et conception d'une situation d'aide à la coordination: le cas de la régulation du trafic du RER*. Sociologie du Travail, 1994. XXXVI(4/94): p. 547--562.
2. Champin, P.-A., A. Mille, and Y. Prié, *MUSETTE: Modelling USEs and Tasks for Tracing Experience*, in *Workshop From structured cases to unstructured problem solving episodes - WS 5 of ICCBR'03, Trondheim (NO)*, B. Fuchs and A. Mille, Editors. 2003, NTNU.
3. Laflaquière, J., et al. *Approche de modélisation de l'expérience: utilisation de systèmes complexes pour l'assistance aux tâches de veille informatiquement médiées*. in *ISKO-France 2005*. 2005. Vandoeuvre les Nancy, France: Presses Universitaires de Nancy.
4. Laflaquière, J. and Y. Prié, *Modélisation d'utilisation de système d'assistance à base de trace: une application de Musette à la tâche de veille documentaire*, in *Actes de l'atelier Traces, Interactions, Co-constructions collectives et relations à la Cognition, AS CoMETE, CNRS STIC RTP38, Paris, 19 décembre 2003*. 2003.

5. Laflaquière, J. and A. Ciaccia. *Facilitation de tâches informatiquement médiées: une approche centrée sur la réflexivité de l'utilisation*. in *6ème colloque des jeunes chercheurs en sciences cognitives*. 2005. Bordeaux.
6. Héraud, J.-M., A. Mille, and J.-M. Jolion, *Les réseaux notionnels: un outil pour guider la navigation dans un cours hypermédia*, in *Proceedings 3ème Colloque Jeunes Chercheurs en Sciences Cognitives*. 1999: Soulac, France. p. 116-121.
7. Héraud, J.-M., *PIXED: une approche collaborative de l'expérience et l'expertise pour guider l'adaptation hypermédia en enseignement à distance*, in *Informatique*. 2002, Université Lyon1: Villeurbanne.
8. Héraud, J.-M., *Pixed: improving adaptable navigation with Case-Based Reasoning*, in *EDMEDIA'2004 16th World Conference on Educational Multimedia Hypermedia and Telecommunications*. 2004: Lugano. p. 656-661.
9. Héraud, J.-M., L. France, and A. Mille. *Pixed: An ITS that guides students with the help of learners' interaction log"*. in *Workshop Analyzing Student-Tutor Interaction Logs to Improve Educational Outcomes*. 2004. ITS 2004, Maceio.
10. Héraud, J.-M. and A. Mille, *Pixed: assister l'apprentissage à distance par la réutilisation de l'expérience*, in *Proceedings de l'Atelier Raisonement à Partir de Cas, Plateforme AFIA'03 Laval*, L. Jean, Editor. 2003.
11. Champin, P.-A., *Ardeco: an assistant for experience reuse in Computer Aided Design*, in *From Structured Cases to Unstructured Problem Solving Episodes for Experience-Based Assistance*, A. Mille and B. Fuchs, Editors. 2003: Trondheim, Norway.
12. Philippon, M., G. Caplat, and A. Mille. *Aide à l'utilisateur: savoir quand intervenir*. in *IHM'05: 17ème conférence francophone sur l'interaction homme-machine*. 2005. Toulouse, France.
13. Stuber, A., S. Hassas, and A. Mille. *Languages games for meaning negotiating between human and computer agents*. in *ESAW 2005: Engineering Societies in the Agents' World*. 2005. Kusadasi, Aydin TURKEY: Springer Verlag.
14. Arana-Lozano, J., Y. Prié, and S. Hassas, *MAZETTE: Multi Agent MUSETTE for Sharing and Reusing Ontologies*, in *WOSE: Workshop on Ontologies, Semantics and E-Learning, OTM Workshops*. 2004, Springer. p. 741--752.
15. Georgeon, O., et al., *Driver behaviour modelling and cognitive tools development in order to assess driver situation awareness*. 2005.

SIGNES ET INTERPRÉTATION DES SIGNES DANS LA FICTION POSTMODERNE : LE SCHÉMA DE L'ENQUÊTE DANS *LE NOM DE LA ROSE*

Ilias YOCARIS, IUFM de Nice et UNSA / CTEL
y.ilias@wanadoo.fr

Mots-clefs : indétermination sémiotique et ontologique, herméneutique, schéma de l'enquête, abduction et méta-abduction, ambiguïté, labyrinthe, rapport signe-source, rhizome deleuzien

Le Nom de la rose d'Umberto Eco donne à voir tout un questionnement sur l'interprétation des signes et l'émergence du sens : le récit de l'enquête policière menée par Adso de Melk et Guillaume de Baskerville dans une abbaye médiévale sert de prétexte à l'auteur pour déconstruire tout le système de références philosophiques qui oriente notre conception du fonctionnement sémiotique depuis l'avènement du rationalisme à la fin du 17^{ème} siècle. L'enjeu fondamental de l'intrigue est d'ordre herméneutique : pour comprendre qui est à l'origine des meurtres successifs qui terrorisent le personnel de l'abbaye, Guillaume doit interpréter une série d'« indices » apparemment épars qui s'offrent à lui au fur et à mesure que progresse son enquête. Sa démarche de travail consistera donc à unifier les données dont il dispose en construisant une histoire cohérente qui soit par définition la seule susceptible de donner un sens à tous ces indices, de rendre l'univers fictionnel du roman intégralement « lisible » comme un livre. Toutefois, le succès d'un tel projet n'est pas garanti d'avance ... En effet, la méthode d'investigation mise au point par Guillaume repose sur trois postulats philosophiques : (1) qu'il existe un strict cloisonnement entre « observant » et « observé », l'enquêteur lui-même n'influant pas sur l'« objet » de son enquête ; (2) que le monde « réel » est par définition conforme à nos représentations mentales ; (3) que le fonctionnement sémiotique est déterministe et univoque. Or, c'est l'effondrement de ces trois postulats qui est décrit en définitive dans Le Nom de la rose : le déroulement des faits observés se trouve perturbé par l'acte même de leur observation, l'assassin se conformant progressivement aux schémas « herméneutiques » de l'enquêteur ; l'histoire que Guillaume a élaborée pour rendre compte de ce qui s'est passé ne coïncide pas avec la « réalité » ; enfin, les indices auxquels il s'est trouvé confronté s'avèrent irréductiblement polysémiques et ambigus. Le malheureux Guillaume finit donc par comprendre que sa démarche de réflexion, et, par extension, toute sa vision du monde, sont battues en brèche par une sémiotique et une ontologie indéterministes, typiquement postmodernes. Dès lors, tout le texte apparaît comme une réflexion sur l'éclatement irrémédiable du monde qu'implique ce changement de paradigme.

Le roman d'Umberto Eco *Le Nom de la rose* (1980) reprend un des schémas narratifs de base du roman policier : celui de l'enquête criminelle. Voici, en quelques lignes, l'intrigue.

En 1327, Guillaume de Baskerville, un moine franciscain d'origine Anglaise, arrive dans une abbaye italienne des Alpes chargé d'une mission diplomatique. Guillaume est accompagné d'un novice Allemand, Adso de Melk, qui est le narrateur du roman. Dès son arrivée, il se voit prié par l'abbé d'enquêter sur la mort d'un de ses moines, dont on a retrouvé le cadavre en bas des murailles du monastère. Aidé par son novice, Guillaume se met à l'ouvrage, alors que les morts de moines se succèdent. Très vite, l'enquête est orientée vers les activités de toutes sortes qui ont lieu nuitamment dans la bibliothèque de l'abbaye : le mystère est d'autant plus épais que celle-ci est un vaste labyrinthe théoriquement interdit d'accès à tous les moines excepté l'abbé et le bibliothécaire. Toujours flanqué d'Adso, Guillaume pénètre par effraction dans le labyrinthe et apprend peu à peu à s'y orienter. Dès lors, les découvertes se succèdent : Guillaume et Adso finissent par comprendre qu'une lutte souterraine est engagée entre plusieurs moines pour récupérer le dernier exemplaire encore existant d'un livre précieux, le deuxième volume de la *Poétique* d'Aristote. Jugeant le livre impie parce qu'il traite de la comédie, l'ancien bibliothécaire de l'abbaye, Jorge de Burgos, en a empoisonné les pages, de sorte que tous les lecteurs qui essaient de les tourner en les humectant des doigts meurent sur-le-champ. Lors de leur affrontement final avec Jorge, Guillaume et Adso provoquent involontairement un incendie qui détruit l'abbaye de fond en comble. Jorge périt dans l'incendie, non sans avoir auparavant profité de l'occasion pour brûler le livre d'Aristote : son contenu restera donc à jamais inconnu à la postérité. L'enquête étant terminée et le désastre consommé, Guillaume et Adso n'ont plus qu'à reprendre la route.

Comme on le voit, Eco s'inscrit délibérément dans la lignée du roman policier « classique », tel qu'il a été codifié dès le 19^{ème} siècle par des auteurs comme Edgar Allan Poe ou Arthur Conan Doyle, le créateur du personnage de Sherlock Holmes. Cette continuité est du reste confirmée par les références intertextuelles implicites qui relient *Le Nom de la rose* à l'œuvre de Conan Doyle : (1) le nom du moine-détective (Guillaume de Baskerville) n'est pas choisi au hasard, puisqu'il renvoie à un des romans les plus célèbres de la série des *Sherlock Holmes*, *The Hound of the Baskervilles* (*Le Chien des Baskerville*) ; (2) le couple Guillaume / Adso (le génial enquêteur et son naïf mais fidèle adjoint) reproduit évidemment le couple Holmes / Watson, ce qui est discrètement signalé au lecteur par le choix du nom propre « Adso » (un quasi-paronyme¹ de « Watson » ...). Il peut sembler à première vue étonnant qu'un sémioticien très brillant

¹ Le paronyme d'un mot donné est un autre mot dont le signifiant est le même à *un ou deux phonèmes près* : ainsi par exemple « classe » [klas] est le paronyme de « crasse » [kras], et vice versa.

comme Eco s'intéresse à ce point au roman policier, un genre paralittéraire généralement considéré comme simpliste ... Toutefois, à bien y regarder, ce choix est parfaitement logique : dans *Les Limites de l'interprétation*, ouvrage théorique de 1990, Eco rappelle que le schéma de l'enquête criminelle pose une série de problèmes purement sémiotiques, portant à la fois sur la nature des « indices » (au sens le plus large du terme) qui s'offrent à l'enquêteur et sur la manière dont il doit rassembler ces « indices » pour faire émerger *un sens*, autrement dit – en l'occurrence – une histoire cohérente : « la séquence d'événements sur laquelle enquête le détective est définissable comme un texte » (Eco 1992 : 261). Or, *Le Nom de la rose* pose une question fondamentale : qu'est-ce qui garantit *a priori* la « lisibilité » et la cohérence de ce texte ? Comme nous allons essayer de le montrer, les personnages (et, bien entendu, l'auteur lui-même) constatent qu'il n'y a aucune réponse à cette question, et cette constatation remet en cause tout le système de références qui oriente notre conception du fonctionnement sémiotique depuis l'avènement du rationalisme philosophique à la fin du 17^{ème} siècle.



En écrivant *Le Nom de la rose*, Eco entend déconstruire à la fois les schémas narratifs du roman policier du 19^{ème} siècle et les postulats philosophiques et ontologiques qui se trouvent à l'origine de ces schémas. Le déroulement du roman policier « classique » est immuable : un crime est commis par un assassin inconnu, et le détective enquête *a posteriori* afin d'identifier le coupable en reconstituant l'enchaînement des faits qui ont abouti à la mort de la victime. Or, comme on prend soin de l'expliquer au profane dans *Les Limites de l'interprétation*, une telle démarche peut être apparentée à deux processus sémiotiques : l'abduction et la méta-abduction.

- (a) **L'abduction.** L'abduction peut être décrite de manière synthétique comme « la *reductio ad unum* d'une pluralité » (Eco 1992 : 262) : confronté à une série décousue de faits (*x, y, z ...*) apparemment inexplicables, l'enquêteur tente d'unifier cette série en trouvant une explication commune qui permet de tout englober dans « une séquence cohérente » (Eco 1992 : 272). C'est ce qui se passe par exemple dans la fameuse nouvelle de Poe « *The Murders in the rue Morgue* » (« Double assassinat dans la rue Morgue »), considérée comme le premier récit policier de l'histoire : en enquêtant sur le meurtre de Mme L'Espanaye et de sa fille, le protagoniste de la nouvelle (C. Auguste Dupin) relève une série d'indices épars (force physique et agilité hors du commun de l'assassin, cris inarticulés entendus par plusieurs témoins, absence de motifs plausibles) à partir desquels il élabore une « hypothèse de travail » qui permet de tout expliquer en même temps : le crime n'aurait pas été commis par un homme, mais ... par un singe.

- (b) **La méta-abduction.** La méta-abduction est définie par Eco comme la confrontation du « monde possible » construit par l'enquêteur avec le « monde actuel » dans lequel il évolue, afin de décider « si l'univers possible déterminé par [ses] abductions de premier niveau est identique à l'univers de [son] expérience » (Eco 1992 : 264). Ainsi, dans la nouvelle de Poe, Dupin finit par retrouver le propriétaire du singe qui a tué les deux femmes (un matelot parisien), et celui-ci confirme le bien-fondé de son hypothèse initiale.

Bien entendu, un tel *modus operandi* a des implications qui dépassent largement le cadre de l'enquête criminelle. On notera ainsi que la méthode de travail élaborée par Dupin et reprise pratiquement telle quelle par ses épigones fictionnels (Sherlock Holmes, Rouletabille, Hercule Poirot ...) se veut « scientifique » : en effet, comme Eco ne manque pas de le souligner, l'usage de l'abduction et de la méta-abduction est monnaie courante dans les sciences de la nature, notamment quand il s'agit d'élaborer une théorie nouvelle à partir de résultats expérimentaux incompatibles avec le paradigme scientifique en vigueur (Eco 1992 : 264 ; cf. Kuhn 1962 : chap. V, VI).

C'est l'ensemble de ce mécanisme sémiotique qui se trouve en quelque sorte « démonté » dans *Le Nom de la rose* : les références de Eco au roman policier du 19^{ème} siècle lui permettent de mettre à nu très précisément les présupposés conceptuels qui rendent possible, à partir d'un « divers » de données factuelles *a priori* irréductibles à toute unification, l'émergence d'un *sens* cohérent. En effet, une lecture attentive des textes de Poe et de Conan Doyle (mais aussi de leurs « épigones » francophones Leroux et Leblanc) montre que ceux-ci reposent sur le postulat de la « lisibilité » absolue du monde qui s'offre au regard de l'enquêteur, d'où un « pansémiotisme » (Gelly 2005 : 105) très frappant qui est souligné dans « The Murders in the rue Morgue » : « tout est pour lui [*l'enquêteur, assimilé à un joueur de whist*] symptôme, diagnostic, tout rend compte à cette perception, intuitive en apparence, du véritable état des choses » (Poe 1989 : 519 ; nous utilisons bien entendu la traduction de Charles Baudelaire). Le « pansémiotisme » revendiqué par Poe consiste à affirmer en substance que l'univers se réduit à un texte cohérent, qu'il suffit à l'homme de déchiffrer : on retrouve là le *topos*² du « grand livre du monde ». Or, comme par hasard, ce *topos* est repris par Guillaume de Baskerville au début de son enquête :

— Mon bon Adso, dit le maître. J'ai passé tout notre voyage à t'apprendre à reconnaître les traces par lesquelles le monde nous parle comme un grand livre. Alain de Lille disait que

² On peut définir sommairement le *topos* (du grec τόπος, « lieu », et par extension « lieu commun ») comme un thème littéraire récurrent qui apparaît dans un grand nombre de textes. Pour une analyse plus approfondie, v. (Rastier 2001 : 216-226).

omnis mundi creatura
quasi liber et pictura
nobis est in speculum³

et il pensait à l'inépuisable réserve de symboles avec quoi Dieu, à travers ses créatures, nous parle de la vie éternelle. Mais l'univers est encore plus loquace que ne le pensait Alain, et non seulement il parle des choses dernières [...] mais aussi des choses proches, et alors là d'une façon lumineuse. (NR, 31-32)

L'idée que le monde est « lisible » comme un texte découle à son tour d'une conjonction de trois principes épistémologiques, directement hérités de la philosophie rationaliste du 17^{ème} et du 18^{ème} siècle (Descartes, Spinoza, Leibniz, Kant) :

(1) Le cloisonnement strict entre « observant » et « observé », qui garantit qu'il n'y a aucune interaction entre le « processus d'observation » (l'enquête à proprement parler) et les « faits » observés (le déroulement de l'histoire que le détective est chargé d'élucider). Ce cloisonnement relève évidemment du « partage dualiste » instauré par Kant et Descartes entre « sujet » et « objet », entre *res cogitans* et *res extensa* (cf. Chevalley 1995 : 161-187). Il correspond, dans le domaine des sciences de la nature, au concept de l'« objectivité forte », autrement dit à la croyance que la « réalité » possède une existence autonome (le « véritable état des choses » dont parle Poe) et des caractéristiques intrinsèques⁴.

(2) L'adéquation du monde à nos représentations, qui garantit le succès du processus méta-abductif quoi qu'il arrive (il est de notoriété commune que les abductions d'Hercule Poirot ou de Sherlock Holmes ne sont *jamais* démenties par l'expérience). Un tel précepte semble « aller de soi », dans la mesure où l'on admet implicitement que *Ordo et connexio idearum idem est ac ordo et connexio rerum*⁵, autrement dit que Dieu « a sculpté en notre âme une faculté de pensée pouvant opérer en accord avec les lois de la nature » (Eco 1992 : 281). On est tenté de le rapprocher de ce que l'on appelle, toujours dans le domaine des sciences de la nature, le « réalisme mathématique » (d'Espagnat 1985 : 24 *et passim*, 1994 : 356), à savoir l'idée que le formalisme mathématique qui fonde les théories physiques n'est pas une simple construction de l'esprit humain, mais renvoie à des vérités intemporelles « dont l'existence est tout à fait indépendante des activités des mathématiciens » (Penrose 1992 : 103 ; cf. 467-469).

³ « Toute créature de l'univers / Est pour nous comme un livre et une peinture / Dans un miroir » [extrait du poème d'Alain de Lille (1128-1202) « Rhythmus de incarnatione Christi » (Patrologie Latine CCX, 579)].

⁴ Cf. d'Espagnat 1985 : 31, n. 1 : « j'appelle "objectivité forte" le fait, pour un énoncé, de ne comporter aucune référence, fut-elle [*sic*] implicite, à la collectivité des "observateurs" » ; v. aussi d'Espagnat 1994 : 34-35.

⁵ « L'ordre et l'enchaînement des idées est le même que l'ordre et l'enchaînement des choses » : v. Baruch de Spinoza, *Éthique*, II, prop. 7 ; cf. Eco 1992 : 281.

(3) Une conception déterministe et univoque du processus sémiotique, qui permet d'éliminer toutes les ambiguïtés en postulant qu'il existe un rapport fixe et bi-univoque entre le signe et sa « source », que chaque *significans* renvoie à un – et un seul – *significatum* qui peut être déterminé de manière précise. Ainsi par exemple, quand le Rouletabille de Gaston Leroux lit par hasard (dans *Le Mystère de la chambre jaune*) une petite annonce signée M.A.T.H.S.N., il en déduit, « sans autre forme de procès », que ces initiales doivent renvoyer *forcément* au nom « Mathilde Stangerson » (le nom de la femme agressée dans la « chambre jaune ») : toute autre possibilité est d'emblée exclue.

Or, c'est l'effondrement de ces trois principes qui est décrit en définitive dans Le Nom de la rose : l'échec de la démarche « pansémiotique » qui caractérise le roman policier « classique » apparaît on ne peut plus clairement si l'on se penche sur la progression de l'enquête menée par Guillaume de Baskerville. Ce dernier est bel et bien un adepte des méthodes rationalistes de réflexion et de travail qui ont fait la gloire de Dupin, de Rouletabille et de Sherlock Holmes, et qui lui valent une certaine notoriété (cf. notamment *NR*, 37). En effet, au début du roman, il réussit un « coup de maître » sémiotique : confronté à un ensemble d'« indices » épars qui se présentent à lui alors qu'il se rapproche de l'abbaye avec Adso (des empreintes de sabots sur le chemin enneigé, des pins dont les branches ont été cassées à la hauteur de cinq pieds, des crins de jais suspendus aux épines d'un buisson de mûres, des moines qui croisent Adso et Guillaume en sens inverse ...), il parvient à reconstituer à partir de ces indices une histoire cohérente : les moines qu'il a rencontrés sont à la poursuite d'un cheval noir de cinq pieds de hauteur. Bien entendu, Guillaume, qui n'a jamais vu ce cheval, est capable de le décrire de manière très précise en utilisant un système complexe d'inférences sémantico-pragmatiques, et son abduction est confirmée dans un deuxième temps par les faits (*NR*, 30-32)⁶. Mais ce triomphe initial ne confirme pas pour autant la validité épistémologique de sa méthode d'investigation *en toutes circonstances* : malgré les succès « locaux » qu'il obtiendra dans la suite du récit grâce à sa virtuosité abductive, Guillaume finit par comprendre que sa démarche de réflexion, et, par extension, toute sa vision du monde, sont battues en brèche par une sémiotique et une ontologie indéterministes, typiquement postmodernes. Dans les dernières pages, le génial enquêteur découvre ainsi – à son grand désespoir – que les postulats initiaux qui fondaient sa méthode « pansémiotique » s'avèrent erronés – ou plutôt inopérants :

(a) Le cloisonnement entre « observant » et « observé » tend à être éliminé, puisque, en l'occurrence, le « processus d'observation » mis en œuvre par l'enquêteur influe directement sur l'« objet » de l'enquête. Qu'est-ce à dire ? Pour expliquer l'enchaînement des morts qui scandent son séjour dans l'abbaye,

⁶ L'épisode du cheval est un « emprunt » intertextuel délibéré : Eco ne fait que reprendre, en la transposant dans l'Italie du 14^{ème} siècle, une scène de *Zadig*, le fameux conte philosophique voltairien (cf. Voltaire 1954 : 10-12).

Guillaume fabrique de toutes pièces un schéma abductif, en postulant qu'il se trouve confronté à une série de meurtres commis par un assassin qui suit « le rythme des sept trompettes de l'Apocalypse » (NR, 475 ; cf. NR 310-311, 371 etc.). Ce schéma est tout à fait erroné, dans la mesure où les trois propositions de base formulées par Guillaume sont fausses toutes les trois (1. tous les moines n'ont pas été assassinés ; 2. il n'y a pas un seul assassin ; 3. à l'origine, les morts successives qui terrorisent le personnel de l'abbaye n'ont rien à voir avec le texte de l'Apocalypse) ; toutefois, par un effet de « prophétie autoréalisatrice », l'assassin « principal » (Jorge de Burgos) finit par s'y conformer « réellement » : afin de confondre Guillaume, il s'arrange pour que son dernier meurtre (celui de Malachie, le bibliothécaire⁷) soit commis en fonction du texte de l'Apocalypse, alors que la similitude entre les morts précédentes et le récit de Jean n'avait été « en réalité » qu'un hasard. De la sorte, le mécanisme de l'enquête se trouve déréglé, parce que l'« observé » est perturbé par l'« observant » : Guillaume comprend qu'il n'a pas découvert « le véritable état des choses » mais qu'il a fini par créer performativement une nouvelle « vérité », au fur et à mesure que progressait son enquête. Un schéma narratif de ce genre, qui rappelle étrangement certaines interprétations de la mécanique quantique avancées à la fin des années '20⁸, apparaît (tel quel ou avec des variantes) dans bon nombre

⁷ Le meurtre de Malachie est plus précisément le dernier commis par Jorge avant d'être découvert : une fois son secret percé, le pugnace vieillard, peu désireux de rendre des comptes à l'abbé, se débrouille pour le supprimer à son tour au moment précis où Guillaume, qui a compris ce qui s'est passé, se lance à sa recherche avec Adso.

⁸ Un des problèmes fondamentaux posés par la mécanique quantique est la description exacte de l'« interaction » qui se développe entre l'« objet » (les particules subatomiques, qui ne sont pas réductibles à une représentation spatio-temporelle unitaire) et le dispositif expérimental mis en place pour l'« observer ». L'interprétation dite « de Copenhague », mise en place progressivement par Bohr et Heisenberg à partir de 1927, a permis de clarifier peu à peu cette question, qui reste toutefois l'une des plus débattues dans le champ épistémologique à l'heure actuelle (v. notamment Bitbol 1996, 1998). En quoi consiste le problème ? Dès les premiers pas de la mécanique quantique, Bohr et Heisenberg se rendent compte qu'ils ont affaire à des « entités » (les particules subatomiques) qui se comportent différemment en fonction du système expérimental mis en place pour les observer : dans certains types d'expériences, ces « entités » se comportent comme des ondes, dans d'autres elles réagissent plutôt comme des corpuscules. Formalisée sur le plan épistémologique par la fameuse notion bohrienne de « complémentarité », cette apparente dualité comportementale ne laisse pas d'intriguer les physiciens de l'époque : comment faut-il comprendre la corrélation entre modification du dispositif observationnel et modification du comportement de l'« objet » observé ? Une première réponse apportée par Heisenberg et par Bohr à la fin des années '20 est que le fonctionnement du dispositif observationnel n'est pas neutre, mais crée toujours une perturbation irrémédiable et incontrôlable du phénomène physique observé. Or, il se trouve (pensent au départ Bohr et Heisenberg) que l'étendue de cette perturbation ne peut jamais être déterminée, car « tenter de l'évaluer amorcerait une régression *ad infinitum* d'interactions elles-mêmes finies » (Bitbol 1996 : 246), comportant à chaque fois l'échange d'une certaine quantité d'énergie (le fameux « quantum d'action »). Ainsi « [u]n appareil "déterminant" [utilisé pour faire une mesure qui départagerait ce qui appartient « en propre » au

d'œuvres de fiction postmodernes retraçant le déroulement d'une enquête, comme *Les Gommages* d'Alain Robbe-Grillet (1953) ou *The Element of crime* de Lars Von Trier (1984) : dans ces deux récits (textuel et filmique), la « ligne de démarcation » entre « sujet observant » et « objet observé » disparaît, puisque l'enquêteur finit par commettre lui-même le crime qu'il était chargé d'élucider (dans *Les Gommages*) ou de prévenir (dans *The Element of crime*) ...

(b) L'adéquation du monde « extérieur » à nos représentations s'avère, au mieux, aléatoire, et ce parce qu'**aucune instance métaphysique n'est plus là pour garantir cette adéquation**. Ce point est très important, dans la mesure où les personnages du *Nom de la rose* sont – tout comme Eco lui-même au début de son parcours conceptuel – profondément imprégnés par la vision du monde et les valeurs philosophiques chrétiennes. Constatant l'échec de sa propre démarche méta-abductive (puisque l'« histoire » qu'il a élaborée ne correspond pas à ce qui s'est « réellement » passé), Guillaume aboutit à la conclusion qu'« il n'est point d'ordre dans l'univers » (NR, 497), que ce dernier n'est donc pas réductible « à l'ordre que notre esprit imagine » (*ibid.*)⁹, et que Dieu ne saurait

« phénomène » et ce qui appartient à l'interaction de ce dernier avec le dispositif expérimental, I. Y.] d'ordre n devrait interagir (de façon nécessairement perturbante) avec l'appareil "perturbant" d'ordre n-1, qui a lui-même interagi avec un objet (ou un appareil-objet) "perturbé" d'ordre n-2 » (Bitbol 1996 : 246-247), et ainsi de suite. Une telle réponse fait désormais partie du corpus de représentations « vulgarisées » des processus subatomiques, notamment parce qu'elle permet de penser le problème de l'interaction entre « observant » et « observé » dans le cadre d'un schéma traditionnel du type dualiste, axé sur une opposition entre deux pôles « clairs et distincts » : d'une part l'instance observatrice, d'autre part le « phénomène » qu'elle « perturbe » par l'acte même de son observation. On notera toutefois que sa pertinence fut remise en cause dès 1935 par Einstein, Podolsky et Rosen : ces derniers démontrèrent que, si l'on s'en tenait à une telle interprétation de la mécanique quantique (qui était pour eux la seule possible), celle-ci ne pouvait offrir du réel qu'une description incomplète (v. Bitbol 1996 : 249-251, Whitaker 1996 : 224-234). En réponse au défi théorique ainsi lancé, Bohr dut radicaliser sa position, pour aboutir à une définition purement holistique de l'« objet » quantique : selon cette définition, la corrélation entre modification du dispositif observationnel et modification du comportement de l'« objet » observé ne résulte pas d'une « interaction mécanique directe » (N. Bohr, cité in Bitbol 1996 : 251), donc d'une perturbation, mais du fait que les « entités » observées ne possèdent pas de déterminations intrinsèques et sont de ce fait purement « contextuelles », autrement dit forment une totalité indivisible avec leur contexte observationnel. Une telle vision de la mécanique quantique, philosophiquement formalisée par le concept d'« objectivité faible » (v. d'Espagnat 1985 : 31, 169-180 *et passim*), semble à l'heure actuelle la plus adaptée pour rendre compte des particularités du « réel » microphysique : on pense notamment aux expériences d'Aspect et de Zeilinger, qui, en mettant au jour des phénomènes d'« intrication quantique » (*quantum entanglement*), ont permis de confirmer la description holistique des phénomènes subatomiques proposée par Bohr à partir de sa seule intuition (v. Zeilinger 1997@, Bouwmeester *et al.* 1997).

⁹ Au vu du déroulement de l'intrigue, une telle conclusion pourrait sembler trop radicale : après tout, la démarche (méta-)abductive de Guillaume n'est pas un échec intégral, puisqu'il a eu raison de supposer que les morts sur lesquelles il enquête découlent toutes (directement ou indirectement) de la lutte pour la possession d'un livre (correctement identifié comme étant le

exister dans la confusion (d'où ses dernières paroles : « *Non in commotione, non in commotione Dominus* », NR, 498). Bref, le malheureux enquêteur découvre qu'il n'évolue pas dans un *cosmos* fondé sur les notions d'Identité et de Récognition mais dans un *chaosmos* miné par les forces de l'Altérité et de la Différence¹⁰, ce qui en fait un personnage plus deleuzien que cartésien.

(c) Une part incompressible d'ambiguïté et d'indétermination apparaît dans le processus sémiotique. En effet, le schéma abductif créé par Guillaume lui a certes permis de construire une « histoire » plus ou moins cohérente et de donner un sens à ce qui s'est passé dans l'abbaye, mais le « prix » épistémologique qu'il a dû payer pour cela est très élevé : en choisissant, après mûre réflexion, de rattacher chacune des morts sur lesquelles il enquête à une des trompettes de l'Apocalypse, Guillaume les traite comme autant de signes qui renvoient chacun à une, et une seule, « source » – le texte de Jean, ou plutôt le projet meurtrier élaboré par l'assassin présumé à partir de ce texte. Or, force est de constater à la fin du livre que chaque mort renvoyait « en réalité » à un enchaînement d'événements et de facteurs différents que Guillaume, obnubilé par son propre schéma interprétatif, n'a pu prendre en considération. Dès lors, un abîme sémiotique s'ouvre sous les pieds du personnage : en effet, il commence à réaliser que chaque *significans*, quel qu'il soit, peut être rattaché à une infinité de *significata*, et qu'il n'existe au fond aucun moyen épistémologique de déterminer la « bonne » connexion – à supposer même que celle-ci existe, parce que les circonstances de certaines morts étaient « en réalité » le fruit du hasard ... Eco résume les impressions de son héros en soulignant qu'il se trouve brutalement projeté dans un monde « déjà structuré en rhizome » (Eco 1985 : 65), le « rhizome » deleuzien étant l'image parfaite d'un labyrinthe sémiotique où chaque point peut être virtuellement connecté à tous les autres (cf. Deleuze et Guattari 1980 : 9-37).

Le résultat conjoint de (a), (b) et (c), c'est que le monde devient irrémédiablement fragmenté et illisible. Loin de se présenter comme un texte cohérent, comme un « grand livre » dans lequel on peut lire *ad libitum*, il ressemble plutôt à la « bibliothèque » reconstituée par Adso à partir de vestiges épars ramassés dans les ruines de l'abbaye brûlée : cette bibliothèque (en grande partie) imaginaire est « composée de morceaux, citations, périodes incomplètes, moignons de livres » (NR, 504). Confronté à la faillite de la démarche

deuxième volume de la *Poétique*) aux pages empoisonnées. Si Guillaume éprouve un sentiment de frustration cognitive, ce n'est pas parce qu'il est submergé par le désordre, mais parce qu'il ne réussit à imposer au réel qu'un ordre *partiel* : il a bien repéré la « cause première » qui est à l'origine de toutes les morts, mais il n'a pas compris que celles-ci se sont enchaînées de manière plus ou moins fortuite, et non point comme des manifestations nécessaires d'un Grand Dessein initial, analogue à la Providence Divine [cf. *infra*, point (c) ; NR, 475].

¹⁰ Le concept deleuzien de *chaosmos* a été forgé pour penser « l'identité interne du monde [cosmos *en grec*, I. Y.] et du chaos » (Deleuze 1968 : 382).

« herméneutique » totalisante de son maître, Adso se focalise sur une série incohérente de *membra disjecta*¹¹ textuels à partir desquels il tente péniblement de dégager un sens, fût-ce de manière « locale ». Mais une telle démarche influe inévitablement sur la perception qu'il a lui-même de son récit : ce dernier perd peu à peu à ses yeux toute signification, et le personnage finit par se réfugier dans un nominalisme au fond assez nihiliste. Les derniers mots du texte (« *nomina nuda tenemus* », « nous tenons des noms nus », NR, 505) montrent que, pour lui, la seule réalité, le seul référent « palpable », c'est *in fine* le signe verbal lui-même.



Le Nom de la rose se présente donc en définitive, non pas seulement comme une réflexion (poursuivie dans tous les romans ultérieurs de Eco, et notamment *Le Pendule de Foucault*) sur le processus d'émergence du sens, mais aussi, plus largement, comme un témoignage historique « indirect » sur la *Weltanschauung* de toute une époque : l'impossibilité de réduire le monde dans sa totalité à un *topic* textuel unitaire immédiatement déchiffrable est un drame épistémologique qui affecte non seulement la sémiotique et la philosophie du langage, mais aussi la philosophie politique (avec la fin des *grands récits* lyotardiens), la physique (avec le « partage des territoires » entre relativité générale et mécanique quantique), les mathématiques (avec le théorème de Gödel), la sociologie des sciences (avec l'émergence de notions méta-théoriques comme l'incommensurabilité ou le conflit entre paradigmes) etc. Le nominalisme autoréflexif typiquement derridien qui se fait jour à la fin du *Nom de la rose* montre que le roman postmoderne glorifie bien la figure de Narcisse, mais d'« un Narcisse aveugle, en quête de ses membres épars, et irrémédiablement voué à la désagrégation » (Dällenbach 1977 : 211).

Ilias YOCARIS
IUFM de Nice et UNSA / CTEL

¹¹ « Membres épars ».

BIBLIOGRAPHIE

- Michel BITBOL (1996) : *Mécanique quantique. Une introduction philosophique*, Paris, Flammarion, coll. « Nouvelle Bibliothèque Scientifique ».
- (1998) : *L'Aveuglante proximité du réel*, Paris, Flammarion, coll. « Champs ».
- Dik BOUWMEESTER *et al.*, « Experimental quantum teleportation », *Nature*, 390, 11 décembre 1997, p. 575-579.
- Catherine CHEVALLEY (1995) : « Nature et loi dans la philosophie moderne », in *Notions de philosophie*, Denis Kambouchner éd., Paris, Gallimard, coll. « Folio Essais », pp. 127-230.
- Lucien DÄLLENBACH (1977) : *Le Récit spéculaire. Essai sur la mise en abyme*, Paris, Seuil, coll. « Poétique ».
- Gilles DELEUZE (1968) : *Différence et répétition*, Paris, Presses Universitaires de France, coll. « Épiméthée ».
- (1969) : *Logique du sens*, Paris, Minuit, coll. « Critique ».
- Gilles DELEUZE et Félix GUATTARI (1980) : *Capitalisme et schizophrénie*, tome 2, *Mille plateaux*, Paris, Minuit, coll. « Critique ».
- Bernard D'ESPAGNAT (1985) : *Une Incertaine réalité. Le monde quantique, la connaissance et la durée*, Paris, Gauthier-Villars.
- (1994) : *Le Réel Voilé. Analyse des concepts quantiques*, Paris, Fayard, coll. « Le Temps des sciences ».
- Umberto ECO (1982 [¹1980]) : *Le Nom de la rose [NR]*, trad. de l'italien par Jean-Noël Schifano, Paris, Grasset/Fasquelle.
- (1985 [¹1983]) : *Apostille au Nom de la rose*, trad. de de l'italien par Myriem Bouzaher, Paris, Grasset, coll. « Le Livre de Poche/Biblio Essais ».
- (1992 [¹1990]) : *Les Limites de l'interprétation*, trad. de l'italien par Myriem Bouzaher, Paris, Grasset/Fasquelle, coll. « Le Livre de Poche/Biblio Essais ».
- Cristina FARRONATO (1999) : « The Theory of abduction and *The Name of the Rose* », in *Semiotics 1998*, C. W. Spinks and John Deely édés, Bern, Peter Lang Publishing, pp. 71-81.
- (2003) : *Eco's chaosmos : from the Middle Ages to postmodernity*, Toronto, University of Toronto Press.
- Christophe GELLY (2005) : *Le Chien des Baskerville. Poétique du roman policier chez Conan Doyle*, Lyon, Presses Universitaires de Lyon, coll. « Champ Anglophone ».
- Jean-Marie KLINKENBERG (1996) : *Précis de sémiotique générale*, Paris, Seuil/De Boeck, coll. « Points Essais ».
- Thomas S. KUHN (1962) : *The Structure of scientific revolutions*, Chicago, The University of Chicago Press.

- Jean-François LYOTARD (1979) : *La Condition postmoderne : rapport sur le savoir*, Paris, Minuit, coll. « Critique ».
- Roger PENROSE (1992 [¹1989]) : *L'Esprit, l'ordinateur et les lois de la physique*, trad. de l'anglais par Françoise Balibar et Claudine Tiercelin, Paris, InterÉditions.
- François RASTIER (2001) : *Arts et sciences du texte*, Paris, PUF, coll. « Formes sémiotiques ».
- Edgar Allan POE (1989 [¹1841]) : « Double assassinat dans la rue Morgue », trad. de l'américain par Charles Baudelaire, in *Contes, essais, poèmes*, Paris, Robert Laffont, coll. « Bouquins », pp. 517-545.
- Baruch de SPINOZA (1988 [¹1677]) : *Éthique*, trad. du latin par Bernard Pautrat, Paris, Seuil, coll. « Points Essais ».
- VOLTAIRE (1954) : *Romans et contes*, Paris, Gallimard, coll. « Bibliothèque de la Pléiade ».
- Andrew WHITAKER (1996) : *Einstein, Bohr and the quantum dilemma*, Cambridge, Cambridge University Press.
- Anton ZEILINGER (1997@) : « On the interpretation and philosophical foundation of quantum mechanics », <http://www.quantum.univie.ac.at/zeilinger/philosop.html>.

Nota : Les ouvrages signalés du sigle @ à côté de la date de parution sont publiés sur support électronique.

Opacité et traduction : le cas des récits étiologiques

Franck Barbin
Laboratoire CRISCO
Université de Caen
Esplanade de la Paix
14032 Caen Cedex 05
franckbarbin@wanadoo.fr

Résumé

Un récit peut être qualifié d'étiologique lorsqu'il matérialise les tentatives du peuple pour expliquer l'inconnu, l'irrationnel (la présence d'un étrange rocher, par exemple). Les récits étiologiques ont pour vocation d'accompagner le public dans sa quête de sens. Le conteur participe à ce travail d'élaboration du sens autour de ce phénomène a priori inexplicable. A la fin de la narration, le phénomène inexplicable de départ a trouvé sa justification et l'auditeur peut repartir rasséréiné. Le traducteur, en tant que nouveau conteur, s'inscrit totalement dans cette démarche et guide pas à pas le lecteur dans son décodage du récit. Il doit combler ce déficit d'informations pour dissiper l'opacité des récits étiologiques.

Leur apparente absurdité peut s'expliquer par la notion freudienne d'« inquiétante étrangeté » (*Unheimlich*). Nous éprouvons ce sentiment d'étrangeté et de familiarité en lisant ou en écoutant ces récits, car l'élaboration secondaire conduit, notamment par l'opération de refoulement, à des manques dans le récit, qui correspondent aux éléments inconscients refoulés. Nous reconnaissons dans ces récits une vérité familière (constitutive de nous-mêmes), mais nous ne pouvons y croire au regard de la raison. Il incombe donc au conteur et au traducteur de reconstruire ces manques et de donner au public les clés nécessaires pour comprendre l'histoire. Jusqu'où faut-il aller pour conserver l'impact du récit et ne pas dénaturer son effet cathartique ?

Mots-clés : catharsis, conteur, folklore, inconscient, mystère, refoulement, sens, traduction.

Introduction

Le thème de l'étrange et du surnaturel est tout particulièrement prégnant dans le Sud-Ouest anglais et dans ses récits populaires. Il suffit de prendre le titre de certains ouvrages écrits ou co-écrits par le seul Michael Williams pour apprécier l'importance de ce thème : *Mysteries in the Devon Landscape*, *Occult in the West*, *Paranormal in the Westcountry*, *Strange Stories from Devon*, *Supernatural Dartmoor*, *Supernatural in the West*. Nous estimons que c'est la topographie de cette région (notamment les landes quasi-désertiques du Dartmoor) qui la prédestine à engendrer des histoires extraordinaires.

Le mystère imprègne cette contrée et ses habitants tentent d'en démêler les fils. Le conteur joue notamment ce rôle, en mettant en scène ces phénomènes inexplicables. Sa narration constitue une tentative d'explication et de résolution du problème. Le traducteur-conteur doit aller plus loin dans cette voie : il doit conduire un lecteur étranger vers la compréhension du phénomène concerné.

Nous allons définir en premier lieu ce que nous entendons par « récit étiologique » et la part de mystère inhérente à ce type de récit. Nous illustrerons ensuite la façon dont le traducteur-conteur doit s'y prendre pour éclaircir ces zones d'ombres et pour amener son lectorat à comprendre le récit. Nous verrons qu'il existe toutefois un danger à procéder ainsi : jusqu'à quel niveau d'explicitation le traducteur doit-il aller pour conserver l'impact du récit et ne pas dénaturer son effet cathartique ? La psychanalyse nous apporte une aide précieuse à ce sujet.

1. Les Récits étiologiques

1.1 Définition

Le « récit étiologique »¹ a pour objet la recherche de causes. Ce type de narration populaire matérialise les tentatives du peuple pour expliquer l'inconnu, les « pourquoi », comme le nomment les folkloristes :

Myths reply to all these questionings, and the answers are always in accordance with that early nebulous condition of thought and reason [...].

¹ Nous n'aborderons pas ici les différences, de portée notamment, qui existent entre les mythes étiologiques et les légendes étiologiques.

All seek to satisfy the eternal human desire to *know*.² (Lang, 1899, 2 : 300–01)

Bernadette Bricout rappelle la finalité de l'étiologie selon Morten Nojgaard :

Exposer une certaine chaîne d'actions dans un passé éloigné, et, ensuite, d'en tirer la conséquence qui explique un phénomène donné de la réalité du lecteur. (Bricout, 2002 : 367)

Les récits à visée étiologique cherchent le plus souvent à expliquer la présence de tel ou tel élément naturel, et tout particulièrement de rochers dans le Devon :

Nearly everyone living or travelling in North Devon has seen Lynton's Valley of Rocks. But how it came into being is the puzzle.³ (Sharman, 1952 : 124)

Long ago he lived, in the thirteenth century, but still his memory lingers on in the beautiful country that lies round Amicombe Hill. The West Ockment tumbles by not far away, and another neighbour is Bridestowe, with St Bridget's Church not three miles off from Lydford Gorge itself. And on the hill stand these curious stones called Bronescombe's Loaf and Cheese.⁴ (Sharman, 1952 : 157)

Umberto Eco déclare que la fonction des mythes est de « donner forme au désordre de l'expérience » :

Nous fuyons l'angoisse qui nous saisit lorsque nous essayons de dire quelque chose de vrai sur le monde réel. Telle est la fonction thérapeutique de la narrativité et la raison pour laquelle les hommes, depuis l'aube de l'humanité, racontent des histoires. (Eco, 1996 : 117)

1.2 Le cas de la légende du prêtre et de son vicaire

L'esprit humain cherche à tout prix à rationaliser l'univers qui l'entoure et le conteur l'aide à donner sens à ce qui était avant mystérieux. C'est notamment la raison pour laquelle la légende du prêtre et de son vicaire (« *The Parson and*

² Traduction proposée : « Les mythes répondent à toutes ces questions et les réponses sont toujours en accord avec une manière primitive et nébuleuse de penser et de raisonner [...] Tous cherchent à satisfaire l'éternel désir de *savoir* de l'homme. »

³ Traduction proposée : « Presque toutes les personnes qui vivent dans le nord du Devon ou qui le traversent ont vu la Vallée des Rochers à Lynton. Mais, comment elle est arrivée là reste un mystère. »

⁴ Traduction proposée : « Il vivait il y a bien longtemps, au XIII^e siècle, mais on se souvient encore de lui dans la belle campagne qui entoure la colline de Amicombe. Les eaux tumultueuses du West Ockment coulent non loin et on trouve aussi à côté Bridestowe, avec l'église Saint Bridget, à cinq kilomètres seulement de la gorge de Lydford. Et sur la colline se trouvent deux rochers bizarres qu'on appelle le *Loaf and Cheese* (le pain et le fromage) de Bronescombe. »

the Clerk ») a été inventée, pour rendre compte de la présence de deux curieux rochers situés au large du port de Dawlish. Nous allons observer d'un peu plus près certaines actualisations que nous avons pu trouver de cette légende⁵. Nous avons choisi les versions qui insistent sur l'étrangeté⁶ de ces deux rochers :

Neither he nor his clerk ever returned to their parish, but that morning the people of Dawlish saw two strange red rocks standing off the cliffs, and later, learning this story they realised that the demons had changed the evil priest and his man into these forms.

Time and weather have wrought many changes in the Parson and Clerk Rocks, not the least curious being to carve upon the Parson Rock the semblance of the two revellers. From certain positions you may see to-day the profiles of both men, the parson as it were in his pulpit, and the clerk at his desk beneath him. (Lyonesse, 1922 : 10–11)

The parson and the clerk never reached the village, and that morning, when the Dawlish fishermen went out in their boats, they found two strange red rocks standing off their cliffs [...]. And there the red rocks still stand to-day, to warn people against evil thoughts and ambitions. (Rutley, 1942 : 80)

At one spot a large part of the cliff had been washed away, leaving two queer shaped rocks of red sandstone as monuments to the power of the sea. One rock was still attached to the mainland while the other rose from the sea a little to the westward. (Coxhead, 1954 : 67)

And, standing in the sea, close to the cliffs, they were astonished to see two large red sandstone rocks that had not been there before.

The two men were never seen again, but the rocks remain there to this day and they are still known as The Parson and Clerk. (Maddock, 1965 : 67–68)

Next morning the people of Dawlish saw to their astonishment that a strange new formation had appeared in the sandstone of their coast. It was the Parson and his Clerk petrified for ever, and if you do not believe this story, go to Dawlish and the people there will show them to you. (Farquharson-Coe, 1976 : 113)

Neither parson nor clerk was ever seen again, although two mysterious rocks appeared from nowhere overnight. The rocks, one bigger than the other, still make a spectacular landmark a mile and a half east of Teignmouth and because of their contrasting size and the circumstances of their arrival, they could only be known as the Parson and the Clerk. (Jones, 1981 : 95)

The townspeople assembled on the beach were filled with wonder, for a curiously formed rock had magically appeared in the sea just below the cliffs. The parson and the clerk had been turned to stone, and there they stand to this very day. (Gibbs, 1983 : 24)

⁵ Pour des raisons évidentes de place, nous n'allons pas présenter ici les dix-sept formules de clôture ou *excipit* que nous avons pu rassembler de cette légende.

⁶ C'est nous qui soulignons les termes relatifs à ce champ sémantique dans les extraits suivants.

Two mysterious rocks appeared over night, one bigger than the other, a mile and a half east of Teignmouth — ever since known as the Parson and Clerk. (Barber et Chard, 1990 : 31)

Toutes les versions de cette histoire conduisent au même châtement final⁷ : pour les punir de leur ambition démesurée, les deux acolytes sont changés en rochers par le diable. Les habitants de la région ont été fort étonnés de trouver un beau matin deux nouveaux rochers dans leur environnement.

Ce sentiment d'étrangeté constitue en fait un champ sémantique dominant dans l'ensemble du récit, comme en témoignent les résultats suivants⁸ :

strange (16), particular (7 sur 8), fancy (5), astonishment (4), curious (4), gloom (3), peculiar (3), queer (3), remarkable (3), unusual (3), mysterious (2), puzzle (2), singular (2), spectacular (2), suspect (2), askance (1), eerie (1), loom (1), magically (1), odd (1), surprise (1), weird (1), wonder (1 sur 3).

De même, lorsque les deux religieux se retrouvent dans la maison du diable, il émane des invités de leur hôte quelque chose d'étrange et de menaçant :

wild (11 sur 22), demons (8), wild-looking (5), scream (4 sur 7), strange (11 sur 16), glee (3), shriek (3), queer (2 sur 3), cackle (2), hideous (2), jeer (2), chilling (1), evil-looking (1), haggard (1), howl (1 sur 2), imps (1), leering (1), pandemonium (1), satanic (1), treacherous-looking (1).

Ce n'est qu'à la fin du récit que les convives se révèlent sous leur vrai jour :

turn (7 sur 17), suddenly (5 sur 7), transform (2), recognize (1), reconnoitre (1), replace (1 sur 3).

L'effet de surprise frappe principalement les deux protagonistes de l'histoire ; le lecteur a très rapidement décelé, par l'entremise du conteur, la vraie nature des invités. Le conteur indique progressivement au lecteur à quoi s'en tenir au sujet de ces convives. Les termes qu'il choisit pour les décrire ne laissent planer aucun doute concernant leur identité réelle.

Cette élucidation de la part de mystère des récits doit toutefois être réalisée avec circonspection : elle recèle des dangers, notamment si nous nous situons dans une optique psychanalytique.

⁷ Nous avons discuté du degré général d'opacité des récits populaires lors de notre étude consacrée à la notion de figement. Nous avons vu combien le rôle du traducteur est déterminant pour rendre le texte accessible au lecteur d'arrivée. (Barbin, 2005)

⁸ Ces résultats ont été obtenus par *Lexico 3*, logiciel de statistique textuelle qui a été développé par André Salem et son équipe de Paris 3 (Cédric Lamalle, William Martinez et Serge Fleury). Nous indiquons entre parenthèses le nombre total d'occurrences de chaque forme, en précisant celles qui sont en accord avec le thème choisi. Les termes sont classés selon un double critère, par ordre de fréquence et par ordre alphabétique.

2. Récits et psychanalyse

2.1 Construction du sens et catharsis

Sigmund Freud suggère que les mythes, corollaires des rêves de l'humanité, nécessite un travail de déchiffrement, d'interprétation, pour être compris : « découvrir les motifs cachés qui ont causé les changements du mythe à partir de son sens original ». (Freud, 1991 : 208)⁹

Dans son article « Le Motif du choix des coffrets », il accorde un rôle important au créateur littéraire, qui doit s'efforcer de retrouver dans des phénomènes résiduels le motif originel du mythe afin que le lecteur puisse éprouver le sens saisissant de celui-ci :

Le remaniement régressif qu'il a ainsi entrepris, au moyen du mythe déformé par une transmutation du désir, laisse affleurer son sens ancien jusqu'à rendre également possible une interprétation allégorique. (Freud, 1999 : 81)

Le traducteur, tout comme le conteur, doit retrouver ou conserver le sens originel du récit, ce qui en fait sa raison d'être.

Jean-Paul Valabrega va plus loin dans ce syncrétisme en développant la notion freudienne de *psycho-mythologie*¹⁰. (Valabrega, 2001 : 163, 166) En rapprochant les contenus mythiques des produits de l'inconscient, il explique ainsi l'apparente absurdité des récits mythiques : nous éprouvons ce sentiment d'incohérence, car l'élaboration secondaire conduit, notamment par l'opération de refoulement, à des manques dans le récit, qui correspondent aux éléments inconscients refoulés.

Sigmund Freud désigne par le terme *Unheimlich*¹¹ ce sentiment à la fois d'étrangeté et de familiarité que nous ressentons en les écoutant ou en les lisant. Le sentiment d'étrangeté est toujours introduit dans le récit par un « tout à

⁹ Karl Abraham a développé cette parenté psychologique entre le rêve et le mythe : « Comme Freud l'a remarqué pour le rêve, le mythe aussi cache un contenu latent derrière le contenu manifeste. Pour l'atteindre, il faut une technique d'interprétation. Comme pour le rêve, ce procédé doit dévoiler l'ensemble du matériel de représentation et de sentiments qui s'exprime dans le mythe. » (Abraham, 1965 : 187)

¹⁰ Sigmund Freud esquisse cette idée dans une lettre à Wilhelm Fliess, datée du 12 décembre 1897. (Freud, 1991 : 210)

¹¹ Sigmund Freud a consacré un article à ce domaine en 1919. Il considère toutefois que, dans les contes, l'impression d'« inquiétante étrangeté » est fondée sur une supercherie de l'auteur (qui nous fait réagir à des expériences qui vont au-delà du vécu possible), réduisant ainsi le sentiment d'angoisse que nous éprouvons. (Freud, 1999 : 260–63)

coup », par un « soudain », qui marque l'irruption et la reconnaissance fulgurante de l'Inconscient dans le champ de la conscience. Ce savoir que l'on détient sans le savoir engendre un sentiment à la fois d'étrangeté et de familiarité. Au regard de la raison, nous ne pouvons y croire (« j'ai fait un curieux rêve ») ; pourtant, nous ne pouvons pas nous empêcher d'y accorder un peu de crédit, car nous y reconnaissons une vérité familière (constitutive de nous-mêmes), qui est celle de l'Inconscient. (Valabrega, 1980 : 144–45)

Jean-Paul Valabrega postule ainsi que l'Inconscient est construit et reconstructible comme un mythe :

Le mythe peut donc, croyons-nous, se définir comme le *conteur de l'Inconscient*. (Valabrega, 2001 : 165)

Lorsqu'un individu ou une collectivité tente de saisir le sens de son rêve ou d'un récit, jugé souvent comme absurde ou déraisonnable, son explication ne peut en être qu'insuffisante, car il ne peut s'inspirer que de son contenu manifeste¹². Il n'a pas accès à son contenu latent¹³, qui est transformé dans le rêve ou dans le récit en contenu manifeste (compromis entre le désir et la censure).

Le travail du traducteur-conteur n'est pas celui d'un analyste, qui cherche à révéler le contenu *latent*. Le conteur Bruno de La Salle amorce déjà cette idée en affirmant qu'il est préjudiciable de vouloir rationaliser les récits populaires :

Il y a évidemment un très grand danger à vouloir tout expliquer, à vouloir tout rendre logique. Malgré des schémas narratifs souvent presque parfaits, le conte ne se réduit pas à des propositions mentalement et moralement logiques. Sa beauté réside dans le mariage apparemment contradictoire d'un formulaire rigoureux avec des personnages et des événements symboliques dont le sens demeure par définition voilé. (La Salle, 1995 : 119)

Karl Abraham parle, au sujet des mythes, de *refoulement collectif*, qui empêche le peuple de comprendre le sens originel de ces mythes. (Abraham, 1965 : 190)

Il doit en rester ainsi dans les récits traduits pour deux raisons. D'une part, le traducteur ne doit pas interpréter le sens du récit à la lumière de ses propres expériences, de ses propres conceptions : ce qui est valable pour lui n'est pas forcément valable pour tout lecteur. Le récit perdrait alors sa valeur universelle.

¹² « Désigne le rêve avant qu'il soit soumis à l'investigation analytique, tel qu'il apparaît au rêveur qui en fait le récit. Par extension on parlera du contenu manifeste de toute production verbalisée — du fantasme à l'œuvre littéraire — qu'on propose d'interpréter selon la méthode analytique. » (Laplanche et Pontalis, 1997 : 101)

¹³ « Ensemble de significations auquel aboutit l'analyse d'une production de l'inconscient, singulièrement du rêve. Une fois déchiffré, le rêve n'apparaît plus comme un récit en images mais comme une organisation de pensées, un discours, exprimant un ou plusieurs désirs. » (Laplanche et Pontalis, 1997 : 100)

D'autre part, le traducteur-ethnologue ne doit faire émerger ce sens caché ou contenu latent, sous peine de le rendre encore plus inaccessible au lecteur. C'est le risque le plus grand qu'encourt le traducteur-conteur : il arrivera toujours à rendre — avec plus ou moins de bonheur — dans la langue d'arrivée le contenu manifeste du récit, mais, il peut, en voulant trop bien faire, vider le récit de sa fonction d'exutoire. En effet, le lecteur saisit, au niveau inconscient, le message véhiculé par le récit, mais s'il l'amène dans le champ de la conscience, il le rend par là même inaccessible au lecteur, privant ainsi la narration de son action bienfaitrice.

Il faut laisser au lecteur la possibilité de s'identifier au personnage principal du récit. L'identification est un genre de comparaison, un « de même que » qui reste étranger au lecteur et qui doit le rester. Les récits populaires permettent alors la décharge de ses désirs refoulés, opérant ainsi un effet cathartique sur le lecteur :

Folklore reveals man's frustrations and attempts to escape in fantasy from repressions imposed upon him by society. [...] But folklore also reveals man's attempts to escape in fantasy from the conditions of his geographical environment and from his own biological limitations.¹⁴ (Bascom, 1965 : 290-91)

2.2 Childe et le pont de la ruse

Nous allons prendre l'exemple de la légende étiologique de Childe le chasseur et nous pencher sur la place du pont dans cette histoire¹⁵. L'appellation *Guile Bridge* semble à première vue logique, car il signifie littéralement « Pont de la Ruse », expression que nous pourrions également traduire par « Pont de la Perfidie » — rappelons-nous que nous avons affaire à un pont anglais...

Or, la situation n'est pas si simple que cela, comme le fait remarquer à juste titre Farquharson-Coe :

Nearly all the stories of Mr. Childe and his horse and his lands, end with the words, "And the bridge is known as Guile Bridge to this day." Well it isn't. I for one don't know it. I have tried long and hard to identify this infernal bridge. A bridge close to Tavistock near Parkwood was suggested. Another informant suggested the bridge near the post office, in the centre of the town.

¹⁴ Traduction proposée : « Le folklore met en scène les frustrations humaines et ses tentatives d'échapper par le fantasme aux répressions que lui impose la société. [...] Mais, le folklore met également en scène ses tentatives d'échapper par le fantasme aux limites imposées par son environnement géographique et par son être biologique. »

¹⁵ En résumé, les moines de deux abbayes distinctes — celles de Tavistock et de Plymstock — veulent hériter des propriétés de Childe. Ils rivalisent de ruse afin de s'emparer de son trésor. Après une longue attente au gué de la rivière Tavy, les moines de Plymstock se rendent finalement compte qu'ils ont été dupés. Ce sont finalement les moines de Tavistock qui enterrent le corps de Childe.

I communicated with the powers regulating bridges and roads and maintaining them, and put the question to them. They did not know.¹⁶
(Farquharson-Coe, 1976 : 100)

Puisque le pont est qualifié d'« infernal », pourquoi ne pas le traduire par le « Pont machiavélique » ?

Le pouvoir des contes est si fort qu'il a réussi à transformer le nom d'un pont (*Guild Bridge*) en un autre (*Guile Bridge*), comme en témoigne Ruth Leger-Gordon :

Etymologists, however, will have none of this fantasy and attribute the building of the bridge to one of the Tavistock medieval trade Guilds, the name having been conveniently corrupted to fit neatly into the legend.¹⁷
(Leger-Gordon, 1972 : 58)

Cette nouvelle dénomination s'accompagne d'un changement étymologique¹⁸, rendant ainsi son origine nettement plus opaque. Bien que la graphie entre les deux termes semble proche pour un francophone, il ne faut pas oublier que ces deux mots ne prononcent pas du tout de la même façon : [gild] et [gail]. Nous ne pouvons pas à proprement parler penser à un jeu de mots ici, puisque personne, ou presque, ne semble faire le lien entre ces deux termes. Seuls William Hoskins et Herbert Finberg tentent de rapprocher ces deux appellations :

No document that has come to light so far refers to a "Guile Bridge" in or near Tavistock, but in 1651 the Tavistock Guildhall is called "Guilehall." If there was a Guile Bridge, therefore, it was probably so called because constructed or maintained by a parish guild.¹⁹ (Hoskins et Finberg, 1952 : 41)

¹⁶ Traduction proposée : « Presque toutes les histoires mettant en scène Childe, son cheval et ses terres se terminent par ces mots : "Depuis ce jour, le pont s'appelle le Pont de la Ruse." Eh bien, non. Moi, par exemple, je ne le connais pas. J'ai essayé durant des années d'identifier ce satané pont. Certains pensent qu'il s'agit d'un pont situé aux alentours de Tavistock, près de Parkwood. Selon un autre informateur, il se trouverait dans le centre ville, près de la poste. J'ai contacté les Ponts et Chaussées pour leur poser la question. Ils n'en ont aucune idée. »

¹⁷ Traduction proposée : « Les étymologistes ne se laissent toutefois pas séduire par cette explication fantaisiste et attribuent la construction de ce pont à l'une des *Guildes* de métier résidant à Tavistock au Moyen Âge. Son nom a simplement été altéré afin de mieux coller à la légende. »

¹⁸ Avons-nous réellement affaire à un cas d'étymologie populaire ? Les avis sont assez divergents à ce sujet, puisque chacun attribue une origine différente au pont en question.

¹⁹ Traduction proposée : « Aucun document à ce jour ne fait référence à un *Guile Bridge* à Tavistock ou dans ses environs, mais, en 1651, le *Guildhall* (hôtel de ville) de Tavistock est appelé *Guilehall*. S'il existait un *Guile Bridge*, il est donc probable qu'il a été nommé ainsi, car il avait été construit ou entretenu par une guilde de la paroisse. »

Le traducteur pourrait donc être tenté d'expliquer cette évolution lexicale en jouant sur l'homophonie en français entre « guilde » et « guigne » : le « Pont de la Guilde » peut facilement devenir le « Pont de la Guigne ».

Mais ce n'est pas le fin mot de l'histoire : James Whinray affirme que ce fameux pont est plus connu sous le nom de *Abbey Bridge* (car il a été construit par des religieux). (Whinray, 1996 : 9) Il reprend certainement le commentaire de Mrs Henry Whitcombe :

The monks, however, were not to be easily outwitted. They hastily changed their course, and throwing a bridge, known to this day as Guile Bridge (but more commonly called the Abbey Bridge), across the river near the Abbey, reached Tavistock in safety, and thus gained possession of the lands.²⁰
(Whitcombe, 1974 : 56)

Comme le récit met en scène des moines qui essaient de récupérer le corps de Childe, nous pourrions donc traduire cette expression par le « Pont des Moines ».

Devons-nous réellement trancher pour savoir quelle hypothèse est la plus plausible ? C'est là tout le charme du folklore. Ce sont autant d'options qui sont aussi valables les unes que les autres ; tout est une question de choix traductologiques et d'impact sur le lecteur.

Conclusion

Nous avons pu observer que la phase de déchiffrement opérée par le traducteur-conteur est non seulement périlleuse, mais aussi lourde de conséquences pour le lecteur : il doit expliciter le message de départ tout en préservant son sens caché. Néanmoins, une fois le problème posé, il est difficile de définir a priori où se situe cette limite ; c'est le traducteur-conteur qui doit la déterminer en fonction de ce qu'il sait du lecteur-cible. Comme l'ethnopsychanalyse l'a bien mis en évidence, un même contenu peut être acceptable dans une culture, mais inacceptable dans une autre : le couple latent/manifeste est régi par une forte détermination culturelle. (Devereux, 1977)

Jean-Paul Valabrega accorde à ce titre une importance toute particulière aux différentes versions d'un mythe :

²⁰ Traduction proposée : « Cependant, les moines n'allaient pas se laisser bernier aussi facilement. Ils changèrent aussitôt leur trajet et construisirent un pont de fortune, connu sous le nom de Guile Bridge (mais plus connu encore sous le nom de Abbey Bridge) sur la rivière près de l'Abbaye. Ils arrivèrent sains et saufs à Tavistock et prirent possession des terres. »

C'est surtout parce que les versions, variantes et variations d'un mythe ouvrent une porte sur ses contenus *latents*. (Valabrega, 2001 : 105)

C'est pourquoi nous leur accordons une place si prépondérante et que leur étude nous semble essentielle pour toute interprétation. Chaque version d'un même récit apporte des clés différentes aux problèmes posés par le récit.

Le traducteur doit être en mesure d'évaluer les besoins du lecteur-cible et d'explicitier seulement les points qui peuvent poser problème :

Le traductologue, lui, constate parfois un manque de méthode dans l'explicitation ; dans un même texte, peuvent coexister des explicitations superflues et certaines obscurités qui auraient pu être éclaircies. (Lederer, 1998 : 164)

Le traducteur s'apparente donc à un équilibriste qui jongle avec le sens du récit, à un fildefériste qui marche sur le fil ténu du sens. En réalité, le traducteur doit opérer un double travail de décodage : il doit non seulement fournir une explication au phénomène insolite, mais également initier son lectorat aux réalités culturelles²¹ du récit d'origine.

Le traducteur-conteur est par conséquent confronté à ces difficultés à double titre : en tant qu'interprète de la tradition et en tant que « co-auteur » du récit. Il est plus qu'un simple « récrivain » du récit de départ ; il contribue à part entière à la création du récit. Comme nous avons pu rapidement l'observer au point 1.2, le traducteur-conteur ne part pas d'un récit-source unique ; il s'inspire au contraire de récits-sources multiples pour parvenir à sa traduction. Il propose un récit qui n'a jamais été énoncé sous cette forme auparavant. Il doit donc avoir pleinement conscience des dangers qui sous-tendent son entreprise.

Bibliographie

Théorie

ABRAHAM, Karl (1965), *Oeuvres complètes I : Rêve et mythe*, Paris, Petite Bibliothèque Payot, Science de l'Homme.

BARBIN, Franck (2005), « Folklore et traduction : quels types de figement ? », In *La Traduction : Identités et altérités*, Christine Durieux, Caen, Cahiers de la Maison de la Recherche en Sciences Humaines de Caen.

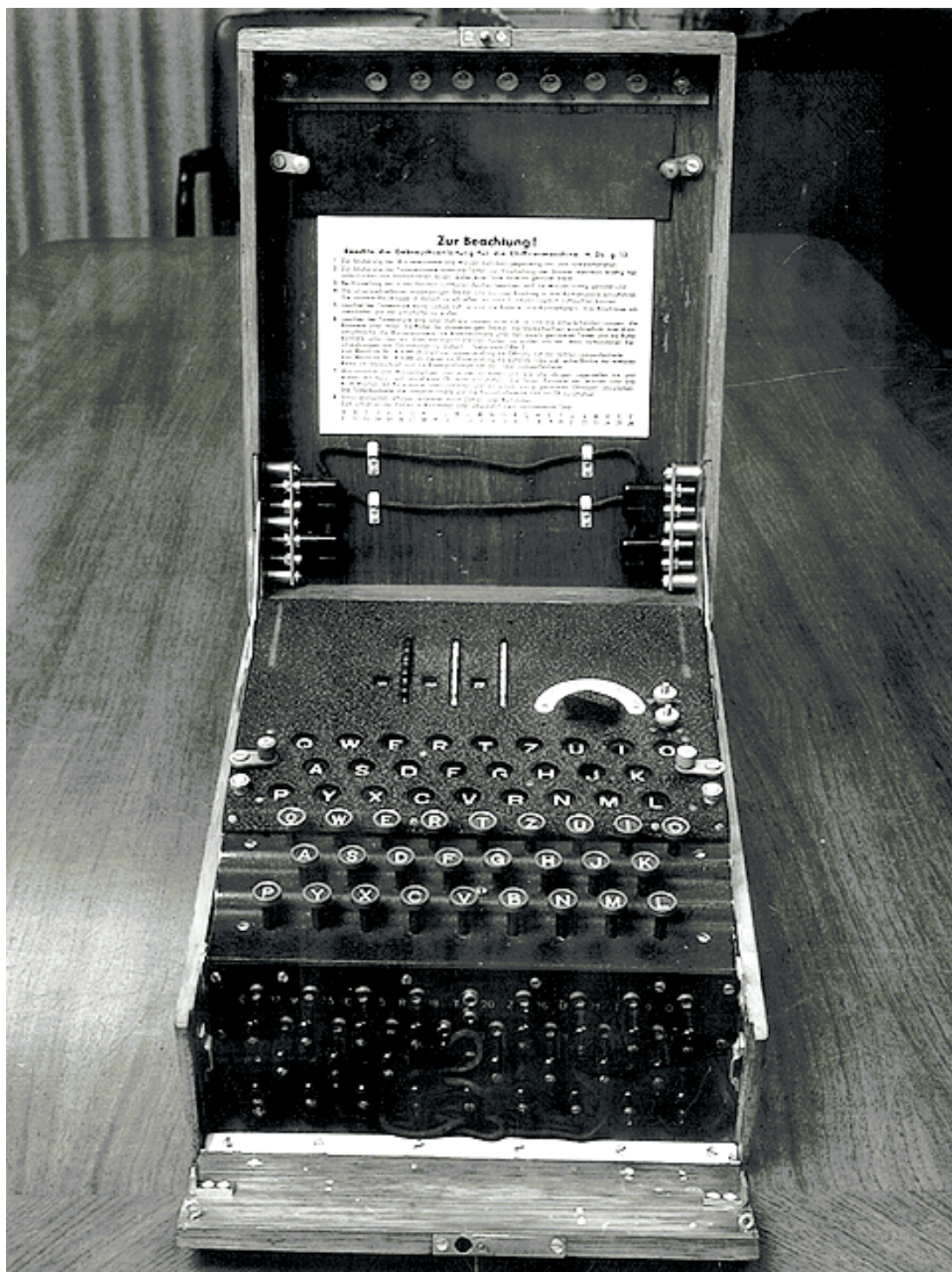
²¹ Nous n'avons pu développer ici ce second aspect. Divers éléments sont sources d'opacité pour le lecteur étranger (réalités géographiques différentes, termes dialectaux, expressions figées, etc.) et demandent un traitement spécifique de la part du traducteur. Pour de plus amples détails, nous renvoyons à notre étude à ce sujet. (Barbin, 2001)

- (2001), « Les Récits populaires du Devon : problèmes de méthode et de traduction », In *Oralité et Traduction*, Michel Ballard, 291–319, Arras, Artois Presse Université, Traductologie.
- BASCOM, William R. (1965), « Four Functions of Folklore », In *The Study of Folklore*, Alan Dundes, 279–98, Englewood Cliffs, Prentice Hall.
- BRICOUT, Bernadette (2002), « Le Conte », *Encyclopaedia Universalis* 6, 365–70.
- DEVEREUX, Georges (1977), *Essais d'ethnopsychiatrie générale*, Paris, Gallimard, Tel.
- ECO, Umberto (1996), *Six promenades dans les bois du roman et d'ailleurs*, Paris, Bernard Grasset.
- FREUD, Sigmund (1991), *La Naissance de la psychanalyse*, Paris, Presses Universitaires de France.
- (1999), *L'Inquiétante étrangeté et autres essais*, Paris, Gallimard, Folio Essais.
- LANG, Andrew (1889), *Myth, Ritual and Religion*, vol 2, Londres, Longmans, Green.
- LAPLANCHE, Jean et Jean-Bertrand PONTALIS (1997), *Vocabulaire de la psychanalyse*, Paris, Presses Universitaires de France, Bibliothèque de Psychanalyse.
- LEDERER, Mariane (1998), « Traduire le culturel : la problématique de l'explicitation », *Palimpsestes* 11, 161–171.
- VALABREGA, Jean-Paul (1980), *Phantasme mythe corps et sens : une théorie psychanalytique de la connaissance*, Paris, Payot, Science de l'Homme.
- (2001), *Les Mythes, conteurs de l'Inconscient : questions d'origine et de fin*, Paris, Éditions Payot et Rivages, Bibliothèque scientifique Payot.

Corpus

- BARBER, Chips et Judy CHARD (1990), *Tales of the Teign*, Exeter, Obelisk Publications.
- COXHEAD, John Ralph Winter (1954), *Legends of Devon*, Westward Ho !, Western Press.
- FARQUHARSON-COE, A. (1976), *Devon : Mostly Ghostly but the Rest is Folklore*, Plymouth, Baron Jay.
- GIBBS, Cory E. (1983), *Tales of Devon*, Dawlish, Flying Basset Publications.
- HOSKINS, William George et Herbert Patrick Reginald FINBERG (1952), *Devonshire Studies*, Londres, Jonathan Cape.
- JONES, Sally (1981), *Legends of Devon*, Bodmin, Bossiney Books.
- LAUDER, Anne et Michael WILLIAMS (1994), *Strange Stories from Devon*, Callington, Penwell.
- LEGER-GORDON, Ruth E. St (1972), *The Witchcraft and Folklore of Dartmoor*, Wakefield, E. P. Publishing.

- LYONESSE, George Basil Barham (1922), *Legend Land*, vol 2, Londres, Great Western Railway.
- MADDOCK, Llywelyn W. (1965), *West Country Folk Tales*, Bath, James Brodie.
- RUTLEY, Cecily M. (1942), *Legends and Folk-Lore of Devonshire*, Londres, Collin Clear Type Press.
- SHARMAN, Valentine Day (1952), *Folk Tales of Devon*, Londres et Édimbourg, Thomas Nelson.
- WHINRAY, James (1996), *Devon Legends*, Penryn, Tor Mark Press.
- WHITCOMBE, Mrs Henry Pennel (1874), *Bygone Days in Devonshire and Cornwall, with Notes of Existing Superstitions and Customs*, Londres, Richard Bentley.
- WILLIAMS, Michael (1979), *Occult in the West*, Bodmin, Bossiney Books.
- (1986), *Paranormal in the Westcountry*, Bodmin, Bossiney Books.
- (1996), *Supernatural in the West*, Bodmin, Bossiney Books.
- (2003), *Supernatural Dartmoor*, Launceston, Bossiney Books.
- WREFORD, Hilary et Michael WILLIAMS (1985), *Mysteries in the Devon Landscape*, Bodmin, Bossiney Books.



Demandes documentaires en contexte professionnel : Analyse cognitive de l'explicitation de la demande

Alexandra Ciaccia

Université Paris X-Nanterre
EA 1588 Psychologie cognitive des conduites complexes
200 avenue de la République 92000 Nanterre

-
Cegos - Direction Recherche et Développement
Convention Cifre
11 rue René Jacques 92798 Issy-Les-Moulineaux

alexciafr@yahoo.fr

Résumé

Dans le cadre d'une étude réalisée dans le centre de documentation de la Cegos, entreprise de formation et de conseil spécialisée dans le management, une analyse des demandes documentaires formulées par les consultants de l'entreprise aux documentalistes du centre a été réalisée. Les situations observées se caractérisent par un échange verbal entre le documentaliste et le consultant, au cours duquel le documentaliste tente de faire expliciter au consultant l'objet de sa demande. Des recherches d'information s'appuyant sur ce précédent échange sont ensuite conduites en vue de constituer un dossier documentaire en réponse à la demande formulée. Quels sont les procédés mis en oeuvre par le documentaliste pour tenter d'élaborer une représentation spécifique de la demande, en vue de la compréhension du problème posé ? Quelles sont les modalités d'explicitation de l'objet de recherche tant du point de vue du demandeur que du documentaliste ? Nous verrons au travers de l'analyse d'enregistrements transcrits de demandes documentaires comment le documentaliste à partir d'un problème qui n'est pas le sien tente d'en comprendre le cadre et le contenu dans le souci de sa résolution. Au travers de différents types d'interventions verbales (écoute, explicitation, apport d'information) de tentatives de mises en relation des informations fournies avec ses propres connaissances, de reformulations, et d'échanges co-construits, l'identification des attributs pertinents de la demande est rendue possible. Du problème initialement étranger pour le documentaliste, se dégage du sens qui permet, dans la mesure du possible, non seulement sa compréhension mais également sa résolution. Enfin, l'identification des informations nécessaires et suffisantes à l'élaboration d'une représentation du problème posé, semble pouvoir être supportée par une démarche de questionnement de type analytique.

Mots clés : Problème, compréhension, attributs, représentation, échange.

Cette étude s'inscrit dans le cadre plus général de travaux en cours sur le traitement de l'information dans des activités de recherche d'information informatisée en contexte professionnel. Dans l'objectif d'étudier les processus de problématisation d'un objet de recherche, que nous supposons mis en œuvre avant et pendant l'accomplissement des recherches, nous avons choisi d'observer une situation en particulier : la formulation orale de demandes documentaires dans le centre de documentation et d'information d'une entreprise. En effet, dans cette situation, le besoin informationnel est verbalisé ; la problématisation de l'objet de recherche n'est plus seulement mental mais verbal : un consultant délègue une recherche d'information à un documentaliste. Pour cela, il explicite verbalement les attributs de son besoin informationnel. La problématisation est par conséquent verbale et construite dans l'interaction avec le documentaliste. Cette situation présente en outre l'avantage d'appréhender les mécanismes en cours avant l'étape de recherche d'information, les circonstances dans lesquelles se construit l'objet d'une recherche, et notamment comment s'effectue le passage de l'identification d'un besoin informationnel à la communication homme-machine dans le cadre d'une tâche de recherche d'information.

L'article présenté ici expose les résultats qui concernent l'analyse des six premiers enregistrements verbaux de demande documentaire au regard des procédures mises en place par le documentaliste pour comprendre et faire émerger du sens à partir de la demande formulée.

Quelques éléments théoriques

Les approches proposées en sciences cognitives sur l'émergence et la construction de sens mettent en évidence des aspects différents mais non contradictoires. Certaines approches développent une conception individuelle de la construction de sens : la signification des perceptions externes (énoncés écrits, actes de langage, ...) se développe en fonction de facteurs internes à l'individu qui perçoit. La conception neurobiologique illustre cette approche : notre système nerveux central est considéré comme « *un système projectif, qui projette en permanence ses hypothèses sur le monde extérieur. Il les met à l'épreuve. Il donne du sens, parfois, à ce qui n'en a pas* » (Changeux, 1998). Le Ny (2005) propose également une conception « interne » de la construction de sens, même s'il reconnaît l'importance du contexte : « *La compréhension s'effectue à partir de deux sources générales d'information, et tout le processus consiste, en définitive à les entrelacer de façon convenable* ». Il propose comme première source, l'information externe : l'information langagière et l'information qui vient de la situation (« *information pragmatique* ») et comme deuxième source, l'information interne : « *il s'agit de l'information qui se trouve dans l'esprit du compreneur (...), sous forme de connaissances de diverses sortes, lexicales, grammaticales et générales : c'est cette information qui sera exploitée pour « donner sens » à l'énoncé.* ». Le sens apparaît ici

comme une représentation complexe se formant dans l'esprit des individus notamment au cours de la compréhension. Les aspects externes ne sont, de son point de vue, qu'une ressource à partir de laquelle le sens va se développer.

D'autres auteurs réservent une place plus importante au contexte, et notamment à la médiatisation dans la construction même de la signification. En effet, Pierce (cité par Bruner, 1991) reconnaît que la signification ne dépend pas seulement d'un signe et d'un référent, mais aussi d'un interprétant : une représentation du monde dans des termes tels que la relation signe-référent est médiatisée. Une étude de Baker et al. (1994) montre en effet que dans le cadre plus spécifique de la modélisation de dialogues qui concernent la recherche d'information et qui se caractérisent par des jeux d'explications entre documentalistes et demandeurs, les dialogues sont constitués « *de phases de raffinements successifs de requêtes et d'explications* », analysés par les auteurs comme des négociations. La signification de la demande d'autrui se développe dans l'échange, la médiatisation et la négociation.

L'étude empirique présentée ici s'inscrit dans cette dernière approche et tente d'illustrer en quoi l'interaction et l'échange permet l'identification d'attributs significatifs d'un besoin informationnel.

Positionnement

Nos hypothèses de recherche sont formulées en référence aux situations observées qui se caractérisent par des échanges verbaux en face à face entre un documentaliste et un consultant, échanges effectués dans le centre de documentation et d'information interne à l'entreprise. L'interaction entre le documentaliste et le consultant est guidée par l'intention suivante : le consultant délègue au documentaliste une recherche documentaire en lien avec son activité professionnelle et l'objectif pour le documentaliste est d'inventorier au cours du discours tous les éléments qui lui permettront au préalable, soit de transférer la recherche à son équipe, soit d'effectuer lui-même les recherches nécessaires à l'élaboration du dossier documentaire.

Les hypothèses reposent sur l'idée générale que l'assimilation d'information, énoncées par un individu en particulier (appelé désormais « le demandeur ») et relevant de contextes référentiels divers, est sous-tendue par l'adoption par le documentaliste de postures variées dans le discours (écoute acquiescante, demande d'explicitation, ...). Ces postures ou interventions sont d'une part, le reflet des procédures mises en place par le documentaliste pour faire émerger du sens à partir de ce que lui fournit le demandeur et, d'autre part sont le reflet de son positionnement quant à la compréhension de l'objet de la demande.

Méthodologie de la recherche

Procédure

L'objectif de la recherche étant d'analyser les demandes réalisées par les consultants au centre de documentation et d'information, nous avons enregistré les demandes avec un dictaphone. Pour chaque demande effectuée, le documentaliste proposait au consultant de participer à un projet d'observation des pratiques de veille dans l'entreprise et d'accepter pour cela l'enregistrement vocal de leur entretien. Les travaux présentés ici portent sur 6 enregistrements de demandes (3 demandes formulées avec le documentaliste SB, 2 demandes formulées avec MG et 1 demande formulée avec CI). Chaque enregistrement a été ensuite transcrit dans l'objectif de réaliser une analyse discursive. Chaque transcription suit une codification qui permet de les comparer entre elles et de rester le plus fidèle à l'échange oral réalisé*.

Les variables manipulées et observées

Chaque protocole a été découpé en **séquences** :

- La Séquence 1 démarre à la première prise de parole (que ce soit le documentaliste ou le demandeur qui prenne la parole) et prend fin lorsque le consultant a présenté le cadre de sa demande, la problématique générale de sa recherche et lorsque la documentaliste effectue ses premières interventions qui ne soient pas de l'ordre de l'écoute*.
- La Séquence 2 comprend les interventions à propos de l'objet de recherche ; elle débute après la Séquence 1 et prend fin lorsque la conversation n'est plus en rapport avec l'objet de recherche.
- La Séquence 3 débute à la fin de la Séquence 2 et se caractérise par des éléments de négociations de délai et de modalités de réception du dossier documentaire ; sa fin marque celle de l'entretien.

D'autre part, les **interventions du documentaliste** ont été catégorisées selon leur contenu :

- Les interventions du type Ecoute (EC) : interventions pour lesquelles le documentaliste marque par un ou deux mots son attention au discours du demandeur (« mm », « d'accord », « ok », ...).
- Les interventions du type Explicitation (EX) : demandes d'éclaircissement faites par le documentaliste sur un terme, sur une situation, ..., fréquemment formulées à la forme interrogative (« donc si c'est sur le circuit de la grande distribution, c'est uniquement pour les particuliers ? », « ça fait combien de temps qu'elle est vendeuse en France ? », ...).
- Les interventions du type Apport d'information (APP) : interventions pour lesquelles le documentaliste propose au cours de l'entretien des éléments de réponse à la demande (« il me semble me souvenir que sur la recherche qu'on avait effectuée sur l'urbanisme (...) on en avait conclu

* Voir en Annexe : transcription d'une séquence 1 et codification

qu'il existait tout un tas de formations », « on va être amené à aller voir du côté des écoles privées », ...).

- Les interventions du type Reformulation (REF) : interventions qui, dans leur intention, peuvent être assimilées à des interventions du type Explicitation, mais qui apparaissent sous la forme déclarative où le documentaliste exprime avec ses propres mots la phrase précédente du demandeur (« *d'accord c'est de l'organisation et de la conception de voyage* », « *quelle valeur je leur donne et quels moyens je veux mettre rapport à cette valeur que j'ai donné euh pour les gérer (...) c'est ça ?* », ...).
- Les interventions qui ne rentrent pas dans cadre de ces 4 catégories ont été classées sous une catégorie appelée Autre (A).

La classification de D. Malrieu (1992) a servi de base pour analyser le contenu du discours du demandeur et du documentaliste. De cette manière, **les mots ou expressions clés** ont été retenus et classifiés en deux catégories : les mots ou expressions clés Thématiques (T) et les mots ou expressions clés Contextuels (C). Les premiers (thématiques) sont caractérisés par un lexique relevant de « *l'approche disciplinaire* » (Malrieu, 1992) et par des termes désignant l'objet de la demande directement réutilisables lors de la recherche d'information, notamment dans les cartouches de requêtes des moteurs de recherche sur Internet (par exemple les mots ou expressions thématiques d'une demande : « *Fongecif* », « *formation* », « *secteur du tourisme plutôt évènementiel* », « *formation en alternance* », « *tourisme d'affaire* », ...). Les seconds (contextuels) spécifient l'objet de la demande par rapport à des situations concrètes ou des exemples (par exemple quelques-uns des mots ou expressions clés contextuels d'une demande en particulier : « *dans le cadre d'un bilan* », « *une fille qui a 38 ans* », « *dans son école de tourisme (...) mais elle a vu en sortant qu'elle était loin de la réalité* », « *sa boîte ne va pas très bien* »).

Hypothèses de recherche

Nous supposons que le documentaliste dans l'exercice de compréhension de la demande formulée par le consultant procède généralement en 2 grandes étapes : *une première étape* au cours de laquelle il saisit le discours du demandeur et construit parallèlement une représentation plus ou moins élémentaire de la demande en fonction des éléments donnés par le demandeur et ceux dont il dispose qui font lien avec le discours ; *une seconde étape* au cours de laquelle le documentaliste va continuer à construire, affiner, étoffer sa représentation de l'objet, en fonction de la résonance ou non des informations perçues et traitées avec ses propres connaissances.

Plus concrètement, nous devrions observer avec l'analyse des données, une fréquence plus importante des interventions de type Ecoute dans la Séquence 1 par rapport à la Séquence 2 (et parallèlement, des interventions de type

Explicitation, Apport d'information et Reformulation plus fréquentes dans la Séquence 2 de l'échange).

Nous supposons également que les interventions du documentaliste, en lien avec le contenu, engendrent un discours plus riche de la part du demandeur. En d'autres termes, les questions et les reformulations poussent le demandeur lui-même à reformuler et expliciter les éléments du discours. Dans cette phase, le documentaliste co-construit sa représentation de la demande avec le demandeur. L'interaction contribue à faire émerger des attributs pertinents pour comprendre la demande. Concrètement, nous devrions observer à l'issue de l'analyse des résultats que les mots ou expressions clés émergent plus fréquemment après des interventions de type Explicitation, Reformulation et Apport d'information. Nous devrions observer également, et ceci de manière anodine, que l'augmentation des interventions va de pair avec l'augmentation de la formulation de mots clés.

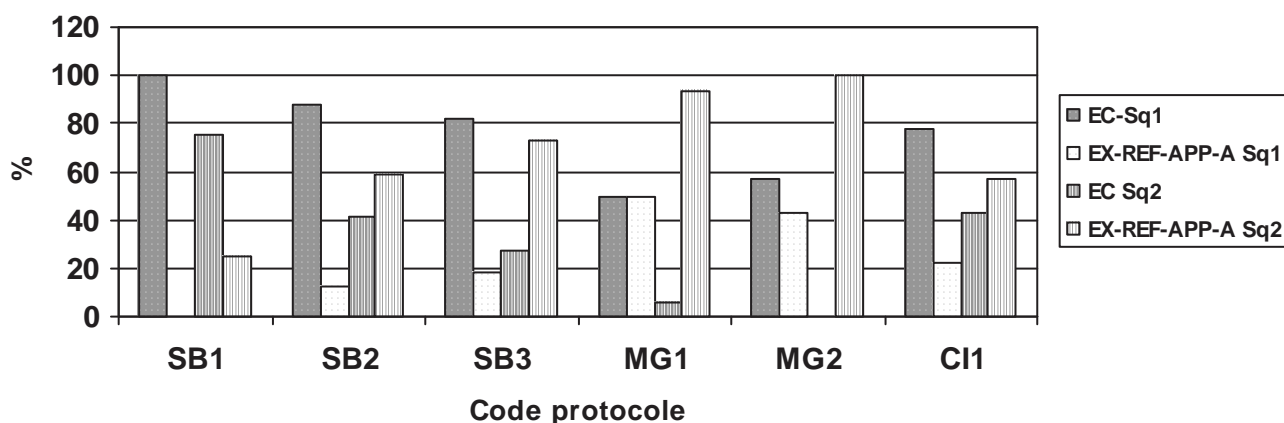
Enfin, nous supposons que le degré de conceptualisation des connaissances liées à un objet influence les modalités de son actualisation, la manière dont elles sont explicitées à autrui. De plus, l'adéquation entre les connaissances de l'énonciateur et du récepteur conditionne également la compréhension du discours de l'un vers l'autre. Les analyses devraient donc montrer que pour les demandes qui manifestent un niveau de conceptualisation élevé, les séquences 2 devraient comporter plus d'interventions du documentaliste (dans le cas où le degré de conceptualisation de l'objet de recherche est moins élevé chez le documentaliste).

Résultats

Hypothèse 1 : une compréhension de la demande en étape

Le Graph1 rend compte des **pourcentages d'interventions classées par type** (EC : Ecoute (en fond gris) ; et, réunis pour ce traitement des données, EX : Explicitation, REF : Reformulation, APP : Apport d'information (en fond blanc) **en fonction des séquences de l'échange** (Sq1 : séquence initiale de l'échange (à pois dans le graph.), Sq2 : séquence sur l'objet de la demande (rayés dans le graph.).

Graph1 Interventions et Séquences

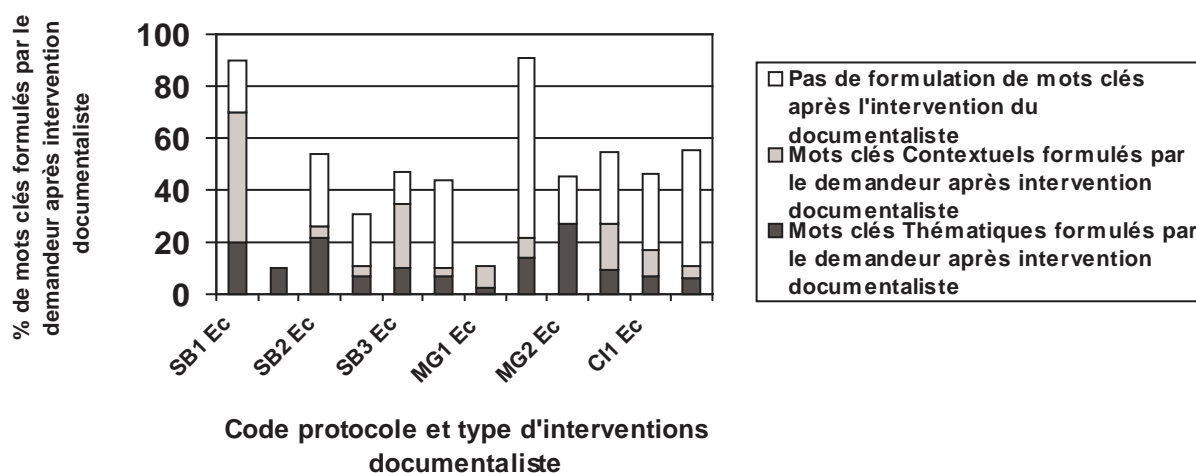


On observe que les interventions du documentaliste relèvent plus souvent du type Ecoute dans la première partie de l'échange alors que dans la seconde partie, les interventions sont plus souvent de l'ordre de l'explicitation, de la reformulation et de l'apport d'information.

Hypothèse 2 : co-construction de la signification de la demande

Le graph2 montre le **pourcentage de mots ou expressions clés thématiques** (les bâtons blancs du graph.) et **contextuels** (les bâtons gris du graph) **en fonction des interventions du documentaliste** (EC : Ecoute ; et réunis pour ce traitement des données, EX : Explicitation, REF : Reformulation, APP : Apport d'information (Appelés Ex sur l'axe des abscisses).

Graph2 Interventions et Mots clés

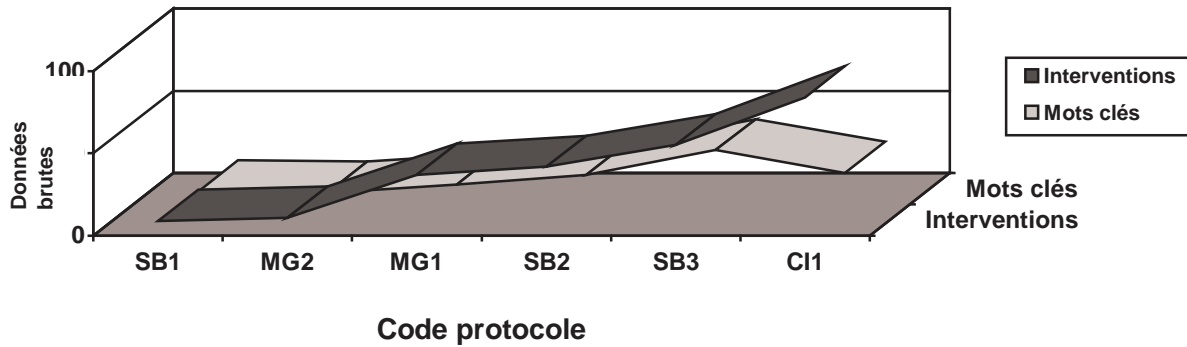


L'analyse du Graph2 montre que l'intervention du type Ecoute est celle qui est le plus souvent suivie de la formulation d'un mot ou d'une expression clé (la moyenne des pourcentages de mots clés est la plus élevée parmi tous les types d'interventions). Plus précisément, l'intervention Ecoute fait suivre la formulation de mots ou expressions clés contextuels.

Les données montrent également que les documentalistes font majoritairement des interventions du type Ecoute, les interventions de type Explicitation, Apport d'information et Reformulation étant utilisées de manière relativement homogène. Si on compare pour chaque protocole le pourcentage de mots clés apparaissant après une intervention Ecoute) et EX (pourcentage de mots clés apparaissant après une intervention Explicitation ou Apport d'information ou Reformulation), on observe que les protocoles MG1, MG2 et CI1 font apparaître une prédominance des formulations de mots clés après les interventions du type Explicitation, Apports d'information et Reformulation, tandis que pour les protocoles SB1, SB2 et SB3, les interventions de type Ecoute sont plus souvent suivies par des mots clés que les autres interventions.

Les courbes1, quant à elles, montrent la tendance à la proportionnalité croissante du nombre d'interventions avec le nombre de mots clés par protocole, toute séquence confondue :

Courbes1 Augmentation des interventions et augmentations des mots clés

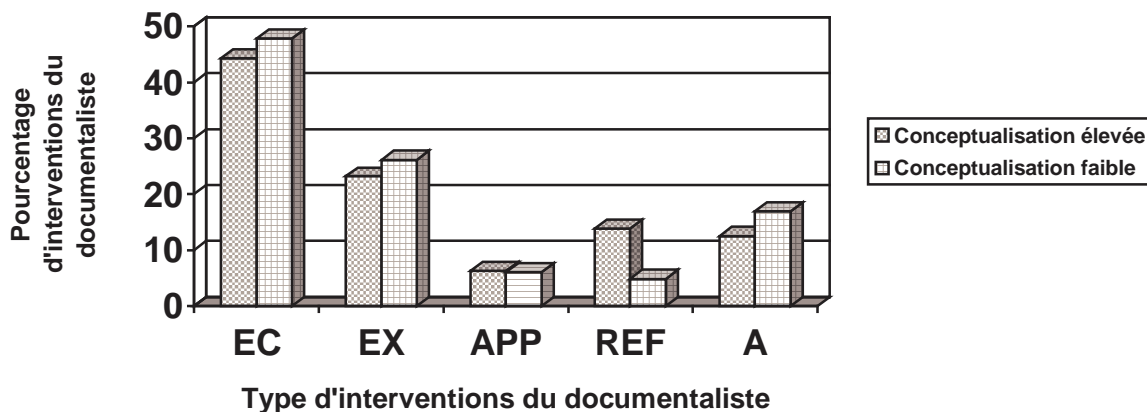


Hypothèse 3 : influence du degré de conceptualisation sur les interventions

Pour mettre à l'épreuve l'hypothèse 3, les protocoles ont été catégorisés selon le degré de conceptualisation de l'objet de la demande pour le demandeur lors de la séquence 1 de présentation initiale de l'objet de recherche. Selon Malrieu (1992), « l'effort de lexicalisation de l'objet d'étude est un indice de définition et d'explicitation des concepts centraux pour décrire le problème au destinataire ». En somme, la verbalisation de mots clés thématiques serait un indice du niveau de conceptualisation d'un objet. De cette façon, les séquences 1 des demandes pour lesquelles le pourcentage de mots ou expressions clés thématiques est supérieur au pourcentage de mots ou expressions clés contextuels, ont été classées dans la catégorie « haut degré de conceptualisation » (les protocoles SB2-MG2-CI1). Les protocoles pour lesquels l'inverse est constaté (mots clés contextuels supérieurs en nombre aux mots clés thématiques) sont catégorisés « faible niveau de conceptualisation » (les protocoles SB1-SB3-MG1).

Le Graph3 rend compte du **pourcentage de chaque type d'interventions du documentaliste** (EC : écoute, EX : explicitation, APP : apport d'information, REF : reformulation et A : autres types d'intervention) **en fonction du degré de conceptualisation de la demande identifié à partir de la séquence 1.**

Graph3 Conceptualisation et Interventions de la séquence2



On observe que le nombre total d'interventions du documentaliste est plus important lorsque le degré de conceptualisation est plus élevé (137 interventions) que lorsque qu'il est plus faible (101 interventions). Plus particulièrement, les interventions pour lesquelles leur nombre est plus important en modalité de conceptualisation élevée sont du type Ecoute et Reformulation. A contrario, les interventions du type Explicitation, Apport d'information et Autres sont plus nombreuses pour les demandes caractérisées par un faible degré de conceptualisation.

Interprétation des résultats

Hypothèse 1

L'hypothèse 1 supposait que le documentaliste adopte deux postures différentes durant l'échange. Une première posture d'écoute est adoptée : les interventions de type Ecoute sont majoritaires dans la première partie de l'entretien. Une seconde posture dans laquelle le documentaliste, en référence à ce qu'il a perçu dans la première partie de l'entretien et à ce qu'il en a compris, interroge le demandeur pour affiner sa compréhension de l'objet (les interventions de type Explicitation, Apport d'information, Reformulation sont plus nombreuses dans la seconde partie de l'entretien). La première posture Ecoute semble sous-tendue par un processus de type ascendant (bottom-up) : la construction de sens se développe et prend forme à partir d'éléments perçus ; et la seconde posture sous-tendue par un processus de type descendant (top-down) : le développement de la signification se prolonge mais cette fois-ci le documentaliste y participe de manière dynamique. En effet, il semble que dans sa démarche de compréhension, le documentaliste construit dans un premier temps une représentation directement liée au discours du demandeur et dans un second temps, en fonction de l'adéquation de ses connaissances avec celles du demandeur, il semble qu'il tente d'obtenir de nouvelles informations thématiques ou contextuelles (selon la présentation du demandeur) pour affiner sa compréhension de la demande.

Hypothèse 2

L'hypothèse 2 supposait que l'interaction facilite l'émergence d'attributs pertinents pour la compréhension de la demande. Les résultats montrent cependant que, pour l'ensemble des entretiens, ce sont les interventions de type Ecoute qui sont le plus souvent suivies de mots ou expressions clés, notamment contextuels. Ce résultat suggère l'écoute acquiescante («oui», «d'accord», «ok», ...) favorise l'émergence d'attributs significatifs. Cependant, l'écoute aurait tendance à conforter le demandeur dans une explicitation contextuelle de son objet de recherche, c'est-à-dire dans la production d'anecdotes situationnelles (« c'est une fille qui a 38 ans », « faire comprendre à la direction marketing

quel est leur job », ...). Ces éléments contextuels sont essentiels pour la compréhension de l'objet mais restent difficiles à transformer en attributs essentiels, directement réutilisables au cours de l'activité de recherche d'information. En outre, les résultats suggèrent des stratégies interindividuelles distinctes de conduites d'entretien à la fois pour le documentaliste et pour le demandeur.

Hypothèse 3

La mise en place de stratégie interindividuelle différente tant dans la conduite d'entretien que dans la formulation de la demande semble être confirmée par les résultats relatifs à l'hypothèse 3. En effet, il semble que le degré de conceptualisation de la formulation de la demande influence la manière dont le documentaliste intervient (un haut degré de conceptualisation entraînerait des interventions du type Ecoute et Reformulation / un faible degré de conceptualisation entraînerait des interventions du type Explicitation, Apport d'information). Cependant le nombre de protocoles analysés étant réduit, nous ne pouvons pas tirer de conclusions plus précises. Il semble néanmoins que la forme de l'interaction soit, pour le demandeur d'une part, dépendante du degré de conceptualisation de l'objet et de sa capacité à verbaliser des connaissances, et pour le documentaliste d'autre part, dépendante des préférences stratégiques en matière de conduite d'entretien et de l'adéquation de son niveau de connaissance avec celui du demandeur.

Discussion générale

Comme dit précédemment, les protocoles analysés sont trop peu nombreux pour pouvoir tirer des conclusions généralisables à d'autres situations similaires. Cependant les résultats suggèrent que la construction de sens dépend non seulement des connaissances des deux intervenants mais aussi de l'interaction qu'ils mènent. On peut supposer que l'interaction et les négociations qui s'en suivent pour comprendre le sens de l'objet de recherche influencent la représentation même de l'objet de recherche. Dans le cas d'un objet de recherche peu conceptualisé ou dont la représentation reste floue, l'interaction contribuerait à en préciser les contours. La signification de l'énonciation prend sa première forme à l'écoute puis se développe et se construit dans l'interaction entre le documentaliste et le demandeur. La médiation par l'interaction est dépendante dans la forme qu'elle prend à la fois des connaissances de l'une et l'autre partie (résonance de l'environnement de connaissance) mais surtout de la forme verbale que prennent les connaissances (de leur actualisation). En d'autres termes, le même demandeur peut formuler sa demande de manière différente selon la personne avec laquelle il interagit et vice-versa.

Les approches socioconstructivistes (notamment celle de Doise et Mugny, 1997) ont développé une approche sociale du développement cognitif dans

laquelle le contexte interactionnel tient une place importante. On le voit ici, l'interaction permet d'alimenter le développement du sens de l'information perçue. Le sens ne se construit pas seulement dans le traitement de cette information mais également dans la prise en compte du contexte de l'information. Selon Bruner (1991), la signification ne réside pas dans l'information. Il considère en effet que la métaphore dominante de l'ordinateur proposée par les sciences cognitives, pour comprendre le fonctionnement humain ne propose plus de comprendre comment se construit la signification mais comment on traite l'information. Selon Bruner, « *l'information ne s'intéresse pas à la signification. En termes informatiques, l'information contient un message précodé dans le système. Le sens précède le message. Il n'est pas produit par l'ordinateur, il n'en relève nullement. Le traitement de l'information ne consiste qu'à inscrire des messages à une adresse dans la mémoire, où aller les y chercher, selon des instructions transmises par une unité centrale (...). L'information précodée est rangée, ordonnée, combinée, comparée. Le système qui gère l'ensemble est indifférent à la nature de ce qui est stocké.* ».

Une analyse discursive plus fine d'un plus grand nombre de protocoles (analyse des séquences d'interaction, par exemple) permettrait de proposer des éléments plus précis sur ce genre de questions.

Les suites de ces travaux consisteront en l'analyse de l'influence de la construction du sens d'un besoin informationnel sur les stratégies d'interrogation de bases informationnelles du point de vue des processus cognitifs mis en jeu.

Bibliographie

Baker, M, Dessalles, J-L, Joab, M., Raccach, P-Y, Safar, B. & Schlienger, D. (1994), Analyse et modélisation d'explications négociées : le cas de la recherche d'informations documentaires,

<http://www.vjf.cnrs.fr/umr8606/FichExt/mbaker/publications/ArticlesBakerPDF/1994/1994EtAl-a.pdf>

Bruner, J. (1991), ... *Car la culture donne forme à l'esprit : de la révolution cognitive à la psychologie culturelle*, Paris : Editions Eshel.

Changeux, J-P., Ricoeur, P. (1998). *Ce qui nous fait penser, la nature et la règle*, Paris : Editions Odiles Jacob.

Doise, W. et Mugny, G. (1981), *Le développement social de l'intelligence*, Paris : Interéditions.

Le Ny, J-F. (2005), *Comment l'esprit produit du sens*, Paris : Editions Odile Jacob.

Malrieu, D. (1992), Les apports d'une étude différentielle de la demande bibliographique pour la modélisation des utilisateurs, *Intellectica*, 15, 187-214.

Annexe

Transcription d'une séquence 1

| SIGNIFICATION | SYMBOLE |
|--|---|
| Interventions inaudibles | _____ |
| Interruptions causées par un autre intervenant | / |
| Courts silences de moins de 15 secondes | ... |
| Longs silences de plus de 15 secondes | |
| Chevauchements de paroles | les paroles concernées par le chevauchement sont soulignées |

Codifications des transcriptions

| INTERVENANTS | VERBALISATIONS |
|----------------|--|
| Demandeur | Oui euh nous travaillons sur euh un projet pour un client qui concerne les machines à café |
| Documentaliste | M |
| Demandeur | expresso filtre et autres euh dans les circuits de la branche de la grande distribution et euh en fait ce que j'aurai voulu euh essayer de trouver c'est euh dans dans les revues de la grande distribution du type euh ELSA Point de Vente Facing etc |
| Documentaliste | mm |
| Demandeur | si il y a eu des articles publiés sur ce thème que ce soient des articles parlant du marché de sa segmentation euh de des expositions des produits euh etc etc etc |
| Documentaliste | donc euh qu'on soit bien d'accord c'est donc les fabricants de machine à café et de filtres euh type euh Espresso ce genre de choses ? |
| Demandeur | Espresso Tasibo Senseo etc etc etc c'est euh en grande distribution donc c'est dans l'univers du petit électroménager |
| Documentaliste | mm |
| Demandeur | donc il y a peut-être eu des études publiées des merchandising sur le petit électroménager |
| Documentaliste | <u>d'accord ça ça peut-être intéressant</u> |
| Demandeur | <u>donc ça ça m'intéresse également</u> |
| Documentaliste | d'accord |
| Demandeur | mais plus spécifiquement ce qui concerne donc euh tout ce qui est machine à café au sens très très large depuis la la machine à café hyper traditionnelle avec euh euh comment ça s'appelle le filtre euh etc etc jusqu'à la machine espre/ euh expresso |
| Documentaliste | d'accord |
| Demandeur | hein ? Donc euh l'univers enfin l'univers le marché disons de la machine à café |
| Documentaliste | <u>d'accord</u> |
| Demandeur | <u>hein ? voilà</u> |

PREPARATION D'UNE EXPERIMENTATION PEDAGOGIQUE EN VUE D'OBSERVATION

Thibault Carron, Jean-Charles Marty, Jean-Mathias Heraud*, Laure France*

Université de Savoie

Laboratoire SYSCOM, équipe Erte Cartable électronique

Bât. Mont-Blanc

73 376 Le Bourget du Lac Cedex

**Ecole Supérieure de Commerce de Chambéry*

prenom.nom@univ-savoie.fr

RESUME

Le travail présenté dans cet article concerne le domaine de l'éducation et plus spécialement l'observation d'une activité pédagogique. Le comportement de l'utilisateur est observé au moyen de traces brutes laissées par son activité et provenant de plusieurs sources. Nous décrivons dans ce papier une étape préalable importante mais généralement peu étudiée ou décrite : la préparation d'une expérimentation observée en sélectionnant et en structurant les critères ou éléments d'observation pertinents. Une proposition de formalisation basée sur les expressions régulières et des langages de description des différentes sources est présentée et illustrée par un exemple.

MOTS CLES

Traces, Expérimentation pédagogique, Observation, Collecte, Formalisation

1. INTRODUCTION

Pour satisfaire au mieux les utilisateurs finaux, les concepteurs de logiciel cherchent à comprendre précisément les besoins des utilisateurs. La tendance actuelle est d'observer et analyser l'activité de l'utilisateur afin de mieux comprendre son comportement [Cooley, 99] [Burton, 01] [Iksal, 05]. Dans le cadre des environnements d'apprentissage en ligne, de tels systèmes ont été développés à la fois par des concepteurs et des enseignants. De nombreuses expérimentations ont été montées avec ces nouveaux systèmes de travail collaboratif [Corbières, 04] [Héraud, 05].

Dans ce contexte, les enseignants d'une part, ont besoin de comprendre l'activité soit en cours d'exécution (notamment dans les expériences d'apprentissage à distance) soit a posteriori (dans une démarche de « débriefing » de ce qu'il s'est passé pendant la séance); les analystes d'autre part, ont besoin de tirer des leçons de ces expérimentations et enfin les étudiants d'avoir des retours sur l'activité

(awareness) pendant leur processus d'apprentissage collaboratif. L'activité d'observation devient ainsi cruciale pour répondre à ces attentes.

Initialement, notre base de travail est constituée des traces représentant le monde dans lequel nous nous situons. Cependant, le nombre de traces est trop important et celles-ci sont difficilement compréhensibles de manière directe, ce qui rend ce monde très difficile à interpréter. Nous avons donc besoin de réduire notre champ d'investigation. Pour cela, il nous est nécessaire de répondre à plusieurs questions classiques : Que doit-on observer ? Où ? Comment ? La formalisation devrait nous permettre d'atteindre ce but.

Une fois le champ d'investigation réduit, nous nous intéressons aux moyens de faire parler les traces. Il s'agit alors de chercher les indices pour décoder une partie des messages qui se cachent à l'intérieur de l'ensemble des traces.

Néanmoins, notre travail ne consiste pas tant à résoudre l'énigme, mais à donner des outils pour aider l'enquêteur, outils qui s'adapteraient à l'énigme. Pour s'adapter, il faut savoir pourquoi on observe.

But général de l'observation

D'une part, l'observation est utile pour la **régulation de l'activité**. Dans un environnement non médié, un enseignant essaie d'être aussi conscient de l'activité que possible, essaie de chercher des indices qui lui permettent de comprendre le niveau / l'état de compréhension d'un étudiant. L'enseignant adapte ensuite sa séance, par exemple en ajoutant des informations supplémentaires ou en gardant un exercice pour une autre session.

D'autre part, l'observation rapporte de nombreuses informations utiles à l'**analyse d'une séance d'apprentissage**, par exemple :

- pour comprendre pourquoi certains apprenants ont réussi une activité où d'autres ont échoué
- pour savoir comment les activités effectives s'appuient sur les préconisations de l'enseignant pour l'enchaînement des activités pédagogiques (scénarios préconisés) [Ferraris, 2005] ou simplement pour connaître la durée que chaque activité a nécessité
- lorsqu'un analyste veut voir si les apprenants ont suivi le scénario préconisé
- lorsqu'un apprenant veut analyser son cheminement pédagogique.

Nous remarquons que les caractéristiques de l'observation dans une expérience donnée sont très différentes en fonction de la diversité des objectifs et des rôles mentionnés ci-dessus. C'est pourquoi nous avons besoin de décrire précisément les attentes de l'observation pour une expérience donnée.

Dans cet article, nous nous focalisons sur la phase de préparation de l'expérience. Nous présentons tout d'abord ce que nous appelons une expérience observée, en donnant un rapide tour d'horizon de chaque étape classique pour

situer notre travail parmi ceux existants. Nous décrivons ensuite ce qui doit être préparé et nous proposons un formalisme pour accomplir cette étape. Nous appliquons enfin notre approche à un exemple. Les perspectives sont discutées en conclusion.

2. METTRE EN PLACE UNE NOUVELLE EXPERIMENTATION EN TENANT COMPTE DES ASPECTS OBSERVATION

A notre connaissance, il n'existe pas de modèle d'observation spécifique permettant de décrire tous les éléments à considérer. Cependant, on s'accorde généralement sur plusieurs étapes communes [Cooley, 99] (bien que les noms associés à ces étapes puissent différer).

2.1 Description du déroulement d'une expérience observée

De notre point de vue, nous considérons qu'une expérience observée peut être décrite de la manière suivante (cf. fig. 1) :

Phase de collecte (« recueil des indices sur le terrain »): Afin d'observer le processus d'une activité d'apprentissage, nous devons collecter des « indices » ou des traces sur l'activité en cours. En fonction d'une expérience donnée, seules certaines traces sont enregistrables. Nous appelons « observables » l'ensemble de toutes ces traces enregistrables. C'est pourquoi la phase de collecte consiste à collecter ces observables pendant l'expérience. Aucune interprétation n'est faite sur les données collectées pendant cette étape.

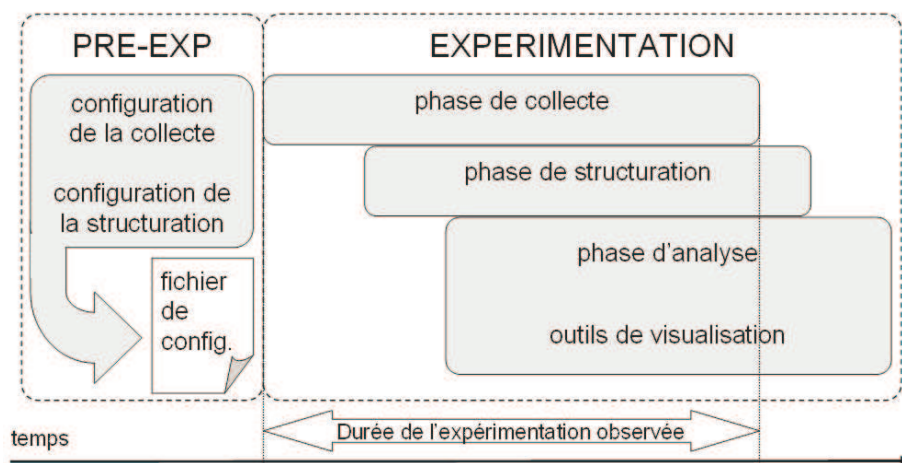


Fig. 1 Déroulement d'une expérience observée

Phase de structuration (« le calepin du détective »): La phase de collecte produit généralement une très grande quantité de données. La bonne compréhension de ces données passe par une bonne structuration. A ce stade, les données collectées sont regroupées, leur présentation est modifiée et il est possible de les annoter.

Phase d'analyse (« l'intelligence du détective aidée par les labos de l'Intelligence Service ») : La plupart du temps, cette étape est réalisée par l'analyste et ne peut être automatisée à cause du haut niveau de complexité de cette tâche. Plus la phase de structuration produit des traces au bon niveau de compréhension, plus leur interprétation en est facile. Des outils spécifiques de représentation ou de visualisation des données [France, 2005] peuvent être utilisés à cette étape.

La phase d'analyse est très difficile à formaliser puisque dans notre contexte, elle dépend trop de critères changeants ou mouvants : l'analyse peut prendre différentes « formes » en fonction de l'analyste. Cependant, toutes les analyses sont basées sur des traces spécifiques, éléments fiables de l'expérimentation : les observables. La question méthodologique à laquelle nous devons répondre lorsque nous mettons en place une expérimentation est : « comment représenter et structurer ce qui a été observé pour une analyse particulière ? »

Ces difficultés soulignent la nécessité d'une description claire de la préparation d'une expérience observée. Pour cette raison, nous nous focalisons dans cet article sur la phase d'initialisation de l'expérimentation (appelée **PRE-EXP**).

2.2 Une vue informelle de la préparation d'une expérience observée

Spécifions pour commencer de manière informelle les deux principales étapes permettant une « bonne » structuration des traces. La structuration sera correcte si l'analyste a toutes les informations nécessaires pour son objectif particulier et si ces informations sont assez compréhensibles pour être interprétées en accord avec l'objectif.

2.2.1 Sélection des observables

Comme nous l'avons décrit précédemment, il est nécessaire de spécifier exactement quelle information peut être observée. Les caractéristiques techniques doivent être prises en compte ici, ainsi que les problèmes éthiques. Une fois cette description obtenue, nous considérons que nous avons à notre disposition pour l'expérience des flux continus de données (flux d'observables), un pour chaque observable.

Un flux d'observable particulier peut être utile pour un objectif spécifique et inutile pour un autre [Marquardt, 04]. Par exemple, un flux d'oculomètre peut être très utile si l'objectif concerne la validation d'une interface utilisateur, mais inutile pour vérifier le degré de collaboration entre les étudiants pendant une session. Il est ainsi absolument nécessaire de « couper » certains flux lorsqu'ils ne sont pas pertinents pour une expérience donnée. C'est un moyen efficace d'éviter des traces non souhaitées parasitant les autres [Zaïane, 01] pendant la phase d'analyse.

2.2.2 Structuration des données collectées

Les données sont généralement collectées à un très bas niveau et dans un état que l'on peut qualifier de "brut", et sont donc difficilement compréhensibles [Héraud, 05]. A partir des logs, il est possible de dater précisément les actions, de les positionner les unes par rapport aux autres, d'en identifier l'auteur, la machine, et de nombreux autres paramètres. Ces données peuvent avoir différents formats, ne pas être dans le bon ordre, ou même être redondantes ou regroupées [Rossi, 05] [Tanasa, 04]. Par exemple, nous pouvons trouver de nombreux indices dans les logs sur l'utilisation d'un service de webmail. Si l'objectif d'observation est basé sur la fréquence d'accès au webmail par les utilisateurs, "connecting to webmail", "retrieving email", "opening a message", "answering a message" sont des éléments adéquats pour cet objectif mais trop précis. Dans ce cas, on peut regrouper tous ces éléments en une simple activité "accès au webmail" afin de synthétiser l'information.

De manière similaire, les données provenant de plusieurs sources d'observation - webmail, messagerie instantanée, VoIP ou fichier envoyé dans le casier¹ d'autrui - peuvent être regroupées sous le terme "outil de communication", plus approprié si nous voulons vérifier les aspects collaboratifs dans une activité d'apprentissage en récupérant l'auteur, le destinataire et le contenu si nécessaire. La même chose est envisageable à l'inverse pour retrouver des informations à d'autres niveaux de granularité (plus fins), par exemple pour savoir quelles activités non préconisées se sont produites entre le début et la fin d'un exercice comme l'accès à un forum.

Ces étapes de collecte et de structuration doivent aussi être préparées ou configurées dans la phase PRE-EXP. Utiliser un formalisme spécifique pour le faire permet de décrire précisément la partie observation de l'expérience. Ceci facilite une séparation claire entre la partie expérience et le côté observation. De plus, cette partie observation peut être réutilisée même partiellement pour une autre expérimentation quand un analyste a le même type d'objectif.

Nous allons proposer dans la partie suivante une formalisation de cette idée.

3. FORMALISATION

La plupart du temps, de nombreuses expériences sont mises en place à partir d'environnements numériques d'apprentissage spécifiques. L'ensemble complet des observables est ainsi généralement le même pour chaque expérimentation. [Champin, 04] définit un modèle d'utilisation qui précise tous les éléments observables du système. Néanmoins, il se révèle difficile et très coûteux en temps de modéliser (ou re-modéliser) les systèmes en spécifiant tous les observables disponibles et leurs relations. Nous souhaitons faciliter ce travail en proposant une phase préliminaire de formalisation.

3.1 Les Observables

¹ service offert par notre environnement numérique de travail collaboratif

3.1.1 Vocabulaire d'observables

Le vocabulaire des observables V_{obs} contient tous les éléments pouvant être observés pendant l'expérimentation. Les observables peuvent avoir différents niveaux de granularité en fonction des sources d'où ils proviennent. En effet, il est possible d'obtenir des logs ou traces brutes mais aussi des éléments de plus haut niveau, collectés à partir d'outils spécifiques. Un observable pourrait être l'état dans lequel se trouve un utilisateur dans un scénario pédagogique. Un outil spécifique peut ainsi fournir à l'utilisateur des informations explicites, telles que, par exemple, "*exercice achevé*", directement interprétables au niveau du scénario pédagogique.

V_{obs} peut ainsi inclure plusieurs autres vocabulaires dédiés à des catégories particulières d'observables permettant d'en faire une classification. Nous pouvons identifier ces derniers par sources, mais aussi par niveau d'abstraction, par sujet. Considérons par exemple, l'ensemble V_{sce} des caractérisations des activités du scénario. V_{sce} regroupe ainsi les activités pédagogiques telles que *exercice achevé*, *consultation de l'aide*, *validation d'exercice*, etc. V_{sce} peut être un sous-ensemble de V_{obs} .

C'est au concepteur de spécifier les catégories si cela est nécessaire. Le meilleur moyen est de prévoir et d'inclure des informations concernant l'observation lorsqu'un nouveau composant ou outil est intégré à l'environnement d'apprentissage.

La définition de l'ensemble des observables permet de sélectionner la « matière première » à étudier. Ainsi, cette dernière doit être collectée en accord avec des conditions particulières et en respectant certaines règles. Nous allons maintenant décrire cette sélection parmi les observables.

3.1.2 Sélection des observables

La sélection des observables peut être exprimée en extension ou en intension. La sélection en extension est faite en énumérant à partir de V_{obs} chaque élément qui doit être collecté, alors que la sélection en intension permet d'exprimer des contraintes sur les observables afin de trouver ceux qui conviennent.

Si V_{obs} est défini par les observables des différentes sources d'observation S_1 , S_2 et S_3 : $V_{obs} = V_{S_1} \cup V_{S_2} \cup V_{S_3}$

Plus généralement, avec n sources d'observation: $V_{obs} = \bigcup_{i=1}^n V_{S_i}$.

L'ensemble des "observés" (observables sélectionnés) est appelé : S_{obs} . S_{obs} est un sous-ensemble de V_{obs} et décrit l'ensemble de tous les observables qui doivent être collectés lors de l'expérimentation observée courante, quelques soient leur source ou leur granularité.

Comme décrit précédemment, les observables sélectionnés doivent être explicites pour devenir plus compréhensibles par l'analyste. L'ensemble des

observables doit être réécrit afin de les rendre compréhensibles : cette “phase de structuration” est réalisée en appliquant certaines règles sur les observables.

3.2 Structuration des observables

Nos règles sont basées sur des expressions régulières permettant la transformation d’observables, généralement de bas niveau (partie gauche de la règle), en une partie droite ayant un plus haut niveau de granularité, bien que différentes transformations puissent être envisagées. La partie gauche des règles est composée uniquement d’éléments de V_{obs} alors que la partie droite comprend des éléments aidant à une meilleure interprétation (éventuellement des éléments de V_{obs}) que l’on appellera V_{ana} (pour vocabulaire d’analyse).

Par exemple, un ensemble spécifique de trois observables (a, b, c) peut être identifié comme réalisant une action particulière complexe pour l’étudiant : répondre à un courrier électronique. Dans ce cas, d’autres observables peuvent apparaître entre le moment de la réception du courrier électronique, la consultation et la réponse. Il est ainsi possible de savoir si la plupart des étudiants ont communiqué par messagerie électronique pendant l’exercice en appliquant des règles de ce type sur l’ensemble des (éléments) observés.

- $a.*bc \rightarrow A$

D’autres règles doivent être définies pour renommer les observables, pour réduire la redondance ou encore pour changer le niveau de granularité et atteindre une signification correspondant à l’axe d’analyse prévu :

- $a_1 | a_4 \rightarrow B$ *renommage*
- $a_2 [^c1]*a_2 b_1 \rightarrow C$ *réduction de la redondance de a_2 si aucun c_1 entre les deux a_2*
- $a_1[b_2 | b_3] \rightarrow D$ *changement de niveau de granularité*
- ...

Le résultat de cette transformation est un texte structuré qui peut être plus facilement analysé, étant donné qu’il est uniquement basé sur des observables pertinents et qu’il est exprimé dans un vocabulaire adapté pour l’analyse. Ce vocabulaire définit un autre vocabulaire V_{ana} correspondant aux actions recherchées par l’analyste.

Il est bien connu en Intelligence Artificielle et en Logique que de telles techniques ont la particularité de donner potentiellement différentes réponses en changeant l’ordre des règles. Le résultat peut ainsi ne pas être un unique « texte structuré » mais plutôt un arbre de textes structurés. Cet arbre peut être exploité grâce à l’expertise de l’analyste pour mettre en lumière de nouvelles pratiques comme nous le suggérerons en conclusion. Néanmoins, toute exploitation des résultats dépend de la phase suivante, l’analyse en elle-même, et ne concerne pas directement cet article.

Nous proposons, dans la partie suivante, un exemple basé sur une expérimentation réelle dans laquelle nous nous étions intéressés à l'étude des nouveaux usages des outils de communication par les étudiants pendant une séance de travaux pratiques (TP) sur machine. La question est : « Comment préparer une telle expérimentation ? »

4. EXEMPLE : USAGE DES OUTILS DE COMMUNICATION PAR LES ETUDIANTS PENDANT UNE ACTIVITE PEDAGOGIQUE

4.1 Description de l'expérience

Notre expérimentation s'est déroulée avec des étudiants de 3ème année à l'Ecole Supérieure de Commerce de Chambéry. L'objet était une séance de travaux pratiques sur le logiciel Adobe Photoshop®. Les étudiants devaient suivre un scénario pédagogique préconisé sur une plate-forme d'apprentissage web. Toutes les ressources nécessaires (cours, images sources, etc.) ainsi que des outils de communication (webmail, forums, messagerie instantanée) étaient disponibles sur la plate-forme d'expérimentation. Chaque poste était appareillé via un logiciel de récupération des événements claviers et souris. Le serveur était équipé d'un module pédagogique appelé pscénario [TECFA, 04] permettant de mettre en place le scénario, et les accès serveur étaient stockés dans des fichiers de logs. Les étudiants avaient pour consigne de ne communiquer que via l'ordinateur. L'enseignant communiquait avec les étudiants grâce à un logiciel de messagerie instantanée et validait ainsi le travail réalisé. Un observateur humain était en charge des prises de notes sur le comportement des étudiants et de l'enseignant en utilisant une grille spécifique.

4.2 Préparation de l'expérience

Le but de cette expérience était d'analyser les usages des outils de communication pour accomplir un scénario pédagogique spécifique. La préparation de l'expérience consiste à sélectionner les observables à collecter. L'ensemble complet des observables (V_{obs}) est connu. La difficulté est évidemment de déterminer les éléments qui seront pertinents. Pour constituer les « observés » ou observables sélectionnés (S_{obs}), nous avons choisi d'une part V_{sce} et d'autre part, de manière extensive, des éléments spécifiques à partir de différentes sources concernant la communication :

$$S_{obs} = V_{sce} \cup V_{Forum} \cup V_{Instant_Messenger} \cup V_{Mail} \cup (V_{Dialogue} - \{\text{conversation sur problèmes techniques}\})$$

Plusieurs sources peuvent être sélectionnées, mais même dans un flux spécifique, certaines informations peuvent être retirées comme on le voit dans

l'exemple précédent : les éléments observables « conversations concernant des problèmes techniques » de la source externe « dialogue » sont retirés. En réalité, il est important d'avoir en tête la corrélation entre les observables (cf. fig. 2 pour les « feuilles » de l'arbre en format simplifié) et le but de l'analyse (ici, les outils de communication) comme le montre la figure 2. Cette figure nous aidera aussi à spécifier les règles pour élaborer la partie structuration.

L'expérience s'est ensuite déroulée en laissant « sur le terrain » les traces que nous avons trouvées pertinentes. Il faut à présent faire parler ces traces en exploitant les indices, l'énigme consistant à découvrir si le scénario préconisé a été suivi ou si l'utilisateur a obtenu son résultat grâce à d'autres activités.

Voici quelques exemples simples de règles d'aide à la recherche de la solution.

- login | send | view | logout → FORUM ❶
- open | IMsend | IMview | close → IM ❷
- receive | eMsend | write | read → E-MAIL ❸
- conversation → DIALOGUE ❹

- FORUM | IM | E-MAIL | DIALOGUE → COM ❺

- (DebEx1) * (FinEx1) → EXERCICE1_ACHEVE ❻

- (DebEx1) [^FinEx1] (DebEx2) → EXERCICE1_ECHOU ❼

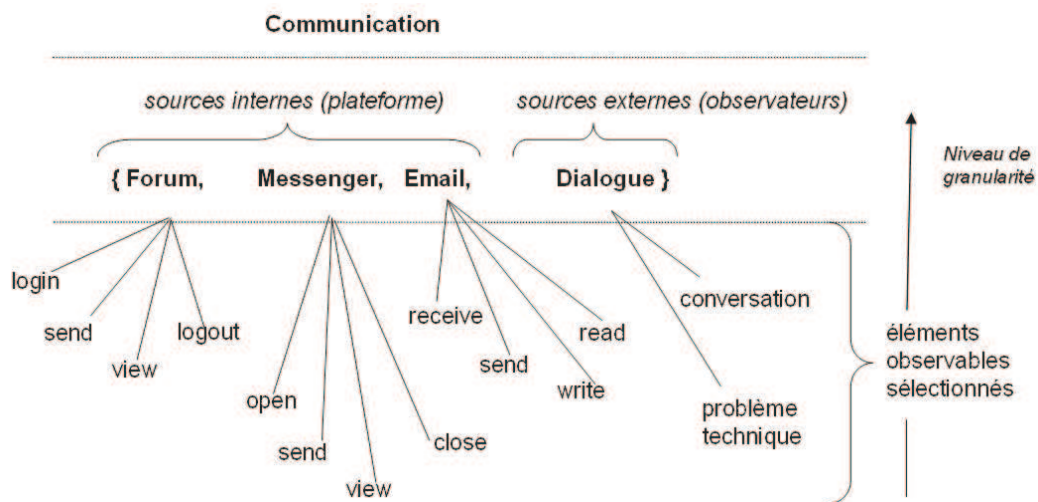


Fig.2 Structuration des observables

Les règles ❶ à ❹ sont directement issues de la définition en extension de V_{Forum} , $V_{\text{Instant_Messenger}}$, V_{Mail} et V_{Dialogue} .

Naturellement, les dernières règles (❺, ❻, ❼) nous amènent au plus haut niveau d'abstraction, rendant très facile l'interprétation des résultats ; les règles 6 et 7

produisant même directement des termes du vocabulaire V_{scc} . Cependant, ces facilités font parfois perdre de l'information qui aurait pu offrir une possibilité de description plus précise de la communication.

4.3 Quelques mots sur les résultats

Le résultat obtenu est un ensemble de textes structurés, un pour chaque étudiant, semblables à celui qui suit (niveau de granularité élevé) :

```
LECTURE_INSTRUCTIONS, EXERCICE1_ACHEVE, EXERCICE2_ECHOUE,  
COM, LECTURE_SUPPORT, DEMANDE_AIDE, COM, EXERCICE2_ACHEVE,  
COM
```

Ce résultat permet d'avoir une vue globale de l'expérimentation, compréhensible sans connaissance particulière, et permet ainsi de pouvoir choisir d'explorer un aspect particulier qui nous intéresse. Par exemple, en observant le temps nécessaire à EXERCICE1_ACHEVE (non représenté ici mais très facile à obtenir puisque les logs sont estampillés), on s'aperçoit que le temps mis pour cet exercice est anormalement long. Ceci est une nouvelle énigme, mais il se trouve que cette fois-ci les traces exprimées sont de trop haut niveau (c'est le détail qui nous intéresse ici).

A partir du moment où l'information obtenue à un niveau de granularité donné ne nous suffit pas pour comprendre ou analyser, il nous est possible d'obtenir des informations plus détaillées en regardant à un niveau de granularité inférieur. Ainsi, en revenant sur l'application des dernières règles (⑥ et ⑦), il est possible de retrouver ce qui s'est passé au niveau de l'exercice 1 :

```
EXERCICE1_ACHEVE = (DebEx1) (LECTURE_INSTRUCTIONS) COM  
(LECTURE_SUPPORT) COM (FinEx1)
```

Ceci permet à un niveau de granularité plus fin d'en déduire que la lecture du support a permis d'aider à trouver la solution.

Cependant, si l'on continue à descendre dans la granularité pour avoir des informations plus précises, on peut regarder ce qui se cache derrière les actions de communication COM. On trouve alors FORUM.send et FORUM.view, nous aiguillant ainsi vers une autre interprétation possible : l'étudiant a posté une demande sur un forum et a attendu de recevoir la réponse pour terminer l'exercice.

Grâce à cette transformation des traces de l'activité, l'expérimentation a pu être interprétée par des non-spécialistes en informatique. Un travail en collaboration avec des chercheurs en Sciences Humaines et Sociales (ici les analystes) nous a donné des résultats intéressants sur l'utilisation des outils de communication : leur utilisation semble être corrélée avec des difficultés (axe d'analyse pédagogique) ou à des besoins de validation (axe évaluation) ou encore à un

besoin d'un outil de vérification (axe re-engineering). Plus généralement, ces résultats nous ont permis de montrer des complémentarités entre les technologies éducatives assistées par ordinateur et les pratiques pédagogiques traditionnelles [Chabert, 05a] [Chabert, 05b]. D'autres résultats sur cette même expérience, permettant d'améliorer le scénario préconisé, sont exposés dans [Héraud, 05].

5. CONCLUSION

Dans cet article, nous nous sommes intéressés à la phase de préparation d'une expérience pédagogique observée. Une première formalisation indépendante du domaine, suffisamment simple pour être mise en œuvre par des non-spécialistes, est proposée pour collecter et structurer les éléments observables. Cette phase est cruciale en vue d'une analyse ultérieure spécifique. Comme nous l'avons montré, l'observation de l'activité devient aussi un point clé pour les enseignants, les analystes ou les étudiants. Une bonne description de ces attentes permettra une configuration correcte des outils d'observation, rendant ainsi possible une observation très précise de l'activité collaborative. Nous pensons ainsi qu'une expérimentation observée doit être très précisément préparée lors d'une phase que l'on appelle « PRE-EXP ».

L'expérience acquise sur les précédentes expérimentations peut ensuite être réutilisée pour la mise en place d'une nouvelle expérimentation à observer. Les perspectives envisagées concernent le développement d'outils et de modèles pour indexer les fichiers de la phase PRE-EXP et pouvoir ainsi capitaliser l'expérience en permettant notamment de choisir correctement l'ensemble des éléments à observer.

De plus, le langage de sélection des observables doit être enrichi afin de le rendre plus puissant, permettant par exemple, d'explicitier des conditions spécifiques sur les observables ou des corrélations entre des observables de différentes sources. Il peut ainsi s'avérer intéressant d'ajouter des contraintes temporelles afin de ne collecter qu'une information spécifique pendant une période déterminée. Plus généralement, notre travail doit permettre d'étudier plusieurs aspects de l'observation dans un environnement numérique de travail (ENT) afin de développer une station d'observation adaptée à l'ENT.

Ce travail ouvre de nombreuses perspectives en relation avec des sujets sensibles de ce domaine. Il est envisageable d'explorer cette méthode afin d'évaluer à la fois les étudiants, les apprentissages ou expérimentations d'un point de vue pédagogique ou même les infrastructures en s'orientant vers une approche qualité [Marty, 04]. Plus généralement, notre travail souligne que l'Observation est une activité en elle-même, et nécessite donc d'être considérée

et prise en compte comme une étape spéciale avec préparation, possibilité de réutilisation et évaluation.

REFERENCES

- Burton, M.C. and Walther, J.B., 2001. A Survey of Web Log Data and their Application in Use-based Design. *Proceedings of the 34th Annual Hawaii International Conference on System Sciences*. Maui, Hawaii, pp. 5026.
- Chabert. G. et al., 2005a. The Electronic Schoolbag: a CSCW workspace. Presentation and evaluation. To be published *In AI & Society: the journal of human centered Systems and Machine Intelligence*, 18 pages.
- Chabert G. et al., 2005b. L'observation des usages dans un contexte éducatif : pratiques pédagogiques et nouvelles formes de communication éducative émergentes ?, Colloque international ISD'05, 14-16 April, Tunis, Tunisia.
- Champin, P.A. et al., 2004. MUNETTE: a Framework for Knowledge Capture from Experience. *12ème Atelier de Raisonnement à Partir de Cas*, Villetaneuse, France.
- Cooley, R. et al, 1999. Data Preparation for Mining World Wide Web Browsing Patterns. *Knowledge and Information Systems*, Vol. 1, No. 1, pp. 5-32.
- Corbière A. and Choquet C., 2004. Re-engineering method for multimedia system in education. *Proceedings of the IEEE Sixth International Symposium on Multimedia Software Engineering*. Miami, USA, pp. 80-87.
- Ferraris et al., 2005, « Modélisation de scénarios d'apprentissage collaboratif pour la classe » Conference EIAH'05, Montpellier, France.
- France L. et al., 2005. Help through visualization to compare learners' activities to recommended learning scenarios. To be published *In Proceedings of IEEE International Conference on Advanced Learning Technologies*, Kaohsiung, Taiwan.
- Héraud J.M. et al., 2005 . Helping the interpretation of Web Logs: Application to Learning Scenario Improvement, *AIED workshop "Usage Analysis in Learning Systems"*, Amsterdam, Netherlands.
- Iksal S., Choquet C., 2005, *An Open Architecture for Usage Analysis in a E-Learning Context*. In: ICAIT 2005, 5-8 July 2005, Kaohsiung (Taiwan).
- Marquardt C.G. et al., 2004. A Pre-Processing Tool for Web Usage Mining in the Distance Education Domain. *Proceedings of the International Database Engineering and Applications Symposium*. Coimbra, Portugal, pp. 78-87.
- Marty JC. et al., 2004. A quality approach for collaborative learning scenarios. *Learning Technology Newsletter of IEEE Computer Society*, Vol. 6, Issue 4, pp 46-48.
- Rossi F. et al., 2005. Visualisation de la perception d'un site web par ses utilisateurs. *Actes de la conférence EGC 2005*. Paris, France, pp. 563-574.
- Tanasa D. and Trousse B., 2004. Advanced Data Preprocessing for Intersites Web Usage Mining, *IEEE Intelligent Systems*, Vol. 19, No.2, pp.59-65.
- TECFA 2004, pscenario, TECFA's Pedagogical Scenario Tool for PostNuke, <http://tecfaseed.unige.ch/door/>.
- Zaïane O. R. and Luo J., 2001. Towards evaluating learners' behaviour in a web-based distance learning environment. in *Proceedings of IEEE International Conference on Advanced Learning Technologies*. Madison, USA, pp. 357-360.

L'énigme de l'articulation entre comportements individuels et collectifs dans les simulations multi-agents

Frédéric Amblard^{*}, Antoine Penciolelli^{**}

[*frederic.amblard@univ-tlse1.fr](mailto:frederic.amblard@univ-tlse1.fr), [**antoine@penciolelli.net](mailto:antoine@penciolelli.net)

1.Introduction

Nous soulèverons ici le problème de l'émergence du sens, dont on souhaiterait qu'elle puisse devenir une construction du sens, au sujet des simulations multi-agents ou plus largement des simulations individus-centrées. Ce type de simulation se focalise sur une spécification souvent exclusive des comportements des individus ou agents et de leurs interactions, entre eux ou avec leur environnement. Les effets observés au niveau macroscopique en sortie de simulations posent dès lors une double question au modélisateur : 1) *qu'est-ce que j'obtiens ?* ou comment qualifier, un peu plus précisément qu'en les qualifiant d'émergents, les résultats au niveau macroscopique de comportements et de dynamiques microscopiques et 2) *pourquoi j'obtiens ces résultats ?* ou, nous le verrons plus en détail par la suite, quelle explication est-ce que je mets en place qui me satisfasse pour comprendre l'articulation entre les comportements individuels spécifiés et leurs résultants macroscopiques.

1.1. *Quels résultats j'obtiens ?*

La stochasticité de ces modèles aussi bien que leur grand nombre de variables (paramètres globaux, distributions utilisées pour initialiser les ensembles d'individus et éventuellement corrélations de ces distributions à des positionnements spatiaux ou topologiques dans le cas des réseaux sociaux), conduisent en pratique à la non-exhaustivité de toute exploration qui serait tentée sur ces modèles. On pourrait évoquer la trop grande dimension de l'ensemble des possibles. Au niveau macroscopique, quand on s'intéresse alors aux formes générées par le modèle, il faut donc se résoudre à considérer que les observations réalisées ne peuvent prétendre à l'exhaustivité. Le manque de preuves formelles liées aux simulations multi-agents et l'inutilisabilité de méthodes de vérifications formelles (de type model-checking) nous rapproche alors irrésistiblement des sciences expérimentales pour caractériser le comportement de notre modèle, comme argumenté par Guillaume Deffuant (Deffuant et al., 2003). Aussi nous devons nous contenter d'un processus de construction du sens fluctuant au grès des observations réalisées, ces observations dépendant 1) *de ce que l'on mesure*, et en ceci fait écho à notre première question, et 2) *des valeurs effectives de ces mesures*.

Notre perplexité augmente encore d'un cran à la lecture de Bergson concernant la création du possible par le réel : « *Au fur et à mesure que la réalité se crée,*

imprévisible et neuve, son image se réfléchit derrière elle dans le passé indéfini ; elle se trouve ainsi avoir été, de tout temps, possible ; mais c'est à ce moment précis qu'elle commence à l'avoir toujours été, et voilà pourquoi je disais que sa possibilité, qui ne précède pas sa réalité, l'aura précédée une fois la réalité apparue. » De manière presque similaire dans le cadre des simulations individus-centrées, les comportements macroscopiques observés sont rendus possibles *ab initio* par leur observation. Nous devons cependant concéder qu'une différence principale concerne le déterminisme de la simulation (la stochasticité n'étant au final qu'émulée par des générateurs pseudo-aléatoires). Cependant, et pour renforcer ce que l'on avait déjà énoncé, l'ensemble des possibles concernant les résultats de la simulation ne peut être dans notre cas ni énuméré (ensemble possiblement infini), ni borné (pratiquement les techniques de model-checking qui pourraient le permettre ne sont pas possibles à mettre en place), on doit donc souvent s'en remettre aux statistiques des expérimentateurs. Au delà de cet ensemble de sorties de simulation non quantifiable, sa caractérisation pose également problème, on pourrait même considérer que le type de mesure ou d'indicateur choisi pour caractériser l'observation multiplie d'autant l'ensemble de ces sorties. Mais nous reviendrons sur le point de la définition ou de la caractérisation de ce qui est observé après avoir abordé la deuxième question, la construction d'une explication sur le comportement du modèle co-évoluant avec les indicateurs choisis pour observer les simulations. Les indicateurs interviennent ainsi en pourvoyeurs de fondements pour l'explication et l'explication en construction orientant le type de mesures à réaliser sur la simulation, soit en imposant son cadre, soit en cherchant à l'éviter (pour chercher, la réfutation, le contre-exemple).

1.2. Pourquoi j'obtiens ces résultats ?

Pour répondre à cette question qui oscille entre construction instrumentalisée et émergence du sens, le modélisateur, au regard des observations qu'il fait de la simulation (on remarquera au passage que ces observations de départ orientent sans doute déjà l'explication qui sera élaborée), va construire progressivement une théorie liant ce qu'il observe. Il va construire une théorie de l'articulation entre niveau individuel et niveau collectif, une théorie pour abstraire, généraliser les simulations particulières qui ont été réalisées. Cette théorie correspond bien à une explication construite, possiblement formalisée, qui tend à s'abstraire du cadre du modèle dans lequel l'explication a été construite pour s'appliquer de manière générique à des classes de modèles équivalents. Sans volonté de construction de cette théorie, on se contente de considérer le modèle comme une boîte noire et d'essayer de construire un lien, dès lors équivalent à un lien statistique, liant entrées et sorties de la boîte, on reste alors sur l'attitude descriptive proposée par les lois physiques, sans proposer d'outils génériques de compréhension des phénomènes, de théories de l'articulation entre

comportements individuels et comportements collectifs à même de produire au niveau du modèle particulier des explications.

Il serait décevant de s'arrêter à ce stade, le modélisateur doit pouvoir, on le sent, pousser plus loin. Les résultats macroscopiques observés ne sont pas émergents, du moins au sens de Morin pour qui « *l'émergence est irréductible phénoménalement et indéductible logiquement* ». Malgré sa longueur, une chaîne logique, un algorithme déterministe, mène dans notre cas des conditions initiales au résultat observé. Même si cette chaîne est complexe, elle existe, *c'est écrit ...*

Un point que nous devons mentionner même si nous ne le discuterons pas, c'est qu'à un moment donné de l'étude, le sens émerge, au sens où l'on est capable de proposer une première explication, ou une nième explication, en accord avec les observations réalisées... Mais si l'on ne sait pas comment émerge ce sens, il nous faut essayer d'instrumentaliser cette émergence pour qu'elle ait davantage de chances de survenir et que l'on soit capable de proposer une explication.

La dynamique de construction de cette explication correspond du reste au schéma classique (Popper, 1972) « *seule l'expérience [dans notre cas l'expérience de simulation] est capable de nous aider à nous faire une idée sur la vérité ou la fausseté des énoncés portant sur des faits* ». Ainsi, si mon explication rend compte des observations réalisées jusqu'à présent, je peux la conserver (éventuellement parmi d'autres), sinon il me faut en construire une nouvelle. Cette réfutabilité est d'autant plus intéressante en simulation qu'il est parfois beaucoup plus réalisable de chercher la réalisation d'un contre-exemple par un processus de simulation que dans les sciences expérimentales classiques. Orienter l'algorithme de simulation pour trouver ce qui ne vérifie pas la théorie, à défaut d'être systématiquement réalisable est tout de même pratiquement envisageable dans certains cas.

1.3. Co-construction de l'explication et de l'observation

Comme nous l'avons souligné, l'indépendance entre observation et construction du sens, posée par les deux questions est évidemment factice. Il nous faut cependant mentionner que comme dans la plupart des domaines scientifiques, il y a interdépendance entre construction du sens et observation empirique *in silico*. Les observations appuyant ou réfutant les théories existantes, voire suggérant de nouvelles explications, et l'explication orientant par le cadre qu'elle impose ou le paradigme dont elle est issue, les mesures qui sont réalisables.

Pour illustrer notre propos, nous allons maintenant présenter l'exploration d'un modèle particulier en insistant rétrospectivement sur la manière dont se sont co-construits les explications sur ce modèle et ses observations.

2. Au départ était ... le modèle ?

Le départ n'est en fait bien évidemment pas le modèle formalisé tel que nous allons le présenter. Il a été précédé par presque trois ans de réflexions sur la manière dont les agriculteurs adoptent des mesures agri-environnementales et en particulier quels en sont les déterminants et quels sont les mécanismes sociaux sous-jacents qui expliquent la diffusion ou non de ces adoptions dans une population. Ce travail a été mené dans le cadre du projet européen IMAGES¹. Le modèle résultant de ce projet étant particulièrement volumineux (en terme de nombre de variables comme de lignes de code), nous nous sommes focalisés un peu après sur une partie de ce modèle, la dynamique d'influence sociale, que nous allons présenter et que nous avons en quelque sorte décontextualisée pour la considérer comme un objet à étudier.

Nous considérons donc une population de N individus, chaque individu i étant caractérisé par deux variables réelles : son opinion x_i et son incertitude u_i . Les opinions sont initialisées à partir d'une distribution uniforme sur $[-1;+1]$ et les interactions par couples d'individus sont sélectionnées aléatoirement parmi la population. Soit $s_i = [x_i - u_i, x_i + u_i]$, le segment d'opinion de i , le changement d'opinion de j sous l'influence de i est proportionnel à la superposition entre leurs deux segments d'opinions (l'accord), divisée par l'incertitude de i . Le changement d'opinion est de plus conditionné par un accord suffisant entre les deux individus. Si les deux individus sont suffisamment proches en opinion, ils s'influencent et se rapprochent davantage, sinon ils sont indifférents l'un à l'autre et ne modifient pas leurs opinions. Le lecteur pourra se référer à (Deffuant et al., 2002) pour plus de détails.

En considérant ce modèle, et avant même d'en exécuter une simulation, déjà, nous sommes en mesure d'anticiper ce qui peut s'y produire. Si les individus ont deux tendances : se rapprocher ou ne pas changer d'opinion, l'on devrait observer dans la simulation la formation de clusters d'opinions. Ce qui, en exécutant la simulation est avéré (cf. fig.1).

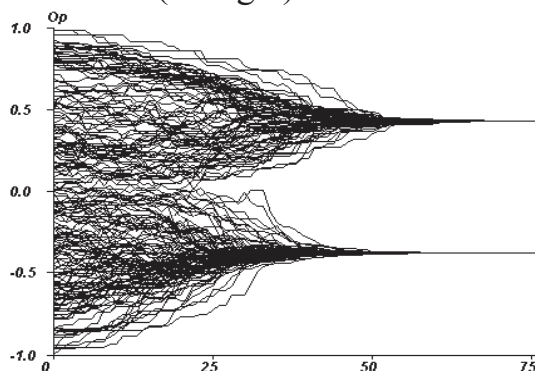


Figure 1 : Trajectoire des opinions des individus dans le temps avec une incertitude identique pour l'ensemble de la population (ici $u=0.4$ et $w=2.0$).

¹ Nous tenons en particulier à remercier ici Guillaume Deffuant et Gérard Weisbuch qui ont dirigé cette recherche.

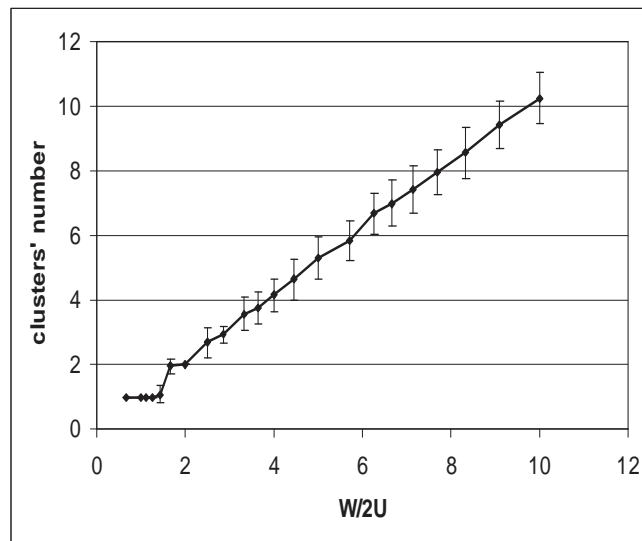


Figure 2 : Tracé du nombre moyen de groupes en fonction de $w/2u$ (w est la largeur de la distribution initiale et u l'incertitude des individus). Chaque point représente le nombre final de groupes moyennés sur 50 simulations (les individus isolés n'étant pas considérés comme des groupes). Le nombre de groupes moyen est proche de $w/2u$ ($r^2=0.99$).

L'explication qualitative fournie précédemment pourrait être raffinée ici (cf. fig.2), mais elle reste valable, l'observation ne donnant pas d'éléments qui soient inexplicables. Les observations présentées en figure 2 permettent cependant d'agrémenter l'explication en donnant la relation statistique qui lie linéairement le nombre de cluster au rapport $w/2u$, la largeur de la distribution divisée par 2 fois l'incertitude. Ce rapport conduit, lui, à affiner l'explication en précisant que les clusters mis en place le sont dans des bandes régulières de largeurs $w/2u$. Par interactions entre les individus un voisinage local d'interactions exclusives se crée progressivement au cours de la simulation. Si l'on reprend la figure 1, on observe du reste ces deux bandes se former assez rapidement au dessus et en-dessous de 0, les individus se positionnant alors exclusivement sur l'une ou l'autre de ces bandes.

3. Ensuite vient ... un autre modèle !

A partir du modèle précédent, nous avons formulé de nouvelles hypothèses que nous souhaitons explorer. Il est à noter ici, que contrairement au cas où si le modèle avait été conçu pour rendre compte précisément d'une réalité particulière et d'un phénomène défini sur lequel nous aurions pu mener des mesures, un raffinement du modèle aurait pût être justifié par un manque explicatif du phénomène observé, une construction du sens trop appauvrie en somme, ici, les hypothèses ajoutées le sont essentiellement

dans un but spéculatif pour tester *in silico* des hypothèses concernant la structuration d'une population.

Pour étendre notre modèle, nous introduisons donc des extrémistes dans notre population en supposant que ces derniers, situés aux extrémités de la distribution d'opinions, sont plus sûrs d'eux (ont une incertitude plus faible). Nous notons p_e la proportion d'extrémistes dans la population, u_e l'incertitude des extrémistes et U l'incertitude des modérés, on suppose donc que $u_e < U$.

Avec ces nouvelles hypothèses et de la même manière que précédemment, nous pouvons produire un comportement attendu des simulations. Les règles d'interaction favorisant les individus plus certains, les extrémistes vont donc se comporter comme des attracteurs aux extrémités de la distribution d'opinions. Un premier cas attendu, en fonction des paramètres choisis évidemment, correspond donc au cas où chaque extrême va attirer à lui une partie de la population. Un second cas attendu, si l'on considère une incertitude de la population centrale plus faible va correspondre au fait que la population va se regrouper en une bande centrale et ne pas être influencée par les extrémistes.

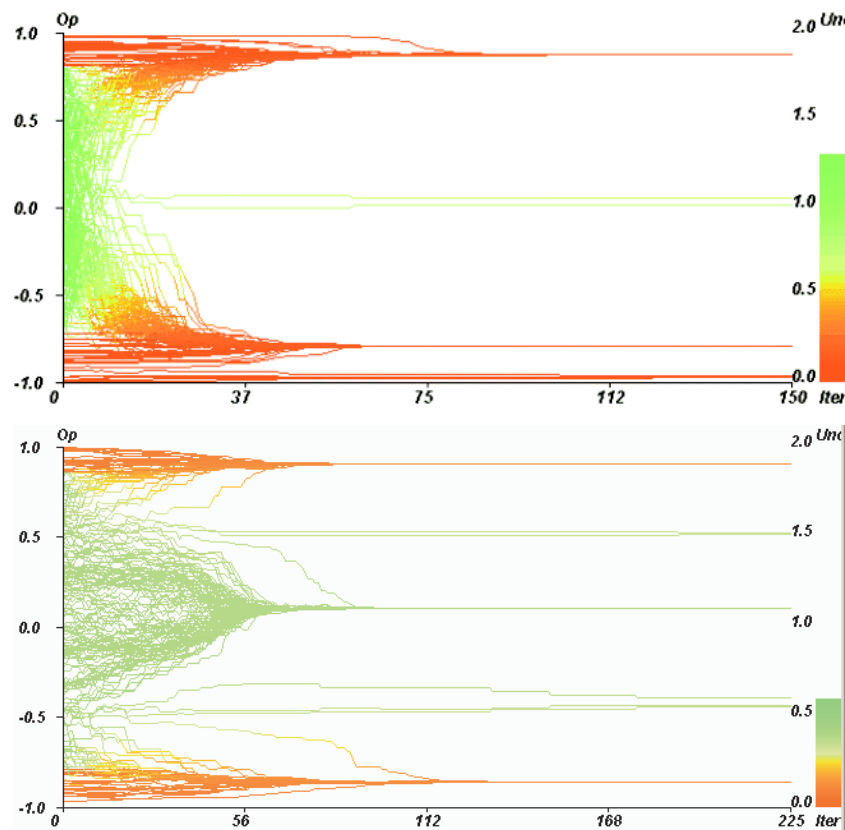


Figure 3 : Deux cas typiques attendus et observés, la convergence double extrême et la convergence centrale. Les figures représentent les trajectoires d'opinions dans le temps des individus de la population, les couleurs représentant l'incertitude des individus de rouge (incertitude proche de 0) à vert (incertitude proche de l'incertitude initiale des modérés).

Cependant, au cours de cette exploration, certains cas typiques et inattendus a priori ont été observés (cf. Fig.4), il s’agit de la convergence massive de l’ensemble de la population vers un seul des extrêmes. La question qui se pose dès lors face à cette observation est : est-ce notre explication est capable de rendre compte de ce phénomène ou bien est-ce que nous devons en réaliser une explication particulière, pour ensuite tenter de la généraliser ?

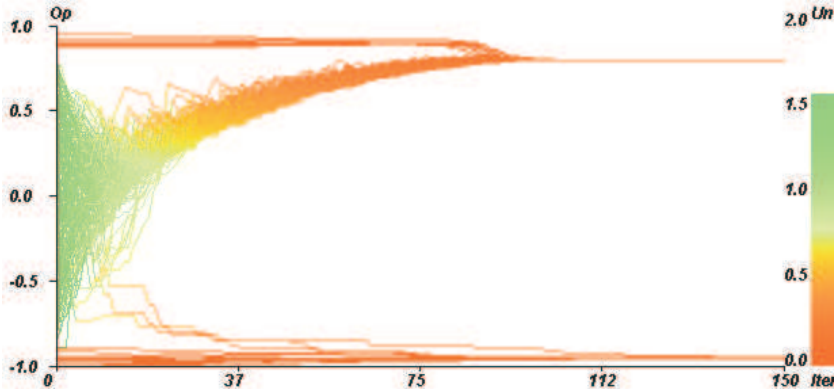


Figure 4 : Le cas inattendu de la convergence vers un seul extrême.

Pour produire une explication, nous devons ici changer notre point de vue sur le système. Il est à noter que recréer la démarche de compréhension du comportement du modèle comme nous sommes en train de le faire est particulièrement périlleux puisqu’il nous faut tenter d’examiner ces résultats avec le regard que nous avons alors, de manière non déformée par l’explication que nous possédons au jour d’aujourd’hui, mais par celle bien imparfaite, que nous avons à l’époque. Pour obtenir de nouvelles informations en jetant un regard nouveau sur ce cas particulier, nous avons donc mis au point un indicateur (cf. eq. 1) qui représente l’influence moyenne à laquelle chaque individu est soumis ou plus exactement la proportion de l’influence à laquelle est soumis l’individu qui provient des extrémistes.

$$E_j = \frac{\sum_{i=extremist} \delta x_{ij}}{\sum_i \delta x_{ij}} \tag{Eq. 1}$$

où δx_{ij} est la modification virtuelle de l’opinion de l’individu A_j sous l’influence de l’individu A_i . En suivant un individu particulier de la population dans ce processus de convergence vers un seul extrême (cf. fig. 5) nous obtenons ainsi de nouveaux éléments qui nous permettent de construire notre explication. Il est à noter que dans ce cadre, le fait de changer de point de vue, de niveau d’organisation et de passer à l’observation d’un comportement individuel particulier participe grandement à la construction réussie d’une explication dans ce cas.

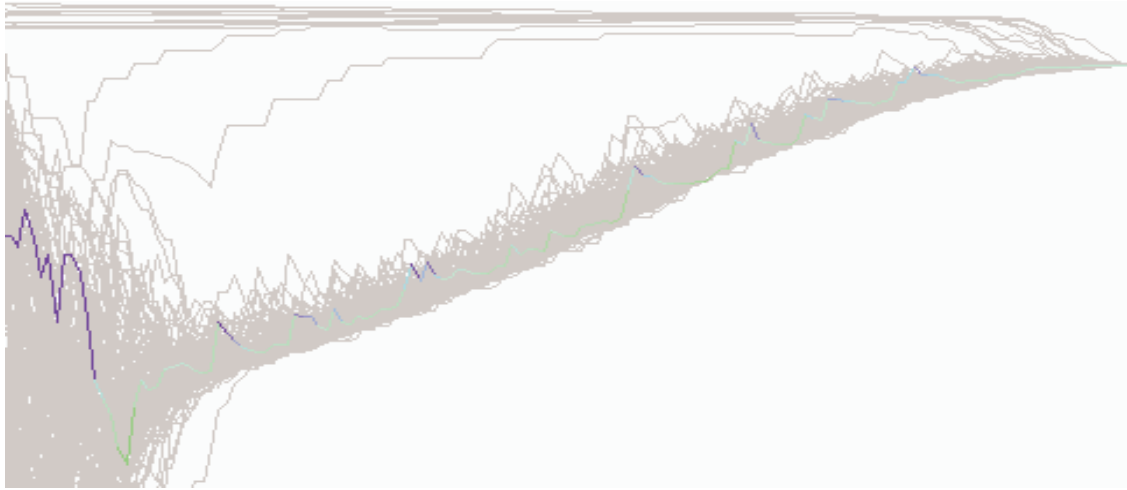


Figure 5 : Zoom sur une trajectoire individuelle correspondant au cas de la convergence vers un seul extrême. Toutes les trajectoires sont en gris à l'exception d'une seule pour laquelle la couleur indique le poids des extrémistes dans l'influence moyenne.

Ainsi, nous observons tout d'abord un processus de regroupement central des modérés. Un mouvement dû à une pression majoritaire des modérés qui pèsent par leur nombre et leur dispersion initiale, dans le processus de regroupement. Une fois ce regroupement central opéré, l'influence des individus du groupe est relativement faible du fait de leur proximité en opinion, ils agiront davantage comme une inertie au mouvement. Après le processus de regroupement central les individus fluctuent au cours de la dérive vers l'extrême. Au cours du temps ils peuvent être attirés par un extrémiste et s'en rapprocher en opinion, puis sous l'influence de la majorité se rapprocher à nouveau du groupe tout en influençant certains éléments de ce dernier. Ce qui conduit à un changement progressif de la majorité vers une opinion extrême. Il y a ainsi oscillation entre influence forte de l'extrême et co-influence entre l'individu et la majorité qui ramène l'individu vers le groupe mais qui tire également progressivement le groupe vers l'extrême.

Au regard de cette explication, des détails de la figure 4 prennent sens, ainsi la diminution progressive de l'incertitude des modérés au cours de la phase de regroupement central, qui est due à quelques interactions des individus avec les extrémistes, est sans doute à l'origine du mouvement vers un seul extrême. Ainsi au cours de ce regroupement central, la diminution de l'incertitude coupe le groupe central d'un des deux extrêmes, ce qui conduit le groupe à être massivement attiré par l'autre. Une hypothèse qui est confirmée par la figure suivante (cf. Fig. 6a) obtenue pour des valeurs de paramètres identiques à ceux de la figure 4 et qui trouverait son explication dans le fait que lors du regroupement central, l'incertitude du groupe diminuant, les individus se coupent des deux extrêmes à la fois. Ainsi que par la figure 6b sur laquelle la représentation

conjointe de l'opinion et du segment d'incertitude (le seuil au delà duquel les individus ne peuvent plus s'influencer) montre effectivement une rupture dans l'influence du regroupement central par un des deux extrêmes.

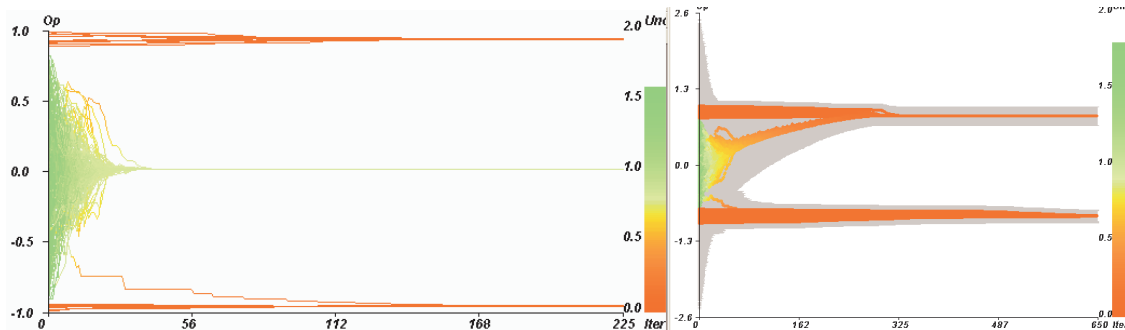


Figure 6 : a) à gauche : une convergence centrale particulière observée pour des valeurs de paramètres identiques à celles de la convergence vers un seul extrême b) à droite, une représentation de la convergence vers un seul extrême visualisant les incertitudes associées aux opinions des individus.

L'observation de ces différents cas et les explications fournies conduisent à se poser une nouvelle question : dans quels cas est-ce que l'on observe quel type de convergence ? Cette question permettrait à nouveau de jeter un regard neuf sur le modèle et de proposer de nouveaux éléments permettant de comprendre plus finement pourquoi l'on observe ces formes.

Pour observer dans quels cas apparaissent chacun des types de convergence, nous devons donc conduire une exploration du comportement du modèle, comportement synthétisé comme étant le type de convergence moyen résultant, dans l'espace des paramètres. Ce problème pose donc celui de la création d'un indicateur permettant de synthétiser le type de convergence observé.

4.Construction de l'indicateur

Pour exprimer les résultats de l'exploration plus aisément, nous utilisons un indicateur de type de convergence, noté y : $y = p'_+{}^2 + p'^-{}^2$ où p'_+ et p'^- sont les proportions d'agents initialement modérés qui deviennent extrémistes (respectivement positifs et négatifs). Ainsi :

- Si aucun des agents modérés ne devient extrémiste (cas de convergence centrale), alors p'_+ et p'^- sont nuls et $y = 0$.
- Si les agents modérés convergent équitablement vers les deux extrêmes, nous avons alors $p'_+ = 0.5$ et $p'^- = 0.5$, d'où $y = 0.5$.
- Si les agents modérés convergent vers un seul extrême, par exemple le positif, nous avons $p'_+ = 1$ et $p'^- = 0$, d'où $y = 1$.

Les valeurs intermédiaires de y correspondant à des situations intermédiaires, la valeur de cet indicateur nous permet aisément de déterminer le type de convergence.

5.Exploration de l'espace des paramètres

La mise au point de cet indicateur nous a permis de catégoriser automatiquement les simulations en différents cas de convergence et qui plus est de manière continue, l'indicateur prenant ses valeurs entre 0 et 1. Pour l'exploration de l'espace des paramètres qui a été ainsi réalisée (cf. Fig.7) nous avons choisi deux dimensions qui nous ont semblé être, suite à l'expérience acquise au cours de l'exploration un peu artisanale qui a précédée, celles qui influaient le plus sur le comportement observé du modèle. Une analyse de sensibilité conduite auparavant nous aurait peut-être permis de choisir ces dimensions de manière plus judicieuse. Toujours est-il que cette exploration conduite sur les dimensions U (incertitude des modérés) et pe (proportion d'extrémistes dans la population) nous a permis de comprendre un certain nombre de points. En particulier le fait que pour des incertitudes faibles des modérés, l'occurrence de convergences centrales à l'exclusion de tout autre, confortait notre explication selon laquelle pour des incertitudes faibles des modérés, ceux-ci sont moins sensibles aux influences des extrémistes.

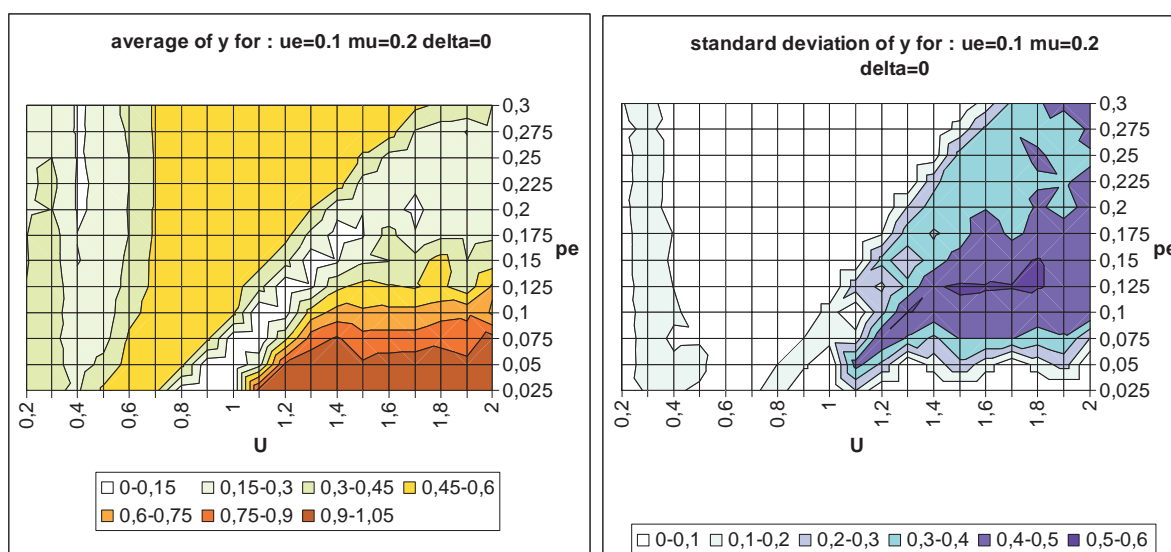


Figure 7 : Formes typiques de la moyenne et de l'écart type de l'indicateur y (50 réplifications en chaque point du graphe) en fonction de l'incertitude des individus modérés (U) et de la proportion globale d'extrémistes dans la population (p_e). Les autres paramètres sont fixés : incertitudes des extrémistes u_e = 0.1, intensité des interactions μ = 0.2. Sur le graphe des moyennes de y, les zones jaunes et blanches à gauche correspondent à la convergence centrale, les zones oranges, typiquement au milieu à la convergence vers les deux extrêmes, et la zone marron en bas à droite correspond au simple extrême.

Une explication plus détaillée des comportements observés est disponible dans (Deffuant et al. 2002). Le point important ici étant que l'on utilise un nouvel outil, un nouveau point de vue, pour comprendre le comportement du modèle et ainsi se donner des outils pour construire du sens et plus exactement pour construire du sens à propos de l'articulation, dans le cadre de ce modèle particulier, entre comportements individuels et comportements collectifs.

Ce processus de construction du sens semble ici sans fin tant nous pourrions solliciter de nouveaux outils qui nous permettraient de discuter nos hypothèses. Il s'agit bien alors de sélectionner pertinemment les outils susceptibles de mettre cette explication en danger, d'utiliser les points de vue susceptibles d'illuminer d'un jour nouveau le comportement du modèle et ainsi d'amener des connaissances qui sont aussi différentes que possibles des observations précédemment réalisées. Ainsi en tendant toujours vers l'asymptote que serait la preuve formelle des propriétés qui sont jusque là uniquement intuitées sur ces modèles, le modélisateur peut solliciter indéfiniment de nouveaux outils, de nouveaux regards sur son modèle, construisant au fur et à mesure son exploration. Comme l'écrit Antonio Machado :

*« Voyageur, le chemin
C'est les traces de tes pas
C'est tout; voyageur,
il n'y a pas de chemin,
Le chemin se fait en marchant »*

Remerciements

Derrière le nous utilisé dans le texte se cachent souvent d'autres personnes qui ont participé au travail plus large qui est présenté : Thierry Faure, Jean-Pierre Nadal, ou qui l'ont dirigé : Guillaume Deffuant et Gérard Weisbuch. Ils n'auront cependant certainement pas vécu le processus de construction du sens de la même manière tant cette construction est subjective et individualisée.

Références

- Bergson, H., *La pensée et le mouvant*, PUF, 2003.
Morin, E., *La méthode tome 1 : La Nature de la Nature*, Seuil, 1977.
Popper, K., *La connaissance Objective*, Flammarion, 1999.

- Deffuant, G., Amblard, F., Duboz, R. et Ramat, E., “Une démarche expérimentale pour la simulation individus-centrée”, *Actes des 10èmes Journées de Rochebrune 2003, Rencontres interdisciplinaires sur les systèmes complexes naturels et artificiels, épistémologie de la simulation*, janvier 2003, Rochebrune, France, pp.45-64.
- Deffuant, G., Amblard, F., Weisbuch, G. et Faure, T., “How can extremism prevail? A study based on the relative agreement model”, *Journal of Artificial Societies and Social Simulation*, vol.5, n°4, 2002.

Enigmes et traces

Le rôle de la dé-contextualisation et de la re-contextualisation dans le processus d'écriture lecture

Marc Nanard, Jocelyne Nanard, Jean-Yves Delort

LIRMM

161 rue Ada, 34392 Montpellier cedex 5

{nanard, delort}@lirmm.fr

04 67 41 85 16 et 04 67 41 85 17

Résumé

Un document ne véhicule pas de sens par lui-même. Lire revient toujours à résoudre l'énigme de la construction d'un contexte interprétatif plausible dans lequel le document prend sens pour un lecteur. L'objectif ici est de présenter quelques mécanismes mis en jeu dans l'élaboration, la préservation, et l'exploration de tels contextes dans le cas de documents électroniques. La forme du document, celle qui est perçue par l'humain avant toute interprétation, constitue une trace qui pré-orienté et peut faciliter ou perturber la lecture. La lecture savante vise une re-construction rationnelle d'un contexte interprétatif, sans toutefois pouvoir prouver l'unicité et la validité de tels contextes. Alors, à défaut de pouvoir reconstruire le contexte interprétatif de l'auteur, peut-on au moins en conserver des traces ? Le dépôt légal, et plus spécialement l'archivage des flux audiovisuels par l'INA est pris comme exemple pour illustrer quelques problèmes généraux de conservation de contextes, avant de regarder les techniques informatiques associées et leur emploi par exemple pour la mémoire d'entreprise. Mais un contexte interprétatif, même celui de l'auteur, n'est pas stable dans le temps. Nous abordons rapidement les techniques informatiques utilisées pour explorer leur évolution et nous montrons sur un exemple de génétique textuelle comment observer la genèse d'une œuvre, et par suite l'évolution de son contexte originel. Enfin la résolution de l'énigme qu'est une lecture ne doit pas conduire le lecteur à réifier ce qu'il a interprété.

Mots clés : lecture, écriture, contexte, dé-contextualisation, re-contextualisation, Mémoire d'entreprise, temps dans les documents.

Dé-contextualisation et re-contextualisation

Dans un conte breton, Aéla, la princesse du fond des mers, retrouve dans une épave un flacon de « vif argent » qui était adressé à « Mr Torricelli ». Les poissons experts de son palais étudient l'objet et concluent que cet élixir

retrouvé au sein d'une « sépulture sous marine d'humains » ne peut qu'être destiné à un de leurs rites funéraires : il doit permettre aux défunts de reprendre vie sur la terre s'ils en boivent. La princesse émerveillée par les perles métalliques souples, brillantes et changeantes de ce merveilleux vif argent, rêve de devenir femme, boit tout le flacon et en meurt... On se trouve ici devant la caricature grossière d'un phénomène de de-contextualisation puis de re-contextualisation. Le fret perdu dans ce naufrage est détaché de son contexte, il n'est plus l'un des objets transportés, destiné à une personne précise. Enigme pour Aéla et ses poissons savants, il est re-contextualisé dans un autre référentiel, celui des experts en archéologie, il y devient un objet merveilleux, réponse erronée à une énigme dont la mauvaise interprétation entraîne la mort d'Aéla. Certains y voient la cause des taux élevés de mercure que l'on trouve dans les poissons carnassiers qui auraient dévoré son cadavre. D'autres craignent qu'elle ne se soit effectivement réincarnée en femme, comme elle en rêvait. Mais ces dernières assertions ne sont sans doute, elles-mêmes, que des re-contextualisations personnalisées du conte.

Lecture et contextualisation

L'objectif de cette première section est de rappeler, en ayant choisi une extension très large au mot « document », quels mécanismes de re-contextualisation sont mis en jeu pour l'élaboration du sens dans les différentes formes de lecture, artistique, technique, ou savante. Nous examinons comment le contexte de lecture est construit, en particulier dans le cas de la lecture savante, et nous montrons quelques phénomènes qui contribuent à biaiser ce contexte.

Dé-contextualisation et re-contextualisation dans le processus d'écriture-lecture

Le mécanisme de dé-contextualisation et de re-contextualisation est un des fondements des mécanismes d'écriture et de lecture [Valenti 2000]. Lire un document revient toujours à résoudre l'énigme de la construction d'un contexte plausible dans lequel l'information prend sens. L'auteur, en produisant une œuvre, dé-contextualise de fait son contenu pour l'offrir à d'autres réinterprétations par les lecteurs, qui ne peuvent résoudre cette énigme que dans leur contexte personnel. Ce mécanisme est poussé à son paroxysme dans le cas de la lecture artistique, par exemple en poésie, où la suggestivité vise à faire émerger des contextes interprétatifs latents chez le lecteur. À l'opposé, la documentation technique doit assurer une similitude aussi grande que possible des contextes interprétatifs de l'auteur et du lecteur pour assurer que la reconstruction du sens soit aussi exacte que possible chez le lecteur. La rhétorique de l'écriture technique suggère donc de placer directement au sein de

chaque document des traces qui facilitent l'élaboration du contexte souhaitable pour l'interpréter.

Entre les deux extrêmes, la lecture de poèmes et celles de manuels techniques, beaucoup d'autres formes de lecture existent, mais toutes partagent la même caractéristique : elles conduisent le lecteur à reconstruire un sens, en interprétant le document dans un contexte qui ne peut être propre qu'à ce lecteur. Faire que plusieurs individus distincts soient capables d'élaborer des contextes interprétatifs assez voisins afin d'échanger des idées au moyen de documents n'est pas une propriété intrinsèque des documents mais seulement un fait sociologique lié au vécu commun des individus concernés [Richaudeau 1987].

Documents

Un document est fondamentalement un support d'information qui doit d'abord être interprété par un humain. C'est en cela qu'un document est par nature une énigme. L'élaboration d'un document peut être réalisée indifféremment par un humain, par un ordinateur, ou même, pourquoi pas, par la nature¹. Le document ne véhicule pas de sens par lui-même [Bachimont 1999]. Il n'est qu'un signifiant qui ne laissera transparaître un signifié que comme résultat de l'activité de lecture, lorsque le lecteur aura résolu l'énigme, en projetant sur elle une part de sa connaissance personnelle, par exemple en annotant le document (« lire, c'est écrire » [Stiegler 1996]). L'interprétation ne peut se faire que dans le référentiel du lecteur.

Document et donnée

La notion de données et celle de documents sont souvent confondues à tort, car toutes deux visent à communiquer des informations. Une donnée est échangée entre partenaires matériels ou humains qui partagent préalablement une même convention pour interpréter celle-ci². Une donnée n'est donc pas énigmatique, elle résulte d'un codage conventionnel et partagé de l'information. La notion document au contraire suppose que l'interprétation par le lecteur puisse se faire sans connaissance de la convention qu'a choisi l'auteur.

¹ La succession des orientations du champ magnétique dans les roches éruptives des rifts mid-océaniques constitue un bien un document relatif à l'histoire de la terre, dans la mesure où l'homme n'a pas, a priori, la clef pour interpréter ce qu'il observe. Enigme, qui, lorsqu'elle est éclairée par d'autres faits connus un contexte de lecture scientifique, prend sens pour nous et nous raconte l'histoire du déplacement des pôles.

² La donnée échangée entre deux machines, ou même entre des partenaires humains, peut être interprétée des deux côtés de façon identique car son protocole d'interprétation a pu être rendu disponible. Le constructeur des machines peut « par construction » faire qu'elles partagent la même façon d'interpréter la donnée. L'auteur d'un protocole de communication en a défini les règles et les a rendues connues des partenaires, la donnée n'est pas une énigme. Pareillement, une fois qu'un physicien a fourni une explication plausible pour le mécanisme de succession des orientations du champ magnétique, une clef qui en permet une interprétation est devenue publique. De ce fait, une séquence particulière de successions d'orientations n'est plus qu'une simple donnée scientifique à interpréter avec cette convention.

Le rôle de la forme comme trace initiale

Négentropie locale nécessaire pour être distinguable dans l'entropie environnante, un document est d'abord une « forme distinguable » qui est perçue par le lecteur comme un écart par rapport à l'état usuel du monde environnant. Cercle tracé sur le sable lisse d'une plage, tache de craie sur un tableau sombre, encre sur un papier blanc, variation temporelle du spectre acoustique, différence de luminosité entre les points d'un écran, le document est toujours matérialisé par une exception dans l'environnement : sa *forme*, au travers de laquelle il est perçu en premier par l'humain. Cette forme a des propriétés qui influent sur le mécanisme de lecture (Voir la théorie du support dans [Bachimont 1999]).

Avant même que le lecteur ne puisse s'appuyer sur son acquis -par exemple la reconnaissance des caractères [Ferrand 2001] ou la connaissance d'une langue- pour analyser le signifiant qu'il perçoit, la forme du document lui a suggéré -de façon étymologiquement « immédiate »- un contexte de lecture et a mis en place, par rappel suggestif, un ensemble de traces qui contribueront à l'orienter dans la démarche d'interprétation de l'énigme. La forme joue un rôle de rappel de connaissances implicites collectives ou individuelles. L'aspect des caractères utilisés, leur placement relatif, les relations géométriques, les couleurs, etc. suggèrent une ambiance propice ou non à l'interprétation correcte du contenu. La forme n'a pas plus ou moins d'importance que le fond, elle en est indivisible. Elle en est simplement le soubassement. Le lecteur la perçoit intuitivement comme une trace de l'activité de l'auteur. Ainsi, une copie d'examen sale, brouillon, raturée,... sera perçue comme telle de prime abord avant même que l'examineur n'ait pu la lire pour juger la qualité du contenu. De même, un graphiste saura, par la composition du document, suggérer de façon immédiate l'ambiance, le sérieux, le dynamisme, la rigueur, la violence, le classicisme ou la modernité d'un site, indépendamment de tout contenu. Un document porte toujours en lui-même, par sa forme, une partie des traces qui permettent de l'interpréter. Un texte de Julius Caesar, finement gravé sur du marbre, porte par sa forme la solennité que l'auteur a voulu donner à son contenu, quel que soit celui-ci. Griffonné sur un papier, il perdrait de sa noblesse initiale, avant peut être de la retrouver lorsqu'un sens lui est affecté.

La forme n'est pas seulement géométrique ou visuelle, même si cet aspect est le plus facile à évoquer. Les premières notes d'un morceau de musique suffisent à mettre en place l'ambiance générale (ex. Beethoven 5ème symphonie). A un niveau plus abstrait, la forme d'une équation suggère une stratégie de résolution avant tout raisonnement. Un champion d'échecs *voit* un coup avant d'avoir analysé le détail de la position des pièces. Les traces activées par la perception de la forme sont toujours subjectives. Elles appartiennent au lecteur, et s'appuient souvent autant sur son sub-conscient et son émotivité que sur son acquis et ses connaissances rationnelles. Elles participent à l'élaboration d'un contexte interprétatif primitif. L'interprétation géniale d'un phénomène

physique encore inexpliqué sera souvent déclenchée chez le chercheur par une suggestivité personnelle, souvent très complexe, issue d’analogies entre des propriétés parfois très abstraites avec celles d’autres faits, traces « vues » et raccordées spontanément par le chercheur, qui facilitent l’élaboration d’un contexte interprétatif pour lire le fait étudié.

La forme, par la force des relations qu’elle suggère implicitement, crée un contexte interprétatif primordial, indépendant de la sémantique réelle du document. Ceci peut même avoir des conséquences dramatiques sur l’interprétation du contenu, y compris dans les domaines où la forme fut longtemps considérée comme secondaire. Dans la décennie 70, un accident ferroviaire fit une soixantaine de morts. Au procès qui s’en suivit, il fut expérimentalement prouvé, en réalisant l’expérience en public sur l’auditoire, que la forme visuelle du document décrivant les actions qu’aurait dû faire le conducteur induisait systématiquement la même erreur d’interprétation sur tous ses lecteurs. Les phrases étaient correctes, un analyseur sémantique en aurait reconstruit parfaitement le sens [Luc 2001]. Mais pour l’humain, leur placement géométrique maladroit (voir figures 1a et 1b) induisait involontairement des rapprochements, des éloignements, traces infimes que le lecteur percevait en premier et qui le conduisaient à une interprétation erronée. Le coupable était le document, ou plutôt sa lecture !

```

Si tel événement se produit
      {précisément celui concerné par l'accident}
alors deux cas sont à envisager
- premier cas si xxxx xxxx xx {condition présente ce jour là}
      alors xxx xxxxx xxxxxxxx x x xxxx
      xxx xxxx xxxx x
      xx xxx xx xxxx xx
- deuxième cas {condition inverse, non présente ce jour là}
      yyy yyy yyyyyy yyy yyy yyy yyy
      yyyyyy yy yyyy yyyyyy yyyyyyyy
      Dans les deux cas, il est impératif de.....
      zzz zz zzzzz zzzzz zzzz zz z zzzz zzzzz
      zzzzz zzz zz zzz zzzzz zzzz zzz z zzzz zzzzz

Si tel autre événement se produit
alors

Si tel événement se produit
      {précisément celui concerné par l'accident}
alors deux cas sont à envisager
- premier cas si xxxx xxxx xx {condition présente ce jour là}
      alors xxx xxxxx xxxxxxxx x x xxxx
      xxx xxxx xxxx x
      xx xxx xx xxxx xx
- deuxième cas {condition inverse, non présente ce jour là}
      yyy yyy yyyyyy yyy yyy yyy yyy
      yyyyyy yy yyyy yyyyyy yyyyyyyy
      Dans les deux cas, il est impératif de.....
      zzz zz zzzzz zzzzz zzzz zz z zzzz zzzzz
      zzzzz zzz zz zzz zzzzz zzzz zzz z zzzz zzzzz

Si tel autre événement se produit
alors
  
```

Figure 1. Influence de la forme sur l’interprétation rapide d’un document : les textes sont identiques. A gauche (figure 1a) : présentation induisant une interprétation erronée. A droite (figure 1b) : présentation induisant une interprétation correcte.

La lecture savante

Sous le terme de lecture savante, on regroupe un ensemble d'activités dans lesquelles le lecteur fait effort pour reconstruire le contexte qui lui permet d'interpréter un document donné au moyen d'un processus récursif reposant sur la lecture d'autres documents [Virbel & Maignien 1996].

Pour le chercheur, l'historien, interpréter un fait consiste à résoudre une énigme. Il se doit d'essayer de reconstruire un sens aussi voisin que possible de celui qu'il suppose que l'information avait dans son contexte initial. Pour cela, il serait indispensable qu'il puisse accéder au contexte de l'information initiale [Maignien 1998]. Hélas, le contexte réel d'une information ancienne est le plus souvent perdu à tout jamais, seuls peuvent exister pour nous des contextes de lecture actuels que leur géniteur réussit à argumenter de façon suffisamment convaincante pour leur donner force de croyance. Ces contextes artificiels résultent de la prise en compte d'autres traces, qui constituent chacune des énigmes, dont l'interprétation repose récursivement sur celle d'autres traces. Construire un contexte revient à assembler un puzzle. S'il paraît achevé, est-on sûr pour autant de son unicité ? Y a-t-il plusieurs solutions ?

Est-ce parce qu'un humain du XXI^e siècle arrive à tailler un silex qui ressemble à celui d'un néandertalien que l'on peut être sûr que ce dernier opérait de cette façon ? La relation à la mort du romain du I^{er} siècle est-elle la même que celle du japonais des années 40 et de l'occidental du 21^{ème} siècle ? Peut-on même espérer en connaître autre chose que notre émotivité personnelle en interprétant d'autres documents, et ce, récursivement. Peut-on juger de la pertinence d'un discours sur un tel thème ?

Ce type de démarche est le propre, par exemple, de l'archéologie, qui, faute de traces réelles suffisantes, ne peut qu'opérer par projection de contextes hypothétiques, puis par vérification de la seule plausibilité de ces contextes. Nul n'a participé aux rites des hommes préhistoriques ni n'a partagé leurs croyances. L'interprétation des œuvres picturales rupestres, au-delà de la simple émotion esthétique qu'elles peuvent procurer, ne peut que les replacer dans des contextes artificiels dont l'exactitude ne peut être prouvée.

La lecture savante ne se limite pas à ces extrêmes. Elle est la pratique courante de la communauté scientifique dans son ensemble, toutes disciplines confondues. Sa démarche repose sur l'utilisation de traces pour élaborer progressivement et récursivement un contexte interprétatif complexe, qui donne une clef possible pour l'interprétation collective des documents ou des indices qui ont servi à la bâtir. Ceci est le propre des démarches explicatives scientifiques : on rassemble un certain nombre de faits dont on montre l'interdépendance et dont l'interprétation donnée à chacun induit l'interprétation souhaitée pour les autres. « La planète se réchauffe ! ». Ceci est un fait mesurable, mais en tant qu'énigme, on lui trouve des dizaines d'explications distinctes, toutes potentiellement cohérentes.

Préservation des contextes

L'objectif de cette seconde section est d'illustrer les difficultés inhérentes à la préservation d'un contexte. A défaut de pouvoir reconstruire le contexte interprétatif de l'auteur, serait-il au moins possible d'en conserver des traces ? A partir de divers exemples, nous en montrons à la fois l'intérêt et les difficultés, avant de regarder les techniques informatiques utiles et leur utilisation pour la mémoire d'entreprise.

Le dépôt légal préserve-t-il un contexte ?

Très tôt, l'Homme a cherché et continue à pérenniser son savoir en le conservant dans des ouvrages et ceux-ci dans des bibliothèques [Merlot 1996]. Le dépôt légal a institutionnalisé et systématisé cette pratique (voir [BnF]). Mais, ce dépôt constitue-t-il la préservation du contexte originel du document ? Quels problèmes se posent ? Quels sont les besoins réels de l'industrie et quels sont les moyens informatiques actuels de préservation des contextes [Thibodeau 2001] ?

Le législateur a assuré la préservation de notre patrimoine intellectuel par la loi sur le dépôt légal qui fut instauré par François Premier. Moyen efficace pour rassembler la connaissance, le législateur l'a étendu aux documents audiovisuels avec la création de l'INA et de la Cinémathèque. Il a toutefois oublié un point capital. Les traces de l'élaboration d'un document par son auteur sont parfois aussi importantes voire plus que le document, car celui-ci, une fois détaché de son contexte perd une partie de son sens. La forme modernisée de cette loi, -« tout ce qui est diffusé doit être archivé »- engendre quotidiennement beaucoup plus d'énigmes pour les générations futures qu'elle n'en résout. Par exemple, chaque jour, chacune des chaînes de télévision adresse à l'INA les cassettes vidéo, copies désormais *mortes* de ce qui a été diffusé dans la journée. Le mot « mortes » est volontairement employé pour qualifier ces cassettes, car elles ont perdu toute relation avec le reste du monde, avec leur contexte. Il ne reste plus qu'un contenu et une côte d'archivage. Ces cassettes ont été arrachées à leur contexte. Quelques mois plus tard, elles sont visualisées pour élaborer les fiches descriptives jetables, qui serviront ensuite à leur indexation normalisée et rationalisée, elle-même réalisée par d'autres spécialistes, des documentalistes, dont la tâche est d'interpréter les fiches et non les cassettes ! Et pourtant, la production de ces émissions avait été faite sous forme numérique ; chaque séquence était directement reliée aux documents qui avaient servi à sa préparation. Ainsi, sous cette forme, on disposait par exemple de documents relatant l'identité de toutes les personnes sur le plateau, de la façon dont le metteur en scène avait préparé son travail, etc. Et malheureusement, tout ce contexte riche, disponible électroniquement, a disparu car lui, ne fait pas partie du dépôt légal. Seul ce qui est parti sur les ondes a été archivé, mais il devient un contenu mort. C'est une des raisons qui entraînent maintenant l'INA dans de

nouveaux programmes de recherche sur la valorisation des contenus permettant l'exploration de contextes associés aux enregistrements. Le slogan de l'INA est d'ailleurs : « nous construisons l'avenir de votre mémoire »

Les traces de l'élaboration d'un objet sont souvent aussi utiles que l'objet lui-même. Elles permettent à l'observateur de se replacer dans un contexte plus proche de celui du créateur. Ainsi, le coffrage d'un pilier en béton s'étant déformé, le résultat final fut une forme des plus inattendues. Le Corbusier voyant cela vérifia que la solidité de la structure n'était pas compromise et demanda que l'on garde, en tant qu'œuvre d'art, cette forme qu'il n'avait pas prévue. Si le célèbre architecte n'avait pas rapporté sur le contexte originel de cet événement, les interprétations de cette particularité pourraient continuer sans fin. Citons un autre exemple. Dans le domaine textuel, la génétique textuelle permet de reconstituer le processus de construction des œuvres à partir des manuscrits [Lebrave 1998].

Comment préserver un contexte ?

Est-ce nécessaire ?

Les volumes considérables de documents accumulés tant par les institutions en charge de l'archivage (BnF, INA, Cinémathèque...) que de façon plus informelle sur le Web, posent et vont poser de plus en plus le problème de leur re-contextualisation. La valorisation du patrimoine audiovisuel, principalement dans le cas de la réutilisation, conduit à extraire de brèves séquences vidéos de leur contexte initial et à les replacer volontairement dans un autre contexte par le mécanisme du montage (voir par exemple Vox Populi [Boconi 2005]). Quel chercheur ne s'est trouvé dans la situation de constater que ses propos ont été « déformés » par un journaliste, alors qu'il ne s'agit que d'assemblage d'extraits authentiques mais pertinemment choisis en vue d'une re-contextualisation différente ? Ce phénomène est inévitable et fait partie intégrante de la communication et d'aspects de son art tels que la rhétorique.

Un jeu classique soumis aux étudiants en multimédia consiste à leur faire réaliser, par exemple à partir d'extraits de films de discours de Jean Jaurès, un montage dans lequel ils lui font prononcer un faux vrai discours conservateur, asocial et xénophobe. De façon similaire, un montage célèbre, fait de pièces authentiques, relatif aux différents ministres des finances ayant exercé depuis le début du siècle dernier les montre tous prononçant la même phrase, avec le même geste emphatique : « je m'engage solennellement à baisser les impôts... » disent-ils véhémentement, en levant élégamment très haut la main droite dans le sens du renforcement de l'idée, à moins que ce ne soit celui des taxes. Ceci n'est qu'une interprétation...

Conserver un environnement

Dans le but de faciliter la lecture savante des archives audiovisuelles, un effort doit être fait pour conserver les informations en relation directe avec leur contexte d'origine. Pour cela, l'INA conserve non seulement les créations

audiovisuelles de toutes natures, en les indexant en tant que pièces isolées, mais aussi des « échantillons complets de journées » qui, par leur enregistrement continu et complet, publicités comprises, constituent une image audiovisuelle du contexte dans lequel a été diffusée telle émission. Grâce à de telles traces, les chercheurs en sociologie peuvent mieux resituer l'ambiance générale du paysage audiovisuel d'une époque donnée et par suite élaborer un cadre plus rigoureux pour l'analyse des propos tenus lors d'une interview, ou des sous entendus d'une publicité. Malgré cet effort, dans quelques centaines d'années, l'interprétation de certaines publicités constituera sans doute encore une énigme. La remise en contexte par des traces du paysage audiovisuel quotidien permettra-t-elle, à elle seule, de la résoudre ?

Conserver les contextes successifs de lecture

Au-delà de telles traces visuelles directes, l'INA pose le problème beaucoup plus complexe de la conservation des archives de faits historiques dont il est dépositaire. Pour la plupart des humains, experts peut-être exceptés, la ré-interprétation d'un fait ne peut se faire ailleurs que dans le contexte courant du lecteur. Ainsi une scène de la guerre de 40 soulève une émotion encore vive par sa relecture au travers d'une mise en contexte actuel. Pourquoi les faits de la guerre de cent ans n'entraînent pas une émotion semblable ? Une réponse est peut-être que nous sommes incapables de résoudre l'énigme qu'est l'interprétation de ces faits dans leur contexte d'origine ! Nous sommes dans l'incapacité de reconstruire correctement le contexte initial. Relus dans notre contexte, ces faits ne sont plus pertinents pour nous, ils ne nous touchent plus autant. Pour éviter ce genre de divergence, un effort très important est fait, par l'indexation des documents et la mise en place de leurs interdépendances, pour conserver et relier les pièces maîtresses du contexte historique des événements et de leurs interprétations successives.

Une des solutions consiste en la conservation de l'évolution des contextes de lecture successifs d'une même information au cours du temps, en reliant entre eux par un réseau hypermédia les éléments qui y participent. Un même fait historique n'est jamais perçu de la même manière dans le feu de l'action et avec une ou plusieurs générations de recul. Ainsi, nul ne trouverait dans l'indexation officielle des archives de l'INA de documents sur la « guerre d'Algérie ». En effet, lorsque ces documents furent archivés et indexés dans leur contexte historique, il n'y avait officiellement pas de guerre déclarée, mais seulement des événements, des troubles de l'ordre public. En faisant coexister les contextes de lecture successifs des mêmes faits et en établissant des relations hypertextuelles entre eux, on rend possible une lecture savante de faits historiques passant par navigation d'un document relatif à l'interprétation d'un fait dans un contexte, à celui de l'interprétation de ce même fait soit dans son contexte de lecture initial, soit dans un autre.

Ce n'est qu'en conservant sous une forme facilement navigable les contextes de lecture successifs de ces faits, qu'au travers de leur évolution on peut constater facilement l'évolution des interprétations.

Les moyens informatiques pour conserver et exploiter un contexte

Le principe général

Les techniques hypermédias assurent le maintien en relation des documents. Rappelons qu'un lien ne doit pas être vu comme un simple « clic pour aller » mais bien comme l'explicitation d'une relation entre deux informations, de préférence sémantique, dénotée par le type du lien. Ceci est important à rappeler car HTML a développé une acception naïve et très restrictive du concept de lien hypertextuel. Cette dérive porte sur plusieurs points majeurs : la non réversibilité des liens, l'absence de typage sémantique, la perte progressive d'usage de l'ancrage dans le document cible, la propriété exclusive des liens par l'auteur du document.

Heureusement, d'autres notations basées sur XML, par exemple Xlink, rendent possibles des constructions plus riches. En particulier, il est possible de spécifier des liens entre documents sans que le propriétaire du lien ait besoin d'être le propriétaire d'un des documents.

Il devient alors possible de faire flotter un réseau indépendant de liens qui s'ancrent sur les documents concernés, sans altérer ni devoir posséder ceux-ci. Cette technique permet en particulier d'attacher une glose moderne à des fac-similes de documents anciens. Le réseau matérialise les diverses relations entre ces documents. En parcourant le réseau, soit explicitement, soit par des outils de raisonnement, il devient possible et même facile d'explorer le contexte, même lointain, d'un document.

Le typage pouvant être quelconque, il n'y a pas de restriction sur la nature des relations ainsi représentées. Dans les cas les plus élaborés, des techniques de représentation de connaissance, basées par exemple sur les graphes conceptuels, rendent possible l'expression de relations sémantiques complexes et le raisonnement logique sur ces descriptions. Un tel procédé a été développé par notre équipe en collaboration avec l'équipe « graphes conceptuels » du LIRMM pour l'indexation vidéo fine de documents de l'INA.

Un réseau hypertextuel constitue un contexte interprétatif en rapprochant pour le lecteur des informations que l'auteur des liens considère comme corrélées. L'exploration de ces relations ou l'accès aux informations concomitantes permet de resituer un travail, une idée, dans son contexte, ou tout au moins dans une partie de celui-ci.

Contexte et mémoire d'entreprise

Le souci de préservation du contexte n'est pas l'apanage des seules institutions en charge des archives nationales. Sur le plan industriel, la mémoire d'entreprise constitue un enjeu économique très important. La re-interprétation de faits industriels, d'ordres ou encore de consignes constitue souvent une

énigme. La conservation des traces des raisons qui ont conduit à une décision, et l'accès à ces traces est fondamental.

Par exemple, Airbus industries a été conduit à déplacer les toilettes de la cabine vers la zone des soutes afin d'augmenter le nombre des sièges pour certaines compagnies. Cette opération implique quelques transformations structurelles mineures de l'appareil, par exemple la pose d'un escalier. Comprendre les raisons de la présence de tel élément structurel, ou de tel détail du protocole d'assemblage conduit à autant d'énigmes qui, certes, peuvent être résolues dans le cadre d'un processus de re-conception, qui situe le problème dans sa globalité dans un contexte réactualisé. La disposition de traces précises sur les motivations et sur le contexte des choix technologiques qui avaient été retenus dans la conception initiale permet de replacer le concepteur de la transformation dans le contexte de travail de son prédécesseur, et facilite ainsi sa tâche.

Interpréter un choix technologique dont on ne dispose plus que du fait en tant que tel, sans le contexte dans lequel il a pris naissance, constitue une énigme dont l'interprétation est incertaine. Son interprétation devient presque évidente, ou tout au moins demande un effort moindre, lorsqu'on peut le replacer dans son contexte original. Ceci implique que les éléments de celui-ci aient été préservés et soient restés attachés aux éléments qu'ils éclairent. La préservation du contexte de travail, son attachement hypermédia à l'objet, au processus dans lesquels il est intervenu, sa mise en mémoire dans l'entreprise est indispensable pour assurer notre compréhension de phénomènes qui deviendront de plus en plus complexes, énigmatiques voire dangereux.

L'utilisation de structures hypermédiées rend possible, pour ne pas dire facile, l'attachement d'informations de toutes natures. En reliant directement les plus futiles traces d'activité aux objets qu'ils concernent, on préserve leur contexte, on permet de l'explorer, et on rend moins énigmatique leur interprétation. Pour l'industrie, ceci est capital.

Cette technique est facilement généralisable à de nombreux domaines. Déjà des appareils de photos équipés de GPS et de boussole électronique sont capables d'indexer les photos en identifiant où elles ont été prises et ce qu'elles sont censées avoir visé. Ce ne sera plus une énigme de savoir où et quand a été prise une photo, mais plutôt de savoir pourquoi l'appareil a indexé cet oiseau volant sur un ciel bleu comme étant le dôme des invalides. Est-ce une bonne chose, nous ne jugerons point.

Ce qui est vrai pour les documents, l'est aussi pour les œuvres matérielles, Ainsi, la cathédrale « la Sagrada Familia » de Barcelone est et serait encore plus une énigme si une partie des plans et des maquettes de Gaudi n'avaient pu être conservée, comme un patrimoine aussi précieux que la cathédrale elle-même [Bonet 2004].

Exploration de l'évolution des contextes

Dans ce qui précède, nous avons parlé du contexte de l'écriture du document, comme si ce contexte était unique et stable. Or tout travail de création a une durée. Son contexte évolue au cours même de la genèse de l'œuvre. L'objectif de cette section est de présenter quelques aspects relatifs à l'évolution temporelle des contextes, ainsi qu'à la mise en évidence et à la prise en compte de cette évolution.

Traces et rôle du temps dans les documents

Outils informatiques pour gérer les contextes temporels

Les techniques informatiques ont permis de prendre en compte ce phénomène, en particulier pour conserver la mémoire des aspects temporels des processus de conception industrielle. Couramment des outils dits de « gestion de versions » sont utilisés professionnellement lorsque des équipes travaillent et font évoluer des objets qu'elles partagent. Chacun peut alors « voir » et utiliser un objet dans l'un de ses contextes temporels. Toutefois il s'agit là d'un temps discontinu, conservant des états identifiés, mais sans leurs intermédiaires.

Des travaux plus spécifiques sur la gestion de version [Shipman 2000] brisent cette barrière de version identifiée pour introduire un parcours continu dans le temps de la genèse de l'œuvre. Le lecteur dispose, en plus des barres de défilement spatial, d'une troisième barre qui explore le temps. Il peut alors, fixant spatialement son intérêt sur un point, explorer de façon continue l'évolution de celui-ci puis, ayant choisi une date donnée, explorer le reste du document tel qu'il était à cette date. Il s'agit ici d'une véritable navigation spatio-temporelle dans le contexte, car les liens vers d'autres documents sont, bien sûr, eux-mêmes temporels et permettent d'explorer l'environnement qu'avait l'œuvre à cet instant.

Ce type de technique s'applique aussi bien à des documents textuels, que multimédias, mais aussi à des documents techniques. Par exemple la prise en compte de l'aspect temporel dans les bases de données géographiques rend possible la visualisation de l'évolution progressive des transformations de parcelles, ou celle des interférences entre des travaux d'infrastructure successifs. La visualisation directe de ces phénomènes temporels donne accès à une catégorie de traces très utiles : l'évolution d'un contexte.

Le temps et l'analyse temporelle des œuvres artistiques

La genèse des œuvres littéraires, ou artistiques de façon générale, est souvent un phénomène long pendant lequel le contexte dans lequel baigne l'auteur peut avoir évolué. L'analyse des couleurs de certains tableaux de Monet resterait une énigme, si l'évolution de ses troubles de la perception des couleurs n'en apportaient l'explication. Pareillement, certains personnages de fresques peintes par Michel-Ange, jugés impudiques à certaines époques, ont été ultérieurement habillés de draperies. Quelle est l'œuvre réelle de l'artiste ? Son évolution et

l'évolution du contexte de celle-ci n'en sont-elles pas des aspects aussi... Et ne parlons pas du Requiem de Mozart...

Les écrivains souvent ne considèrent de leur œuvre que la forme finale. Le laborieux travail de genèse est le plus souvent masqué. Toutefois, des collectionneurs, des musées, des investisseurs, rachètent des manuscrits et des brouillons de manuscrits d'auteurs célèbres. Ces pièces disparates, souvent revendues, parfois dramatiquement éparpillées par petits lots chez des antiquaires au gré de successions, constituent un ensemble de traces précieuses pour retracer la genèse d'un œuvre [Lebrave 1992]. Contrairement aux techniques informatiques de gestion de versions, qui sont facilitées par la saisie homogène des données à la source, ici la reconstruction de l'historique d'une œuvre s'apparente à un travail de police scientifique, à la résolution d'une véritable énigme [Ferrer 1998]. L'auteur a écrit ces différentes pages, mais quand et dans quel ordre ? Il faut réordonner les feuillets pour mettre en évidence des séquences, retrouver les analogies entre plusieurs versions d'un même passage, identifier les ajouts, les retraites, les déplacements. Il faut trouver des indices pour la datation, le type de papier, sa marque, la composition de l'encre, le type de plume, etc. Par exemple, l'apparition de retouches au stylo bille de la main de l'auteur permet de dater ces retouches comme postérieures à l'invention de cet instrument. Une experte du domaine racontait même, qu'après un long travail de ré-ordonnancement, elle avait eu la surprise de constater que, en prenant le paquet de feuilles ainsi reconstitué, se trouvait sur un côté une première histoire sur des pages rayées, et à leur dos une autre version de l'histoire, plus complète, mais dont l'enchaînement des pages était précisément dans l'ordre inverse. Enigme ? Non, l'auteur avait simplement repris globalement le paquet de feuilles de la première version de l'histoire dont le verso était encore vierge. De ce fait, le verso de la dernière page était devenu le recto de la première de la nouvelle version. La mise en évidence de l'évolution de l'œuvre au cours de sa genèse et la mise en relation de faits externes et des variations apportées par l'auteur à son document créent un contexte interprétatif beaucoup plus proche de celui de l'auteur pour comprendre l'œuvre.

Interprétation et sens

Avant de conclure, il nous semble important de rappeler dans cette section qu'une lecture n'est qu'une interprétation contextualisée et qu'il serait dangereux de réifier une interprétation. La mécanique quantique nous fournit un exemple où une interprétation de phénomènes, une lecture du monde, est parfois assimilée à sa pure description.

La construction d'un contexte, au sens étymologique, ce qui va avec le texte, ce qui le rend compréhensible, donne au lecteur l'impression d'avoir résolu l'énigme posée par le document qu'il lit. Le lecteur d'un poème, l'admirateur d'un tableau, l'auditeur d'une sonate se prend rarement au jeu. Il a sa propre

interprétation et cela lui suffit. Mais il existe des interprétations qui ont un goût de vérité. De vérité absolue. D'universalité. C'est le cas lorsqu'une lecture savante a persuadé le lecteur que l'élaboration complexe d'un contexte, auquel il vient de parvenir, ne peut être qu'unique. Sa lecture devient alors un fait de science. Ce phénomène se rencontre dans toutes les disciplines. Le savant *lit* le monde, mais très vite il oublie que sa lecture n'en est qu'une interprétation, qu'elle conduit à l'élaboration d'une connaissance, que le monde n'est pas réductible à cette connaissance. Très vite on apprend à raisonner sur cette connaissance en tant que telle, cela est merveilleux, puissant, mais on finit par croire que l'on raisonne sur l'objet de cette connaissance.

La physique quantique est arrivée à un très haut degré de précision dans sa capacité à décrire et par suite à prédire des faits du monde. On peut se demander si elle ne serait pas seulement une description de ce que nous pouvons savoir du monde. Lorsque les raisonnements quantiques indiquent qu'un objet est simultanément dans plusieurs états et que le fait de l'observer le précipite dans l'un et un seul de ces états, est-ce l'objet lui-même qui a cette propriété encore énigmatique ou seulement notre connaissance de l'objet ? Quand les raisonnements quantiques indiquent qu'en observant l'un de deux objets liés mais inconnus, situés en des lieux différents, on fixe à distance les propriétés de l'autre, n'est-ce pas en réalité de la connaissance de ces faits dont nous parlons. Ne serait-il pas plus naturel de dire que la connaissance des propriétés de l'un nous renseigne en même temps sur les propriétés de l'autre ? Ces connaissances ne constituent-elles pas plutôt une lecture particulière du monde ?

La lecture d'un document, fut-il le monde, n'en est qu'une interprétation. Réifier cette lecture, assimiler le document à son interprétation est peut être une facilité qui fait vite oublier la complexité du mécanisme de lecture.

Conclusion

La lecture sous toutes ses formes implique la re-contextualisation d'un document. Que ce contexte soit élaboré émotivement dans une lecture artistique, transmis méthodiquement dans la lecture technique, reconstruit rationnellement dans une lecture savante, ce contexte ne peut appartenir qu'au lecteur. Préserver le contexte initial de l'auteur est certes un besoin, en particulier pour l'industrie, où la mémoire d'entreprise a pour objectif d'éviter de transformer en énigmes complexes l'interprétation de faits et de documents. Des techniques hypertextuelles constituent un moyen technique pour conserver et explorer de tels contextes, ainsi que leur évolution temporelle. Mais ces contextes eux-mêmes ne sont disponibles que comme des documents, recontextualisés lors de leur lecture. L'archivage des œuvres audiovisuelles atteste de la complexité de conservation de contextes interprétatifs. Le présent article, plutôt que de discuter doctement de chacun de ces points, a cherché à les présenter au travers

d'exemples, d'anecdotes, de traces suggestives placées là pour laisser au lecteur le soin d'élaborer lui-même le contexte interprétatif qui lui plaira.

Références

- Bachimont. B. De l'hypertexte à l'hypotexte : les parcours de la mémoire documentaire. *Technologies, Idéologies, Pratiques*. (Mémoire de la technique et techniques de la mémoire, sous la direction de Charles Lenay et Véronique Havelange), 1999, 195-225.
- Bélisle, C. *La lecture numérique : réalités, enjeux et perspectives*. Ouvrage collectif. Presses de l'enssib, 2004, ISBN 2-910227-51-0.
- BnF, dépôt légal, <http://www.bnf.fr/pages/infopro/depotleg/depotleg.htm>
- Bocconi, S., Nack, F., Hardman, L. Supporting the Generation of Argument Structure within Video Sequences. In *Proc. Hypertext '2005*, ACM Press. (To appear 2005).
- Bonet, J. *The Essential Gaudi*. 3^{ème} édition, Portic, 2004, ISBN84-7306-729-0.
- Ferrand, L. *Processus de base de la reconnaissance des mots écrits chez l'adulte*. Deboeck Université. 2001.
- Ferrer, D. Le matériel et le virtuel : du paradigme indiciaire à la logique de mondes possibles. In M. Contat et D. Ferrer eds., *Pourquoi la critique génétique ?*, 1998.
- Lebrave, J.-L., La critique génétique : une discipline nouvelle ou un avatar moderne de la philologie? *Genesis* N° 1. Avril 1992.
- Lebrave, J.-L., Du visible au lisible : l'édition génétique électronique. *Actes du Colloque Genèses*. 1998.
- Luc, C. Une typologie des énumérations basées sur les structures rhétoriques et architecturales du texte. In *Proc. Conf. TALN'2001*, Tours 2001.
- Maignien Y, La bibliothèque virtuelle ou de l'ars memoria à Xanadu in *Bulletin des bibliothèques de France*, 1995, t.40, n°2, p. 8-17.
- Merlot, M. *Nouvelles Alexandries. Les grands chantiers de bibliothèques dans le monde*. Ouvrage collectif. Editions du Cercle de la Librairie, 1996.
- Richaudeau, F. Mais qu'est-ce que lire?, dans *Communication et langages*, 74, 1987, p. 110-113.
- Shipman III F.M., Hsieh H. Navigable history: a reader's view of writer's time. *New Review of Hypermedia and Multimedia*, Vol. 6, 2000, Taylor Graham, 147-168.
- Stiegler, B. (1996). *La technique et le temps ; Tome 2 : la désorientation* », Galilée.
- Thibodeau, K. Building the Archives of the Future: Advances in Preserving Electronic Records at the National Archives and Records Administration. *D-Lib Magazine*, February 2001, Volume 7 Number 2.
- Valenti, J. Lecture, processus et situation cognitive, *Recherches sémiotiques*, 2000, p. 289-331.

Virbel, J. & Maignien, (1996). Encyclopédisme et hypermédias : de la difficulté d'être à la complexité du dire. *Tous les savoirs du monde. Encyclopédies et Bibliothèques de Summer au 21ème siècle. BnF – Flammarion,*



La confrontation des systèmes sémio-linguistique et musical en chanson : l'autre voie du sens

Martine GROCCIA,
UMR 5191 - Laboratoire ICAR (Interactions, Corpus, Apprentissages, Représentations)
Université Lumière Lyon 2
Martine.Groccia@univ-lyon2.fr

Mots clés : *sémiotique, chanson, syncrétisme, énonciation, expression, contenu.*

Résumé :

Notre travail de thèse porte sur la chanson en tant qu'objet syncrétique, réalisant simultanément deux systèmes sémiotiques distincts : le système sémio-linguistique, et le système musical. Notre postulat de départ est qu'une chanson ne se réduit ni à son instance verbale, ni à son instance musicale, et que l'étude de son sens doit prendre en compte non seulement ces instances isolées, mais aussi les aboutissants de leur interaction.

Dans le cadre des 13^{ème} Journées de Rochebrune, portant sur la thématique « Traces, Enigmes, Problèmes : Emergence et construction du sens », nous choisissons d'étayer l'hypothèse selon laquelle la saisie du sens d'une chanson, et partant son interprétation par un auditeur, s'élabore notamment dans des tensions entre le textuel et le musical, tensions qui déconstruisent potentiellement la sémosis instaurée par le verbal, et offrent à l'interprétation de l'auditeur un substrat de signifiante, à investir d'éventuels interprétations secondes. Ces interprétations secondes, accessibles à l'auditeur par l'intermédiaire de sa saisie sensible de l'objet, restent néanmoins énigmatiques et virtuelles, dans la mesure où elles ne résultent pas de l'opération d'une sémosis, et ne sont pas appréhendables comme contenus actualisés. La signification de l'objet reste alors corrélée à sa saisie perceptive par un auditeur, et pose par conséquent les questions, et de la position de l'énonciataire en tant que Sujet de l'Énonciation, et de la clôture du texte en tant que tout de signification.

1. La construction du sens dans la composante verbale

L'analyse impose la déstructuration de l'objet selon ses deux composantes principales, le texte et la musique. Cette déstructuration peut avoir pour conséquence l'élaboration de fragments d'analyse qui ne peuvent rendre compte de l'appréhension de l'objet en tant qu'il est perçu par un auditeur-énonciataire, qui lui, reçoit syncrétiquement l'ensemble des données. Il s'agit donc dans un premier temps d'explicitier quelques phénomènes discursifs qui construisent la cohérence du discours, et partant sa signification, et non pas de rendre compte de la construction du sens par un auditeur-énonciataire. Nous proposons donc une lecture globale des contenus verbaux à travers deux points succinctement abordés : la mise en place d'un univers axiologique binaire ; la construction de la relation entre les acteurs.

Couplet 1

1 Si t'allais dans les musées
Tu t'srais bien amusé
Cette peinture de Manet
C'est qu'ça vaut d' la monnaie
5 C'est l'déjeuner sur l'herbe
Lève pas le nez en l'air
Des filles nues ça d'vrait t'plaire
Et les mecs en imper

Refrain 1

Des femmes en dessous et des hommes
10 Et des hommes en pardessus et des femmes
Car l'homme l'homme est bête et mauvais
Ça oui ça oui on le sait
Mais sa femme vaut pas mieux qu'lui
Ça y est ça y est j'l'ai dit
15 J'te l'ai dit j'te l'aurais bien dit

Couplet 2

T'es bête à manger d'l'herbe
Tu m'fiches ma vie en l'air
Tu passes tes nuits dehors
Mais si tu crois qu' je dors
20 Toi et ta pouf dans l'herbe
Quand tu t'envoies en l'air
Dans les étages sans bruit
Ici on s'reproduit

Refrain 2

Les femmes les femmes en dessous des hommes
25 Et les hommes les hommes par dessus les femmes
Car l'homme l'homme est bête et mauvais
Ça oui ça oui on le sait
Mais sa femme vaut pas mieux qu'lui
Ça y est ça y est j'l'ai dit
30 J'te l'aurais bien dit

Pont

On s'aime on s'chamaille
On s'aime on déraile
Tu piques ta crise
Toutes ces horreurs ça résiste à l'analyse
35 Mais l'amour finit toujours

Refrain 3

les femmes qu'ont le dessous et les hommes
Et les hommes qu'ont le dessus et les femmes
Car l'homme l'homme est bête et mauvais
Ça oui ça oui on le sait
40 Mais sa femme vaut pas mieux qu'lui
Ça y est ça y est j'l'ai dit
J'te l'ai dit j'te l'aurais dit

Refrain 4

Les femmes qu'ont le dessous et les hommes
Et les hommes qu'ont le dessus et les femmes
45 Car l'homme l'homme est bête et mauvais
Ça oui ça oui on le sait
Mais sa femme vaut pas mieux qu'lui
Ça y est ça y est j'l'ai dit
J'te l'ai dit j'te l'aurais dit

1.1. Instauration d'un univers axiologique binaire

Dans la mesure où cette chanson est prise en charge par un interprète féminin, qui chante et offre donc une expression « incarnée » au narrateur, et dans la mesure où dans un discours à la première personne, le narrateur (ou Sujet de l'énonciation énoncée) et le héros (Sujet de l'énoncé) sont, par définition, coréférentiels et renvoient à une même variable actorielle, *ego*, deux positions actérielles sont distribuées dans ce texte, et s'investissent de deux identités antinomiques : le « je » acteur féminin, et par complémentarité socialement et thématiquement normée, le « tu », acteur masculin.

La disjonction, dès le premier couplet, entre l'acteur féminin, qui fréquente les lieux culturels, et l'acteur masculin, qui ne les fréquente pas, pose pour chacun des sujets deux modes d'existence au monde, et par là-même, permet d'instituer pour l'ensemble de la chanson deux univers de valeurs antinomiques : les valeurs ici potentialisées par la convocation d'un objet d'art (finesse d'esprit, intelligence, beauté, sensibilité, etc.) deviennent constitutives de la figure féminine, et par opposition paradigmatique entre les deux acteurs, la figure masculine s'investit de valeurs contraires. Se construit ainsi une axiologie où la figure de l'homme supporte une valorisation /négative/ : vulgarité, trivialité, laideur, et celle de la femme une valorisation /positive/ : distinction, intelligence, beauté.

Par ailleurs, l'ensemble de la chanson développant les thématiques du conflit et de l'incompréhension entre les hommes et les femmes, et l'énonciation étant prise en charge par le personnage féminin, les valorisations effectivement réalisées au niveau discursif, seront celles déterminant l'univers axiologique du sujet masculin, le corollaire féminin se construisant *in absentia*, par l'intermédiaire de valeurs virtualisées. L'acte même d'assomption du discours, par l'opération de débrayage par rapport à l'*hic et nunc* du sujet-acteur, fonde le Sujet de l'énonciation énoncée, ici la figure féminine, comme détenant un savoir et un croire supplémentaire par rapport au narrataire, et permet d'appréhender ces valorisations comme un savoir assumé, ou un jugement.

Les figures convoquées dans le discours sont alors investies de ces valorisations, et permettent dans le texte de construire une lecture dysphorique des isotopies de l'argent, du pouvoir et du sexe. Ainsi par exemple, dans le premier couplet, les figures « monnaie », « filles nues », « mecs en imper », tout en assurant la permanence de la représentation du tableau de Manet, s'investissent de significations triviale et sexuelle.

1.2. La relation entre les acteurs

Les deux positions d'acteurs, l'homme vs la femme, s'actualisent dans des rôles actanciels, localisables dans l'organisation textuelle, et subissent une transformation progressive aboutissant à la construction d'acteurs génériques. Le parcours actoriel dessine en effet trois positions pour les acteurs, chacune cumulant le rôle actanciel précédemment actualisé dans la position précédente :

« je » vs « tu » → « les hommes » vs « les femmes » → « l'homme » vs « sa femme »

Chacune de ces positions correspond à un développement discursif particulier, et permet d'aboutir dans les refrains à un traitement argumentatif de la thématique principale : la relation entre les deux sexes.

1. « je » vs « tu » ou la mise en discours d'une relation de couple.

Ces deux actants interviennent plus spécifiquement dans les couplets, dans lesquels ils assument le rôle des protagonistes féminin et masculin d'une relation de couple individualisée.

- Les deux acteurs prennent place dans le récit par le biais du programme narratif mis en place dans le couplet 1, et introduit par la phrase hypothétique des vers 1 et 2 "si t'allais dans les musées, tu t'srais bien amusé". Ce programme narratif pose d'emblée la disjonction entre l'acteur féminin, qui fréquente les musées, et l'acteur masculin, qui ne les fréquente pas. Cependant, les thèmes de l'argent et de la sexualité, actualisés par les figures « monnaie », « filles nues », et « mecs en imper » signalent une relation de familiarité, voire d'intimité entre les deux personnages.

- La relation entre les acteurs est spécifiée dans le couplet 2. Le développement d'un discours de reproches, syntaxiquement repérable par la répétition, presque à chaque début de vers, du pronom personnel "tu", correspond à une mise en accusation du personnage masculin par le personnage féminin. La thématique de l'adultère, développée aux vers 18 à 21, et de la femme blessée au vers 16, inscrit les deux personnages dans une relation sentimentale, et donc de couple en crise.

- Le pont permet de réunir les deux protagonistes. Les deux acteurs se réalisent dans un seul actant "on" (vers 31/32). L'actant révélant ainsi la conjonction des acteurs, jusque là disjoints, les procès actualisent une tentative de résolution des conflits, à chaque fois repositionnée dans une situation d'échec (« on s'aime », réitéré, mais corrigé par les deux procès suivant "on s'chamaille", vers 31, et "on déraile", vers 32). La négociation par le discours s'avère ici inefficace, et est sanctionnée par une nouvelle séparation des deux acteurs à travers la reprise

du "tu" aux vers 33 et 34. L'issue positive des négociations est définitivement évincée au vers 35 "mais l'amour finit toujours", qui se pose comme une sanction.

2. « les femmes » vs « les hommes » ou la mise en place des relations de pouvoir

Les deux premiers vers de chaque refrain mettent en place un actant non plus individuel mais collectif. « les femmes » vs « les hommes » sont manifestés dans des syntagmes nominaux qui figurent une répartition des places, et sanctionne le rapport relationnel entre les acteurs :

- refrain 1 : "des femmes en dessous et des hommes / et des hommes en pardessus et des femmes" (vers 9/10). Les figures actualisées, « dessous » et « pardessus », placent les deux acteurs aux antipodes d'un axe /- vêtu/ vs /+ vêtu/, et par métaphore, actualisent, entre autres, les sèmes /démuni/ vs /paré/.

- refrain 2 : "les femmes les femmes en dessous des hommes / et les hommes les hommes par dessus les femmes". Les figures actualisées dans le refrain précédent prennent ici valeur de métaphore spatiale, et manifestent une représentation de l'acte sexuel.

- refrains 3 et 4 : "les femmes qu'ont le dessous et les hommes / et les hommes qu'ont le dessus et les femmes". Ici, la nominalisation de "dessous" et "dessus" permet d'actualiser une isotopie du pouvoir, déjà implicitement présente au refrain 2. et répartit les places de pouvoir entre les deux acteurs.

L'ensemble des syntagmes permet l'actualisation dans le discours des lieux communs disponibles dans le savoir encyclopédique du lecteur : l'homme est le sexe fort vs la femme est le sexe faible.

3. « l'homme » vs « sa femme » ou le développement argumentatif.

Aux vers 11 et 13, les deux acteurs acquièrent un statut d'acteur générique, et sont manifestés comme actants d'un développement argumentatif. « car l'homme l'homme est bête et mauvais » (vers 11) est articulé aux constats précédemment exposés par un lien de causalité, tout en se posant comme un argument d'autorité, fondé avant tout sur une opinion communément admise ("ça oui ça oui on le sait", vers 12). Il permet ainsi de justifier l'ensemble des valeurs négatives attribuées à l'acteur masculin dans le discours. Dans la continuité de l'opposition des acteurs jusque là construite dans le discours, il provoque l'attente de l'actualisation de son corollaire déductif : "la femme est intelligente et bonne". Or, la proposition effectivement manifestée "sa femme vaut pas mieux que lui" (vers 13) détruit cette opposition systématique. Tout d'abord, la détermination de l'acteur féminin par le possessif « sa » modifie le traitement relationnel entre les deux acteurs : le rapport de possession ainsi manifesté permet de reconstruire la figure du couple, et rétablit le lien entre les deux acteurs, jusque là disjoints et dans l'impossibilité de rétablir ce lien. La conséquence en est que le statut générique de l'acteur féminin subit une restriction : non plus la femme, mais l'épouse de l'homme. Par ailleurs, par équivalence syntagmatique, cette seconde proposition acquière elle aussi le statut d'argument d'autorité, alors que son contenu ne correspond pas à une opinion communément admise. L'ensemble de cette rhétorique, en projetant brutalement la figure féminine dans le même système de valeurs que l'homme, produit un effet de chute.

Nous avons pointé ici quelques phénomènes locaux de la construction de la signification de cette chanson, dans son discours verbal. qui permettent de proposer une lecture globale de son sens. Le discours construit ici propose un simulacre d'argumentation, sensé soutenir la thèse selon laquelle l'homme assoit son statut social et individuel grâce à sa vilénie, ce qui non seulement maintiendrait la femme dans un certain avilissement, mais signerait également l'incompatibilité entre les deux sexes dans la mesure où le mode d'existence au monde de la femme s'appuie sur un régime de valeurs antinomique. Or, la chute du développement

argumentatif actualise une thèse en sus, non pas contraire, mais qui épouse justement la première, transformant en quelque sorte le discours jusque là produit en une vaste syllepse qui ouvre un autre univers de discours.

2. La perception de l'objet sonore, et la saisie du sens

Nous avons affirmé que le texte de chanson, en tant que manifestation sémio-linguistique, est prééminent quant à la manifestation des contenus de l'objet, et donc quant à la construction de son sens. C'est en effet lui qui offre les fondements de la signification à partir desquels l'auditeur construit sa lecture. Cependant, la dimension sonore et musicale de l'objet inscrit cette sémosis du verbal dans une saisie perceptive globale de l'objet syncrétique. Dans la mesure où il est impossible de séparer les étapes logiques de la perception, de l'élaboration sémiotique et de la représentation, il apparaît évident que l'univers musical est susceptible d'infléchir la lecture de l'auditeur, et de lui fournir, dans ses tensions avec les données linguistiques, non pas des significations, mais un substrat de signifiante à partir duquel il pourra construire sa signification..

Nous proposons ici un développement en trois temps : dans un premier temps, nous posons une description succincte, mais suffisante ici, de la composante musicale. Nous mettrons en évidence dans un second temps quelques éléments musicaux susceptibles d'être prégnants pour l'auditeur, afin d'engager ensuite une réflexion sur les conséquences quant à la saisie qu'il peut faire du sens de l'objet.

2.1. Description de la composante musicale

1. Structure globale de la chanson :

Décrire la structure globale permet de mettre en évidence les phénomènes de répétition mis en place dans la chanson. Or, les structures répétées assurent différentes fonctions pour l'appréhension sensible de l'objet par un auditeur : elles articulent, et hiérarchisent entre eux les éléments sonores, en découpant dans le continuum sonore des unités discrètes et identifiables ; elles posent des équivalences paradigmatiques entre ces unités ainsi identifiées ; elles mettent en évidence les éléments qui eux, ne sont pas répétés, et acquièrent ainsi un statut spécifique ; et enfin, elles posent une norme d'exécution des éléments répétés, rendant ainsi saillant aux oreilles de l'auditeur tout écart à cette norme.

Dans cette chanson de Maurane, nous pouvons donc identifier la structure suivante :

- une introduction : elle se divise en deux parties différentes :

- une première partie lente, jouée en rubato, c'est-à-dire dans une interprétation du tempo assez libre, et sans mesure proprement dite. Sa mélodie constitue une annonce du thème qui va suivre. Cependant, cette mélodie n'est pas vraiment perçue en tant que telle mais constitue plutôt un fond sonore sur lequel se développe un brouhaha de conversations, avec à peine plus en avant de la masse sonore, la voix de l'interprète, qui discute. Nous appellerons cette partie de l'introduction "prologue", en référence à l'épilogue qu'on trouve à la fin, et qui se démarque lui aussi par l'utilisation d'un rythme particulier.

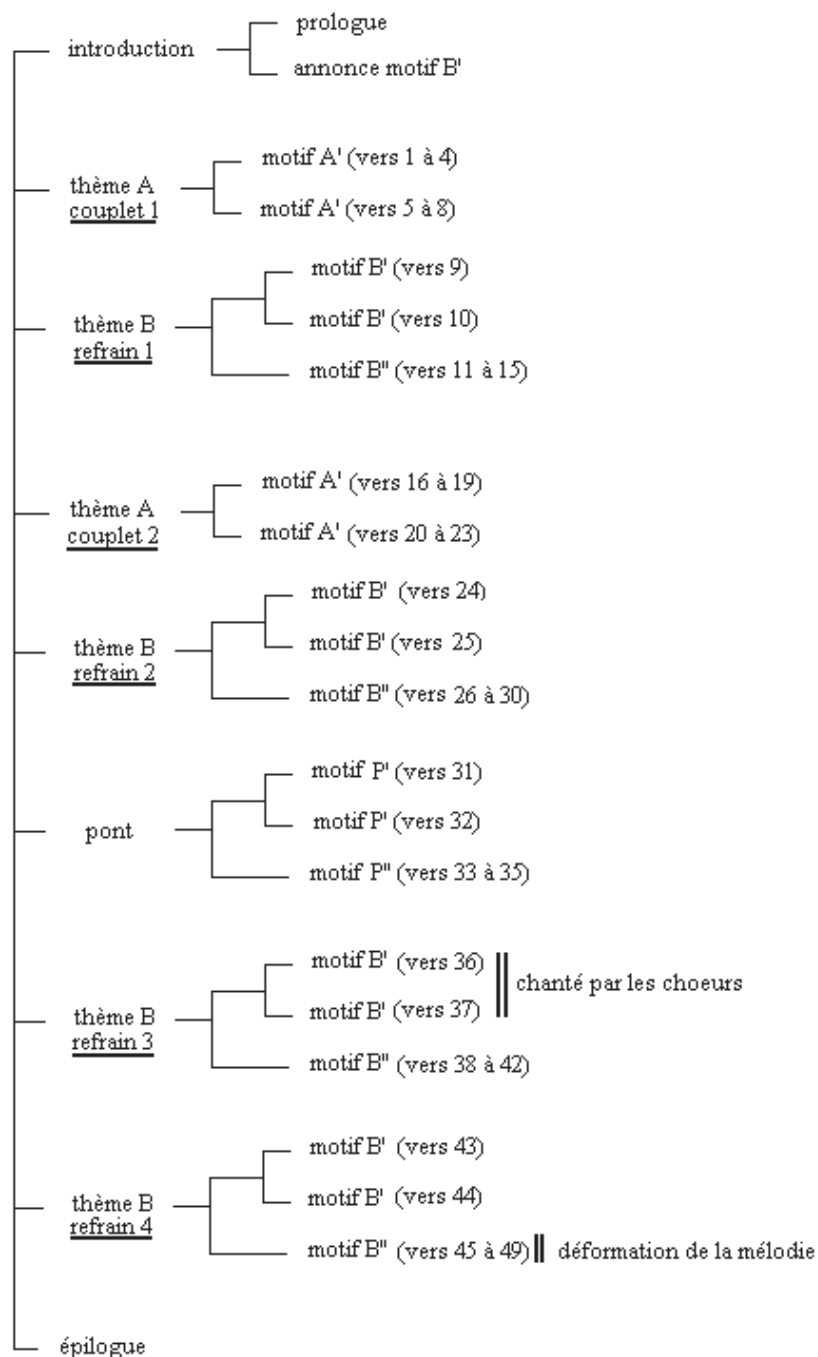
- une seconde partie plus rapide (sur 4 mesures) qui impulse le tempo du reste de la chanson, et qui introduit un thème chanté (qui s'avèrera correspondre au motif B' du thème B). Cette seconde partie fonctionne comme la véritable introduction du développement musical qui suit, notamment à cause de son tempo et de son développement mélodique.

- des thèmes chantés A, B et C, qui correspondent respectivement aux couplets, refrains et pont.

Il y a quatre refrains dans cette chanson : les deux premiers sont répétés à l'identique, les deux suivants présentent des spécificités d'exécution : le troisième est marqué par la prise en charge des deux premiers vers par les chœurs, le quatrième est marqué par une déformation de la mélodie.

- un épilogue qui ne reprend aucun motif déjà présent dans la chanson. Il est musicalement complètement innovant par rapport au reste de la chanson.

La chanson s'organise donc globalement selon le schéma suivant :



2. Instrumentation, rythme, mélodie

L'instrumentation, le rythme et la mélodie sont les trois données principales à prendre en compte pour la description de l'univers musical d'une chanson. Elles sont constitutives du

paysage sonore dans lequel est plongé le sujet lorsqu'il écoute la chanson. Ce paysage, en tant que son, investit le corps écoutant, et constitue l'enveloppe sonore dans laquelle se vit l'intimité entre soi et la sensation de l'expérience du son.

→ instrumentation : elle est très riche et hétéroclite. De nombreux instruments sont utilisés dont les principaux sont : la batterie, complétée sur certains passages par des percussions, la clarinette basse et le saxophone basse, le piano, le violon. A cet ensemble s'ajoute une présence vocale assez importante : la voix de l'interprète, mais aussi un ensemble de chœurs, qui en plus de prendre en charge une partie de refrain, participe à la construction des bruits de la chanson.

→ rythme : il est complexe et joue un rôle particulier dans l'orchestration. Nous retiendrons ici uniquement le fait qu'il permet de délimiter trois grandes parties de la chanson :

- le prologue, joué en rubato, et donc sans aucune mesure particulière
- le corps même de la chanson joué sur un tempo assez rapide, 120 à la noire environ, globalement¹ sur une mesure binaire à quatre temps (une mesure à 4/4).
- l'épilogue : son tempo est deux fois moins rapide que la chanson elle-même (60 à la blanche pointée) et surtout, son rythme est construit sur une mesure ternaire².

→ mélodie : construite sur deux tonalités principales (si♭ majeur et fa♯ majeur), la mélodie suit la complexité de l'ensemble des éléments musicaux, et exploite de nombreuses harmonies. Nous retiendrons seulement ici qu'elle est marquée par un changement assez extraordinaire, qu'on a noté "déformation de la mélodie" au refrain 4. La chanteuse, dans cette partie de la mélodie, effectue un glissando ascendant qui la mène pratiquement deux tons plus haut que la mélodie initiale, et, performance vocale, elle tient cette transposition sur presque tout le motif B" (la tonalité initiale est récupérée sur le dernier vers) alors que l'accompagnement des instruments maintient l'harmonie initiale, produisant ainsi un effet de dissonance très marqué.

3. Interprétation

Dans la saisie perceptive de l'auditeur, pour laquelle nous tentons de fournir ici des éléments descriptifs, l'interprétation est le siège d'enjeux bien spécifiques (que nous aborderons plus loin) dans la mesure où c'est par elle que s'énonce le corps chantant de l'interprète, véritable Sujet de l'énonciation incarné. Dans cette chanson, l'interprétation est principalement prise en charge par Maurane, excepté sur les vers 36 et 37 où elle est assumée par des chœurs.

La voix de Maurane est une voix pleine, puissante et assez grave pour une voix de femme, c'est une véritable voix d'alto. Cependant, Maurane possède une tessiture assez large, et elle effectue des sauts d'intervalles importants (le plus grand étant un écart de 9^{ème} au début de B" sur "car l'homme"). Par ailleurs, la chanteuse joue également sur de nombreux apports de voix non chantée : bruitages, rires sarcastiques, etc.

Quant aux chœurs, ils se divisent en un chœur de femmes et un chœur d'hommes. Le plus important à noter ici est que la mélodie est interprétée une octave au-dessus de l'interprétation de Maurane pour les femmes, et une octave en dessous pour celle des hommes, ce qui a pour effet d'agrandir encore l'écart mélodique dans lequel se développe la chanson.

2.2. Tensions et distorsions entre texte et musique

¹ Un passage particulier abandonne cette mesure à quatre temps, mais nous ne l'aborderons pas ici.

² Une mesure ternaire correspond, par exemple, à un rythme de valse : le temps, au lieu de se diviser en deux et multiples de deux comme dans les mesures binaires, se divise ici en trois et multiples de trois. On a ainsi un premier temps accentué + deux temps faibles : 123/123 etc.

Parmi les éléments musicaux rapidement exposés, nous retiendrons trois phénomènes sonores susceptibles d'entrer en tension avec les contenus de la composante verbale, et par conséquent de distordre le processus sémiotique à l'œuvre dans la construction du sens par l'auditeur.

1. Abondance et diversité des données musicales vs univers discursif binaire

L'univers musical de cette chanson est complexe : les tonalités utilisées sont nombreuses³, la mélodie truffée d'altérations et les rythmes nombreux et variés. L'utilisation des voix de chœur et des bruits participe encore de cette complexité musicale.

Par ailleurs, le prologue et l'épilogue, qui assurent l'ouverture et la fermeture de l'objet, présentent des univers musicaux qui leur sont propres, et se détachent donc de l'ensemble des structures répétées. Dans le prologue, sur le brouhaha des voix et la conversation qu'on devine de l'interprète, se greffe un développement mélodique constitué de quatre notes (ré-mi-b-do-ré^b), formant des intervalles de seconde mineure. La superposition des bruits, à valeur référentielle, et de la cellule mélodique, construit l'image mentale d'une scène de suspense, à caractère cinématographique, et déclenche des scénarios possibles dans l'imagination de l'auditeur. Ce traitement sonore a pour effet d'ouvrir le champs des possibles quant à la prédiction des développements ultérieurs..

L'épilogue, quant à lui, développe un univers sonore tout à fait étrange. L'entrée dans cet épilogue se fait de manière soudaine : la cadence parfaite du dernier refrain est coupée dans sa résolution, le discours lui-même subit cette coupure (le texte de la chanson se termine sur "j'te l'aurais", le syntagme restant inachevé, sur cette rupture agrammaticale). L'auditeur se trouve ainsi comme aspiré dans le nouvel espace sonore proposé. Les développements mélodiques ne correspondent à rien de déjà entendu dans la chanson, et la mesure à trois temps, en rupture avec le reste de la chanson, donne à l'ensemble sonore un mouvement de tourbillon. S'ajoutent encore des voix inquiétantes et des rires sarcastiques.

Ainsi, l'ensemble des données, instrumentales, rythmiques et mélodiques, construit un paysage sonore foisonnant et hétéroclite, qui apparaît comme non homologue à la représentation mentale d'un univers binaire, tel que le propose le discours.

2. Des traitements isotopiques divergents

La musique établit d'emblée une échelle de perception verticale dans la mesure où elle se caractérise par la hauteur des sons /grave/ vs /aigu/. Cette échelle verticale rend disponible, sur le plan de l'expression, une isotopie bas/haut, susceptible d'entrer ou non en correspondance avec une isotopie homologue, sur le plan des contenus verbaux. Dans cette chanson, le traitement des acteurs dans le discours les positionne justement dans une répartition des places /position basse/ (pour les femmes) vs /position haute/ (pour les hommes). Or, la composante musicale peut malmener ce traitement isotopique, en faisant émerger de la matérialité même du signifiant musical, une lecture antiphrastique de l'isotopie manifestée en discours. C'est l'effet produit ici dans la reprise du refrain par les chœurs (refrain 3). Les vers 36 et 37 sont chantés par les chœurs d'hommes et de femmes selon l'organisation suivante :

"les femmes qu'ont le dessous" ⇒ chanté par les femmes

"et les hommes / et les hommes qu'ont le dessus" ⇒ chanté par les hommes

"et les femmes" ⇒ chanté par les femmes

³ Nous avons repéré au moins quatre tonalités différentes : si^b majeur, fa[#] majeur, do⁷, mi^b majeur.

Précisons que les voix de femmes sont très haut perchées (une octave au-dessus de la mélodie initiale) et que les voix d'hommes sont très graves (une octave en dessous). On a donc les deux plans suivants manifestés syncrétiquement :

| Contenu verbal | Expression musicale |
|--------------------------|---------------------|
| « femmes » + « dessous » | Voix très aiguës |
| « hommes » + « dessus » | Voix très graves |

L'antiphrase naît de la confrontation des places affectées aux acteurs dans le discours avec la matérialité des voix elles-mêmes, qui placent les acteurs dans un rapport inverse : les femmes en haut de l'échelle musicale, et les hommes en bas de cette même échelle. Par ailleurs, cette antiphrase est soutenue par la réalisation sonore du syntagme (dans le déroulement de l'exécution, les voix de femmes enserrant celles des hommes), et par l'interprétation des voix féminines, nasillardes et moqueuses. On peut noter également que l'intervalle mélodique entre les deux voix produit, par son ampleur, un effet de caricature.

3. Interprétation de la chanteuse

La chanteuse intègre, dans son interprétation purement musicale, des marqueurs à caractère intonatifs, que l'on peut qualifier de marqueurs thymiques dans la mesure où ils révèlent un état affectif du Sujet énonçant dans le *hic et nunc* de l'exécution de la chanson. On a par exemple au vers 35 une manifestation flagrante de ce phénomène : "mais l'amour finit toujours" est interprété sur un ton ostensiblement grotesque, et se réalise en . L'ajout de la semi-consonne, la duplication de la syllabe, le roulement du R, confèrent au vers une sonorité à la fois burlesque et railleuse. Autre illustration de ce phénomène : le rire sarcastique que la chanteuse émet à la fin du refrain 1. Enfin, la déformation mélodique du refrain 4 peut s'analyser comme une sorte de débordement de ces marqueurs thymiques sur l'exécution de la mélodie, débordement qui a pour effet de désagréger le pattern mélodique, qui s'actualise alors dans une exécution non conforme. Cette déformation crée une forte dissonance, et apparaît comme l'actualisation, sur le plan sonore, d'un débordement d'ordre passionnel, qui place au premier plan les affects du Sujet.

2.3. Saisie du sens par l'auditeur-énonciataire

Les quelques phénomènes sonores que nous venons de mettre en évidence sont prégnants pour l'auditeur, dans sa saisie sensible de l'objet. Or, cette saisie perceptive influe directement sur le processus sémiotique qu'il engrange pour la construction de sa signification, par l'intermédiaire de deux phénomènes principaux, liés à la dimension sonore de l'objet.

1. Le statut particulier de la voix chantée

Comme nous l'avons signalé, la voix en chanson a la particularité d'offrir une expression incarnée au Sujet de l'énonciation. Or, cette caractéristique est susceptible de déclencher un processus de conversion des instances énonciatives, et offre à l'énonciataire la possibilité d'occuper la place de l'énonciateur. En effet, la voix qui chante, en incarnant un sujet du monde, construit lors de l'écoute, ce que Fontanille et Zilberberg⁴ nomment un champ de présence, c'est-à-dire un domaine spatio-temporel où s'exerce la perception de l'objet. Ce champ de présence est investi des caractéristiques de la voix du sujet énonciateur, et de sa réalisation effective. Se développe ainsi un rapport particulier entre l'énonciation de la

⁴ Fontanille J., Zilberberg C., *Tension et signification*, Ed. Mardaga, Liège, 1998. p92

chanson chantée et le je/ici/maintenant de l'auditeur. Ce rapport tient dans le fait que l'acte d'énonciation ainsi effectué installe dans une certaine coprésence temporelle les deux acteurs de la communication : à l'instant T où l'auditeur écoute une chanson, la seule propagation des ondes constituant la voix du sujet chantant, saisies dans la perception de l'auditeur/énonciataire, donne corps, au sens littéral du terme, à ce protagoniste absent de l'échange, l'énonciateur. Par ailleurs, dans la mesure où *"l'expression 'incarnée' installe le sujet qui parle dans un rapport in praesentia avec sa sensation, et par empathie, cherche à installer celui qui écoute dans un rapport identique."*⁵, ce corps chantant de l'énonciateur, auquel l'auditeur a accès grâce à sa performance vocale, l'incite, par empathie, à s'identifier à lui.

Par conséquent, la voix invite celui qui écoute à sentir dans son propre corps ce que le corps de l'autre exprime, ce qu'il rend manifeste dans la matérialité même du signifiant vocal, et offre ainsi la possibilité à l'auditeur de l'investir de ses propres affects. Dans cette optique, la signification de l'objet reste inévitablement corrélée à l'interprétation de l'énonciataire qui, bien qu'étant orienté dans sa saisie de la signifiante musicale par certains éléments objectivables du matériau sonore, implique forcément dans l'élaboration de la signification, son identité propre, elle-même sans cesse dépendante de son état de Sujet au monde, dans l'espace-temps de la réception de l'objet.

2. Le caractère illimité de l'objet sonore

Dans son article traitant du sensible et des modalités de la sémiotique⁶, Jean Fisette expose la thèse selon laquelle la logique de l'audible échappe au point de vue. Il explique que les caractères diffus et difficilement localisables de l'origine du son rendent la notion de point de vue inopérante pour les objets sonores, et par-là même, ces objets s'en trouvent illimités. Or, cette illimitation met à mal l'exhaustivité même de l'objet, nécessaire pour la construction d'une totalité signifiante et partant pour la clôture de sa signification. En d'autres termes, un sujet placé dans un environnement à prédominance sonore se trouve dans l'impossibilité d'une saisie exhaustive de l'objet, et est alors placé dans la nécessité d'assumer, dans l'ordre du savoir, l'incertitude.

En conclusion, l'analyse ciblée de cette chanson de Maurane nous aura donc permis de rendre compte de quelques phénomènes en jeu dans la construction du sens d'une chanson par un auditeur-énonciataire. Au delà des données textuelles qui relèvent de la sémiotique du verbal, la composante musicale intervient également, et à plusieurs titres, dans cette construction de la signification. Elle structure notamment le continuum sonore à travers la gestion des formes répétées, et elle assigne à ces formes des patterns mélodiques, rythmiques, et interprétatifs.

Cependant, son rôle diverge de celui assumé par la composante verbale, dans la mesure où elle ne construit pas une correspondance entre expression et contenu. Mais elle peut, dans le cadre de la modalité que nous avons abordée ici, déstabiliser la sémiotique du verbal, notamment en rendant prégnants pour l'oreille de l'auditeur des écarts et des tensions. Ces écarts et tensions constituent un ensemble d'indices pour accéder à des degrés d'interprétation supplémentaires, indices qui fonctionnent pour l'auditeur comme autant de brèches qu'il pourra investir de significations liées à sa perception propre de l'objet, perception unique et individuée de sujet ultime de l'énonciation. Cependant, cette fonction du musical laisse néanmoins l'auditeur dans l'impossibilité de résoudre l'énigme de la signification de l'objet.

⁵ De Chanay Hugues, *La voix d'opéra : sémiologie et rhétorique*, in « Puissance de la voix : Corps sentant, corps sensibles », Nouveaux Actes Sémiotiques, PULIM, 2001.

⁶ J. Fontanille, J. Fisette, *Le sensible et les modalités de la sémiotique*, Revue Tangence n°64

Des trajectoires individuelles à la ville en mouvement : reconstruire les mobilités urbaines quotidiennes

Arnaud Banos

SET (UMR 5603 CNRS)

Université de Pau et des Pays de l'Adour

Avenue du Doyen Poplawski, 64000 Pau

arnaud.banos@univ-pau.fr

<http://www.univ-pau.fr/~banos/banos.html>

La ville change, s'étale, se dilue, prend ses aises. Au sein des nouveaux « *archipels urbains* » (Beaucire, 1997), discontinus et multipolarisés, les distances s'allongent, les densités baissent. Le mouvement devient la norme, dans cet « *espace transactionnel* » sillonné par des « *nomades-urbains* » pressés (Beauchard, 1999), aux trajectoires toujours plus complexes. L'avènement des « *quatre roues de la fortune* » a considérablement amplifié cette tendance, ouvrant de nouveaux territoires de conquête à ces aventuriers du quotidien, autorisant des pratiques individuelles plus riches, renouvelant les termes de cette « *lutte contre l'espace* » dont parlait Fernand Braudel. Toujours plus nombreux, les mobiles urbains irriguent quasiment en permanence cette trame urbaine faite pour et par l'automobile, consommant des espaces plus étendus qu'auparavant. Les trajectoires individuelles se lient et se délient, au gré des pérégrinations effectuées par chacun, ajoutant à la complexité des motifs spatiaux tissés en permanence comme par une main invisible, particulièrement habile et facétieuse.

Confronté à cette incroyable effervescence, cette vie bouillonnante et grouillante, le géographe est désarmé. Appréhendée à un niveau microscopique, la « fourmilière » urbaine semble en effet n'être qu'amalgame de comportements individuels particuliers, spécifiques, non réductibles les uns aux autres. Et pourtant, des structures plus ou moins nettes, plus ou moins durables, semblent émerger par moment de cette invraisemblable diversité, pour peu que l'on accepte de prendre de la hauteur, de s'éloigner temporairement des individus pour observer le territoire. L'objectif est alors de reconstruire ces multiples traces individuelles qui conduisent à l'émergence de la ville en mouvement. En créant des mondes artificiels, informatiques, peuplés d'agents artificiels autonomes agissant au sein de réseaux d'interaction à géométrie variable, le géographe se dote dès lors de plateformes d'expérimentation d'un genre nouveau, lui permettant de reproduire *in silico* des situations observées afin de mieux les étudier. Mieux comprendre le monde qui nous entoure, à

travers l'exploration de mondes artificiels, devient à ce titre l'un des enjeux clés de la discipline géographique.

1. Reconnaître la complexité des phénomènes étudiés

Mieux connaître la mobilité quotidienne à l'échelle d'une ville est un enjeu de taille. Nulle politique d'aménagement urbain ne saurait en effet aujourd'hui faire l'impasse sur ce sujet brûlant qui, au-delà de sa dimension « transport », renvoie à la fonction essentielle et fondamentale de la ville : permettre et favoriser les interactions sociales entre ses habitants. Mais c'est également un défi, scientifique et technique. Comment en effet appréhender une telle diversité, une telle richesse ? Comment prétendre embrasser du même regard cette multitude de comportements individuels si différents ? Le parti pris traditionnel est de réduire cette complexité de manière systématique. Tout d'abord en réduisant l'objet même d'étude et en le ramenant à des dimensions maîtrisables. La mobilité est ainsi souvent simplifiée, réduite à ses composantes de base (une série de déplacements, souvent considérés comme indépendants, entre deux activités). L'espace géographique est également souvent simplifié, par une opération de zonage spatial, permettant de raisonner sur des sous-ensembles spatiaux moins nombreux et plus homogènes, donc plus facilement maîtrisables. Les données d'enquêtes, recueillies auprès d'échantillons d'individus, sont elles-mêmes le plus souvent simplifiées au moyen d'opérations d'agrégation, permettant de raisonner non plus sur des individus mais sur des groupes sociaux ou spatiaux, au moyen d'indicateurs statistiques généraux. La diversité (le terme statistique clé est variabilité) des comportements individuels se trouve ainsi souvent réduite à un comportement moyen, plus simple à modéliser, mais qu'on ne retrouve évidemment jamais dans la rue. « L'homme moyen » de Quételet, qui fut décrié dès le départ comme ne pouvant être qu'un monstre, perdure toujours dans les sciences humaines et sociales mais également économiques, sous la forme d'un « individu moyen », abstrait, aux caractéristiques parfois invraisemblables, mais aptes à être formalisées et modélisées. Enfin, la mobilité est très rarement appréhendée dans sa composante dynamique, alors même que le mouvement est au cœur du phénomène.

Ces critiques doivent évidemment être replacées dans un canevas plus large, par exemple celui du fonctionnement particulier de la recherche en tant qu'activité sociale inscrite dans un contexte historique donné et alternant des phases de stabilité (ce que le philosophe Thomas Kuhn a proposé d'appeler des phases de science normale) avec des phases de transition, de changement de « paradigme ». Or, le paradigme réductionniste, justement centré sur l'idée d'une simplification par réduction et décomposition de tout objet d'étude (ici la mobilité quotidienne), est sérieusement questionnable, et ce d'autant plus que le renouvellement des questionnements scientifiques, des stratégies d'investigation et des options technologiques ouvrent la voie à d'autres approches de la

fourmière urbaine. En particulier, nous défendons depuis plusieurs années l'idée d'une approche exploratoire qui, mise en œuvre à bon escient, permettrait de révéler la complexité inhérente des phénomènes étudiés (ce qui suppose évidemment d'accorder une telle qualité aux phénomènes en question), afin de mieux l'étudier, la comprendre et au-delà en tirer partie, dans l'aide à la décision et l'action. Quelques exemples permettront d'imager notre propos.

2. Révéler la complexité !

L'un des exemples les plus classiques de simplification par agrégation dans le domaine de la mobilité et du transport est illustré par les deux premières étapes de la figure 1 : 1) le découpage de l'espace en un nombre réduit de zones supposées homogènes en terme de mobilité ; 2) la création de matrices origines/destination, estimant le nombre de déplacements émis et reçus par chacune de ces zones.

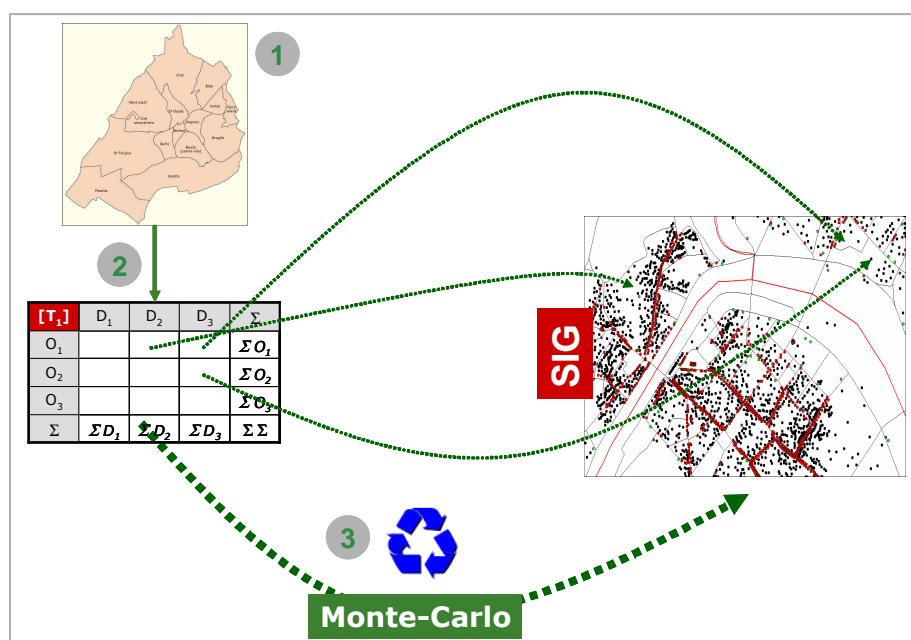


Figure 1. Désagréger l'agrégé par simulation

Toute tentative de désagrégation d'une telle information implique par conséquent la mise en œuvre d'un processus de modélisation et de simulation, à l'image du processus stochastique suggéré à l'étape 3. L'objectif est ici d'affecter chacun des 800 000 déplacements quotidiens estimés à Besançon à l'intérieur des zones définies, au bâti urbain, en fonction des données de la matrice et ce pour chaque période d'un quart d'heure, de manière à produire une image plus fine de la manière dont l'espace urbain est occupé au cours de la journée¹. Sur la figure 2², on observe ainsi ces fluctuations continues de l'espace urbain, qui se vide et se remplit en permanence au cours de la journée, en

¹ Pour une présentation détaillée de la démarche, consulter (Banos et Thévenin, 2005)

² L'animation complète est visible à l'adresse suivante : <http://www.univ-pau.fr/~banos/geovisualisation.html>

fonction des déplacements des individus qui l'occupent. Révéler et identifier les respirations d'une agglomération, d'une région urbaine, à travers l'agencement sans cesse renouvelé des espaces « pleins » et « vides » au cours d'une même journée, constituent indéniablement des enjeux urbains forts, qui justifient un investissement méthodologique supplémentaire. Toutefois, à se focaliser ainsi sur l'espace, on en vient malheureusement souvent à oublier les individus ! La question qui se pose est alors : est-il possible de dépasser ce premier niveau d'analyse, de manière à retrouver la trace de ces individus toujours plus mobiles qui sillonnent les artères urbaines ?

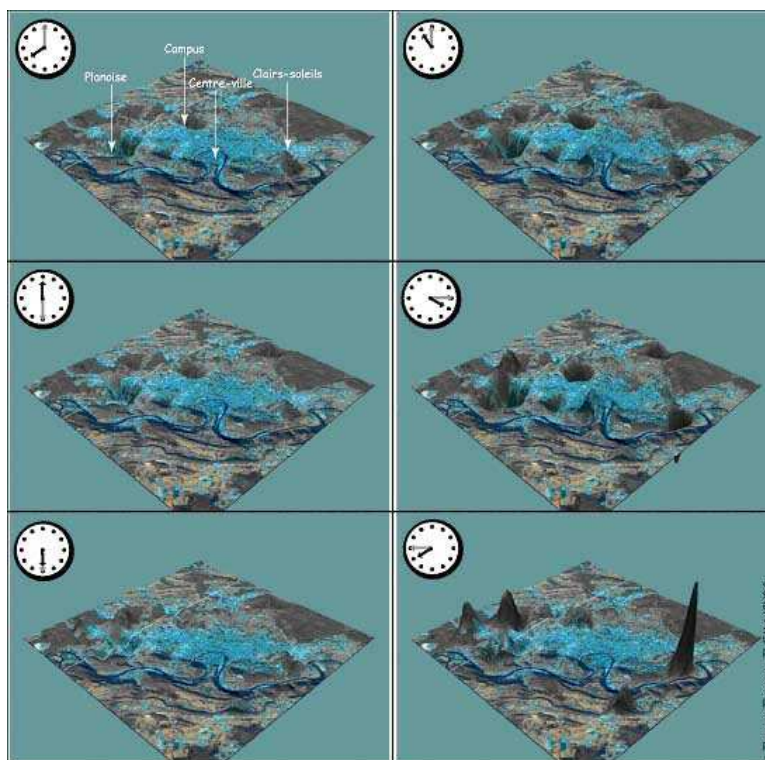


Figure 2. Reconstruire par simulation les « palpitations » du territoire (ici Besançon)

C'est précisément ce genre de questionnement qui nous a guidé, dans le cadre de travaux menés en collaboration avec deux chercheurs en informatique du Laboratoire d'Informatique de l'UPPA, Bruno Jobard et Julien Lesbegueries (Lesbegueries, 2005). Notre objectif était d'exploiter de manière originale une autre base de données classique dans le domaine de la mobilité quotidienne, l'Enquête Ménage-Déplacements, disponible dans la plupart des grandes agglomérations françaises. Cette base de données regroupe des informations particulièrement détaillées sur l'ensemble des déplacements réalisés par un échantillon représentatif d'individu, un jour donné. La figure 3 illustre le type de données disponibles pour chaque individu.

On notera la richesse des informations fournies sur les activités réalisées, les modes de transport utilisés, ainsi que les horaires. On notera également que la dimension spatiale est plus problématique : la localisation exacte des individus n'est pas connue. Chaque activité est ainsi localisée approximativement dans

une zone de déplacement, tandis que les trajets réellement effectués par les individus dans le réseau routier ne sont pas recensés.

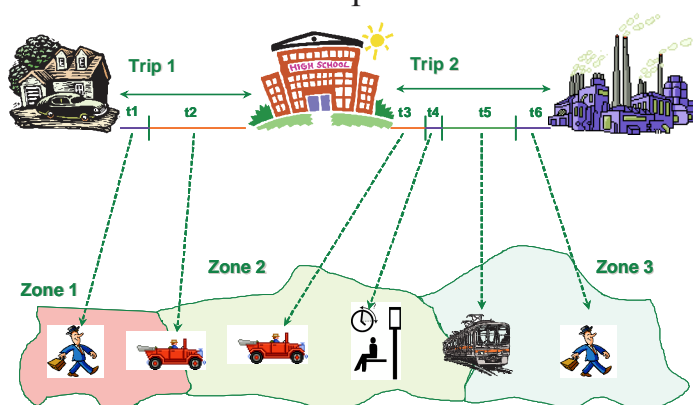


Figure 3. Les données individuelles de mobilité fournies par l'Enquête Ménage-Déplacements (Source : Zeitouni et Wang, 2005)

A partir de cette base de données, et en comblant les lacunes identifiées, est-il possible de reconstruire les déplacements effectués par les individus enquêtés ? Une telle entreprise suppose de formuler des hypothèses de travail permettant de compenser les données manquantes. Ainsi, à titre d'exemple, nous assumons que les individus empruntent toujours l'itinéraire le plus court en temps entre une origine et une destination données. Cette hypothèse, qui peut paraître audacieuse voire irréaliste, se retrouve pourtant dans la quasi-totalité des modèles de déplacement mis en œuvre dans la plupart des pays depuis les années 60-70. Une telle hégémonie trouve une double justification :

- ✓ Théorique : assumer que chaque individu est capable d'un tel tour de force revient à l'inscrire *ipso facto* dans le cadre théorique aujourd'hui encore dominant de l'acteur rationnel, ou Homo Oeconomicus : la variabilité des comportements individuels peut être résumée par un comportement synthétique moyen, qui serait celui d'un individu moyen, parfaitement rationnel, en situation d'information illimitée, doté de capacités de calcul infinies et cherchant toujours à maximiser son utilité en adoptant un comportement optimisateur. Bref, capable de toujours trouver le plus court chemin dans un réseau routier.
- ✓ Technique : une telle hypothèse trouve enfin une justification technique immédiate, grâce aux travaux menés de longue date en recherche opérationnelle dans le domaine du parcours de graphe. Il existe en effet différentes familles d'algorithmes aux comportements et propriétés parfaitement connus, permettant d'identifier le plus court chemin entre deux points dans un réseau.

La mise en œuvre de ce genre de modèle au sein d'un Systèmes d'Information Géographique, environnement informatique permettant de gérer, d'exploiter et d'analyser l'information territorialisée, permet ainsi de reconstruire les déplacements et itinéraires « probables » (au sens mathématique du terme) des

individus enquêtés...soit près de 13 000 à Lille ! Se pose alors la question de la visualisation des résultats obtenus. Car révéler la complexité possède un coût, entre autre lié à notre capacité à la montrer ! La figure 4 montre un exemple de solution retenue dans le cadre de ce projet³.

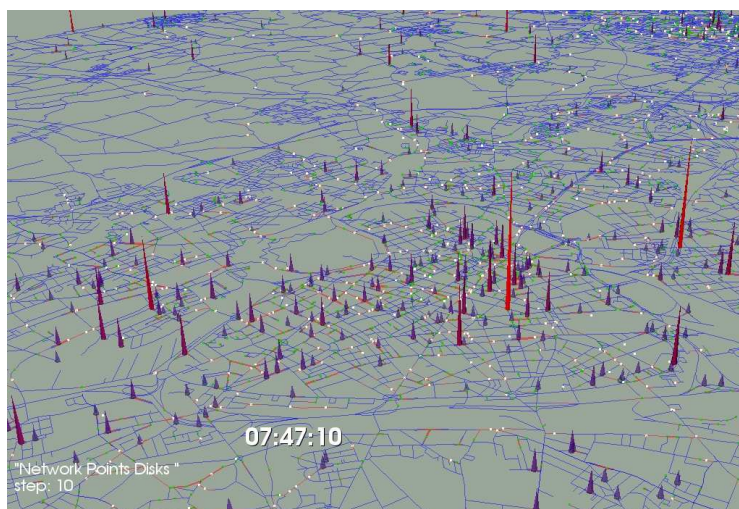


Figure 4. Reconstruire les traces individuelles (ici à Lille)

Le réseau routier figure en bleu, les individus mobiles (se déplaçant) sont représentés sous forme de points blancs, ceux immobiles à l'instant t sous forme de points verts. Les pics expriment une mesure du nombre d'individus présents sur le réseau à proximité, pour un seuil de densité donné. Cet arrêt sur image correspond à la situation simulée à 07h47.

Il est légitime de s'interroger sur la portée d'un tel travail. A l'évidence, les capacités visuelles de l'être humain, même le plus doué ou le mieux entraîné, sont insuffisantes pour suivre les déplacements de 13 000 individus...quant à parler du million d'habitants de l'agglomération lilloise...Disons le clairement, à l'heure actuelle, ce type de réalisation a une vertu avant tout démonstrative et pédagogique : pointer du doigt la complexité inhérente de l'objet d'étude, afin d'éviter que la batterie d'hypothèses toutes plus lourdes et contraignantes les unes que les autres (et parfois posées sans même y penser) soit égrenée *a priori*. Toutefois, il paraît évident qu'une telle vue, forcément partielle, ne prend son sens qu'au sein d'un éventail de vues complémentaires⁴. Diversifier les points de vue et les éclairages sur un même problème est en effet l'un des principes de base de l'analyse exploratoire des données, telle qu'elle a été définie dans les années 1970 par le mathématicien américain John Tukey (Tukey, 1977).

Ces deux exemples qui auront permis, nous l'espérons, d'illustrer le fond de notre propos, auront peut être éveillé chez le lecteur une question insidieuse quoique légitime : une fois la complexité des phénomènes révélés,

³ L'animation complète est visible à l'adresse suivante : <http://www.univ-pau.fr/~banos/etudiants-travaux.html>

⁴ Une série d'animations complémentaires est visible à partir de la même adresse.

qu'elle est l'étape suivante ? Deux options s'offrent à nous. L'option analytique, très liée au parti pris réductionniste cité précédemment, nous incite à déconstruire cette complexité, en décomposant le tout en sous-ensembles plus simples (les parties), jusqu'à ce que le cœur même du ou des phénomènes étudiés apparaisse, tel un diamant débarrassé de sa gangue. Une intuition aujourd'hui largement répandue, plus connue sous le credo systémique « le tout est plus que la somme des parties », suggère toutefois que cette course à la décomposition analytique est par essence vouée à l'échec : même si nous connaissions l'ensemble des facteurs intervenant dans les actions individuelles, nous ne serions pas pour autant capables de prédire le comportement d'un groupe d'individus, compte tenu des multiples interactions à l'œuvre. Cette limite fondamentale nous a alors poussé, dans le cadre du projet SAMU, à nous inscrire dans une famille théorique que nous qualifierons de large, regroupant des travaux variés autour des systèmes complexes, de la simulation de sociétés artificielles et de la simulation à base d'agents et visant à mieux comprendre les phénomènes complexes non plus en les déconstruisant, mais en les reconstruisant (Epstein J., Axtell R., 1996 ; Resnick, 1997 ; Schelling, 1978).

3. Reconstruire la complexité !

Le projet SAMU, pour Simulation Agents et Modélisation Urbaine (Banos et al., 2005), mené en collaboration avec Sylvain Lassarre de l'INRETS⁵ et Abhimanyu Godara de l'Université Banaras Hindu, en Inde, trouve son origine dans un constat simple : en dépit de son rôle fondamental dans le système des mobilités quotidiennes urbaines, la marche à pied reste le mode de déplacement le moins bien connu. De manière générale, si la France dispose depuis les années 1970 d'outils d'observation standardisés des mobilités à l'échelle urbaine (enquêtes ménages-déplacements, enquêtes emploi du temps, enquêtes nationales transport, enquêtes cordons), force est de constater la faible adéquation de ces dispositifs à l'appréhension des déplacements piétons. Diffus, quasi isotropes, souvent de courte portée et dans une grande mesure stochastiques, ces derniers n'en constituent pas moins une part considérable de la mobilité quotidienne, la marche à pied s'intégrant forcément au cœur de toute chaîne de déplacement. Longtemps délaissée, la marche à pied est aujourd'hui revalorisée, portée par l'aspiration générale à des modes de vie urbains moins agressifs, protégés des externalités négatives engendrées par l'invasion automobile. Mieux comprendre les stratégies de déplacement des piétons dans un environnement urbain mouvant, caractérisé à la fois par un ensemble de contraintes et d'opportunités, mais également de nuisances et de dangers, est devenu l'un des principaux enjeux actuels de la planification urbaine. Un certain nombre de travaux s'attachent depuis peu à explorer ce domaine, notamment dans la perspective des systèmes complexes, mais rares sont ceux abordant le

⁵ Institut National de Recherche sur les Transports et leur Sécurité

problème de la dynamique des déplacements piétons en interaction avec le trafic routier, qui est pourtant l'une des dimensions clés de l'environnement urbain vécu. C'est cette lacune que le projet SAMU cherche à combler. L'hypothèse centrale formulée dans ce projet est la suivante :

en dépit de la complexité avérée des mécanismes cognitifs en jeu, la dynamique d'ensemble des déplacements piétons dans un environnement mouvant doit pouvoir être approchée à partir de règles individuelles simples, mais intégrées dans un réseau d'interactions en évolution permanente.

Dans cette perspective, le projet SAMU⁶ vise à créer un environnement urbain virtuel et à le peupler d'entités autonomes, piétons et automobilistes, animées par des motifs propres et agissant dans le cadre d'une architecture distribuée de type multi-agents. Ce projet s'inscrit par ailleurs dans un cadre théorique large, la théorie des systèmes complexes. Cette théorie, qui a été vulgarisée par plusieurs ouvrages quelque peu romancés et/ou romantiques (Benkirane, 2002 ; Johnson, 2001 ; Waldrop, 1992), regroupe une variété considérable de travaux le plus souvent transversaux, reposant toutefois sur un certain nombre de principes communs. Un système dit complexe est ainsi constitué d'un grand nombre d'entités en interaction, situées dans un environnement. Les entités agissent dans et sur cet environnement, qui les influence en retour.

Par ailleurs, dans SAMU chacune de ces entités, dénommé « agent » d'après la terminologie en vigueur dans le domaine, est guidée par des règles de fonctionnement le plus souvent simples, de type stimulus-réaction. Le fait que ces stimuli soient eux-mêmes situés dans l'environnement, suffit à produire des réactions individuelles différenciées, à partir d'un canevas invariant de règles. Un tel système ne bénéficie pas, de plus, d'un mode de contrôle global, centralisé. Le pouvoir d'action des agents est réduit à une dimension très locale, et certaines structures globales observées sont le fait de processus d'auto-organisation. Dans une telle perspective, les multiples interactions, qui plus est localisées, entre agents peuvent conduire à l'apparition de propriétés à un autre niveau d'observation ou d'agrégation, intermédiaire ou global, non déductibles à partir des simples propriétés des agents. Ces propriétés sont dites émergentes et leur identification constitue l'un des principaux enjeux des théories de la complexité. Enfin, certains systèmes complexes sont également qualifiés d'adaptatifs, afin de souligner leur capacité à évoluer, essentiellement par mutation ou par apprentissage, soit au niveau des règles de fonctionnement des agents, soit au niveau des agents eux-mêmes.

L'informatique permet de donner corps à ce cadre théorique, par la création de mondes artificiels, au sein desquels évoluent des agents également artificiels, et qui puissent devenir observables au même titre que le monde réel. Ainsi que le

⁶ SAMU est développé au sein de la plateforme NetLogo : <http://ccl.northwestern.edu/netlogo/>

souligne Alexis Drogoul, chercheur en informatique spécialiste des Systèmes Multi-Agents, « *ce ne sont plus seulement les résultats de la simulation qui deviennent objets d'attention, mais également les processus souvent complexes par lesquels ces résultats ont été obtenus* », ce qui introduit une composante expérimentale novatrice dans l'étude des phénomènes sociaux. D'après ce même auteur, « *observer un monde artificiel offre de grands avantages du point de vue de la simulation : tous les paramètres, depuis les lois physiques de l'environnement jusqu'à la définition des acteurs qui le peuplent, peuvent être modifiés, simplifiés ou complexifiés. Tout en lui permettant de conserver sa méthode de travail, l'environnement expérimental va donc permettre au chercheur d'adopter une démarche spéculative, fondée sur la constitution d'hypothèses difficilement vérifiables en nature* » (Drogoul, 2000).

Dans la première version de SAMU⁷, l'environnement urbain est simplifié à l'extrême : des rues et des trottoirs. Des voitures sillonnent les rues au hasard, en respectant le sens de circulation, tandis que des piétons localisés au hasard convergent tous vers le même point par le plus court chemin. Des curseurs permettent de spécifier le nombre d'agents et leur comportement : capacité d'accélération et de décélération pour les voitures (en fonction de la vitesse du véhicule devant et/ou de la présence de piétons dans un champ de vision variable) ; portée de la vision pour les piétons, à laquelle s'ajoute une capacité d'accélération et de décélération en phase de traversée, en fonction de la présence de voitures dans le champ de vision. Les agents évoluent ainsi en parallèle et adaptent leur vitesse en fonction des interactions locales auxquelles ils prennent part.

Des graphiques mis à jours en permanence permettent de suivre l'évolution de certains paramètres agrégés tels que la distribution des vitesses minimum et maximum des véhicules ou celle de la proportion de piétons "écrasés". Enfin, un module cartographique permet de suivre en permanence l'occupation des cellules par les agents, mais également de visualiser la répartition des accidents.

Cette idée de reconstruire (générer) la complexité d'un phénomène par l'identification d'un ensemble réduit (nécessaire et suffisant) de briques élémentaires, insérées dans un réseau d'interactions approprié, soulève évidemment toute une série de questionnements épistémologiques, logiques, méthodologiques et même techniques dont nous ne prétendons pas ici ne serait-ce qu'esquisser les contours.

⁷ Cette première version de SAMU peut être testée en ligne depuis le site <http://www.univ-pau.fr/~banos/sma.html>

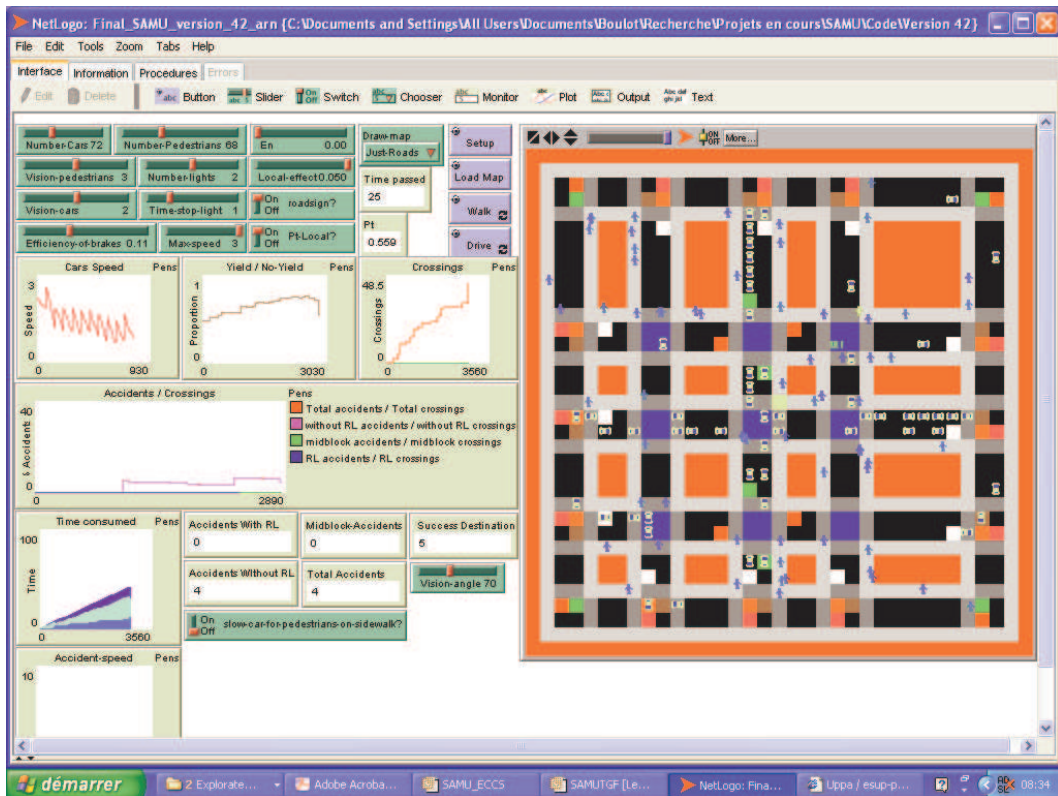


Figure 5. Le prototype SAMU

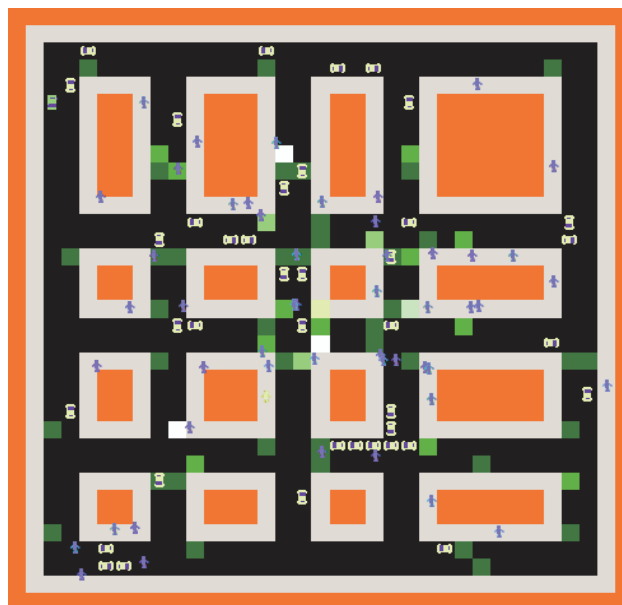


Figure 6. Un exemple de résultat : les accidents de piétons

Nous nous contenterons de souligner un point essentiel : une telle entreprise peut elle aboutir sans une feuille de route cohérente et suffisamment détaillée ? Or, force est de constater qu'un tel cadre n'existe pas vraiment à l'heure actuelle. Le modélisateur peut, au mieux, se laisser guider par un nombre réduit de principes, afin d'éviter de créer des modèles trop riches et compliqués, dont les résultats seraient aussi difficile à analyser et comprendre que le phénomène initialement

étudié. Citons ici le principe d'économie de moyens et celui de parcimonie qui occupent une place centrale dans le domaine. Le principe d'économie de moyens exprime ainsi une idée clé, déjà évoquée : au lieu d'intégrer autant de détails que possible dans le modèle, il vaut mieux fonctionner à l'économie et chercher à identifier les détails strictement nécessaires et suffisants permettant de générer des structures émergents (par exemple un embouteillage ou une distribution statistique d'accidents). Le principe de parcimonie, également dénommé rasoir d'Occam en référence aux travaux du philosophe William d'Occam, suggère de retenir, parmi toutes les explications possible d'un phénomène, la plus simple. Ainsi, dans SAMU, les embouteillages sont des phénomènes émergents, non programmés en tant que tels dans le prototype, et qui naissent de l'interaction locale des véhicules : chaque voiture accélère ou décélère en fonction du comportement de la voiture située devant elle, dans son champs de vision, mais avec des micro variations aléatoires, liées au fait que chaque conducteur peut difficilement réagir exactement de la même façon que les autres (pensons au temps de réaction par exemple). Ce principe simple, traduit par une règle de comportement au niveau des agents « voiture », permet néanmoins de générer des embouteillages (phénomène émergent à l'échelle de la rue ou de la ville virtuelle), pour certains paramètres d'accélération ou de décélération.

4. Conclusion

Les enjeux actuels de la gestion des villes incitent à la mise en œuvre de stratégies d'investigation multi-scalaires, à même de prendre en compte les territoires de la mobilité, mais également de remonter jusqu'aux traces spatiales laissées par les acteurs mêmes de cette mobilité. Or, un tel projet s'inscrit dans un cadre largement contraint par les données existantes. La désagrégation de données agrégées est ainsi une nécessité, qui implique la production de protocoles de simulation spatiale au sein de systèmes d'information géographique. Utiles, ces protocoles ne sauraient toutefois permettre de remonter réellement jusqu'aux traces individuelles, à moins de gonfler déraisonnablement le stock d'hypothèses de base. L'exploitation de données plus désagrégées, recueillies auprès d'échantillons réduits d'individus, et *a priori* mieux adaptées aux objectifs fixés, se heurte toutefois à d'autres difficultés, dont la généralisation n'est pas la moindre. Enfin, une troisième voie complémentaire consiste à générer de toute pièce ces traces individuelles, au sein de laboratoires informatiques virtuels permettant de spécifier les caractéristiques et règles comportementales de base minimales susceptibles de déboucher sur l'émergence de comportements collectifs et de macro-structures qui fassent sens. Cette approche séduisante n'en demeure pas moins particulièrement délicate à mettre en œuvre, notamment du point de vue de la validation des sorties du modèle. Ces trois approches complémentaires illustrent toutefois, nous l'espérons, la richesse des solutions qui s'offrent aujourd'hui au

géographe dans son exploration de la diversité des trajectoires individuelles à l'origine de la ville en mouvement

5. Références Bibliographiques

- Banos A., Chardonnel S., Lang C., Marilleau N., Thévenin T., 2005 : Simulating the swarming city : a MAS approach, Proceedings of the CUPUM'05 Conference on Computers in Urban Planning and urban Management, London, 29-30 Juin, 17 p.
- Banos A., Godara A., Lassarre S., 2005 : Simulating pedestrians and cars behaviours in a virtual city : an agent-based approach, Proceedings of the European Conference on Complex Systems, Paris, 14-18 Novembre, 4 p.
- Banos A., Thévenin T., 2005 : La carte animée pour révéler les rythmes urbains, Revue Internationale de Géomatique, Vol 15, n° 1, pp. 11-31
- Beauchard J., 1999 : La bataille du territoire, L'Harmattan, Paris, 144 p.
- Beaucire F., 1997 : « Il faut relancer la planification urbaine », Transport Public, n° 959, pp. 40-44.
- Benkirane, 2002 : La complexité, vertiges et promesses, Le Pommier, Paris
- Drogoul A., 2000 : Systèmes multi-agents distribués, Habilitation à Diriger des Recherches, Université de Paris 6, 117 p.
- Epstein J., Axtell R., 1996 : Growing artificial societies : social science from the bottom up, Brookings Institution Press, MIT Press, Washington DC
- Johnson, S., 2001 : Emergence: The Connected Lives of Ants, Brains, Cities, and Software, Scribner, New York.
- Kuhn T., 1983 (1962 pour la première édition) : La structure des révolutions scientifiques, Flammarion, Paris, 284 p.
- Lesbegueries J., 2005 : Reconstruction et visualisation des déplacements d'une population urbaine, Mémoire de DEA en Sciences Informatiques, Université de Pau, 38 p.
- Resnick M., 1997 : Turtles, Termites, and Traffic Jams: Explorations in Massively Parallel Microworlds, MIT Press
- Schweitzer F., 2003 : Brownian Agents and active particles Springer-Verlag
- Schelling T., 1978 : Micromotives and Macrobehavior, W.W. Norton & Company
- Tukey J.-W., 1977 : Exploratory data analysis, Addison-Wesley, Reading, Massachusetts, 688 p.
- Waldrop M., 1992 : Complexity : the emerging science at the edge of chaos, Simon and Schuster, New-York

Dealing with the *not yet known* as a leverage for innovative policy making: the case of a regional ICT program

Sylvie Occelli

IRES Istituto di Ricerche Economico e Sociali del Piemonte,
via Nizza 18,
10125 Turin, Italy
Occelli@ires.piemonte.it

ABSTRACT. One widespread perception about changes in today human organizations is a diffuse sense of incapacitated ability to act as determined by a tension between the unexpected or unintelligibility traits of many situations and an unconcerned, unaware and careless attitude on the part of the observer. These issues are particularly evident in today policy on innovation, where ontological uncertainties accompanying the attributions of new system functionalities required by innovation have shaken the conventional approaches at the roots. In this paper an effort is made to sharpen some of these questions by considering as a reference, the case of an Information and Communication Technologies policy program which the Regional Government of Piedmont has recently launched as a leverage to foster regional development.

Key words: ICT, innovation, policy initiatives, pro-active attitudes, appreciative settings.

1. Introduction

One widespread perception about changes in today human organizations is a diffuse *sense of incapacitated ability to act* as determined by a tension between the unexpected or unintelligibility traits of many situations and an unconcerned, unaware and careless attitude on the part of the observer.

Although acknowledged in the domain of complexity approaches, the debate about the developmental paths of cities and regions has paid only scanty attention to that perception and its underlying tension. At most, it has made it possible to probe for their existence while investigating some of their phenomenological traits, i.e. uncertainties, insecurity, increasing risky situations, and the links with the unfolding of events occurring at the different spatio-temporal scales of system evolution.

This has also affected policy making where, in the last decade, increasing attention has been paid to the realization of decision support activities, either in the form of computer based methodological aids or within institutionalized analytical procedures, i.e. national/ regional observatories and project impact analysis such as those required by some EU policies.

This sense of incapacitated ability to act is particularly evident in today policy on innovation, where as some authors have recently emphasized, the acknowledgement of ontological uncertainties accompanying the attributions of new system functionalities required by innovation have shaken the conventional approaches at the roots (Lane, 2004).

In this paper, an effort is made to sharpen some of these questions by considering as a reference, the case of an Information and Communication Technologies (ICT) policy program which the Regional Government of Piedmont has recently launched to foster regional development.

In the remainder of the text discussion develops as follows. First, a reference is made to a conceptual model of innovation recently put forward by Pollone and Occelli (2005) in which novelty in the conception of innovation itself is both a drive and a major dimension of its deployment.

Then, a few consequences of the suggested model are emphasized as far as ICT policies in Piedmont are concerned. Their discussion shows that whereas, on the one hand, the effectiveness of many existing procedures for policy implementation is increasingly scrutinized and in some cases severely criticized, on the other one, the acknowledgment of dealing with unknown situations raises a sense of *incapacitated ability* to act which may hold back engaging into new actions.

While acknowledging this conundrum, the latest ICT regional program seems to have taken up the challenge. Unknown situations are viewed as a lever for innovative courses of policy action. Some general implications this lever which are briefly sketched in the last part of the paper are:

- the fact that the unknown is perceived in a different way. It is not a source of uncertainty to be got rid of, rather, it is a situation to be dealt with permanently. The unknown is simply what is not yet known;
- the fact that dealing with the unknown entails to engage into actions. It calls for a search into new ways to apprehend a situation relatively to a given context (i.e. the *milieu* and territorial features characterizing a certain area), requiring in certain cases new procedures and a change in the behaviours of the stakeholders;
- the fact that the above search does not take place in a vacuum. It has its major references in a given context, spatially, institutionally, and cognitively delimited, but needs to extend beyond it.

2. Challenging features of Information Communication Technologies

In the recent debate about ICT emphasis has been put to the fact it is that they entail innovation, and, more specifically, a peculiar type of innovation.

Underlying it, in fact, are two inter-linked components (see Occelli, 2001, 2003, Occelli and Staricco, 2001):

- the *hard component*, which is the tangible part of the innovation, i.e. a type of commodity, a technological artefact, an idea, a behaviour;
- and the *soft component*, a bundle of intangible aspects related to its operation, which enable the creation of knowledge potentials for both the users, i.e. novel possibilities for agents to learn about their surrounding world and the organizations users belong to, i.e. new roles and functionalities for firms, institution, cities and regions.

This is schematically represented in Fig.1 which maintains that innovation is made by up of three interlinked elements:

- Information and Communication Technologies (ICT), which consist of the set of artefacts set forth by the technological progress (computers, telecommunication networks, cellular phones and the hybrids resulting from the convergences of the former);
- Information, broadly understood here as the whole continuum of transferable knowledge which can be conveyed by means of ICT, i.e. all the information related activities deemed to affect the value-chain (performance) of a transaction;
- Functionality, here defined as the operating activities making it possible the working of organization, which is supported and/or enabled by the coupling of ICT and Information.

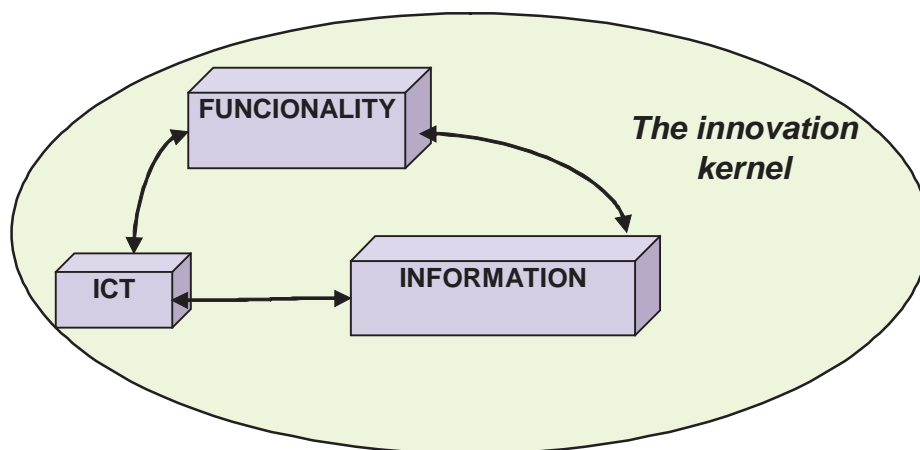


Figure 1 The innovation kernel

A foremost fundamental aspect of innovation is that it is inherent to the evolution of human organizations. This would explain why: a) it is searched for in spite of its usually costly investment and b) positive expectations about new

functionalities of innovation are recognized as major drivers of this search (Witt, 2004)¹.

In addition, as emphasized in Tab. 1, novel situations require interlinked processes of expression and evaluation, i.e. the former being the manifestation of novelty, the latter referring to both the acknowledgment and assessment of its realization. Evaluation, in particular is essential for: a) recognizing the appearance of a certain innovation, (i.e. attributing meaning to it), b) favouring its diffusion and c) creating the conditions for its encoding (consolidation) within the existing socio-economic structure².

Associated with the manifestation of novel situations is the production of a whole set of modifications, whose outcomes ultimately depend on how the innovation kernel of Fig.1, will be operated upon in the evolution path of systems³. As far as the operation of the evolutionary path is concerned, three main stages have been identified in the literature (see Graham, 2002), which are related to certain perspectives of analysis of the relation between ICT and regional systems.

Table 1. Features of innovation in relation to its generation and evaluation processes

| | <i>Evaluation process</i> | | |
|-------------------------------|---|--|--|
| <i>Generation process</i> | Usefulness | Viability | Critical aspects |
| Dissatisfaction of status quo | A) Socio-economic upgrading according to reference criteria in the activity system. Improving activity performances | B) Easing constraints and assessing compatibilities of targeted upgraded socio-economic situations. Supporting competitiveness of cities and regions | E) Intrinsic area non-homogeneity of ICT potentials. Worsening existing disparities. Exposures to external exploitations |

¹ According to Witt (2004), these expectations can be explained in terms of hypotheses related to “the individuals’ past present disposition as the motivation for action” (p.10). Major complementary hypotheses about the driving disposition would concern:

- dissatisfaction with the status quo relative to a variable aspiration level, which in turn might reflect earlier successes or failures. According to this hypothesis, therefore, the search for novelty, would be most likely triggered in situation of crisis;
- deprivation of novel mental experiences, relative to the appreciation of cognitive stimuli provided by an environment. Whenever there is a lack of such stimuli, then, individuals may be motivated to create them anew. In this case, the driver for innovation would depend on the preferences for experiencing novelty. Depending on the fluctuating rate of preferences, therefore, innovation is likely to be looked for at a lower but constant basic rate.

² In many cases, undertaking such an evaluation process would also mean adopting a cultural attitude prone to search for mental novelty in order to deal with the solution of current problems, or removing constraints to changes.

³ In this respect, as already pointed out in earlier innovation studies (see Brown, 1980) the spatial and temporal dimensions are fundamental references for the innovation uptake and spreading in socio-economic systems. On the conceptual ground, the need of accounting for the generation and novel roles of innovation is a major challenging issue (Witt, 2002, Lane, 2004).

| | | | |
|-------------------------------|---|---|---|
| Deprivation of mental novelty | C) Identifying behavioural principles underlying socio-economic upgrading. Aligning activity changes with context changes | D) Stimulating agents' responsible engagement in socio-economic development. Cooperative behavioural trends within functional hierarchies | F) Difficulties to reconcile perceived local needs with external requirements and constraints (i.e. UE policies, international trade) |
|-------------------------------|---|---|---|

These stages can be briefly summarized as follows:

- a) substitution, according to which ICT applications would mainly aim at replacements in order to improve efficiency;
- b) co-evolution, according to which as a result of ICT increasing spreading, production of virtual action spaces would occur together with physical counterparts in socio-economic and spatial organizations;
- c) recombination, according to which ICT introduces a further relational dimension which adds up to the already existing social, economic, spatial and institutional ones. Actually, the technological dimension does not simply plug in an additional layer, but turns out to be generative of new ways of dealing with those relational dimensions (see Mckenzie, 2003, Dodge and Kitchin, 2004).

3. A few implications on ICT policy initiatives

Aspects of the above questioning are also detectable in the recent debate about innovation in a knowledge-based society. Building on a critical appraisal of the conventional view that innovation proceeds along a linear developmental path views which underpinned past policies, a 'third view' is called for as a necessary determinant for building the European knowledge-based society (Louis Lengrand & Associés European, 2002).

If one admits that there exists a correspondence between the views on innovation which have been proposed in the literature and the perspectives above mentioned about the relationships between ICT and regional development, then the framework shown in Tab. 2 can be outlined.

Table 2. Views of innovation and perspectives on the relationships between ICTs and regional systems

| | <i>Prevailing view of innovation</i> | <i>Perspectives on the ICT-urban system relationships</i> |
|----------------|--|--|
| Earlier view | Innovation proceeds along a chain of development stages, i.e. from laboratory science to commercial applications and their diffusion in the economic system | <i>Substitution:</i> technologies can replace the many place-based relationships and dynamics of human organizations |
| 2nd generation | There exist many feed-back loops between the different stages of the innovation chain Recognition of the existence of innovation systems: national and regional systems as recipients and activators of innovations | <i>Co-evolution:</i> electronic spaces and organisation / spatial spaces are produced together Co-evolution results from the on-going restructuring of the political-economy system |
| 3rd generation | Changing nature of innovation. | <i>Recombination:</i> socio-technical hybrids are |

| | | |
|--|---|--|
| | <p>Understanding the rational and regulator processes underlying the generation of innovation</p> <p>Articulation and fusion of knowledge.</p> <p>Search of new ways of collaborating by different agents</p> | <p>produced</p> <p>Social and spatial life become continuously recombined</p> <p>Fully relational view of technology</p> |
|--|---|--|

Of course, the suggested framework is likely to be very simplistic. The need to take a ‘third view to innovation’ does not necessarily mean that we would be capable or willing to adopt a recombination perspective. In addition, the working of the substitution and co-evolution perspectives might apply as well.

On a conceptual ground, therefore, two challenging questions for a regional system having to deal with the introduction of ICT would be: a) how to gain as much advantages as possible from the deployment of the different perspectives and b) how to proceed most effectively along the different stages in order to favour the different regional sub areas. Both these questions, in fact, highlight the intrinsic uncertainties resulting from the spreading of ICT in regional systems and more generally into society.

That intrinsic indeterminacy of ICT outcome also affected the design of the initiatives which were carried out in Piedmont in the nineties. It had two major influences:

- on the one hand, it progressively called for more articulated development programs in order to better anchor ICT positive effects onto local areas (see Occelli, 2003, Pollone and Occelli, 2005). In addition, it was realized that, in order to avoid the worsening of existing regional disparities, the provision of ICT infrastructures, i.e. the building of an ubiquitous broadband telecommunication network, should be accompanied by initiatives aimed to improve users’ capability and literacy. The crucial argument here is that while gaps in the provision of ICT will be eventually filled, as the market of ICT products expands and/or targeted policy initiatives are put in place, additional gaps are likely to be opened, just because of users’ lack of access to ICT, due to budget constraint, cultural barriers, lack of competence and unwillingness to engage into innovative activities. In certain contexts, these gaps may combine with pre-existing social and spatial disparities thus leading to the formation of new forms of socioeconomic and spatial inequalities which have been called digital divides (see Warschauer, 2002);
- on the other one, it contributed to raise an awareness that faltering policy-making was caused not only by the difficulties to identify value-laden future development scenarios worth being socially shared. Rather, an additional major source of problems was recognized in the impossibility to find in the existing set of rules and procedures the institutional and operational scaffolding structures allowing for the policy actions to be successfully accomplished.

4. The innovation kernel as a lever of ICT policies

The above mentioned influences also permeate the latest ICT policy program, called Wireless-PIEdmont (WI-PIE), which the Regional Government has launched at the end of 2004 to boost socioeconomic development in Piedmont. This is a foremost initiative in terms of the amount of financial resources, duration (it is a three year program) management efforts and overall involvement of societal bodies, i.e. firm associations, local governments and social organizations.

Compared with the earlier ones, the program is much broader in scope and its design has several institutional, social and organizational aspects the description of which goes beyond the scope of this paper.

It consists of seven so-called strategic projects, mentioned in Fig. 2, for each of which a specific goal and a set of accomplishments have been specified. The overall goal of the policy program is the provision of a broadband connectivity infrastructure covering all areas of Piedmont. Related objectives are:

- tackling the risks of ICT lagging areas and digital divides likely to be produced by market-led ICT grow process occurring in Piedmont as a result of the deregulation occurring in the telecommunication sector;
- adopting a pro-active attitude in policy making to assist regional socio-economic transitions through and by means of ICT applications.

Associated with the wide ranging aims of the program are several encompassing benefits, i.e. increasing regional competitiveness, improving administrative efficiency, improving technological transfer from research centres to Small and Medium Enterprises, purporting to respond to a number of concerns in regional development raised by the current trends of societal, political and institutional changes.

What it is worth underlining here is the very conception of the overall program. Much as a consequence of the acknowledged indeterminacy of ICT policies as well as of an increased familiarity of the governmental bodies with Business Intelligence based management, a knowledge basis has been put at the core of the regional program. Although, increasingly, data-gathering and assessment activities are undertaken in policy programs as relevant but still additional procedures, in this case, they are explicitly acknowledged as internal components of the whole program. This represents an innovative trait for a policy program which distinguishes it from earlier regional ICT initiatives.

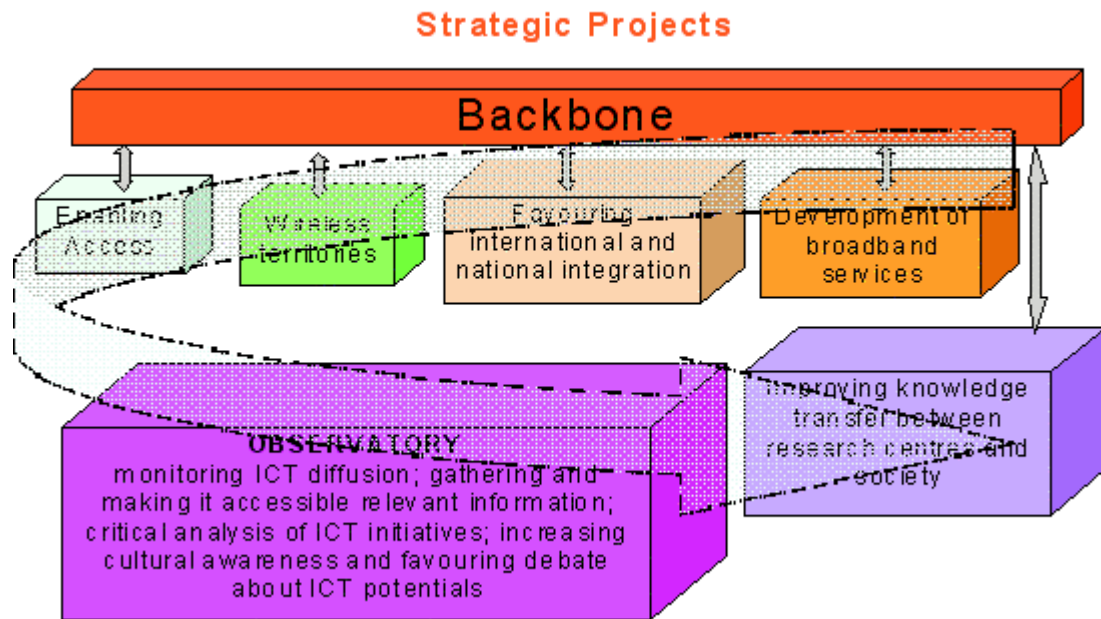


Figure 2 The strategic projects of the Piedmont ICT Policy Program (WI PIE)

Some noticeable features are:

- the fact that it has been instantiated into a Project, see the box indicated as Observatory in Fig.2, with explicitly stated objectives to fulfil, and, namely, the satisfaction of information need of the other projects and the formulation of steering information for the advancement of the overall program;
- the fact that it is understood as a permanent research activity on the relationships between ICT and regional development. As a kind of repository keeping traces of the outcome of overall policy implementation.

To get some clues about the kind of activities entailed by the Observatory, one can view them as related to the deployment of the innovation kernel shown in Fig. 1. In this respect, they may be considered as drivers of a learning process, allowing the perspectives on the relationships between ICT and regional development, referred to above, to evolve. This is graphically summarized in Fig. 3.

Initially, substitution would make it possible to establish only one component of the innovation kernel, i.e. the hard component of ICT. As its effects spread out and become increasingly perceived, co-evolution can subsequently take place. At this stage, the information component would come be fully active, and foster a range of transformations which can finally result in new system functionalities.

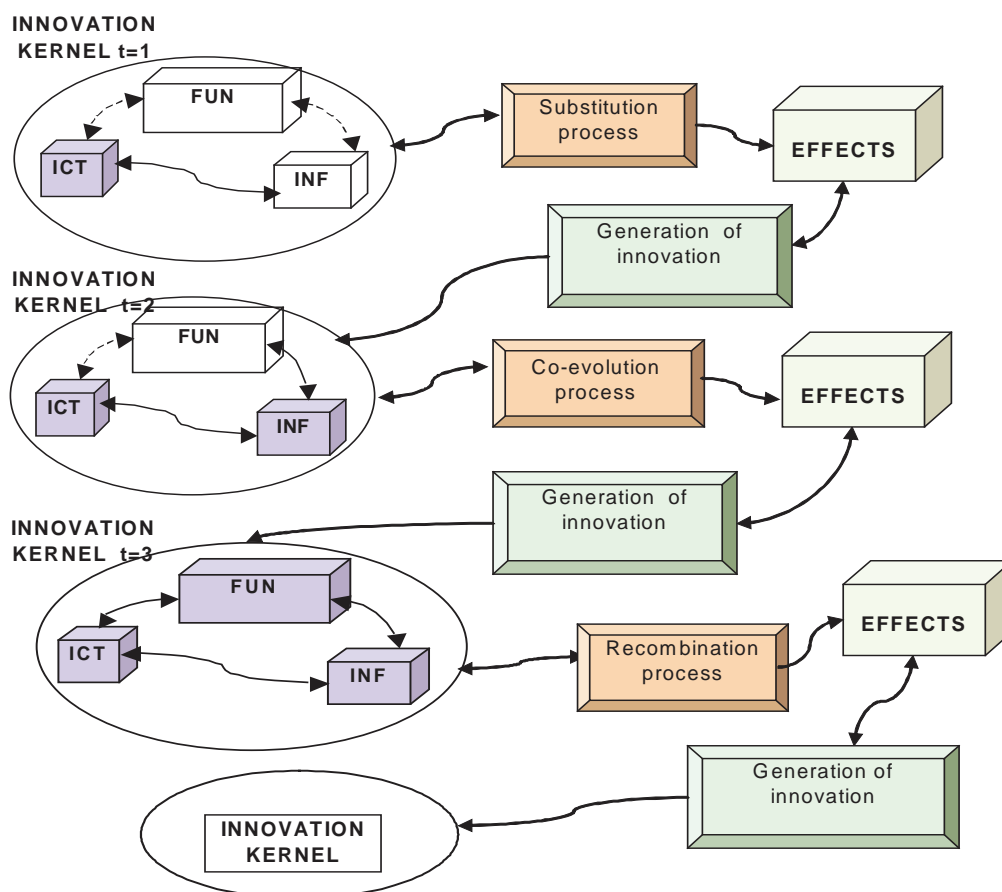


Figure 3 The innovation kernel in action: learning to innovate

At all stages, processes of generation and evaluation of innovations are entailed and differentially put in place. Knowledge, i.e. transferable information and aptitudes to mental novelty, is required to both activate and support the working of the innovation kernel, while, in the process, other types of knowledge are produced and stocked away. These can be exploited to feed other innovation kernels already existing at the current stage, or to activate new ones in subsequent stages of evolution.

5. Concluding remarks: challenging features of the application of the innovation kernel

Undertaking the activities assigned to the Observatory Project, i.e. applying the innovation kernel sketched above, does not simply imply the unfolding of learning stages but entails the viability of a certain setting. This is made up of three interlinked elements, graphically mapped in Fig. 4:

- a) system context. This consists of a network of relationships between people having various roles, backgrounds and competences and belonging to different institutions, i.e. researchers in several domains, project managers,

stakeholders and decision-makers at different levels of the regional government. This is embedded in a rather complex web of systemic links which underlies the whole ICT regional program. In the observatory project this network primarily connects those actors directly operating in it (these are listed on the left side of Fig. 4 and include the institute of socioeconomic research of Piedmont, IRES, research centres on technologies, BOELLA and CSP, the regional computer centre CSI, the Turin Polytechnic⁴) and the actors responsible of the other projects (these are shown on the right side of Fig. 4) ;

- b) the drives of the observation activities, that are responsible of the pattern of researches, their implementation and responsive advancements. To cope with the acknowledged indeterminacy of ICT impacts and the need to sustain a pro-active attitude to policy making these drives are based on two knowledge processes, indicated in Fig. 4 as the exploration-exploitation and coding-decoding cycles. These are two crucial processes well known in the complexity debate. The former reflects the intrinsic tension between, on the one hand, the need to anticipate the unknown, and, on the other one, the opportunity to take advantage of what already known. An appropriate modulation of this tension is essential for the efficacy of the Observatory project in the regional program. The latter echoes the stages of a modelling process, such as conceptualized for example by Rosen (1985). Here, the cycle hints more to the role that a modelling activity as a cognitive mediator can play in linking a problem situation and its interpretation, within the overall human activity system (see Occelli, 2002). Modelling in this respect, is an extraordinary cognitive artefact through which selectively unveil clues and marks from a situation.
- c) the inscription of the results of the research activities into knowledge artefacts in order to consolidate them into the system structure. These consist of several products, i.e. reports, data-bases, seminars, that are instantiated on two levels of the scaffolding structure associated with the observatory:
- an internal level, labelled in Fig.4 as back office (implicit and private knowledge), being operated and accessed only by the participants to the ICT overall program. While constituting a working environment for knowledge sharing and transfer, it will represent a first experience of ICT application for policy implementation;
 - an external level, indicated in Fig.4 as front office (explicit and public knowledge), open to the general public through a web site.

⁴ The network is also open to other regional centres undertaking research on ICT, such as the regional Internet Access centre TOP-IX and other university departments.

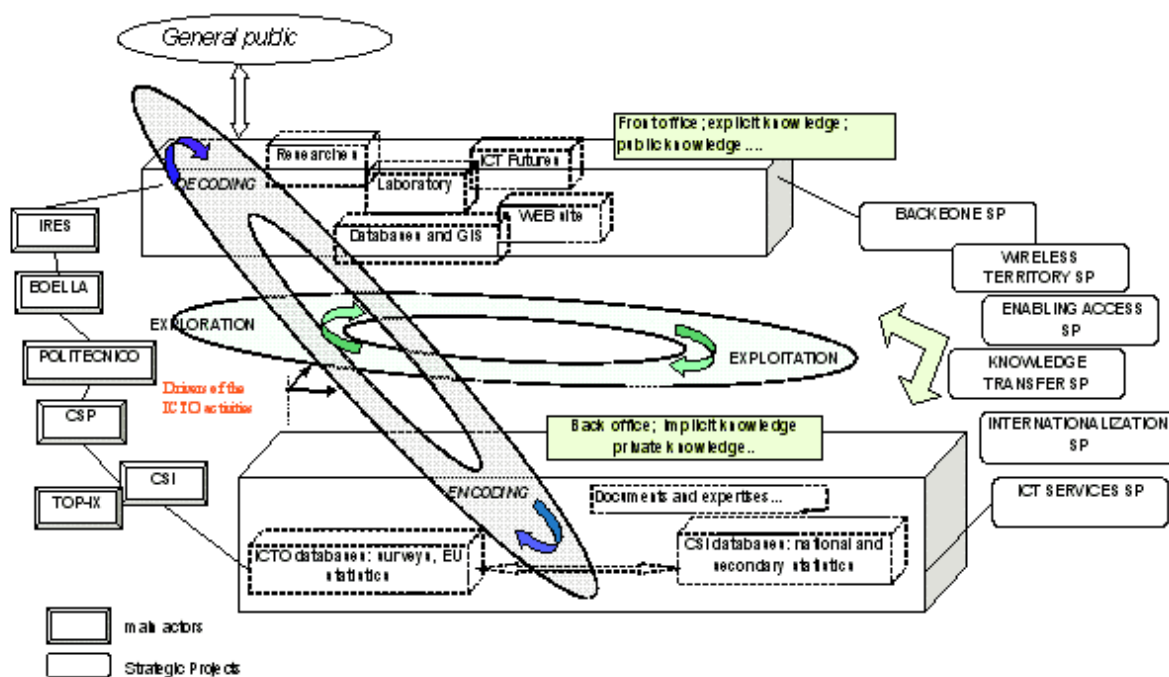


Figure 4 Designing the scaffolding of the observation activities

To a closer inspection, the above elements and their operations reveal a number of features belonging to what has been called by a system thinker *an appreciative setting* to mean a human based process intended to provide an epistemology for making sense of social processes (we refer to the work by Vickers as discussed in Checkland, 2005 and Winder, 2005). Relevant traits of such a setting are: knowledge openness, self-creation and responsibility in giving meaning to human processes.

The extent to which the scaffolding structure sketched in Fig. 4, is likely to result in an appreciative setting represents a challenging issue in the future development of the observatory activities on ICT.

References

- Brown L.A (1981) *Innovation Diffusion: A New Perspective*, Methuen, London.
- Checkland P. (2005) Webs of Significance. The work of Geoffrey Vickers, *Systems Research and Behavioral Science*, 22, 4, 285-290.
- Dodge M., Kitchen R.(2004) Code, space and everyday life, WP 81, CASA, London.
- Graham S. (2000) Conceptualizing Electronic Space, in Wilson M.I., Corey K.E. eds., *Information Tectonics*, Wiley, New York, 7- 28.
- Lane D. (2004) Innovation and industrial districts: Some considerations towards a policy for local development, Paper presented at the XXV AISRE Conference, Novara, October 6-8.

- Louis Lengrand & Associés European (2002) *Innovation tomorrow. Innovation policy and the regulatory framework: Making innovation an integral part of the broader structural agenda*, PREST (University of Manchester) ANRT – France, *Innovation papers No 28* Directorate-General for Enterprise, EUR 17052, European Commission Luxemburg.
- Haglund H., Nevalainen R. (2004) *Sustainable Information Society Powered by Diversity*, Baltic Palette II, Action Group 3 Information Society.
- Mackenzie A. (2003) *Transduction: Invention, Innovation and Cultural Life*, Institute for Cultural Research, Lancaster University, Lancaster.
- Occelli S. (2001) *New Information Technologies and the City: Re-discovering people and places*, *Netcom*, 15, 3-4, 207-230.
- Occelli, S. (2002) *Facing urban complexity: towards cognitive modelling. Part 1: modelling as a cognitive mediator*, Paper presented at the XII European Colloquium on Theoretical and Quantitative Geography, St.Valery-en-Caux, 7-11 September, 2001. www.cybergeog.fr
- Occelli S. (2003) *Digital communities as collective agents of urban systems. Some evidence for the Turin Metropolitan Area*, Paper presented at Digital Communities, 2003, Stockholm, June 4-6.
- Occelli S., Staricco L. (2001) *Nuove tecnologie di informazione e di comunicazione e la città. Elementi di riflessione*, Angeli, Milano.
- Pollone M., Occelli S. (2005), "Leveraging ICT for Regional development: the Case of Piedmont", Paper presented at The International Conference Digital Communities 2005, June 6-9, Benevento.
- Rosen R. (1985) *Anticipatory System*, Pergamon, Oxford.
- Williams G. (2005) *Geoffrey Vickers: Philosopher of Responsibility*, *Systems Research and Behavioral Science*, 22, 4, 291-298.
- Wilson M.I., Corey K.E. eds. (2000) *Information Tectonics*, Wiley, New York.
- Warschauer M. (2002) *Reconceptualizing the Digital Divide*, *First Monday*, *Firstmonday*, 7, 7, http://firstmonday.org/issues/issue7_7/warschauer/index.html
- Witt U. (2004) *On Novelty and Heterogeneity*, *Papers on Economics and Evolution*, 5, Max Planck Institute for Research into Economic Systems, Evolutionary Economics Group, Jena.

Tracer laisse-t-il des traces ?

Comment la représentation spatiale de territoires de projets contribue à la conception collaborative de projets de territoire ?

Sylvie Lardon,

UMR Métafort, INRA & ENGREF POP'TER, 24 avenue des Landais, 63171 Aubière Cedex 9 (lardon@engref.fr)

RÉSUMÉ : Dans les démarches de diagnostic de territoire, la production de représentations spatiales est le support des interactions entre acteurs pour le projet de territoire. Nous avons conçu un jeu d'expression pour articuler différentes méthodologies et élaborer une vision partagée du territoire, par la production collective de représentations spatiales. L'analyse du processus de conception collaborative d'objets géographiques peut s'appuyer sur un dispositif de plate-forme méthodologique, pour garder trace des interactions et vérifier le rôle du tracé dans l'émergence de sens.

MOTS-CLÉS : représentation spatiale, conception collaborative, jeu, territoire.

1. Introduction

L'usage des représentations spatiales est de plus en plus courant dans les projets de développement de territoires, tant urbains que ruraux. Cartes, images, systèmes d'information géographique... sont mobilisés à différentes étapes d'un processus de construction collective. D'autres outils font l'objet de développements plus récents : cartes dynamiques, dessins d'acteurs, simulations informatiques. Cependant, dans les démarches participatives de développement territorial, se pose la question de l'appropriation par les acteurs des méthodes proposées, pour mettre en œuvre un raisonnement spatial et pour construire collectivement des projets de territoire (Debarbieux & Vanier, 2002).

Nous faisons l'hypothèse que les représentations spatiales sont des objets intermédiaires (Vinck, 1999) qui facilitent les interactions entre acteurs et la coordination des actions de développement. Il s'agit de contribuer à un processus de construction collective. Plus qu'une simple activité de visualisation, le tracé des différentes représentations du territoire permet leur ancrage dans l'espace, la confrontation des points de vue des acteurs et la production d'une vision partagée du territoire. Il importe donc que les acteurs participent eux même à cette activité de tracé.

Dans le cadre de formations à l'aménagement du territoire, nous avons conçu une démarche de diagnostic de territoire basée sur l'usage des outils de représentation spatiale (2). Dans le but de faire participer les acteurs au diagnostic et à la prospective du territoire, pour construire un projet, nous avons élaboré et testé un dispositif de jeu de territoire (3). Les premières analyses à partir de regards pluridisciplinaires nous amènent à proposer une plate-forme de conception collaborative d'objets géographiques (4). L'objectif est de permettre aux utilisateurs de tracer leurs représentations du territoire et de garder trace de leurs raisonnements et ce faisant, de vérifier si elles laissent des traces dans leurs actions (5).

2. Une démarche de diagnostic de territoire

Dans une perspective de développement territorial, vu comme « l'augmentation de la capacité des acteurs à maîtriser les évolutions qui les concernent », l'usage des représentations spatiales constitue une modalité privilégiée de compréhension des structures et des dynamiques spatiales (Lardon *et al.*, 2001) et de conception de projets de territoire (Debarbieux & Lardon, 2003). C'est un outil pour le diagnostic et l'organisation d'une démarche réflexive avec des acteurs voulant mener individuellement et collectivement un raisonnement spatial. Elle constitue un instrument de visualisation et de mise en discussion des points de vue d'acteurs et de leurs pratiques territoriales, contribuant à la transformation des territoires.

La démarche de diagnostic de territoire est fondée sur l'analyse des principes organisateurs de l'espace (Piveteau & Lardon, 2002). Elle combine des informations issues de documents cartographiques et statistiques existants avec des informations apportées par des enquêtes auprès des acteurs concernés. Elle extrait les principales structures et dynamiques du territoire ainsi que les projets des acteurs pour construire des scénarios d'évolution à mettre en débat avec les acteurs.

La modélisation graphique sert de fil directeur pour confronter et intégrer les connaissances produites tout au long de la démarche. « *Les chorèmes sont des structures élémentaires de l'espace qui se représentent par des modèles graphiques... Toute configuration spatiale relève de la combinaison éventuellement très complexe de mécanismes simples* » (Brunet, 1986). La grille simplifiée (Figure 1) élaborée pour répondre à des questions d'aménagement du territoire reprend les sept principes organisateurs de l'espace comme autant de questions à se poser et de modèles élémentaires auxquels se référer pour construire et interpréter des formes d'organisation spatiale observées dans les territoires.

Elle est mise en œuvre dans un itinéraire méthodologique qui est « la façon d'articuler différentes méthodes de traitement des informations mais aussi de mobilisation des acteurs pour construire progressivement une vision partagée et stratégique du territoire ». Il s'appuie sur la production de représentations spatiales. Leur filiation rend compte de la construction du raisonnement, des points-clés et des chemins alternatifs possibles (Lardon & Mainguenaud, 2005). Il est restitué sous la forme de la séquence des représentations spatiales élaborées et de la succession des étapes à passer, en explicitant les objectifs et les connaissances produites (Figure 2).

Pour aider à la conception de projets de territoire, il est nécessaire de fournir aux acteurs les outils de raisonnement sur l'espace. Il s'agit de mettre en adéquation la capacité des acteurs à coordonner leurs actions avec les caractéristiques du territoire qu'ils transforment. Comprendre l'organisation spatiale du territoire, dans ses modalités d'articulation interne, mais aussi en externe avec les territoires voisins, préciser les modèles de développement souhaités pour le territoire et valoriser ses spécificités pour se positionner dans les dynamiques existantes, constituent les points forts du raisonnement à construire (Lardon & Piveteau, 2005). Il passe par l'identification de zones-charnières, comme marqueur, lieu de passage, ressource, front ou confins, qui sont autant d'enjeux pour le développement des territoires.

Pour que les acteurs participent aux projets de territoire, il faut non seulement qu'ils s'approprient les démarches, mais aussi que s'exprime la volonté de collaborer. La participation ne se décrète pas, elle se construit en amont du processus de communication, dès le processus de conception (Roche, 1998). Si la diffusion des informations, la visualisation des scénarios, la prise de décision sont utiles, il importe également de produire des connaissances, construire des scénarios, élaborer des stratégies avec les acteurs. Pour cela, il est nécessaire de proposer des dispositifs où les acteurs puissent exprimer leurs territorialités plurielles et où les différentes spatialités sont intégrées dans les modèles (Lardon, 2003). Le jeu de territoire est ainsi proposé comme support à la participation des acteurs.

3. Un dispositif de jeu

Dans la perspective d'impliquer les acteurs dans le diagnostic et la prospective du

territoire, nous avons conçu un jeu visant à construire une vision partagée du territoire. Il a été mis au point avec des chercheurs, lors du séminaire « Territoires qui se dessinent. Représentations graphiques et prospective territoriale », organisé par le groupe de prospective de la DATAR et L'ENGREF Clermont-Ferrand (Angeon et Lardon, 2003). Depuis, il a été adapté à différentes situations dans le cadre d'ateliers participatifs tels que ceux menés avec les acteurs du Sud-Aveyron dans le cadre du programme d'accompagnement du viaduc de Millau¹.

Le jeu se déroule en trois séquences :

- Construction collective d'une maquette du territoire, à partir du choix de cartes des structures et des dynamiques du territoire, selon un tour de jeu où chacun argumente et dessine, sur la maquette, l'objet spatial retenu. Elle correspond à une reconstruction collective du diagnostic de territoire.
- Elaboration individuelle de scénarios d'évolution pour le territoire, à partir des connaissances construites lors de la séquence précédente, donnant à voir les transformations du territoire liées à des choix d'actions à mener. C'est une phase de prospective.
- Présentation collective des différents scénarios et débat sur les enjeux auxquels ils répondent et sur les actions concrètes proposées. Elle aboutit à l'élaboration de stratégies pour le territoire.

Le jeu Millau s'est déroulé en trois ateliers participatifs en parallèle, avec une quinzaine d'acteurs du territoire. En première séquence, les maquettes produites mettent en évidence les principaux enjeux du territoire, liés à la concentration des dynamiques et à la valorisation des ressources (Figure 3). Les différentes visions du territoire portées par les acteurs (Figure 4) sont confrontées dans l'élaboration des scénarios en seconde séquence et leur argumentation en troisième séquence (Figure 5). Il y a émergence d'une représentation partagée du territoire et émergence de sens, comme orientation du projet de territoire. Le décalage apporté par le dispositif de jeu permet de dépasser les éventuels conflits de représentation entre les acteurs pour s'attacher à la construction d'un sens collectif, qui reste à consolider.

Le jeu constitue un dispositif de collaboration collective. Il favorise la dynamique de groupe, en mettant les acteurs en position de collaborer. Il offre un cadre collectif de production de règles et d'appropriation et de diffusion des informations. Il facilite l'apprentissage collectif et met en scène les raisonnements des acteurs, en particulier dans la création graphique et dans l'explicitation de modèles d'aménagement (concentration, diversification, développement endogène, ouverture sur l'extérieur). Comme *compositeur*, le jeu rend compte d'un diagnostic et le fait partager à différents acteurs. Comme *médiateur* pour l'analyse de problèmes ou de conflits, le jeu aide à trouver des solutions. Comme *coordonnateur* d'actions sur le territoire, le jeu facilite la mise en œuvre de stratégies.

Des premières expérimentations, on peut tirer trois résultats principaux :

- Les objets géographiques, en tant que structures, portent également en eux-mêmes les fonctions qui leur sont attribuées. Dessiner un objet, c'est aussi parler de ses relations avec les autres objets. En cela, les objets sont révélateurs des activités. Inversement, un objectif assigné ne peut pas être rempli par n'importe quel objet. Ceux-ci sont porteurs de sens. L'iconographie du territoire est en elle-même révélatrice des représentations des acteurs. Ainsi, le dessin de la brebis dans la maquette de l'atelier 1 de Millau (Figure 4) exprime l'importance attachée à la dynamique de production agricole et à sa valeur symbolique en terme de développement territorial.
- La résolution des problèmes, dans des projets de développement peut s'appuyer sur les objets géographiques qui servent de référents spatiaux et de points d'ancrage. De par leur matérialité, ils circonscrivent les domaines du possible et tiennent compte de la spatialité des processus. De par leur valeur symbolique, ils focalisent l'attention et expriment la

¹ Le jeu du territoire Auvergne est disponible sous forme de Cdrom (Angeon *et al.*, 2003), celui du territoire de Millau également (Lardon *et al.*, 2005), ainsi que le site de l'ENGREF (<http://www.engref.fr/millau>)

diversité des territorialités. Ils permettent de confronter des points de vue et de trouver des accords. Ainsi, les différentes visions du territoire portées par les acteurs du territoire de Millau peuvent s'articuler dans la mise en œuvre d'actions concrètes, telles que la création de zones d'activités, en mobilisant chaque acteur selon sa propre marge de manœuvre, comme entrepreneur, gestionnaire de l'espace, élu local ou technicien.

- Le processus de production d'objets géographiques se fait dans le jeu selon une alternance de phases d'ouverture et de fermeture : fermeture par les consignes, les outils mis à disposition ou la contrainte des jeux d'acteurs ; ouverture par les marges de création laissées, la formulation des légendes ou les interactions permises entre les acteurs. Ce dispositif permet l'émergence de nouvelles idées, tout en fournissant un cadre de référence pour la comparaison des expériences. Ainsi, lors des ateliers participatifs de Millau, l'idée d'une attractivité du territoire, basée sur la valorisation de l'image du viaduc en lien avec la qualité des ressources et des produits, a pris de la consistance.

Il importe d'avancer des propositions méthodologiques pour approfondir ces premiers résultats et hypothèses sur le rôle des représentations spatiales dans la conception de projets de territoire. Elles font appel à des compétences en conception d'outils informatiques et en analyse cognitive des interactions. Les outils informatiques doivent permettre l'échange d'informations spatialisées, l'intégration de vues (Claramunt & Mainguenaud, 1995), la gestion de l'aspect temporel et la modélisation des connaissances spatiales (Le Ber *et al.*, 2002). L'analyse des interactions recherche la production de sens dans les objets et la gestuelle, c'est-à-dire dans l'intrication des dire et des faire (Brassac, 2004 ; Mondada, 2004, 2005). Il est indispensable de mettre en place des dispositifs pour capter, le plus finement possible, l'histoire de la production des traces, labiles et pérennes, produites par les acteurs dans de tels processus collaboratifs.

4. Une plate-forme méthodologique de conception collaborative

Nous avons croisé différents points de vue disciplinaires au sein de l'AS144² sur un jeu de conception collaborative d'objets géographiques. Ils ont convergé vers la proposition d'un dispositif de saisie des interactions avec un équipement de captation et un protocole d'aide à la conception (Le Ber *et al.*, 2005). Il s'agit de constituer une plateforme mobile pour automatiser les informations spatiales manipulées. Ce besoin correspond à trois objectifs principaux :

- faciliter l'expression des connaissances des divers acteurs en présence et la manipulation des informations disponibles, telles que les représentations spatiales produites ;
- garder trace des versions successives et des filiations des représentations spatiales dans un itinéraire méthodologique, pour reconstituer le raisonnement suivi ou le remettre en cause ;
- réaliser des traitements pour comparer et extraire les formes fortes des « histoires interactionnelles » entre les personnes et les représentations spatiales et préciser ainsi les modalités concrètes de production de sens.

Cette plateforme serait constituée d'un ensemble d'appareillages tels que caméras, tablettes graphiques et microphones. Elle aurait le double rôle de capter un processus de conception collaborative, en vue de son analyse et d'aider à la mise en œuvre de ce processus, grâce à son outillage technique et à son protocole.

La fonction de la plateforme mobile est précisément d'enregistrer et ainsi de pérenniser la mobilisation des représentations spatiales. Elle se décline en trois points :

- Conforter et tester les hypothèses avancées relatives au rôle des inscriptions dans la conception collaborative : les objets géographiques sont révélateurs d'activités et porteurs de sens, ce sont des points d'ancrage pour la résolution de problèmes et des objets intermédiaires pour la représentation, la traduction et la médiation.

² L'action spécifique du CNRS « Construction collaborative d'objets géo-graphiques » s'est déroulée durant l'année 2004. Elle avait pour objectif l'étude interdisciplinaire des outils et des méthodes de représentation de l'espace, utilisés ou utilisables dans les démarches de conception collaborative (<http://www-engees.u-strasbg.fr/~fleber/AS144>).

- Mettre au point une méthodologie de description des séances de conception collaborative : l'objectif n'est pas ici de capter de façon complète une session de travail, en contraignant la position des participants, mais de saisir quelques événements clés, qui fournissent des traces, au double sens de mémoire externe et de piste explicative, de la conception collaborative.
- Acquérir de façon *routinière* des corpus d'interactions et d'inscriptions ayant trait à la conception collaborative de projets territoriaux. Ces corpus ont un intérêt pour les psychologues, comme matériau expérimental de leur activité de recherche, mais aussi pour les informaticiens, comme mémoire des séances d'acquisition de connaissances et donc comme moyen d'affiner la modélisation de ces connaissances, et pour les agronomes et géographes, comme cristallisation de l'élaboration de leurs propres modèles des territoires.

La plate-forme, utilisée dans un cadre interdisciplinaire a la particularité de permettre de croiser et d'articuler les méthodologies classiquement utilisées dans les champs disciplinaires concernés : ethnologie des processus cognitifs pour les psychologues, recherche dans l'action et la formation pour les agronomes et géographes, acquisition des connaissances et usage des outils pour les informaticiens.

Elle fournit ainsi une aide, technique et méthodologique, à l'activité de conception collaborative dans le cadre de la gestion de territoire. Cette aide inclut un aspect réflexif de retour sur cette activité, pour évaluer le processus de production de sens et analyser les interactions entre les représentations spatiales produites et les représentations des acteurs.

Finalement ce type d'outillage s'inscrit dans une triple approche : technique, par l'expérimentation de l'usage d'outils numériques dans les différentes situations visées ; méthodologique, grâce à la définition d'un protocole d'aide à la conception *via* ces outils numériques – protocole de construction et d'analyse des données ; et enfin archivistique, par la constitution d'une base de corpus qui pourront faire l'objet d'une étude systématique en vue d'une capitalisation, d'une généralisation et d'une diffusion des observations.

5. Conclusion

La construction de projets de territoire fait appel à des capacités de conception collaborative et de création de sens par les acteurs eux-mêmes. Leur participation au tracé des représentations spatiales du territoire modifie leurs propres représentations et génère des interactions qui laissent des traces dans les actions qui en découlent.

Dans un itinéraire méthodologique de diagnostic de territoire, les représentations spatiales sont des objets intermédiaires qui facilitent les interactions entre acteurs. L'analyse de leur filiation reconstitue le raisonnement spatial mené, leur expression sous forme de graphes et de chorèmes permet la traduction entre disciplines. Insérées dans des outils informatiques sous forme de jeu, elles servent de media à l'échange et à la production de connaissances.

Le jeu constitue un dispositif de collaboration collective. Comme *composeur*, il rend compte d'un diagnostic partagé entre différents acteurs. Comme *médiateur* pour l'analyse de problèmes ou de conflits, il aide à trouver des solutions. Comme *coordonnateur* d'actions sur le territoire, il facilite la mise en œuvre de stratégies.

La plate-forme permet de tester le rôle des représentations spatiales dans l'action collective. Elle supporte trois activités essentielles : tracer des représentations spatiales pour les partager, garder trace des représentations successives pour vérifier leur contribution au processus et identifier comment la transformation des représentations des acteurs laisse des traces dans l'élaboration des actions.

Une telle plate-forme dédiée à la conception collaborative d'objets géographiques, s'appuyant sur un dispositif de jeu articulant différentes représentations spatiales pourrait être utilisée plus généralement par des acteurs du développement territorial. Elle s'inscrit dans le contexte social de la participation des acteurs aux projets de territoire³ et ouvre sur de nouvelles perspectives d'outils et de dispositifs pour la gouvernance des territoires.

³ Voir articles Joliveau « Les usages participatifs de la géomatique dans la gestion et la planification

6. Références bibliographiques

- V. Angeon , S. Lardon. Dessiner et comprendre le territoire : quand le jeu devient un processus collectif d'apprentissage et de création. In : (*Debarbieux et Lardon, 2003*), pp. 245-257, 2003.
- V. Angeon, S. Lardon, C. Mahé, G. Warrot. "Dessinez l'Auvergne : Comprendre le territoire". Cd-Rom ENGREF Clermont-Ferrand, 2003.
- C. Brassac. Action située et distribuée et analyse du discours : quelques interrogations. *Cahiers de Linguistique Française*, n°26, pp. 251-268, 2004.
- R. Brunet. La carte-modèle et les chorèmes. *Mappemonde*, vol. 86, n°4, pp. 3-6, 1986.
- C. Claramunt, M. Mainguenaud. Dynamic and Flexible Vision of a Spatial Database, Database and EXpert Application - DEXA (Workshop), Londres, Royaume-Uni, 4-9 Sept., 1995.
- B. Debarbieux, M. Vanier (coord.). *Ces territorialités qui se dessinent* . Éditions de l'Aube, DATAR, 267p, 2002.
- B. Debarbieux, S. Lardon (coord.). *Les figures du projet territorial*. Éditions de l'Aube, DATAR, 270p, 2003.
- S. Lardon. Diagnostic de territoire et représentations spatiales : les chorèmes, graphes et jeux. In Debarbieux B., Lardon S. (dir.), *Les figures du projet territorial*. Editions de l'Aube, Datar, Bibliothèque des territoires, pp 109-129., 2003.
- S. Lardon, P. Maurel et V. Piveteau (coord.). Représentations spatiales et développement territorial. Paris, Editions Hermès, 437p, 2001.
- S. Lardon, P. et V. Piveteau. Méthodologie de diagnostic pour le projet de territoire : une approche par les modèles spatiaux. *Revue Géocarrefour*, 2005 (à paraître).
- S. Lardon, V. Angeon, P. Moquay, Y. Poss, P. Aumasson, F. Brau, A. Caron. " Les territoires de Millau : élaborer vous-même le diagnostic". Cd-Rom ENGREF Clermont-Ferrand, 2005.
- F. Le Ber, C. Brassac, et J.-L. Metzger. Analyse de l'interaction experts - informaticiens pour la modélisation de connaissances spatiales. In : *Journées francophones d'ingénierie des connaissances, IC'2002, Rouen*, pp. 29-37, 2002.
- F. Le Ber, S. Lardon, C. Brassac, M. Mainguenaud, J. M. Préau, Construction collaborative d'objets géo-graphiques. Retour sur une recherche collective. Colloque SAGEO, 2005.
- L. Mondada. Temporalité, séquentialité et multimodalité au fondement de l'organisation de l'interaction : le pointage comme pratique de prise de tour. *Cahiers de Linguistique Française*, n°26, pp. 269-292, 2004.
- L. Mondada,. L'analyse de corpus en linguistique interactionnelle : de l'étude de cas singulier à l'étude de collections. In : *Sémantique et corpus*, coord. par A. Condamine, pp. 76-108. Hermès, 2005.
- V. Piveteau, S. Lardon. Chorèmes et diagnostics de territoire : une expérience de formation. *Mappemonde* 68 (2002/4), pp 1-6, 2002
- S. Roche. L'appropriation sociale des technologies de l'information géographique. *L'espace Géographique*, vol. 4, pp. 317-327, 1998.
- D. Vinck. Les objets intermédiaires dans les réseaux de coopération scientifique. Contribution à la prise en compte des objets dans les dynamiques sociales, *Revue Française de Sociologie*, vol. XL, n°2, pp. 385-414, 1999.

territoriales » et Lardon *et al.* « Jeu collaboratif de construction de territoire », à paraître dans la revue *Canadian Geographer*.

Figure 1. La grille simplifiée des chorèmes pour l'aménagement des territoires

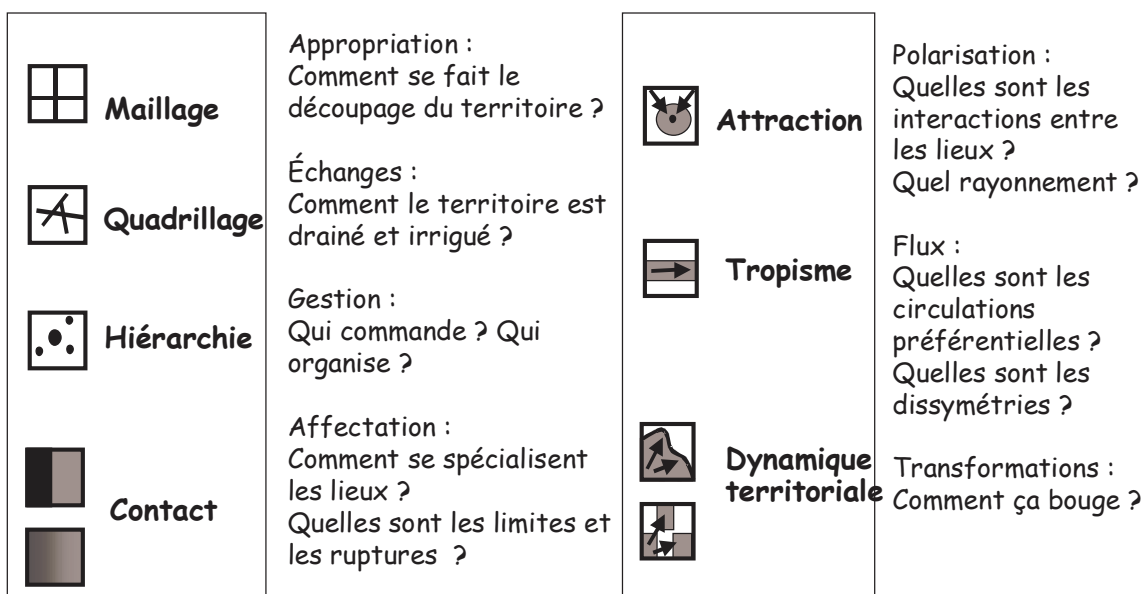


Figure 2 - Itinéraire méthodologique de la constitution des pays en Auvergne

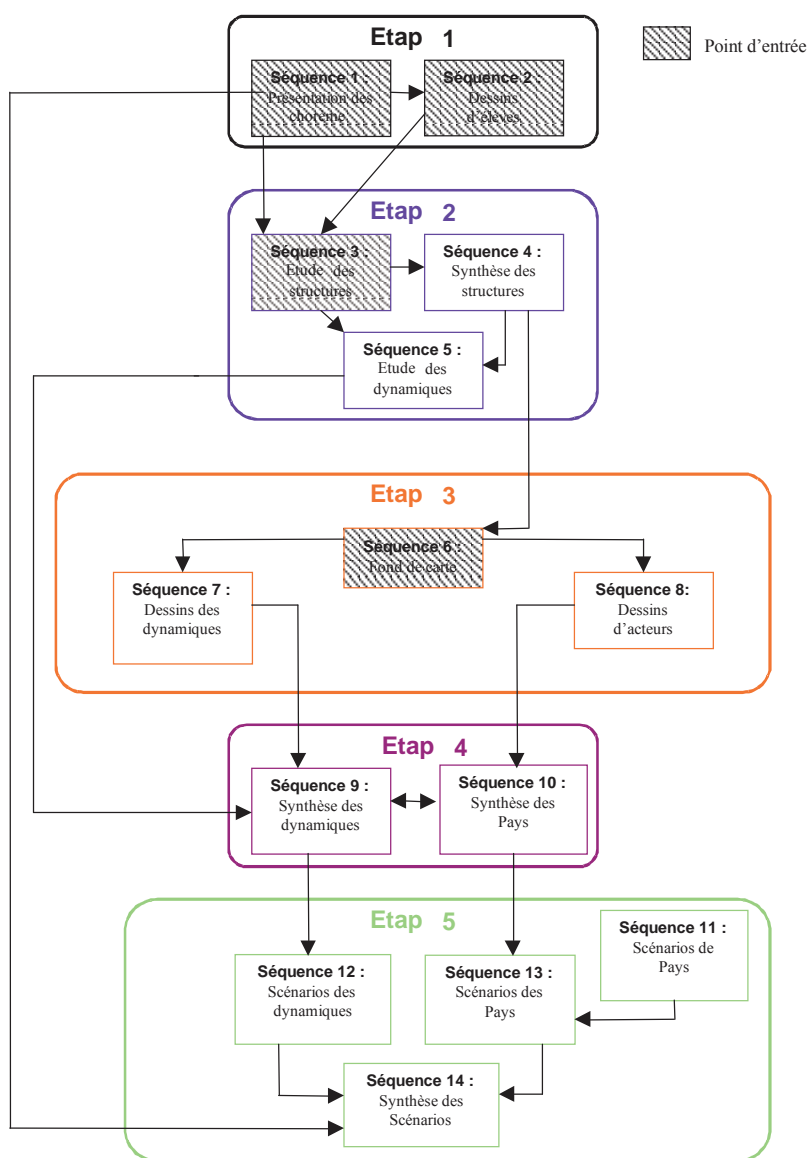


Figure 3. Une maquette du Jeu Millau

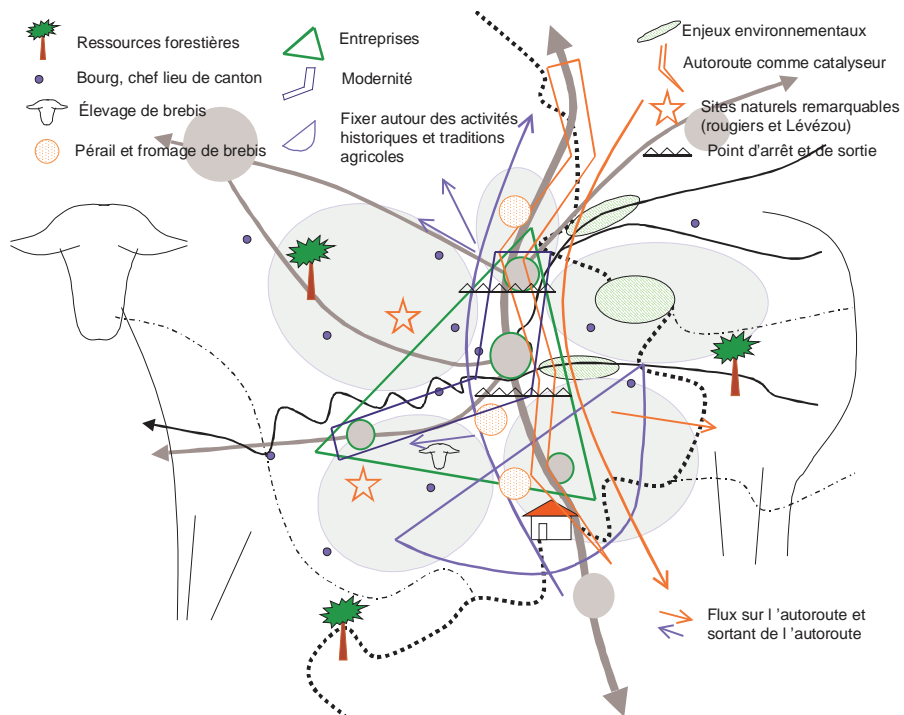


Figure 4. Différentes visions du territoire de Millau

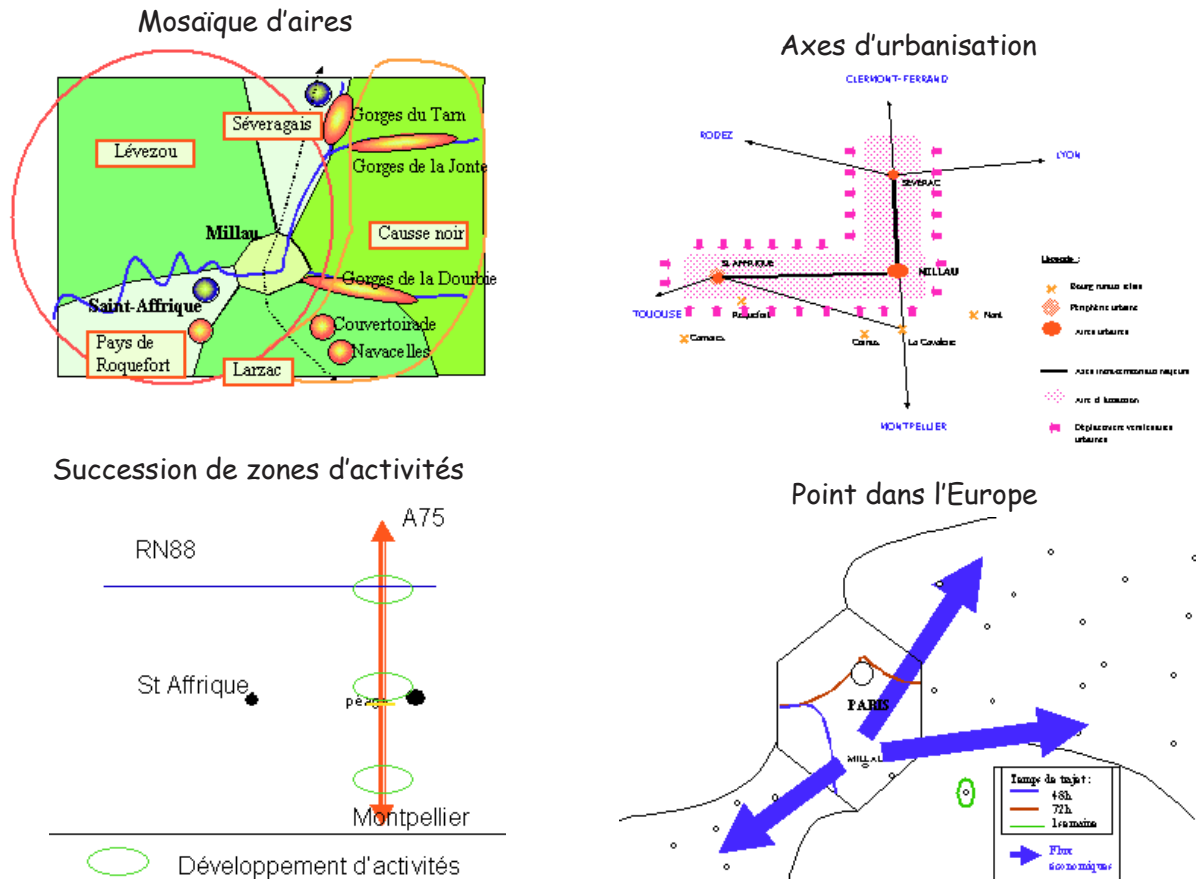
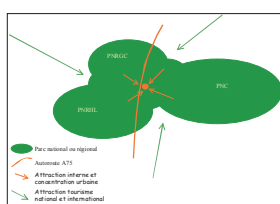
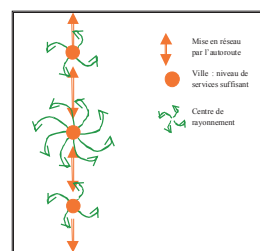


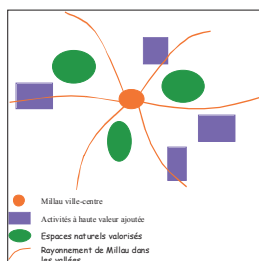
Figure 5. Scénarios d'évolution



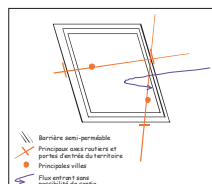
« Yellowstone »
Un grand parc naturel



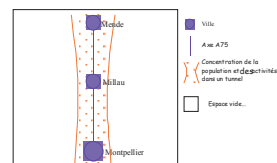
« Passage de relais »
Un réseau de territoires



« Tarn Valley »
Un bassin attractif



« Aveyron for ever »
Repli sur les limites



« Osons le désert »
Concentration sur l'axe

Les processus morphodynamiques des villes : énigmes et traces.

Dominique Badariotti

Laboratoire SET – UMR 5603 du CNRS
Université de Pau et des Pays de l'Adour
Dominique.badariotti@univ-pau.fr

Dans cet article, nous nous intéressons aux processus morpho-dynamiques des villes, appréhendés sous l'angle des énigmes et des traces. Dans ce questionnement nous nous intéresserons davantage aux centres urbains médiévaux de nos villes, qui nous semblent être les éléments morphologiques les plus intéressants dans cette optique.

La forme physique de la ville s'est constituée progressivement au cours du temps, par l'action aménagiste et urbanistique des sociétés qui nous ont précédées et des individus qui les composaient. Ces sociétés ont réglementé de façon plus ou moins contraignante l'organisation spatiale des villes, en laissant des traces de leur action dans la morphologie urbaine. Par ailleurs, les individus qui composaient ces sociétés ont également apporté leur contribution à la formation de la trame urbaine actuelle, chacun à son niveau et avec ses moyens, en tant qu'habitant, entrepreneur, promoteur, édile, ce qui a également laissé des traces.

La ville bâtie apparaît ainsi comme le résultat physique de processus morphogéniques très différents : certains sont fondés sur l'application plus ou moins ferme d'une règle, alors que d'autres semblent tirer leur origine dans des dynamiques auto-organisées de type complexe, impliquant tous les citoyens. Qu'en est-il réellement ?

Pour répondre à cette question, nous allons dans un premier temps considérer la formation des villes, pour voir dans quelles mesure leurs tissus sont imputables à des plans préétablis, puis nous verrons des positions alternatives, qui considèrent davantage les villes comme des organismes autorégulés.

1. Les traces de l'organisation initiale de la ville

La conception de la ville, qui en fait une construction raisonnée et maîtrisée par les hommes, s'appuie non pas tant sur les conditions d'apparition de la ville dans les sociétés humaines¹ que sur le contexte connu de création d'un grand nombre de villes au cours des siècles d'expansion des civilisations urbaines. En effet, si certaines villes sont apparues « naturellement », phénomènes émergents d'une société complexe naissante, d'autres – et

¹ Ces conditions restent toujours mystérieuses car si on connaît bien les justifications de l'apparition de la ville - qui sont au demeurant contradictoires, la révolution agricole du néolithique étant tour à tour présentée comme une conséquence (Jacobs, 1969) ou une cause de l'apparition des villes (Bairoch, 1985) - il ne faut pas oublier que ces justifications ne reposent que sur des hypothèses.

elles sont fort nombreuses - ont été créées *ex-nihilo par* la seule volonté de l'homme et ont donc bénéficié de tracés établis antérieurement à leur construction, qui ont parfois été dévoyés.

1.1. Les traces de l'organisation directive de la ville

Il n'est pas possible ici d'évoquer toutes les villes construites *ex-novo* et selon un plan préconçu. Retenons simplement que les phénomènes d'expansion de l'urbanisation sont apparus à plusieurs reprises dans l'histoire de l'humanité, et qu'ils s'appuyaient en général sur des phénomènes de colonisation eux-mêmes liés à des phases d'expansion économique ou culturelle de certaines civilisations. A titre d'exemple, on peut évoquer les principales phases de développement de l'urbanisation en occident, qui a connu des phases d'essor important à l'époque gréco-romaine, mais aussi à d'autres reprises notamment au cours du moyen-âge (bastides romanes, villes teutoniques ou hanséatiques, comptoirs vénitiens ou gênois, ...) et aux époques modernes et contemporaines (colonisation américaine, villes nouvelles spécialisées, industrielles, militaires, touristiques,).

Pour les périodes les plus récentes, nous disposons bien entendu de plans qui prouvent que ces villes n'ont pas été créées – ou recrées - selon un principe aléatoire mais bien selon un ordre préconçu. Dans la plupart des cas, la création repose sur un découpage de l'espace en mailles : celles-ci constituent à la fois la base du réseau viaire et les limites des îlots. L'adoption d'un plan en damier, plus ou moins régulier, semble dominante – même si elle n'est pas exclusive - de l'antiquité à nos jours. Penchons nous sur le cas des villes antiques, qui sont à l'origine de beaucoup de villes médiévales.

Ainsi, la période gréco-romaine a connu à la fois l'apparition de l'urbanisation en Europe mais aussi sa diffusion tout autour du bassin méditerranéen d'abord, puis dans l'ensemble de la zone dominée par la *pax romana* dans un second temps. Le système défensif (enceinte), le plan caractérisé par des îlots et des rues, le temple, le marché, le bâtiment administratif et le bâtiment d'habitat, ... , tous les éléments essentiels de la ville sont déjà présents dans ces premières fondations urbaines comme nous le rapporte Mumford (1964,119).

Au demeurant, ces fondations ne se sont pas faites par hasard, mais selon des principes bien raisonnés. Leurs localisations ont été soigneusement choisies en fonction de contraintes diverses – parmi lesquelles figurent des contraintes socio-géo-techniques (les contingences des sites, les possibilités défensives, les impératifs commerciaux de l'époque, les capacités d'aménagement, ...) - mais aussi selon d'antiques principes directeurs (Platon, -370, Livre II ; Aristote,-340, Chap VII) et leurs plans reposent sur des logiques rigoureuses de découpage et d'affectation de l'espace qui divisent le territoire selon des tracés rectilignes (parcelles, routes, ...).

Par exemple dans le cas des villes romaines, l'espace est découpé à grande échelle selon le principe de la *Centuratio*². Des îlots de 70 à 150 m. de côté composent l'espace urbain, qui est quadrillé de rues de largeurs hiérarchisées allant de 2,8 m. à plus de 6 m. d'emprise (Benevolo, 1994). Les *itinera* sont d'étroites venelles piétonnes, les *actus* sont des voies carrossables uniques permettant le passage d'une charrette tandis que les *viae* sont les

² Il s'agit d'un principe de lotissement de l'espace rural qui est composée de *decumani* (voies parallèles au côté le plus long) et de *cardines* (perpendiculaires et plus courts), distants de 20 *actus*, soit env. 700 m., délimitant ainsi les *centuria* (lots carrés 49 000 m² soit 490 ares).

pénétrantes de l'époque, laissant la voie à deux charettes. L'ensemble était dessiné grâce à la *groma*, sorte d'outil composé d'une croix, fixée sur une hampe et ornée de quatre fils à plombs permettant de tracer les axes majeurs. Les îlots sont eux-mêmes divisés en parcelles, de tailles régulières, destinées à accueillir les monuments et bâtiments administratifs ainsi que les immeubles d'habitation, collectifs (les *insula* sur des parcelles de 3 à 4 ares) ou individuels (les *domus*, villas méditerranéennes construites sur des parcelles de 8 à 10 ares).

Toutefois ces plans, soigneusement et rigoureusement dessinés à l'origine, ont ensuite été livrés aux constructeurs, qui n'ont pas toujours scrupuleusement respecté les volontés qui s'y exprimaient, ainsi qu'au temps où l'accumulation des ans a fini par faire oublier les intentions de départ.

1.2 . Les traces d'un développement non-maîtrisé et le dévoiement des plans par l'usage

Au Moyen Age, en l'absence de règles d'urbanisme conservatoire, la tentation était en effet forte d'envahir l'espace public de la rue par des constructions qui n'étaient même pas illicites, en l'absence de réglementation ! Un autre aspect a été la modification ou la densification de ces tissus, parfois à la faveur d'une catastrophe urbaine (incendie, tremblement de terre, ...) mais plus fréquemment par l'effet de décisions individuelles peu, pas ou mal cadrées par les pouvoirs publics.

Les conséquences de ces processus ont conduit à un apparent désordre des centres villes, encore visible aujourd'hui dans leur trame viaire. Ce désordre explique pourquoi on trouve dans la littérature certains écrits qui présentent les centres médiévaux de nos villes comme étant le résultat d'un processus d'occupation de l'espace non maîtrisé, si ce n'est non voulu (Chabot, 1958 ; Pelletier et Delfante, 1989 ; Wackermann 2000, ...). Il serait ainsi la preuve du caractère anarchique du développement des villes, qui obligea les pouvoirs publics à prendre des mesures visant à éviter la fermeture progressive de certains espaces par l'invention d'un droit de l'usage du sol urbain permettant de réguler ces aspects (Sully, 1607).

C'est cette évolution apparemment chaotique des centre-villes qui a abouti au concept de ville spontanée (*gewachsene Stadt*), souvent appliqué à la ville médiévale, mais qui est pourtant dénoncé par d'autres auteurs mettant ainsi l'accent sur l'aspect très raisonné des désordres apparents qui frappent globalement la ville médiévale. En effet, contrairement à une idée parfois répandue, le Moyen age n'a pas fait preuve du laisser-aller le plus total en matière de croissance urbaine. Des règles d'urbanisme existaient, bien qu'elles ne fussent pas toujours écrites et qu'on n'en ait pas gardé trace pour cette raison. A Strasbourg par exemple, les saillies, surplombs et encorbellements étaient limités par un étalon gravé dans la pierre du portail sud de la cathédrale, où l'on peut encore voir aujourd'hui son échelle graphique et lire « *Das ist das Mass des Ueberhaenges* »³. D'autres villes avaient sans aucun doute leurs propres systèmes de porter à connaissance des règles d'urbanismes à respecter : mais il est vrai que celles-ci ont rarement été représentées comme aujourd'hui sur un plan.

Le fait que fort peu de plans médiévaux de tracés de rue ou d'extension de villes soient parvenus jusqu'à nous ne signifie pas que ces rues aient été construites sans plan et sans règles. Les plans de nombreuses cathédrales ne nous sont pas non plus parvenus et pourtant il

³ « Ceci est la mesure de l'encorbellement ». Cette inscription n'est toutefois pas datée.

est difficile d'admettre que de tels édifices aient pu être construits en l'absence de tout plan. Ces plans ont donc sans aucun doute existé, mais sous une forme qui a empêché leur transmission à travers les siècles. On peut en effet supposer que ces plans n'étaient pas dessinés sur du papier, mais tracés à même le sable ou le sol et effacés sitôt l'ouvrage fini, selon la tradition opérative des architectes, ce qui garantissait le secret de l'oeuvre.

D'autre part, la logique d'interprétation habituelle de la tortuosité des rues médiévales, considère que les rues plus ou moins rectilignes seraient héritées de plans tracés, datant la plupart du temps de la création grecque ou romaine des villes, alors que les rues curvilignes seraient d'origine « spontanée ». C'est oublier que même les Romains traçaient parfois des rues courbes, en fonction des contraintes des lieux. C'est oublier aussi que la tradition des constructeurs faisait moins usage de la règle que du compas.

Schémas à l'appui, deux auteurs allemands, Humpert et Schenk (2001), ont pu montrer, à partir de l'étude de dizaines d'exemples de villes médiévales allemandes, alsaciennes ou italiennes, que les dessins de rues supposées « spontanées » correspondaient en fait à des tracés géométriques précis et élaborés. En général il s'agit de tracés d'arcs de cercles, dont les origines correspondent à des points particuliers de la ville, qui ont été choisis selon une logique assez différente de la logique contemporaine.

Les origines de ces arcs nous semblent tout à fait quelconques aujourd'hui : ils ne correspondent pas forcément aux points forts de la ville, à ses haut-lieux urbains (cathédrale, château, centres...). Il en va de même des axes de symétrie utilisés dans ces tracés, qui forment des lignes ne nous paraissant pas aujourd'hui particulièrement significatives. Ces tracés reposaient apparemment sur une logique géométrique de construction et non sur une logique sur-valorisant certains lieux et rendant par contre-coup les autres insignifiants, comme le fait la perspective⁴.

La logique de ces tracés nous échappent donc en partie aujourd'hui⁵. Jusqu'à présent, nous pensions que la ville médiévale était davantage sensible à l'ordre vertical, qui l'élevait dans la foi, qu'à celui horizontal (Blanquart, 1997). Nous regardions les plans de nos centre-villes médiévaux, avec leurs rues tortueuses et inégales, comme des créations d'une époque incapable d'affirmer une organisation horizontale. Aujourd'hui, nous découvrons que des traces mesurables semblent témoigner d'une réelle pensée ordonnatrice horizontale, même si nous ne comprenons pas encore cette pensée. Si nous voulons à nouveau mieux comprendre la ville médiévale, il nous faudra réintégrer ces modes de pensée et rechercher ces organisations, à l'instar de Humpert et Schenk, afin d'en débusquer les traces dans les centre-villes médiévaux.

*Pour conclure sur ce point,
il semble donc certain que la plupart des villes, ainsi que leurs extensions, ont été réellement
construites sur la base de plans préétablis. Toutefois, il est tout aussi certain que ces plans
ont été progressivement dévoyés, sous la pression de décisions individuelles mal maîtrisées
par les pouvoirs publics. La ville ancienne médiévale, qui forme le coeur de la plupart de nos*

⁴ C'est moins la mise en scène d'un lieu particulier que l'équilibre de l'ensemble qui était recherché au Moyen-âge ; les urbanistes de ce temps ont donc davantage travaillé la proportion que la perspective.

⁵ On peut y voir une conséquence de l'invention de la perspective, qui a profondément transformé notre approche esthétique et symbolique de la ville, nous faisant oublier les principes de composition du Moyen-âge.

cités actuelles (du moins en Europe) est donc un savant mélange de voulu et de vécu, d'éléments pré-conçus et d'objets surajoutés sans ordre ni contrôle au cours du temps.

2. L'énigme de l'auto-organisation urbaine

Le dévoiement des plans initiaux des villes aurait donc conduit celles-ci vers une sorte de « désordre organisé », ou les éléments de structuration ne découlent pas entièrement des pouvoirs publics mais apparaissent aussi comme une émergence du système urbain et de la société urbaine. On voit bien que cette idée conduit, plus ou moins directement, à la formulation d'une conception de la ville en tant que système complexe. La ville ne serait pas un pur produit de l'esprit et du calcul, mais serait aussi en partie, ou totalement, l'expression d'une dynamique complexe, qui agence des efforts nombreux et isolés en un tout néanmoins ordonné. Cette idée de la ville a été explorée théoriquement et concrètement.

2.1. Une conception auto-organisée des villes

D'un point de vue théorique, la vision de la ville en tant que système auto-organisé repose sur une pensée bien précise de la ville, à savoir sur une conception de la ville en tant que système complexe.

De fait, depuis les années 1960 / 1970, parallèlement à la formulation de la théorie générale des systèmes (von Bertalanffy, 1950, 1973) et suite aux travaux de précurseurs géographes comme Brian Berry (1964) ou Jay Forrester (1969), un courant de recherches s'est spécialisé en géographie urbaine dans l'analyse de la ville en tant que système. Ce courant de recherches, qui s'appuie sur une théorie systémique de la ville, a intégré les diverses évolutions de la théorie des systèmes, dont les plus récentes, aboutissant à la formulation de la notion de système complexe.

Selon Dauphiné (2003), on observe trois formes de complexités en géographie : *une complexité structurelle* émergeant d'une multitude d'agents individuels en interaction ; *une complexité de niveau*, liée à une imbrication d'échelles spatiales ou de niveaux d'organisation différents ; *une complexité dynamique* liée à l'existence d'évolutions non linéaires, entraînant l'imprédictibilité des effets du système dans le futur et ce même lorsque les facteurs sont simples et bien connus. Toutefois, ces trois types de complexités ne sont pas exclusives, et un objet géographique complexe peut répondre simultanément à l'une et à l'autre définition : c'est le cas pour la ville.

Ainsi, les géographes ont pu mettre en évidence le caractère auto-similaire de la ville à diverses échelles (Batty, Longley, 1994) tout en soulignant le fait qu'elle était soumise à des dynamiques non linéaires (Pumain, Sanders, Saint-Julien, 1989). Ils ont également montré qu'elle se caractérisait par un certain degré d'emboîtement d'échelles : de l'immeuble à l'îlot, de l'îlot au quartier, du quartier à la ville, et de la ville à la conurbation (Frankhauser, 1994). Il en est de même pour les niveaux d'organisation multiples qui pèsent sur la ville, depuis les décisions désagrégées des agents individuels jusqu'à celles, diverses, des entités collectives qui nous regroupent et nous gouvernent (Reymond, Cauvin, Kleinschmager, 1998).

Nous l'avons déjà dit, penser la ville comme un système complexe revient à la penser comme un organisme qui dispose d'un ordre auto-généré, malgré son apparence chaotique ou désordonnée. C'est notamment cette vision précise de la ville, en tant que système complexe, qui légitime les analyses fractales qui sont faites de sa morphologie.

2.2. Les conséquences morphologiques d'une conception complexe de la ville

La ville est donc considérée comme un système social et spatial complexe. Or, selon Moulin (1992, 144) « la dynamique des systèmes – complexes - est caractérisée par des transferts de matière, d'énergie et d'information à travers des interfaces fractales ».

D'un point de vue morphodynamique, certains auteurs vont même plus loin et considèrent que tout système vivant complexe adopte une forme fractale car elle lui permet d'optimiser ses relations avec son environnement, et notamment de maximiser les échanges (Frontier, Pichot-vialle, 1992). Il est vrai qu'en théorie, à masse égale, une structure fractale offre une surface d'interface avec son environnement bien supérieure à celle que peut offrir une structure euclidienne. Si nous considérons que la ville est un système vivant complexe, il semble dès lors logique qu'elle se plie à cette stratégie et se perpétue au travers d'une géométrie fractale, qui permet d'augmenter les densités humaines tout en offrant à tout-un-chacun un accès aux espaces libres, à la lumière et à l'air, pour reprendre les éléments clés de la Charte d'Athènes (Le Corbusier, 1957).

L'étude morphométrique a confirmé, à ce jour, la théorie morphogénique de la ville. Du point de vue morphologique, l'observation partagée par plusieurs auteurs et confirmée par des mesures, atteste bien que la géométrie des villes répond à une logique fractale (Frankhauser, 1994). La géométrie des villes semble donc effectivement être de nature fractale, à la fois simple dans son détail, complexe dans son ensemble et générant des surfaces bâties dont les périmètres précis s'étendent à l'infini.

Mieux, les chercheurs travaillant sur la fractalité de la ville considèrent que celle-ci ne répond pas à une seule et unique loi fractale, mais bien à un ensemble de lois fractales différentes, selon les quartiers, les tissus urbains, voire les époques. Cette observation vient conforter l'idée d'une ville complexe, organisée dans son ensemble, mais dont les parties sont formées d'entités variées, issues de processus fort différents. L'étude de la fractalité des structures urbaines permettrait de travailler au moins indirectement sur leur complexité.

Les villes sont ainsi le résultat à la fois d'un plan pré-conçu et d'un phénomène d'auto-organisation. Ce fait peut, à première lecture, sembler paradoxal, mais c'est un paradoxe apparent qui ne résiste pas à l'analyse. Si la ville s'est effectivement organisée selon certaines règles, édictées par les pouvoirs publics et représentées sur des plans, cela ne suffit pas pour rejeter totalement l'hypothèse d'auto-organisation et considérer que l'ordre urbain ne serait pas une réalité émergente de la société urbaine, en effet :

- d'une part, le fait que la ville réponde, en partie au moins, à un plan préconçu, ne l'empêche pas d'être aussi en partie auto-organisée ;
- d'autre part, on peut très bien considérer que les règles et les plans d'urbanisme eux-mêmes ne sont au fond que la matérialisation d'une émergence du système, tout comme les pouvoirs publics qui spécifient ces règles et les font appliquer peuvent eux aussi être considéré comme des éléments organisateurs émergents.

Une ville peut donc très bien être à la fois construite selon des plans et être auto-organisée. On peut avoir affaire à des organismes dont la trame a été dessinée selon un but précis à une époque, mais dont les éléments ont été apportés à d'autres époques dans un certain désordre auto-raisonné par les habitants. C'est d'ailleurs une position partagée par les auteurs historiens de la ville : les plans des centres villes médiévaux ont été hérités de l'époque antique, mais leur réoccupation par les constructions médiévales se serait faite sans respecter les tracés viaries hérités, dans une certaine anarchie (Giovannoni, 1931 ; Lavedan, 1926). Cette observation est également cohérente avec le principe dialogique des systèmes auto-organisés, qui précise que des notions antagoniques (par exemple ordre et désordre dans ce cas) sont en général associées dans l'émergence d'organisations complexes.

On en arrive à ce point à penser la ville ni comme une construction uniquement raisonnée, issue de l'esprit d'une minorité de dirigeants et d'urbanistes, ni comme une construction totalement auto-organisée, dont le dessin n'aurait aucun dessein. La ville tiendrait à la fois des deux principes : elle serait quelque chose d'intermédiaire entre le planifié et l'auto-organisé, d'où sa grande complexité.

Mais à ce niveau, deux questions se posent :

- *Si les villes sont bien de système complexes, elles sont forcément générées par une multitude d'interactions simples, combinées à l'infini : quels peuvent être ces éléments basiques qui répétés et combinés à l'infini donnent son organisation actuelle à la ville médiévale, ou à la ville contemporaine ?*
- *Si la ville est auto-organisée, sur quoi repose son auto-organisation ? N'y a-t-il pas d'autres éléments à prendre en compte dans la réalisation d'une ville que ses seuls tracés ?*

3. Les traces d'un autre ordre sous-jacent ?

Au fond, que les plans des villes anciennes aient existé ou pas, qu'ils aient été tracés sur des supports périssables ou non, qu'ils aient été peu, prou ou pas respectés par les habitants, n'est finalement pas si fondamental. La pensée préalable de l'agencement des lieux et son respect ne sont pas les seuls éléments organisateurs de l'espace urbain. En effet, bien que l'ordre des villes ait pratiquement toujours été étudié à travers leur plan, d'autres éléments génériques existent dans la réalisation des villes et de leurs tracés.

Pour construire une ville, que faut-il ? Certes, il faut une idée préconçue, que l'on matérialisera par un plan, mais il faut aussi et surtout construire des immeubles et des rues, qui eux aussi – individuellement – doivent être tracés pour être maîtrisés. Une ville n'est pas que son plan, elle est aussi la somme des immeubles et des constructions qui y ont été réalisées, de même qu'elle est l'ensemble des hommes et des sociétés qui y ont vécu. Si on se limite aux aspects morphologiques, on peut supposer que les techniques des hommes de l'art qui ont supervisé la construction des immeubles urbains – techniques qui ont été transmises d'homme à homme depuis le début de l'architecture - constituent aussi un élément générique de la ville.

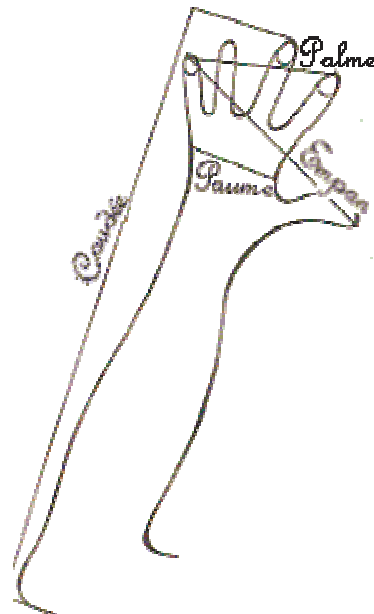
Effectivement, nous relevons que les techniques de tracé des plans et des constructions utilisés par les citadins n'étaient pas livrées au hasard, y compris lorsque l'agencement des rues ne se faisait pas en damier. Ceci est lié aux outils traditionnels utilisés par les architectes et les techniciens du bâtiment, et qui se limitaient à un morceau de bois (règle) et à une ficelle, permettant de construire des cercles, des angles droits et des rectangles. L'instrument majeur

de ces tracés était le compas ou le cordeau à nœuds, associé à un étalon de mesure, la canne des bâtisseurs ou canne des maîtres d'œuvre, qui permettait de construire des tracés harmonieux, complexes et géométriquement basés sur le nombre d'or (Φ), en termes de proportion tout au moins.

Le cordeau était déjà utilisé dans la plus haute antiquité par les architectes égyptiens ou romains pour faire leurs tracés. Avec une simple règle et une corde on peut construire un triangle rectangle, et donc construire un angle droit sans té. Avec un cordeau à 5 nœuds, on peut aussi construire le triangle rectangle familier 3-4-5, qui permet de construire le rectangle d'or dont le rapport du grand coté au petit coté correspond à Φ (Vincent, 2002).

La canne des bâtisseurs est une règle qui représente la quine, étalon des 5 unités de mesures de l'époque (paume, palme, empan, pied et coudée royale (Vincent, 2002, 35). Celles ci sont basées sur des valeurs anthropiques qui sont, fait remarquable, organisées à la fois selon une suite de Fibonacci (une mesure étant égale à la somme des deux mesures précédentes) et selon une progression géométrique de raison Φ (une mesure est égale au produit de la mesure précédente par Φ).

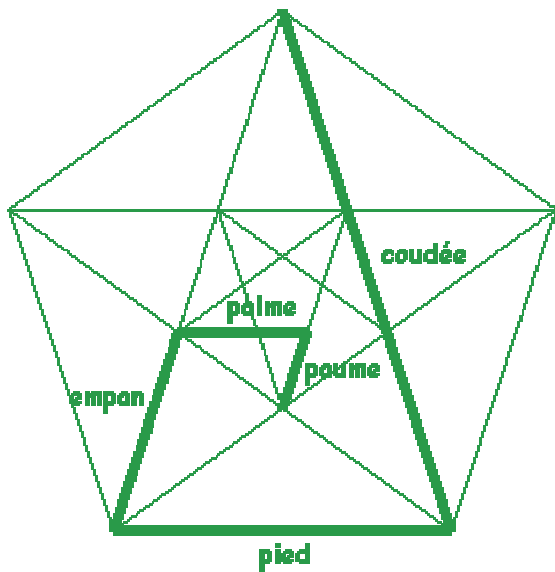
Fig 1 : la quine



La coudée royale égyptienne, dénommée quine ou « mesure de l'initié » est estimée au 19^{ème} siècle à 52,36 cm.

Avec ces deux outils, on peut construire d'autres outils nécessaires à la construction, comme un té ou une *groma*, ainsi que toute une foule de figures géométriques régulières et symétriques, des plus simples (cercles, carré, triangles, rectangles,...) aux plus complexes (polygones réguliers, polyèdres réguliers, ogives, étoiles, mandorles, spirales, ellipses, oeufs...) (Vincent, 2002) en conservant une proportion de ces objets par rapport à la quine.

Les tracés antiques, tout comme les tracés médiévaux, mis en évidence par Humpert et Schenk (2000), ont été réalisés avec ces outils et conservent donc en point commun une proportion particulière qui est celle du nombre d'or.



Cette suite est aussi géométrique puisque le rapport entre deux mesures consécutives est le nombre d'or.
 Nous le visualisons sur la suite des pentagones du schéma de gauche.

Source : http://perso.wanadoo.fr/therese.eveilleau/pages/truc_mat/textes/nombre_dor.htm

L'utilisation de ces mesures et de ces proportions ne s'est toutefois pas limitée au Moyen Age : l'intégration de ces mesures a survécu à l'adoption du système métrique, pour des raisons multiples. D'une part on peut citer le caractère traditionnel de la transmission du savoir dans les ordres et corporations de constructeurs et d'architectes (le modulator du Corbusier était lui aussi basé sur le nombre d'or : Le Corbusier, 1950 ; Vincent, 2002, 31) ; d'autre part on peut évoquer le souci esthétique partagé par la masse populaire, comme l'a prétendument montré le philosophe Fechner (Cleyet-Michaud, 2002, 98), le rectangle d'or serait en effet considéré par une majorité de personnes comme un rectangle particulièrement équilibré et esthétique.

Par ailleurs, nous observons aussi que les techniques de tracé évoquées possèdent des propriétés qui peuvent nous amener à des formes fractales ou complexes. Les formes simples que l'on manipule, déclinées à l'infini, peuvent en effet engendrer des formes complexes et fractales à partir de leur répétition. Elles permettent aussi de faire un lien entre les constructions anciennes et modernes de la ville, qui reposent finalement sur des proportions similaires, puisqu'elles sont en partie calées sur la proportion du corps humain.

On peut ici tenter de relier ces différents aspects en formulant l'hypothèse que ce qui fait le lien entre la ville ancienne et la ville moderne, que l'élément répétitif qui organise tout cela, au delà du temps et des plans différents, c'est tout simplement l'homme et sa proportion résumée ou symbolisée par le nombre d'or.


4. L'énigme du nombre d'or dans l'urbanisme des villes médiévales

Si on pose l'hypothèse que le nombre d'or est à l'origine de la dimensions de la plupart des constructions de nos centre-villes moyenâgeux, on devrait pouvoir retrouver par des mesures d'analyse spatiale des références à Φ dans les centres villes médiévaux.

Trouver une telle référence reviendrait à montrer que les principes utilisés pour construire le détail de la ville se retrouvent à un niveau plus général, dans une relation d'invariance d'échelle qui est conforme à la théorie fractale et au principe hologrammatique des systèmes complexes. Cela permettrait aussi de vérifier qu'un ordre local, répété à l'infini, permet d'obtenir un ordre global, même lorsque celui-ci n'a pas été formulé, ce qui est conforme à la notion de paysage adaptatif développée dans la théorie des systèmes complexes (Lewin, 1994, p. 63).

On peut dès lors se demander s'il existe un lien entre la dimension fractale des villes et le nombre d'or : si tel était le cas, ce lien devrait se matérialiser plus fortement dans les centres villes médiévaux - où l'on sait que les tracés et les édifices ont été réalisés selon le nombre d'or - que dans les villes modernes où cette référence n'a pas forcément toujours été appliquée. Nous proposons donc de voir si on peut identifier des mesures de fractalité se rapprochant de 1,618034 (Φ) : nous effectuerons ce test dans le centre urbain médiéval de Strasbourg (Tab 1).

Tableau 1 : Analyse fractale du centre médiéval de Strasbourg

| Centre médiéval de Strasbourg | Analyse fractale radiale |
|--|--|
|  | <p>Centre de comptage : barycentre (135, 102) / Bornes : 1-179 Progression quadratique</p> <p>Résultats</p> <p>dim : 1,616</p> <p>a : 3,3916 c : -326,72</p> <p>Rapp, de corr, : 0,996825 Rap, cor, ajus, : 0,996752</p> |

Si on pratique une analyse fractale radiale du centre médiéval de Strasbourg, que l'on seuille en fonction des limites du tissu urbain, on observe que l'on trouve une estimation de dimension $D_r = 1,616$, ce qui correspond à 2 millièmes près au nombre d'or (1,618). Certes, en changeant le bornage de l'analyse, la dimension fractale radiale changerait aussi, mais que l'on obtienne le nombre d'or en partant du barycentre et juste au moment où le bornage correspond à la limite de l'image selon un des axes, est tout de même remarquable.

Evidemment, ce petit exemple ne démontre rien. Popper nous dirait : vous avez trouvé ici un pigeon noir, ce qui prouve simplement qu'il en existe, mais ne prouve pas qu'ils soient tous noirs ! Il faudrait pour valider ce test, le multiplier en effectuant d'autres mesures sur d'autres centres médiévaux - de préférence ayant peu changé depuis le Moyen Age - ou sur des plans anciens, par exemple les plans-reliefs.

Mais ce test nous donne envie de continuer à explorer cette voie et de poursuivre les recherches basées sur l'hypothèse que les techniques de tracés et de constructions des

immeubles marquent la forme globale de la ville, vérifiant ainsi l'idée que la ville serait un système complexe dont la forme serait un paysage adaptatif ou un attracteur morphogénique (Roger Lewin, 1994). Ce marquage, ou cette trace que nous venons de déceler, correspond bien à une pensée complexe des villes, dans la mesure où un élément de détail, répété à l'infini fini par faire émerger un ordre global. Cet ordre global reste de surcroît dans des rapports de proportion similaires, qu'il s'agisse de la ville entière ou d'une simple bâtisse, ce comportement rappelant la propriété d'invariance à travers les échelles des fractales.

*Notre recherche des traces
de l'organisation ou de l'auto-organisation des villes médiévales
nous a ainsi amené à constater que le nombre d'or, qui a été à la clef de maintes et maintes
constructions, anciennes comme modernes, se retrouve également dans l'urbanisme des
centres villes anciens, du moins dans celui de Strasbourg.
Ce constat nous encourage à poursuivre la recherche de traces
dans la ville nous permettant de mieux lever le voile
sur son principe d'organisation.*

Conclusion

Notre contribution se situe donc sur un terrain qui oppose au fond deux théories morphodynamiques de la ville, à savoir une conception de la ville en tant que système auto-organisé émergent *versus* une conception de la ville, en tant que construction raisonnée. En les caricaturant, on peut considérer que l'une est en faveur d'un urbanisme adaptatif, qui suit tant bien que mal le mouvement de la société, alors que l'autre soutient davantage un urbanisme directif, qui prétend maîtriser l'ensemble des constructions urbaines.

Ces deux théories sont actuellement toujours en débat, même si on voit bien ici que les villes n'ont pas totalement été livrées au hasard, lors de leur mise en place. D'une part des tracés, parfois complexes, préexistaient souvent à la construction des nouveaux quartiers, et d'autre part, l'homogénéité des techniques de construction donnait un ordre implicite aux espaces construits, un ordre basé sur une proportion elle-même commandée par le nombre d'or des architectes.

La question posée ici n'est d'ailleurs pas qu'une opposition formelle entre adeptes du rôle organisateur des urbanistes et partisans du laisser-aller : elle a une dimension scientifique et philosophique. En effet, la ville est une forme d'organisation spatiale et humaine qui est apparue - par hasard ou par nécessité - à un moment de l'histoire de l'humanité, voici environ 10 000 ans, et qui se maintient et se développe depuis selon une organisation fortement similaire. Tout le débat réside dans la part de hasard et de nécessité qui a présidé à la création de la ville et à son évolution. La ville est-elle hautement improbable, est-elle apparue par hasard et se maintient-elle uniquement parce qu'elle a une utilité (hypothèse darwinienne) ? Ou au contraire aurait-on pu prédire son apparition, compte tenu des transformations technologiques, sociales, spatiales... fortement probables, qui ont affecté les groupes humains et leurs habitats (hypothèse complexe) ?

Notre recherche visait à relever ce vaste débat et à le faire avancer en proposant des réflexions, des méthodes et des outils susceptibles de produire des informations, des mesures ou des observations permettant de le recadrer. Un des outils utilisés consiste à effectuer des mesures de fractalité dans les divers quartiers de la ville. Dans les secteurs réguliers et similaires nous cherchons des mesures de fractalité similaires, et dans les zones connues pour

s'être historiquement auto-organisées, nous cherchons un autre principe d'ordre, un ordre caché, régulé peut-être par le nombre d'or, avec le résultat que l'on sait !

- Aristote, -340 (?), *La politique. L'homme animal civique*. Paris, Gonthier, PUF, 1980, 288 p.
- Bairoch Paul, 1985, *De Jéricho à Mexico, villes et économie dans l'histoire*. Paris, Gallimard, 707 p.
- Batty Michael, Longley Paul, 1994, *A geometry of form and function*. London, Academic press, 394 p.
- Benevolo Leonardo, 1994, *Histoire de la ville*. Paris, Parenthèses, 1^{ère} éd. 1975, 509 p.
- Bertalanffy Ludwig von, 1950, An outline of general system theory. *British journal of philosophie of science*, n°1, 139-164
- Bertalanffy Ludwig von, 1973, *Théorie générale des systèmes*. Paris, Dunod, 296 p.
- Berry Brian Joe Lobley, 1964, Cities as systems within systems of cities. *Papers of the regional science association*, n°13, p. 147-163
- Blanquart Paul, 1998, *Une histoire de la ville. Pour repenser la société*. Paris, La Découverte, 193 p.
- Chabot Georges, 1958, *Les villes Aperçu de géographie humaine*. Paris, Armand Colin, 1^{ère} éd. 1948, 224 p.
- Chalmers Alan, 1987, *Qu'est ce que la science ?* Paris, La découverte, 287 p.
- Cleyet-Michaud Marius, 1973, *Le nombre d'or*. Paris, PUF, Q-S-J1530, 126 p.
- Dauphiné André, 2003, *Les théories de la complexité chez les géographes*. Paris, Anthropos, 248 p.
- Frankhauser Pierre, 1994, *La fractalité des structures urbaines*. Paris : Anthropos - Economica, 291 p.
- Forrester Jay Wright, 1969, *Urban dynamics*. Cambridge Mass., Wright Allen press, 285 p.
- Frontier Serge, Pichot-Viale Denise, 1992, *Ecologie et systémique*, in Le Gallou, Bouchon-Meunier, Paris, Technique et documentation, p. 226-240
- Giovannoni Gustavo, 1998, *L'urbanisme face aux villes anciennes*. Paris, seuil, 1^{ère} ed 1931, 349 p.
- Jacobs Jane, 1969, *The economy of cities*. New York, Random house, 268 p.
- Le Corbusier, 1951, *Le Modulor. Essai sur une mesure harmonique à l'échelle humaine, applicable universellement à l'architecture et à la mécanique*. Paris, Ed. de l'Architecture d'Aujourd'hui, 240 p.
- Le Corbusier, 1957, *La Charte d'Athènes*. Paris, Editions de minuit, 1^{ère} éd. 1941, 190 p.
- Le Gallou Francis, Bouchon-Meunier Bernadette, 1992, *Système. Théorie et applications*. Paris, Technique et documentation, 341 p.
- Lewin Roger, 1994, *La complexité. Une théorie de la vie au bord du chaos*. Paris, Interéditions, 223 p.
- Moulin Thiébaud, 1992, *Outils mathématiques. Fractals, relateurs mathématiques*. in Le Gallou, Bouchon-Meunier, p. 140-151
- Mumford Lewis, 1964, *La cité à travers l'histoire*, Paris, Seuil, 781 p.
- Pelletier Jean, Delfante Charles, 1989, *Villes et urbanisme dans le monde*. Paris, Masson, 200 p.
- Platon, -370 (?), *La république*. Paris, Garnier Flammarion, 1966, 507 p.
- Pumain Denise, Sanders Léna, Saint-Julien Thérèse, 1989, *Villes et auto-organisation*. Paris, Economica, 191 p.
- Reymond Henri, Cauvin Colette, Kleinschmager Richard, 1998, *L'espace géographique des villes. Pour une synergie multistrates*. Paris, Economica, 1998, 557 p.
- Vincent Robert, 2001, *Géométrie du nombre d'or*. Marseille, Chalagam, 127 p.

Comité de programme :

- Frédérick Amblard (Informatique)
- Dominique Badariotti (Géographie & Urbanisme)
- Pierre Beust (Informatique & Linguistique)
- Daniele Bourcier (Droit & Sciences Politiques)
- Paul Bourgine (Ecole Polytechnique, Paris)
- Roger Cozien (Criminalistique & art)
- Pierre de Coninck (Design)
- Guillaume Deffuant (Sciences Cognitives)
- Jean-Louis Desalles (Sciences Cognitives)
- Francis Jauréguiberry (Sociologie)
- Serge Mauger (Linguistique)
- Jean-Pierre Muller (Informatique & Multi-agent)
- Jocelyne Nanard (IHM & Ergonomie)
- Anne Nicolle (Informatique)
- Denis Phan (Économie)
- Violaine Prince - *Présidente du C.P.* (Informatique)
- Sylviane Schwer (Langages formels)
- Olivier Soubeyran (Aménagement du territoire & Histoire)

Comité d'organisation :

- Serge Mauger - *Président du C.O.* (GREYC CNRS, Université de Caen, Caen)
- Pierre Beust - *co-organisateur* (GREYC CNRS, Université de Caen, Caen)
- Roger Cozien - *co-organisateur* (Gendarmerie Nationale)

avec la contribution de :

- Dominique Badariotti (Laboratoire SET, UPPA, Pau) - *organisateur 2004*
- Danièle Bourcier (Cersa, CNRS, Paris)
- Jerzy Karczmarczuk (GREYC CNRS, Université de Caen, Caen) - *Responsable d'édition des Actes*
- Serge Stinckwich (GREYC CNRS, Université de Caen, Caen)- *Organisateur 2005*

