



Du collectif au social

Actes des journées de Rochebrune

28 janvier au 3 février 1996

Organisées avec le soutien
d'ECAL
et de l'ENST

Sous le patronage
de l'Association pour la Recherche Cognitive
et du Comité Technique Vie Artificielle - AFCET

ENST 96 S 002

Ecole Nationale Supérieure des Télécommunications

À la mémoire de Joëlle Biondi



Journées de Rochebrune 96

Haute-Savoie

« Du collectif au social »

Texte des contributions

assemblés par Jean-Louis Dessalles

organisation :

John Stewart
Jean-louis
Dessalles
Thierry Fuhs

comité de programme :

Hugues Bersini, Eric Bonabeau, Paul Bourguine,
Geoffroy De Schutter, Guillaume Deffuant, Wolf Eberwein,
Charles Lenay, Anne Nicolle, Pierre-Yves Raccah,
Guy Theraulaz, Bernard Victorri

Avec le soutien d'ECAL et de l'ENST

**Sous le patronage de l'Association pour la Recherche Cognitive
et du Comité Technique Vie Artificielle - AFCET**

ENST 96-S-002

Ecole Nationale Supérieure des Télécommunications
46 rue Barrault - 75634 paris Cedex 13



Journées de Rochebrune 96

Haute-Savoie

« Du collectif au social »

Ces cinquièmes journées de Rochebrune ont réuni, du 28 janvier au 3 février 1996, des chercheurs de disciplines très diverses (biologie, linguistique, sociologie, Intelligence Artificielle, Sciences physiques, etc.) qui, ensemble, ont abordé la question de savoir quelle(s) définition(s) théorique(s) on peut proposer pour la distinction entre « collectif » et « social ». Une capacité individuelle de se représenter la société, et d'agir délibérément afin de modifier l'organisation sociale, est-elle nécessaire et/ou suffisante ? Une organisation collective suffisamment complexe pour qu'on la qualifie de « sociale » peut-elle être composée d'individus très simples ? Si ce n'est pas le cas, la complexité de l'individu reflète-t-elle celle de l'organisation collective ? Quelles contraintes fonctionnelles doit-on postuler entre l'individu et le collectif pour expliquer la viabilité d'une organisation sociale ? Au delà de ces contraintes, peut-il y avoir dans certains cas contradiction ou opposition entre les principes de fonctionnement individuel et social ?

Ces questions concernent toutes les formes collectives, telles qu'elles sont modélisées dans les sciences physiques, biologiques, humaines ou dans les sciences de l'artificiel. Les contributions assemblées dans ces Actes abordent ces questions sous l'angle de l'organisation collective en tant que telle, ou à propos de certains de ses aspects, sans restriction de domaines disciplinaires.

Sommaire

Michèle Courant, Patricio Lerena <i>Bio-machines</i>	9
Anne Nicolle, Ludovic Delépine <i>Les systèmes de signes comme fondateurs des sociétés. Le cas des systèmes multi-agents</i>	21
Michèle Courant, Antony Robert <i>Du collectif au social: vers une distinction fonctionnelle quantitative</i>	31
Juan Gonzalez <i>Vers une théorie écologique de l'amour</i>	39
Jean-Louis Dessalles <i>Pensée privée et communication sociale</i>	49
Tiziana Zalla <i>Conscience de soi et conscience d'autrui dans les théories représentationnelles de l'esprit</i>	61
F. de Gaulejac, A. Gallo <i>Les comportements interindividuels des animaux sont-ils des comportements sociaux ?</i>	71
Christian Brassac, John Stewart <i>Le sens dans les processus interlocutoires, un observé ou un co-construit ?</i>	85
Anne-Marie Toniolo, Thierry Fierville, Didier Desor <i>Du collectif au social : illustrations expérimentales de l'animal à l'homme</i>	95
Laurent Toubiana, Jean-François Vibert <i>De la contagion à l'épidémie: modèle neuromimétique de propagation de maladies transmissibles</i>	105
Hervé Glotin, Rafael Laboissiere <i>La Vie Artificielle d'une Société de Robots Parlants: Emergence et Changements du Code Phonétique</i>	113

<i>Evelyne Andreewsky, Danièle Bourcier, Christophe Parisse, Herminia Peraita</i> <i>Une société de mots : le texte</i>	127
<i>James De Almeida</i> <i>Une analyse de conversation au service de la conception d'un système de production automatisé</i>	135
<i>Nils Ferrand</i> <i>De l'apport potentiel de la sociologie pour l'ingénierie des systèmes sociaux artificiels</i>	143
<i>Guillaume Deffuant</i> <i>Mécanismes empathiques et complexité sociale: autour de J.P. Dupuy</i>	163
<i>Marie-Noelle Sarget</i> <i>Complexité individuelle et complexité sociale : antagonismes et complémentarités</i>	179
<i>John Stewart</i> <i>L'intelligence collective des êtres humains est-elle négative?</i>	185
index des auteurs	191

Bio-Machines

Michèle Courant, Patricio Lerena
Institut d'Informatique de l'Université de Fribourg, 3 Chemin du Musée,
CH-1700 Fribourg, Suisse.
E-mail: michele.courant@unifr.ch, patricio.lerena@unifr.ch

Résumé

Nous proposons d'étudier la différence entre systèmes collectifs et systèmes sociaux selon la méthodologie de la vie artificielle, en nous plaçant au niveau de l'organisation biologique et de sa dynamique. Pour ce faire, nous définissons une machine formelle, que nous appelons Bio-machine, conçue comme une extension des automates cellulaires et des machines de Turing. Cette machine est destinée à synthétiser divers types d'entités, allant des unicellulaires au multicellulaires simples, et à reproduire divers phénomènes biologiques, dont l'évolution et l'adaptation individuelle d'organismes durant leur existence.

Ceci nous munit d'un cadre formel et expérimental susceptible de contribuer à une certaine caractérisation des systèmes collectifs et sociaux en termes de propriétés structurelles et fonctionnelles.

mots-clés: vie artificielle, automates cellulaires, évolution, adaptation, génétique, cellules, neurones, colonies, multicellulaires, machines de Turing.

1. Introduction

Dans cet article, nous proposons d'aborder la transition du collectif au social au niveau des organisations biologiques et de leur dynamique, en focalisant sur le niveau des interactions entre les entités.

L'approche adoptée est celle de la vie artificielle [Levy 1992, Langton 1988]. Elle est basée sur un nouveau modèle de machines, appelé Bio-machines, qui sur le plan formel dérive conjointement des machines de Turing et des automates cellulaires [Turing 1936, Ulam 1950] et qui intègre un certain nombre d'idées issues de la génétique et des réseaux de neurones. La force essentielle du modèle est de supporter la définition d'entités, d'une part de premier ordre (unicellulaires vivant en liberté ou organisés en colonies) mais aussi de deuxième ordre (multicellulaires), et de permettre à la fois la phylogenèse et l'adaptation des individus durant leur existence. La dynamique du système est contrainte par des lois physiques simples, en particulier des lois énergétiques, dont il semble qu'elles jouent un rôle fondamental dans l'émergence des formes et des processus biologiques [Ben-Jacob 1995].

Globalement, la vocation du modèle est de constituer un outil simple et théoriquement bien fondé, qui soit à même de déboucher sur des réalisations matérielles faciles à construire et efficaces [Hillis 1989]. Pour cette raison, il s'enracine largement dans un principe de localité des interactions. L'optique choisie est de minimiser les lois explicites pour laisser la place à un maximum d'émergence. Le modèle se démarque ainsi de systèmes comme ECHO [Holland 1992, Jones 1993] dont par exemple, les contraintes d'interactions sont plus explicitées et dont la façon de modéliser l'espace et ses contraintes matérielles est très différente. Un autre modèle assez comparable au nôtre est PolyWorld [Yaeger 1994]. Outre par sa simplicité, notre travail en diffère surtout par le niveau modélisé, mais également par le fait que le mode d'apprentissage est explicité dans PolyWorld tandis qu'il est laissé ouvert en faveur de l'émergence dans les bio-machines. Un parallèle peut également être établi avec l'environnement Tierra [Ray 1991, Ray 1995], la différence fondamentale avec celui-ci se situant au niveau de la topologie spatiale.

Concernant les résultats attendus en relation avec le thème des journées, notre espoir est que les réflexions sous-jacentes à la conception du modèle et consécutives aux expérimentations menées puissent contribuer à une certaine caractérisation des systèmes collectifs et sociaux en termes de propriétés structurelles et fonctionnelles.

L'article est dédié à la présentation des bio-machines. Il commence par quelques brefs rappels sur les machines de Turing et les automates cellulaires. Un modèle formel intermédiaire, appelé bi-machine, opérant un couplage entre ces deux modèles classiques est ensuite défini. Partant de cette bi-machine, les bio-machines de premier ordre et de deuxième ordre sont finalement décrites.

2. Fondements

Pour des raisons de complétude de la présentation, ce paragraphe est consacré aux fondements des bio-machines. Dans la mesure où les bi-machines, dont elles dérivent, résultent d'une fusion entre les machines de Turing et les automates cellulaires, nous rappelons successivement les traits essentiels de ces deux modèles de base. Nous en profitons pour les rapprocher et munir les machines de Turing d'une nouvelle terminologie.

2.1. Machines de Turing

Le modèle des machines de Turing fut conçu par Alan Turing en 1936. Il constitue le modèle de base du concept d'algorithme.

Une machine de Turing est composée d'une *bande infinie* divisée en *cellules* et d'une *tête*. Chaque cellule de la bande peut contenir un symbole issu d'un ensemble fixé contenant notamment un symbole particulier appelé *symbole blanc*. La tête est un automate d'états finis qui peut à chaque transition se déplacer le long de la bande, lire et écrire des symboles sur elle, et changer son état interne selon son état courant et celui de l'entrée constituée par le symbole de la bande placée sous elle. Formellement, elle se définit comme un quintuplet (Q, q_0, S, I, f) où:

- Q est un ensemble fini formé des états de la tête,
- q_0 est un état distingué de Q appelé état initial,
- S est un ensemble fini de symboles appelé alphabet de la bande, contenant le symbole blanc noté B ,
- I est un sous-ensemble de $S - \{B\}$ appelé alphabet d'entrée,
- f est la fonction de transition définie comme une fonction partielle
 $f: Q \times S \dashrightarrow Q \times S \times \{Gauche, Droite, Immobile\}$.

La position initiale de la tête détermine la cellule de la bande qui reçoit le premier symbole d'entrée. Une fois les symboles d'entrée placés sur la bande, la tête se trouve dans son état initial q_0 et la machine commence à fonctionner. A chaque pas, la fonction de transition prend en compte le symbole situé sous la tête de lecture ainsi que l'état interne de la tête et détermine le nouvel état interne de la tête, le symbole qui est écrit sur la bande et le mouvement de la tête: ce mouvement peut être un déplacement vers la gauche (noté *Gauche*), un déplacement vers la droite (noté *Droite*) ou une absence de mouvement (notée *Immobile*). La machine s'arrête quand elle ne peut plus fonctionner, c'est-à-dire lorsque sa tête se trouve dans un certain état q tandis que le symbole d'entrée est s , et que la fonction de transition f n'est pas définie pour (q, s) . Il est à noter que l'utilisation des machines de Turing pour la reconnaissance de langages amène en outre à considérer certains états comme des états finals: on dit que la machine a reconnu une entrée lorsque cette entrée a conduit la tête dans un état final et s'arrête.

Par rapport au modèle standard de machine de Turing, diverses versions modifiées ont également été conçues. Nous en citons et décrivons brièvement quelques unes (pour plus de détails, cf [Sudkamp 1988, Wood 1987, Hopcroft 1979, Hopcroft 1984]), ce sont les machines de Turing:

- unidirectionnelles (la bande est limitée d'un côté);
- non déterministes (le domaine de la fonction de transition est l'ensemble de toutes les parties de $Q \times S \times \{Gauche, Droite, Immobile\}$);
- multi-pistes (plusieurs bandes parallèles, mais une seule tête les parcourt simultanément);
- multi-bandes (plusieurs bandes et une tête par bande; les têtes agissent de façon synchrone);
- multi-dimensionnelles (la bande est un tableau de cellules de dimension d ; les mouvements sont possibles vers chacune des $2.d$ cellules voisines canoniques);
- multi-têtes (plusieurs têtes indépendantes l'une de l'autre opèrent de façon synchrone);
- sensibles au contexte (la fonction de transition prend comme extra-arguments des symboles

du voisinage de la cellule courante).

Toutes ces versions sont équivalentes au modèle standard. Elles reconnaissent toutes par exemple la même classe de langages, à savoir la classe des langages récursivement énumérables. Malgré cette équivalence en puissance, elles diffèrent cependant par leur expressivité: certains problèmes sont beaucoup mieux décrits dans l'une ou l'autre des versions.

On peut concevoir une machine de Turing standard comme un système dans lequel une entité (la tête) existe dans un espace discret de dimension 1 (la bande), en relation avec un environnement (l'entrée). L'entité possède et encapsule certaines règles dictant la manière dont elle interagit avec son environnement. Cette entité très primitive est munie d'un seul senseur, peut se déplacer dans son environnement et le modifier. Nous appelons cette entité une cellule mobile, et nous parlons désormais des senseurs, des états et des actions de cette cellule mobile. Le symbole blanc peut être vu comme un espace vide. La métaphore est encore plus convaincante si l'on considère des versions modifiées du modèle standard comme une machine multi-têtes, sensible au contexte et multi-dimensionnelle, disons à deux ou trois dimensions. Dans ce cas, les entités possèdent plusieurs senseurs et habitent un environnement plus réaliste peuplés d'autres entités.

2.2. Automates cellulaires

Les automates cellulaires sont un modèle proposé en 1950 par Stanislaw Ulam [Ulam 1950] et John Von Neumann [Von Neumann 1950]. Von Neumann les utilisa pour démontrer que l'auto-reproduction d'automates était possible.

Un automate cellulaire est un système dynamique discret. Une grille uniforme infinie de dimension d (d étant un naturel positif) joue le rôle de son espace. Chaque élément de la grille est appelé cellule. Chaque cellule peut contenir un symbole quelconque (ou de manière équivalente, se trouver dans un état quelconque) élément d'un ensemble fini donné. Ce symbole est mis-à-jour selon une règle locale (appelée fonction de transition), qui est la même pour toutes les cellules. Tandis que le temps discret avance, toutes les cellules de la grille sont mises-à-jour de façon synchrone, exactement une fois par unité de temps.

Formellement, un automate cellulaire est défini comme un quadruplet (d, S, N, f) où:

- d est la dimension de la grille infinie;
- S est un ensemble fini de symboles cellulaires ou états de cellules, aussi appelé alphabet de l'automate cellulaire;
- N est un voisinage fini;
- f est une fonction de transition $f: S^N \rightarrow S$.

A chaque unité de temps et pour toute cellule, la fonction de transition détermine le nouvel état de la cellule d'après l'état de son voisinage cellulaire, y compris éventuellement la cellule elle-même.

Il convient de remarquer que, dans la mesure où le voisinage est fini, les cellules n'interagissent que *localement*. De même l'existence d'une fonction de transition unique forme une sorte de loi universelle gouvernant les comportements des cellules: cette propriété est appelée *uniformité* [Margolus 1987].

Une propriété intéressante des automates cellulaires est qu'ils supportent le calcul universel: il est possible de construire des automates cellulaires capables d'émuler des machines de Turing universelles. Un exemple très simple de tel automate est le *Jeu de la vie* de John Conway [Gardner 1970].

3. Bi-Machines

Le but de ce paragraphe est de fusionner les machines de Turing et les automates cellulaires afin d'obtenir une machine plus expressive que nous appelons bi-machine. Cette machine constitue la base de la définition des bio-machines.

L'idée conduisant à la définition des bi-machines est d'enrichir les automates cellulaires de quelques primitives leur permettant de supporter le mouvement [Stephenson 1992, Eckart 1995, Sipper 1994, Goel 1988]. Pour des raisons de clarté de l'exposé, nous jugeons néanmoins préférable de partir des machines de Turing pour les modifier de façon à ce qu'elles acceptent des

entrées dynamiques. Il est clair cependant, que les Bi-machines restent équivalentes en puissance aux machines de Turing.

3.1. Des machines de Turing aux bi-machines

Avant d'aborder la construction du modèle proprement dit, munissons-nous encore de quelques compléments terminologiques: comme dans les machines de Turing, nous considérons la bande comme un *environnement*, composé de *cellules d'environnements*, tandis que la tête est appelée *cellule mobile*. Les cellules mobiles se déplacent *au-dessus* (et non pas *dans*) l'environnement: cellules mobiles et environnement ne sont pas dans le même plan. Quand nous voulons nous référer à l'unité d'espace qui peut être occupée par une cellule, dans l'un des deux plans ou les deux, nous parlons de *cellules spatiales*.

Nous apportons maintenant un certain nombre de modifications au modèle standard de machines de Turing. Parmi ces modifications, certaines sont classiques ou se rapprochent de modifications classiques:

- *Orientation et mouvement*: L'intérêt de cette modification, seulement formelle, est de renforcer la métaphore d'une entité se déplaçant dans son environnement. Des actions standard de la tête comme "*aller-à-droite*", "*aller-à-gauche*" ne sont pas appropriées au cadre dans lequel nous voulons utiliser notre machine, où ces actions ne sont pas primitives. Nous fragmentons donc ces actions en des actions comme "*tourner-à-gauche*", "*tourner-à-droite*", "*se-déplacer*", i.e. avancer, ou éventuellement "*commencer-déplacement*", "*cesser-déplacement*";
- *d dimensions*: l'un des paramètres de la bi-machine détermine sa dimension d , où d est un entier positif. Quand $d > 1$, ceci rend plus intéressantes certaines contraintes dont nous voulons munir la machine, comme l'existence de symboles *solides* ou *impénétrables*;
- *Plusieurs senseurs*: Une cellule mobile peut avoir plusieurs senseurs (comme dans les machines de Turing sensibles au contexte). Ces senseurs sont solidaires d'elle-même: ils tournent quand la cellule tourne, etc. La position des senseurs d'une cellule définit son voisinage (comme dans les automates cellulaires). Dans ce papier, nous considérons qu'un voisinage peut inclure la position de la cellule elle-même;
- *Action d'écriture*: Durant une unité de temps, une cellule mobile peut écrire un symbole sur une cellule d'environnement de son voisinage. Ainsi que nous le verrons ultérieurement, l'écriture d'un symbole donné est une action; la machine est paramétrée par les coordonnées relatives de la cellule où l'on écrit. En général, la cellule d'environnement préférée est celle qui est située sous la cellule mobile elle-même.
En cas de conflit entre une action d'écriture et la fonction de transition de l'environnement, l'action d'écriture domine, autrement dit l'action d'écriture peut écraser un symbole écrit par l'environnement, excepté dans le cas où celui-ci est solide;
- *Contraintes matérielles*: Les symboles sont divisés en deux catégories, les pénétrables et les impénétrables (ou solides). La sémantique de ces derniers est telle que si une cellule tente de se mouvoir ou d'écrire sur lui, son action échoue. Les cellules mobiles elles-mêmes sont considérées comme des symboles solides;
- *Plusieurs cellules mobiles*: La machine supporte l'existence simultanée de plusieurs cellules mobiles disposant chacune leurs propres lois de comportements. Ceci introduit donc du parallélisme au niveau des cellules mobiles, avec deux approches envisageables, l'une synchrone, l'autre asynchrone, débouchant respectivement sur deux familles de bi-machines: nous nous concentrons ici sur la version synchrone du modèle. La possible multiplicité des cellules mobiles amène avec elles divers types de contraintes: (1) à l'initialisation du système, toutes les cellules mobiles doivent être disposées sur des cellules d'environnement toutes différentes et libres de tout symbole solide. (2) si deux cellules mobiles tentent de se déplacer ou d'écrire simultanément sur une même cellule de l'environnement, l'une d'entre elles, choisie de façon non déterministe, réussit tandis que l'autre échoue. (3) une cellule mobile qui tente de se déplacer sur une position déjà occupée par une cellule mobile échoue, et ce même si cette dernière la quitte à la même unité de temps;
- *Symbole d'une cellule mobile*: une cellule mobile est munie d'un porteur de symbole, ou signaliseur. Le symbole joue le rôle d'état perceptible ou état externe: il peut être perçu par les autres cellules mobiles, voire la cellule elle-même. Ce symbole a pour position celle de la cellule. Il ne peut être modifié que par sa propriétaire. Une cellule mobile qui tente d'écrire

sur le signaliseur d'une autre cellule échoue. Le signaliseur cache la cellule d'environnement sur laquelle est localisée son propriétaire; son symbole est toujours considéré comme étant solide et ne peut pas être le symbole blanc;

- *Actions d'une cellule*: "tourner-à-droite", "se-déplacer", "cesser-déplacement", "écrire-symbole(...)", etc sont des actions de cellules. La création d'une bi-machine exige de définir un ensemble fini d'actions disponibles pour les cellules mobiles. Un tel ensemble peut être muni d'une relation d'ordre, éventuellement partiel. A chaque unité de temps, la fonction de transition de la cellule détermine l'ensemble, éventuellement vide, des actions qu'elle doit exécuter. L'ordre d'exécution des actions à une étape donnée doit respecter la relation d'ordre définie sur l'ensemble des actions;
- *Division des cellules mobiles*: Une cellule mobile peut créer une copie d'elle-même. Ceci s'effectue par l'exécution d'une action appelée "division-cellulaire". La position, relative à la cellule mère, dans laquelle naît la cellule fille est un paramètre de la machine. En cas d'impossibilité (position déjà occupée), l'action échoue;
- *Arrêt d'une cellule mobile*: Quand la fonction de transition d'une cellule mobile n'est pas définie pour la configuration courante (états interne et externe, et symboles lus), la cellule mobile s'arrête et "meurt". Elle disparaît de l'environnement et simultanément un symbole est écrit à sa position. Une fonction appelée *fonction de dégradation* détermine le symbole à écrire selon la configuration courante;
- *Environnement dynamique*: Alors que dans une machine de Turing, l'entrée i.e. l'environnement, est fixée a priori, l'idée est d'autoriser ici un environnement dynamique. Une dynamique d'environnement doit être fournie comme entrée. Ceci est réalisé comme dans les automates cellulaires, notamment via la fonction de transition de l'environnement, à la seule différence que la présence des cellules mobiles influence la dynamique de l'environnement. Etant donnée une cellule d'environnement à mettre-à-jour, si l'une de ses voisines est occupée par une cellule mobile, la fonction de transition ne voit pas la cellule d'environnement du fait qu'elle est occultée par le signaliseur de la cellule mobile. La seule exception se produit pour la position de la cellule d'environnement en train de se mettre à jour: pour cette cellule et à cette position, les deux sont visibles, à savoir le symbole d'environnement qu'elle contient et le symbole de la cellule mobile.

3.2. Définition

Formellement une bi-machine est spécifiée par $(d, S, N, f_e, Q, A, <_A, p_w, p_d, n_0, \{f_i\}_{i \in \{1,2,\dots,n_0\}}, q_0, s_0, f_d)$, où:

- d est la dimension de la grille infinie¹,
- S est un ensemble fini de symboles de cellules ou états externes, aussi appelé alphabet,
- N est un voisinage fini,
- f_e est la fonction de transition (de l'environnement); $f_e: S^N \times S \rightarrow S$,
- Q est une ensemble d'états internes (de cellules mobiles),
- A est un ensemble fini d'actions (de cellules mobiles),
- $<_A$ est une relation d'ordre (éventuellement partiel) définie sur A ,
- $p_w \in N$ est la position relative où un symbole est écrit,
- $p_d \in N$ est la position relative où est placée la copie d'une cellule après division cellulaire,
- n_0 est le nombre initial de cellules mobiles,
- f_i sont les fonctions de transition (éventuellement partielles) des cellules mobiles; $f_i: (Q \times S) \times S^N \rightarrow (Q \times S) \times P(A)$,
- $q_0 \in Q$ est l'état interne initial pour toutes les cellules mobiles,
- $s_0 \in S$ est l'état externe initial pour toutes les cellules mobiles,
- f_d est la fonction de dégradation des cellules mobiles; $f_d: (Q \times S) \times S \rightarrow S$,

Dans ce formalisme nous voyons que la fonction de transition d'une cellule mobile détermine seulement un nouvel état (interne et externe) de la cellule ainsi qu'un ensemble d'actions, telles que "tourner-à-droite", "division-cellulaire", etc, à exécuter. Ces actions ne prennent en compte que les états interne et externe et les symboles du voisinage. Elles peuvent échouer, auquel cas leur

¹De façon classique, au niveau de l'implémentation, cette grille infinie est remplacée soit par une grille toroïdale ou par une grille finie délimitée par des symboles solides.

effet est celui d'une action vide. Les actions sont notamment le moyen d'écrire des symboles: la seule manière pour une cellule mobile d'écrire un symbole est de le faire via une action: contrairement à ce qui se passe dans les machines de Turing, la fonction de transition ne détermine pas ici directement un symbole à écrire. Il est également possible de définir des actions qui modifient l'état de la cellule.

4. Bio-machines d'ordre 1

Ce paragraphe se consacre à la présentation du modèle des bio-machines. Suità une présentation générale, nous décrivons plus précisément les fonctions de transition.

4.1. Présentation générale

Nous contruisons maintenant les bio-machines par l'adjonction aux bi-machines d'un certain nombre de contraintes et capacités orientées vers les problèmes visés:

- *Contraintes énergétiques*: Une première contrainte importante concerne l'énergie. L'état interne d'une cellule est enrichi d'un élément représentant son énergie disponible. Chaque action d'une cellule mobile est affectée d'un certain "prix énergétique", quelle que soit cette action (cf aussi § reproduction). Ce prix doit être fixé pour toute action dans une machine donnée; il est indépendant des cellules (ce qui peut s'interpréter comme le fait que toutes les cellules ont, entre autres, même "masse"). Une cellule qui exécute une action doit en payer le prix. Si la cellule n'a pas suffisamment d'énergie pour exécuter cette action, l'action échoue. Le fait d'exister pour une cellule a aussi un prix. Quand l'énergie d'une cellule mobile tombe à zéro, la cellule meurt (i.e. elle s'arrête) et la fonction de dégradation est appelée. Une action appelée "*manger*" permet à la cellule d'augmenter son énergie: cette action efface un symbole d'une classe spéciale de symboles appelée "*nourriture*" et écrit à sa place le symbole blanc. En cas de succès de cette action, l'énergie de la cellule est augmentée d'une valeur qui dépend exclusivement du symbole "mangé".
- *Dynamique de l'environnement*: La contrainte d'uniformité de l'environnement est ici relâchée, ce qui a pour effet de déboucher sur des automates cellulaires non uniformes (ou hétérogènes). Comme cela a déjà été signalé dans [Sipper 1994], ceci préserve la plupart des caractéristiques essentielles du modèle original: localité des interactions, parallélisme massif et simplicité des cellules (automates d'états finis). La raison de cette modification est que l'uniformité n'est pas absolument nécessaire dans le contexte qui nous intéresse: les automates cellulaires sont ici utilisés pour modéliser des environnements et non pas l'univers. La contrainte sur le déterminisme des fonctions de transitions est aussi relâchée. En général, la dynamique de l'environnement est supposée écrire –i.e. créer– des symboles "*nourriture*", générer et transmettre des perturbations, etc. Néanmoins, il est également permis d'avoir des dynamiques d'environnement qui créent de nouvelles cellules mobiles. Pour ce faire, un symbole spécial "*création-de-cellule-mobile-aléatoire*" est mis à disposition. Lorsqu'une cellule d'environnement atteint cet état, une cellule mobile de génome aléatoire est générée à sa position.
- *Reproduction*: Les cellules peuvent se reproduire. Par opposition aux bi-machines, dans lesquelles la reproduction est toujours sans erreur, les cellules mobiles des bio-machines peuvent subir des mutations à la reproduction. Dans les bio-machines, les cellules sont dotées d'un niveau de représentation génotypique et deux types de reproduction sont prévues, respectivement asexuel et sexuel.

Dans le cas de la reproduction asexuelle, une copie éventuellement mutée du génome est réalisée et son interprétation (lorsque la cellule se divise) donne lieu à la création d'une nouvelle cellule. Si la reproduction réussit, la cellule mère transmet à la cellule fille une partie de son énergie, qui devient son énergie utilisable initiale.

Dans le cas de la reproduction sexuelle, la cellule qui exécute l'action, doit avoir devant elle (i.e. dans la position où elle se retrouverait si elle avançait) une autre cellule mobile jouant le rôle de partenaire. Le matériel génétique des deux est combiné et le résultat est utilisé pour créer une nouvelle cellule. L'opérateur de recombinaison est un paramètre de la machine.

Dès lors que la reproduction est autorisée, nous pouvons parler d'individus et d'espèces.

L'évolution par sélection naturelle permet la découverte des "bons" comportements de cellules mobiles (cf "fonction de fitness a posteriori" dans [Packard 1988]). Ces comportements sont donc changés par l'interaction des espèces avec l'environnement, via les individus.

- *Adaptation individuelle*: L'interaction d'un individu avec son environnement influence la manière dont il se modifie durant sa vie et éventuellement s'adapte à cet environnement. Cette possibilité d'adaptation des cellules mobiles est concrètement supportée par la mise à disposition d'actions spéciales permettant les cellules mobiles à changer leur fonction de transition (Cf. § 4.2).

4.2. Représentation des fonctions de transition

Un point fondamental des bio-machines concerne la représentation des fonctions de transitions. Deux mécanismes doivent être considérés: l'évolution et l'adaptation individuelle. Le premier opère au niveau génotypique, le second au niveau phénotypique. Pour les comprendre, certaines précisions sur les représentations choisies sont nécessaires. Nous ne discutons ici que la manière dont la fonction de transition est encodée dans le phénotype, autrement dit de la représentation opératoire de la fonction de transition.

La représentation adoptée pour les fonctions de transition des cellules mobiles est un réseau de neurones non récurrent à une seule couche et à non-linéarité fortement limitée. Les entrées du réseau sont formées des états interne et externe et des symboles du voisinage, tandis qu'un neurone de sortie est associé à chaque action. Pour chaque neurone de sortie, une somme pondérée des entrées paramétrée par un vecteur de poids représentant les synapses spécifie son activation, et le résultat subit un filtrage non linéaire. Ce choix n'est pas seulement motivé par le mimétisme biologique qui en découle, mais aussi par la représentation concise et robuste qu'il permet d'obtenir. Il fournit une solution simple au problème de l'adaptation individuelle des fonctions de transition, par simple ajustement des poids du réseau (ceci se fait à l'aide d'actions). Un autre avantage du choix effectué, qui se révèle dans l'association de cellules proposée ci-après (cf § 5), est que la non-linéarité injectée dans la définition des fonctions de transitions rend leur composition beaucoup plus intéressante.

Une conséquence de ce choix de représentation est qu'il requiert d'associer à tout symbole de S et à tout état de Q une valeur numérique servant d'entrée aux senseurs. La solution que nous avons adoptée consiste à représenter les symboles et les états par des vecteurs de nombres. Différents senseurs sont ensuite spécialisés sur différentes composantes de ces vecteurs. Un symbole ou un état peut être sensé par un ou plusieurs senseurs.

5. Bio-machines d'ordre 2

5.1. Introduction au modèle

Une forme particulière de bio-machine est ici présentée. Dans ce modèle, les cellules mobiles incluent dans leur répertoire d'actions des actions permettant à des cellules adjacentes de s'associer pour former des composites. De cette possibilité d'associations de cellules en entités de deuxième ordre résulte précisément l'appellation du modèle.

Une association a pour effet de créer des *liens* entre les cellules. Ceci leur impose certaines contraintes, comme l'obligation de se mouvoir ensemble et de diffuser leur énergie au sein de l'association. Ainsi que nous le détaillons plus loin, un lien autorise également le transfert d'information génétique.

Nous distinguons deux types d'associations appelées respectivement "en colonies" et "en multicellulaires". Une association en colonie ne concerne que le niveau phénotypique des individus; elle n'est donc pas héritée. En revanche, une association en multicellulaire ajoute à l'association en colonie une association génétique qui confère à l'organisme la capacité de se reproduire, et donc d'évoluer, en tant que tel.

L'association en colonie est aussi appelée *association-p* (p pour phénotype), tandis que l'association en multicellulaire est appelée *association-g* (g pour génotype). L'association-g implique l'association-p. Dans la version actuelle du modèle, une association-g a comme post-

condition une association-p, cette dernière n'étant pas une pré-condition de la première. L'étude de la version plus primitive, dans laquelle une association-g a effectivement comme pré-condition une association-p, est cependant prévue.

Une action permettant la dissociation est également fournie. Celle-ci n'opère qu'au niveau des associations-p; rien ne peut pour l'instant dissocier une association-g.

5.2. Terminologie

Une terminologie relative au rôle joué par les cellules dans les entités composites nous est à présent nécessaire.

Une cellule qui exécute avec succès une action d'association, est appelée *initiatrice de l'association*. La cellule qui est l'objet de l'action d'association est appelée la *cellule subordonnée*. La subordination est une relation transitive: nous appelons aussi subordonnée d'une cellule, toute subordonnée d'une subordonnée de cette cellule, et ce récursivement. La relation réciproque de la subordination est la *super-ordination*.

Une entité composite est originellement initialisée par une cellule, qui s'associe avec d'autres, lesquelles peuvent, de même que l'initiatrice, s'associer encore etc. La cellule initiatrice originale est appelée la tête du multicellulaire; c'est aussi la seule cellule du multicellulaire qui n'est subordonnée à aucune autre.

5.3. Liens entre cellules

Un lien entre cellules force ces cellules à se diffuser mutuellement leur énergie disponible et à se mouvoir ensemble.

En ce qui concerne la diffusion d'énergie, une solution très simple a été choisie. La diffusion ne dépend pas du comportement de la cellule: elle est uniquement gouvernée par des lois physiques. Pour toute cellule mobile, la différence d'énergie utilisable par rapport à toutes ses cellules associées (subordonnées ou super-associées) est calculée: un processus de diffusion (dont le sens est déterminé par le signe de la différence) est activé entre la cellule et l'une des cellules maximisant la différence choisie arbitrairement. Ceci évite les grandes oscillations.

Le déplacement groupé constitue une exigence plus problématique. Il est tout d'abord nécessaire d'observer que dans le contexte de notre machine synchrone et uniquement basée sur des interactions locales, des liens parfaitement rigides sont inconciliables avec cette possibilité, du moins si l'on tient à opérer de façon simple et élégante et sans introduire de limitation drastique quant à la taille maximale du multicellulaire ou la vitesse maximale de déplacement par exemple. La raison à cela est que la cellule subordonnée ne remarque que sa cellule "super-ordonnée" s'est déplacée qu'au pas suivant le déplacement. Dans cet univers, la transmission de l'information a une vitesse maximale ("la vitesse de la lumière"). Ceci implique qu'une distance fixe ne peut être maintenue entre les cellules associées. Par conséquent, une certaine flexibilité est requise.

Dans la solution adoptée, non seulement la flexibilité mais aussi l'élasticité est nécessaire. Considérons le cas d'une translation. Quand la distance entre deux cellules associées dépasse une valeur fixée (définie comme un paramètre de la machine), les cellules ne sont plus autorisées à s'éloigner davantage. La cellule subordonnée est forcée au pas suivant d'aller dans une direction qui restaure la distance normale, c'est-à-dire minimise l'énergie potentielle. La cellule super-ordonnée doit attendre cette restauration: elle n'est pas autorisée à avancer jusqu'à ce que la cellule subordonnée n'ait atteint la position d'énergie potentielle minimale. Une conséquence de cette règle est que seule la tête peut librement décider d'un déplacement, bien qu'elle soit souvent stoppée par le lien. Une règle analogue permet également à une cellule de "pousser" une subordonnée. D'autres règles s'appliquent également en cas de blocage d'une cellule subordonnée par un solide; dans ce cas, celle-ci peut influencer le mouvement de sa "super-ordonnée". Des considérations similaires non développées ici, sont aussi valables pour les rotations.

Notons qu'un lien entre deux cellules maintient les senseurs de voisinage direct de chaque cellule en perception mutuelle. Les senseurs du voisinage direct peuvent ainsi être conçus comme situés sur la membrane sensible et *flexible* d'une cellule mobile. Ceci préserve la topologie des connexions des senseurs entre des cellules directement associées. Cette préservation est très importante car une grande partie de la fonctionnalité de l'entité multicellulaire pourrait être enracinée dans sa topologie. Certaines cellules interconnectées de l'entité peuvent en effet même évoluer pour jouer le rôle de réseaux de neurones; le fait de trop perturber la topologie serait donc

a fortiori fatal pour l'entité.

5.4. Reproduction et développement embryonique

Les bio-machines de premier ordre n'exigeaient qu'une fonction permettant de déterminer un phénotype à partir d'un génotype. Aucun développement embryologique n'y avait lieu. Dans les bio-machines d'ordre 2, qui permettent l'existence d'entités multicellulaires, la situation est plus complexe. Le fait que seules des interactions locales sont possibles implique, comme nous allons le voir, qu'une sorte de processus de développement embryologique soit nécessaire.

Dans une bio-machine de deuxième ordre, chaque cellule ne connaît que sa propre part du génome global. Cette part inclut le codage de sa relation avec les cellules auxquelles elle est directement associée. Outre cela, elle peut mémoriser temporairement une partie du génome d'une autre cellule. Ce mécanisme est ce qui permet la transmission de matériel génétique entre cellules. Il rend possible la construction d'une copie complète du génome, qui distribué à travers l'entité, peut parvenir à un embryon (cette transmission est réalisée en un temps proportionnel au nombre de cellules traversées).

Considérons le cas d'une reproduction asexuelle. La cellule exécutant l'action menant à une reproduction asexuelle envoie une requête de code génétique à toutes ses associées directes qui, à leur tour, retransmettent cette requête à leur propres associées, etc. Sans entrer dans les détails algorithmiques, la première portion du génome qui arrive à cette cellule après un certain nombre de pas est celle qui correspond à la tête. En utilisant cette portion du génome, un œuf est ensuite généré. Aux pas suivants, une copie des portions du génome des cellules subordonnées à la tête sont transmises à l'œuf. Cet œuf utilise cette information pour créer ses propres cellules associées. Par l'application répétée de cette procédure (en pipeline), l'entièreté du génome parvient finalement à l'embryon. Durant ce processus, la tête de l'embryon, c'est-à-dire sa cellule initiale, est considérée comme p-associée à l'entité mère, d'où la subordination de toutes ses cellules à cette entité mère. La tête ne devient indépendante qu'au terme de la gestation: le "cordon ombilical" est alors rompu, la diffusion d'énergie entre la mère et l'enfant cesse et l'enfant peut s'éloigner de sa mère.

Dans le cas d'une reproduction sexuelle, le processus est similaire, excepté que le matériel génétique arrive à l'embryon depuis ses deux parents. La seule précision à apporter ici est que des deux parents, l'initiateur de la reproduction sexuelle conserve son indépendance durant le développement embryologique, tandis que l'autre parent reste subordonné.

6. Conclusion

Dans la mesure où nous considérons comme légitime le fait de parler de société y compris au niveau du biologique, la problématique du collectif au social a été abordée sous l'angle de la modélisation en vie artificielle. Une nouvelle famille de machines appelées bio-machines, destinée à capter certaines propriétés d'organismes simples comme des unicellulaires et des multicellulaires primitifs a été définie; celle-ci tente d'éclairer la transition du collectif au social par le biais d'une étude systématique sur les types et niveaux d'interactions à envisager entre des entités biologiques.

Le modèle actuel a été élaboré en trois étapes. Tout d'abord un modèle formel intermédiaire appelé bi-machine, a été défini. Ce modèle, conçu comme une fusion des machines de Turing et des automates cellulaires, supporte la création de cellules mobiles capables de division. Sur cette base, les bio-machines d'ordre 1 ont ensuite été construites. Celles-ci rajoutent essentiellement aux bi-machines un niveau de description génotypique assorti d'un mécanisme de reproduction sexuelle par combinaison de matériel génétique et mutation, ainsi qu'une possibilité d'adaptation des individus durant leur existence. Enfin, les bio-machines d'ordre 2, qui permettent aux cellules de s'associer en colonies et multi-cellulaires, ont été amenées.

Notons que diverses extensions et raffinements du modèle sont en train d'être considérés, en particulier des extensions visant à enrichir les modalités de transition vers les multicellulaires et permettant d'appréhender certaines formes de parasitisme.

Sur le plan pratique, un prototype permettant de simuler des bio-machines d'ordre 1 et 2 en dimension 2 a été réalisé. Ce logiciel a donné lieu à quelques expérimentations préliminaires portant sur l'évolution et l'adaptation et touchant à des questions comme la formation des espèces, la variabilité inter-espèces, ainsi que l'apparition de colonies et de multicellulaires. Les résultats

sont encourageants – parfois même surprenants! –, mais demeurent encore à analyser finement. Le développement d'une version parallèle et plus générique du logiciel de simulation est également en cours.

Concernant le thème du collectif au social, l'étape de conception du modèle et les réalisations qui en découlent ont occasionné un certain nombre de réflexions sur les types de dépendances susceptibles d'exister entre entités biologiques. Notre position actuelle est ainsi de considérer comme indissociables les concepts de société et d'autonomie des entités: ni des liens trop faibles entre des individus, ni des liens trop forts, impliquant une diminution de l'autonomie des individus, ne sont appropriés pour parler de société. Le social circonscrit donc pour nous une sorte de zone médiane, autour de laquelle les comportements peuvent être qualifiés de strictement collectifs. Métaphoriquement projetée sur les états de la matière, le social serait l'analogie de l'état liquide, par opposition aux états respectivement solide –trop contraint– et gazeux –insuffisamment contraint–. Dans les bio-machines ces idées sont concrétisées par les différents niveaux d'interaction proposés: depuis l'absence d'interaction jusqu'aux très fortes associations de type "g", en passant par les niveaux intermédiaires des interactions sensorielles et des associations moyennes de type "p".

Suite à ces réflexions, de prochaines expérimentations portant l'apparition de certains comportement sociaux fascinants tels que "l'altruisme" sont prévues.

Références

- Ben-Jacob, E. (1995). *Physics of Biological Systems: From Molecules to Species*. European Nordic Summer School and Workshop, Humlebak, Denmark.
- Dana Eckart, J. (1995). *A Cellular Automata Simulation System*. Radford University.
- Gardner, M. (April 1970). The Fantastic Combination of John's Conway New Solitaire Game 'Life', *Scientific American* 223:4, 120-123.
- Goel, N. and Thompson, L. (1988). Movable Finite Automata: A new Tool for Computer Modeling of Living Systems. *Artificial Life*, Addison-Wesley Publishing Company.
- Hillis, D. (1989). *The Connection Machine*, Cambridge, MIT Press.
- Holland, John (July 1992). The Echo Model. Proposal for a Research Program in Adaptive Computation, Santa Fe Institute.
- Hopcroft, J.E. (May 1984). Turing Machines. *Scientific American* 250, pp. 86-98.
- Hopcroft, J. and Ullman, J. (1979). *Introduction to Automata Theory, Languages and Computation*, Addison-Wesley Publ. Company.
- Jones, Terry and Forrest, Stephanie (1993). An Introduction to SFI Echo. Santa Fe Institute TR 93-12-074.
- Langton, C. (1988). Preface. *Artificial Life*, Addison-Wesley Publ. Company.
- Levy, S. (1992). *Artificial Life*. Jonathan Cape, London.
- Toffoli, T. and Margolus, N. (1987). *Cellular Automata Machines*. The MIT Press.
- Packard, N. (1988). Intrinsic Adaptation in a Simple Model for Evolution. *Artificial Life*, Addison-Wesley Publishing Company.
- Ray, T. (1991). An approach to the synthesis of life. *Artificial Life II*, Addison-Wesley Publishing Company, 371-408.
- Ray, T. (1995). *An Evolutionary Approach to Synthetic Biology: Zen and the Art of Creating Life*. *Artificial Life: An Overview*, The MIT Press.
- Sipper, M. (1994). Non-Uniform Cellular Automata: Evolution in Rule Space and ... *Artificial Life IV*, MIT Press.
- Stephenson, I. (1992). Creature Processing: An Alternative Cellular Architecture. Technical Report ASEG92.04, Department of Electronics, Univ. of York.
- Sudkamp, T. (1988). *Languages and Machines*. Addison-Wesley Publishing Company.
- Turing, A.M. (1936). On Computable Numbers with an Application to the Entscheidungs Problem. *Proc. London Math. Soc.*, 42 (Series 2), 230-265.

Ulam, S. (1950). Random Processes and Transformations. Proceedings of the International Congress in Mathematics of 1950, Vol. 2, 264-275.

Von Neumann, J. (1966). Theory of Self-Reproducing Automata. Completed and edited by Arthur Burks. Univ. of Illinois.

Woods, D. (1987). Theory of Computation, Harper & Row Publishers, New York.

Yaeger, Larry (1994). Computational Genetics, Physiology, Metabolism, Neural Systems, Learning, Vision, and Behavior or PolyWorld: Life in a New Context. Artificial Life III, Addison-Wesley.

Les systèmes de signes comme fondateurs des sociétés. Le cas des systèmes multi-agents.

Anne Nicolle, Ludovic Delépine

GREYC URA CNRS 1526 - Université de Caen - 14032 Caen Cedex

Mail : Anne. Nicolle@info. unicaen. fr

Tel : 31 45 59 65

Fax : 31 45 58 14

Résumé

Une analyse des conditions dans lesquelles des systèmes informatiques distribués peuvent être vus comme formant des sociétés avec les personnes qui les utilisent est proposée dans cet article : la question centrale du passage des collectivités de machines aux sociétés d'agents est la nature du langage employé pour les interactions, seule une langue naturelle peut permettre une interaction assez souple pour faire émerger des interactions sociales entre les participants - personnes et machines - d'une collectivité d'agents.

mots clés : multi-agents, sémiotique, interaction

1- Les réseaux de machines du collectif au social.

Mettre des machines en réseau comme le fait Internet crée un collectif de machines mais pas une société. Ce n'est pas une société parce qu'il n'y a pas autonomie des machines pour deux raisons principales :

- les informations données sur la "toile" (ou WWW) sont entièrement à la charge de l'humain qui les y met. Les machines ne font que présenter des informations qui leur ont été données sans les comprendre et sans même les contrôler.
- le langage de communication est un code partagé, exécuté de manière algorithmique sans décisions à prendre de la part de la machine.

Les machines sont alors des outils ou des instruments perfectionnés. Peut-être pourrait-on dire qu'elles font partie des sociétés humaines, comme les voitures ou les chiens en font partie, et sans doute de manière plus profonde car les réseaux d'informations structurent en profondeur les sociétés. Mais ils ne forment pas en tant que tels une société dans la société, comme une classe d'une école, un laboratoire ou un bureau en forment une.

Les systèmes multi-agents ou SMA conçus en intelligence artificielle, au contraire insistent sur l'autonomie des agents logiciels. Ils mettent en oeuvre des principes d'interaction variés entre des machines et des personnes. Il est donc raisonnable de se poser la question de savoir s'ils forment des sociétés, ou de se demander à quelles conditions ils peuvent en former. Le point de vue défendu dans cet article est que la nature du langage d'interaction est le point central du passage du collectif au social dans les systèmes multi-agents.

2- L'intelligence artificielle distribuée

L'IA distribuée est une branche récente de l'IA qui propose de remplacer les logiciels conçus de manière centralisée par des logiciels basés sur l'interaction d'agents. Un agent logiciel est une entité informatique qui contient ses propres données et un protocole d'exécution élémentaire. Ce protocole peut être un calcul, mais le plus souvent il consiste à tester des conditions sur ses données et à réaliser les actions qui correspondent aux conditions les plus spécifiques qui sont vérifiées. Les actions les plus courantes qu'un agent peut accomplir sont de modifier ses propres données et d'envoyer des messages à d'autres agents. Ils peuvent aussi modifier l'environnement via les ports de sortie de la machine hôte, l'écran de la machine par exemple fait partie de l'environnement. Un agent ne peut pas agir directement sur les représentations internes d'un autre agent : quand il envoie un message, l'agent qui le reçoit va l'interpréter et y répondre selon son propre protocole. Les agents qui communiquent par messages doivent connaître les agents auxquels ils peuvent envoyer des messages. Les agents peuvent aussi communiquer en plaçant des messages dans une mémoire accessible par plusieurs agents, ils peuvent alors communiquer sans se connaître, où les donner à un agent postier qui peut faire une diffusion en fonction de conditions vérifiées par les agents et pas en fonction de leur nom.

L'IA distribuée s'est développée pour répondre à plusieurs préoccupations :

- d'abord, il est difficile de concevoir, de valider et de maintenir des logiciels qui dépassent une certaine taille. Concevoir des petits agents qui interagissent est une application du principe « diviser pour régner » qui peut faciliter la mise au point, le test et permettre une meilleure réutilisabilité.
- ensuite certains problèmes, comme la simulation, sont mieux résolus de manière décentralisée, parce que c'est la manière dont ils se présentent naturellement. Il est très difficile d'envisager à priori tous les cas qui peuvent se produire dans un environnement où il y a beaucoup d'agents différents car les cas possibles sont trop nombreux. Il est beaucoup plus facile de décrire ce que chaque agent doit faire au cas par cas, car le nombre de cas qui sont différents du point de vue d'un agent particulier est faible. C'est la combinatoire d'ensemble qui est trop importante.
- enfin ce paradigme est bien adapté aux architectures distribuées, quand il s'agit de résoudre un problème dont les données et les utilisateurs sont répartis sur plusieurs machines en réseau.

3- Les agents logiciels

La notion d'agent est apparue en intelligence artificielle pour décrire des entités logicielles autonomes exécutant une boucle de calcul visant à se maintenir en état de fonctionnement dans un environnement changeant (viabilité). On appelle donc agent au sens propre une entité informatique logicielle ayant des données locales - la mémoire - un ou plusieurs points d'entrée extérieurs scrutés en permanence - la perception - certaines capacités d'action et un protocole de comportement correspondant à une boucle infinie

perception -> action

fondée sur une catégorisation des perceptions : toute perception est classée dans une catégorie de la mémoire et déclenche les actions correspondant à cette catégorie.

On distingue la plupart du temps trois sortes d'agents, bien que cette classification ne soit pas stricte et puisse être raffinée :

- **les agents réactifs** qui n'ont pas d'intelligence propre, mais dont une population pourra produire un comportement intelligent par émergence, à l'exemple des sociétés d'insectes.

- **les agents rationnels** qui sont spécialisés pour résoudre un problème particulier avec un certain point de vue. Ils ont des buts explicites qu'ils cherchent à atteindre, et leur protocole de comportement est une boucle

perception -> délibération -> action.

Entre la perception et l'action s'ajoute une activité rationnelle de délibération qui permet de prendre des décisions pour agir de manière optimale par rapport à des buts. La condition d'optimalité suppose qu'on puisse évaluer les choix issus de la délibération. Les agents rationnels peuvent coopérer pour résoudre des problèmes complexes, en décomposant les problèmes en sous problèmes, ou en essayant en parallèle des méthodes correspondant à des points de vue différents. Chacun d'eux fonctionne avec ses propres représentations et sa propre logique, mais il faut supposer qu'ils ont en commun un principe sous-jacent de rationalité. Ils communiquent avec un langage symbolique partagé.

Un agent rationnel doit agir de manière optimale par rapport à ses buts, ce n'est possible que si le problème admet une formalisation suffisamment précise, en particulier si le problème admet des situations énumérables et que les solutions du problème sont reconnaissables de manière intrinsèque. Mais que se passe-t-il quand il n'y a pas de manière intrinsèque pour caractériser les situations d'un problème qui sont des solutions ? Si on peut cependant comparer deux solutions entre elles pour savoir laquelle est la meilleure, on est dans un problème d'optimisation et la qualité de la solution va dépendre des compétences de l'agent. Dans ce cas il faut faire un compromis entre le temps passé à la recherche et la qualité de la solution obtenue et H. Simon a introduit la notion de « solution satisfaisante » pour décrire l'arrêt de la recherche des solutions pour ces problèmes [Simon 83]. Il parle alors de rationalité limitée car la solution obtenue dépend de l'environnement et des compétences de l'agent qui la produit. Par exemple, si on cherche une aiguille dans une botte de foin, on sera moins regardant sur la taille de l'aiguille que si on la cherche dans une trousse de couture, et l'agent qui a un aimant aura sans doute de meilleurs résultats. La rationalité limitée introduit des agents d'un nouveau type, capables de gérer un compromis entre la recherche d'une solution encore meilleure et le temps nécessaire pour l'atteindre, compte tenu de leurs propres compétences.

Un agent rationnel peut être généraliste ou spécialiste. Les agents généralistes connaissent des méthodes générales de recherche et savent agréger les solutions partielles produites par des agents spécialistes. Ils ont des métaconnaissances plutôt que des connaissances [Pitrat 90]. Les agents spécialistes ont des connaissances approfondies d'un domaine. Ces agents rationnels sont utilisés pour construire des systèmes multi-experts dont le comportement global est obtenu par négociation entre les experts à partir des solutions proposées par chacun.

- **les agents cognitifs** ont une plus grande autonomie, leur mémoire est fondée sur un principe de catégorisation au lieu d'être fondée sur un principe de reconnaissance de modèles. Ils ont une intelligence propre, plus ou moins développée, au sens plus traditionnel de l'intelligence artificielle de capacité à résoudre des problèmes, mais aussi, ils ont des capacités d'adaptation et d'argumentation... Ils ont plutôt des intentions que des buts, et ce sont les intentions qui sont le ressort de l'interaction langagière [Jullien & al. 95]. Ce sont ces agents qui vont pouvoir construire des sociétés d'agents.

Les agents les plus complets et les plus autonomes sont des agents réflexifs : ils ont un modèle de soi et un modèle des autres, ils peuvent apprendre par l'expérience, par imitation, ils peuvent utiliser les langues naturelles pour communiquer et apprendre par usage de la langue. A part les êtres humains, il n'existe pas à ma connaissance d'agents ayant toutes ces capacités, mais c'est un objectif que d'en concevoir et d'en construire.

4- L'interaction entre les agents

Jacques Ferber donne la définition suivante de l'interaction : « Une interaction est la mise en relation dynamique de deux ou plusieurs agents par le biais d'un ensemble d'actions réciproques. » Et il voit dans l'interaction le fondement des organisations sociales : « Les interactions sont non seulement la conséquence d'actions effectuées par plusieurs agents en même temps, mais aussi l'élément nécessaire à la constitution d'organisations sociales. » [Ferber 95 p. 67]. S'il décrit les différentes sortes d'agents comme ici dans le §2, il n'en tire pas de conséquences sur les différentes formes d'interaction possibles et donc sur les différentes formes de groupes (collectivités ou sociétés) qui peuvent en résulter parce qu'il renonce à prendre le langage dans sa dimension essentielle, la co-référentiation : « Malheureusement, cette conception de la communication comme source de construction du modèle cognitif des interlocuteurs et du langage lui-même semble un peu trop en avance pour les possibilités techniques dont nous disposons... » [Ferber 95 p. 339].

Or les systèmes multi-agents peuvent être vus de deux façons différentes. Ils sont vus le plus souvent comme une organisation d'agents logiciels en interaction. Mais ce sont des humains qui les ont conçus et qui les utilisent et les faire sortir du système est réducteur. Nous élargirons donc ce point de vue en appelant système multi-agents un collectif comportant des agents personnes et des agents logiciels installés sur des machines en réseau. On fait alors ressortir l'importance de l'interaction des agents logiciels avec leurs utilisateurs et on fait apparaître trois formes d'interaction structurant ce collectif :

- entre des personnes via les machines (mail, collecticiels...)
- entre les machines
- entre une personne et une machine.

Cette vision des choses ne permet plus de faire l'économie de la co-référentiation dans le dialogue pour être bien adaptée aux agents humains. C'est justement ce problème que nous attaquons à travers l'étude des dialogues humains-machine [Nicolle & Bricon-Souf 94], et il est indispensable de le prendre en compte dans les systèmes multi-agents si on considère l'interaction humain-machine comme essentielle.

De la définition d'agent donné dans [Ferber 95, p13], nous ne conservons que les caractéristiques essentielles liées à l'interaction, qui s'appliquent aussi bien aux agents humains qu'aux agents logiciels. Un agent est une entité physique ou virtuelle qui :

- a) est capable d'agir dans un environnement,
- b) peut communiquer directement avec d'autres agents.

Les agents logiciels vont être dotés de comportements par leurs concepteurs. Un comportement se manifeste par une collection d'actions qu'un agent peut réaliser pour satisfaire ses intentions, et par une stratégie de choix de la meilleure action à faire dans un contexte donné. Cette stratégie est basée sur l'évaluation de critères et sur leur ordonnancement ou leur pondération. C'est ensuite la dynamique du système qui va permettre la réalisation des actions à partir du choix d'un état de départ du système multi-agents. Une fois fixé le comportement de l'agent, il doit fonctionner de manière autonome, sans intervention extérieure autres que les actions d'interaction qui peuvent provoquer des transferts. Ce sont les transferts entre l'agent et le monde extérieur qui permettent l'évolution de l'agent. Cette évolution peut consister à tester certaines conditions qui ne l'étaient pas avant, à acquérir de nouveaux critères de choix ou à modifier l'ordonnancement ou la pondération des critères.

Les actions d'interaction mettent en jeu au moins deux agents, les pôles, dont les positions possibles sont : émetteur (celui qui envoie le message) et récepteur (celui qui reçoit le message). Prenons à titre d'exemple deux types de comportements : passif et actif. Les agents passifs ne sont pas des agents au sens propre, ce sont plutôt des objets dans l'environnement, mais les agents actifs peuvent très bien les voir comme des agents. Les agents actifs peuvent être réactifs, rationnels ou cognitif, les différences ne seront prises en compte que plus tard. En déployant pour les deux

positions possibles, émetteur et récepteur ces deux comportements, on obtient la représentation suivante :

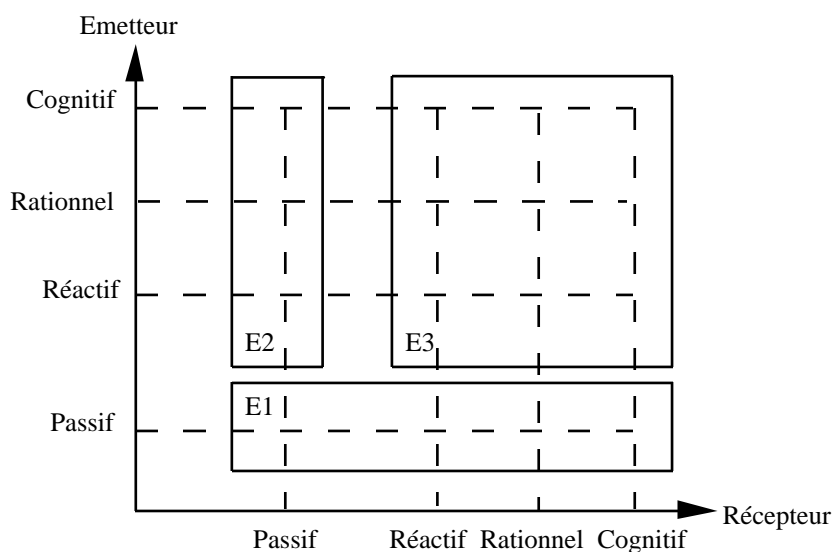


figure 1 : les 3 espaces d'interaction

Cette représentation obtenue permet de mettre en évidence trois types d'espaces d'interaction à partir des comportements des agents impliqués :

E1 - il n'y a pas d'action car l'émetteur n'envoie pas de message,

E2 - espace représentant les actions de type communication indirecte. Il s'agit du partage d'un espace de données commun, la communication étant alors réalisée par un mécanisme de filtrage (pattern matching). Cette structure peut correspondre à l'espace de résolution d'un problème, éventuellement partitionnée en niveau d'abstraction (par exemple, un tableau noir) ou bien représenter un espace physique dans lequel évolue l'agent (par exemple, communication par l'environnement comme dans le modèle des fourmis ou le code de la route),

E3 - espace représentant les actions de type communication directe. Celles-ci peuvent se faire par envoi point à point, en désignant le récepteur par son adresse ou par diffusion de messages à la ronde, la détermination des destinataires pouvant s'établir par adresse ou par calcul.

Seul l'espace E3 rend possibles des interactions, c'est à dire des actions réciproques de deux agents l'un vis à vis de l'autre. Voir l'interaction comme une action réciproque amène à introduire une Unité Minimale d'Interaction (UMI) pour rendre compte de l'alternance des positions d'émetteur et de récepteur des pôles impliqués dans l'interaction et pour structurer la communication. L'énoncé, qui est l'objet de la pragmatique¹, ne peut pas être cette unité minimale d'interaction car il ne rend pas compte de la présence des deux actants, qui distingue le dialogue du discours. Dans les premiers modèles du dialogue, les UMI reposent sur la notion de paire adjacente, elles regroupent un énoncé et la réponse qui le suit [Roulet 85]. Le dialogue est alors structuré par la mise en évidence des relations entre ces UMI. Mais Brassac a montré l'importance du troisième énoncé dans la co-référenciation [Brassac 94] et Luzzati, en introduisant le niveau d'incidence, a montré que la portée d'un énoncé pouvait dépasser de beaucoup l'énoncé qui le suit [Luzzati 95]. Toutefois, ce qui est commun à ce premier paradigme de modélisation, c'est que l'interaction est observée par quelqu'un d'extérieur à l'interaction. Les UMI sont reconstituées à partir de ce qui est observé. Dans un deuxième paradigme de modélisation [Lehuen & al. 96], il ne s'agit pas de décrire un dialogue qui a eu lieu mais de donner aux machines les moyens de tenir un

¹Alors que la phrase est l'objet de la linguistique, et qu'elle est caractérisée par sa structure syntaxique et sa signification, l'énoncé est le produit de l'énonciation d'une phrase, ou plus généralement d'une expression linguistique, correcte ou incorrecte, dans un contexte de communication. [Moeshler et Reboul, 1995 p.. 22]

dialogue. Les modèles descriptifs ne suffisent pas, il faut des modèles effectifs et l'interaction se représente du point de vue d'un agent. De ce point de vue, il n'y a pas de symétrie entre les actions de l'un des actants et celles de l'autre, il y a donc deux découpages possibles des suites d'énoncés en fonction de celui des actants qui se représente le dialogue en cours. Prenons par exemple le dialogue homme machine, soient les agents M et H et les énoncés successifs M1, H1, M2, H2. Du point de vue de M, on peut décrire deux UMI regroupant les énoncés M1 et H1 puis M2 et H2 alors que du point de vue de H, il y a 3 UMI, f et M1 (initialisation de la communication), puis H1 et M2 et H2 qui est la première partie de l'UMI en cours.

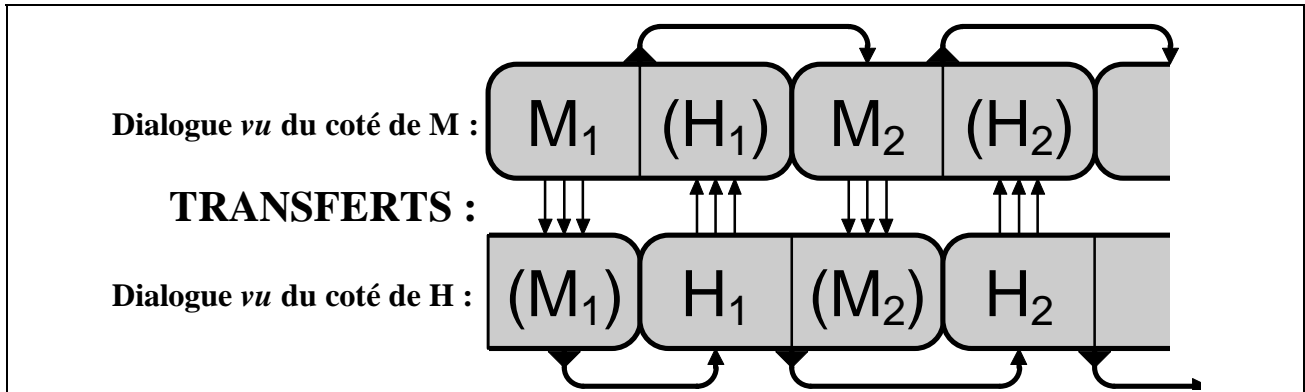


Figure 2: représentation des UMI

Les agents disposent de compétences différentes pour manipuler les UMI en fonction de leurs capacités. La structure et l'utilisation de ces UMI ne va pas être la même si on considère le point de vue des agents ayant un comportement réactif, rationnel ou cognitif. Dans le premier cas, l'agent fonctionne avec une boucle de comportement perception -> action et les deux parties de l'UMI se définissent relativement aux intervalles perception-action et action-perception. Dans les autres cas, l'agent fonctionne avec une boucle de comportement perception -> délibération -> action et les deux parties des UMI comportent un énoncé et des attentes :

- att_j et att_{H_i} sont les attentes provoquées respectivement par les actions M_j et H_j ,
- H_j et att_{H_i} sont les interprétations de la perception de l'agent considéré.

<i>UMI_i</i>	
M_i : Intervention de la machine { actes de dialogue } { objets du discours }	1
Att_{M_i} : Attentes de la machine { actes de dialogue } { objets du discours }	
H_i : Intervention de l'utilisateur { actes de dialogue } { objets du discours }	2
Att_{H_i} : Attentes de l'utilisateur { actes de dialogue } { objets du discours }	

figure 3 : une UMI pour un agent rationnel ou cognitif.

L'interaction est donc un objet complexe qu'il est possible d'aborder soit en observant celle-ci de l'extérieur, soit en participant à celle-ci de l'intérieur. Dans le cadre de l'approche systémique de l'interaction, [Delépine & al. 95] proposent trois niveaux indétachables d'analyse qui sont les niveaux physique, logique et sémiologique. Le niveau physique traite de la transmission de signaux, le niveau logique de l'alternance des positions de chacun des pôles et le niveau sémiologique d'une représentation psychologique. L'importance de ce dernier niveau, que Vernant appelle dialogal [Vernant 94] varie en fonction de la capacité de l'agent à se représenter le monde et les autres agents. Pour les agents rationnels, on peut ramener ce niveau à un niveau logique car le langage est réifié (voir paragraphe suivant). Pour des agents cognitifs, c'est à partir de ces trois niveaux irréductibles que les UMI pourront être manipulées et permettre à l'interaction de se mettre en place. Ces interactions permettront alors le transfert d'informations dans un cadre de diffusion, de coordination, de collaboration ou encore de coopération entre agents humains et machines, qui permettront de leur définir des rôles sociaux. Le passage du collectif au social est là, dans la différence entre les langages d'interaction, purement logiques ou pleinement sémiologiques, et dans la richesse des fonctionnements collectifs que cette interaction permet. Il faut donc maintenant préciser ce que nous entendons par système de signes et montrer les différences de degré de complexité qu'on peut y trouver.

5- Les systèmes de signes (et en particulier le langage) comme fondateurs des sociétés

Une organisation de SMA émerge à partir des spécificités des agents et de la dynamique de l'interaction dans les transferts, « l'organisation désigne à la fois le processus d'élaboration d'une structure et le résultat même de ce processus » [Ferber 95, p98]. Pour caractériser les différentes formes d'organisation par le langage d'interaction, il faut revenir à la distinction entre les différentes sortes de signes proposée par Peirce selon la seconde tricotomie [Peirce 78] et l'appliquer aux interactions entre agents humains ou machines :

- les indices ou les traces renvoient à l'objet qu'ils dénotent parce qu'ils sont directement affectés par cet objet, comme la fumée est indice du feu, ou les traces de phéromone indices du passage des fourmis. Ils permettent l'interaction dans les sociétés animales via des processus de type stimulus-réponse. Les agents réactifs utilisent ce mode d'interaction.

- les icônes renvoient à l'objet qu'elles dénotent en vertu de ressemblances entre le signe et l'objet. Elles peuvent donc être utilisés comme signes sans apprentissage et sans co-référentiation préalable en vertu de leurs propriétés intrinsèques. Elles sont donc bien adaptées aux interactions entre les humains et les machines quand il s'agit de tâches simples, comme les commandes d'un système d'exploitation via une interface graphique (bureau du MacIntosh, Window 95...)

- les symboles renvoient à l'objet qu'ils dénotent en vertu de conventions sociales. Ils nécessitent au préalable un accord entre les parties pour être interprétés utilement. Tout comme les langues naturelles, les langages de programmation ou de commandes de machines sont des systèmes de signes symboliques. Mais les seconds sont réifiés alors que les premières ne le sont pas.

Un langage réifié est un langage dont on a défini de façon précise le lexique, la syntaxe et la sémantique. L'interprétation des énoncés d'un langage réifié n'est pas l'objet de négociation car elle n'offre pas de jeu, et le langage ne peut pas évoluer de par son usage, mais seulement par une négociation extérieure, au sein des organismes de normalisation. On ne peut pas fonder une *société* d'agents sur un langage de ce type car la réification du langage empêche l'apprentissage des agents et l'évolution de la société. Ce fut un grand progrès pour l'informatique que la réification des langages de programmation évolués à partir des développements de la théorie des langages, à cause de la sûreté de fonctionnement et de l'indépendance au matériel qu'elle a permis. Il y a eu d'abord

réification de la syntaxe à la fin des années 60 avec Algol, puis réification de la sémantique à la fin des années 70. Il n'est pas question de revenir sur ces acquis mais de construire langage d'interaction entre agents au niveau conceptuel, au dessus du niveau symbolique géré par les langages de programmation, l'indépendance entre ces niveaux ayant été établie par [Newell 82]. Cette préoccupation n'est pas partagée, car l'objectif actuel des travaux autour de KQML, standard de communication de haut niveau initialisé par la DARPA aux États Unis, est de réifier le langage de communication entre les agents. Ils proposent une utilisation curieuse de la théorie des actes de langage, qui encapsule les actes de langages dans les énoncés [Ferber 95 p. 358] et donc rend inutiles les processus de co-référentiation. Mais c'est à travers ces processus que l'apprentissage est possible et que le langage peut évoluer [Vivier 93] et avec lui la société. La normalisation est la forme ultime de la réification par toutes les contraintes qu'elle pose sur l'évolution, celle de KQML peut être utile pour des agents rationnels, mais elle n'améliorera pas la communication homme machine.

Les langues naturelles au contraire, malgré les tentatives des grammairiens et du système scolaires, ne sont pas réifiées. Leur apprentissage est progressif et utilise la langue elle-même, et la métaphore est constitutive de leur emploi [Lakoff & Johnson 80]. Il n'y a nul besoin de réifier un langage pour se comprendre sinon comment les enfants apprendraient-ils à parler ? Mais comment apprendre à parler aux machines alors qu'elles n'ont pas de désir propre ? Il faut construire une amorce de machine dont le comportement soit analogue en définissant un modèle de l'intention pour une machine [Andrès 95] relativement aux états de la mémoire. Si nous revenons à l'analyse que Peirce fait des systèmes de signes, nous pouvons remarquer qu'il ne définit pas le signe comme clos : « [En bref, un signe est] tout ce qui détermine quelque chose d'autre (son interprétant) à renvoyer à un objet auquel lui-même renvoie (son objet) de la même manière, l'interprétant devenant à son tour signe et ainsi de suite ad infinitum. » [Peirce 78, p124]. Il ne définit pas l'interprétant comme un humain et prend bien soin de donner des exemples où il ne l'est pas, ce qui est très encourageant pour développer le rôle d'interprétant d'une machine.

Chaque sujet peut réaliser des opérations de catégorisation pour son propre compte, du fait qu'il est un être vivant qui cherche à maintenir sa viabilité. La mise en commun des résultats de cette catégorisation passe par les systèmes de signes, à travers les opérations de co-référenciation nécessaires au dialogue. Ils vont constituer la connaissance commune, construire des représentations partagées dans l'inter subjectivité des interprétants. La connaissance commune n'est donc pas donnée à priori et ne constitue pas un préalable à l'installation d'un dialogue, elle vient ensuite par les effets de la co-référentiation. Elle s'est constituée petit à petit dans l'histoire de l'humanité et dans l'histoire de chaque homme, de résultats partagés sur les catégories - les modèles - mais aussi d'une argumentation permettant à chacun de se les approprier par transfert, et pas avec des arguments d'autorité. Le but psycholinguistique de Compèrobot est d'utiliser le dialogue pour amener les enfants à paraphraser et permet donc justement d'étudier ce processus de négociation - sur la désignation des objets, de leur position, et de leur localisation par rapport à d'autres objets - qui construit la connaissance commune [Vivier 93]. La question suivante est de savoir si des agents logiciels peuvent participer à ce processus et quels seraient pour les humains les avantages à interagir ainsi avec les machines.

6- La conception d'un agent cognitif.

Tout d'abord, il s'agit de mettre en place et d'étudier des situations où les agents ont une activité en rapport avec des choses et des événements réels et pas avec des modèles pré construits des choses et des événements. En effet « Le signe ne peut que représenter l'objet et en dire quelque chose. Il ne peut ni faire connaître ni reconnaître l'objet; car c'est ce que veut dire dans le présent volume objet d'un signe; à savoir ce dont la connaissance est présupposée pour pouvoir communiquer des informations supplémentaires le concernant » [Peirce 78, p. 123]. Nous nous

plaçons donc dans le courant de la cognition située, et nous cherchons à faire interagir les agents à partir de leurs perceptions, de leur mémoire et de leurs actions.

Pour concevoir une machine autonome qui puisse dialoguer en langue naturelle il faut installer une boucle infinie *perception - délibération - action*.. Cette boucle est classique dans son principe [Peireira et Grosz 93], elle est procédurale et s'applique à la situation de référence. Dans le modèle proposé pour Compèrobot [Nicolle 93] elle s'applique à trois niveaux : à la situation de référence, au dialogue lui-même, et au contrôle du processus tout entier. Pour le niveau de contrôle elle doit pouvoir s'observer elle-même et donc avoir une double représentation : déclarative et procédurale. Un modèle réflexif, qui se décrit lui-même, est utilisé. On appelle réification la représentation déclarative de la boucle - qui est par essence un processus - et cette représentation déclarative est en relation de dépendance causale [Maes 87] avec la représentation procédurale, ce qui permet l'auto-modification. Il faut pour cela décomposer le processus concret en deux parties : un processus abstrait, très général, de nature procédurale, et une représentation déclarative du processus concret. Le processus abstrait va fonctionner de manière inconsciente en interprétant la représentation du processus concret. Toute modification de cette représentation va alors modifier le processus qui s'exécute.

Cette double nature déclarative et procédurale de la boucle d'interaction est à mettre en relation avec la capacité d'un sujet à se construire des représentations du monde perçu. L'activité de perception est d'abord une activité individuelle et inconsciente de catégorisation des sensations [Dubois 91], en retour aux activités sensorielles et motrices. Le langage oblige à réifier cette catégorisation pour la mettre en mots, il permet alors de se mettre d'accord sur la catégorisation, donc de la préciser et de l'argumenter [Jullien 95]. Par le transfert des expériences d'autrui via la parole, la catégorisation faite par chaque individu s'améliore car elle peut tenir compte de davantage d'arguments, fondés sur les expériences des autres. Entrer dans un dialogue sémiotique suppose seulement qu'on ait la capacité de manifester son accord et son désaccord et des capacités d'argumenter, fondées sur ce que Trognon appelle « la psychologie populaire » c'est à dire la représentation d'autrui à l'image de soi. Le sujet qui dialogue construit sa propre image à travers l'image que les autres lui renvoient, et la réflexion est alors possible. Les représentations, en étant dites, prennent leur autonomie par rapport au réel perçu. Les jeux avec le langage mettent en place la réflexion et l'imagination vient jouer un rôle essentiel dans l'activité mentale et langagière.

7. Les premiers résultats

Cet article propose une synthèse entre plusieurs projets de recherche sur le dialogue homme-machine, sur les systèmes réflexifs, et sur les systèmes multi-agents. Il est donc principalement prospectif. Il propose un modèle d'agent cognitif et réflexif, communiquant en langue naturelle, qui n'est pas implanté actuellement. Toutefois, il s'appuie sur plusieurs réalisations :

- le langage réflexif Airelle utilisé à Caen pour enseigner l'intelligence artificielle en maîtrise, DEA, à ENSI et qui a servi et sert de langage de développement pour plusieurs thèses.
- les expériences de Compèrobot [Nicolle & Vivier, à paraître].
- le système de simulation multi-agents réflexif proposé par Benoit Durand [Durand & Nicolle 95] qui est actuellement opérationnel. Il est construit sur la base d'une boucle réflexive avec une double représentation, déclarative et procédurale, comme ce qui est proposé ici.
- le logiciel de dialogue de Jérôme Lehuen [Lehuen & al. 96] qui commence à interagir et à apprendre à partir de ses interactions, en utilisant les UMI présentées ici. Il devrait être bientôt expérimenté avec des utilisateurs de la médiathèque du Mans.

Références

- Brassac Christian (1994). Modéliser l'enchaînement conversationnel. *Actes des journées de Rochebrune : Autonomie et interactions fonctionnelles*.
- Delépine Ludovic, Nicolle Anne et Coursil, Jacques (1995). Transfer Frameworks for modelling Human-Machine Dialogue. *International Symposium "Language and Consciousness", Varna, Bulgaria, 12-15 September*.
- Durand Benoit et Nicolle Anne (1995). Un modèle de systèmes multi-agents réactifs pour la représentation de situations et de processus. *Journées francophones IAD-SMA*.
- Ferber Jacques (1995). Les systèmes multi-agents. Paris : InterÉditions.
- Jullien Yannick (1995). Communication expérimentateur-machine sur une expérience psychologique. *Actes 01'Design 95, Autrans*.
- Lakoff George et Johnson Mark (1980). Les métaphores dans la vie quotidienne. Éditions de Minuit 1985 pour la traduction française.
- Lehuen Jérôme, Nicolle Anne et Luzzati Daniel (1996). Un modèle hypothético-expérimental dynamique pour la gestion des dialogues homme-machine. *Actes RFIA'96*.
- Luzzati Daniel (1995). Le dialogue verbal homme-machine. Paris : Masson
- Maes et Nardi Eds (1987). Meta-level architectures and reflection. Amsterdam : North-Holland.
- Newell Allen (1982). The knowledge level, *Artificial intelligence*, 18, 1, p. 87-127.
- Nicolle Anne (1993). Towards a natural language dialog with machines. *Actes du Deuxième congrès européen de Systémique, AFCET-CSCI, Prague*.
- Nicolle Anne (1994). Conception du logiciel par objets : réification, réflexivité et amorce, *Actes Journée d'étude du GRASCE, « De l'Objet au Système », Aix En Provence*.
- Nicolle Anne, Bricon-Souf Nathalie (1994). Machines Et Échanges Langagiers Dans Le Projet Compèrobot. *Actes du colloque "Le Dialogique", colloque international sur les formes philosophiques, littéraires, linguistiques et cognitives du dialogue.*, Le Mans.
- Nicolle Anne et Vivier Jean, Questions de méthode en dialogue homme-machine, l'expérience Compèrobot. À paraître.
- Moeshler Jacques et Reboul Anne (1994). Dictionnaire encyclopédique de pragmatique. Paris : Seuil.
- Peirce Charles S. (1978). Écrits sur le signe rassemblés, traduits et commentés par Gérard Deledalle. Paris : Seuil.
- Peireira F. C. N. et Grosz B. J. (1993). Introduction. *Artificial Intelligence special volume Natural Language Processing*, vol. 63 n° 1-2, Octobre 93, New York : American Elsevier.
- Roulet Eddy (1985). L'articulation du discours en français contemporain. Berne : Peter Lang.
- Simon Herbert A. (1983). Search and reasoning in Problem Solving. *Artificial Intelligence* n° 21
- Vernant Denis (1994). Dialectique, Forme dialogale et dialogique. *Actes du colloque "Le Dialogique", colloque international sur les formes philosophiques, littéraires, linguistiques et cognitives du dialogue*. Le Mans.
- Vivier Jean (1993). Comperobot, when dialog enlarges capacity. *Actes du Deuxième Congrès Européen de Systémique, AFCET-CSCI, Prague*.

Du collectif au social: vers une distinction fonctionnelle quantitative

M. Courant et A. Robert

Institut d'Informatique de l'Université de Fribourg

Chemin du musée 3, CH-1700 Fribourg

E-mail: {michele.courant,antony.robert}@unifr.ch

Résumé

Partant des mesures de complexité existantes, l'article vise à dégager deux aspects duaux de la complexité: complexité structurelle et fonctionnelle. Cette vision nous permet alors d'élaborer une théorie de l'évolution basée sur deux principes antagonistes, à même de supporter un principe d'économie gouvernant l'évolution des systèmes naturels. Ce cadre nous permet de fonder une distinction entre collectif et social. En effet, sur le plan structurel, on retrouve le lien entre les deux termes. Sur le plan fonctionnel en revanche, le collectif manque de signification, alors que le social implique une orientation coordonnée de l'activité des individus.

mots clés: complexité, mesure, évolution, antagonisme.

1. Introduction

Le thème choisi pour les journées, à savoir “du collectif au social” traduit intuitivement une double idée. D'une part il évoque une différence de nature entre des systèmes susceptibles d'être qualifiés de collectifs et de sociaux: il appelle donc la recherche de critères de distinction entre eux. D'autre part il suggère qu'il puisse exister une transition depuis le stade du collectif vers le stade du social. Autrement dit, il tend vers une vision selon laquelle un système social peut être considéré comme une émanation d'un système collectif, produite par l'évolution. Précisons qu'outre le fait de souscrire à cette vision, nous proposons même de l'élargir, pour conférer un statut systémique à la distinction entre collectif et social. Ce faisant, nous lui accordons une opérationnalité transversale, englobant les systèmes physiques et biologiques (allant des colonies de cellules et organismes multicellulaires aux systèmes sociaux proprement dits) [23].

Sur cette base, notre contribution au thème réside en la recherche d'une caractérisation formelle de la distinction entre les stades du collectif et du social. Essentiellement, nous proposons de l'aborder via la notion de complexité, notion que nous chargeons ainsi de présider à l'élaboration d'une théorie générale de l'évolution. Cet objectif, très ambitieux, est en fait supporté par une idée intuitive de distinction entre des types de complexité et de son implication dans un principe d'économie s'exprimant dans l'évolution.

La suite de l'article s'organise en trois sections. Nous procédons tout d'abord à un bref tour d'horizon du champ de la complexité. Nous amorçons ensuite une réflexion critique sur la base des mesures de complexité introduites, tandis que la conclusion recentre l'étude sur le thème des journées. Globalement, notre contribution se conçoit essentiellement comme l'ouverture d'un débat sur la portée et la pertinence d'une approche quantitative des systèmes collectifs et sociaux en particulier, et de l'organisation du vivant en général.

2. Principe d'économie

Nous évoquons ici brièvement l'intuition qui sous-tend notre étude de la complexité. Celle-ci s'exprime grossièrement comme un principe d'économie qui semble gouverner les systèmes naturels.

Si l'on observe les structures de la nature, on s'aperçoit en effet que les domaines bien compris par la science le sont grâce à une hypothèse de simplicité. Du reste, bon nombre de théories scientifiques, qui approchent tout de même assez bien notre expérience, révèlent une grande régularité depuis les structures d'atomes jusqu'au mouvement des galaxies. En partant de quelques lois et quelques objets atomiques, il se forme de nouveaux objets de base régis par les mêmes lois, qui vont encore se regrouper en objets constituants d'un autre niveau [16].

La structure du vivant, elle, de par la richesse de ses composants et leurs inextricables relations, paraît parfois échapper à la logique simple de ces structures. Elle en viendrait même à se caractériser, surtout à ses stades les plus évolués tels qu'à l'échelle humaine où domine l'appréhension cognitive et sociale des phénomènes, par son affranchissement vis-à-vis des lois physiques. Pourtant, on peut se demander si de manière sous-jacente il n'existerait pas dans le vivant un principe d'économie de moyens identique à celui régissant les autres types de structures physiques, mais dont l'application reste difficile dans ce cas.

Le biologique en suggère en tout cas l'intuition. Si l'on n'entre pas dans les détails moléculaires de certaines plantes, par exemple, on parvient très bien à décrire leurs structures en utilisant des constructions fractales self-similaires. L'interprétation d'E. Ben-Jacob [4], en termes d'optimisation du rendement énergétique des formes fractales obtenues par la culture de colonies de bactéries confère même un rôle actif au principe d'économie: les formes fractales émergeraient d'un couplage entre diffusion induite par la nourriture et rappel exercé par les forces de surface.

Il convient également de souligner la grande variété de structures obtenues par l'application récurrente de règles simples dans les automates cellulaires et dans l'étude de la criticalité auto-organisée [3]; et ces systèmes possèdent des comportements très difficilement descriptibles dans certaines conditions. Il semble donc que le goût de la nature pour la récursivité soit conservé même dans les formes vivantes: et quel autre mécanisme peut se targuer d'une plus grande économie de moyens?

Par ailleurs, l'intuition d'un principe d'économie se trouve également confortée par les émergences du biologique que sont le cognitif et le social. Dans le cadre de la linguistique du discours, les maximes de Grice [12] et la théorie de la pertinence en général, sont l'expression assez directe d'une économie de moyens au service de fins. De même certains traits du social, qu'ils soient structurels, comme la spécialisation, ou fonctionnels, comme la coopération, semblent aussi indiquer un principe d'économie, que celui-ci soit causal ou émergent. Enfin, quelle autre institution que l'économie de marché qui, par l'actualisation de la conceptualisation du monde qu'elle propose, l'érige en loi sociale, manifeste mieux l'hypothèse d'un principe d'économie ?

3. Différentes mesures de complexité

Comme en témoignent les multiples instituts de la complexité récemment nés de par le monde, l'espoir de trouver une mesure universelle de complexité, qui capte notamment l'essence de l'organisation et la dynamique du vivant, anime une recherche très active. Cependant ce problème n'a pu à ce jour cristalliser de consensus, et l'on doit pour l'instant se contenter de différentes mesures dont le champ d'application reste visiblement limité. Nous présentons maintenant quelques unes des mesures de complexité proposées, en indiquant brièvement leur contexte de développement, ainsi que les hypothèses s'y rattachant (pour une étude plus complète, voir [9, 11, 18]). Nous examinons, tout d'abord deux mesures de type informationnelles, puis la notion de profondeur logique.

3.1. La quantité d'information

La théorie de la communication développée par Shannon [25] s'attache à l'étude des transmissions de messages sur un canal de communication qui relie un émetteur à un récepteur. Pour pouvoir définir la capacité d'un canal de communication, il convient de déterminer la taille minimale d'un message donné, celui-ci étant codé sur un alphabet fixé.

La formule de Shannon donne la quantité d'information contenue dans un message en termes de l'incertitude levée par celui-ci lorsqu'il arrive au récepteur. De manière formelle, le récepteur connaît a priori l'ensemble des messages possibles. La réception d'un message particulier lève donc l'incertitude sur cet ensemble et c'est cette incertitude qui est mesurée.

L'application de cette mesure se généralise aux systèmes physiques si l'on considère qu'un canal de communication existe entre un observateur mesurant les états d'un système physique et le système lui-même [7]. Elle conduit alors à une mesure de complexité d'un système définie comme la quantité maximale d'information que l'on peut extraire du système par des mesures sur ses états internes. Les recherches conduites par H. Atlan, pour tenter d'appréhender en termes quantitatifs l'organisation biologique, relèvent en particulier de cette approche [1].

3.2. L'information algorithmique

Une autre mesure informationnelle, qui considère également la description d'un système sous forme d'un message est la complexité dite de Kolmogorov ou Information Algorithmique [8, 13]. Cette mesure dérive de celle de Shannon, tout en l'affinant et la généralisant puissamment à la fois. L'approche consiste à considérer la description d'un système dans un langage donné, par exemple celui d'une machine de Turing universelle.

L'idée est de minimiser les redondances de description, sans perte d'information, pour juger de la complexité d'un système en termes de longueur de sa description. Pour appréhender le phénomène des redondances, la mesure se base sur la réutilisation de descriptions partielles du système, en simulant en quelque sorte d'autres machines pour lesquelles ces descriptions feraient partie du langage.

Un reproche qui a été adressé à cette mesure est sa non-unicité dans le cas de descriptions de longueurs finies. Si l'on se donne comme langage de description celui d'une machine de Turing universelle, la longueur d'une description sur une autre machine de Turing universelle sera en effet généralement différente. En revanche, lorsque la longueur d'une description devient grande, les différences s'atténuent. En pratique, cette mesure permet l'élaboration d'algorithmes de compression de données tels que l'algorithme de Lempel-Ziv [17]. Globalement, on peut dire que cette mesure est fondée sur les probabilités d'apparition de séquences de lettres dans la description. Elle est par conséquent une mesure de l'aléatoire et l'aléatoire seul n'est pas nécessairement suffisant pour cerner l'idée intuitive de la complexité.

3.3. La profondeur logique

Une mesure relativement bien adaptée à l'étude de systèmes dynamiques est la profondeur logique de Bennett [5]. Celle-ci aborde les systèmes auto-organisés sous un angle génératif. La profondeur logique se définit comme le temps d'exécution du programme de longueur minimale donnant la description du système (i.e. l'algorithme optimal pour la mesure d'information algorithmique).

L'apport essentiel de cette mesure est que tout système est désormais muni d'une origine. La profondeur logique estime le temps de reconstruction du système à partir de cette origine, en partant du principe que l'origine la plus probable est la plus courte en termes de longueur de description. Cependant, il s'agit d'une mesure strictement dépendante de la mesure d'information algorithmique. Elle ne s'oriente donc même pas vers l'idée d'une compensation possible entre algorithme et structures de données.

4. Complexité: structure et fonctionnalité

4.1. Le problème de la description

Les diverses mesures existantes reflètent quelques unes des caractéristiques essentielles évoquées par la notion intuitive de complexité et signalent les principaux obstacles à l'établissement d'une mesure universelle. Elles indiquent premièrement que toute mesure de la complexité d'un système doit passer par sa description, que ce soit dans un langage algorithmique ou logique. Ce passage obligé par la description renvoie donc au choix d'un langage de référence, puis au degré d'universalité de celui-ci. L'approche employée suggère en fait de parler de complexité relative au moyen de description choisi, voire au modèle scientifique sous-jacent à ce moyen.

Les mesures précédentes utilisent des descriptions physiques généralement codées sous la forme de chaînes de bits. On voit donc que le choix du langage de référence entraîne également celui des éléments atomiques du système, implicitement attachés au formalisme de la description via son axiomatique. Ainsi le choix du langage décide de la granularité de la description et fixe l'échelle de l'observation. Par ailleurs, lorsqu'on traite d'un système dans son environnement et par rapport à l'observateur, le problème de la description s'achoppe au problème fondamental et bien connu en théorie générale des systèmes [6] de la délimitation fonctionnelle du système par rapport à cet environnement et l'observateur.

4.2. De la nécessité d'un couplage

Un aspect fondamental révélé par les différentes approches citées est la distinction première qu'il convient de faire entre une vue externe et une vue interne du système.

La vision interne tente de déterminer la complexité d'un système uniquement par une description de ses divers composants, ce qui est typiquement l'objet de la mesure de l'information algorithmique. L'élimination des redondances en est caractéristique.

De son côté la vision externe inclut la dynamique de fonctionnement du système et prend en compte son aspect situé. La profondeur logique rend quelque peu compte de ce point de vue. En fait, L. Löfgren [19] avait fait cette observation en associant la complexité aux processus de description et d'interprétation d'un système, où l'interprétation et la description sont les processus de traduction de la description vers le système et vice-versa. Il associait alors l'information algorithmique à la complexité descriptive et une forme de profondeur logique à l'interprétation.

Nous pensons que la dualité proposée par Löfgren est fondamentale et qu'elle permet l'expression de finesses inabordables par une mesure unique. Poursuivant un but applicatif, orienté d'un côté vers l'interprétation du darwinisme et de l'autre vers la construction de systèmes artificiels évolutifs, notamment robotiques, nous préférons toutefois la reformuler en termes de complexité structurelle et complexité fonctionnelle. A un niveau de description donné, la structure traite du côté statique du système en décrivant les éléments qui le composent et leur ordonnancement, alors que la fonctionnalité explicite le côté dynamique du système en décrivant les interactions entre les éléments et avec l'environnement, voire l'observateur.

4.3. Structure et fonctionnalité, propriétés

Pour expliquer les propriétés que l'on rattache aux termes de structure et de fonctionnalité, il est utile de prendre des exemples.

La fonctionnalité apparaît essentiellement comme une manifestation coordonnée des lois régissant le système et son environnement. Prenons par exemple un morceau de fer: pour une description structurelle opérant au niveau des ions ferreux, la complexité structurelle est fixée. Toutefois, le fer peut être aimanté ou non, selon que la polarisation des molécules constitutives est globalement dirigée dans

un sens ou que les polarisations s'annulent mutuellement. On dit que la fonctionnalité d'un aimant est plus grande que celle d'un morceau de fer non magnétisé. En général, on s'intéresse simultanément à plusieurs fonctionnalités et c'est là qu'intervient la notion de complexité fonctionnelle. On peut alors avoir différentes sortes de propriétés associées à cette caractérisation de la fonctionnalité, telles que la propriété d'additivité de champs. Les parties d'un système peuvent donc agir de manière coordonnée, auquel cas elles tendent à générer de la fonctionnalité, ou désordonnée, auquel cas leurs fonctionnalités tendent à s'annuler.

Au contraire, la réutilisation d'une structure dans un système, comme cela s'obtient par division cellulaire dans un organisme multicellulaire, amène à une simplification structurelle du système. Ceci est dû au fait que l'on désire une description minimale unique du système en question, indépendamment du langage de description utilisé. La mesure de complexité structurelle évalue ainsi la taille de la description exempte de redondances dans un langage, ce qui est le cas limite de la mesure de l'information algorithmique pour des messages de longueur infinie.

Pour justifier les deux mesures duales suggérées, il faut observer les paradoxes encore mal compris d'une nature allant toujours vers une entropie plus élevée, mais qui admet localement l'existence et la conservation de sous-systèmes à basse entropie et qui progresse par transitions brutales. Diverses tentatives ont été faites pour expliquer ces phénomènes, que ce soit par la théorie des systèmes dissipatifs et du troisième état [2, 16, 22], par les systèmes chaotiques [10] ou la théorie des catastrophes [24]. Dans la mesure où notre objectif n'est pas spécifiquement d'expliquer la genèse de la dualité, par exemple par l'enracinement causal d'une complexité dans l'autre, mais plutôt d'en étudier l'incidence sur l'évolution, nous préconisons quant à nous une démarche axiomatique. L'idée est de postuler l'existence d'un antagonisme fondamental des lois gouvernant la complexification des structures et l'augmentation de fonctionnalité. Ce faisant, nous rejoignons fortement l'antagonisme de processus homogénéisant et hétérogénéisant de S. Lupasco [20], ainsi que l'idée de projet du vivant acceptée par J. Monod [21]. Nous nous reposons donc sur l'existence de deux principes, de complexification des structures et d'augmentation de la fonctionnalité. Le premier coïncide dans la nature avec le second principe de la thermodynamique, via la dérive entropique. Une croissance locale de fonctionnalité s'y opposant, mène en particulier au darwinisme. Toute structure étant "porteuse" de fonctionnalité, la recherche d'un équilibre antagoniste passe soit par une optimisation de la structure se rapportant à une fonctionnalité définie, par exemple au travers de hiérarchisations, soit par une optimisation de la fonctionnalité pour une complexité structurelle donnée, par exemple par la découverte de nouvelles configurations structurelles. Ceci ramène aux principes élémentaires de sélection et mutation. On en revient ainsi à un principe d'économie tel qu'à complexités fonctionnelles égales, les systèmes les plus simples structurellement (et les plus probables par conséquent) sont conservés. Dans cette optique, le vivant est une tentative d'optimisation entre la fonctionnalité maximale pour une structure la plus simple possible, ce que Langton [14, 15] voulait sans doute exprimer par le titre *Life at the edge of Chaos*.

4.4. Deux exemples

L'ambiguïté de la notion de complexité d'un système apparaît clairement lorsqu'on se réfère aux exemples habituellement choisis pour en former une intuition. Considérons d'abord la classification par ordre de complexité croissante des motifs donnés sur la figure 1. Nous prétendons que l'approche utilisée en général consiste à confondre dans la procédure de classification les deux aspects duaux de la complexité. Usuellement l'image de gauche est immédiatement qualifiée de simple puisque sa structure est simple. La vraie question consiste en fait à évaluer la complexité des deux images de droite. Le critère de comparaison devient alors celui de la complexité fonctionnelle. On observe que le motif de droite a probablement été généré aléatoirement, alors que celui du milieu rappelle clairement le portrait de phase d'un système dynamique (chaotique?). Dans la mesure où la classe des motifs aléatoires est facilement décrite en termes de bruit, la figure du centre s'affirme finalement comme la plus

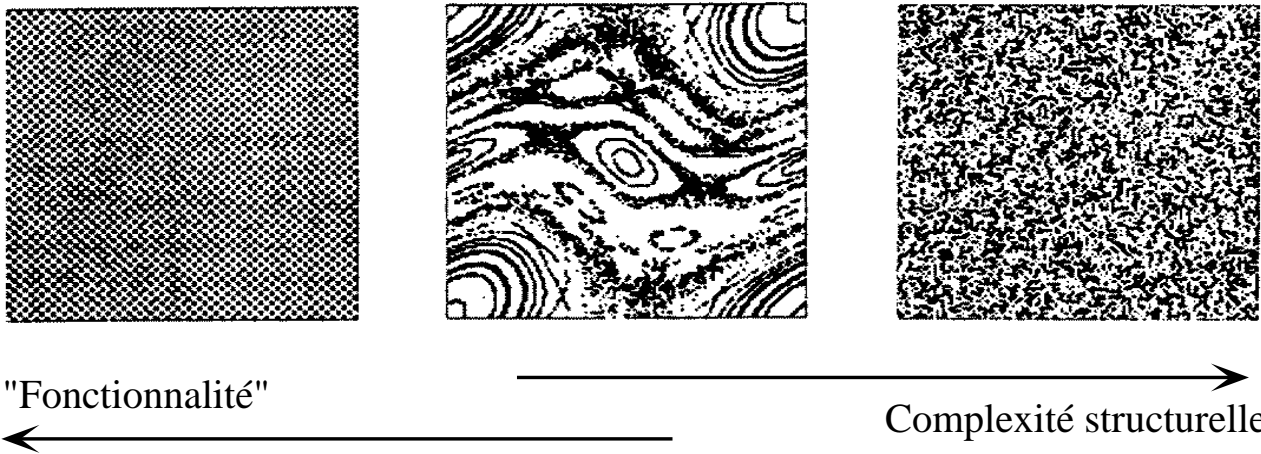


Figure 1: Trois motifs complexes?

complexe des trois. La complexité fonctionnelle ne pouvant en principe pas être déduite d'une image statique, l'exemple exige de recourir au critère de la complexité structurelle. Dès lors que le niveau de description est le pixel, la croissance en complexité s'observe ainsi de gauche à droite. En revanche, si l'on cherchait à donner un sens aux images, i.e. à les classer d'un point de vue fonctionnel, on pourrait les traiter comme des portraits de phase, d'où une complexité fonctionnelle maximale au centre, mais il serait tout aussi justifié de les interpréter comme des particules animées d'un mouvement imaginaire, ce qui donnerait une fonctionnalité croissante de droite à gauche comme indiqué par la flèche.

Un test moins évident et qui n'est que rarement proposé est la comparaison de systèmes d'échelles très différentes. Soit la comparaison du système atmosphérique terrestre avec un organisme vivant tel que l'homme: sans émettre d'hypothèses supplémentaires, il devient tout à fait difficile d'ordonner les exemples selon leur complexité. Dans le contexte que nous proposons, après avoir fixé comme niveau de description le niveau moléculaire, on constate que les particules atmosphériques n'entretiennent pas de relation structurelle régulière entre elles, alors que l'organisation biologique est faite de redondances et que le nombre de composants du corps humain est nettement inférieur à celui de l'atmosphère, d'où une complexité structurelle plus élevée pour l'atmosphère terrestre. En revanche, les compétences d'un être humain sont d'une telle diversité que même les propriétés chaotiques du comportement de l'atmosphère semblent d'une complexité fonctionnelle inférieure. L'ambiguïté du cas est donc symptomatique de la dualité inhérente à la notion de complexité.

5. Conclusion

Sur la base de différentes mesures de complexité déjà en vigueur, nous avons premièrement distingué deux types de complexités, respectivement structurelle et fonctionnelle, et posé comme une axiomatique destinée à présider à l'élaboration d'une théorie de l'évolution utilisant ces mesures, un antagonisme fondamental de lois réalisant leur couplage. Selon cette approche, on retrouve au niveau des structures le lien entre collectif et social, les deux termes caractérisant des systèmes formés d'agents dotés d'une relative autonomie. Le plan de la fonctionnalité montre en revanche que le collectif manque de signification, alors que le social implique une orientation coordonnée de l'activité des individus. L'articulation des deux mesures de complexité permet de considérer le social comme la marque d'une étape de transition vers un optimum de fonctionnalité pour une structure établie. Cette transition s'observe par la spécialisation progressive des éléments du système, qui perdent progressivement leur complexité fonctionnelle individuelle en faveur d'une augmentation de la fonctionnalité du tout.

Le collectif et le social ont déjà fait l'objet de diverses distinctions, plus ou moins explicites. En opposition à certaines démarches qualitatives ou informelles, le propos de l'article a été d'envisager si

et comment une approche formelle quantitative pouvait contribuer à l'affinement d'une telle distinction, et d'inviter au débat sur la question.

En ce qui nous concerne, même si toute mesure est limitée par construction, et si, vu l'enjeu des mesures ici proposées, il est d'autant plus important de ne pas tomber dans le piège d'interprétations hâtives conduisant à des aberrations sur le plan éthique, nous aimerions défendre l'idée qu'une approche quantitative de l'évolution et de l'organisation du vivant est aujourd'hui souhaitable.

Références

- [1] H. Atlan, *L'organisation biologique et la théorie de l'information*. Hermann, 1972.
- [2] H. Atlan, *Entre le cristal et la fumée, essai sur l'organisation du vivant*. Seuil, 1979.
- [3] P. Bak, C. Tang, K. Wiesenfeld, *Self-organized Criticality*. Physical Review A, 38, pp 364-374, 1988.
- [4] E. Ben-Jacob, I. Cohen, A Czirók, *Smart Bacterial Colonies.*, Physics of Biological Systems: From Molecules to Species, Nordita 14-27, August 1995.
- [5] C. Bennett, *Dissipation, information, computational complexity and the definition of organisation*. Emerging syntheses in science, Pines Ed., Addison-Wesley, Redwood City, CA.
- [6] L. von Bertalanffy, *Théorie générale des systèmes*. Dunod, Paris, 1993.
- [7] L. Brillouin, *Science and Information Theory*. Academic Press, New York, 1962.
- [8] G.J. Chaitin, *Algorithmic Information Theory*. rev. 3rd printing Cambridge Univ. Press.
- [9] B. Edmonds, *What is complexity? The philosophy of complexity per se with application to some examples in evolution*. <http://alphard.cpm.aca.mmu.ac.uk/>
- [10] J. Gleick, *La théorie du chaos, vers une nouvelle science*. Albin Michel.
- [11] P. Grassberger, *Problems in Quantifying Self-generated Complexity*. Helvetica Physica Acta, 62, pp.489-511, 1989.
- [12] H.P. Grice, *Logic and conversation*. Mimeo, 1968.
- [13] A.N. Kolmogorov, *Three Approches to the Quantitative Definition of Information*. Probl. of Information Theory 1, 3 ,1965
- [14] C.G. Langton, *Life at the Edge of Chaos*. Artificial Life II, SFI studies in science of complexity, 10, Addison-Wesley, 1991.
- [15] C.G. Langton, *Computation at the Edge of Chaos*. Emergent Computation, S. Forrest Ed., MIT/North-Holland, 1991.
- [16] E. Laszlo, *La cohérence du réel: évolution, coeur du savoir*. Bordas, Paris 1989.
- [17] J. Ziv, A. Lempel, IEEE transactions on Information Theory, 24, 530, 1978.
- [18] S. Lloyd, H. Pagels, *Complexity as Thermodynamic Depth*. Annals of Physics 188, pp.186-213, Academic Press, 1988.

- [19] L. Löfgren, *Complexity of Descriptions of Systems: a Foundational Study*. Int. J. General Systems 3, 197-214, 1974.
- [20] S. Lupasco, *L'énergie et la matière vivante*. Le Rocher, 1987.
- [21] J. Monod, *Le hasard et la nécessité*. Editions du Seuil, Paris, 1970.
- [22] I. Prigogine, I. Stengers, *La nouvelle alliance: métamorphose de la science*. Editions folio essais, Gallimard, 1986.
- [23] E. Schwarz, *A Metamodel to Interpret the Emergence, Evolution and Functioning of Viable Natural Systems*. R. Trapp Ed., Proceedings of the 12th European meeting on cybernetics and systems research 94, World Scientific, Singapore, pp.1579-1586, 1994.
- [24] R. Thom, *Stabilité structurelle et morphogénèse*. Interditions, 1977.
- [25] W. Weaver, C.E. Shannon, *Théorie mathématique de la communication*. Retz - C.E.P.L., Paris, 1975.

Vers une théorie écologique de l'amour

(Essai sur une société permettant l'autodétermination à l'individu en matière d'amour et d'action dérivée)

Juan Gonzalez

C.R.E.A., Ecole Polytechnique
1, rue Descartes, 75005 Paris
E-mail: gonzalez@poly.polytechnique.fr

Résumé

Malgré ma conviction que l'on ne peut pas énoncer, au sujet de l'amour, des propositions du type de celles des sciences naturelles (Wittgenstein), cet essai se veut un effort pour mettre au jour quelques enjeux impliqués dans et par les actions motivées par ce que l'on nomme 'amour' dans une société occidentale, actuelle, et industrialisée. Dans un premier temps, un cadre est proposé pour cerner les enjeux: il s'agit d'un volet de la théorie de la motivation [égoïsme vs. altruisme] et d'un autre volet de l'éthique [déontologie vs. conséquentialisme]. Ensuite, j'offre une définition de l'amour, et je rappelle deux acceptions opérationnelles, bien que traditionnellement concurrentielles, de celui-ci : éros et agapè. Dans un troisième temps, je mets en évidence quelques traits du paradigme socio-cognitif installé 'par défaut' —qui structure et conditionne nos rapports sociaux— et je critique de manière succincte l'aliénation qui entraîne un tel paradigme aujourd'hui vis-à-vis des individus se voulant rationnels, libres et motivés par l'amour. Enfin, je propose un modèle alternatif qui a pour but de conceptualiser nos rapports sociaux de manière égoïste et néanmoins holiste et 'écologique', et d'ouvrir ainsi des possibilités à l'individu pour s'autodéterminer en matière d'amour et d'action dérivée au sein d'une société accueillante et fertile pour cela.

mots clés: amour (sentiments), éthique, action, société, théorie sociale, écologie, égoïsme.

"Moral character is assessed not by what a man knows but by what he loves"
St. Augustine, Epistulae

1. Introduction

Je compte aborder le thème de l'amour en tant qu'idée, affection ou sentiment reliant des individus, c'est-à-dire, en tant que notion de caractère social. J'évite donc, volontairement, d'aborder le thème dans son acception mystique ou religieuse, tout en soulignant que l'ombre de cette acception se trouve souvent inscrite, de par notre histoire, dans l'horizon d'utilisation de l'acception à caractère social. Dans ces pages, il s'agit de comprendre, ou au moins d'ébaucher une critique, de quelques enjeux impliqués par les notions d'amour et d'action dérivée, et d'évaluer la pertinence ou la viabilité de ces notions dans le contexte d'une société occidentale actuelle, industrialisée et urbaine. Ceci m'amènera à rejeter le modèle courant 'standard' (par défaut) d'application d'une telle notion et à proposer un autre modèle, compatible avec des principes libertaires d'autoorganisation sociale, qui se fonde sur un égoïsme 'positif' et dans lequel l'amour est un principe organique et constructif de développement collectif. Bien entendu, j'aborde cette question de manière arbitraire quant à sa thématization. En outre, je développe cette question moins comme une analyse soucieuse des formalismes et d'une méthodologie que comme une présentation de la genèse et du développement de la problématique telles que je les vois.

2. Egoïsme vs. Altruisme

Dans la mesure où le thème de l'amour fait intervenir des raisons qui fournissent des motifs et/ou des justifications pour agir et penser relativement à autrui, je crois approprié de faire un détour par la théorie de la motivation et par des considérations éthiques, ne serait-ce que pour nous accorder sur certains termes et pour poser une base à la discussion qui suivra.

De manière générale, on oppose l'altruisme à l'égoïsme. Ce dernier habituellement désigne un principe postulant que l'action humaine consciente est motivée par (ou que par) l'intérêt ou bénéfice de l'agent, dans une situation quelconque¹. Or, si l'on admet que les motifs ne sont pas des causes (et sont donc des états mentaux), on peut élargir ce principe de motivation au domaine de notre vie mentale et postuler que nos pensées (désirs, aversions, croyances, interprétations, etc.) —si elles sont conscientes— sont guidées ou mues par un intérêt égoïste. Cet intérêt ou bénéfice peut être direct ou indirect, immédiat ou futur, explicite ou implicite.

¹Savoir si ce principe s'applique à toute situation ou bien aux situations de concurrence ou de compétition seulement est discutable.

Les questions de savoir si toute action ou pensée motivée doit être forcément consciente, ou si le bénéfice auquel on se réfère lors de la justification d'une action ou d'une pensée est toujours le même qui nous a motivé effectivement, ou s'il y a un sens à distinguer entre l'action/pensée inconsciente et l'action/pensée non-consciente, ces questions s'avèrent difficiles et ouvertes au débat. Pour simplifier donc, disons que l'égoïsme se résume comme le principe suivant lequel toute action ou pensée consciente est motivée et justifiée par, ou que par, le bénéfice de l'agent lui-même dans une situation quelconque.

Par opposition, l'altruisme généralement désigne le principe suivant lequel toute action ou pensée motivée se justifie, d'après l'agent lui-même, par (ou que par) l'intérêt ou le bénéfice d'autrui dans une situation sociale quelconque; ce bénéfice pouvant être direct ou indirect, immédiat ou futur, mais attention, uniquement explicite (voir plus bas).

Il est important de remarquer trois points dans cette opposition: 1) la complexité supplémentaire de la position altruiste, du fait que —selon cette thèse— toute motivation et justification impliquent (des états mentaux) d'autres personnes, tandis que dans la position égoïste ce n'est pas forcément le cas; 2) la version forte de chaque position: celle qu'introduit le terme d'exclusivité 'que par'; et 3) l'asymétrie entre l'égoïsme et l'altruisme quant à leur condition de conscience: on peut être égoïste de manière explicite ou implicite, alors qu'on ne peut être altruiste que de manière explicite.

Le dernier point pose suffisamment de problèmes pour nous amener à l'examiner de plus près. Voyons. D'un côté, ce n'est pas pareil de *savoir que* l'on fait ou pense quelque chose, que de *savoir pourquoi* l'on fait ou pense quelque chose. J'assimile le premier cas à la conscience présente ou concomitante, tandis que le deuxième cas je l'assimile à la conscience réflexive ou justificative. Or, le type de conscience dont il est question dans l'égoïsme et l'altruisme, en tant que principes de motivation, ne m'apparaît pas clairement. Alors, sans me prononcer sur la nature de la relation entre ces deux types de conscience, j'opte ici pour une lecture de 'conscience' du deuxième type, c'est-à-dire, pour une lecture justificative du terme². Dans cette optique, est consciente toute action ou pensée dont les raisons, concomitantes ou pas, sont explicites à l'agent. 'Etre explicite' exige donc une conscience de deuxième ordre (i.e., [*savoir que* + *savoir pourquoi*]).

D'un autre côté, il y a le domaine des faits qui contraint la description sur laquelle la justification se base. De cette description, même rétrospective, dépend la justification. Cela revient à dire que là où il y a justification il y a description et, deuxièmement, que l'on ne peut pas justifier nos actions ou pensées avec n'importe quelle raison. Si Lola gifle Lalo et si elle peut produire une justification pour son action, alors elle peut produire une description des faits et au moins une raison ou motif pour son acte qui se base sur une telle description. Évidemment, qui dit description dit aussi interprétation³, mais ce qui importe maintenant c'est de souligner que la description porte sur —et est donc déterminée aussi par— les faits. Enfin, si l'on accepte ces postulats, on peut déduire que, face aux faits, il n'y a que la position égoïste qui est susceptible d'accueillir (ou de se faire attribuer) des motivations implicites; ici surgit donc l'asymétrie de justification: la position altruiste exige que les motivations pour agir soient explicites. Il se peut donc que nous soyons égoïstes, mais non pas altruistes, à notre insu⁴.

3. Déontologie vs. Conséquentialisme

Passons maintenant au domaine de l'éthique, c'est-à-dire, au domaine qui étudie les critères pour savoir si les raisons qui justifient nos actes sont acceptables socialement ou non. Autrement dit, sans jeu de mots, passons à l'étude des justifications des justifications. Mais, d'abord, considérons la petite scène suivante:

Il était une fois deux amis affamés se trouvant devant une assiette contenant deux parts, de taille inégale, d'un gâteau délicieux. L'un d'entre eux prit l'initiative de servir et donna la plus petite part à l'autre, se réservant la plus grande.

- Pourquoi t'es-tu servi la plus grande part? C'est incorrect! —se plaignit son ami.
- Qu'aurais-tu fais alors dans mon cas ? —lui demanda le premier.
- Et bien, naturellement j'aurais pris la petite part pour moi! —répondit l'autre d'un ton indigné.
- Je le savais. C'est pourquoi je te l'ai donnée —dit son ami, en croquant sa part.⁵

² Ceci nous évite de buter sur des apories ou, pour le moins, sur des problèmes trop complexes ou confus.

³ Les questions de l'objectivité de la description et de la vérité des propositions pourraient se poser ici, mais elles ne sont pas directement pertinentes pour notre discussion.

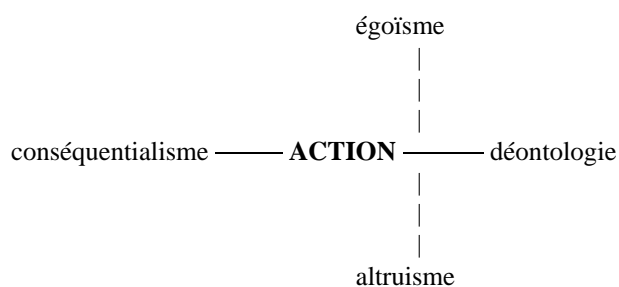
⁴ Rappelons que: a) seul sont conscientes les actions ou pensées qui présentent, de manière explicite, des raisons d'agir ou de penser, b) l'action ou pensée dont le bénéfice reste implicite est non-consciente, et c) l'altruisme comporte une complexité supplémentaire (cf. #1 plus haut).

⁵ Merci à Teófilo Ibarra pour me raconter cette petite histoire.

Dans l'optique traditionnelle des sociétés judéo-chrétiennes l'altruisme ou abnégation est la voie 'naturelle' pour l'expression de l'amour entre les individus. Et l'égoïsme, par conséquence, est considéré comme une voie 'non-naturelle' et donc inapte, voire contraire, à l'expression de l'amour entre les individus. Dans une telle optique, et si l'on adopte une version extrême et stricte de l'altruisme, les deux amis de la petite histoire, s'ils sont vraiment amis, sont destinés à mourir de faim ou, au mieux, à rester paralysés devant le gâteau: si l'on s'en tient est à cette position, même l'initiative de servir est théoriquement exclue, car cela entraînerait un conflit décisionnel insurmontable et donc une négation de toute action motivée subséquente. On pourrait objecter qu'en pratique ce genre d'impasse théorique se résout aisément en situation, tout comme les paradoxes de Zenon. Il n'empêche que, dans les cas de la motivation et de la justification, les critères éthiques de qualification (de l'action) restent inévitablement théoriques et indépendants des faits. C'est pourquoi on est obligé de considérer le problème de la justification dans toute sa profondeur et complexité —théorique et pratique —ce que devient rapidement un casse-tête impossible.

Je vais donc me contenter de faire une brève présentation du problème des critères éthiques pour la justification de nos actions (et pensées) par rapport à deux doctrines traditionnelles et concurrentielles: la déontologie (principes abstraits ou formels, universels, à priori) [cf. Kant] et le conséquentialisme, sous sa forme utilitariste (principes prenant en compte les circonstances et les conséquences des actions) [cf. Épicure, Hobbes, Hume, Bentham, Mill].

D'abord, notons que l'action peut se placer au milieu d'un carrefour conceptuel assez orthodoxe, où le paramètre de motivation (vertical) est perpendiculaire au paramètre de justification:



Je placerais Hobbes, bon égoïste et utilitariste, en haut à gauche, tandis que mère Thérèse serait en bas, à droite. La position altruiste s'accompagne généralement d'une coloration qui évoque, lorsqu'il s'agit de justifications, le principe de *l'Impératif Catégorique* de Kant. C'est-à-dire que, lorsqu'il s'agit de justifier l'action, la justification prend la forme d'un dogme ou d'une vérité apodictique, ou bien d'une certitude atteinte par réduction à l'absurde de sa négation (c'est-à-dire, de la position égoïste). Notons aussi que l'égoïsme s'inscrit plutôt dans un cadre utilitariste lorsqu'il s'agit de justifier l'action.

Ensuite, remarquons que la notion de bénéfice personnel est absente de toute justification déontologique: il faut faire ceci ou agir comme cela, indépendamment de ma situation ou du résultat de mon action (justification dogmatique). La justification utilitaire, par contre, repose sur les conséquences ou sur le résultat de mon action (justification téléologique), bien que ceci ne soit pas forcément dans un but de profit personnel. Néanmoins, puisque la position utilitaire repose sur un jugement concernant les conséquences de mon action, et donc sur une subjectivité, on est obligé d'accepter, d'une part, l'intervention de la notion de responsabilité (i.e., conscience des conséquences, pouvoir de décision effectif, et capacité de justification) et, d'autre part, l'origine inévitablement égocentrique de tout jugement et de toute action responsable. En d'autres termes, être utilitariste nous force à assumer notre point de vue et notre responsabilité, en les saturant pleinement.

Dans la position déontologique le jugement est exocentrique et donc aveugle vis-à-vis de l'action (le devoir n'exige pas le discernement personnel lorsqu'on justifie l'action: p.ex., lorsqu'on s'explique en disant "Ce sont les ordres"; "C'est la loi"; "C'est l'Oracle qui le veut ainsi"; "C'est la tradition, c'est comme ça"), alors que dans la position utilitariste le jugement et la justification personnels sont des conditions nécessaires à toute action. C'est pourquoi il ressort que la position utilitariste peut accueillir plus 'naturellement' une motivation égoïste.

Enfin, la question de la faillibilité peut toujours se poser pour les deux camps: se peut-il que, même si je suis convaincu que mon devoir suprême est x , le bénéfice réel en faisant x soit nul, voire négatif? Ou bien, se peut-il que, même si je suis convaincu que je maximise mon bénéfice personnel, ou celui de nous tous, en choisissant de faire ou de ne pas faire x , les faits ne correspondent pas à mes prévisions et que, donc, x entraîne une diminution de bénéfice personnel ou collectif? Oui, évidemment le problème de la faillibilité se pose dans les deux thèses. Mais, encore une fois, une asymétrie apparaît et fait pencher la balance de la rationalité: pour une même situation, il n'y a que la position utilitariste qui permet à l'individu d'assumer sa capacité de justification pleine et responsable. C'est-à-dire que, même s'il s'avère qu'il a tort d'avoir pensé ou agit comme cela, l'utilitariste pourra toujours produire des raisons pour défendre ses motivations, pour assumer sa responsabilité et prouver sa rationalité; alors que celui qui agit en fonction d'un dogme ou d'une foi, ne pourra pas produire des raisons d'agir: il ne pourra ni assumer un point de vue ni une capacité de justification responsable vis-à-vis de l'action —il s'arrête donc au bord de la rationalité et disparaît en tant qu'individu éthique.

4. Éros et Agapè

Maintenant, après ce fastidieux mais nécessaire parcours, nous pouvons aborder la question de l'amour. Rappelons que cette notion désigne ici une idée, affection, ou sentiment qui s'inscrit dans, et qualifie, certains rapports entre individus; ici elle sera conçue en tant que sentiment. Traditionnellement, l'amour est quelque chose qui s'éprouve, qui se donne, et qui —s'il est réciproque— lie deux ou plusieurs personnes. Par exemple, si Roméo et Juliette s'aiment, on suppose que l'amour les lie et qu'ils (se) donnent et éprouvent de l'amour. Pour renforcer cette conception, je remarque que le verbe 'aimer', transitif, fonctionne comme le véhicule du sentiment. Ainsi, dire qu'une mère aime son enfant, c'est dire qu'elle lui donne ou lui transmet son amour. Et, dès que l'enfant reçoit cet amour, il est (considéré) aimé. Donc, généralement, 'aimer quelqu'un' c'est:

- a) éprouver un sentiment (l'amour) à l'égard de quelqu'un,
- b) l'exprimer de façon verbale ou comportementale,
- c) si a) et b) sont réciproques: établir un lien entre les participants.

Savoir quelle est la fonction de l'amour dans une société et savoir s'il y a plusieurs types d'amour à l'intérieur de celle-ci (et, le cas échéant, comment les classifier) sont des questions très controversées depuis Platon. C'est pourquoi j'ai choisi de concentrer mon attention seulement sur deux acceptions du terme qui concernent surtout le secteur adulte d'une société: l'amour à dimension érotique⁶ (éros) et l'amour à dimension fraternelle (agapè) [cf. Höffe, Fromm].

Éros est un sentiment qui se distingue d'abord par l'implication de l'élément sensuel ou corporel. C'est l'amour dont il est question, typiquement, dans une relation affective où l'élément sexuel est manifeste. Néanmoins, il est important de souligner que, bien qu'une relation sexuelle implique évidemment l'élément corporel, l'inverse n'est pas vrai⁷. Ensuite, éros implique un niveau important de confiance en l'autre, ne serait-ce que par la vulnérabilité qui entraîne une proximité physique et psychologique réciproque. Enfin, éros se présente, surtout sous son aspect sexuel, en tant que vecteur à destinée unique, qui serait complémentaire et correspondant à un autre dans un rapport exclusif.

Or, la confiance impliquée dans une telle relation, essentielle pour l'ouverture, le maintien et le développement de la relation, peut être conçue comme un capital, qui croît au fur et à mesure que le lien se renforce. Mais, en même temps, puisque ce capital n'a de la valeur que sous l'égide de cette relation, il rapportera des bénéfices seulement aux membres participants. La notion de 'privilège' que procure ce capital émerge alors et s'impose à la relation comme une conséquence logique. Mais qui dit 'privilège' dit aussi 'droits et obligations particulières', lesquelles, à leur tour, deviennent la condition du privilège. Ainsi, le privilège propre à une relation érotique devient —notamment lorsqu'elle est sexuelle— condition et conséquence en même temps. Autrement dit, tant le capital que la relation qui le génère deviennent mutuellement captifs.

Agapè, par contre, est un sentiment qui se distingue d'abord par son caractère de solidarité ou de soutien fraternel. C'est un amour qu'on trouve dans des relations affectives typiquement non-sexuelles (famille et amis) mais aussi au sein des relations marquées par une convergence idéologique ou spirituelle. Le capital de confiance que agapè génère n'atteint pas —en tout cas pas aussi rapidement— le niveau de celui d'éros: ils se situent dans des échelles cognitives et temporelles différentes. Pourtant, comme dans le cas d'éros, la notion de 'privilège' émerge ici aussi, mais, en général, seulement en tant que conséquence (et non pas en tant que condition ou contrainte): peut être parce qu'ici les droits et les obligations sont moindres, ou aussi parce qu'il y a un certain seuil de 'captivité' qui n'est pas franchi. Toujours est-il que agapè offre plus de liberté et se présente, principalement, comme un vecteur à destinées multiples et à correspondances diverses.

Toutefois, bien que ces deux acceptions de l'amour aient plusieurs points en commun et qu'elles soient compatibles (du moins en principe), je constate que, dans la pratique, elles sont rendues incompatibles, voire contraires. Cet antagonisme de fait, j'en suis persuadé, est le résultat d'une conception anachronique, souvent psychologisante et cynique, des rapports entre individus, qui empêche ceux-ci et la société, à leur détriment, de se concevoir et de se vivre mutuellement de manière organique, synchronique et synergique (donc de manière écologique).

A l'exception de la section suivante, où j'ébauche une critique de l'anachronisme en question, le reste de ce texte sera consacré à proposer un cadre théorique du social permettant, d'abord, le mélange ou la fusion optimale d'éros et d'agapè (dans une optique écologique) et, ensuite, un moyen éthique, rationnel et constructif de mettre en pratique cette nouvelle forme d'amour.

5. Le modèle pyramidal

⁶Ici je réserve l'adjectif 'érotique' pour qualifier exclusivement les relations entre individus.

⁷ Malgré les différences culturelles, je crois qu'il est possible de se toucher, et même de se faire plaisir (sensuellement), sans le faire pour autant sexuellement (désir de copulation, exacerbation psycho-sensori-motrice, génitaux impliqués). Cependant, il est vrai que l'élément sensuel (dans une relation) peut conduire à l'élément sexuel.

Il y a, dans nos sociétés (telles que spécifiées plus haut), un paradigme social établi depuis longtemps qui conditionne notre façon de concevoir l'amour et d'avoir des rapports amoureux (au sens large) avec autrui. Je vais donc décrire un modèle de prescription ou de détermination socio-cognitive que nous trouvons installé (par défaut) dans la société où nous naissons.

Ce modèle standard je l'appellerai 'le modèle pyramidal'. C'est un modèle qui place l'individu au sommet d'une pyramide à couches hiérarchiques et au centre d'une configuration catégorielle concentrique. Ces deux caractéristiques structurent nos conceptions et nos rapports de façon verticale et égocentrique, de manière à rendre progressivement négligeable tout ce qui s'éloigne du sommet et du centre. En outre, si l'on conçoit le modèle pyramidal comme un sous-ensemble de la structure sociale générale, on peut dire que nos rapports et nos liens amoureux sont des rapports institutionnalisés: la structure sociale les accueille et les détermine. Le modèle standard est donc égocentrique, vertical, institutionnel et structurant (voir figures 5.1 et 5.2).

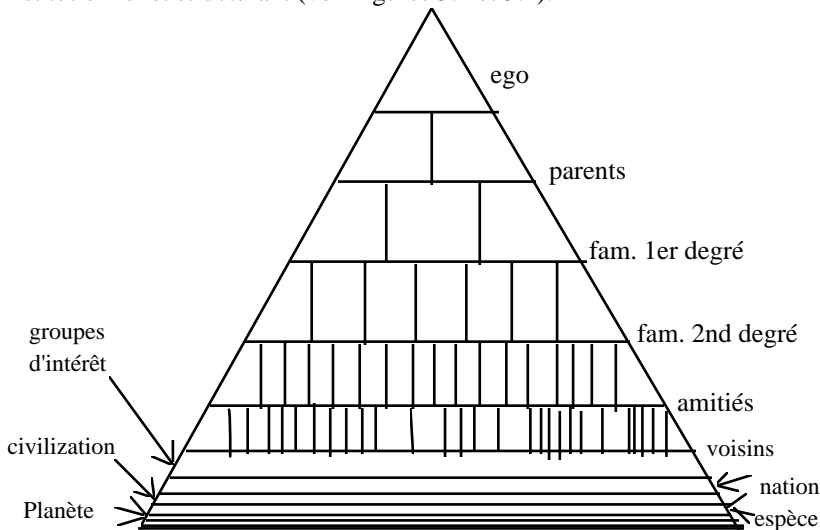


FIG. 5.1
MODELE PYRAMIDAL DE DÉTERMINATION SOCIO-COGNITIVE:

La structuration sociale relative à 'X' correspond à la structuration cognitive de 'X'. Cette structuration s'organise essentiellement de manière égocentrique et verticale. Néanmoins, le contenu effectif, l'ordre hiérarchique, et le poids des couches sont contingents et relatifs à 'X' et au contexte.

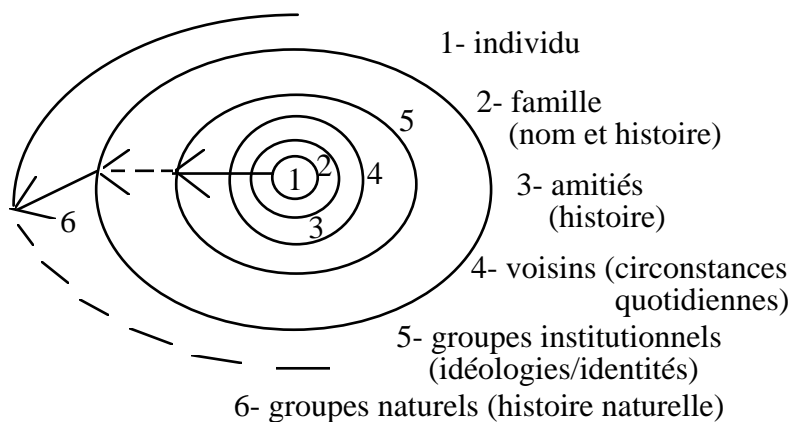


FIG. 5.2
CATÉGORISATION HIÉRARCHIQUE DE RELATIONS: ÉGOCENTRIQUE ET CONCENTRIQUE. (Résultant de l'évolution basée sur le modèle pyramidal). Exemple d'une configuration relative à 'X' et à son contexte en $t = m$. Noter le saut qualitatif d'identité exigé entre '4' et '5', ainsi que l'identité fléchie relative à '6'.

Il me semble vérifié que nos facultés cognitives et nos actions, au sein des sociétés en question, exigent l'existence d'une entité —physique et morale— constituée autour des notions de <point de vue>, d'<agent> et de <responsabilité>. La présence de 'personnes' (agents égocentriques et éthiques) s'impose alors, et s'annonce déjà dans le domaine de la jurisprudence. Rappelons, en passant, qu'une telle présence demande une rationalité d'arrière plan partagée au préalable, qui prescrit les conditions de motivation et de justification à tout acte social se voulant rationnel⁸. A leur tour, motivations et justifications vont impliquer nécessairement un agent narrateur (à la première personne) et au moins une raison d'agir. C'est donc parce que nous sommes rationnels qu'individualité et éthique s'impliquent mutuellement.

Or, dans le modèle pyramidal, la soumission de l'amour à un régime institutionnel, hiérarchique, déontologique et à prétentions altruistes, tout en postulant un agent égocentrique, responsable et libre, semble inefficace et inacceptable. Je suis de l'avis qu'une telle contradiction nous aliène cognitivement et existentiellement, et empêche donc l'émergence d'une vision et d'une pratique écologique de l'amour. Aliénation cognitive, parce que l'on oppose ou scinde artificiellement les notions d'individualité et de rationalité, et existentielle, parce que l'on enlève la capacité

⁸ Définition personnelle: toute acte est social ss'il établit, affecte ou implique des rapports/liens, entre au moins deux individus, de façon réciproque.

d'autodétermination et de plénitude à l'individu. Et cette aliénation est engendrée et articulée par un égoïsme 'négatif' que j'assimile à l'égoïsme fort et anti-social traditionnellement critiqué.

L'égoïsme négatif, sous certains aspects, tient éros et agapè comme incompatibles, le privilège (dérivé du capital de confiance) comme un bien exclusif, les institutions comme l'espace naturel pour concevoir et canaliser les rapports interpersonnels, les catégories cognitives comme des hiérarchies naturelles d'identification et d'appartenance. De plus, l'égoïsme négatif demeure largement tributaire de l'optique altruiste de motivation et de celle déontologique⁹ de justification, bien que, dans les faits, la pensée et l'action ne correspondent pas à leur prétendue genèse théorique de motivation, mais plutôt à une logique égoïste (forte) et non-consciente, voire hypocrite. Enfin, l'égoïsme négatif issu du modèle pyramidal conçoit l'individu comme une entité monolithique (et non comme une personne en devenir) et le milieu social comme un environnement hostile et concurrentiel, surtout dans la mesure où celui-ci s'éloigne —de manière physique et figurée— de l'Ego¹⁰.

Sans être forcément d'accord avec cette exposition, il reste que des questions se posent à propos, d'une part, de la viabilité de l'amour (et de l'action dérivée) dans le cadre du modèle standard, ainsi que de son développement éthique, rationnel et constructif; et, d'autre part, à propos de la possible élaboration d'autres modèles ou cadres alternatifs qui favoriseraient un tel développement. Quant à la première question, et étant donné les constatations exposées plus haut (et à la suite, entre autres, d'Épictète, de Kierkegaard, de Mill, de Nietzsche, de Tolstoï, de Russell), ma réponse est courte et nette: le cadre courant (standard) d'insertion pour l'amour est périmé. Quant à la deuxième question: il est possible, en effet, de concevoir un cadre alternatif que j'expose maintenant.

6. Le modèle écologique

Une opération cruciale à réaliser, dans le cadre alternatif que je propose, c'est la réhabilitation d'un certain égoïsme ou individualisme 'faible' que j'appelle 'égoïsme positif' et que je définis plus bas comme étant compatible avec une vision écologique, mais surtout, avec une vision éthique, rationnelle et constructive de l'amour. Il faut donc rejeter le mépris ou aversion que suscite au premier abord le terme ('l'égoïsme'), et ensuite envisager la possibilité d'un usage prolifère de celui-ci, compatible avec l'amour. Ceci s'oppose à certaines positions psychologisantes et mystifiantes des sentiments qui considèrent, entre autres, qu'il n'y a pas de place pour l'égoïsme dans l'amour¹¹.

Une autre opération cruciale à réaliser consiste à comprendre que, en raison de l'absence de choix au niveau de la structure et des institutions qui accueillent et qui permettent notre développement cognitif précoce, une manière alternative de concevoir l'amour sera possible, initialement, seulement chez l'individu adulte, conscient, qui est susceptible d'avoir un regard critique sur ses pensées et ses actions. Ces conditions satisfaites, c'est au niveau de la détermination (que les institutions imposent à l'individu) que l'on peut agir pour la contrecarrer. Rappelons ici que l'on peut placer l'action au carrefour d'un paramètre de motivation et d'un autre de justification (voir ¶ 3) pour la rendre rationnelle et constructive (ou éthique). On peut donc dire que nous pouvons déterminer nos propres actions et pensées dans la mesure où nous pouvons produire (en nous appropriant ou en créant) des motivations et des justifications pour celles-là.

Or, une telle opération est tout à fait en harmonie avec une perspective égoïste et utilitariste, à laquelle je me rallie. Mais cette perspective semble insuffisante pour accueillir et cultiver une vision et une pratique écologique de l'amour. C'est pourquoi il me semble nécessaire de nuancer une telle perspective, en adoptant une position d'égoïsme 'faible' (que j'appelle 'égoïsme positif'; adjectif: égotif, égotive), et qui peut s'insérer dans une structure alternative pour concevoir nos rapports sociaux de manière égotive et écologique:

⁹Je remarque que, très souvent, les codes déontologiques sont exprimés dans des termes négatifs, comme les commandements et les réglementations.

¹⁰ Par ailleurs, je remarque que, fréquemment, une certaine attitude (souvent cynique ou détachée par rapport à ce que sont 'les amis') semble se réserver le droit de distinguer et de mystifier 'l'Amour' au dépens de 'l'amitié'. Par exemple, R. Nozick distingue entre 'amitié' et 'amour' en disant que "Quand un malheur frappe un ami, c'est lui que le malheur frappe...En revanche, quand un malheur frappe un être aimé, vous aussi êtes touché." Nozick nuance quand même la distinction en disant "...je ne veux pas...dire par là que l'on ne puisse pas aussi *aimer un ami d'amour*." Mes italiques. p. 73. Un autre exemple, pris au hasard: F.Alberoni constate que "L'amitié ne connaît pas la mise à l'épreuve, à l'inverse de l'amour qui met constamment en jeu, sans la moindre hésitation, la solidité du sentiment." p.23

¹¹ Par exemple, j'ai trouvé les passages suivants: "L'égoïste est, en principe, un homme incomplet, qui ne peut pas s'insérer harmonieusement dans son milieu social; il est, au fond, un désadapté. Le progrès psychologique de l'homme implique une collaboration intelligente et bienveillante avec les autres..." p.92, C. Konczewski. Plus choquant encore: "Ce que nous révèle en effet l'amour, c'est que nous ne sommes que par autrui et que le monde même n'existe que pour ceux qui sont au moins deux"; "[l'amour] n'est point la perfection close de deux êtres, mais leur commune participation à une réalité supérieure par laquelle seule ils sont véritablement." (sic); et, encore, "ce ne sont pas les êtres singuliers qui produisent l'amour, mais l'amour qui les produit."; "Le sens de l'amour est donc clair: nulle autre expérience ne permet de comprendre aussi nettement la conciliation de l'un et du multiple dans la vie de l'esprit" p. 47-50, J. Lacroix.

MODELE ECOLOGIQUE:
STRUCTURE VIRTUELLE DE BASE, AVEC
INDIVIDU 'X' AU CENTRE

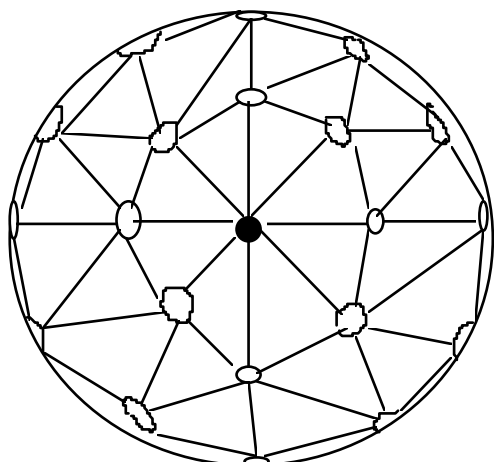


FIG 6.1

MODELE ECOLOGIQUE:
EXEMPLE D'UN ETAT STRUCTUREL
POUR 'X' EN $t=n$

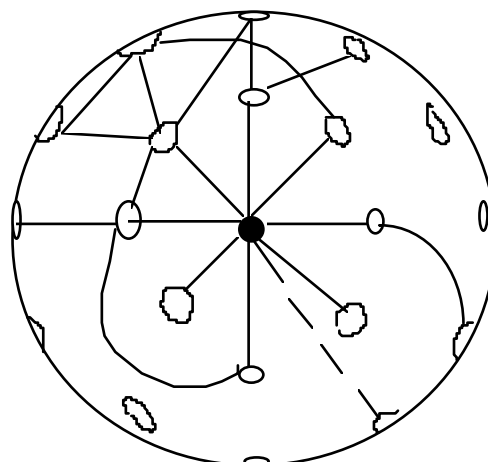
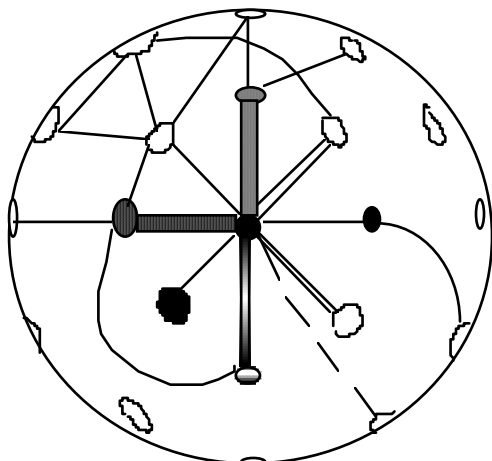


FIG.6.2

Inserée dans ce modèle, la position égotive se définit surtout en fonction des paramètres cognitifs, sociaux, éthiques et politiques qu'elle peut optimiser sur la base de certains traits du modèle: 1) le modèle est égocentrique et permet donc à chaque individu d'en être 'le centre', 2) le modèle introduit obligatoirement le milieu social comme variable à considérer dans toute motivation et justification d'acte social: l'action individuelle est locale, complexe et a des répercussions globales (principe organique), 3) puisque les effets semi-globaux (connexions de 1er degré) et globaux entrent en ligne de compte pour l'action, ce modèle interdit la version forte de la position égoïste et cherche à fusionner la notion de bénéfice individuel avec celle de bénéfice collectif, tout en octroyant une priorité minimale vitale à l'individu, car il est source et véhicule de l'acte rationnel et responsable, et 4) dans ce modèle, la condition de conscience est exigée à toute action ou pensée égotive, en éliminant par là la possibilité de bénéfice implicite ou non-conscient, et en permettant à autrui de nous percevoir clairement et ouvertement: on doit donc assumer et afficher nos motivations de manière explicite.

Dans ce nouveau cadre, l'amour écologique ou holiste (caractérisé déjà par l'assimilation d'éros et d'agapè dans un même registre) a non seulement une place à occuper, mais aussi un rôle important à jouer. La place à occuper se trouve dans un réseau égocentrique, horizontal, organique, inclusif, et transitif (non-hiérarchique, non-catégoriel, non-institutionnel); et le rôle à jouer est évidemment à déterminer par les individus au cours des échanges dans le réseau à **construire**. Néanmoins, je peux déjà proposer deux rôles pour l'amour: a) en tant que capacité affective permettant l'inclusion, l'exploration et le développement —dans la mesure des capacités et des désirs individuels— de plusieurs relations en même temps (quoique pas forcément dans le même espace) de sorte à optimiser un bénéfice personnel (et donc, comme nous l'avons vu précédemment, réciproque et global), tout en sous-entendant dans cela l'élaboration d'une formule ou équation particulière à chaque relation (principes de synchronie et de synergie de groupe); et b) en tant que privilège non-captif permettant la libre circulation dans le réseau, et aussi la transitivité de valeur (vis-à-vis des relations non-contiguës —de 2ème degré et au delà) du capital de confiance accumulé lors des échanges avec des relations contiguës. On peut visualiser un tel développement ainsi (FIG. 6.3):

**MODELE ECOLOGIQUE:
EXEMPLE D'UN ETAT STRUCTUREL
DEVELOPPE PAR 'X' EN $t=n+m$**



LEGENDE (du point de vue de 'X'):

———— = relation établie

— — — = proto-relation

Largeur des traits = Robustesse de la relation =
Importance du capital de confiance généré

Couleur des traits = Intensité/fréquentation
(statistique) de la relation

Texture des traits = Formule ('chimie') particulière
à la relation

Finalement, je reste persuadé que l'émergence d'une nouvelle forme de vie due à la mise en oeuvre d'une telle conception, nous permettrait, enfin, de mettre la relation entre effet et cause à l'inverse (cf. Nietzsche, *Twilight of the Idols*) et de devenir pleinement éthiques et conséquents. Dans une telle structure, la seule déontologie initiale nécessaire serait celle consistant en des voeux de bonne volonté et de loyauté (dans le but d'atteindre une robustesse et une maturité fonctionnelle dans les liens) et, dans un deuxième temps, le seul voeu à renouveler serait celui de rester égotif.

Par ailleurs, rien n'interdit l'établissement des liens —ou même d'un seul lien— à titre préférentiel si les conditions d'individualité et d'écologie sont remplies. Ce(s) témoin(s) préférentiel(s) peu(ven)t jouer un rôle d'abri, de famille, de foyer émotionnel, de ressource passionnelle, de nid de progéniture; mais, et ceci est important aussi, un témoin préférentiel peut être un outil précieux, un étalon d'excellence, un miroir magique.

* * *

Si je n'ai pas réussi même à esquisser un projet capable d'installer correctement, dans nos pratiques, ce dont les chansons d'amour parlent, c'est peut-être parce que je suis incapable de comprendre les chansons d'amour...ou bien,

Références:

- Alberoni, Francesco (1984). *L'amitié*. Paris: POCKET, 1995 éd.
 Épictète. *Maximes*. Arles: Collection Babel, ACTES SUD, 1993 éd.
 Bretall, R. (ed). *A Kierkegaard anthology*. Princeton: Princeton University Press, 1973 ed.
 Bernhardt, Jean (1989) *Hobbes*. Col. Que sais-je? Paris: PUF
 Chadwick, Henry (1986) *Augustine*. Oxford: Oxford University Press, p.52 (citation)
 Fromm, E. (1956) *L'art d'aimer*. Paris: Desclée de Brouwer, 66-68/ 72-76, 1995 éd.
 Gorée, G. and Barbier, J. *Love without boundaries: Mother Theresa of Calcutta*. Indiana: Our Sunday Visitor, Inc. 1979 éd.
 Höffe, Otrfied (ed). *Petit Dictionnaire d'Ethique*. Paris: éd. du Cerf, 1993
 Hume, David. An enquiry concerning human understanding, sec. VIII. In *The Empiricists*. New York: Anchor Books, 1976 éd.
 Kant, I. *Fondements de la métaphysique des moeurs*. Paris: Librairie Delagrave, 1994 éd.
 Kierkegaard, S. *The diary of Søren Kierkegaard*. New Jersey: Citadel Press, 1960 éd.
 Kierkegaard, S. *The present age*. New York: Harper Torchbooks, 1962 éd.
 Konczewski, C. (1951) *La Sympathie comme fonction de progrès et de connaissance*. Paris: PUF
 Lacroix, Jean (1956). *Les sentiments et la vie morale*. Paris: PUF, 1976 éd.
 Mill, John S. *L'Utilitarisme*. Paris: Flammarion, 1988 éd.
 Nietzsche, F. Thus spoke Zarathustra. In Walter Kaufmann ed., *The Portable Nietzsche*. New York: Penguin Books, 1982 éd.
 Nietzsche, F. Twilight of the idols. In Walter Kaufmann ed., *The Portable Nietzsche*. New York: Penguin Books, 1982 éd.
 Nozick, Robert (1995 tr). *Méditations sur la vie*. Paris: Odile Jacob.
 Peters, R.S., article 'Hobbes, Thomas', in *The Encyclopedia of Philosophy*. New York: Macmillan Publishing Co. Vol. IV, 39-44

Plato, The Symposium. In *The Collected Dialogues*, Hamilton and Cairns ed. New Jersey:
Princeton University Press, 1985 éd.

Russell, Bertrand (1949). *Autoridad e Individuo*. Mexico: Fondo de Cultura Económico, 1973 éd.

Tolstoy, Leo. *The death of Ivan Ilych*. New York: New American Library, 1960 éd.

Mes remerciements profonds à Priscille Glad, courageuse et nourrissante exploratrice et compagne (bien que pas assez égotive). Merci à Elisabeth Vroemen pour ses réflexions. Merci à Olivier Fontaine pour ses corrections et pour sa synergie toujours éclairante.

Pensée privée et communication sociale

Jean-louis Dessalles ¹

¹ Ecole Nationale Supérieure des Télécommunications

46 rue Barrault - 75013 Paris

E-mail : dessalles@enst.fr

Résumé

Le fonctionnement cognitif d'un individu et le fonctionnement de la communication sociale sont deux phénomènes soumis à des contraintes assez différentes. Il n'est donc pas évident d'imaginer un parallélisme étroit entre ces deux processus. Pourtant la question de la parenté entre la pensée et le langage a été maintes fois abordée : pensons-nous avec des mots, la pensée est-elle un langage intériorisé, le langage précède-t-il la pensée chez l'enfant, etc. ? Je propose aussi d'aborder cette question, mais sous un angle original. La caractéristique du langage qui est retenue ici n'est pas la faculté d'agencer des mots pour produire des phrases qui évoquent une signification. Si l'on regarde le langage au niveau pragmatique, en retenant seulement de la faculté langagière le fait qu'elle permet d'enchaîner des arguments pertinents, alors force est de constater, dans le détail, une forte similitude entre le déroulement du flux de la pensée et celui de la communication sociale par excellence : la conversation. En d'autres termes, je suggère que l'enchaînement des pensées obéit aux contraintes de pertinence. Si l'on accepte de considérer une telle hypothèse, alors on doit envisager la possibilité que la structure de la pensée privée, constitutive de l'intelligence des individus humains, soit phylogénétiquement une conséquence des exigences de la communication sociale.

mots clés : pensée, communication, conversation, pertinence, conscience

1. Introduction : présence du social dans le fonctionnement cognitif individuel

La pensée est un phénomène privé, concernant l'individu. Le langage est par essence social : une langue n'existe que parce qu'un groupe la parle, l'existence de la fonction langagière est intimement liée au caractère social de l'espèce humaine. Tout lien que l'on peut établir entre pensée et langage devrait donc sembler un peu paradoxal. Il ne va pas de soi, en effet, qu'un lien étroit puisse exister entre un aspect fondamental du fonctionnement cognitif individuel, la pensée, et un dispositif permettant aux individus de communiquer au sein de leur groupe. Pour prendre une image technologique, les exigences liées à la transmission des données sur les lignes de télécommunications n'ont rien à voir avec les exigences du traitement de l'information dans le processeur central d'un ordinateur. Dans le premier cas on parle de débit, de bruit, de modulation, de trame, etc. Dans le second cas il est plutôt question de traitement multitâche, de mémoire, de parallélisation, d'interruption... Les paramètres qui joueront sur l'efficacité et sur la fiabilité ne seront pas les mêmes lorsqu'il s'agit de communiquer l'information et lorsqu'il s'agit de la traiter. Si la pensée joue un rôle psychologique important, si le fonctionnement efficace de l'individu en dépend, comment expliquer que sa structure soit calquée sur l'organisation d'un système qui sert simplement à échanger des informations ? Réciproquement, les contraintes de la communication verbale : pauvreté du canal acoustique, nécessité d'introduire de la redondance et de la rétroaction, etc. sont telles qu'il semble *a priori* difficile d'accepter que le langage soit une simple émanation de la pensée. Pourtant, je vais suggérer ici qu'un lien extrêmement fort, une quasi identité, existe entre pensée et langage, que le langage est premier et que la pensée en est, en quelque sorte, un mécanisme dérivé. Je serai donc amené à montrer que la pensée obéit à des contraintes qui ne résultent pas toutes d'une exigence d'efficacité du fonctionnement psychologique individuel. Autrement dit le fonctionnement cognitif individuel est, peut-être par « accident », une conséquence des contraintes liées au caractère social de l'espèce humaine.

2. Langage et pensée

De nombreux auteurs ont souligné les liens étroits qui unissent pensée et langage. A commencer par le rapport entre la pensée et les mots. Minsky, par exemple, suggère le fait que la pensée est amplifiée par la langue :

« Si la parole implique la pensée, on peut alors se demander : “*Quelle proportion de la pensée ordinaire nécessite l’emploi de mots ?*” Il est certain que bon nombre de nos méthodes de pensée les plus efficaces font très peu appel à nos services linguistiques. Peut-être ne recourons-nous aux mots que lorsque d’autres méthodes échouent. Cependant l’usage du langage peut ouvrir des univers de pensée totalement nouveaux, parce qu’une fois que l’on peut représenter des choses sous forme de chaînes de mots, il devient possible de les employer de façons infiniment variées [Minsky 1985 p. 453]. »

Fodor, dans *The language of thought*, observe que le fonctionnement cognitif suppose la manipulation des symboles d’un langage, et réfute le fait que ce langage soit une langue naturelle [Fodor 1975 p. 56]. Cette question du rapport entre pensée et *langue* ne sera pas abordée ici. Ce qui nous intéresse est le rapport entre la pensée et le *langage*, considéré au niveau *pragmatique* comme la capacité à reconnaître et à engendrer des arguments pertinents. Par exemple, D. Sperber et D. Wilson [1987] critiquent le point de vue de J. Katz selon lequel toute pensée peut être exprimée par une phrase dans n’importe quelle langue naturelle. Ils rappellent qu’au delà des mots, de la syntaxe et de la sémantique, le contexte et les processus inférentiels jouent nécessairement un rôle fondamental dans l’organisation des pensées.

C’est bien à ce niveau que le langage est supposé jouer un rôle dans la mise en place des connaissances et des capacités de raisonnement de l’enfant. Il est habituel, à cet égard, d’opposer les conceptions de Piaget [*Le langage et la pensée chez l’enfant*, 1923], [*Th. du langage, Th. de l’apprentissage*, in Piatelli-Palmarini 1979] et celles de Vygotski [*Pensée et langage*, 1962]. Ces deux auteurs, à partir de l’observation du même phénomène, parviennent à des conclusions diamétralement opposées : voyant des enfants se parler à eux-mêmes pendant des heures tout en jouant ou en vivant des scènes imaginaires, les deux auteurs conviennent de l’importance du langage dans la formation de la pensée et du raisonnement. Mais alors que pour Piaget l’autodialogue constitue une étape dans le processus de décentration qui mène l’enfant d’un point de vue égocentrique vers un point de vue où l’autre, l’interlocuteur, joue un rôle symétrique au sien, Vygotski considère que le processus inverse est à l’œuvre : lors de ses verbalisations solitaires l’enfant intériorise les interactions sociales qu’il transforme peu à peu en processus de pensée. L’enfant acquiert ainsi ses idées sur le monde à travers ses dialogues, y compris ceux qu’il conduit avec son compagnon imaginaire (*the Virtual Other* [Braten 1992]).

Ces auteurs, et beaucoup d’autres qui abordent cette même question des rapports entre notre faculté de penser et notre comportement langagier, considèrent la pensée d’une manière assez générale au point qu’il est difficile de distinguer chez eux la pensée du fonctionnement cognitif dans tout ce qu’il a de spécifiquement humain. Par exemple, Davidson [*Thought and talk*, 1975] emploie le mot *thought* pour désigner les croyances et les désirs que le sujet humain peut avoir à un moment donné. La question que je souhaite aborder ici est beaucoup plus spécifique : il s’agit de la pensée en tant qu’aspect très particulier du fonctionnement cognitif, la pensée consciente, celle qui, comme on dit banalement, « a un fil » que l’on peut parfois perdre, la pensée en tant qu’activité à laquelle nous ne pouvons nous empêcher de nous adonner dès que nous n’avons rien à faire.

3. Le fil de la pensée consciente

3.1. Flux de pensée et conscience

Le concept de pensée évoque immédiatement celui de conscience. Certes, il existe des phénomènes de pensée inconscients, comme l’indique le témoignage suivant :

« Je remets des feuilles vierges dans le bac à feuilles de mon imprimante. Je me mets à penser vaguement à un chevalier et à son cheval, juste assez nettement pour me demander d’où me vient cette

pensée parasite. Il m'apparaît alors que le lien est le mot "page" : page blanche et page écuyer. (noter que mon voisin de bureau s'appelle Chevalier). » [Restitution juste après l'épisode]

Ces processus inconscients, généralement de type associatif, ne sont pas pris en compte dans la description que je veux donner ici. Lorsqu'un individu reste un instant sans rien faire qui occupe sa conscience (ce que certains appellent la mémoire de travail [Baddeley 1992]), il se met à penser. Appelons ce phénomène *flux de pensée* (*stream of thought*). Si on lui demande « A quoi pensez-vous ? », il est capable, au prix d'un petit effort, de faire un compte-rendu qu'il estime fidèle de ce à quoi il vient de penser. En revanche ce genre de témoignage n'existe quasiment pas lorsque la mémoire de travail est occupée à une tâche comme l'énoncé de l'alphabet à l'envers*.

Le flux de pensée semble donc tellement lié à la conscience que l'on peut se demander s'il s'agit d'un phénomène caractérisable que l'on peut isoler de l'ensemble du fonctionnement cognitif conscient. Or si la conscience semble être une propriété nécessaire de la pensée, la réciproque est certainement fautive. En ce moment, je suis conscient de la position de mon corps, de la douleur qui ne me quitte pas dans cette molaire, de la luminosité de l'espace ambiant, etc. Tous ces aspects ne font cependant pas partie du flux de pensée qui m'habite en ce moment. Intéressons-nous ici aux processus liés au flux de pensée, indépendamment de leur rapports avec la conscience.

La question qui se pose alors avec force est donc de savoir si le flux de pensée correspond à un phénomène psychologique bien caractérisé. En effet, nous ne le connaissons que par l'introspection ou par les comptes-rendus linguistiques qui en sont faits en direct ou en différé. Comment peut-on être certain qu'il ne s'agit pas d'un artefact lié à la verbalisation ?

3.2. La thèse de l'artefact verbal

Lorsque l'on demande soudain à un sujet oisif à quoi il pense en le priant d'écrire le contenu des pensées qu'il vient d'avoir, on obtient un texte. Dans quelle mesure peut-on affirmer que ce texte représente le contenu d'un flux de pensée ? On pourrait aisément prétendre au contraire que le texte est une invention, causée par la requête de l'expérimentateur, sans que l'on puisse savoir dans quelle mesure cette invention est inspirée par les processus cognitifs qui ont précédé dans l'esprit du sujet. C'est ce que j'appelle ici la thèse de l'artefact verbal. Le même problème se pose dans les expériences de « pensée à voix haute ».

Cette question de la fiabilité des comptes-rendus verbaux se pose de la même façon dans d'autres contextes, notamment en résolution de problème. En effet, lorsqu'on demande à des sujets de résoudre un problème tout en commentant ce qui leur « passe par la tête », on obtient des verbalisations dont il s'agit d'estimer la validité en tant qu'indices des processus de raisonnement. Cette question de l'authenticité des verbalisations est depuis longtemps débattue [Caron-Pargue & Caron, 1989]. Certains montrent que les verbalisations des sujets à propos de leurs propres processus cognitifs sont largement incomplètes et parfois erronées [Nisbett & DeCamp Wilson 1977]. D'autres répondent que ces erreurs ou ces omissions ne concernent que les informations auxquelles les sujets n'ont pas été contraints d'être attentifs par la tâche [Ericsson & Simon 1980, 1984].

Pour contrer la thèse de l'artefact, j'invoquerai deux arguments. Le premier est le sentiment très fort d'authenticité manifesté par le sujet qui vient de fournir ses pensées par écrit[†]. Quelle en est la valeur ? Cette valeur existe, puisqu'en dernier ressort, l'activité scientifique repose sur un tel sentiment. En effet, la seule différence entre une expérience introspective reproductible et une expérience scientifique standard, c'est que la deuxième est médiatisée par un support matériel externe que nous supposons commun à tous les observateurs présents. Nous sommes convaincus de l'authenticité de ce que nous avons vu au cours d'une expérience de mécanique (par ex. deux corps

* Un sujet avoua tout de même avoir eu la pensée suivante en arrivant vers la lettre K : la tâche devient plus facile lorsque l'on approche du début de l'alphabet.

[†] Le lecteur, avant de se prononcer, est invité à réaliser l'expérience par lui-même : écrire l'un de ses propres épisodes de pensée.

de poids différent qui tombent à la même vitesse), et c'est là-dessus que nous fondons notre notion d'objectivité. Pourtant, cette authenticité n'est qu'un sentiment : nous pouvons être trompés par un illusionniste habile. Nous ne pouvons même pas invalider le raisonnement qui prétendrait que nous n'avons jamais réellement vu cette expérience, puisque nous venons d'être créés, avec nos souvenirs, en même temps que le reste du monde, il y a juste quelques secondes (voir *Le nombril d'Adam* dans [Gould 1985]). Le modèle défendu ici à propos des contenus de pensée possède toutes les caractéristiques de scientificité (en particulier, chacun peut reproduire l'expérience sur lui-même et tenter ainsi de réfuter le modèle) sauf une : l'objet particulier (le flux de pensée) n'est pas directement partageable. Chacun doit donc mettre cette théorie à l'épreuve par lui-même. Mais n'est-ce pas ce que nous devrions toujours faire, idéalement, en Science ?

Le deuxième argument pour défendre la validité des comptes-rendus de pensée est plus spécifique : il concerne la structure formelle de ces comptes-rendus. Nous allons voir en effet qu'ils possèdent une propriété caractéristique, que les sujets ne soupçonnent pas, et qui fait défaut aux comptes-rendus artificiels de pensée : ceux qu'on demande aux sujets d'inventer ou ceux que l'on trouve dans les romans. Cette propriété est celle de la pertinence des pensées.

3.3. La séquentialité du flux de pensée

On peut s'étonner du fait que ces processus de flux de pensée, malgré leur importance psychologique, aient été fort peu étudiés pour eux-mêmes. Comme le note K. Pope [1978], le phénomène du flux de la pensée a été abondamment décrit dans la littérature et dans le cinéma, mais pendant longtemps il n'a pas été considéré comme un sujet scientifique sérieux, surtout dans le contexte de la psychologie américaine. Or le phénomène de pensée que nous sommes en train de considérer est vécu subjectivement comme un flux séquentiel, et ce phénomène est sans doute suffisamment inattendu pour qu'on s'en préoccupe. En effet, le cerveau se caractérise surtout comme un système parallèle qui n'est en rien soumis aux contraintes de séquentialité que connaissent par exemple les ordinateurs à processeur central. Pourquoi notre pensée serait-elle astreinte à cette mono-dimensionnalité ? Ce pourrait être un aspect artefactuel lié à la verbalisation, à un système de monitoring qui surveille des processus parallèles inconscients ou même une simple contrainte liée à la taille de notre mémoire de travail :

« Our introspective conceptions are accordingly very different from their unconscious counterparts : we tend to force intrinsically parallel notions into a serial strait-jacket. » [Johnson-Laird 1983 p. 470]

« Pourquoi, par exemple, les pensées semblent-elles si souvent jaillir en flots séquentiels ? C'est parce que chaque fois qu'il n'y a plus de place, les enregistrements de nos pensées récentes doivent repousser ceux de nos pensées plus anciennes. » [Minsky 1985 p. 283]

Envisageons tout de même avec un peu de sérieux l'hypothèse selon laquelle il y aurait réellement, et pour une raison fondamentale à découvrir, un enchaînement strictement séquentiel des pensées. Pour comprendre le mécanisme de ces enchaînements, il faut disposer de données fiables. S. Freud, et d'autres après lui, ont étudié en détail quelques aspects de l'enchaînement de certaines pensées entre elles par des mécanismes associatifs. Au delà de ces enchaînements associatifs au déterminisme parfois inconscient, il est fait mention dans [Freud 1907 p.419] de ces épisodes de flux de pensée appelés fantasmes ou rêves diurnes, et dont la structure est censée ressembler à celle des rêves. Mais les témoignages écrits de contenus de flux de pensée restent rares dans la littérature. E. Klinger [1978] mentionne les difficultés méthodologiques liées au recueil de ces témoignages : les questionnaires, l'échantillonnage, la pensée à haute voix, la restitution, toutes ces méthodes restent très imparfaites. Pourtant, l'exercice est relativement facile et chacun peut l'exercer sur soi ou sur d'autres. L'échantillonnage, par exemple, consiste à interrompre une personne, de préférence quelqu'un qui n'est pas engagé dans l'accomplissement d'une tâche non routinière, et à lui demander le contenu actuel de ses pensées. Cette personne peut aussi accepter de livrer, en se les remémorant, les contenus de pensée qui ont précédé. Un sentiment de fiabilité élevé accompagne généralement ces témoignages, ce qui est très loin d'être le cas lorsqu'on demande par

exemple aux mêmes personnes de livrer le contenu de leurs rêves. Voici quatre témoignages donnés par deux sujets. Les deux premiers ont été obtenus par la méthode « pensée à haute voix », les deux autres ont été restitués par le sujet juste après avoir été vécus.

« I'm sitting here thinking that I have to be thinking. . . I'm singing a song in my head, that new Ringo Starr song that I heard on the radio. It must have been the last thing that I heard when I went out of the house. It's funny I haven't been thinking about it all morning, I just, It just came to me. » (transcript from [Pope 1978]).

« No, but I'd rather be in California. Any time there's a day like today, I can't imagine why anybody would choose to stay on the East Coast anywhere from October to February or March. If Yale wasn't here, I mean, I'm glad I'm at Yale, I wouldn't want to be anywhere else only because of, the only reason that I would want to go anywhere is the climate. If this could be moved to California, that would be ideal. But it's enough, I don't have to worry. I mean this is four years. I'm not losing anything. I'm not missing out. Everything will be there when I get out and I'll just go d o what I want to then. » (transcript from [Pope 1978]).

« Je me retrouve seul à attendre sans livre. Je pense à mon programme informatique (en C). Je me dis que je peux débayer de tête. Je me dis que ça doit être difficile. Je repense à la fin du film "*La diagonale du fou*" où ils jouent aux Echecs de tête. Je revois la scène. Là, ils s'entendaient très bien comme si il y avait une grande amitié entre eux. Alors qu'avant tous leurs rapports étaient sous forme de lutte. » [Restitution après l'épisode]

« J'entends mon fils (20,5 mois) pleurer. Je me dis qu'il y aura bien quelqu'un pour le calmer. Effectivement, je l'entends se calmer. Je n'aurai pas à y aller. Je m'imagine en train de le consoler. Il aurait continué à pleurer quelques secondes dans mes bras. Il aurait compris qu'il était quand même content et qu'il allait se calmer. Un animal n'aurait peut-être pas compris. J'imagine une lionne consolant ses lionceaux en les léchant ; elle se fout qu'ils crient. Pourtant, elle pourrait partir en les abandonnant. Je pense au film ("*Vivre Vite*" de Carlos Saura). A la fin la fille part avec le butin en abandonnant son copain blessé alors que la police arrive. Comment faisait-elle pour échapper à la police ? Celle-ci cherchait un homme. Peut-être est-elle sortie par le sous-sol alors que la police arrivait juste. Ou plutôt, elle était déjà à 100m quand on entend la police. Et le mec ? Il tire et se fait tuer ? Qu'est-ce que j'aurais fait à sa place ? Le film s'arrête là. Il aura pu se rendre. Sa situation était désespérée : il est blessé, la police arrive, il a le choix entre la mort et d'être soigné en prison. En prison on vit. On peut travailler. Je repense à un téléfilm avec un meurtre au musée du jardin des plantes. Un grand patron du labo est accusé : "En prison, vous pourrez continuer vos travaux de recherche,..." Je trouve bizarre qu'il soit accusé ainsi, d'autant qu'après c'est un autre qui est reconnu coupable. » [Restitution après épisode]

La forme de ces témoignages va nous permettre de faire quelques observations sur les rapports entre le flux de pensée et la pragmatique du langage.

4. Pensée et pertinence

4.1. Le dialogue intérieur

Les rapports entre le flux de pensée et le langage, considéré au niveau pragmatique, sont souvent évoqués, mais rarement étudiés. Bien des auteurs ont comparé le flux de pensée à un dialogue intérieur. Par exemple G. Court évoque le processus de pensée qui précède le processus d'expression linguistique :

« Dans une situation d'expression, on est amené à former hors de la présence réelle ou imaginée des interlocuteurs et pour son usage propre, dans une sorte de dialogue avec soi-même, une pensée dont l'organisation interne pourra s'inscrire entre les pôles d'une part d'une rationalité claire, précise, objective, et d'autre part d'une affectivité qui se décante, s'affine, s'humanise en s'exprimant : cette possibilité constitue le fondement de la culture » [Court 1971].

Cette mention du dialogue intérieur est fréquente, mais la thèse que je voudrais défendre ici est beaucoup plus radicale et n'a jamais, à ma connaissance, été envisagée. Il ne s'agit pas simplement de faire une analogie entre le flux de pensée et une forme de dialogue, mais de considérer le flux de pensée comme une véritable conversation. En d'autres termes, ma suggestion

est qu'on ne peut pas distinguer sur des critères formels qualitatifs l'enchaînement des contenus de pensée de celui des arguments échangés au cours d'une conversation. Une conséquence très importante de cette hypothèse est que les contenus du flux de pensée sont soumis aux contraintes de pertinence.

Avant d'examiner cette question d'un point de vue technique, il est permis de se demander : la pensée est une conversation, oui, mais avec qui ? Avec *l'autre virtuel* de Braten [1992] ? En fait, je suggère que cette question peut très bien rester sans réponse. Lorsque l'on étudie l'enchaînement des arguments dans une conversation, on peut très bien ignorer le fait que ces arguments sont émis par des personnes physiques distinctes (tant que cela n'interfère pas avec la connaissance du contexte) [Dessalles 1993]. Il suffit de considérer le phénomène d'auto-réplique illustré dans l'extrait suivant :

L- *J'ai lu que le Maroc était le premier fournisseur en textile de la France. C'est un tout petit pays. T'imagines, tout ce qu'on importe ? ! On [ne] fabrique plus de textile. Hé bien le pays auquel on achète le plus, c'est le Maroc. Remarque, le marché est peut-être morcelé.*
P- *C'est parce que... c'est des entreprises françaises qui se sont installées là-bas. Elles se délocalisent.*

Ici, l'argument « le marché est peut-être morcelé » est donné par *L* comme une tentative d'explication de son propre étonnement. Si le marché est morcelé entre de nombreux fournisseurs, il n'est plus très étonnant que le plus important d'entre eux soit un petit pays. De même, l'extrait n'aurait rien de pathologique si l'argument donné par *P* avait été émis par *L*. Nous pouvons donc tout à fait nous intéresser aux règles d'enchaînement des arguments, ou des contenus de pensée, sans nous soucier de savoir qui parle.

4.2. La structure argumentative du flux de pensée

Lorsque l'on regarde les contenus de pensée donnés plus haut, on observe une structure argumentative caractéristique. Donnons quelques exemples, sans analyser finement chaque extrait. Pour ce faire, il faut passer par une interprétation de certains éléments de ces témoignages. Le lecteur est invité à vérifier la pertinence de ces interprétations, et peut réaliser la même opération sur ses propres contenus de pensée.

Premier extrait. L'argument « *It must have been the last thing that I heard...* » est une explication typique. Elle vient résoudre la question non mentionnée *D'où me vient cette chanson ?*. Vient ensuite un étonnement explicite : pourquoi cette mélodie surgit-elle seulement maintenant ?

Deuxième extrait. Le sujet a pensé précédemment à l'éventualité d'aller à New York. Il repense à un autre de ses projets, celui d'aller en Californie, comme réponse au problème du climat. La phrase suivante « *I can't imagine why anybody...* » montre qu'il transforme ce désir de climat doux en étonnement. Il revient ensuite au réseau de ses contraintes de préférence : il veut aller en Californie pour trouver un climat doux, et en même temps il veut rester à Yale pour ses études. La solution : déplacer Yale ! Le sujet semble renoncer à cette éventualité pour admettre finalement que rester quatre ans à Yale est finalement supportable, ce qui rend la situation acceptable.

Troisième extrait. Le sujet est dans la situation indésirable de perdre son temps. La solution : penser à son programme informatique. Mais il se heurte à une contrainte de faisabilité. Toutefois, dans le film qu'il évoque, les personnages en étaient capables. Intervient alors un étonnement : pourquoi sont-ils devenus brusquement amis ?

Quatrième extrait. Le lecteur est invité à vérifier que le quatrième extrait, comme les autres, peut être interprété de la manière suivante : le sujet est sensible à des contraintes de cohérence (dans la scène finale du film, la police avait bouclé le quartier, or la fille a pu lui échapper) et à des contraintes de préférence (mourir est indésirable, vivre en prison aussi). Le sujet, au cours de ses pensées, oscille sans cesse entre la perception de ce type de contraintes et leur solution : il est indésirable de se déranger pour calmer l'enfant. L'enfant se calme quand on le console. Pourquoi ?

Parce qu'il comprend. Ca ne marcherait donc pas pour des animaux. Comment se fait-il que la lionne n'abandonne pas ses enfants ? Etc.

Cette oscillation entre la perception de contraintes et les tentatives pour y échapper est caractéristique de la conversation spontanée [Dessalles 1993]. Dans [Dessalles 1991] j'ai montré comment la satisfaction de contraintes de cohérence (celles qui provoquent des étonnements) et de contraintes de préférence (liées aux désirs et aux indésirabilités) pouvait expliquer le phénomène de pertinence ainsi qu'une partie du déterminisme du comportement conversationnel. Dans une situation de communication spontanée, l'exigence de pertinence s'exerce de manière extrêmement forte. Une intervention non pertinente est immédiatement dénoncée (« Pourquoi dis-tu ça ? »). La conséquence de cette exigence se traduit par le fait que toute intervention à contenu propositionnel ne peut avoir que deux rôles : poser un problème (incohérence, indésirabilité par ex.) ou résoudre un tel problème. Il en résulte une dynamique d'oscillation entre la perception de problématiques et leur satisfaction. Il est tout à fait spectaculaire de retrouver cette dynamique dans les contenus de pensée, et de constater finalement que *chaque élément de pensée est correct du point de vue de l'exigence de pertinence*. Mettons cette alternance en évidence dans le cas du troisième extrait :

problème : le sujet attend, il n'a pas de livre, il va perdre son temps, ce qui est indésirable
solution : le sujet va tenter de déboguer son programme de tête
problème : déboguer un programme de tête est difficile
transition analogique : les personnages de *La diagonale du fou* peuvent jouer aux Echecs de tête
problème : Pourquoi étaient-ils brusquement devenus amis ?

Il est clair que l'alternance problèmes/solutions n'est pas la seule particularité que l'on peut relever dans les contenus de pensée. Mais ce que l'on peut observer par ailleurs vient renforcer l'analogie avec les conversations ! Lorsque le sujet pensant réussit à échapper aux contraintes logiques (cohérence et préférence) qu'il venait de percevoir (ou lorsqu'il renonce), son esprit vagabonde par analogie et par association, *exactement comme on passe d'un sujet de conversation à un autre*. Par exemple la pensée de la lionne qui part en abandonnant ses petits déclenche l'analogie avec la scène du film de Saura. Les règles qui régissent l'enchaînement des sujets de conversation [Vuchinich 1980] [Dessalles 1988] sont de même des mécanismes d'association et d'analogie, que l'on résume parfois en disant que les interlocuteurs s'interdisent de « sauter du coq à l'âne ». Encore une fois, il est spectaculaire que les mêmes contraintes de continuité associative opèrent dans les contenus du flux de pensée.

Il faut pour finir tempérer un peu le propos, en reconnaissant que le flux de pensée possède des caractéristiques absentes de la conversation. Par exemple, nous pensons aussi avec des images, nous créons mentalement des scènes (la lionne qui lèche ses petits), et en cela nous sommes plus près du rêve que du langage, comme S. Freud l'avait noté. De telles scènes doivent être décrites à l'aide de mots ou de mime au cours des conversations, alors qu'il suffit de les « voir » et de les « vivre » au cours du processus de pensée. Mais de telles différences ne sauraient masquer les ressemblances importantes que nous avons soulignées entre les deux processus.

4.3. L'influence de la verbalisation

Il nous faut revenir sur ce que j'ai appelé la thèse de l'artefact. Peut-on considérer que les trois caractéristiques que l'on trouve dans les comptes-rendus de pensée : (1) le caractère séquentiel, (2) la contrainte de pertinence logique qui se traduit par une oscillation entre problématiques et solutions, (3) les transitions associatives, soient toutes créées au moment de la verbalisation ? Le sujet aurait simplement l'illusion d'avoir eu des pensées ainsi ordonnées ? Si tel est le cas, on doit s'attendre à ce que ces trois propriétés soient attachées au langage, non à la pensée. Il est pourtant assez facile de vérifier que la propriété (2) (qui implique la propriété (1)) est rarement présente dans les verbalisations écrites. Il est quasiment impossible de la trouver dans les

romans, même si l'auteur nous donne accès au flux de pensée de son personnage. A titre d'illustration, voici un extrait de *Tu ne t'aime pas* [Sarraute 1989]. Ce roman est intéressant pour le sujet traité ici, puisqu'il présente le flux de pensée comme un dialogue entre les différentes personnalités qu'héberge un même individu :

- Comment font-ils pour se sentir si nets, si simples ?
 - Ils doivent s'y entraîner très tôt... ils y sont dès leur plus jeune âge puissamment aidés... Les mieux doués, les plus précoces se voient déjà eux-mêmes tels que tout le monde les voit : en bébés... puis en petits garçons, en fillettes, en garçons manqués...
 - Une fois qu'ils ont pris ce pli de se sentir tels qu'on les voit, ils le gardent toujours... à chaque étape de leur vie, ils se sentent être des femmes, des hommes...
 - Et rien que cela. De "vraies" femmes, de "vrais" hommes... le plus conformes possible aux modèles...
 - Oui "vrais" jusque dans leurs moindres gestes, dans leur voix, leurs intonations...
 - Et dans leurs passions, leurs affections...
- [Sarraute 1989, p. 31]

Un tel texte présente certainement des qualités littéraires et une grande finesse psychologique. Lorsqu'on l'analyse du point de vue de la pertinence, en revanche, on constate que les contenus propositionnels ne sont pas ancrés les uns sur les autres comme dans l'alternance problème/solution. La première intervention pose certes un problème, la première ligne de la deuxième intervention (Ils doivent s'y entraîner [...]) fournit une explication, mais ensuite c'en est terminé : aucune problématique n'est plus explicitée, tous les contenus semblent aller dans le même sens, celui d'une critique des personnalités simples et conformes. On ne constate pas l'alternance caractéristique des contenus de pensée. Cet extrait, malgré sa valeur, est typique à cet égard. Les dialogues et les contenus de pensée relatés dans les oeuvres de fiction respectent rarement les contraintes d'enchaînement pertinent des contenus propositionnels.

Cette illustration suggère le fait que les propriétés que l'on détecte dans les comptes-rendus verbaux du flux de pensée sont loin d'être automatiquement liées au processus de verbalisation. Il semble bien qu'il s'agisse là de propriétés intrinsèques du processus de pensée. On peut difficilement imaginer que les sujets soient capables d'introduire un enchaînement pertinent lors de la verbalisation, et qu'ils en soient incapables lorsqu'ils veulent représenter par écrit des pensées fictives. Ceci semble être un argument fort contre la thèse de l'artefact. Le témoignage verbal des flux de pensée trahit la présence d'une dynamique de satisfaction de contraintes : cette alternance problème/solution, alors que cette caractéristique est le plus souvent absente des productions textuelles.

5. La phylogénèse sociale de la pensée

La structure du flux de pensée semble ainsi, par plusieurs de ses aspects, être calquée sur celle de la communication spontanée. Répétons-le, ce fait devrait nous surprendre. Si le flux de pensée a un rôle cognitif important en permettant au sujet de mieux « traiter l'information », alors il n'y a pas *a priori* de raisons pour qu'il s'agisse d'un processus séquentiel, soumis aux exigences de la pertinence et de la continuité associative. En revanche, le déterminisme de ces trois propriétés (séquentialité, pertinence, continuité) peut être trouvé dans les contraintes biologiques liées à la communication sociale.

Dans [Dessalles 1995], je propose l'hypothèse selon laquelle les exigences de pertinence peuvent être justifiées d'un point de vue phylogénétique. Elles résulteraient d'un système d'échange asymétrique entre les membres d'un groupe social : la pertinence du locuteur est « utile » à l'auditeur parce qu'elle lui permet de repérer un problème nouveau ou de connaître la solution d'un problème déjà perçu. En retour, le locuteur pertinent se voit accorder un « bonus » en terme de statut. De cette manière, on peut parvenir à un système plausible qui rend la communication évolutivement stable. Un bon modèle de ce système d'échange asymétrique nous est fourni par l'échange d'information dans la communauté scientifique. Dans un tel contexte, on comprend que

l'exigence de pertinence soit imposée par le locuteur. Les locuteurs que l'on écoute préférentiellement, et qui sont ainsi récompensés en terme de statut, sont ceux dont on perçoit d'emblée l'utilité des propos. Ce schéma nous offre un scénario possible pour expliquer l'apparition des propriétés fondamentales de la communication.

Comment, en revanche, expliquer la présence de ces mêmes propriétés dans le processus de pensée ? Elles ne contribuent en rien à l'idée qu'on peut se faire de l'efficacité d'un tel processus. On peut bien imaginer que le brassage d'éléments d'informations disparates, un peu comme celui que l'on observe dans le rêve, pourrait être au moins aussi efficace pour le fonctionnement cognitif de l'individu. Prenons maintenant une comparaison technologique. Si l'on demande à un ingénieur d'imaginer à quoi un robot doit occuper son unité centrale quand il n'a rien à faire, l'ingénieur lui fera explorer des problèmes fortement combinatoires, par exemple des tâches de planification qu'il n'a pas le temps d'explorer suffisamment lorsqu'il est pris dans l'action. Le robot pourrait aussi faire du « réapprentissage » en analysant ses expériences d'un point de vue statistique, ce qui lui permettra de généraliser les exemples et les contre-exemples et de compiler certaines procédures. Il pourra aussi essayer de calculer de nouvelles représentations de son environnement, passant d'une représentation linéaire (chemins parcourus) à une carte bidimensionnelle, etc. Tous ces processus de pensée artificielle, que le robot peut faire sans se déplacer, peuvent être dans une certaine mesure parallèles. Pourquoi les humains pensent-ils séquentiellement ? Quelle est la fonction remplie par leur mécanisme de flux de pensée ?

Pour expliquer la présence des propriétés de séquentialité, de pertinence et de continuité dans le flux de pensée, sans qu'elles y soient *a priori* nécessaire, on peut émettre l'hypothèse suivante : phylogénétiquement, la pensée serait un sous-produit de la communication. Elle jouerait pour l'individu le même rôle positif que celui que joue la conversation pour les interlocuteurs. La conversation est utile aux interlocuteurs parce qu'elle leur permet de repérer des incohérences dans leurs croyances et des incompatibilités entre leurs désirs*. La pensée joue, c'est notre hypothèse, exactement ce même rôle pour l'individu : améliorer la compatibilité des croyances et des désirs. On comprend alors pourquoi ce mécanisme, logique à la base, est complètement séquentiel. Les algorithmes informatiques qui assurent la même fonction de débogage dans les bases de connaissances sont tous des algorithmes séquentiels. On comprend aussi comment un mécanisme de pensée calqué sur la communication a pu être retenu par la sélection naturelle. Si l'individu, en situation conversationnelle, est capable d'effectuer les deux phases de la production d'arguments pertinents : le repérage des problématiques et la recherche de solutions, il est facile d'imaginer qu'il lui soit devenu possible, étant seul, d'enchaîner ces deux phases. Si l'auto-conversation qui en résulte peut lui permettre d'améliorer la cohérence logique de ses connaissances et de ses motivations, son comportement (y compris son comportement conversationnel) ne pourra que s'en trouver amélioré. Il semble qu'en cette occasion, la nature ait fait ce qu'elle fait souvent : faire du neuf avec du vieux [Jacob 1970]. Le système obtenu n'est peut-être pas optimal. Il aurait peut-être été plus efficace d'utiliser notre temps d'inactivité à d'autres processus cognitifs. Simplement ce dispositif que nous appelons pensée consciente était disponible, et il a été retenu par la sélection.

Avec cette façon de voir, une partie importante de notre fonctionnement cognitif est une conséquence directe et en partie arbitraire de notre passé phylogénétique social. Les exigences de la vie sociale de type humain, structurée par le comportement langagier, ont « pénétré » dans notre esprit au point de nous doter d'un organe supplémentaire, la pensée. Nous sommes tellement sociaux que nous transportons la société avec nous, dans notre tête, sous la forme d'une conversation perpétuelle.

La façon traditionnelle de présenter le processus d'hominisation consiste à dire que les premiers hominidés sont devenus plus intelligents grâce à l'avantage que leur conférait l'utilisation

* Nous n'avons pas parlé ici des contraintes liées à la probabilité des événements, mais les conversations servent de même à réestimer les probabilités *a priori* attribuées aux événements [Dessalles 1993].

efficace d'outils. Ils se seraient mis à penser, puis, tout naturellement, ils auraient eu des idées à communiquer. Un tel scénario explique certes la ressemblance entre pensée et langage, mais il n'explique pas le déterminisme des propriétés communes (séquentialité, pertinence, continuité). Dans un tel cadre, l'adéquation de ces propriétés aux exigences de la communication altruiste [Dessalles 1995] tiendrait du miracle. La thèse défendue ici est exactement inverse : elle fait de la pensée un produit « dérivé » de la communication.

Si l'on accepte le fait que notre forme de communication est spécifique à notre espèce, alors on doit, si l'on suit la thèse de cet article, accepter le fait que notre forme de pensée est elle aussi spécifiquement humaine, et ne ressemble en rien aux processus cognitifs animaux*. Cette caractéristique particulière qui différencie notre fonctionnement cognitif de celui des animaux est isomorphe à une interaction de type social. Lorsque nous pensons, nous sommes (pour ainsi dire) comme le chimpanzé qui s'épouille lui-même : nous jouons deux rôles sociaux à la fois. On voit rarement un individu se serrer la main ou se saluer lui-même, et pourtant nous passons le plus clair de notre temps, lorsque nous pensons, à converser avec nous-mêmes en respectant soigneusement les contraintes liées à la communication sociale. La question de savoir si le même type d'organisation sociale pourrait exister chez des êtres qui n'auraient pas ce mode auto-conversationnel, qui stopperaient leur activité communicationnelle dès qu'ils sont seuls, reste ouverte.

Références

- Baddeley Alan (1992). Working memory: the interface between memory and cognition. *Journal of Cognitive Neuroscience* vol. 4 n° 3, 1992, pp. 281-288
- Braten Stein (1992). *Born with the other in mind - Infant's minds and social feelings (The Dialogic Mind)*. à paraître, 1992
- Caron-Pargue Josianne, Caron Jean (1989). Processus psycholinguistiques et analyse des verbalisations dans une tâche cognitive. *Archives de Psychologie* vol. 57, 1989, pp. 3-32
- Court Georges (1971). *Cours de linguistique 783*. CNTE - Texte série 1 - AAA - LGL 131, Vanves 1971
- Davidson Donald (1975). Thought and talk. In Guttenplan Samuel, *Mind and language : Wolfson College Lectures 1974*, Oxford University Press, Oxford 1975, pp. 7-23
- Dessalles Jean Louis (1988). *La logique des conversations quotidiennes*. Non publié, 1988
- Dessalles Jean Louis (1990). The simulation of conversations. In Kohonen Teuvo, Fogelman-Soulié Françoise, *COGNITIVA 90 - Proceedings of the Third Cognitiva Symposium (Madrid)*, North Holland, Amsterdam 1991, pp. 483-492
- Dessalles Jean Louis (1993). *Modèle cognitif de la communication spontanée, appliqué à l'apprentissage des concepts*. Thèse de doctorat, TELECOM-Paris 93E022, Paris 1993
- Dessalles Jean Louis (1995). Contraintes sur l'évolution naturelle de la communication. In Theraulaz Guy, Dessalles Jean-Louis, et al., *Evolution et organisation: hasard et contraintes dans la genèse des formes collectives*, 4èmes journées de Rochebrune, ENST 95-S-001, 1995, pp. 113-118
- Ericsson K. Anders, Simon Herbert A. (1980). Verbal reports as data. *Psychological Review* vol. 87 n° 3, 1980, pp. 215-253
- Fodor Jerry A. (1975). *The language of thought*. Harvard Univ. Press, Oxford 1975
- Freud Sigmund (1907). *L'interprétation des rêves*. PUF, Paris 1973
- Gould, Stephen Jay (1985). *The flamingo's smile*. Penguin Books.

* Ceci ne préjuge en rien de la capacité des animaux à « penser ». Simplement leur forme de pensée ne peut être que très différente.

- Jacob François (1970). *La logique du vivant - une histoire de l'hérédité*. Gallimard, Paris 1970
- Johnson-Laird P.N. (1983). *Mental Models*. Cambridge University Press, London 1983
- Klinger E. (1978). Modes of Normal Consciousness Flow. In Pope K.S., Singer J.L., *The Stream of Consciousness*, J.Wiley, 1978, pp. 225-257
- Minsky Marvin (1985). *La Société de l'Esprit*. InterEditions, Paris 1988
- Nisbett Richard E., DeCamp Wilson Timothy (1977). Telling more than we can know: verbal reports on mental processes. *Psychological Review* vol. 84 n° 3, 1977, pp. 231-259
- Piaget Jean (1923). *Le langage et la pensée chez l'enfant*. Delachaux & Niestlé, Neuchâtel 1988
- Piattelli- Massimo, Palmarini (1979). *Théories du langage - Théories de l'apprentissage*. Seuil, Paris 1979
- Pope Kenneth S. (1978). How gender, solitude, and posture influence the stream of consciousness. In Pope K.S., Singer J.L., *The Stream of Consciousness*, J.Wiley, 1978, pp. 259-299
- Sarraute, Nathalie (1989). *Tu ne t'aimes pas*. Paris : Gallimard.
- Sperber Dan, Wilson Deirdre (1986). *La pertinence*. Les Editions de Minuit, Paris 1989
- Vuchinich Samuel (1980). Logical Relations and Comprehension in Conversation. *Journal of Psycholinguistic Research* vol. 9 n° 5, 1980
- Vygotski L.S. (1962). *Pensée et langage*. Editions Sociales (Thought and Language, 1962), Paris 1985

Conscience de soi et conscience d'autrui dans les théories représentationnelles de l'esprit

Tiziana Zalla¹

¹ Crea - Ecole Polytechnique
1, Rue Descartes - 75005 Paris
Zalla@poly.polytechnique.fr

Résumé

Certains aspects de la conscience réflexive des individus humains semblent liés à la capacité d'attribuer des états mentaux à autrui. Cette aptitude suppose que l'individu dispose d'une théorie naïve de l'esprit qui lui permet de comprendre et de prévoir les états mentaux des autres individus, ainsi que les siens propres. Dans le cadre d'une conception modulaire de l'esprit, certains aspects de la conscience pourraient être considérés comme un produit d'un mécanisme cognitif sélectionné pour des raisons de nature sociale.

1. Introduction

Certains aspects de la conscience de soi sont liés à des capacités perceptives et proprioceptives de l'individu. D'autres, en revanche, sont liés à une psychologie naïve, à la capacité de s'attribuer à soi-même comme à d'autres des états mentaux. D'autres, enfin, sont liés à la structure de notre mémoire et aux représentations des événements qu'elle contient qui impliquent l'individu dans son rapport avec le monde. Pour comprendre ce qui, dans la conscience de soi, est en rapport avec le caractère social de l'espèce humaine, il faut commencer par constater que la conscience n'est pas un phénomène unitaire. Elle peut être caractérisée par un certain nombre de *modules* dont il est important de comprendre la structure et le développement ontogénétique et phylogénétique.

2. Le concept de modularité

Les termes "module" et "modularité", largement utilisés par les sciences cognitives, désignent un type d'architecture fonctionnelle de l'esprit-cerveau. Néanmoins, ces concepts présentent des modes d'application et des significations qui changent selon les auteurs et les théories de référence. F.J. Gall a été un des précurseurs de la conception de systèmes de mémoire multiples. Gall divisait ces facultés en domaines particuliers selon le contenu, par exemple, la musique, les mathématiques et supposait que chaque faculté spécialisée avait sa propre mémoire. J. Fodor (1983) a repris l'idée d'une multiplicité de fonctions, notamment la distinction entre facultés "horizontales" et "facultés verticales", en proposant d'opposer les "systèmes centraux" et les "systèmes modulaires". Les facultés "horizontales" sont identifiées avec des processus centraux de jugement, de raisonnement et de planification. Fodor définit la modularité par neuf propriétés, parmi lesquelles la spécificité de domaine, le caractère obligatoire de l'analyse, l'automatisme et la rapidité des systèmes d'input sont les attributs fondamentaux. Par spécificité de domaine, Fodor entend le fait que chaque analyseur d'input possède des mécanismes physiologiques distincts qui correspondent à des domaines distincts de stimulations proximales. Étant donné une certaine variété des systèmes perceptifs, seule

une classe restreinte de stimuli pourra activer chaque système computationnel. Le fonctionnement des modules est informationnellement cloisonné, ce qui signifie que les opérations des systèmes d'input ne sont pas influencées au cours de leur activité d'analyse par des informations rétroactives provenant du corps des connaissances d'arrière-plan du sujet. Seules les représentations finales des processus d'analyse d'input sont disponibles et accessibles aux processus cognitifs centraux et à la conscience. Le développement des structures neurologiques sous-jacentes aux modules suivent des schèmes spécifiques génétiquement déterminés et activés par la stimulation de l'environnement. Les systèmes d'input sont donc associés à une architecture neurale fixe. L'ontogenèse des systèmes d'input présente un rythme et une séquentialité caractéristiques. A la différence des modules, les processus centraux, correspondant aux facultés globales et holistiques généralement conscientes, traversent les domaines multiples de connaissances .

Dans le cadre d'une conception modulaire, je propose une révision de la conception fodorienne de la modularité fortement innéiste en essayant de l'intégrer avec une perspective développementale. Cet élargissement conceptuel permet d'appliquer ensuite le principe de modularité aux processus centraux.

Selon la théorie modulaire proposée par les neuropsychologues Moscovitch et Umiltà (1989), il existe trois types de systèmes modulaires: 1) le module simple de type fodorien; 2) les modules assemblés de façon innée et 3) les modules assemblés après l'expérience. Ces types de structures modulaires diffèrent quant à la complexité et à la composition. Les modules ont la propriété de constituer une seule unité fonctionnelle, même quand ils sont l'assemblage de plusieurs modules de base.

Les modules de type I correspondent aux modules de bases qui accomplissent une seule fonction simple, telle que la détection des caractéristiques sensorielles pour chaque modalité, la perception des couleurs, de la fréquence acoustique, du mouvement, de la profondeur, ou bien des informations sensorielles plus complexes et structurées, telles que la perception des visages et des expressions émotionnelles.

Les modules de type II sont constitué des modules assemblés de façon innée dont l'architecture fonctionnelle est fixée endogénétiquement et dont l'output est intégré par des processus non modulaires consacrés au traitement d'un groupe restreint de stimuli. Comme les modules de base, les modules de ce type ont un domaine plus étendu et complexe que le domaine des modules composants, mais restreint à une seule modalité.

La troisième organisation modulaire, les modules de type III, consiste en un assemblage de modules qui devient modulaire en fonction de l'apprentissage. Comme dans le cas de la lecture ou de la parole, la structure fonctionnelle est fixée par l'expérience. Les processus acquis et automatisés sont exécutés par une organisation modulaire construite par l'expérience, et l'attention est nécessaire pour assembler les sous-routines automatiques et les modules de type II.

Il est assez facile de comprendre la valeur adaptative des caractères obligatoires et rapides du traitement sensoriel de l'information. On ne réfléchit pas sur la présence d'un danger : la simple existence d'un danger active le système d'analyse des inputs qui sont capables de réflexes intelligents. Le même argument, basé sur des considérations évolutionnistes, envisage l'émergence de modules cognitifs de plus en plus complexes, caractérisés par un grain plus fin et une capacité inférentielle plus élaborée¹. Les modules de type II refléteraient un processus de modularisation progressive des systèmes cognitifs supérieurs au cours de l'évolution. La spécialisation fonctionnelle au sein des systèmes cognitifs centraux et l'apparition de dispositifs de résolution de problèmes plus rapides, informationnellement cloisonnés et à spécificité de domaine ont été décisifs pour une amélioration des performances et le succès adaptatif. Les modules de type III dont l'architecture fonctionnelle est fixée par apprentissage individuel, sont voués à faire face aux problèmes et aux contingences donnés par l'interaction avec l'environnement, tout en gardant en même temps une plasticité cérébrale.

¹ Cf. Cosmides et Tooby (1994) pour la discussion approfondie de cet argument.

3. Modularité et conscience

Sur la base de cette taxinomie des systèmes modulaires, je propose de considérer trois concepts différents de conscience de soi. D'abord, un premier concept de conscience minimale de soi qui est très proche à la notion de conscience phénoménale². A ce type de conscience correspondent des capacités perceptives, proprioceptives et kinesthésiques de traitement de l'information. La deuxième notion de conscience de soi, qui correspond à la conscience réflexive, requiert d'autres dispositifs cognitifs, notamment des capacités métareprésentationnelles et intentionnelles. Je propose enfin un troisième concept de conscience de soi distinct à la fois du concept "d'expérience phénoménale" et de celui de conscience réflexive. Par cette troisième notion de conscience de soi j'entends les connaissances structurées sous une forme narrative concernant soi-même. Ce dernier sens de conscience qui porte sur une idée plus sophistiquée du moi, implique une mémoire autobiographique et des capacités narratives (Cf. Dennett, 1992; Flanagan, 1992).

Les études chez l'enfant montrent que la conscience suit d'un point de vue ontogénétique des étapes de développement (Cf. Parker et Mitchell, 1994). D'un point de vue phylogénétique, une conscience liée à des capacités inférentielles qui portent sur la connaissance des autres semble être une forme de conscience exclusive des êtres humains (Cf. Premack et Woodruff, 1978).

Une forme minimale de conscience de soi repose sur ce que Head (1920) avait appelé "schéma du corps". La proprioception joue un rôle fondamental dans la formation du soi et le schéma corporel fournit le premier niveau d'abstraction pour la représentation du corps à partir de ses bases purement sensorielles et kinesthésiques. C'est le niveau auquel le sujet est capable de reconnaître les expériences, les sensations telles que la douleur ou le plaisir et les stimuli endogènes liés à l'effort musculaire associé à l'exécution des actions, et de les identifier comme des expériences subjectives et privées du moi. L'hypothèse d'un schéma corporel et donc d'une cognition spatiale spécifique pour les corps humains a trouvé un soutien empirique dans les syndromes d'autopagnosie, c'est-à-dire l'incapacité de situer les parties du corps l'une par rapport à l'autre, les capacités spatiales générales et sémantiques restant relativement intactes. La détection d'analogies entre expériences corporelles d'autrui apprises de manière perceptuelle et ses propres sensations proprioceptives est possible grâce à un même schéma corporel amodal permettant d'établir des équivalences entre représentations à modalités différentes (Cf. Reed et Farah, 1995). L'expérience visuelle des autres corps est donc appariée de manière innée avec ses propres sensations kinesthésiques. L'appariement des stimuli kinesthésiques, proprioceptifs et visuels est préalable à l'apparition chez le nouveau-né d'un mécanisme d'imitation et de la capacité de se reconnaître dans un miroir qui chez l'enfant se manifeste environs à l'âge de 18 mois. On pourrait avancer l'hypothèse selon laquelle ce type de conscience phénoménale est le résultat de processus tout à fait modulaires, au sens fodorien.

Un domaine de connaissances important pour la constitution d'une conscience de soi est le système de mémoire épisodique³. La distinction entre mémoire épisodique et mémoire sémantique est due à Tulving (1983; 1985). Sherry et Schacter (1987) ont donné des arguments évolutifs en faveur d'une conception non-unitaire de la mémoire. Leur hypothèse concernant l'existence de systèmes multiples de mémoire est basée sur le principe d'incompatibilité fonctionnelle. Ce principe se définit en termes de contraintes environnementales et de spécialisation adaptative. La différenciation et la spécialisation des capacités d'apprentissage, voire la complexité de l'organisation fonctionnelle des systèmes cognitifs, est le résultat d'une réponse adaptative à de nouveaux problèmes posés par l'environnement. Le système de mémoire est défini par Sherry et Schacter en termes de règles opératoires des mécanismes d'encodage, d'acquisition et de récupération. Quand on se réfère à des systèmes multiples de connaissance, on fait appel à des règles opérationnelles différentes. En ce sens, un domaine de mémoire ne se définit pas par le type

² Par "conscience phénoménale" on entend le caractère qualitatif, privé et subjectif de la conscience.

³ Ce système est constitué par l'hippocampe et les structures limbiques corrélées

d'information, mais par ses règles opératoires. La sélection naturelle a produit ainsi l'apparition de nouvelles capacités d'apprentissage et d'un système de mémoire capables de répondre à des problèmes particuliers, de même que chez les oiseaux l'empreinte et de l'apprentissage du chant correspondent à l'évolution de deux systèmes de mémoire spécialisés. Parmi les exemples de spécialisation et d'incompatibilité fonctionnelle, on peut mentionner la mémoire procédurale, sur laquelle se basent les capacités motrices et perceptives, et la mémoire épisodique. La mémoire procédurale est basée sur la détection et la préservation des invariants extraits des épisodes qui sont importants pour l'accomplissement d'un comportement automatisé. Quant à la mémoire épisodique, sa fonction principale est de préserver les détails perceptifs, contextuels, temporels et spatiaux des événements. Sherry et Schacter font donc l'hypothèse que des modifications évolutives et des adaptations diversifiées et préservées par une incompatibilité fonctionnelle se sont produites à l'intérieur d'un système de mémoire originellement unifié. La mémoire épisodique a émergé à la suite d'une pression sélective qui a produit une multiplicité de systèmes de mémoire adaptés à différents problèmes de l'environnement. La spécialisation des mécanismes d'apprentissage serait alors à l'origine d'une amélioration des performances et des capacités de résolution de problème.

L'importance de ce système de mémoire pour la conscience de soi réside dans le caractère contextuellement et temporellement intégré de ses représentations et dans la richesse des détails perceptifs, alors que la mémoire sémantique concerne une connaissance générale et abstraite des faits.

4. Conscience réflexive et théorie de l'esprit

Je propose de considérer un type conscience de soi qui est lié à l'apparition d'une théorie de l'esprit. Les processus mentaux de compréhension et de communication postulés par Grice (1957) sont de type métacognitif, c'est-à-dire définis en termes de capacités à se représenter des représentations primaires⁴. Si tu vois que j'ai une veste verte, tu as seulement une représentation de ce fait; mais si je sais que tu sais que j'ai une veste verte, alors il faut que je sois capable de me représenter non seulement le fait, mais aussi ta représentation mentale. La métareprésentation est un processus récursif, à savoir qu'une représentation peut se référer à une autre représentation qui, à son tour, se réfère à une autre. Les capacités inférentielles d'attribution d'états mentaux, tels que les croyances, les désirs, les intentions, requièrent des capacités métareprésentationnelles, à savoir un corps de connaissances sur la cognition et sur la motivation des comportements (Premack et Woodruff, 1978, Leslie, 1987; Whiten, 1991, Perner, 1991). On distingue donc les métareprésentations des représentations primaires, définies en termes de leur relation sémantique directe avec le monde.

Une théorie de l'esprit (*Theory of Mind*) développée, comme celle dont disposent les adultes, s'établit chez l'enfant entre quatre et cinq ans. A cet âge l'enfant est capable de raisonner sur la base de fausses croyances ainsi que le montre le test de Wimmer et Perner (1983). Avant, à l'âge de trois ans, la capacité de "faire semblant"⁵ (ainsi que d'autres jeux symboliques), qui est un premier signal de l'apparition d'une psychologie naïve, est encore absente. Néanmoins, les enfants de deux ans ont déjà acquis la capacité de se reconnaître dans un miroir (Cf. Woodruff et Premack, 1979; J.

⁴ D'après Gomez (1994), le nombre de représentations imbriquées dans la communication humaine, ce qu'il appelle "la boucle griceenne", demande un minimum de cinq états mentaux imbriqués de la part du locuteur et six de la part de l'auditeur. La théorie de la communication de H. P. Grice (1957) semble impliquer la combinaison d'auto-conscience - la capacité de réfléchir sur ses propres états mentaux internes - et une théorie de l'esprit, c'est-à-dire l'attribution d'états mentaux à autrui.

⁵ Dans la situation consistant à faire semblant de boire de l'eau d'un verre vide, le terme "métareprésentation" désigne une représentation interne d'une relation épistémique entre deux personnes, une situation réelle (le verre vide) et une situation imaginaire (le verre contient de l'eau).

Perner, 1991). Les termes auto-référentiels chez l'enfant n'apparaissent pas avant 19 mois; ils sont en revanche largement employés vers l'âge de 27 mois (Kagan, 1989, p.233). La relation entre la reconnaissance dans un miroir et la théorie de l'esprit est controversée. Les expériences originales faites avec miroir par Gallup (1977) montrent que les singes ont une image du corps alors que les macaques n'en ont pas - les chimpanzés en se regardant dans le miroir touchent la tâche rouge peinte sur leur front, tandis que les macaques ne la remarquent guère. Avant l'âge de 18 mois, l'enfant ne sait pas que l'image des ses apparences physiques reflétées par le miroir correspond à l'image proprioceptive interne de son corps; cette découverte est faite à partir de l'âge de 2 ans. Les enfants autistiques qui ne possèdent pas la capacité d'attribuer des croyances à autrui, montrent toutefois la capacité de se reconnaître dans un miroir (Baron-Cohen et al., 1985). Ces observations suggèrent donc que la reconnaissance de soi-même dans un miroir requiert des capacités cognitives distinctes de la "théorie de l'esprit". On peut supposer que la capacité de se reconnaître dans un miroir est une condition nécessaire, mais non suffisante pour le développement de la théorie de l'esprit. Premack et Woodruff (1978) affirment que les capacités de base de la théorie de l'esprit s'expriment par la capacité d'empathie ou par une sorte de simulation (de) ou d'identification (avec) l'autre. Selon cette interprétation, dite "théorie de la simulation", l'attribution de croyances, et d'états mentaux en général, ne requiert pas des capacités métareprésentationnelles, mais tout simplement la capacité de s'identifier avec la personne et d'adopter son point de vue. Selon les défenseurs de la "théorie de la théorie" (Gopnik et Wellman, 1994), par "théorie de l'esprit" il faut entendre un véritable domaine de connaissances sur la cognition et des capacités inférentielles (nomologico-déductives) sur la base desquelles on peut attribuer des états mentaux et prédire le comportement d'autrui. Dans ce contexte, les processus décrits par la théorie de l'esprit peuvent être considérés comme étant à l'origine à la fois d'une forme représentationnelle d'auto-conscience et de conscience d'autrui. Cette capacité s'exprime par la prise en compte de ses propres états internes et de ceux des autres dans le processus de résolution de problème, de planification consciente du comportement et dans la communication.

Le domaine de connaissances concernant l'attribution des états mentaux semble être distinct d'autres domaines de connaissances, tels que ceux de la physique et des nombres (Cf. Carey et Spelke, 1994). La différence cruciale entre ces deux types de domaines paraît être la notion d'être animé par opposition avec la notion d'objet physique (Cf. Wellman, 1990). M. Jeannerod (1994) a proposé une hypothèse portant sur le type de signaux qui permettent au sujet de se reconnaître comme agent de ses actions. Des signaux endogènes, proprioceptifs et kinesthésiques relatifs à l'effort physique qui accompagnent l'exécution et la représentation des actions seraient à l'origine de l'intentionnalité. Dans l'imagination motrice, la représentation reproduit les aspects subjectifs liés à l'acte musculaire, à savoir que le rythme cardiaque, la respiration et le taux de CO₂ dans le sang simulent la situation d'effort physique qui caractérise l'exécution réelle de l'action. En ce sens, les capacités intentionnelles seraient plutôt expliquées dans le cadre d'une théorie simulativiste.

La notion de conscience réflexive que j'ai adopté dans ce contexte requiert néanmoins des capacités inférentielles qui s'ajoutent aux stratégies de simulation. En outre, la conception réaliste des états intentionnels et représentationnels, exprimée par la "théorie de la théorie", permet d'adopter la même attitude vis-à-vis des états de conscience.

Dans ce qui suit, je voudrais discuter davantage la structure fonctionnelle de la théorie de l'esprit et la manière dont elle pourrait rendre compte du développement des fonctions conscientes. D'une manière générale, il me semble que les résultats issus de différentes disciplines montrent que la conscience, loin d'être un phénomène discret et unitaire, se définit par rapport à différentes capacités cognitives.

5. Un module pour la psychologie naïve ?

Selon Leslie (1987) et Fodor (1992) la compréhension de l'esprit et l'interprétation des comportements en termes de croyances et de désirs sont dus à un module inné qui consiste en des mécanismes permettant de créer des métareprésentations à partir de représentations primaires. Leslie (1987; 1991) soutient qu'un module spécialisé inné voué à la connaissance psychologique, qui est apparu au cours de l'évolution phylogénétique, permet l'acquisition des concepts et la construction des systèmes de représentations internes par la détection obligatoire de certains types d'inputs. Les troubles cognitifs des enfants autistiques serait essentiellement dus à un déficit des capacités métareprésentationnelles et donc de la théorie de l'esprit⁶. Leslie propose que ces changements sont dus à la maturation, à l'âge de 18 mois, d'un module initial de métareprésentation, suivi par l'émergence d'un deuxième module dédié à la théorie de l'esprit vers 3/4 ans. La théorie avancée par Leslie et Fodor postulent l'existence d'un répertoire riche de compétences innées qui s'exprime en particulier dans l'émergence du phénomène de "faire semblant" à un âge précoce (entre 18 mois et 2 ans). La compréhension de ce phénomène est, dans un sens, une forme élémentaire de compréhension des états mentaux d'autrui. Cela représente, selon les auteurs, l'état initial dans la formation d'un module de la théorie de l'esprit.

L'idée d'un module pour la théorie de l'esprit est toutefois controversée. Selon Gopnik et Wellman (1994, p.281) deux hypothèses sont à la base de la conception de Leslie et Fodor : la première est que la connaissance de l'esprit (le sien et celui d'autrui) est innée; la deuxième est que cette compétence dépend d'un système modulaire. Gopnik et Wellman acceptent la première hypothèse, mais ils refusent la deuxième. Ils admettent l'existence des structures initiales innées, d'un système initial de connaissances psychologiques qui rend possible le développement d'une théorie complète de l'esprit, mais ils soutiennent que la définition de module, telle que l'a donnée Fodor, oblige à concevoir la théorie de l'esprit en termes de structures cloisonnées, spécifiées de façon innée, dotées des processus obligatoires et rapides. La différence entre la théorie modulaire et la "théorie de la théorie", que Gopnik et Wellman proposent, réside essentiellement dans le rôle que l'expérience joue dans la formation de ces compétences chez l'enfant. La théorie de l'esprit change par l'accumulation de nouvelles données et par un remaniement interne de la théorie même par des processus dynamiques internes du système. Ce qui est inné selon la théorie de Gopnik et Wellman n'est qu'un état initial de la théorie psychologique. En s'appuyant sur le développement et sur les aspects dynamiques de la théorie, ils affirment que la capacité innée de métareprésentation joue un rôle dans le développement d'une théorie de l'esprit, mais celle-ci ne constitue pas un module. Selon ces auteurs, une flexibilité de ces structures de connaissance semble être nécessaire; ils soulignent que la possibilité de changer et de modifier ce type de représentations mentales caractérise ces structures conceptuelles. En revanche, une conception modulaire stricte n'admet pas que l'expérience puisse changer les représentations conceptuelles. Leslie attribue les changements observés à des processus externes à ce système de représentations, qui sont liés à la survenue de la maturation d'autres systèmes représentationnels et d'autres capacités cognitives.

Baron-Cohen, Leslie et Frith (1986) en menant des expériences sur les autistes, ont démontré que ceux-ci ne présentent des difficultés de compréhension que quand il s'agit de prendre en considération ce que les autres connaissent ou attendent, et non dans la compréhension de situations socio-comportementales et affectives. Ces données semblent écarter l'hypothèse selon laquelle l'autisme est un trouble de nature essentiellement affective. Les désordres affectifs, sociaux et communicatifs ne seraient que des effets secondaires d'un déficit cognitif de base touchant les capacités métareprésentationnelles. Ce fait montre aussi qu'il s'agit d'un système modulaire.

⁶Les enfants autistiques montrent une compréhension très limitée des situations impliquant des états mentaux; en revanche, ils montrent de bonnes performances dans des tâches impliquant la connaissance mécanique des événements (Baron-Cohen et al., 1986). La pauvreté du vocabulaire mental semble confirmer une telle hypothèse.

Pour ma part, je crois que si un système inférentiel d'attribution de croyances, de désirs, d'intentions doit admettre la capacité de modifier ses états internes, ce système ne peut pas être un module de type fodorien. Il devrait en fait rendre compte des changements internes au système de croyances, ainsi que de l'introduction et de l'élimination des anciennes, ce qui n'en fait pas un module informationnellement cloisonné, mais un système qui admet des développements conceptuels. Selon la taxinomie des organisations modulaires proposée ici, le système de la théorie de l'esprit peut être considéré comme un module de type III, à savoir que ce système fonctionne de manière modulaire par rapport à d'autres fonctions exécutives et de raisonnement. Il n'est pourtant pas modulaire au sens strict car sa structure fonctionnelle est fixée au cours de l'expérience.

6. L'aspect social de la conscience réflexive

Comme je l'ai déjà dit, selon la théorie de la simulation, la découverte que les corps des autres personnes sont aussi dotés d'un esprit se fait grâce à un raisonnement par analogie à partir de soi-même : nous apprenons par l'expérience personnelle que nos états mentaux causent nos comportements, et de ce fait nous dérivons que des comportements similaires chez les autres sont causés par états mentaux similaires. Un schéma corporel commun aux représentations de soi et d'autrui permet d'attribuer par empathie des intentions et des états mentaux à soi-même et à autrui, comme causes des comportements. La pratique constante de la simulation permet d'acquérir un ensemble de règles générales nécessaires à la prise de décision. Cette hypothèse semble être confirmée par les recherches sur la représentation du corps. En suivant Reed et Farah, (1995), il semble y avoir un schéma corporel qui permet d'apparier ses gestes et ses comportements avec ceux des autres dans un même cadre de référence. Selon Gopnik et Meltzoff (1994), la capacité innée d'appariement kinesthésique-visuel qui permet la reconnaissance de soi dans le miroir est à la base des formes tardives et plus évoluées de la théorie de l'esprit. Gopnik (1993) tout en soulignant le rôle de l'imitation sociale et d'une disposition innée, permettant à l'enfant d'apparier ses propres sensations corporelles internes avec les apparences physiques d'autrui, n'ignore pas l'existence des capacités de raisonnement très développées. Nombre d'expériences montrent qu'il y a un large parallélisme chez les enfants entre la compréhension de leurs propres états mentaux et la compréhension des états mentaux d'autrui. Gopnik et Meltzoff ont remarqué que beaucoup d'études concordent sur le fait que les aptitudes à comprendre les états psychologiques, pour soi-même et les autres, se développent parallèlement. Ils en concluent que la compréhension de ses propres états mentaux ne s'accomplit pas par la perception interne et directe, mais comme pour l'attribution d'états intentionnels à autrui, par des processus inférentiels. Le développement d'une théorie de l'esprit plus sophistiquée permet d'intégrer les connaissances provenant de l'expérience sociale et culturelle, ainsi que de fixer des limites entre le moi et les autres personnes.

A mon avis, tout en admettant l'importance de l'expérience personnelle et des capacités de simulation des états d'autrui, il paraît difficile d'expliquer les capacités intentionnelles et la communication sans avoir recours à un véritable répertoire conceptuel psychologique.

Baron-Cohen (1994) a présenté un modèle à plusieurs composants modulaires, dont un détecteur de mouvements auto-générés et un détecteur de la direction des yeux. Sans entrer dans les détails du modèle, ce qui est intéressant, c'est que le système cognitif qu'il décrit semble correspondre à un module de type III. Leslie (1991, p.74) affirme que le module de la théorie de l'esprit peut avoir une architecture interne complexe et modulaire, ce qui va à l'encontre de mon hypothèse.

Les réflexions sur l'existence d'un système modulaire de la psychologie naïve et sur les relations de ce système avec la conscience de soi, permettent d'émettre certaines hypothèses concernant l'origine d'un tel mécanisme. L'attribution d'états mentaux à autrui est une compétence d'une extrême importance pour la viabilité dans l'interaction sociale : c'est cette compétence qui permet de prévoir le comportement des autres ou de communiquer avec eux. La disposition à

développer des compétences psychologiques — c'est-à-dire un module de la théorie de son propre esprit et de celui d'autrui — a pu être sélectionnée par l'évolution de manière à permettre à l'individu de développer une communication intentionnelle⁷. Selon Leslie (1991, p.75), la compréhension de l'intention de communiquer est une compétence préalable à tout acte de communication. Car le contenu du message dans la communication n'est pas déterminé par l'acte même, mais par la compréhension de l'intention de l'agent. Les autistes, à qui cette compétence sociale fait défaut, sont inadaptés à la vie de groupe. On imagine aisément que cette compétence sociale, lorsqu'elle s'est présentée, a été retenue par la sélection naturelle, dans la mesure où elle confère un avantage décisif à celui qui en dispose. On peut donc proposer le scénario selon lequel la part essentielle de notre conscience réflexive, qui réside dans notre aptitude à nous attribuer à nous-mêmes des états mentaux, s'est développée parallèlement à des compétences sociales.

7. Conclusion

J'ai voulu distinguer entre trois concepts de conscience. Il y a un sens selon lequel nous sommes conscients de nos expériences. C'est une acception minimale de conscience qui ne requiert pas de concept du soi sophistiqué ni de capacités auto-référentielles. Le sentiment d'appartenance, caractéristique essentielle des expériences phénoménales, tient de la permanence des représentations proprioceptives. Un deuxième sens de conscience de soi est lié au développement d'un répertoire de connaissances psychologiques, dit "théorie de l'esprit". En ce sens-là, des capacités métareprésentationnelles permettant d'attribuer à nous-mêmes et à autrui des états intentionnels sont à l'origine d'une forme de conscience réflexive. Le troisième type conscience est lié à une connaissance de soi sous la forme narrative. Cette forme de conscience requiert l'évocation explicite d'un ensemble de connaissances concernant le vécu de la personne, à savoir un type de mémoire épisodique,

Pour comprendre quel type de pression sélective a rendu possible la propagation de la conscience de soi, on peut rechercher ses manifestations socio-écologiques dans plusieurs espèces et regarder si elles sont corrélées avec une structure ou une compétence particulière. L'évolution convergente de structures fonctionnellement équivalentes constituerait une argumentation forte en faveur de sa valeur adaptative. En ce sens-là, le troisième type de conscience de soi, "le soi narratif", pourrait n'être qu'un effet secondaire, une utilisation culturellement déterminée d'une structure cognitive sélectionnée pour résoudre des problèmes adaptatifs spécifiques à l'espèce. La large variété des types de problèmes posés par l'environnement est donc à l'origine chez l'homme d'une architecture cognitive fonctionnellement spécialisée, dans laquelle les capacités de métareprésentation jouent un rôle central à la fois dans l'interaction sociale, dans la communication et dans la constitution d'une conscience de soi qui atteint un haut degré de complexité chez l'homme grâce à l'existence d'une activité narrative et d'un système de mémoire épisodique.

Bibliographie

Baron-Cohen S., (1994) : "How to build a baby that can read mind:Cognitive mechanisms in mindreading", *Child Development*, 62, pp.460-467.

Baron-Cohen S., Leslie A. M., Frith U., (1985) : "Does the autistic child have a 'theory of mind'?", *Cognition*, 21, pp.37-46.

⁷Selon Leslie (1991), d'ailleurs, les mécanismes qui sous-tendent la capacité de "faire semblant", notamment, la théorie de l'esprit, le raisonnement contrefactuel et la communication intentionnelle auraient été sélectionnés au cours de l'évolution humaine.

Baron-Cohen S., Leslie A. M. et Frith U., (1986) : “Mechanical, behavioral and intentional understanding of picture stories in autistic children”, British Journal of Developmental Psychology, 4, pp.113-125.

Carey S. et Spelke E., (1994) : “Domain-specific knowledge and conceptual change” in L. A. Hirschfeld & S.A. Gelman (Eds), Hirschfeld L.A. et Gelman S.A. (Eds), Mapping the mind. Domain specificity in cognition and culture. Cambridge University Press, pp. 169-199.

Cosmides L. et Tooby J., (1994) : “Origins of domain specificity: The evolution of functional organization”, in L. A. Hirschfeld & S.A. Gelman (Eds), Hirschfeld L.A. et Gelman S.A. (Eds), Mapping the mind. Domain specificity in cognition and culture. Cambridge University Press, pp. 85-116.

Dennett D.C., (1992) : “The Self as a center of Narrative Gravity”, Self and Consciousness. Multiple Perspectives, K.S. Kessel & P.M. Cole, D.L. Johnson (Eds), LEA, pp.103-115.

Flanagan O., (1992) : Consciousness Reconsidered, MIT Press.

Fodor J.A., (1983) : The Modularity of Mind. An essay on Faculty Psychology, The MIT Press, Cambridge, Mass., and London, England; trad. fr. La modularité de l'esprit, Paris, Minuit, (1986).

Fodor J., (1992): “A theory of the child’s theory of mind”, Cognition, 44, pp.283-296.

Gallup G. G. Jr, (1994) : “Self-recognition : Research strategies and experimental design”, S.T. Parker, K. W. Mitchell, M. L. Boccia (Eds), Self-awareness in Animals and Humans. Developmental Perspectives, Cambridge University Press, pp. 35-50.

Gomez J. C., (1994) : “Mutual awareness in primate communication : A Gricean approach”, S.T. Parker, K. W. Mitchell, M. L. Boccia (Eds), Self-awareness in Animals and Humans. Developmental Perspectives, Cambridge University Press, pp. 61-80.

Gopnik A., (1993) : “How we know our minds: The illusion of first person knowledge on intentionality”, Behavioral and Brain Sciences, 16, pp.1-14.

Gopnik A. et Wellman H. M., (1994) : “The theory theory”, in L. A. Hirschfeld & S.A. Gelman (Eds), Hirschfeld L.A. et Gelman S.A. (Eds), Mapping the mind. Domain specificity in cognition and culture. Cambridge University Press, pp. 257-293.

Gopnik A. et Meltzoff A. N., (1994) : “Minds, bodies, and persons : Young children’s understanding of the self and others as reflected in imitation and theory of mind research”, S.T. Parker, K. W. Mitchell, M. L. Boccia (Eds), Self-awareness in Animals and Humans. Developmental Perspectives. Cambridge University Press, pp.166-186.

Grice H.P., (1957) : “Meaning”, Philosophical Review, 64, pp. 377-388.

Hughes C., Russell J., Robbins T. W., (1994) : “Evidence for executive dysfunction in autism”, Neuropsychologia, vol. 32, N. 4, pp.477-492.

Jeannerod M., (1994) : “The representing brain : Neural correlates of motor intention and imagery”, Behavioral and Brain Sciences, 17, pp.187-245.

Kagan J., (1989) : Unstable ideas : temperament, cognition, and self, Cambridge, MA: Harvard University press.

Leslie A.M. (1987) : “Pretense and Representation : The Origins of “theory of Mind””, Psychological Review, vol. 94, n.4, pp.412-426.

Leslie A. M. (1991) : “The Theory of Mind Impairment in Autism: Evidence for a Modular Mechanism of Development?”, Whiten A. (Ed), (1991) : Natural Theories of Mind, Basil Blackwell, pp. 63-78.

Moscovitch M., Umiltà C., (1989) : “Modularity and Neuropsychology : The organisation of attention and memory”. in Schwartz M. (Eds), Modular processes in dementia. Cambridge, MA: MIT/Bradford.

Parker S. T. et Mitchell R.W, (1994) : “Evolving self-consciousness”, S.T. Parker, K. W. Mitchell, M. L. Boccia (Eds), Self-awareness in Animals and Humans. Developmental Perspectives. Cambridge University Press, pp.413-427.

Perner J., (1991) : Understanding the representational mind, MIT Press, Cambridge Mass.

Premack D. et Woodruff G., (1978) : “Does the chimpanzee have a theory of mind?”, Behavioral and Brain Sciences, 1, pp.515-526.

Reed C. L. et Farah M. J., (1995) : “The Psychological Reality of Body Schema : A Test With Normal Participants”, Journal of Experimental Psychology : Human Perception and Performance, vol. 21, No. 2, pp. 334-343.

Tulving E., (1983) : Elements of episodic memory, Oxford. The Clarendon Press.

Tulving E., (1985) : “Memory and Consciousness”, Canadian Psychology, 26, pp.1-12.

Wimmer H. et Perner J., (1983) : “Beliefs about beliefs: Representation and constraining function of wrong beliefs in young children’s understanding of deception”, Cognition, 13, pp.103-128.

Whiten A. (Ed), (1991) : Natural Theories of Mind, Basil Blackwell.

Whiten et Perner J., (1991) : “Fundamental Issues in the Multidisciplinary Study of Mindreading”, Whiten A. (Ed), Natural Theories of Mind, Basil Blackwell, pp.1-18.

Les comportements inter-individuels des animaux sont-ils des comportements sociaux ?

Fabienne de Gaulejac et Alain Gallo

Laboratoire de Neurobiologie et Comportement

Université Paul Sabatier

118, route de Narbonne 31062 Toulouse Cedex

Tel : 61 55 69 12 Fax : 61 55 84 44 e-mail : gallo@mail.cict.fr

Résumé

L'éthologie qualifie généralement les regroupements d'animaux conspécifiques de sociétés animales dont les caractéristiques sont souvent implicitement assimilées à celles des sociétés humaines. A contrario, une approche psycho-éthologique, qui accepte le principe selon lequel, au delà de la continuité évolutive entre l'homme et les animaux, il existe une discontinuité fondamentale dans leurs modes de fonctionnement psychologique, et partant, dans leurs conduites respectives, permet de montrer l'originalité et la spécificité des collectivités animales en regard des sociétés humaines. En effet, un examen approfondi des mécanismes d'interaction à l'oeuvre dans les sociétés indique que les collectivités animales ne se plient pas aux conditions du social énoncées chez l'homme, de par l'inscription spatiale de leurs comportements.

Mots clés : social, collectif, légalité, localité, objectivité, vécu, psycho-éthologie

Dans la pratique éthologique, le comportement inter-individuel se définit comme une action dirigée vers un individu de même espèce, comportement dont la structure et l'organisation seraient différentes des comportements orientés vers des individus appartenant à d'autres espèces et, a fortiori, vers des objets inanimés. Cela revient ainsi à conférer au conspécifique, à l'autre, un statut particulier.

Il est possible de s'interroger sur l'existence de tels comportements, exclusivement dirigés vers des conspécifiques. En effet, on relève, dans la littérature éthologique, plusieurs observations de comportements dits sociaux orientés vers des objets inanimés (léchage d'une pierre dans la continuité des mouvements de toilette du congénère, tentative de nourrissage d'un oeuf, tentative d'accouplement avec un rocher...) et, inversement, des comportements habituellement produits sur des objets inanimés et appliqués à autrui (couchage sur un congénère, transport d'un congénère par une partie de son corps lors de la construction du nid...).

Mais à supposer même que l'on reconnaisse la nature particulière des interactions comportementales entre conspécifiques, comme semblent le montrer les phénomènes d'imitation, de communication..., pour décider de qualifier de social un collectif d'individus, encore faut-il s'assurer que certaines conditions fondamentales sont remplies. La distinction entre collectif et social dans le règne animal s'éclairerait si on pouvait la faire reposer sur un principe unitaire expliquant, par exemple, la différence de fonctionnement psychique entre un être susceptible d'élaborer des contraintes légales et de s'y soumettre et un être soumis à des contraintes purement empiriques. Dans un premier temps, en s'appuyant sur la conception ordinairement énoncée du social chez l'homme, on essaiera de placer en regard la collectivité animale pour dégager un principe général. Dans un second temps, on montrera que les comportements des animaux, parce qu'ils sont inscrits spatialement, ne se plient pas à ce principe et, en particulier, à la nature des contraintes qui jouent sur les comportements sociaux. En conclusion, on reviendra sur la conception du social chez l'homme en se demandant s'il ne s'agit pas d'un artefact réaliste.

1. La condition du social

Pour chacun des phénomènes que nous allons invoquer (l'inventaire ne se veut pas exhaustif), la mise en oeuvre du social dépend de l'exercice d'une capacité d'abstraction qui peut se traduire par l'usage du symbole, la représentation d'un alter ego doué de représentations, la faculté d'élaborer le concept d'un individu et l'élaboration de lois générales.

1.1. Les relations triadiques

Egalement appelées spécularité (Vullierme, 1988 ; Dupuy, 1992) ou conscience réflexive (Zalla, 1996). Pour Lacadee (1992) qui exprime la différence essentielle entre l'homme social et l'animal en termes de triade symbolique, la référence est clairement lacanienne. « *C'est l'autre du langage qui est responsable de l'altériorité de l'expérience humaine, qui a déjà agi bien avant l'interaction précoce mère-enfant en détachant le petit d'homme à naître du déterminisme présumé de son équipement biologique pour le produire comme un sujet. Le symbolique n'est pas une fonction qui se déduirait de l'interaction mère-enfant ou des compétences précoces du bébé, mais en tant que lieu de l'autre, il est ce champ de la détermination, de la spécification du bébé comme petit d'homme comme parlêtre (être parlé par l'autre)* ». La parole prononcée avant que n'existe l'être biologique s'impose donc comme constitutive de l'être. Cette théorie instaure donc une différence radicale de constitution entre l'homme et l'animal.

Selon Moscovici (1995) d'autre part, il existerait en psychologie sociale humaine un regard spécifique, se traduisant aussi par une lecture ternaire, et non binaire, des faits et des relations. Cette 'tiercéité', selon l'expression de Peirce (1956), suppose une constante médiation, via la relation fondamentale à autrui, dans les rapports qu'entretient le sujet avec la réalité (objet physique ou social, réel ou symbolique). Vidal (1987) souligne ainsi qu'en repérant la place qu'un tiers tient pour sa mère, l'enfant ne peut que repérer du même coup la distance qui le sépare d'elle. Cette distance, sans doute plus symbolique que réelle, suscite chez l'enfant diverses tentatives pour la réduire, sur le seul registre où elle peut effectivement l'être, celui du symbolique, du désir. L'espace serait donc médiatisé symboliquement chez l'humain. Au contraire, lorsqu'il veut réduire la distance qui le sépare de sa mère, le jeune animal le fait dans le seul registre dont il dispose, l'espace physique selon Vidal. Nous verrons ultérieurement si l'on ne peut pas qualifier autrement l'espace pour l'animal. Mais en tout état de cause, il ne serait pas capable de se représenter l'objet d'un conspécifique, le compagnon, le lieu ou quelque autre objet que cet autre investit, c'est-à-dire qu'il serait dans l'incapacité d'élaborer une 'theory of mind' (Premack & Woodruff, 1978 ; Perner, 1991). « *L'animal n'a pas affaire à l'autre de l'autre* » (Vidal, 1994). Ainsi, l'attention que manifeste un enfant (dès six mois) pour fixer du regard le même objet que sa mère n'a-t-elle pas pu être mise en évidence chez l'animal, pas même chez les primates (Bruner, 1983 ; Bard & Vauclair, 1984) : bien qu'une récente expérience menée par Povinelli et ses collaborateurs (1990) conclut en la possibilité d'attribution de savoirs à autrui par des chimpanzés, la méthodologie employée a été jugée critiquable et l'interprétation anthropomorphique ; les travaux de Cheney et Seyfarth (1990) sur l'attribution de savoirs à leurs congénères par des macaques révèlent leur absence de connaissance des états mentaux de ces derniers : dans une situation-test, les auteurs placent une mère et son jeune dans un couloir menant à une arène et les deux animaux assistent à la position de nourriture dans l'arène et simultanément à l'entrée, dans cette même arène, d'un animalier muni d'un filet de capture, dans l'autre situation-test, seule la mère peut voir la scène. Les résultats indiquent que les mères ne réagissent pas différemment vis-à-vis de leurs jeunes dans l'une et l'autre des conditions expérimentales ; elles les laissent aller, montrant ainsi qu'elles ne connaissent pas leurs états mentaux, à savoir qu'ils ne sont pas renseignés sur la présence de l'animalier et ne connaissent donc pas le risque encouru. Cependant, il faut garder quelque prudence en ce qui

concerne les singes anthropoïdes car il existe actuellement une profusion de travaux contradictoires sur le sujet. On peut néanmoins se demander si les structures sociales complexes, plus ou moins fortement hiérarchisées et relativement stables, souvent décrites chez les vertébrés et invertébrés, n'existent pas simplement au niveau de la description que fait l'observateur des comportements inter-individuels d'animaux en groupe. Pour Vidal (1992), « *la question essentielle est de savoir si les animaux, impliqués dans les relations avec les autres membres de leur groupe social, peuvent eux-mêmes élaborer quelque représentation (...) de ces mêmes relations* » ou bien s'il ne s'agit que de l'activité de l'observateur. Les auteurs s'accordent généralement pour reconnaître qu'il est peu probable qu'un animal, nouvellement introduit dans un groupe déjà constitué, prenne en compte les relations hiérarchiques établies entre ses congénères pour trouver sa propre place au sein du groupe, de sorte qu'ayant, par exemple, dominé un individu X et ayant été dominé par Y, il ne pourrait inférer qu'il a quelque chance de dominer les individus dominés par X, et quelque risque d'être dominé par les dominants sur Y (cf. Kummer, 1968 chez les Babouins). On citera cependant la position de Zayan (1987) selon lequel il est vraisemblable que la transitivité perceptive existe chez les gallinacés. Le même phénomène pourrait exister chez les primates qui, expérimentalement, réalisent des inférences transitives en passation individuelle (non sociales). Cette conception n'est pas encore expérimentée. Certains rappelleront aussi, comme critique à la conception dyadique des relations interindividuelles animales, des expériences comme celles de Cheney & Seyfarth (1980) sur la reconnaissance vocale des singes vervets : alors que les jeunes sont hors de vue de leurs mères respectives regroupées entre elles, les auteurs diffusent, à partir d'un haut-parleur dissimulé, les vocalisations préalablement enregistrées de l'un d'entre eux. Une analyse de la conduite des femelles non concernées par ces vocalisations indique qu'elles regardent en direction de la mère impliquée quand celle-ci ne se tourne pas vers le haut-parleur. Les auteurs concluent en la capacité de ces femelles à faire la relation entre l'une d'entre elles et son jeune. De cette « *démonstration d'une véritable catégorisation des individus en termes d'appartenance filiative* » (Vauclair, 1990), on pourrait effectivement conclure en la présence, pour cette espèce, de relations triadiques puisque ces primates sont, semble-t-il, en mesure de discerner l'autre de leurs congénères. Cependant, les auteurs reconnaissent que les données disponibles sur la cognition sociale sont encore trop peu nombreuses et trop contradictoires pour aboutir à des conclusions solides sur la formation ou non de catégories sociales abstraites chez les primates.

1.2. La reconnaissance individuelle

La reconnaissance individuelle est un préalable à toute reconnaissance sociale. En effet, la connaissance et la maîtrise, par les animaux, de relations du type liens d'affiliation ou échelle des hiérarchies sociales, de même que leur capacité de coopération, de coordination..., si elles sont effectives, impliquent nécessairement une aptitude de reconnaissance individuelle des congénères. Le principe des expériences consiste à faire discriminer à un animal des diapositives représentant un individu particulier parmi d'autres. Une fois la différenciation réalisée, de nouvelles diapositives de ce même individu sont présentées. Les oiseaux sont capables de généraliser ce qu'ils ont appris (Trillmich, 1976 chez des perruches ; Ryan, 1982 chez des gallinacés) ainsi que les primates (Dasser, 1987). Cependant, de nombreuses observations révèlent une liaison entre sphères de fonctionnement et congénère, telle que celui-ci ne garderait pas le même sens selon les fonctions. Ces observations posent le problème de la permanence de la signification du conspécifique chez l'animal : par exemple le jeune de goéland, objet de tous les soins de la part de ses parents, mais qui peut, en certaines circonstances, être pris pour une proie par eux (cf. infra) ; celles de mâles chimpanzés adultes, dont le comportement d'épouillage envers des jeunes est généralement considéré comme une marque d'affection, et qui, lorsqu'ils sont en colère, saisissent indifféremment des branches ou un jeune et les frappent à même le sol. A cette 'polysémie' de l'autre chez l'animal s'ajoutent les nombreuses observations de Lorenz (1970) sur le défaut de reconnaissance du compagnon (compagnon-enfant, parent, conjoint ou social), chez les oiseaux de

différentes espèces : pour les petits du héron bihoreau, par exemple, le parent n'est pas un animal déterminé, mais quelque chose qui se pose sur le bord du nid et s'y comporte d'une manière spécifique, en exécutant, par exemple, une parade de salutation ; qu'un parent pénètre dans le nid sans cette parade, il sera violemment repoussé, mais qu'une main humaine l'imite et elle sera immédiatement investie par les jeunes, au point d'être défendue contre le véritable parent arrivant au nid dans une position menaçante à son égard. Des études approfondies montrent d'autre part que la présence de jeunes étrangers dans une couvée est généralement rapidement découverte par les parents, non pas parce qu'ils auraient une connaissance personnelle de tous leurs jeunes, mais parce que l'intrus se comporte souvent différemment des autres. La disparition d'un ou plusieurs poussins n'est en revanche jamais remarquée par les parents, la présence d'un seul suffisant généralement pour qu'aucune différence de comportement soit observée et ce, même chez des espèces nidifuges dont la taille des couvées est abondante. Plus encore, la mort d'une couvée entière de choucas provoque leur retrait du nid par les parents, mais ceux-ci continuent à couvrir le nid plusieurs jours durant. Des cygnes tuberculé peuvent attaquer leur conjoint lorsque ce dernier fouille du bec le fond de l'eau parce qu'ils ne voient pas sa tête et le prennent pour un cygne étranger. Dans le même ordre d'idées, chez certains oiseaux, la formation des couples indique qu'ils ne reconnaissent même pas les partenaires de sexes différents : le mâle exécute sa parade devant n'importe quel congénère et les femelles, de même que les mâles de force moindre, y répondent par la fuite, tandis que les mâles de force égale y répondent par une conduite analogue ; dans le premier cas de figure, une poursuite a lieu, suivie d'un accouplement, quel que soit le sexe du partenaire, dans l'autre un combat s'engage. Le mâle considère donc comme une femelle tout congénère ne provoquant pas sa réaction de combat. Enfin, un oiseau imprégné à l'homme, ne le reconnaît que lorsque tous les signes émanant de lui correspondent à son image habituelle. Lorenz explique ce phénomène par une 'qualité de complexité', selon laquelle l'animal ne divise pas sa perception d'ensemble en propriétés isolées : que l'homme auquel il est habitué se vête d'un long manteau ou d'un chapeau, et il se montrera craintif à son égard. Et lorsque l'auteur observe enfin un cas de reconnaissance d'un individu de leur colonie, après plusieurs mois d'absence, par des choucas, il note aussitôt que cette reconnaissance n'a lieu qu'à un endroit déterminé, l'emplacement de la nichée, et à un moment donné, uniquement en période de reproduction (cf. infra).

1.3. La communication

Havelange (1994) considère que, chez l'homme, l'organisation sociale est fondamentale tant comme condition de possibilité que comme conséquence de la communication. Chez l'animal, selon Cherry (1961), « *la communication sociale est le processus par lequel le comportement d'un individu affecte celui des autres* ». Un certain accord semble établi sur l'existence d'un système de communication animale. « *Cependant, il serait faux de dire que les animaux parlent* » (Stewart et al., 1994). Si les anthropoïdes sont capables, par apprentissage conditionné, de retenir une liste de mots, substitués abstraits dont la fonction est de remplacer les objets ou événements qu'ils désignent, les éthologistes ne leur reconnaissent pas l'usage d'une structure langagière ni d'une articulation syntaxique des signes (Sebeok, 1974 ; Terrace, 1985) : par exemple, l'analyse détaillée des conditions et du contexte de la production de 'phrases' construites par différents chimpanzés n'a pas permis de repérer des combinaisons régulières de signes. Or ce sont ces règles syntaxiques qui permettent l'attribution d'un sens aux phrases prononcées. Plus fondamentalement, les comportements de différents animaux peuvent être parfaitement coordonnés, sans pour autant qu'ils relèvent d'une communication sociale. Deux individus ajustent leurs actions dans la réalisation d'une tâche commune. Cette coordination observée est le plus souvent interprétée sous une forme téléologique, chacun des partenaires ayant l'intention de coopérer. Cette intention ne peut se concrétiser que grâce à la communication et la différenciation des points de vue. Mais cette définition des opérations qui président à la réalisation de comportements coordonnés (cf. niveaux 3 -coordination- et 4 -collaboration- dans la classification de Boesch & Boesch, 1989) ne s'applique

pas nécessairement à tous les animaux. Même quand les actions sont remarquablement coordonnées, cela n'implique pas l'existence d'un centre de décision intelligent : chez l'insecte social, la construction du nid ne relève pas d'une représentation ni d'un effort de coordination, mais résulte d'informations élémentaires distribuées sur tous les organismes concernés. Ainsi, selon Kraft (1994), les araignées coloniales coopèrent parce qu'elles s'ignorent, dans la construction de leur toile, pourtant remarquablement structurée. Par ailleurs, la définition de Cherry (1961) n'est pas satisfaisante car le comportement d'un individu peut affecter celui des autres sans qu'il soit nécessaire de se référer à un phénomène de communication sociale. En biologie, l'extension de la notion de communication est telle qu'elle dépasse largement le cadre de la communication intentionnelle. Nous défendons l'idée que l'usage de cette notion devrait être réservée à des phénomènes intentionnels. De plus, l'existence d'une communication intentionnelle suppose que l'animal comprenne l'intention de son partenaire (cf. 'theory of mind'). Dans la communication vocale, les sons émis par les animaux sont considérés, au moins partiellement, comme des signaux d'événements. Ainsi, les singes Vervets émettent trois sortes de cris d'alarme très stéréotypés, pour prévenir, semble-t-il, leurs congénères (Strushaker, 1967) : l'un est annonciateur de l'événement 'aigle', un autre de l'événement 'serpent', et un troisième de l'événement 'léopard'. Les primatologues discutent de la dimension sémantique et du caractère intentionnel de ces cris d'alarme (cf. de Lannoy, 1987, par exemple). Cependant, nous pensons qu'il faut revenir sur la définition de départ du son ou du cri comme substitut de l'événement, définition qui conduit presque fatalement à considérer que l'animal qui entend ce cri lui donne son sens et s'attend à la venue de l'événement, qui conduit aussi presque fatalement à décider que l'animal qui pousse le cri a l'intention de communiquer et de prévenir ses congénères. Il faudrait éprouver l'idée suivante qui, à notre avis, modifierait considérablement le regard porté sur le phénomène de communication dite sociale : le cri qui, pour nous dénonce l'aigle, est, pour le vervet, l'aigle tombant du ciel. Méthodologiquement, on devrait démontrer que cette relation d'identité ne suffit pas à expliquer le phénomène pour pouvoir parler de sémantisation des cris. En utilisant le vocabulaire de Thom, nous pourrions dire que les émissions sonores en usage dans la collectivité sont des prégnances, et non des vecteurs de prégnances, des substituts de l'événement.

1.4. Les règles sociales

Enfin, pour ce qui est de l'existence de règles chez l'animal, on cite généralement l'existence de transmissions culturelles (ouverture de bouteilles de lait par les mésanges bleues ; lavage des patates douces chez les macaques du Japon...). Les derniers travaux montrent qu'il s'agit en fait de découvertes individuelles multiples (Zentall & Galef, 1988). Si rôle du congénère il y a dans l'apprentissage, il s'agirait davantage de facilitation sociale (l'attention étant attirée sur un objet ou un lieu particulier, par exemple) que d'imitation (apprentissage à propos d'un comportement et non d'un stimulus) ou de véritable acte pédagogique (Boesch & Boesch, 1989). Pour ce qui est des règles dites morales comme les inhibitions d'agression entre prédateurs conspécifiques, l'opinion qui prévaut aujourd'hui semble être celle qu'avait développé Darwin (1871) : « *l'homme est l'être unique auquel on puisse avec certitude reconnaître la faculté morale ; ce fait constitue la plus grande de toutes les distinctions que l'on puisse faire entre lui et les animaux* ». Rien ne permet d'affirmer qu'un animal dispose des notions de bien et de mal et se conduise selon quelque jugement de valeur, de justice ou d'équité.

Dans le domaine des Sciences humaines, les comportements inter-individuels au sein d'une collectivité peuvent être qualifiés de sociaux, et non pas seulement de collectifs, parce qu'ils dépendent de facteurs comme les institutions, les lois, les règles, les normes, les valeurs (éthique, esthétique, politique...) et les tabous (prohibition de l'inceste, par exemple). Exception faites des normes et des valeurs (variables en fonction des classes sociales...), ces facteurs ou contraintes vécus par une société humaine présentent la particularité d'être valables pour tous les individus appartenant à la même société, quelles que soient les circonstances, c'est-à-dire d'être objectivés.

C'est cette nature particulière et spécifique de ces contraintes qui autorise à parler de social chez l'homme, même si leur contenu peut être remis en cause. Qu'en est-il chez l'animal ? Les comportements des individus les uns envers les autres ne dépendent pas de tels facteurs. Nous proposons de qualifier les contraintes de strictement locales et empiriques dans le cas d'une collectivité et d'essentiellement légales dans le cas d'une société. La différence est analogue à celle entre signal motivé et signe arbitraire : ainsi le symbole, au sens de de Saussure (1916), ou l'indice ont des rapports avec ce qu'ils sont censés représenter, contrairement au signe arbitraire. Notons que cette référence à la distinction entre signal motivé et signe arbitraire n'est pas pertinente si les contraintes légales sont considérées comme une pure formalisation des contraintes empiriques (les gènes tiennent en laisse la culture). Quoi qu'il en soit, qu'il y ait jeu de contraintes et effet de modulation du comportement individuel en groupe ne signifie donc pas que l'on puisse nécessairement parler de social.

2. L'inscription spatiale des comportements animaux

S'il est exact que le social implique différentes formes d'abstraction, une des méthodes qui pourrait permettre d'extrapoler, éventuellement, leur existence dans les collectivités d'animaux consiste en un examen de la nature profonde du comportement animal.

Nous avons réalisé une série d'expériences sur des meutes de loups captifs, mettant en évidence une inscription spatiale des comportements individuels et inter-individuels (de Gaulejac et al., 1995b sous presse, 1995c et 1995d soumis pour publication). Par exemple, la priorité d'accès à la source alimentaire (considérée dans la littérature éco-éthologique comme un bon indice de l'organisation sociale du groupe), mais également la fréquence des approches de la nourriture, la fréquence des consommations et l'ordre dans lequel les animaux commencent à manger varient en fonction du lieu de dépôt de la nourriture : en situation habituelle de nourrissage, la nourriture étant déposée dans une zone de l'enclos très fréquentée par les sub-adultes des deux sexes, ce sont ces individus (et non le couple ou le mâle reproducteur, comme on aurait pu s'y attendre), qui profitent le plus de la situation selon les indices retenus ; en revanche, lorsque la nourriture est déposée dans une zone très fréquentée par les adultes - ou bien par les plus jeunes - ce sont alors ces individus qui bénéficient des conditions maximales de consommation. De la même façon, cette inscription spatiale des comportements inter-individuels régit les réponses des meutes de loups à l'introduction et au déplacement d'objets biologiquement neutres (naturels comme artificiels) dans leur domaine de vie. L'histoire des comportements de ces animaux serait donc spatialisée. Cette caractéristique, spécifiquement animale, peut être rapprochée des effets de contexte : un stimulus apparaissant généralement dans une situation particulière, laquelle joue un rôle de fond non neutre, ne provoquera pas la même réponse du sujet lors du changement de situation.

Nous proposons alors d'opérer un renversement de la conception habituelle (Bon, 1991 ; Vidal, 1992) selon laquelle l'occupation et l'utilisation de l'espace par les animaux est déterminée par l'attachement dit social entre congénères et leur organisation définie comme sociale. Selon Washburn (1978), « *le fait social constitue un niveau d'élucidation des phénomènes au même titre que les mécanismes physiologiques ou neurobiologiques* ». Chez les mammifères, c'est la femelle qui nourrit sa progéniture. Il en résulte une unité sociale stable, constituée par la dyade mère-jeune. Pour Wilson (1975), la socialisation, initiée par cette relation mère-jeune, consiste en la somme des expériences sociales qui vont modifier le développement des individus, influencer leurs futures tendances sociales et, en conséquence, leur répartition spatiale. En montrant que l'organisation de l'espace, de même que la dynamique des déplacements importent davantage que les liens dits sociaux dans la vie d'une population de moutons, Lawrence et Wood-Gush (1988) illustrent ce renversement de la formulation, de même que Waser et Jones (1983) qui considèrent la philopatrie comme un attachement à un site plutôt qu'à des conspécifiques.

Pour ce qui est précisément de l'ontogenèse de cet investissement spatial, l'attachement en est le premier moment et ce que nous nommons communément sociabilité trouverait son origine non pas dans cette relation à la mère en tant qu'être social, mais dans la relation à la mère comme lieu. Initialement, tous les éléments retenus par le jeune (la forme, l'odeur et les cris de la mère (Novak, 1991), toute figure (paroi d'un dispositif, mangeoire...) présente dans l'environnement au cours de la période précoce (Salzen, 1967 ; Vidal, 1976)) constitueraient, pour le jeune, ce que nous nommons la 'mère'. Le rapprochement des jeunes entre eux et la formation de crèches correspondraient ensuite à un phénomène ontogénétique commun du monde propre de chaque individu (en termes d'investissement-désinvestissement de formes saillantes) et pourrait s'expliquer non pas en termes de ségrégation inter-âge due à une différence de maturité sociale, mais en termes de partage du même espace vécu, de la même hiérarchie de prégnances (chez le singe Rhésus, Colvin, 1983).

Si la formation observée de rassemblements plus ou moins grands d'animaux dépend de la cartographie des investissements individuels ('actospatialité'), les concepts de choix, de décision de se regrouper ne sont peut-être pas pertinents. Et si la collectivité observée n'est que la résultante de la coaction de différents paramètres caractérisant les individus à un moment donné et dans un contexte particulier, des notions comme celles de social, de sociabilité, de socialisation... ne sont plus pertinentes et cela, même si le congénère s'inscrit dans la hiérarchie de prégnances d'un individu. Si ce contexte change, les rapports entre individus se modifient souvent. En témoigne la variabilité observée dans la hiérarchie de dominance des loups captifs en fonction de leur lieu de présence. De plus, chez ces animaux captifs, on note une grande discrétion dans les interactions comportementales (peu d'attitudes de menace, d'intimidation, de monopolisation ou de combats...) à l'occasion des nourrissages, généralement dispensés en un lieu défini. Inversement, à l'état sauvage, après poursuite et mise à mort de la proie par la meute, la prise alimentaire est réalisée en un lieu chaque fois différent. Tous les enregistrements montrent alors que les comportements interindividuels des loups présentent une intensité particulièrement importante et chaque fois renouvelée. Outre l'hypothèse de l'excitation due à la capture d'une proie vivante, il est possible d'avancer celle d'une compétition pour l'investissement de la zone de consommation (zone généralement vierge de toute signification antérieure à celle de la proie). Dans le même ordre d'idée, Oates (1976) signale que les échanges agonistiques habituels relevés entre différents groupes de colobes noir et blanc cessent lorsque ces singes arboricoles quittent la canopée pour se nourrir autour des zones humides. Par conséquent, si les regroupements présentent effectivement des avantages observables incontestables, par exemple vis à vis de la prédation, ce caractère adaptatif émergerait de la tendance certes commune, mais individualiste, des animaux à investir un même espace au même moment.

Le fait de vivre en groupe, ou simplement le rapprochement temporaire entre individus, serait donc essentiellement déterminé par l'occupation d'un même espace et la congruence de ce qui est prégnant pour chacun. On peut ainsi comprendre l'attitude du mouflon corse mâle qui ne s'accouple qu'avec des femelles présentes en un certain lieu malgré la présence d'autres femelles sur son chemin. Il n'existerait donc pas, chez l'animal, de rencontre véritable sans partage psychologique du même espace et la proximité physique ne saurait, à elle seule, signifier ce partage : des individus proches les uns des autres peuvent parfaitement investir différemment le même emplacement. Wickler (1976), ainsi que Freeman et Grossman (1992), par exemple, soulignent le risque qu'il y aurait à inférer une inter-attraction, un lien ou même une familiarité entre des individus en se basant sur leur agrégation.

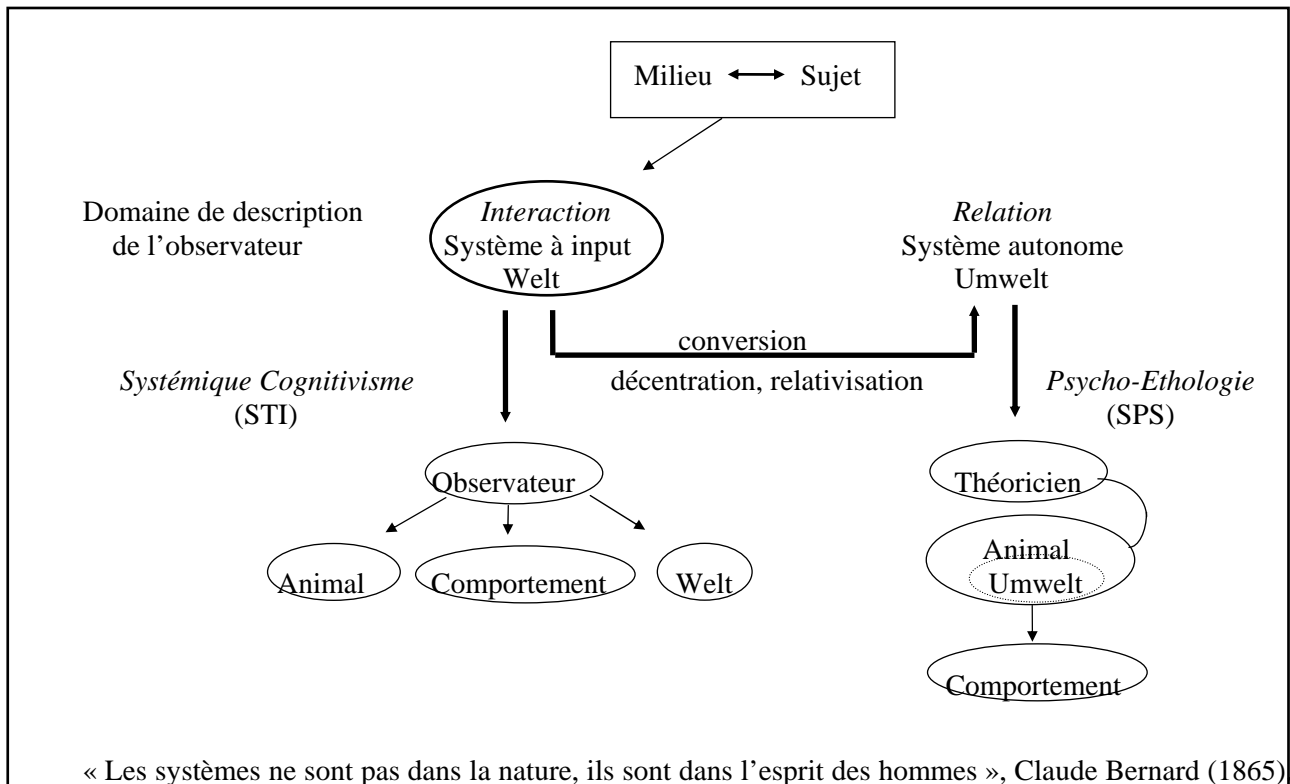
Dans ce contexte théorique, il est ainsi proposé que le rapport des animaux à leurs congénères dépende d'un processus fondamental d'investissement et non pas que ces rapports régissent leurs conduites en certains lieux. Cependant, il ne nous semble pas non plus que l'autre prenne généralement sens à partir de l'espace, qu'il émerge de la prégnance (valence) du lieu, au

même titre qu'un objet neutre, naturel ou artificiel (cf. de Gaulejac et al., 1995c, 1995d) et ce, parce que l'autre est un objet biologiquement signifiant. Nous avons montré, en effet, l'impact de la valeur biologique des objets-tests dans la réaction des loups à leur déplacement dans l'enclos : alors qu'un objet biologiquement neutre (bassine, pierre...) provoque l'intérêt des animaux lorsqu'il est placé dans une zone investie et son désintérêt lorsqu'il est déplacé vers une zone non investie (ou l'inverse), dans le cas où l'on dépose la nourriture non pas à l'endroit habituel mais dans une zone non investie, on observe au contraire, au bout d'un certain temps, qu'elle finit par les attirer temporairement. On peut cependant s'interroger sur les conduites infanticides de différentes espèces animales, « *reconnues comme un comportement naturel et non nécessairement pathologique* » (Hausfater & Hrdy, 1984). On connaît l'exemple des Goélands qui, lorsque le jeune est au nid, l'entourent de tous les soins 'maternels' mais qui, lorsqu'il est accidentellement trouvé en dehors du nid, le considèrent comme une proie au même titre qu'une autre (Baerends, 19XX). Si plus tard, le jeune goéland quitte naturellement le nid, il n'est pas la cible des attaques de ses parents. Autre exemple d'infanticide souvent constaté mais rarement décrit dans la littérature éthologique, celui des louves à l'état naturel et en captivité : c'est par exemple le cas de la louve reproductrice d'une des meutes que nous avons étudiées, laquelle a tué sa portée après que celle-ci ait été déplacée par les soigneurs, hors de sa présence, et sans que les jeunes aient été imprégnés de l'odeur de l'homme ; à cause de travaux dans l'enclos et par souci de non abandon par la mère, les louveteaux avaient été déplacés de la tanière vers un endroit très fréquenté par la femelle ; malheureusement, il s'agissait de sa zone de consommation alimentaire. Sans trop insister, je rapporterai pourtant l'observation détaillée, faite par McLeod (1990), d'une louve multipare : bien qu'ayant elle-même une propre portée au moment des faits, elle a tué celle d'une autre louve qui avait mis bas dans une autre tanière : le dispositif vidéo installé dans cette dernière tanière a permis de mettre en évidence le comportement ambivalent de la femelle, interrompant sa conduite infanticide par des épisodes de comportement 'maternel'. Malgré l'issue fâcheuse de cette scène, il nous semble, qu'en certains courts instants, la grossesse des louveteaux a donné sens de tanière à ce lieu qui ne l'avait pas, *a priori*, pour cette femelle. Si les exemples précédents, comme de nombreux autres, laissent à penser que les jeunes ne sont compris comme tels par leurs parents que lorsqu'ils se trouvent dans le nid, la tanière..., qui elle-même est investie du comportement dit maternel à un moment donné de l'existence des parents, ce dernier cas suggère que la grossesse des jeunes peut perturber, voire modifier en certains cas, la conduite infanticide. Chez les rapaces, par exemple, il arrive que le premier oisillon de la couvée soit tué par ses parents, le temps que la grossesse adéquate s'instaure (observation personnelle). Notons enfin que dans tous les cas rapportés d'infanticide commis par des louves, les jeunes ne sont que rarement dévorés après avoir été tués (malgré leur présence dans la zone alimentaire, par exemple), ce qui laisse à penser qu'ils ne sont complètement pas assimilés à de la nourriture.

En résumé, dans la littérature éthologique, les différentes interprétations des phénomènes d'interaction sociale ne sont pas toujours concordantes. Certaines se réfèrent au paradigme cognitiviste (cf. les récents travaux de Rumbaugh sur le langage des Bonobos), d'autres, au contraire, en arguant du canon de Morgan, renvoient au paradigme behaviouriste et les discussions restent ouvertes. Quoi qu'il en soit, si on accepte ces phénomènes comme conditions du social, il semble que l'intégration de la parole de l'autre et l'adhésion à un ensemble de règles, normes, valeurs, préceptes..., traduisant une certaine conception culturelle et morale (Darwin, 1871 ; Kummer, 1978 ; Vidal, 1992) de l'individu, ne caractérisent que la société humaine. La collectivité animale semble, elle, pouvoir être définie par la soumission à des exigences empiriques, des contraintes locales et ponctuelles (ici et maintenant). Il arrive que l'éthologiste comme le profane confonde ces notions de collectif et social, parce que les contraintes empiriques et les lois générales ont la même incidence : elles définissent ce que l'individu peut faire et ce qu'il ne peut pas faire. Il en résulte, pour l'observateur d'interactions comportementales entre différents animaux, une démarcation apparente entre ce qu'il est possible et impossible de faire avec (ou contre) les autres

dans des régions déterminées de l'espace. Mais cette limitation ne proviendrait pas d'une règle sociale et ne serait pas l'expression de l'existence d'une organisation sociale parce qu'elle ne régit pas les individus où qu'ils soient.

En conclusion, il est aisé de comprendre qu'en psychologie sociale humaine, selon une perspective classique, via le processus de socialisation - qui permet au sujet d'acquérir « *en relation avec d'autres personnes, les connaissances, les compétences, les normes et les valeurs avec lesquelles elle agira comme membre d'une société donnée* » (Monteil, 1991) - et la socialité résultante, concernant les comportements associatifs et coopératifs (Armitage, 1988), le sujet est considéré comme un être qui traite et organise l'information : les interactions entre individus sont appréhendées comme des inputs et des outputs. Dans ce contexte de Système de Traitement de l'Information (STI), les relations sociales relèveraient d'une structure paradigmatique où le spectacle offert aux yeux de l'observateur est susceptible d'être découpé en divers éléments (individu, congénère, environnement), éléments analysés en envisageant la part respective des parties et leurs interactions (cf. schéma). L'alternative à cette conception est d'en rester à la conception d'un collectif où chaque individu est défini comme un Système de Production de Significations (SPS) : contrairement à l'orientation cognitiviste représentationnelle, cette conception pose la cognition, non pas comme une représentation du monde extérieur pré-déterminé, mais comme une action productive, à savoir l'émergence d'un monde propre.



Dans cette optique, le chercheur ne part pas d'observations parallèles d'éléments différents du Welt, mais de la construction des règles de fonctionnement du monde propre (von Uexküll, 1956) de l'animal. Ce monde psychologique propre, monde-pour-l'être que l'on étudie, fonctionnerait comme un système autonome dans lequel n'existeraient que des réalités subjectives. Au niveau empirique, ce qui se décrit d'une certaine manière dans le Welt, par exemple la rencontre entre deux individus, est donc à transcrire en termes subjectifs. La question dans ce contexte est : « qu'est ce l'autre en l'animal ? ».

L'homme comme l'animal peut être appréhendé comme un système autonome, un SPS. La construction active du monde propre n'est pas une caractéristique proprement animale. Mais contrairement à l'animal, l'homme disposerait d'une faculté lui permettant de construire un espace neutre (espace vide) où situer ses projections, c'est-à-dire son monde en général (Welt) et ses conspécifiques en particulier. Ainsi, la perturbation est transformée en message existant à distance, c'est-à-dire en information. L'homme se libérerait alors de la fascination de l'autre par des processus d'abstraction (cf. condition du social), par exemple, en lui donnant un nom (cf. Thom, 1988). Ce phénomène est fondateur du fonctionnement en termes de STI. En toute logique, on ne peut parler de social que dans un tel contexte théorique autorisant la prise en compte du comportement de l'autre comme n'appartenant qu'à l'autre. Le paradigme du SPS, au contraire, ne peut que renvoyer à la notion de collectif d'individus. Ainsi l'homme fonctionne comme STI en croyant avoir échappé à l'autre et à l'espace comme vécus. Il prétend donner à l'autre ce qu'il appelle une 'autonomie' en énonçant qu'il supprime la sienne ('hétéronomie'). On peut se demander s'il ne s'agit pas seulement d'un artefact de construction réaliste. Auquel cas, le social n'existerait pas plus chez l'homme que chez l'animal.

Références

Armitage, K.B. (1988). Resources and social organization of ground-dwelling squirrels. In C.N. Slobodchikoff, *The ecology of social behavior*. San Diego : Academic Press, 131-155.

Bard, K.A. & Vauclair, J. (1984). The communicative context of object manipulation in Ape and Human adult-infant pairs. *Journal of Human Evolution*, 13, 181-190.

Boesch, C. & Boesch, H. (1989). Hunting behavior of wild chimpanzees in the Taï national park. *American Journal of Physical Anthropology*, 78, 547-573.

Bon, R. (1991). *Trajectoires sociales chez le Mouflon de Corse : rôle des processus d'attachement dans la sociogenèse*. Thèse de 3ème cycle, Toulouse : Université Paul Sabatier.

Bruner, J.S. (1983). *Le développement de l'enfant : savoir faire, savoir dire*. Paris : PUF.

Cheney, D.L. & Seyfarth, R.M. (1980). Vocal recognition in free-ranging Vervets Monkeys. *Animal Behavior*, 28, 362-367.

Cheney, D.L. & Seyfarth, R.M. (1990). *How monkeys see the world. Inside the mind of another species*. Chicago : University of Chicago Press.

Colvin, J. (1983). Description of sibling and peer relationships among immature male Rhesus monkeys. In R.A. Hinde, *Primates social relationships. An integrated approach*. Oxford : Blackwell Sci. Publ., 20-27.

Darwin, C. (1871). *The descent of Man and selection in relation to sex*. Londres : J. Murray. Trad. fr. *La descendance de l'homme*. Paris : Reinwald, 1872-1874, réed. Bruxelles : Ed. Complexe 1981.

Dasser, V. (1987). Slides of group members as representations of real animals. *Ethology*, 76, 65-73.

Dupuy, J.P. (1992). *Logique des phénomènes collectifs*. Paris : Ellipse.

- de Gaulejac, F. & Gallo, A. (1995a). Des interactions entre l'animal et le monde à l'énaction d'un monde propre. In XXXX, INRA. G. Theraulaz & F. Spitz (Eds.). A paraître.
- de Gaulejac, F., Gallo, A. & Legendre, X. (1995b). Préliminaire à l'étude du lien entre rang ou rôle social et réaction à la nouveauté chez l'animal. *Bulletin SFECA*, sous presse.
- de Gaulejac, F., Gallo, A. & Le Pape, G. (1995c). Relationships Between Space Usage and Behavioural Reaction of Captive European Wolf (*Canis lupus lupus*) to new Biologically Neutral Objects. Soumis pour publication
- de Gaulejac, F., Gallo, A. & Le Pape, G. (1995d). Relationships Between Space Usage and Behavioural Reaction of Captive European Wolf (*Canis lupus lupus*) to novelty : Effect of the Biological Importance of the Test Objects. Soumis pour publication.
- de Lannoy, J.D. (1987). A propos des expériences mentales chez les animaux. In M. Siguan, *Comportement, cognition, conscience : la psychologie à la recherche de son objet*. Paris : PUF, 89-100.
- de Saussure, F. (1916). *Cours de linguistique générale*. Paris : Payot
- Freeman, M.C., & Grossman, G.D. (1992). Group foraging by a stream minnow: shoals or aggregation? *Animal Behaviour*, 44, 393-403.
- Gallo, A. & de Gaulejac, F. (1996). La condition animale. In B. Cyrulnik, XXX.
- Hausfater, G. & Hrdy, S.B. (1984). *Infanticide*. New York : Aldine Publishing Company.
- Havelange, V. (1994). Sciences cognitives et tradition sociologiques. *Revue Internationale de Systémique*, sous presse.
- Kraft, B. (1996). Les araignées coloniales coopèrent parce qu'elles s'ignorent. *Bulletin SFECA*, (à paraître).
- Kummer, H. (1968). *Social organization of Hamadryas Baboons*. Chicago : University of Chicago Press.
- Kummer, H. (1978). Analogs of morality among nonhuman Primates. In G.S. Stent, *Morality as a biological phenomenon*. Life research report 9, Dahlem Konferenzen, Berlin, Verlag Chemie, 35-52.
- Lacadee, P. (1992). Le bébé ou le petit d'homme. In Etudes du CEREDA, *L'autre du bébé*. Toulouse : Presses Universitaires du Mirail, 25-37.
- Lawrence, A.B., & Wodd-Gush, D.G.M. (1988). Home range behaviour and social organization of scottish Blackface sheep. *Journal of Applied Ecology*, 25, 25-40.
- Lorenz, K. (1970). *Essais sur le comportement animal et humain, les leçons de l'évolution de la théorie du comportement*. Paris : Seuil.
- McLeod, P.J. (1990). Infanticide by female wolves. *Canadian Journal of Zoology*, 68, 402-404.

- Monteil, J.M. (1991). Socialisation. In H. Bloch, R. Chemama, A. Gallo, P. Leconte, J.F. Le Ny, S. Moscovici, J. Postel & M. Reuchlin, *Grand Dictionnaire de la Psychologie*. Paris : Larousse, 729.
- Moscovici, S. (1995). *Psychologie sociale*. Paris : PUF.
- Novak, R. (1991). Senses involved in discrimination of merino ewes at close contact and from a distance by their newborn lambs. *Animal Behaviour*, 42, 357-366.
- Oates, J.F. (1976). The guereza and its food. In T. Clutton-Brock, *Primate Ecology*. London : Academic Press, 276-371.
- Peirce, C.S. (1956). Logic as semiotic : the theory of signs. In J. Buchler, *The philosophy of Peirce, selected writings*, 3^{ème} éd., Trad. fr. *Ecrits sur le signe*, Paris : Seuil.
- Perner, J. (1991). *Understanding the representational mind*. MIT Press, Cambridge Mass.
- Povinelli, D., Nelson, K.E. & Boysen, S.T. (1990). Inferences about guessing and knowing by chimpanzee, *Journal of Comparative Psychology*, 104, 203-210.
- Premack, D. & Woodruff, G. (1978). Does the chimpanzee have a theory of mind ?, *Behavioral and Brain Sciences*, 1, 515-526.
- Ryan, C.M.E. (1982). Concept formation and individual recognition in the domestic chicken. *Behav. Anim. Letters*, 2, 231-220.
- Salzen, E.A. (1967). Imprinting in birds and primates. *Behaviour*, 28, 232-253.
- Seboek, T.A. (1974). Comment un signal devient un signe. In E. Morin & M. Piatelli-Palmarini, *L'unité de l'homme I. Le primate et l'homme*. Paris : Seuil, 64-77.
- Stewart, J., Scheps, R. & Clement, P. (1995). Une interprétation biologique de l'interprétation : II la phylogénèse de l'interprétation. *Colloque de Cerisy*.
- Struhsaker, T. (1967). Auditory communication among Vervet Monkeys. In S.A. Altmann, *Social communication among primates*, Chicago : University of Chicago Press, 281-324.
- Terrace, H.S. (1995). In the beginning was the Name. *Journal of American Psychological Association*, 40, 1101-1028.
- Trillmich, F. (1976). Learning experiments on individual recognition in Budgerigars. *Z. Tierpsychology*, 41, 372-395.
- Vauclair, J. (1990). Les images mentales. *La Recherche*, 224, 1006-1014.
- Vauclair, J. (1996). *La cognition animale*. Paris : PUF, Que sais-je ? n°3061.
- Vidal, J.M. (1976). *Empreinte filiale et sexuelle : réflexions sur le processus d'attachement d'après une étude expérimentale sur le coq domestique*. Thèse de docteur d'état, Université de Rennes.
- Vidal, J.M. (1987). Naissances et images animales et humaines. *Bull. Off. Soc. Franç. Psycho. Prophyl. Obsét.*, 109, 7-16.

- Vidal, J.M. (1992). Evolution des psychismes et évolution des organismes. In P. Tort, *Darwinisme et société*. Paris : PUF, 519-548.
- Vida, J.M. (1994). 'Theory of mind' ou 'theory of love' ? Un éclairage à partir des symptômes autistiques. In R. Zazzo, *Croissance de l'enfant, Genèse de l'homme*. Paris : PUF, 143-151.
- Vullierme, J.L. (1988). Spécularité morale et spécularité anthropologique. In *Individu et justice sociale - autour de John Rawls*, 145-157.
- von Uexküll, J. (1956). *Mondes animaux et monde humain*, suivie de *Théorie de la signification*. Paris : Gonthier.
- Waser, P.M., & Jones, W.T. (1983). Natal phyloptry among solitary mammals. *Quarterly Review of Biology*, 58, 355-390.
- Washburn, S.L. (1978). Human behavior and the behavior of other animals. *American psychologist*, 33, 405-418.
- Wickler, W. (1976). The ethological analysis of attachment. Sociometric, motivational and sociophysiological aspects. *Z. Tierpsychol.*, 42, 12-28.
- Wilson, E.O. (1975). *Sociobiology. The new synthesis*. Cambridge : Belknap Press.
- Zalla, T. (1996). Conscience de soi et conscience d'autrui dans les théories représentationnelles de l'esprit. *Actes des Journées de Rochebrune 1996*, ENST-96S00x
- Zayan, R. (1987). Recognition between individuals indicated by aggression and dominance in pairs of domestic fowl. In R. Zayan & I.J.H. Duncan, *Cognitive aspects of behaviour in the domestic fowl*. Elsevier, 321-438..
- Zentall, T. & Galef, B.J. (1988). *Social learning : Psychological and biological perspectives*. Hillsdale, Lawrence Erlbaum Associates.

Le sens dans les processus interlocutoires, un observé ou un co-construit ?

Christian Brassac¹, John Stewart²

¹ Groupe de Recherches sur les Communications
Université de Nancy 2, B.P. 33-97. 54015 Nancy Cedex
E-mail : brassac@clsh.u-nancy.fr

² Institut Pasteur
29 rue du Docteur Roux. 75015 Paris
E-mail : stewart@pasteur.fr

Résumé

Les principaux modèles de la communication interhumaine s'inscrivent dans une perspective objectiviste. L'intercompréhension y est envisagée comme l'ajustement entre deux cognitions individuelles. Le traitement d'un énoncé est un mécanisme qu'un individu met en œuvre pour découvrir le sens que son interlocuteur intentionne de communiquer. Ce sens préexiste au traitement, que ce dernier soit un simple décodage ou un travail inférentiel. Cette façon de voir ressort du paradigme de la communicativité que Francis Jacques oppose à celui de la communicabilité.

Dans ce dernier, dont nous défendrons le bien fondé, l'intercompréhension est un processus de construction conjointe d'un sens qui se stabilise interactionnellement au cours du déroulement de l'interlocution. Considérée ainsi, l'intercompréhension est l'émergence, dans l'entre-deux conversationnel, d'une cognition distribuée. Notre proposition s'inscrit de la sorte dans une perspective constructiviste de la communication interhumaine.

Mots clés : Conversation, objectivisme, constructivisme, description, sens.

Introduction

Les travaux classiques consacrés à l'étude des échanges langagiers interhumains sont tous fondés sur un postulat de départ :

Lorsqu'un énoncé est produit par un des locuteurs, il est *traité* par l'auditeur.

Ce traitement est le fait du seul auditeur. Son objectif est de trouver, de découvrir, de détecter le contenu que le locuteur a pour intention de communiquer. Les modèles codiques et inférentiels de la communication ont pour ambition de rendre compte du mode de traitement mis en œuvre par l'auditeur. Qu'il soit de l'ordre du décodage ou du travail inférentiel, le traitement s'ancre sur l'énoncé produit.

Nous voudrions interroger cette thèse en exhibant le fait qu'elle présuppose l'existence de ce sens à communiquer. Pour ce faire, nous allons montrer dans un premier temps que ce présupposé est très largement répandu et ensuite qu'il est inutile à la modélisation de la communication langagière interhumaine. Dans un second temps nous affirmerons qu'il est utile de distinguer deux niveaux de description des phénomènes interlocutoires. Nous montrerons en quoi, selon nous, décrire "de

l'intérieur" les mécanismes d'intercompréhension en situation d'interaction langagière permet d'intégrer l'étude des processus cognitifs dans le paradigme constructiviste.

1. Le traitement d'un énoncé en conversation

1.1 La communicativité : un paradigme dominant

L'idée selon laquelle, en conversation, un locuteur parle dans l'intention de "faire passer" quelque chose et que l'auditeur traite cet énoncé pour trouver ce "quelque chose" est extrêmement répandue. Elle constitue le fonds du "sens commun", comme le révèle une analyse des métaphores que nous utilisons, en langue courante, pour parler à propos de la communication Reddy, [1979]. Elle est présente autant dans les modèles codiques que dans les modèles inférentiels de la communication. Et l'idée selon laquelle un sens bien défini préexiste à son expression linguistique est centrale à la conception "cartésienne" de l'esprit [Dennett, 1993].

Dans un article intitulé "La métaphore du tuyau", Reddy établit une liste de plusieurs centaines d'expressions courantes qui, toutes, reflètent une théorie spontanée très particulière concernant les processus à l'œuvre dans la communication langagière. Sommairement, selon cette théorie, l'émetteur possède une idée claire et précise ; il "encode" cette idée dans un message ; le message est physiquement transporté (par des ondes sonores, par téléphone/e-mail, etc.) au récepteur qui, dans un troisième temps "décode" le message. Une conséquence de cette théorie est que, si personne ne fait de bêtes erreurs, il n'y a pas de perte d'information dans la transmission, l'idée d'arrivée peut idéalement être isomorphe à l'idée de départ, et donc que la communication peut être parfaite. Les malentendus ou les quiproquos par exemple seraient donc pathologiques. Ce modèle est très proche d'une "ingénierie des télécommunications", et la théorie informationnelle de Shannon et Weaver [1949] l'exprime clairement.

Pour le modèle inférentiel, il ne faut pas se cacher que le calcul qui prend pour point de départ l'énoncé produit par un locuteur a, idéalement, pour point d'arrivée le sens que ce dernier a voulu communiquer. Autrement dit, l'interprétation est une opération dont le produit est un objet qui lui préexiste : le sens intentionné par le locuteur. Ce n'est pas parce que l'on est gricéen [Grice, 1979] ou adepte de la théorie de la pertinence de Sperber et Wilson [1985] que l'on échappe à cette vision des choses.

Le mécanisme est le suivant :

1. Le locuteur possède une certaine intention communicative,
2. Il produit un énoncé porteur d'un sens,
3. L'auditeur perçoit cet énoncé et le "traite",
4. A l'issue du traitement, il récupère, il trouve, il atteint le sens intentionné.

Le traitement est "bon" lorsqu'on observe une certaine relation entre le sens intentionné et le sens découvert. Remarquez qu'il n'est pas dit qu'il s'agit là du sens littéral ou non littéral de l'énoncé. Le travail peut concerner, sauf pour le modèle codique mais peu importe, les actes indirects, les implicatures conversationnelles, etc. L'implicite est évidemment dans la portée du traitement de ce sens préexistant. On dit alors que les interlocuteurs "se comprennent". Tout le monde est content : le locuteur a communiqué ostensivement un sens, l'auditeur a découvert ce qu'il fallait ; ils se sont compris.

Quel que soit le mode de traitement, cette façon de voir place au centre du dispositif d'intercompréhension l'intention de sens du locuteur, **pré-existante au travail d'interprétation-décodage de l'auditeur.**

Finalement, il est utile de dire explicitement que tous ces modèles (à l'exception de ceux d'entre eux qui sont fondés sur la notion d'"indication pour la construction du sens", [Raccah 1990])

reposent sur une conception "cartésienne" de l'esprit. Prenons l'exemple d'un acte de parole. Dennett a plaisamment illustré les difficultés de cette conception en explicitant ce qu'il a appelé le "modèle bureaucratique". Selon ce scénario, il y aurait l'esprit, installé dans le siège de son théâtre cartésien plus ou moins à l'instar du Président des Etats-Unis dans son bureau ovale. Au début du processus, l'esprit décide de parler. Il convoque alors le "bureau des formulations", qu'il charge de mettre sa pensée en paroles. Ce travail réalisé, le bureau des formulations convoque à son tour le "bureau opérationnel", chargé de traduire les mots en mouvements musculaires de l'appareil vocal. Les commandes sont alors envoyées aux muscles qui s'exécutent... et le son est produit. Toute la difficulté réside dans les modalités de la communication effective entre l'esprit et le premier "bureau", et des bureaux entre eux. Si l'esprit ne peut pas *déjà* parler, comment peut-il communiquer son idée aux bureaux ? Et s'il reste enfermé dans son bureau ovale, quelle garantie peut-il avoir que les sons qui sortent en fin de parcours correspondent le moins du monde à ce qu'il voulait dire ? On est en présence d'une régression à l'infini, avec des sous-interfaces indéfiniment emboîtées les unes dans les autres, particulièrement vicieuse.

La distinction, proposée par Francis Jacques [1995] entre communicativité et communicabilité est très précieuse à cet endroit. L'approche que nous venons d'évoquer s'inscrit dans le paradigme de la communicativité. Même s'il ne s'agit pas que de transmission, on a affaire à un phénomène d'intercompréhension conçu comme une suite de couples production-interprétation, asynchrone et asymétrique. Le locuteur est propriétaire d'un sens, explicite ou non, l'auditeur a pour tâche de le découvrir avant d'être, à son tour, propriétaire responsable de son interprétation dont il produit la trace langagière qui sera elle-même l'objet de traitement. On a là une chaîne de mécanismes cognitifs **individuels**. L'intercompréhension est considérée comme un processus d'adéquation entre les cognitions personnelles des interactants.

1.2 La communicabilité : un paradigme concurrent

On peut concevoir un paradigme de la communicabilité qui va à l'encontre de cette vision. Il place au centre de ses préoccupations une intercompréhension vue comme une co-construction processuelle d'un sens des énoncés successifs, produit par une **cognition distribuée**. Dans cette optique, on ne considère pas la conversation comme le lieu où s'organise l'adéquation de deux cognitions individuelles. Elle est conçue comme le lieu de **l'intercommunicabilité des cognitions** en présence (l'expression est d'Alain Trognon). L'intention de sens est une production conjointe, distribuée sur l'espace intersubjectif que constitue l'entre-deux conversationnel.

Afin de comprendre pourquoi il est possible de soutenir une telle position, nous ferons un détour qui pourrait apparaître comme étant d'ordre méthodologique, mais qui nous semble-t-il, est très important épistémologiquement. Il concerne le point de vue de l'observateur.

2. Décrire une conversation : quelle méthode ?

2.1 Comment décrire un phénomène [description₁ vs description₂]

La position qu'occupe un analyste d'interactions verbales de face à face est une position difficile. En mettant en œuvre un quelconque appareillage d'investigation des phénomènes qui surviennent dans l'échange parlé, le conversationnaliste (terme pris dans le sens général de chercheur s'intéressant à l'interlocution) est amené à observer, de l'extérieur, une suite de productions langagières transcrites. Sa première tâche consiste à décrire "ce qui se passe".

Il s'attache à rendre compte d'activités produites par une dyade d'individus humains. Ces derniers se sont parlé, mutuellement. Le corpus est l'ensemble des traces langagières des mécanismes cognitifs à l'œuvre dans la relation interlocutoire qui les réunit. Le travail de l'analyste dépend des catégories théoriques qu'il utilise pour mener à bien son travail de recherche. Ceci est bien évident.

Il est aussi important de souligner que cet analyste, en tant que participant régulièrement à des interlocutions, dans sa vie quotidienne, professionnelle ou autre, est un conversant. Bien sûr, il n'est pas un des producteurs des énoncés mis en séquences dans l'extrait conversationnel qu'il a devant les yeux mais il est un expert en la matière qu'il travaille. En cette qualité, il a une tendance à attribuer à des énoncés donnés des catégories données.

Intéressons nous à un corpus donné, commençant par l'énoncé [1]

[1] Sais-tu quelle est la distance entre ici et la maison ?

Soit l'analyse débutant par l'affirmation : "la première intervention de cette conversation est une question portant sur une distance". On a affaire là, typiquement, à une attribution qui accole de façon externe une catégorie à un énoncé. Ce geste est extrêmement significatif de la démarche d'analyse du chercheur. Il est en effet symptomatique d'un présupposé épistémologique très puissant et qui sous-tend une vision dominante de la communication interhumaine. On peut exprimer ce présupposé sous la forme suivante : le sens de cet énoncé existe, est le contenu de ce que veut signifier son locuteur et l'interlocuteur va, dans la suite le traiter. Et la suite de l'exposé consiste à présenter ce traitement, que ce dernier soit un simple décodage ou qu'il soit le résultat d'un mécanisme inférentiel plus ou moins compliqué.

Nous aimerions dans cet article nous inscrire en faux face à cette façon de faire. Nous le ferons car nous prétendons que le présupposé cité ci dessus est inadéquat pour une description correcte de l'intercompréhension qui survient dans l'échange parlé considéré.

Mais peut-être tout le problème est-il dans le terme **description**.

Une position éthologiste consiste à observer, d'un point de vue extérieur, un ensemble de comportements, animaux ou humains. Ce faisant, l'observation se dit à l'aide de catégories, appartenant à l'appareil cognitif de l'observateur, humain. Cette description₁ Stewart, [1995] n'est pas dépendante de ce que font les observés pour eux mêmes. Plus précisément, en décrivant les comportements des fourmis, le chercheur applique des catégories de chercheur et non pas de fourmis. Une description₂ consisterait à exhiber les mécanismes qui génèrent effectivement les comportements des fourmis. Il s'agirait de décrire les comportements en les termes des observés.

Les deux types de description sont complètement différents au niveau phénoménal. Par exemple, selon la description₁ on dira peut-être que chaque fourmi "suit une trajectoire" (c'est effectivement ce qu'on voit du point de vue de l'observateur). La description₂ s'attachera à identifier les bouclages sensori-moteurs (les stimulations sensorielles, la manière dont celles-ci sont utilisées pour guider les actions de la fourmi, et les conséquences de ces actions pour les sensations ultérieures...) qui génèrent effectivement ce comportement. Le but ultime d'un travail scientifique est de relier les deux niveaux de description par une modélisation qui *explique* la description₁ comme conséquence émergente des mécanismes identifiés au niveau 2. Il est important de noter que dans une perspective de ce type, il y a de multiples "descriptions₁" possibles, mais qui ne sont pas toutes équivalentes. En fonction de la description₂ retenue, certains aspects du comportement seront plus ou moins pertinents. Prenons un exemple, cette fois-ci d'un comportement linguistique.

Reddy [1979] a proposé une métaphore alternative à celle du "tuyau" pour comprendre la nature de ce qui se passe quand des êtres humains se parlent. Imaginons la situation d'un certain nombre de personnes naufragées qui ont échoué sur une île déserte. Imaginons en outre que chacun est dans une partie différente de l'île, séparé par des murs. Les murs empêchent les uns et les autres de se voir et de s'entendre, mais il existe quelques fentes dans lesquelles on peut glisser des signaux - brindilles d'herbe, poussières de terre, etc. Peut-on communiquer dans ce cas? Eh bien, peut-être que oui; mais il faut un acte de foi qu'il existe effectivement quelqu'un de l'autre côté du mur, et que

son expérience est suffisamment semblable à la sienne pour qu'il puisse *deviner* ce que les indices glissés peuvent vouloir dire. Ainsi faut-il des trésors d'imagination créatrice de part et d'autre, pour constituer les indices d'abord, et pour les interpréter ensuite. Dans ces conditions, on comprend aisément qu'un aspect très important de la communication consistera à vérifier si on se comprend ou pas. Et en effet, dans la conversation ordinaire en face à face, des phrases comme "tu vois ce que je veux dire?", "non je comprends pas, redis-le moi", ou alternativement "oui, oui, continue", etc., jouent un rôle très important. Et en effet, une grande partie de cette communication consiste en des signes non-verbaux - hochements de tête, froncement de sourcils, prosodie, des silences éloquents, etc. Selon cette théorie (description₂) concernant la nature du processus, ces signes non-verbaux constituent une partie essentielle de la communication linguistique, et une description₁ appropriée devra les inclure.

2.2 ... conversationnel ?

Lorsque le chercheur dit, en observant la conversation que le premier énoncé est une question portant sur une distance, il décrit₁ le phénomène. Pour parler plus techniquement, il attribue à cet énoncé le statut de directif dont le contenu propositionnel est la distance. Le point clé est le suivant. De quel droit un observateur peut-il affirmer qu'il s'agit là d'une question ? La réponse est la suivante, et elle est, nous semble-t-il inacceptable : parce que l'observateur l'interprète comme une question !! Ceci est inacceptable car, c'est son statut de conversant qui lui fait dire ceci et non sa position d'observateur. Plus, c'est sa façon de concevoir la communication qui le mène à assigner à [1] une fonction interrogative à propos de la distance. Et si les conversants, les vrais, ceux qui sont observés, ceux qui ont produit la séquence, n'étaient pas d'accord ?? Et si l'auditeur de [1] ne l'entendait pas comme une question ? Par exemple l'auditeur de [1] pourrait répliquer comme suit :

- [2'] Est-ce que je te pose des questions moi ?
- [2''] Oui bien sûr que je sais !
- [2'''] Arrête de me reprocher toujours mon ignorance
- [2'''''] Ecoute je ne suis pas géomètre
- [2'''''''] Cent trois mètres quarante.

Selon sa réplique, comment l'auditeur a-t-il compris est-ce le bon terme ?) l'énoncé [1] ? Seul l'examen de la suite de l'interaction peut permettre de réfléchir à la question.

En attribuant à [1] un statut de directif, le descripteur₁ ne rend pas compte du mécanisme qui génère effectivement l'intercompréhension. Plus, il impose son point de vue d'analyste sur ce corpus avant de vérifier sa pertinence analytique. L'analyse qu'il propose de la suite de l'échange dépend du fait qu'il, l'observateur, a assigné à [1] un statut de question.

Autrement dit, il ne décrit₂ pas le phénomène interlocutoire. Pour nous, décrire₂ l'agencement processuel de cette interlocution nécessite d'essayer de rendre compte de la façon dont les acteurs s'intercomprennent, et non pas comment l'interlocuteur comprend ce que l'observateur a compris comme question ! C'est dire que la description₂ a pour objet le déroulement de la relation interindividuelle et non les conduites individuelles des interactants.

Essayer d'expliquer comment un auditeur comprend le sens qu'a voulu communiquer le locuteur, c'est adopter une posture "description₁". Il s'agit d'une posture égologique, égophorique qui s'inscrit dans le paradigme de la communicativité.

Essayer d'expliquer comment la paire des interlocuteurs produit un sens conjointement, processuellement, c'est adopter la posture "description₂". Il s'agit d'une posture relatiologique, duophorique qui s'inscrit dans le paradigme de la communicabilité.

3. Vers une approche constructiviste de la conversation

3.1 L'intercompréhension comme processus [analyse d'un exemple]

Les nombreux travaux empiriques qui ont menés l'un d'entre nous, chercheur au Groupe de Recherches sur les Communications, à proposer un modèle de l'enchaînement conversationnel, nous conduisent à penser qu'effectivement, l'intercompréhension est un phénomène beaucoup plus subtil qu'une simple coïncidence de cognitions en co-présence. Nous avons exposé en plusieurs endroits en quoi et pourquoi on doit envisager l'intercompréhension comme un processus de co-construction du sens Trognon et Brassac, [1995, Brassac, 1994, Ghiglione et Trognon, 1994, par exemple]. L'idée principale que nous défendons est que le sens d'un énoncé n'est pas contrôlé par le locuteur, ni par l'auditeur. Il n'appartient ni à l'un ni à l'autre. Il émerge dans l'entre-deux. Il est construit conjointement par les deux interlocuteurs.

Il n'y a pas un sens qui est communiqué (le locuteur n'a pas une intention, celle de communiquer un sens). Il y a, peut-être, un sens qui est communicable. Plus sûrement, l'énoncé produit initialement est porteur d'un potentiel de sens. L'échange se déroule autour de l'actualisation d'un élément de ce potentiel, actualisation qui est le produit de la dyade, qui est l'œuvre conjointe des interactants.

Soit l'extrait conversationnel suivant [Dessalles, école de l'ARC de Bonas, 1995].

- E₁ Oh c'est marrant ! J'ai exactement ce tableau chez moi,
il a la même taille, il représente la même chose
L₁ On te l'a peut-être volé
E₂ Non le mien il est plus sombre, il est plus beau.

Voici clairement ce que nous aimerions montrer.

- E₁ est une forme linguistique, porteuse d'un potentiel de sens.
- L₁ est une proposition d'actualisation d'un des éléments de ce potentiel.
- Le couple (E₂, (L₁--->E₁)) est une prise de position, à **deux**, par rapport à cette proposition d'actualisation. Cette prise de position est négociable.
- S'il y a un sens il est en suspens, par exemple au temps t₃ de l'énonciation de L₂.

a. Il est clair que le locuteur E exprime sous forme langagière "quelque chose". L'énoncé E₁ est une forme linguistique, porteuse de sens. Nous ne nous permettrons pas dire ici qu'il s'agit d'une affirmation, de l'expression d'un regret, d'une inquiétude, d'une question ou autre action langagière. Nous dirons simplement que cet énoncé **est susceptible** de porter tout ceci, et peut-être autre chose que nous analystes "ne voyons pas". La question n'est sûrement pas de privilégier l'un ou l'autre de ces possibles.

b. En énonçant L₁, L propose d'actualiser un des sens du potentiel. Il s'agit d'un sens que l'on pourrait dire "au pied de la lettre". En effet en disant "Oh c'est marrant", "exactement" et "ce", E exprime son étonnement que **son** tableau soit là devant lui, dans cette vitrine. Pourquoi peut-on dire ceci ? Parce que L apporte une explication à cet étonnement : "c'est le tien, on te l'a volé". Il satisfait les conditions préparatoires de l'acte de langage expressif d'étonnement.

c. En énonçant E_2 , E invalide cette proposition de sens. En effet, il donne un élément supplémentaire qui prouve que "ce" tableau n'est pas le sien : c'est la tonalité. Par conséquent, il dit à son interlocuteur : il ne faut pas prendre ce que j'ai dit au pied de la lettre. Le couple ($E_2, (L_1 \rightarrow E_1)$) a pour fonction de suspendre, dans ce cas, le mécanisme d'ajustement des cognitions.

d. A cet instant, l'observateur ne peut en aucun cas dire quelque chose comme : E a voulu dire ceci, L ne l'a pas compris. Le processus est en cours, son produit fini (s'il existe) est encore indéterminé, il est négociable.

Remarquez que dire tout de go que L_1 est ironique, tombe sous la même critique : de quel droit l'observateur peut-il pénétrer dans le système cognitif de L pour y "voir" qu'il ne croit pas à ce qu'il dit en L_1 . La réponse sur la question du caractère ironique de ce "on te l'a peut-être volé" est à chercher dans la suite de l'interaction. Peut-être d'ailleurs ne pourra-t-on pas la résoudre.

L'extrait étant limité à trois tours de parole on ne peut "savoir" si ce phénomène d'intercompréhension se résoudra en un point de fixation sur *le* (?) sens en jeu ici. En tout état de cause, nous aimerions insister sur le point suivant. L'interlocution est un endroit où se déploie un flux de sens qui peut porter la mésentente, le malentendu et qui est marqué par une indétermination radicale (voir à ce sujet les pages très éclairantes de Pierre Livet [1994, chapitre 1 particulièrement]). Cette indétermination ne constitue en aucun cas un obstacle à l'intercompréhension. En effet, c'est précisément parce que le processus est producteur d'indétermination qu'il permet l'intercommunicabilité des cognitions. Le déroulement de la conversation se satisfait pleinement de la possibilité d'une simple entente provisoire. Il s'actualise même précisément sur ce provisoire. De ce point de vue, il est intéressant de se pencher sur le cas où les interlocuteurs se rendent mutuellement compte que jusqu'alors ils étaient en train de "surfer" sur un malentendu. Il se produit alors un hiatus dans le déroulement linéaire de la communication. Mais loin d'être le signe d'un échec (comme on aurait tendance à penser dans l'optique de la "métaphore du tuyau"), un tel événement est en fait la marque d'une communication *réussie*. En effet, se rendre compte que l'on croyait se comprendre, mais qu'on se trompait, signifie "en relief" que la communication peut être, sinon parfaite, néanmoins réelle. Ce point est important, car il montre ce en quoi le constructivisme n'est pas relativiste (car dans ce cas, si "n'importe quoi" était possible et tous les possibles équivalents, les malentendus n'existeraient pas ou bien, ce qui revient au même, tout serait malentendu).

3.2 Objectivisme et constructivisme en conversation

Autrement dit, si E voulait, intentionnait de communiquer une cognition, son expression linguistique, E_1 , est susceptible de le permettre ; la cognition en question est communicable, même si non communiquée. Mais E peut très bien se satisfaire de ne pas "avoir été entendu", pour autant que son interlocuteur coopère à une construction conjointe. L'important c'est la communicabilité et non la communicativité.

Selon nous il y a un danger majeur à parler de l'activité de compréhension réalisée par les sujets en interlocution. En effet on peut soutenir que lorsqu'un énoncé est produit par un des locuteurs, il n'est pas *compris* par l'auditeur. Ce dernier ne met pas en place une procédure de traitement qui le mènerait à découvrir ce que voulait dire son interlocuteur, qui le conduirait à exhiber un sens qui préexisterait à ce traitement.

Le travail de participation conjointe à l'émergence de sens ne porte pas sur un objet qu'il faut observer de l'extérieur. Ce travail est un façonnage, une sculpture, qui se réalise à deux, d'un objet

en perpétuel devenir : le flux de sens des expressions langagières produites par les conversants. Le nom de ce travail ? L'intercompréhension.

En fait, pour avoir une théorie complète de l'intercompréhension (et surtout une explication en termes d'une description₂), il nous faudra certainement une bonne théorie de ce que c'est que la compréhension tout court. La compréhension peut être *déclenchée* par un signal, linguistique ou autre (dans le cas de la perception, ce sera non pas une stimulation sensorielle simple, mais l'engagement dans une boucle sensori-motrice, cf. ci-dessus) ; mais quelle est la nature du processus qui est ainsi déclenché ? Nous croyons qu'elle n'est pas à chercher dans la structure du signal lui-même ; pour illustrer ce point, il sera intéressant de raconter un événement dans l'histoire de la biologie.

Dans les années trente, les embryologistes étaient très excités car ils avaient découvert un phénomène passionnant. Au cours des mouvements des tissus qui se produisent au long de l'embryogenèse, il se trouve qu'un tissu vient buter sur un autre - par exemple, la cornée de l'oeil touche l'ectoderme. Il se produit alors le phénomène dit d'"induction" : l'ectoderme touché par la cornée réagit en se différenciant pour former une rétine. La fonctionnalité de ce phénomène est évidente : c'est ainsi que la rétine se trouve *exactement* en face de la cornée. Plus généralement, l'induction est un mécanisme clé dans (l'auto-)organisation de l'embryogenèse. Etant donné l'importance de ce phénomène, les embryologistes étaient ravis de découvrir que le phénomène pouvait être reproduit *in vitro*, c'est à dire dans une situation expérimentale en boîte de Pétri. Spemann a obtenu un Prix Nobel pour ces premières découvertes ; aussitôt ce fut la ruée, car les embryologistes croyaient qu'il y aurait d'autres Prix Nobel à ramasser facilement en identifiant la ou les substances chimiques responsables du phénomène. L'expérience-type était la suivante. On broie le tissu qui déclenche l'induction (en l'occurrence, la cornée) ; on sépare ses composants biochimiques ; et on voit lequel de ces composants est celui qui réalise l'induction. Les techniques de séparation étaient assez fastidieuses (il fallait séparer des molécules biochimiques telles les protéines), et procédaient par étapes successives, mais la réussite en fin de compte semblait garantie d'avance. Eh bien, il n'en fut pas ainsi. Plusieurs laboratoires commençaient à obtenir des substances "purifiées", actives dans le test d'induction. Mais - consternation - il n'y avait pas deux laboratoires qui trouvaient la même chose. Il y avait autant de "substances actives" que de laboratoires. Devant cette situation incongrue, un chercheur s'est dit : "Puisque tant substances différentes sont toutes actives, pourquoi se fatiguer à les produire par purification fastidieuse ?" Il a donc pris des substances chimiques déjà existantes dans des flacons sur les rayons de son laboratoire ... et en effet, plusieurs d'entre eux se révélèrent efficaces dans le test d'induction. Le comble était qu'une de ces substances n'était autre que le bleu de méthylène, une substance banale ; pire encore, il se trouvait que l'argile de kaolin, substance humble s'il en est, était également efficace.

Cet événement fut vécu comme une débâcle par les embryologistes concernés. Mais ne peut-on pas considérer qu'à sa façon, la nature essayait de "dire" quelque chose ? A savoir, que le "secret" du phénomène d'induction n'était pas à trouver dans la structure de la substance qui déclenchait l'induction, mais dans la capacité du tissu induit de réagir ? Si on transpose l'analogie, cela veut dire que le "secret" de la compréhension ne réside pas dans une quelconque "information", "encodée" dans la structure du message, mais dans la capacité créatrice du sujet à construire une interprétation. Cela s'accorde bien avec le point de vue constructiviste selon lequel une perception n'est nullement une "représentation" d'un objet externe préexistant, mais une création du sujet lui-même.

Dennett [1993] a suggéré l'esquisse de ce que pourrait être un modèle alternatif de la compréhension. Contrairement au modèle du "théâtre cartésien", Dennett considère que le sens d'une pensée ne préexiste pas à son expression. Plutôt, dans la situation d'un sujet en présence d'un signal, ce qui est déclenché est une *multitude* d'interprétations potentielles qui se développent en parallèle. Au bout d'un certain temps, l'une de ces interprétations "prend le dessus" par rapport aux

autres, et constitue un "point fixe endogène". Les entrées sensorielles seront alors focalisés sur ce qui fait sens par rapport à cette interprétation. Mais "l'objet" qui vient ainsi à naître ne préexistait pas au processus d'interprétation. La preuve en est que d'autres sujets peuvent y voir d'autres "objets" et qu'il y a là place à la négociation intersubjective.

Conclusion

Nous prétendons qu'il est nécessaire de distinguer deux niveaux de description de tout phénomène. Nous prétendons aussi que la description du point de vue de l'extérieur (description₁) est inadéquate pour atteindre la compréhension des mécanismes qui engendrent le déroulement du phénomène à décrire. L'adoption d'une description₂, "de l'intérieur", de ce phénomène est indispensable pour rendre compte de son déploiement.

La conversation, conçue comme processus interlocutoire, est le lieu de l'intercommunicabilité des cognitions. Forme prototypique de l'interaction verbale, elle met en scène des acteurs qui, en interagissant, co-construisent le sens des formes linguistiques qui en constituent la trame. On a là un phénomène intersubjectif qui mérite d'être décrit.

Notre credo consiste à dire deux choses. S'en tenir à une description₁ de l'enchaînement conversationnel ne peut permettre d'en atteindre la logique interne. Et ce, même si la démarche adoptée est résolument interactionniste. Seule une description₂ peut conduire à rendre compte correctement des mécanismes qui génèrent l'émergence de sens dans l'interlocution. La logique interlocutoire est l'outil technique au service de cette description₂.

En reposant sur l'idée essentielle que le sens est co-construit, de façon processuelle et radicalement dialogique, cette façon de modéliser la conversation participe d'une perspective constructiviste. Point n'est besoin de postuler l'existence d'un sens préexistant et à l'expression et à la compréhension de la forme linguistique proférée en contexte. Il suffit d'accepter l'idée simple selon laquelle les interactants, immergés dans un potentiel subtilement impermanent de sens, le façonnent conjointement et processuellement et ainsi en font émerger un sens provisoire et toujours négociable. L'intercompréhension, que nous tenons pour un objet central des recherches en sciences cognitives, est cette co-opération intersubjective.

Références

Brassac, Christian, (1995). L'imprévisibilité de la conversation, une constructibilité par défaut. Séminaire interdisciplinaire de pragmatique, Institute for Advanced Studies, The Hebrew University of Jerusalem, 24-26 mai 1995.

Brassac, Christian, (1993). Analyse de conversations et théorie des actes de langage. Cahiers de Linguistique Française 13: 62-76.

Brassac, Christian, (1994). Speech acts and conversational sequencing. Pragmatics and Cognition Vol 2(1): 191-205.

Brassac, Christian et Alain Trognon, (1990). "Implications" entre énoncés, "engagements" entre actes illocutoires. Verbum XIII (4): 195-221.

Dennett, Daniel, (1993). La conscience expliquée. Paris : Editions Odile Jacob.

Dennett, Daniel, (1990). La stratégie de l'interprète. Le sens commun et l'univers quotidien. Paris : Gallimard.

Dessalles, Jean-Louis, (1995). Les rôles du contexte et de la situation dans la cognition. Actes de l'école d'été de l'ARC, Bonas, volume C, 6-25.

- Ghiglione, Rodolphe et Alain Trognon, (1994). Où va la pragmatique ? De la pragmatique à la psychologie sociale. Grenoble: Presses Universitaires de Grenoble.
- Grice, Paul, (1979). Logique et communication. Communications
- Jacques, Francis, (1995). Dialogue, dialogisme, interlocution. Séminaire interdisciplinaire de pragmatique, Institute for Advanced Studies, The Hebrew University of Jerusalem, 24-26 mai 1995.
- Livet, Pierre, (1994). La communauté virtuelle. Action et communication. Combas: Editions de l'éclat.
- Raccah, Pierre-Yves, (1990). Quelques remarques sur la sémantique linguistique et la construction du sens. Travaux de linguistique et de philologie n°30, Klincksieck, 71-81.
- Reddy, M.J., (1979). The conduit metaphor - a case of frame conflict in our language about language. in : Ortony, A. (éd.), Metaphor and Thought. Cambridge: Cambridge University Press, 284-324.
- Shannon, C.E., Weaver, W., (1949). The mathematical theory of communication. Urbana, University of Illinois Press.
- Sperber, Dan et Deirdre Wilson, (1985). La pertinence. Communication et cognition. Paris : Editions de Minuit.
- Stewart, John, (1995). Les rôles du contexte et de la situation dans la cognition. Actes de l'école d'été de l'ARC, Bonas, volume C, 103-115.
- Trognon, Alain et Christian Brassac, (1988). Actes de langage et conversation. Intellectica 6 (2): 211-232.
- Trognon, Alain et Christian Brassac, (1995). Formalizing the theory of intentionality. Journal of Pragmatics 23, 555-562.
- Trognon, Alain et Christian Brassac, (1993). L'enchaînement conversationnel. Cahiers de Linguistique Française 13: 76-107.

Du collectif au social : illustrations expérimentales de l'animal à l'homme

Anne-Marie Toniolo¹, Thierry Fierville¹, Didier Desor²

¹ Laboratoire de Psychologie - Université de Nancy 2

B.P. 3397 - 54015 Nancy Cedex

fiervill@clsh.u-nancy.fr toniolo@sbiol.u-nancy.fr

² Laboratoire de Biologie et Physiologie du Comportement (URA CNRS 1293)

Université Henry Poincaré - B.P. 239 - 54506 Vandoeuvre-les-Nancy Cedex

desor@sbiol.u-nancy.fr

Résumé

Nous présentons trois expériences mettant en jeu des groupes restreints en situation de résolution de problème. La première situation expérimentale a pour originalité de s'appliquer à des mammifères, petits rongeurs de laboratoire placés en situation de contrainte alimentaire, et de mettre en évidence un processus systématique de différenciation comportementale. La seconde est une adaptation de cette situation à des groupes humains. Elle vise à catégoriser des individus, aux compétences potentiellement équivalentes, face à l'accomplissement d'une tâche individuelle. La troisième introduit une approche nouvelle dans les méthodes de conception industrielle. Fondée sur la méthode des jeux de rôles, elle cherche, à travers l'analyse des interactions verbales et non-verbales déployées autour d'un produit en cours d'élaboration, à expliciter les mécanismes d'ajustement sollicités par la mise en situation de conception distribuée chez les représentants de trois modules techniques complémentaires. Ces trois situations contribuent à analyser les phénomènes d'émergence à l'oeuvre dans un contexte social.

Mots clés : émergence, dynamique de groupe, intelligence distribuée, résolution de problème

Quand des individus se rencontrent, se rassemblent ou sont rassemblés, ils produisent via leurs comportements, des interactions. Ces interactions permettent à la communication de s'établir et constituent les bases observables du processus de socialisation. Assujetties à des lois internes [Toniolo & coll., 1994], les interactions engendrent des relations dont le réseau peut évoluer vers l'élaboration d'une structure [Hinde, 1976]. Le système ainsi constitué, en l'occurrence social, peut se maintenir, évoluer sous le jeu de ces interactions, et ce, d'autant plus si les frictions de l'environnement auquel il est soumis se font contraignantes. Cette dernière considération évoque les perturbations que l'environnement peut produire sur le système, de façon générale. Il convient néanmoins de distinguer deux cas. La pression du milieu peut s'adresser aux éléments qui composent le système (problème individuel), puis, de proche en proche, l'affecter dans son ensemble. Inversement, elle peut s'exercer sur le système lui-même (problème collectif), puis se répercuter sur ses composants. Dans les deux cas, la perspective se déplace de la dynamique interne d'un fonctionnement stabilisé au principe organisateur qui en explique la genèse. Sous l'action de ce principe, quelque soit la nature initiale du problème, sa résolution fait passer le système d'un état dé- ou faiblement organisé (collectif) à un état organisé ou autrement organisé (social) [Le Moigne, 1990].

Par ailleurs, selon que les éléments du système sont, au départ, différenciés ou non, la réponse générale du système aux pressions de l'environnement se fait soit par un processus d'homogénéisation, soit par un processus de différenciation. Ces processus se trouvent sous le contrôle d'un déterminisme individuel lié aux caractères intrinsèques des éléments du système ou sous celui d'un déterminisme lié au contexte situationnel dont les paramètres définissent les degrés de liberté dont disposent les éléments pour y évoluer [Chase, 1993]. Adopter exclusivement l'une ou l'autre de ces deux options conduit néanmoins à un réductionnisme qui n'a pas lieu d'être. En effet, l'émergence de structures sociales par saut qualitatif, à partir du niveau individuel, résulte autant d'interactions entre individus potentiellement différents que d'interactions entre individus et environnement [Beckers, 1992]. De plus, l'enrichissement du système par apparition de propriétés nouvelles peut, en retour, réduire la marge d'expression des individus : « il y a toujours, et dans tout système, et même chez ceux qui y suscitent des émergences, des contraintes sur les parties, qui imposent restrictions et servitudes. Ces contraintes, restrictions, servitudes leur font perdre ou leur inhibent des qualités ou propriétés » [Morin, 1980].

Cette problématique est au centre d'une recherche initialement développée autour d'un phénomène de différenciation comportementale qui s'opère parmi de petits groupes d'individus confrontés à une contrainte vitale de l'environnement. Ces individus ont la singularité d'appartenir à une espèce phylogénétiquement éloignée de la nôtre, puisqu'il s'agit de petits rongeurs de laboratoire. Cette recherche se déploie dans une dimension interdisciplinaire qui, après avoir impliqué des équipes de psychologues et d'éthologues, s'ouvre actuellement à des chercheurs d'une discipline éloignée des Sciences de la Vie : celle des Sciences Pour l'Ingénieur. Ces chercheurs se trouvent réunis au sein d'un groupe autour de la problématique relative aux processus interactifs complexes (Groupe de Recherche Interdisciplinaire sur les Processus Interactifs Complexes. Coordination : toniolo@sbiol.u-nancy.fr). Nous présenterons dans cet exposé le modèle expérimental à l'origine de la réflexion et les principaux résultats qui en découlent. Nous évoquerons l'extension de la problématique à deux expérimentations humaines, en cours de réalisation. L'une émane directement du modèle animal, l'autre, conceptuellement proche, s'inscrit dans le domaine de la conception distribuée (Concurrent Engineering).

1. Autonomie et dépendance chez l'animal soumis à une contrainte vitale de l'environnement

1.1. Historique

Deux sources d'inspiration, l'une naturaliste, l'autre expérimentale, sont à l'origine de la situation précédemment évoquée. En 1973, deux écologistes italiens observent des colonies de rats sur les berges du Pô. Ils constatent que certaines d'entre elles plongent pour pêcher des mollusques sur le fond du fleuve, alors que les membres de colonies voisines, dans le même environnement, ne plongent pas [Gandolfi & Parisi, 1973]. Les auteurs concluent en faveur d'une transmission de l'information de type "acculturation". La seconde source d'inspiration est de nature expérimentale et remonte à la parution d'un article de Mowrer en 1940. Celui-ci décrit une situation de "social problem" où, pour obtenir de la nourriture, des rats testés en groupe sont obligés d'appuyer sur un levier. Levier et distributeur étant situés aux extrémités opposées de la cage du dispositif, la situation produit une différenciation comportementale entre rats travailleurs, qui appuient sur le levier, et rats non-travailleurs, qui consomment la nourriture ainsi fournie. Ce phénomène est qualifié de "parasitisme social" ou de "division du travail" par Oldfield-Box en 1967. Plus tard, Galef [1980] montre que dans un dispositif assurant l'obtention d'un supplément appétitif de nourriture par une plongée au fond d'un aquarium, tous les rats concernés ne plongent pas.

1.2. Cadre expérimental

La situation présente reprend le principe de contrainte vis-à-vis de l'acquisition de nourriture, en s'appuyant sur le répertoire comportemental naturel du rat et en provoquant des interactions directement quantifiables. Elle s'appuie sur un dispositif composé d'une cage où sont regroupés des rats mâles de souche Wistar. A cette cage est annexé un tunnel conduisant à un aquarium couvert et rempli d'eau, à l'extrémité duquel se trouve un distributeur de nourriture. Après une phase de familiarisation à sec (deux jours), les animaux doivent affronter la montée progressive de l'eau dans l'aquarium (six jours) jusqu'à son immersion complète. Ils doivent alors plonger et nager en apnée pour atteindre la mangeoire sous laquelle ils n'ont aucune possibilité d'appui. Pour consommer la nourriture, il leur faut regagner la cage. Les rats sont placés dans ce dispositif pendant une durée quotidienne de trois heures [Colin & Desor, 1986]. Ces paramètres maximisent ainsi le coût de la satisfaction d'un besoin individuel par le biais d'une contrainte physique et sociale : elle agit sur l'émotivité des sujets, par l'éloignement de la nourriture et sa consommation différée ; elle induit une pression sociale inévitable du fait du regroupement d'individus tous autant concernés.

1.3. Structuration des groupes

Lorsque les animaux sont confrontés seuls à la situation, ils finissent tous par plonger et se ravitailler. En revanche, quand ils sont groupés par six, ils se différencient en trois catégories. Environ 50% plongent et rapportent la nourriture dans la cage. Les autres ne plongent jamais et survivent en volant leur nourriture aux précédents. Une observation détaillée fait apparaître deux types distincts d'individus parmi les rats "transporteurs" : si certains d'entre eux ravitaillent effectivement leurs congénères "non-transporteurs", d'autres, qualifiés d'autonomes conservent, en revanche, la nourriture rapportée, au prix d'une résistance efficace. Cette différenciation comportementale est un phénomène robuste. Ces trois profils sont très stables dans le temps. Ils restent inchangés après huit mois d'entraînement quotidien. Ce phénomène est apparu, à ce jour, dans plusieurs centaines de groupes testés dans cette situation. Il se manifeste également chez d'autres souches de rats, voire chez d'autres rongeurs [Helder & coll., 1995]. Deux arguments supplémentaires confortent son caractère émergent : le rôle de l'effectif et les différenciations après remaniement expérimental des groupes. En effet, plus l'effectif du groupe augmente, plus le taux de différenciation s'affirme [Thullier & coll., 1992]. Par ailleurs, le regroupement, exclusivement de "transporteurs" ou de "non-transporteurs" préalablement différenciés, conduit à de nouvelles différenciations. Ces re-différenciations témoignent de la flexibilité du phénomène et viennent confirmer sa dépendance à l'égard d'un facteur social prégnant.

1.4. Scénario descriptif de la différenciation

Par le jeu des interactions, trois solutions s'offrent aux individus pour résoudre leur problème [Desor & coll, 1993]. C'est l'articulation de ces trois solutions, assimilées à des rôles, qui garantit leur survie et, par là même, assure la viabilité du groupe. L'état de santé satisfaisant de l'ensemble du groupe témoigne de l'efficacité de cette régulation. Celle-ci repose sur la connaissance que chaque individu a de son propre rôle et du rôle de ses congénères [Desor & coll., 1992]. Les individus investissent ces rôles en fonction de leurs légères différences initiales, puissamment amplifiées par le contexte situationnel [Toniolo & coll., 1994]. Chaque rôle s'exprime par des attributs qui s'affinent au fil du temps, favorisant l'apparition d'un réseau de relations privilégiées, puis le consolidant. Le comportement des "non-transporteurs" ne se réduit pas au vol de nourriture. Ceux-ci sont capables d'orienter le comportement de leurs congénères, en les incitant à plonger [Desor & Toniolo, 1992]. Si les "ravitailleurs" essaient d'échapper aux sollicitations de leurs

congénères, ils finissent toujours par plonger. Mais, au fil des séances, ils résistent de moins en moins aux attaques pour se résigner (au sens pharmacologique du terme) en cédant immédiatement leur butin. Les individus "autonomes", eux, ne sont jamais inquiétés.

Face à la limitation des ressources, à leur accès difficile, sous la pression du groupe, la dynamique des interactions interindividuelles fait évoluer les rôles en spécialisations comportementales. Celles-ci peuvent être interprétées en terme de fonctionnalités émergentes [Steels, 1991] qui assurent la viabilité et la pérennité du groupe dès lors constitué et structuré. Ces éléments permettent d'appréhender le phénomène dans une perspective systémique en adoptant une approche de type "agent" [Erceau & Ferber, 1994] élaborée dans le cadre de l'Intelligence Artificielle Distribuée, mais dont les concepts débordent largement le cadre original et pénètrent les Sciences Humaines et Sociales ainsi que l'Ethologie. Dans cette approche deux courants coexistent. Ils correspondent aux deux pôles d'une conception des organisations sociales. Pour l'un, les sujets sont des agents "intentionnels", de type cognitif, dans le sens où ils exploitent des plans explicites pour atteindre un but ou accomplir un objectif. C'est typiquement le cas des sociétés d'Humains. Pour l'autre courant, l'intelligence des agents considérés comme réactifs, est limitée ; l'organisation émane du système envisagé dans sa globalité. C'est typiquement le cas des sociétés d'insectes [Deneubourg & coll, 1991].

Dans la situation expérimentale évoquée, qui ne répond ni à une logique de compétition, ni à une logique de collaboration, les deux aspects se manifestent conjointement : les rats se conduisent à la fois comme agents "intentionnels" quand ils font preuve de cognition et comme agents réactifs dans la mesure où la résolution de leur problème est médiatisée par le groupe en formation. Pour reprendre Bonabeau [1994], nous pouvons dire que la situation témoigne d'un phénomène émergent de « cognition collective qui dépend des relations entre les composants : ceux-ci doivent instancier une organisation particulière que la science cognitive a pour tâche d'identifier ». Parce que les paramètres de cette situation rappellent certains contextes empiriquement observés ou artificiellement créés chez l'Humain (situations d'isolement ou de confinement, par exemple), il nous est apparu judicieux d'envisager une situation expérimentale qui mette en jeu des sujets sur la base des significations conceptuelles dégagées dans le modèle animal et qui, par analogie, en reproduise la pertinence expérimentale.

2. Différenciations comportementales en situation humaine de confinement

2.1. Principe de l'expérimentation

L'expérimentation se déroule dans une salle de 50 m² où sont réunis 5 sujets de même sexe issus d'une population de même origine (étudiants en psychologie et biologie) et qui ne se connaissent pas au début de l'expérience. Au cours de l'exposé d'une consigne collective, les sujets sont informés de l'objectif à atteindre : chacun d'eux doit essayer de s'approprier un maximum de points disponibles en quantité limitée. Ces points se capitalisent au fil de plusieurs séances consécutives auxquelles sont conviés les sujets. L'appropriation des points se réalise en deux temps. Pour obtenir des points, le sujet doit affronter une épreuve d'habileté manuelle. La réussite à ce test psychomoteur, contrôlé par ordinateur, conditionne la distribution d'un nombre fixé de points par écriture automatique sur la disquette personnelle du sujet se soumettant à l'épreuve. Les points gagnés ne deviennent propriété définitive du sujet qu'après avoir été ultérieurement validés auprès d'un des 5 ordinateurs disponibles à l'autre bout de la pièce. Là encore, un programme spécifique gère le processus de validation, qui consiste en une conversion des points rapportés par le sujet en points définitifs, qui se capitalisent sur sa disquette. Cette opération se déroule au rythme d'un point par seconde, le sujet maintenant appuyée une touche du clavier dévolue à cet effet.

L'organisation de la salle d'expérimentation s'articule autour des dispositifs que nous venons d'évoquer (figure 1).

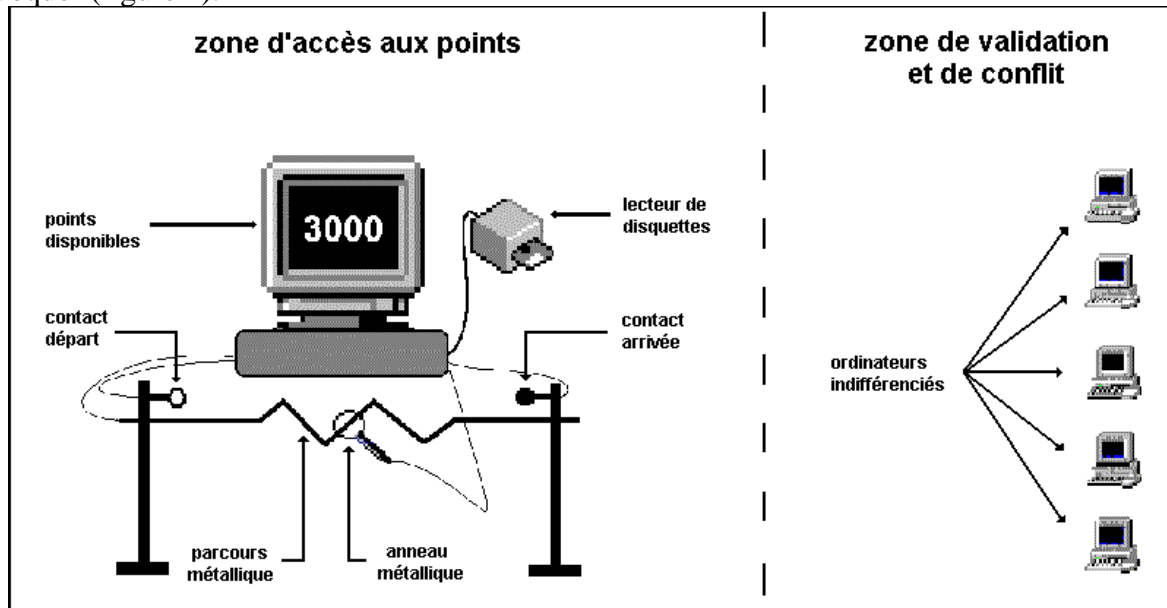


Figure 1 : Disposition générale de la salle d'expérimentation

2.2. Dynamique de la situation

Ces conditions particulières - source unique de distribution et validation différée des points -, ménagent aux sujets une possibilité supplémentaire pour constituer leur capital : l'interception, sur le lieu de validation, des points en cours de conversion rapportés par autrui. Ainsi, la phase de validation peut, à tout moment, être interrompue par un condisciple qui revendique alors la propriété des points en appuyant sur une touche destinée à cet effet. Cette action provoque l'ouverture d'un conflit. Dès lors, l'attribution des points en suspens à l'un des deux protagonistes est déterminée par l'issue d'un duel simulé par l'ordinateur. Le système est conçu de telle sorte que le sujet provoqué ne puisse se dérober : le conflit engagé est inéluctablement décisif. Cependant, chacun des deux sujets conserve, à tout moment, la possibilité d'abandonner le duel. Celui-ci prend la forme d'une épreuve d'identification d'anagrammes, requérant des aptitudes d'ordre cognitif. La réponse du candidat le plus rapide est prise en compte. Le premier des deux protagonistes parvenu à 10 points sort vainqueur et s'octroie le potentiel de points en jeu. Il ne lui reste plus alors qu'à les valider à son profit, tandis que le perdant peut, soit retourner tenter sa chance au test psychomoteur, soit essayer de s'approprier les points d'un troisième individu. Ce cycle d'actions et d'interactions se prolonge jusqu'à épuisement total du stock de points disponibles en début de séance.

2.3. Hypothèses de travail

Ainsi, la singularité de la situation expérimentale consiste, ici aussi, à maximiser les paramètres d'une double pression, physique et sociale qui s'exerce sur les sujets. Pour satisfaire son objectif, chaque sujet doit franchir l'obstacle que constitue l'épreuve d'habileté motrice. Cette épreuve a la particularité d'être pénible, non seulement par la dextérité qu'elle requiert, mais également par le caractère anxiogène qui l'accompagne : d'une part, les erreurs sont scandées par l'émission d'un signal sonore désagréable ; d'autre part, le caractère éprouvant est renforcé par la présence environnantes de condisciples tous autant intéressés par l'acquisition des points et/ou par la pression qu'ils sont en mesure d'exercer lors de la validation des points. Agissant sur l'émotivité différentielle des sujets, cette épreuve peut être surmontée ou évitée. Dans ce dernier cas, elle peut

être suppléée par un second mode d'acquisition sollicitant des compétences d'une autre nature. Le schéma suivant (figure 2) résume l'articulation des deux modes de satisfaction du besoin.

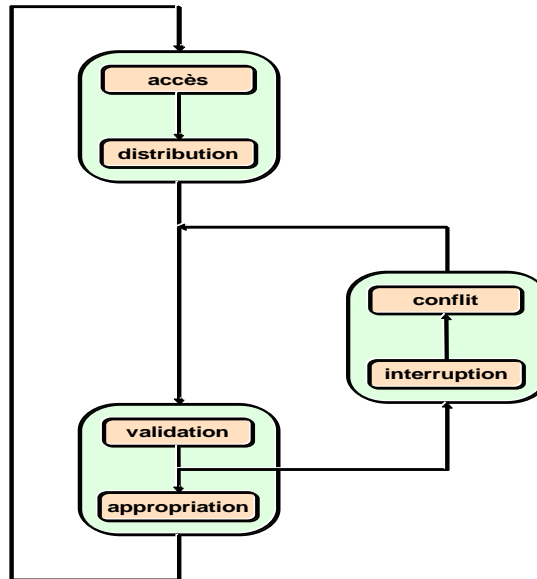


Figure 2 : Dynamique fonctionnelle de la situation expérimentale

Nous faisons l'hypothèse d'une différenciation comportementale produite par un scénario dont la dynamique rappelle, par analogie, la situation animale. En effet, lorsqu'un sujet affronte l'épreuve manuelle et qu'il n'effectue pas correctement la tâche, deux éventualités s'offrent à lui : soit recommencer, soit se décourager. Dans ce dernier cas, il peut abandonner ou recourir aux compétences d'un condisciple qui, plus adroit, plus patient ou/et moins émotif, finira par franchir l'obstacle avec succès. Cet individu qui persiste, exerce sa dextérité et par apprentissage devient plus habile, voire expert dans cette épreuve. Il se spécialise, ainsi, dans un rôle d'approvisionnement dont saura tirer profit son condisciple. Inversement, ce dernier, face à ses propres échecs, verra son appréhension augmenter et se trouvera renforcé dans l'alternative d'obtenir des points en les capturant sur le lieu de leur validation. Les sujets se trouvent ainsi pris dans un processus supra-individuel qui les conduit à adopter, inévitablement, l'un des trois profils que la situation fait émerger. Nous suggérons que cette approche permettrait de reconsidérer l'analyse conceptuelle de la dynamique des petits groupes humains. La première étape de cette recherche a consisté à calibrer et à valider les paramètres des dispositifs définissant le cadre opérationnel. Les groupes sont actuellement constitués et leur soumission à la situation est imminente.

3. Structuration de groupe médiatisée par la conception d'un objet

3.1. Position du problème

Le second volet de la présente recherche s'intéresse à la résolution d'un problème dans le domaine de la conception. Il répond à un projet financé par le C.N.R.S. dans le cadre d'un plan d'action national où les départements des Sciences de l'Homme et des Sciences Pour l'Ingénieur collaborent pour la première fois. Ce projet présente une complémentarité par rapport aux travaux précédents, tant au niveau des concepts que des processus requis. Il s'appuie sur la même méthodologie d'analyse des interactions sociales qui se développent au sein d'un groupe restreint d'individus réunis pour élaborer un produit. Celui-ci est le témoin tangible d'une structuration de groupe sous-jacente.

La conception de l'objet dans une perspective distribuée présente des avantages considérables en terme de coût/bénéfice par rapport à une démarche classiquement séquentielle de son élaboration. Ces bénéfices se mesurent en particulier par la réduction des délais et par une meilleure adéquation entre la demande et le produit fini. Cette constatation suppose une collaboration des parties engagées, depuis la définition du cahier des charges, jusqu'à la fabrication du produit. Cette collaboration procède d'une mise en commun, en permanence ajustée, des savoirs et des savoir-faire. Certaines faiblesses discréditent néanmoins la méthode. Elles proviennent pour une grande part de l'interférence du facteur humain (conflits, dilution des responsabilités). Pour aborder ce problème, un modèle d'étude a été développé où la conception d'un produit est envisagée comme la construction d'un sens commun à partir d'échanges d'information entre différents modules, en l'occurrence : un module fonctionnel, un module structurel et un module de fabrication.

3.2. Mise en oeuvre d'un modèle d'étude expérimental

L'objectif est de mettre en évidence les meilleures conditions possibles qui favorisent la convergence des connaissances en éprouvant puis en validant un modèle de conception distribuée. La réalisation de cet objectif se fonde sur une expérimentation de conception d'un système mécanique de complexité relativement faible, sous forme de jeu de rôles avec des intervenants faisant office de module. Des tentatives préliminaires ont été menées afin de cerner au mieux les conditions les plus favorables pour une telle démarche. La première a consisté à concevoir un objet (toret à meuler) par des intervenants isolés, ne communiquant que par supports écrits et par l'intermédiaire d'un médiateur [Charpentier & coll., 1995]. La seconde investigation a porté sur la conception d'un objet tout aussi mécaniquement simple (bras oscillant) par des intervenants géographiquement éloignés, reliés par réseau informatique. Dans les deux cas, les résultats ont donné des indications quant à la dynamique des interactions, leur impact dans l'articulation des solutions trouvées. Cependant, les conditions ont été éprouvantes pour les participants. Elles se sont soldées par l'apparition de conflits dans la première situation, et par une lassitude démotivante et une perte d'intérêt dans la seconde.

3.3. Dernière phase de la mise au point expérimentale

Ces constats nous ont conduits à élaborer une troisième forme d'expérimentation. Elle consiste à réunir autour d'une table pour un temps limité (six heures) les représentants des trois modules. Le système à concevoir, plus complexe ici, doit permettre la mise et le maintien en position de pièces en bois de forme quelconque, présentant de légers défauts de surface, en vue de leur correction sur la table d'une défonceuse. Les contraintes sont définies par le cahier des charges mis à la disposition de chacun des trois modules et explicitement détaillé au début de l'expérimentation par le module fonctionnel. Les six heures d'intervention ont été fractionnées en trois demi journées de travail, pendant lesquelles les intervenants n'ont pas quitté la salle d'expérimentation. Au terme de ces séances, l'objet a effectivement émergé, avec potentialité de fabrication immédiate. Les intervenants ont été directement observés par un psychologue, un éthologue et un sociologue. Ils ont été filmés et leurs interactions, verbales et non-verbales, dépouillées selon une grille reconstituant l'historique des échanges. Les données recueillies sont actuellement soumises à un traitement statistique apte à formaliser la dynamique simultanée de toutes les interactions exprimées. Il conviendra ensuite de mettre en relation les séquences d'interactions significatives avec l'évolution technique de l'objet et d'évaluer dans quelle mesure elles témoignent d'une progression, d'une régression, d'une stagnation voire d'une remise en cause totale de l'objet en cours de conception. Il s'agit de dégager les "moments" de l'émergence et d'identifier les processus qui les sous-tendent.

Cette étude, en cours d'analyse, demandera, dans une étape ultérieure (2^{ème} tranche du projet), validation des conclusions. Elle apparaît complémentaire de l'étude relative au modèle animal et à son extension humaine développée au début de cet exposé. Dans les deux cas, il s'agit de la résolution collective d'un problème sous fortes contraintes environnementales auxquelles est soumis un groupe d'individus ou d'entités. Mais, dans le premier cas, le problème est individuel alors que, dans le second, il est d'ordre collectif. Au départ, dans le premier cas, les entités sont homogènes, non différenciées : elles bénéficient de compétences équivalentes ; dans le second cas elles sont hétérogènes, bien différenciées : elles sont dotées de compétences spécifiques. Dans les deux cas, il y a convergence des interactions dans le sens de l'émergence d'une organisation sociale. Mais le passage du collectif (regroupement en terme de sommation initiale des entités) au social (structuration de relations) se fait par un processus de différenciation et selon une logique de compétition dans le premier cas ; il s'effectue par un processus d'homogénéisation et selon une logique de collaboration dans le second cas. Le matériel et des protocoles expérimentaux ont été conçus pour objectiver des interactions qui conduisent, du fait de leurs itérations, à l'établissement de relations et ainsi à la constitution de la structure d'un système en formation et en voie de stabilisation. Cette approche expérimentale nous ouvre ainsi de façon directe aux interrogations relatives à l'émergence et à la régulation d'un système.

Références

- Beckers, R. (1992). L'auto-organisation - Une réponse alternative à la complexité individuelle ? Thèse de Doctorat, Université Paris-Nord (Paris XIII).
- Bonabeau, E. (1994). Intelligence collective ? In Bonabeau, E. & Theraulaz, G. (Eds). *L'intelligence collective*. Paris : Hermès, 13-27.
- Charpentier, P., Garro, O. & Toniolo, A.-M. (1995). Vers l'émergence d'un produit en conception distribuée. 4^{ème} Colloque PRIMECA sur la Conception Mécanique Intégrée. La Plagne.
- Chase, I.D. (1993). Generating societies : collective social patterns in humans and animals. In Actes du colloque E.C.A.L. (European Conference on Artificial Life). 24-26 mai, Bruxelles.
- Colin, C. & Desor, D. (1986). Différenciations comportementales dans des groupes de rats soumis à une difficulté d'accès à la nourriture. *Behav. Proc.*, 13, 85-100.
- Deneubourg, J.L., Goss, S., Beckers & R., Sandini, G. (1991). Collectively self-solving problems. In Babloyantz, A. (Ed.) *Self-organization, emergent properties and learning*. Plenum.
- Desor, D. & Toniolo, A.M. (1992). Incentive behavior in structured groups of rats : about the possible occurrence of social cognitive processes. In Actes C.A.C.S. 92 (Comparative Approaches to Cognitive Science). 5-17 juillet, Aix en Provence.
- Desor, D., Toniolo, A.M., Dickes, P. & Krafft, B. (1993). Trois solutions pour un problème : structuration de groupes de rats en réponse à une difficulté d'accès à la nourriture. Actes du Symposium international : Universel et Différentiel en Psychologie. Aix-en-Provence.
- Desor, D., Toniolo, A.M., Krafft, B. & Dickes, P. (1992). Cognition sociale chez l'animal : à propos du comportement d'incitation chez le rat. Actes du 5^{ème} colloque de l'ARC : Percevoir, Raisonner, Agir : articulation des modèles cognitifs. Nancy : CRIN-CNRS.

- Erceau, J. & Ferber, J. (1994). Intelligence artificielle distribuée et systèmes multi-agents. In Bonabeau, E. & Theraulaz, G. (Eds). *L'intelligence collective*. Paris : Hermès, 157-179.
- Galef, B.G. (1980). Diving for food : analysis of a possible case of social learning in wild rats. (*Rattus Norvegicus*). *J. Comp. Physiol. Psychol.* 94, 416-425.
- Gandolfi, G. & Parisi, V. (1973). Ethological aspects of predation by rats, *Rattus Norvegicus* (Berkenhout), on bivalves, *Unio pictorum L* and *Cerastoderma lamarcki* (Reeve). *Bull. Zool.*, 40, 69-74.
- Helder, R., Desor, D. & Toniolo, A.M. (1995). Potential Stock Differences in the Social Behavior of Rats in a Situation of restricted Access to Food. *Behavior Genetics*, Vol. 25, 5.
- Hinde, R.A. (1976). Interactions, relationships and social structure. *Man*, 11, 1-17.
- Le Moigne, J.L. (1990). La théorie du système général : théorie de la modélisation. (2^{ème} édition). Paris : P.U.F.
- Morin, E. (1980). *La méthode : la vie de la vie*. Paris : Seuil.
- Mowrer, O.H. (1940). Animal studies in the genesis of personality. *Trans. N.Y. Acad. Sc.* 3, 8-11.
- Oldfield-Box, H. (1967). Social organization of rats in a "social problem" situation. *Nature*, 213, 533-534.
- Steels, L. (1991). Toward a theory of emergent functionality. In Meyer, J.A. & Wilson, S.W. (Eds). *Simulation of Adaptive Behavior : From Animals to Animats*. MIT Press, Bradford Books.
- Thullier, F., Desor, D., Mos, J. & Krafft, B. (1992). Effects of group size on social organization in rats with restricted access to food. *Physiol. Behav.*, 52.
- Toniolo, A.M., Desor, D. & Dickes, P. (1994). Différences individuelles au cours de l'ontogénèse et rôles sociaux à l'âge adulte : une étude chez le rat Wistar. XI^{èmes} Journées de Psychologie Différentielle. Montpellier.

De la contagion à l'épidémie: modèle neuromimétique de propagation de maladies transmissibles

Laurent Toubiana¹, Jean François Vibert¹

¹B3E, INSERM U444, ISARS
Faculté de médecine Saint-Antoine
27, rue Chaligny
75571 Paris Cedex 12, France.
Tel: 44 73 84 52
Fax: 44 73 84 54
e-mail: toubiana@b3e.jussieu.fr

Résumé:

Les épidémies de maladies transmissibles sont l'expression de phénomènes complexes au sein de collectivité. Depuis 1984, Le Réseau National Téléinformatique de Surveillance et d'Information sur les Maladies Transmissibles (RNTMT) collecte les données épidémiologiques sur huit maladies transmissibles. Nous utilisons les données de morbidité de syndromes grippaux pour étudier la complexité des dynamiques spatio-temporelles d'épidémies. Nous proposons un modèle permettant la simulation de la propagation d'épidémies de maladies transmissibles en développant une analogie neuromimétique. Le taux de contact, paramètre fondamental de la dynamique des épidémies, est représentatif de comportements sociaux au sein de collectivités. Il est représenté par le nombre et le poids des connexions entre les éléments de notre modèle.

Mots clés: Réseau de Neurones Artificiel; maladies transmissibles; épidémies; dynamiques spatio-temporelles; grippe.

1. Introduction

Une contagion est la transmission d'un agent pathogène d'un individu malade à un individu sain. Un phénomène épidémique est engendré lorsqu'un malade contamine plus d'un individu. Le sens commun établit intuitivement la causalité entre le constat d'un phénomène à une échelle individuelle, "transmission d'un agent pathogène" et un effet à une échelle collective, "l'épidémie".

Les virus se déplacent dans un espace en utilisant comme vecteur les individus qu'ils infectent. En conséquence, les relations entre unités d'espace qui permettent le déplacement des individus, autorisent de la même manière, la diffusion des virus. L'évolution des incidences de maladies transmissibles est l'expression d'un système complexe: diffusion dans le temps et dans l'espace de populations de virus sur d'autres populations (par exemple humaines).

Nous proposons un modèle permettant la simulation de la propagation d'épidémies de maladies transmissibles en exploitant une analogie neuromimétique. Dans un réseau de neurones, chaque neurone transmet un potentiel d'action à tous les neurones avec lesquels il est connecté. Chaque neurone intègre les signaux qu'il reçoit, réagit selon son état et retransmet un potentiel d'action. Au sein d'un groupe, un individu transmet un virus aux individus avec lesquels il est en relation. Ceux-ci développent la maladie selon leurs état immunitaire, puis la retransmette. Le modèle est constitué d'objets élémentaires interconnectés en réseau, susceptibles de recevoir et de transmettre un signal. Chaque objet réagit de manière non linéaire aux entrées qu'il intègre. Chaque réaction engendre une phase de moindre réceptivité. Ce modèle simule le comportement d'un ensemble d'éléments en ne définissant que les caractéristiques des éléments qui le compose et les

relations qui les lient. Le comportement d'un réseau de ce type est transposé à celui d'une collectivité vis à vis d'une maladie contagieuse.

Les comportements sociaux au sein de collectivités sont représentés implicitement dans notre modèle par le nombre et le poids des connexions entre éléments. Ces paramètres définissent la notion de "taux de contact", paramètre fondamental de la dynamique des épidémies.

Ce modèle utilise une méthodologie développée en modélisation neuronale à plausibilité biologique [Vibert 1994].

2. Le modèle

2.1. Les éléments du modèle

Le modèle est constitué par l'assemblage en réseau d'éléments définis ci-dessous. Par la suite nous identifions le comportement de chaque élément à celui d'une entité spatiale. Ce principe d'identification de l'activité d'un élément à celle d'une entité spatiale est valable quelque soit la nature et la taille de l'entité considérée; celle-ci pouvant correspondre soit à un individu soit à une de population plus ou moins importante: par exemple unités territoriales administratives telles que les départements.

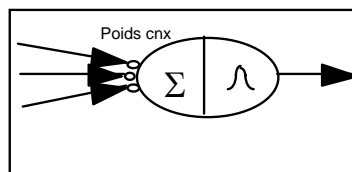


Figure 1 : Structure élémentaire du modèle.

La structure élémentaire du système est un objet très simple ayant trois propriétés fondamentales:

- Son "état" intérieur est susceptible d'évoluer en fonction de signaux reçus de son milieu extérieur.
- Il reçoit des signaux émanant de son milieu extérieur c'est à dire d'autres objets de même nature.
- Il a la capacité de communiquer son état à d'autres éléments de même nature.

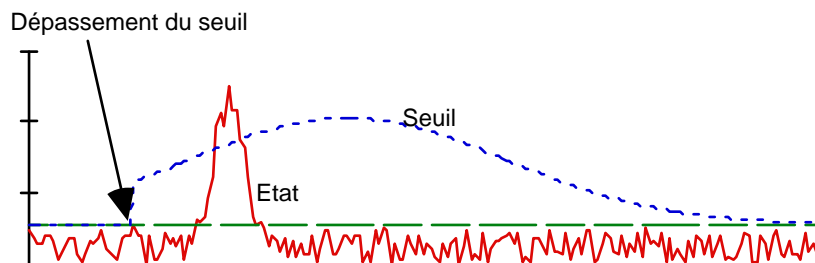


Figure 2: Évolution de l'état et du seuil d'un élément.

Sa dynamique est définie par l'évolution non linéaire de son état: il réagit brutalement lorsque les signaux reçus dépassent un seuil de réceptivité. Ce seuil évolue lui aussi. Il est fonction de l'état de l'entité. La valeur du seuil définit des périodes réfractaires qui correspondent à une moindre réceptivité.

2.2. Le réseau

Les éléments définis ci-dessus ont la capacité de transmettre leur état aux autres éléments lorsque ceux-ci sont connectés. L'ensemble connecté forme un réseau. Chaque élément d'un réseau peut avoir des connexions sortantes et entrantes pour émettre et recevoir l'état des autres éléments.

Chaque connexion est affectée d'un coefficient qui pondère la valeur de l'état transmis. Nous créons des réseaux en spécifiant les connexions et leur poids pour chaque élément. Les poids des connexions sont calculés en fonction de données relatives à l'entité représentée, par exemple des données spatiales.

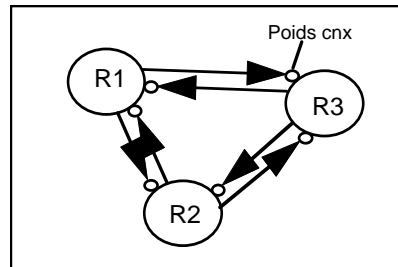


Figure 3 : Un réseau est un ensemble d'objets élémentaires connectés.

Chaque élément d'un réseau peut avoir des connexions sortantes et entrantes pour émettre et recevoir l'état des autres éléments. Chaque connexion est affectée d'un coefficient qui pondère la valeur de l'information transmise.

3. Application aux épidémies de grippe

La grippe est une des huit maladies transmissibles surveillées de manière systématique par le Réseau National Téléinformatique de surveillance et d'information sur les Maladies Transmissibles (RNTMT)[Valleron AJ., 1986]. La grippe est une pathologie aiguë virale de l'arbre respiratoire. La guérison est habituellement spontanée en deux à sept jours. La période d'incubation est courte: de 24 à 72 heures. La période contagieuse est probablement limitée à trois jours. Elle se transmet de personne à personne par des gouttelettes de "Flügge": sécrétions respiratoires exhalées par un individu infectieux et inhalées par les autres [Cliff 1986]. Nous utilisons les données d'incidences produites par ce réseau (fig.4).

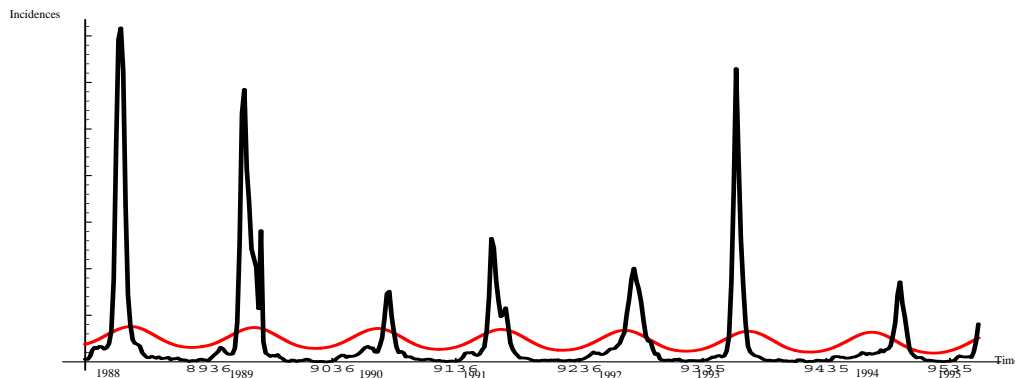


Figure 4 : Évolution de l'incidence des syndromes grippaux entre 1984 et 1993.

Cette figure montre la courbe des incidences de syndromes grippaux (nombre

de nouveaux cas rapporté à la population au cours d'une semaine) de 1988 à 1995. La courbe des incidences est accompagnée de la courbe de seuils épidémiques (Serfling). Tout dépassement de la courbe seuil par la courbe des incidences définit un début de période épidémique. Chaque pic d'incidence correspond à une période épidémique.

Depuis le début de la surveillance en 1984, il est possible d'observer de brutales élévations de la valeur de l'incidence, chaque année en période hivernale. Elles correspondent à des épidémies de grippe qui durent environ six semaines et peuvent atteindre un nombre total de cas estimés allant de 1 à 4 millions. Les cartes des épidémies de syndromes grippaux produites chaque semaine mettent en évidence la complexité de la dynamique spatio-temporelle du phénomène (fig.5).

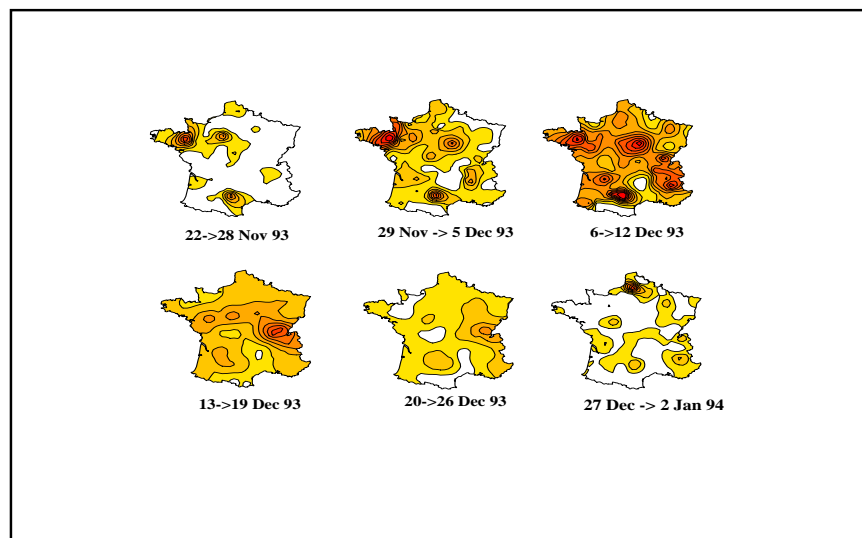


Figure 5: Evolution spatio-temporelle de la saison épidémique de syndromes grippaux sur la France correspondant à l'hivers 1993/94

4. Identification du modèle à un espace de propagation

A partir de ces données de la morbidité de la grippe, il est possible d'ajuster les paramètres des éléments de notre modèle pour effectuer des simulations de la propagation d'épidémies. L'espace définissant le territoire français est découpé en 94 sous-espaces correspondant aux départements (définition administrative). Nous établissons des coefficients d'interactions (couplage) entre entités spatiales (ici le département) sur la base d'une équation gravitaire [Reilly W. 1931; Converse PD. 1949, Wilson AG. 1967]. Cette équation permet d'évaluer les flux migratoires interurbains à partir de données sur les positions des unités spatiales et leur démographie. Chaque département est représenté par un élément défini plus haut. Les poids des connexions correspondent aux coefficients de couplage. L'état de l'élément correspond à la valeur de l'incidence de la grippe dans le département auquel il est associé (fig.6).

A tous les temps de la simulation, chaque élément fait la somme pondérée des états qui lui parviennent. Lorsque cette somme dépasse le seuil de réceptivité, l'élément entre dans une phase d'activation. Le parallèle est le suivant: toute région de l'espace reçoit un nombre d'immigrés de l'ensemble des autres espaces. Connaissant l'état épidémique de l'espace d'origine, nous pouvons estimer en chaque espace l'influence de l'immigration en terme épidémique. Le nombre d'infectés autochtones auquel on additionne les infectés immigrés, est rapporté à la population de l'espace considéré. Si le taux d'infectés dépasse un seuil alors le phénomène épidémique est déclenché.

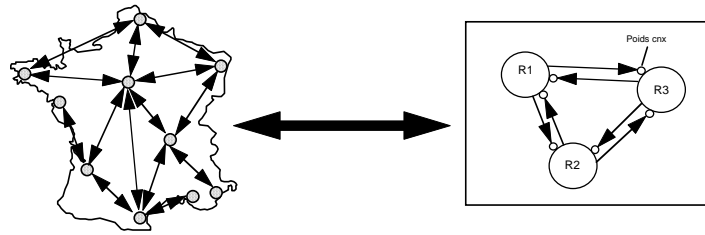


Figure 6 : modèle pour la simulation des dynamique d'épidémies

La figure représente le modèle que nous utilisons pour la simulation des épidémies sur le territoire français avec une résolution départementale. Chaque élément reçoit la valeur des sortie de tous les autres (ici 94 entrées). chaque entrée est affectée d'un coefficient appelé **poids** qui correspond à la part des immigrants infectés. Le calcul de ces poids s'inspire du modèle gravitaire.

La figure 7 présente sous la forme d'une série de cartes, le résultat d'une simulation dans les conditions suivantes. Le réseau est constitué de 94 éléments identifiés aux départements français. Le réseau est entièrement connecté. Le poids de chaque connexion est calculée comme décrit ci-dessus, ceci correspond à définir l'espace physique en fonction de ces caractéristiques, en l'occurrence la répartition spatiale de son peuplement. Les conditions initiales sont générées par une variable aléatoire centrée à 1 et d'écart type 2. La valeur initiale du seuil est 3, ce qui correspond approximativement à celle retenue pour la déclaration officielle des épidémies de grippe par la méthode du Serfling [Costagliola et al. 1991]. Un bruit gaussien centré et d'écart type 1 est ajouté à l'état de chaque élément comme nous pouvons l'observer sur les séries temporelles de la morbidité de la grippe depuis onze ans. La valeur du pas de simulation est la semaine, qui est aussi le pas de temps utilisé pour le recueil des incidences de la grippe. La simulation est faite sur 500 semaines, soit environ dix ans (le RNTMT est entré en fonction depuis la fin 1984). Etant donné les conditions initiales et le bruit ajouté, l'état de quelques éléments fini par dépasser le seuil. Dès lors, des réactions locales sont engendrées qui, par le jeu des connexions, envahissent peu à peu le territoire. Tout espace touché voit son seuil de réceptivité augmenter, il ne peut donc être touché de nouveau au cours de la même période et l'épidémie fini par s'évanouir d'elle même. Le réseau reste stable a un niveau bas pendant une durée qui correspond à la période réfractaire des éléments. Cette période dépassée, le réseau renouvelle un cycle d'activité. La période réfractaire est un des paramètres de la simulation ajusté au niveau de chaque élément. Elle est fixée ici à une valeur de 50 semaines, spécifique d'une périodicité circannuelle, telle qu'elle est observée sur les données (Fig. 4). Les cartes de la figure 7 montre l'évolution spatio-temporelle du phénomène. Elles correspondent à un scénario épidémiques que l'on peut rapprocher qualitativement des cartes des saisons épidémiques telle qu'elles sont observées par le RNTMT (fig.5)

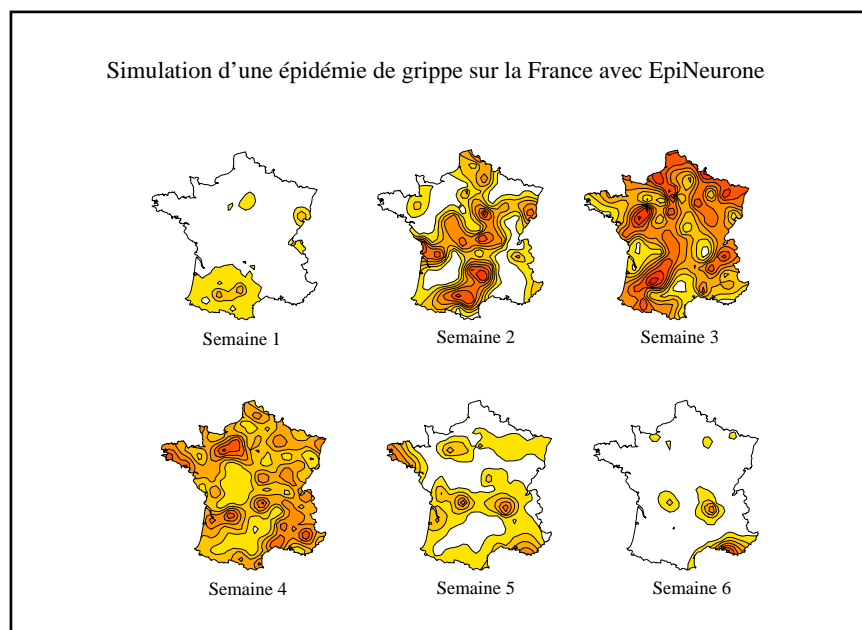


Figure 7 : Cartes simulées de l'évolution spatio-temporelle d'une épidémie de grippe.

5. Discussion

Il existe une importante littérature traitant de la dynamique des épidémies; Le Pont F. et al. [1989] font une revue de cette littérature à l'occasion d'une étude sur la diffusion de l'épidémie de SIDA. Des modèles mathématiques et statistiques [Anderson RM. & May RM., 1991; Longini IM. 1988; Baroyan OV. 71; Cliff AD. & Haggett, 1992] ont été mis au point en vue de décrire des diffusions à partir de données épidémiologiques.

Nous proposons une nouvelle approche méthodologique pour la simulation dynamique d'épidémies qui s'inspire des travaux effectués dans le domaine de neurophysiologie numérique. Le modèle permet de générer des comportements globaux de population. Il reprend ce qu'écrivait Bailey en 1975 "*les épidémies dans les grandes populations sont une sommation de plus petites épidémies se produisant à des endroits géographiquement séparés. Les courbes et les figures observées ne sont alors que le résultat d'une superposition d'autant de petites épidémies se produisant simultanément à l'intérieur de petits groupes. En réalité, ces groupes se recouvrent et interagissent tous.*". Contrairement à bien d'autres modèles, nous ne définissons pas une loi comportementale au niveau d'une population mais les propriétés simples et fondamentales des éléments qui la composent. Il s'agit donc de valider les hypothèses faites à un niveau individuel en observant un résultat émergent à un niveau collectif. Sans que nous l'ayons explicitement mis en oeuvre et a fortiori démontré, il nous semble qu'implicitement nous intégrons des caractéristiques sociales dans les paramètres de notre modèle pour en évaluer les effets sur la propagation d'épidémies. Pour ne citer qu'un exemple, nous pensons que le taux de contact (paramètre fondamental de la dynamique des épidémies), représenté dans notre modèle par le nombre et le poids des connexions, est très dépendant des comportements sociaux au sein d'une collectivité.

Dans ce cadre, nous avons simulé des situations épidémiques évoluant avec une dynamique spatio-temporelle voisine de celle recueillie par le RNTMT en identifiant les connexions interneuronales aux coefficients d'interactions entre les départements français. Les résultats obtenus montrent comment le comportement erratique des phénomènes épidémiques observés peut être lié à la fois aux interactions entre groupes de population et leur état de réceptivité. Ce résultat a été

choisi parmi de nombreuses simulations pour son côté "spectaculaire". En effet, il reproduit qualitativement le phénomène réel observé à partir des valeurs de paramètres très proches de ceux utilisées en épidémiologie pour définir les épidémies (Serfling) d'une part et d'autre part de données démographiques et géographique donc lié à l'espace physique de propagation.

Cependant, le travail présenté ici doit être considéré comme une étude de faisabilité. Evidement une étude statistique qui établirait par exemple des corrélations spatiales entre les dynamiques du phénomène étudié et celles issues des simulations de notre modèle, le validerait. Le problème est extrêmement complexe et nous étudions la méthodologie pour le résoudre.

Si la nature des éléments de notre modèle reste simple, le modèle possède plus de trente paramètres qui permettent de simuler une combinatoire importante de situations. L'identification et la validation de ces paramètres ont été longuement étudiés. Ceci a permis d'étudier la diversité des comportements possibles et la sensibilité à la fois aux valeurs des paramètres et aux conditions initiales. L'aspect temporel que nous n'avons pas présenté ici, a été abordé au cours de cette étude. Nous avons proposé un modèle explicatif pour l'étude de la question de la périodicité des épidémies de grippe. nous avons montré qu'un couplage critique entre neurones est susceptible de générer et entretenir un rythme [Toubiana et al. en préparation].

Notre objectif est de faire de cette méthode un système prédictif et donc une méthode d'alerte directement utilisable pour les épidémiologistes. Ceci n'est pas encore réalisé, cependant les idées générées au cours de son élaboration ont permis de dresser une liste technique et conceptuelle visant à son amélioration.

Par ailleurs, ce modèle développé pour simuler des épidémies peut être transposé à d'autres phénomènes se propageant dans un espace complexe. Ce modèle pourrait être utilisé dans le cadre de la diffusion des innovations ou encore pour étudier les phénomènes de modes comme celle de la dynamique spatio-temporelle des prénoms pour ne citer que ces exemples.

6. Références

Anderson RM, May RM (1991). *infectious diseases of Humans. Dynamics and control*. Oxford Sciences Publications.

Baroyan OV, Rvachev LA, Basilevsky UV, Ermakov vv, Franck KD, Rvachev MA, Shashkov VA (1971). Computer modelling of influenza epidemics for the whole country (USSR). *Adv Appl Prob* 3: 224-226

Bailey NTJ: (1975) *The mathematical theory of infectious diseases and its applications*. Edited by Griffin, London, 2nd edition.

Cliff AD, Haggett P, Ord JK (1986). *Spatial aspects of influenza epidemics*. Edited by Pion Limited, London

Cliff AD, Haggett P (1992). *Atlas of Disease Distributions, analytic approaches to epidemiological data*, Blackwell Publishers, Oxford

Converse PD. (1949) New laws of retail gravitation, *The J. Of marketing* 14:379-384

Costagliola D., A. Flahault, D. Galinec, Ph. Garnerin, J. Menares, A-J. Valleron. (1991). A routine tool for detection and assessment of epidemics of influenza-like syndrome in France. *Am. J. of Pub. Health*. 81: 97-99.

Le Pont F., A. Flahault, NP. Chau, M. Guiguet, AJ. Valleron (1989) Modèles mathématiques de diffusion de l'épidémie de SIDA. Revue de la littérature. Retrovirus T II, 4: 107-122

Longini, I.M. (1988). A Mathematical model for predicting the geographic spread of new infectious agents. Mathematical biosciences 90:367-383

Reilly W. (1931) The laws of retail gravitation

Toubiana, L., E. Av'Ron, J.-F. Vibert (en préparation) An artificial neural network model for the spread of communicable diseases.

Valleron, A-J., E. Bouvet, Ph. Garnerin, J. Menares, I. Heard, S.Letrait, J. Lefaucheux.: (1986). A computer network for the surveillance of communicable diseases : the French experiment. Am. J. of Pub. Health. 76: 1289-92.

Vibert J-F (1994) La biologie numérique: une approche nouvelle de l'étude du système nerveux. Pathologie Biologie: Editorial invité, sous presse.

Wilson AG. (1967) A statistical Theory of spatial distribution Models. Transportation Research 253-269

La Vie Artificielle d'une société de robots parlants: émergence et changements du code phonétique.

Hervé Glotin¹, Rafael Laboissière²

^{1&2} Institut de la Communication Parlée

INPG, 46 avenue Félix Viallet 38031 Grenoble Cedex, France.

¹E-mail : glotin@icp.grenet.fr ²E-mail : rafael@icp.grenet.fr

Résumé

La parole est un parcours optimisé d'une séquence d'objectifs perceptifs négociés entre un locuteur et un auditeur. Comment sont négociés les objectifs et optimisés les parcours ? Ces questions sont à la base de l'élaboration de futurs systèmes de reconnaissance automatique et de synthèse adaptative de la parole. A partir d'un modèle articulatoire nous avons simulé l'émergence d'un code phonétique dans le cadre d'interactions acoustiques entre robots parlants nommés carls. La seule contrainte de convergence des lexiques des carls porte sur l'évaluation de la coopération entre les robots et sur leurs contraintes articulatoires. Les systèmes qui émergent s'approchent des systèmes les plus fréquents dans les langues du monde. Les algorithmes de type « vie artificielle » ont été mis en oeuvre pour simuler l'évolution d'une telle société d'agents sensori-moteurs.

mots clés : phonétique, autoorganisation, vie artificielle, émergence, communication, coopération, production de la parole.

1. Introduction

Supposez qu'un robot parlant soit capable d'apprendre à produire des gestes articulatoires à partir de sons entendus. Imaginez qu'une assemblée de ces robots essaie d'établir un code de communication orale. Nous serions alors en mesure de trouver des réponses à des questions qui sont au coeur de la recherche en communication parlée. Quelles unités phonétiques choisiront les robots pour communiquer ? En supposant que les simulations soient assez sophistiquées, comment le code phonétique changera au cours du temps et comment apparaîtront les dialectes ? Plusieurs études ont été consacrées à la compréhension de l'émergence du code linguistique à partir des connaissances articulatoires et acoustiques de la production de la parole (Boë et al., 1994). Elles partent toutes de la même question: comment via des interactions sensori-motrices, émergent les systèmes linguistiques ? Ces études sont fondées sur des principes d'optimisation de critères liés à la fois à l'efficacité perceptive et à l'économie motrice. Lindblom (1995) conclue même que : « La parole est au fond un phénomène biologique, et, par conséquent, on ne peut pas regarder la variabilité de la parole comme un accident. Notre examen des données expérimentales nous mène [...] à proposer qu'elle est une conséquence inévitable de son organisation adaptative. »

Notre étude porte sur les changements phonétiques résultants d'échanges entre instances de robots parlants (des modèles articulatoires) au sein d'interactions en production et en perception. Nous avons conçu un cadre de modélisation suffisamment puissant pour que les phénomènes de changements phonétiques soient mis en évidence, ainsi que les mécanismes d'émergences des systèmes phonologiques. Dans cette modélisation, les ensembles de recherche sont purement continus et les interactions sont locales¹. Ces deux aspects sont novateurs dans ce domaine de recherche en phonétique.

Commentaire [mn1] : 1

¹ Contrairement aux travaux de Boë et al., 1994 (voir fin de la section 2).

Le robot parlant et sa fonction d'inversion permettant d'inférer les séquences articulatoires en fonction des buts acoustiques donnés, étaient déjà élaborés à l'Institut de la Communication Parlée (ICP) (Laboissière, 1992). Nous nous sommes donnés des critères et des mécanismes d'une bonne communication (meilleure coopération entre agents), en modélisant l'interaction entre production et perception dans le contexte d'une robotique anthropomorphique de la communication parlée. Nous avons construit un cadre d'interaction entre deux robots qui se « parlent » et déterminé comment les robots adaptent en conséquence leurs codes internes avec un nombre fixe d'items dans leur lexique. Nous avons structuré une société de robots, agents sensori-moteurs, en forçant les échanges entre tous, et en forçant l'utilisation de tous leurs items. Nos algorithmes sont du registre de la « vie artificielle ». Nous présenterons tout d'abord les recherches en phonétique afin de montrer comment nous pouvons les transposer dans notre approche. Puis nous aborderons notre monde AGORA et ses algorithmes de vie artificielle ainsi que les résultats des simulations. Ensuite nous concevrons un système d'émergence de dialectes basé sur la théorie des niches écologiques en vie artificielle. Enfin nous ouvrirons une perspective sur l'étude des changements phonétiques à travers la théorie de la variabilité adaptative (Lindblom, 1987).

2. Les systèmes phonétiques

Lors de la production des sons de la parole, le conduit vocal peut être excité selon un mode d'excitation quasi-périodique. L'appareil articulatoire renferme des cavités résonnantes et à ce titre vient renforcer certaines fréquences, appelées formants. On caractérise les voyelles par leurs formants classés par ordre croissant de fréquences (F1, F2, F3 etc ...). Les contraintes articulatoires, dépendantes de la géométrie du conduit vocal humain, restreignent les formants dans un certain domaine clos. Dans le plan F1 et F2 ce domaine s'approche d'un triangle, les voyelles [a i u] occupant ses sommets. Chaque langue, possède un certain ensemble de voyelles, dites canoniques qui constituent un système vocalique. Le critère de qualité des systèmes vocaliques repose sur la maximisation des distances perceptives inter-voyelles (Liljencrants & Lindblom, 1972). Mais ce contraste est limité par une exigence d'économie articulatoire qui est antagoniste à la demande en contraste perceptif. Les données expérimentales en production montrent l'existence de deux « forces » antagonistes spécifiant le contrôle du système moteur de la parole : l'exigence de distinction perceptive (du point de vue de l'auditeur) et de la demande d'économie articulatoire (du point de vue du locuteur) (Schwartz, 1995). Dans le modèle de prédiction de Bosch (1986), *les voyelles préférentielles dans les systèmes sont celles qui au niveau articulatoire sont produites avec une déviation minimale par rapport à une position neutre, tout en conservant une distinction perceptive maximale.*

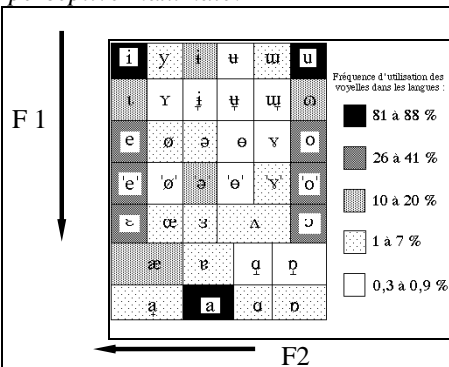


Figure 1 : occurrences des voyelles dans les langues (Vallée, 1994), placées suivant leur deux premiers

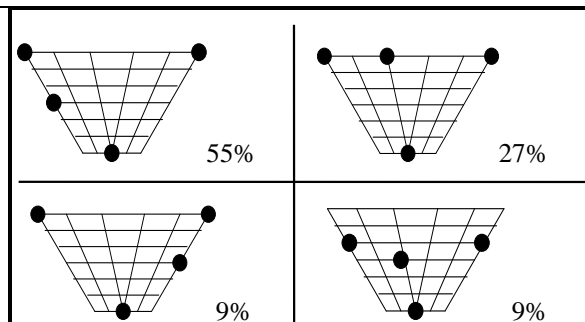


Figure 2 : les systèmes vocaliques à 4 voyelles relevés dans la base de 317 langues UPSID avec leur occurrence (Boë et al., 1994)

Par exemple la plupart des langues du monde qui sont composées de trois voyelles, ont leurs voyelles qui sont placées aux extrêmes du triangle vocalique. Mais les langues aborigènes à trois voyelles ne suivent pas cette règle : leurs voyelles sont plus centrées dans le triangle vocalique (Butcher, 1994). Néanmoins les grandes tendances des systèmes vocaliques (cf. Fig. 1 & 2), sont expliquées par des critères de perception et de production.

Liljencrants et Lindblom (1972) proposent un modèle de prédiction des systèmes vocaliques suivant le critère global de maximisation de la somme de toutes les distances inter-voyelles dans l'espace perceptif, ce qui assure une efficacité perceptive maximum. Cette théorie de dispersion maximale, améliorée par un coût intra-vocalique (Boë et al., 1994) donne des prédictions adéquates.

3. Un modèle d'évolution de coopération

MacLennan et Burghardt (1994) présentent un modèle d'évolution de coopération (Axelrod, 1984). Ils étudient le comportement de créatures artificielles appelées *simorgs*. Ils montrent comment à partir d'interactions locales entre les *simorgs* dans un environnement simple, un langage propre à la communication peut émerger, sur le seul critère de coopération.

Dans ces simulations, la coopération prend la forme d'une action pertinente sur un environnement régi par un automate interne. Dès lors, il est possible de comptabiliser les bénéfices d'une bonne coopération entre deux créatures. Ce qui donne lieu au calcul d'une fitness qui est utilisée pour choisir les créatures qui seront élevées et celles qui seront tuées. Une simulation consiste en des «cycles d'élevage», qui sont constitués de «cycle d'environnement», eux mêmes composés de «cycles d'action». Après un certain nombre de cycles «d'environnement et d'action», selon la valeur de leur fitness, les créatures sont tuées, ou choisies comme parents pour créer une nouvelle population en croisant (combinaison des chaînes de symboles), les tables de transitions de leur automate par un algorithme génétique (Goldberg, 1994). Donc plus la table de transition d'un *simorg* est adaptée, plus celui-ci se reproduit par le biais de l'algorithme génétique que nous décrivons dans le paragraphe suivant. Les enfants possèdent des propriétés issues de la combinaison de celles des parents, et sont plus ou moins bien adaptés. Finalement toute la société atteint un fort niveau de coopération grâce à un langage émergent.

Ce travail nous a inspiré pour nos recherches, même si la notion de communication par langage n'est pas explicitement introduite dans nos simulations. Les organismes coopérants de notre société, à l'instar des *simorgs*, n'attachent pas de sens aux symboles produits. Nos recherches ne portent pas sur la genèse du langage, mais seulement sur le choix des symboles phonétiques, en supposant le langage pré-existant.

4. AGORA et les carls

AGORA est un ensemble de robots parlants, chacun nommé *carl*², qui interagissent les uns avec les autres. Le but est de les faire converger vers des symboles phonétiques communs propres à la communication en simulant les mécanismes de base de perception et de production de la parole.

4.1. Le modèle articulatoire

Un *carl* est une instance d'un modèle articulatoire issu d'une analyse statistique de la forme du conduit vocal (Maeda, 1988). Avec huit paramètres de commande (écartement de la mâchoire, le dos et le corps de la langue, le déplacement horizontal et vertical du bout de la langue et celui des lèvres et la hauteur du larynx), le modèle articulatoire génère la coupe sagittale du conduit vocal

² Cerveau Analytique de Recherche en Linguistique / Cooperative Agent for Research on Linguistic

(cf. Fig. 3). A partir de celle ci, on calcule la fonction d'aire et la fonction de transfert du conduit vocal qui fournit les formants correspondants. Un *carl* est donc défini par la matrice des paramètres articulatoires (état interne), et des formants associés à chaque item vocalique (état externe accessible à chaque *carl*) donnés par la fonction directe du modèle.

Certes, le travail décrit ici pourrait être étendu à de vrais robots utilisant des médias de communication arbitraires, mais nos *carls* sont manifestement anthropomorphiques.

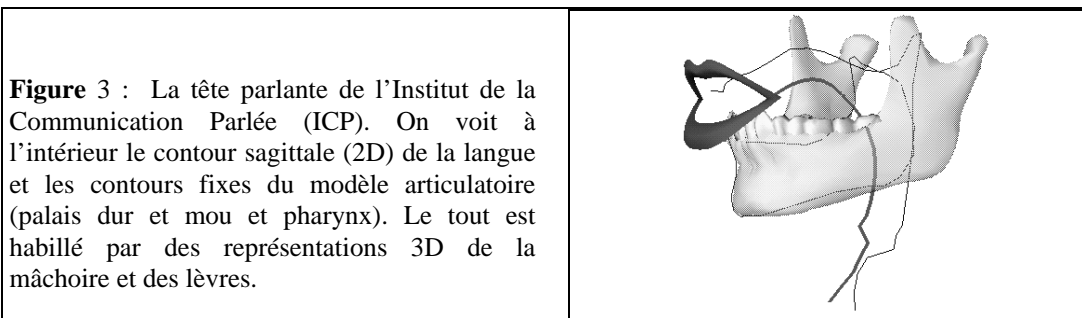


Figure 3 : La tête parlante de l'Institut de la Communication Parlée (ICP). On voit à l'intérieur le contour sagittale (2D) de la langue et les contours fixes du modèle articulatoire (palais dur et mou et pharynx). Le tout est habillé par des représentations 3D de la mâchoire et des lèvres.

4.2. Les interactions phonétiques des carls

Les *carls* établissent des communications orales deux par deux en utilisant leurs symboles internes, chaque *carl* ayant son lexique d'items vocaliques, initialement tirés au hasard autour de la voyelle neutre (obtenue par le relâchement du conduit vocal, paramètres articulatoires nuls). Lors d'une interaction (cf. Fig. 5), le *carl* émetteur produit un symbole interne choisi au hasard qui est perçu par le *carl* récepteur. Le *carl* récepteur compare au niveau acoustique ce symbole perçu avec ceux de son lexique et il les adapte de la façon suivante: l'item le plus proche est approché, les autres sont éloignés, suivant des forces d'attraction et de répulsion (cf. Fig. 4). Cette règle simple d'interaction locale peut faire émerger des codes phonétiques comme nous le montrerons plus tard. Notre espace perceptif a pour unité le bark qui est l'échelle physiologique des fréquences. Dans cet espace la distance perceptive Φ , entre deux items vocaliques d'un *carl* A et un *carl* B est (Boë et al., 1994):

$$\Phi = \sqrt{\left((F1_a - F1_b)^2 + \lambda^2 (F2'_a - F2'_b)^2 \right)}$$

$F1_x$ et $F2'_x$ sont les formants en bark (Schroeder et al., 1979). $F2'$ le deuxième formant effectif est calculé d'après $F2$, $F3$ et $F4$ (Schwartz, et al., 1995) et pondéré par λ (Boë et al., 1994).

Chaque item du lexique du *carl* récepteur subit donc un changement acoustique. Puisque les symboles internes sont représentés en terme de position articulatoire il faut que le *carl* soit capable d'inverser une variation formantique en une variation articulatoire. Or la fonction inverse formants vers articulatoires n'est pas unique, car il s'agit d'un problème à excès de liberté (4 entrées pour 8 sorties). Nous implémentons

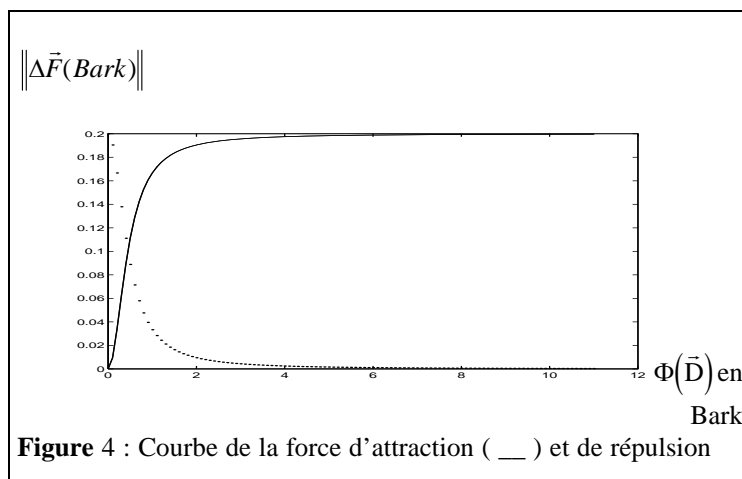


Figure 4 : Courbe de la force d'attraction (—) et de répulsion (···)

une fonction d'inversion qui assure un minimum de changement articulatoire.

(...) dans l'espace acoustique entre deux items. Chaque vecteur force a pour support la droite (item_locuteur, item_auditeur).

Dans le cadre de notre travail, étant donné que tous les *carls* sont identiques au niveau des capacités de production, nous ne traitons pas le problème de normalisation (Payan, 1992), à savoir comment l'auditeur ramène les symboles produits par le locuteur à son propre système de production.

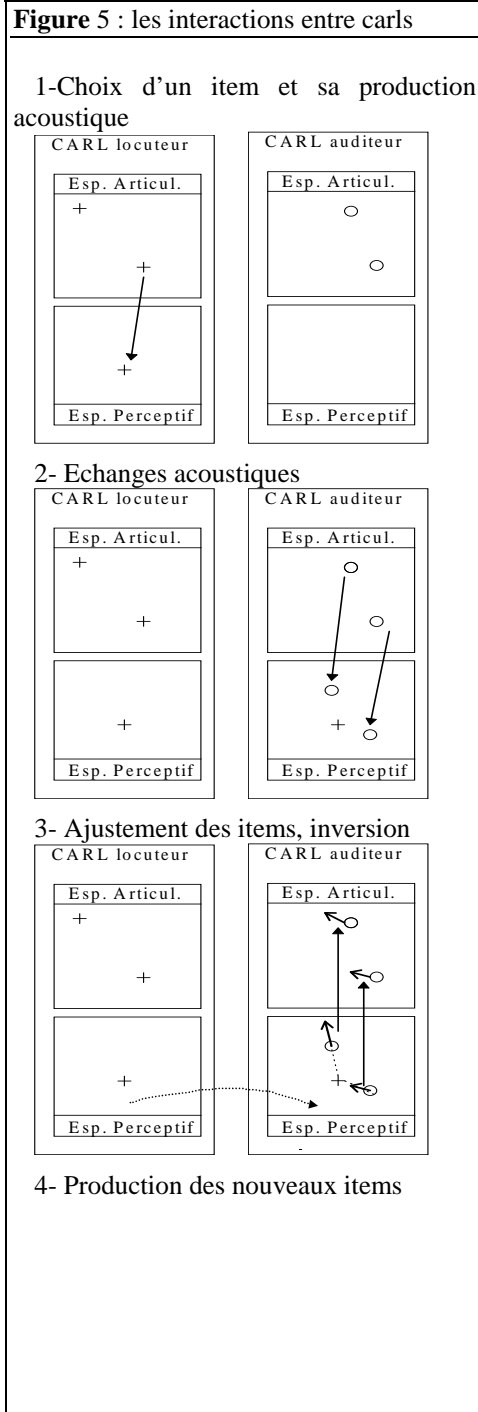
4.3. L' algorithme d'apprentissage

Nous définissons une ronde d'interactions comme une suite aléatoire de couples de tous les *carls* qui vont communiquer successivement. On tire aléatoirement autant de rondes que de nombre de cycles de la phase d'apprentissage. Le cumul des normes des vecteurs de variation articulatoire qu'un *carl* effectue à chaque interaction nous donne son coût d'adaptation. Moins un *carl* change son lexique au cours des interactions, plus son langage est adapté pour communiquer avec la société. Ceci est le coeur de notre approche. Nous pouvons dès lors calculer une fonction rendant compte du degré de coopération d'un *carl* dans la société : sa *fitness*.

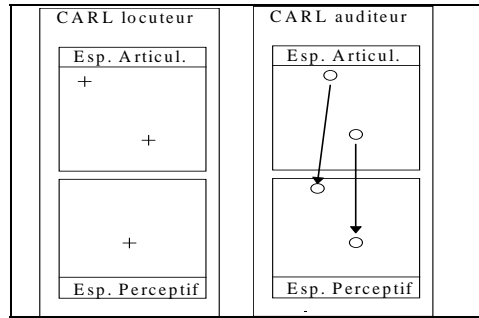
$$\text{fitness}(\text{carl}) = \frac{1}{1 + \text{cout}(\text{carl})} \quad \text{avec}$$

$$\text{cout} = \sum_{\text{interactions}} \sum_{\text{items}} \|\Delta \vec{Articu}\|$$

Nous infligeons à notre société une pression de sélection, pour une convergence vers une coopération croissante entre tous les *carls*. On favorise la perpétuation des *carls* les plus adaptés dans une nouvelle population de *carls* pour un nouvel apprentissage (cf. Fig. 6). L'outil d'exploration pour notre modèle, sera un algorithme génétique ou pseudo-génétique, en ce sens que c'est un algorithme qui ne contient pas de notion de mutation ni d'autre type d'opérateurs aléatoires et que la taille des populations reste faible. Cet outil ne reflète aucun aspect anthropologique dans nos simulations, et notre technique génétique qui est du type Lamarckien n'a pas pour vocation de refléter

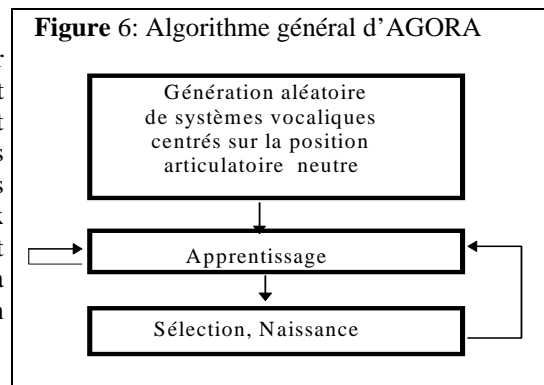


un aspect évolutif, mais seulement prospectif.



Une fois l'apprentissage terminé, les *carls* sont classés par ordre croissant de *fitness* et sont dupliqués dans la nouvelle génération. Notre fonction de duplication des meilleurs *carls* est appliquée en traitant les cas par ordre décroissant de *fitness*, et tant que le nombre d'individus est insuffisant pour obtenir une population de taille identique à la précédente. La règle est qu'un individu est dupliqué en un nombre de descendants proportionnel à sa *fitness* (Baker 1985). Puis une moitié de cette sous population de meilleurs *carls*, est croisée entre elle, l'autre moitié reste intacte, suivant la technique de l'élitisme qui accélère la convergence (Goldberg, 1994). Le croisement consiste en un échange de la moitié des items des lexiques de deux *carls*. Par exemple, pour quatre items on fixe le point de croisement entre le second et le troisième item. Une fois née, la nouvelle génération reprend le même cycle d'apprentissage que celui de ses parents. On arrête la simulation lorsque l'on juge la convergence suffisante.(cf. Fig. 6).

De façon à assurer que l'apprentissage acquis par les parents sera perpétué chez les enfants, il faut que les régions du système vocalique qui sont occupées par les parents apparaissent aussi dans le lexique des enfants. Pour cela nous avons conçu un algorithme d'appariement de deux lexiques avant leur croisement. L'appariement produit les couples d'items inter-lexiques, la distance perceptive entre chaque item d'un même couple étant minimale.



4.4. Validation de l'algorithme d'apprentissage

Il nous est apparu nécessaire de démontrer la cohérence de notre algorithme en supprimant les fonctions articulatoires pour tester le sous-problème « acoustique seule ». Nous avons donc réalisé une version d'AGORA où les *carls* n'évoluent que dans le seul espace des formants, leur coût d'adaptation étant dans ce cas particulier la somme des variations acoustiques pour chaque item à chaque échange, ce qui garantit la coopération (convergence des lexiques) des individus dans l'espace acoustique. En effet, on minimise ainsi le déplacement des voyelles à chaque génération car on sélectionne les *carls* dont le lexique a peu bougé, donc ceux dont le niveau de coopération est élevé car leur demande d'adaptation a été faible.

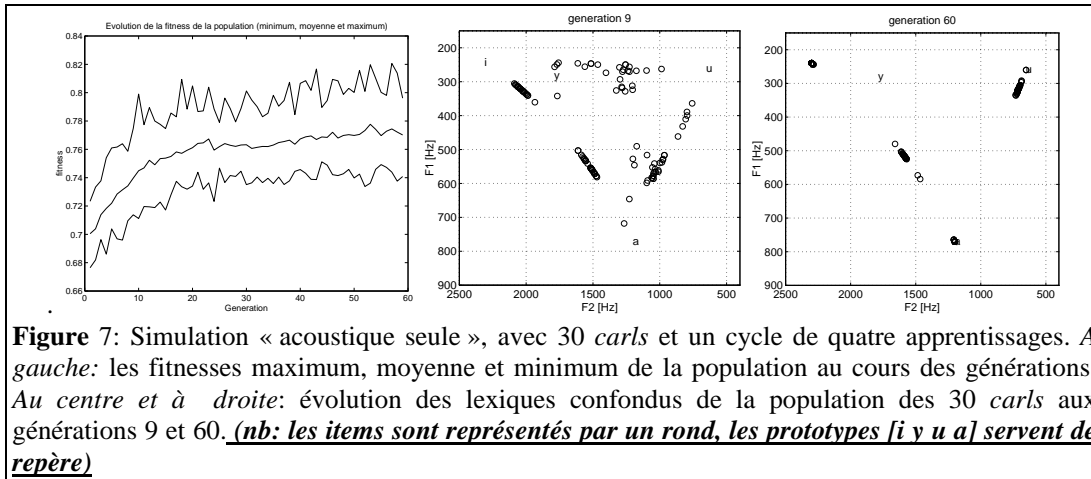


Figure 7: Simulation « acoustique seule », avec 30 *carls* et un cycle de quatre apprentissages. *A gauche:* les fitnesses maximum, moyenne et minimum de la population au cours des générations. *Au centre et à droite:* évolution des lexiques confondus de la population des 30 *carls* aux générations 9 et 60. (nb: les items sont représentés par un rond, les prototypes [i y u a] servent de repère)

Dans cette version d'AGORA l'espace acoustique a deux dimensions, les réalisations acoustiques sont bornées a priori par un triangle dont les sommets sont les formants F1 et F2 pour [i a u]. Durant l'apprentissage tout point qui sort du triangle est projeté orthogonalement sur le bord, ce qui simule les contraintes de production. La fitness (cf. Fig. 7) suit l'évolution prédite dans la section 4.3 : la coopération entre les agents *carls* croît. La convergence des items est observée pour *chaque* lexique des individus vers les voyelles [i a u] et [e'] dès la soixantième génération (cf. Fig. 7), ce qui correspond au système vocalique à quatre voyelles le plus fréquent dans les langues du monde (cf. Fig. 2). Nous pouvons maintenant explorer l'espace articulatoire.

5. Résultats

Nous passons maintenant à des simulations perceptuo_motrices, c'est à dire que nous nous plaçons dans les espaces acoustique et articulatoire comme le décrit la section 4.2. La simulation réalisée comporte 4 *carls* possédant chacun 4 items, pour 6 cycles (6 rondes de dialogues durant l'apprentissage). Nous avons filtré la moyenne de la fitness au cours des générations (cf. Fig. 8) pour ne rendre compte que des variations significatives des populations. Aux générations correspondantes à un pic de la fitness moyenne filtrée, les *carls* convergent vers un lexique commun.

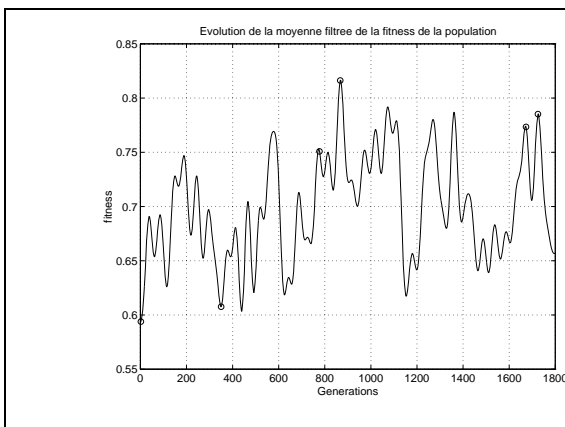


Figure 8: Evolution du degré de coopération des *carls*: la fitness moyenne filtrée de la population sur les 1800 premières générations. Les ronds indiquent les fitness moyennes des générations détaillées par la suite. Le filtre utilisé est du type Butterworth d'ordre 10 avec une fonction de coupure de 0.05 génération⁻¹.

Les chutes de la fitness peuvent être attribuées à une « catastrophe génétique » (MacLennan et al. 1994), ou à une instabilité des fonctions articulatoires. La dispersion acoustique des items (cf. Fig. 9) est observée à la génération 351. La fitness à cette génération est faible (cf. Fig. 8), les *carls* coopèrent pas. Le pic de fitness moyenne pour la 777^e génération correspond déjà à l'agrégation des lexiques qui tendent vers [i] [a] [u] [y] (cf. Fig. 9). Finalement à la génération 867, qui a une forte fitness moyenne, chaque *carl* a un lexique qui a convergé vers des voyelles adéquates : le [a] et le [u] y sont approchés et les deux autres voyelles s'approchent du [i] et d'une voyelle centrale assez avancée. Ceci est proche du système [i, a, u, ε] qui est un système vocalique existant dans les langues du monde (cf. Fig. 2).

La chute de la fonction fitness vers la 1150^e génération correspond à une réorganisation des systèmes (cf. Fig. 9). Nous observons ce phénomène sur l'ensemble de la simulation. Le système de la génération 1670 (cf. Fig. 10), est sensiblement identique à celui de la 867^e, mais le [ε] a basculé vers le [o]. En mettant en évidence les coupes sagittales des *carls* (cf. Fig. 3) à la génération 867 (cf. Fig. 11), nous montrons que chaque *carl* a développé sa propre stratégie articulatoire: par exemple une même voyelle peut être plus ou moins labialisée. A la figure 12, nous montrons que les *carls* 1 et 3 de la génération 1725 coopèrent sur le système [i a u 'o'] (d'après

Fig. 1).

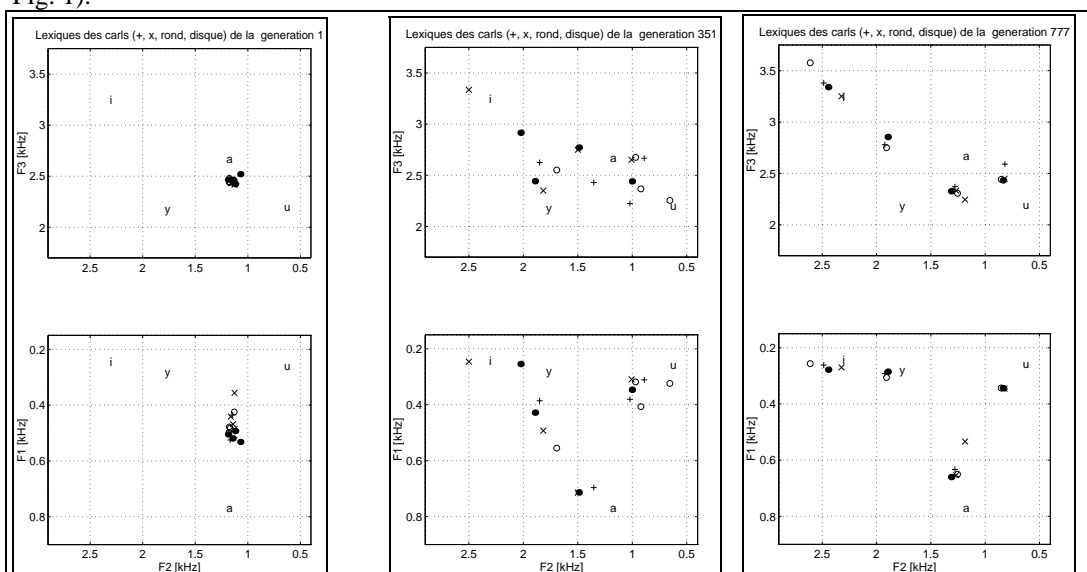


Figure 9: évolution des lexiques confondus de la population aux générations 1, 350 et 777. Les différents lexiques sont représentés par +, x, o, et un disque. Les prototypes [i y u a] servent de repère.

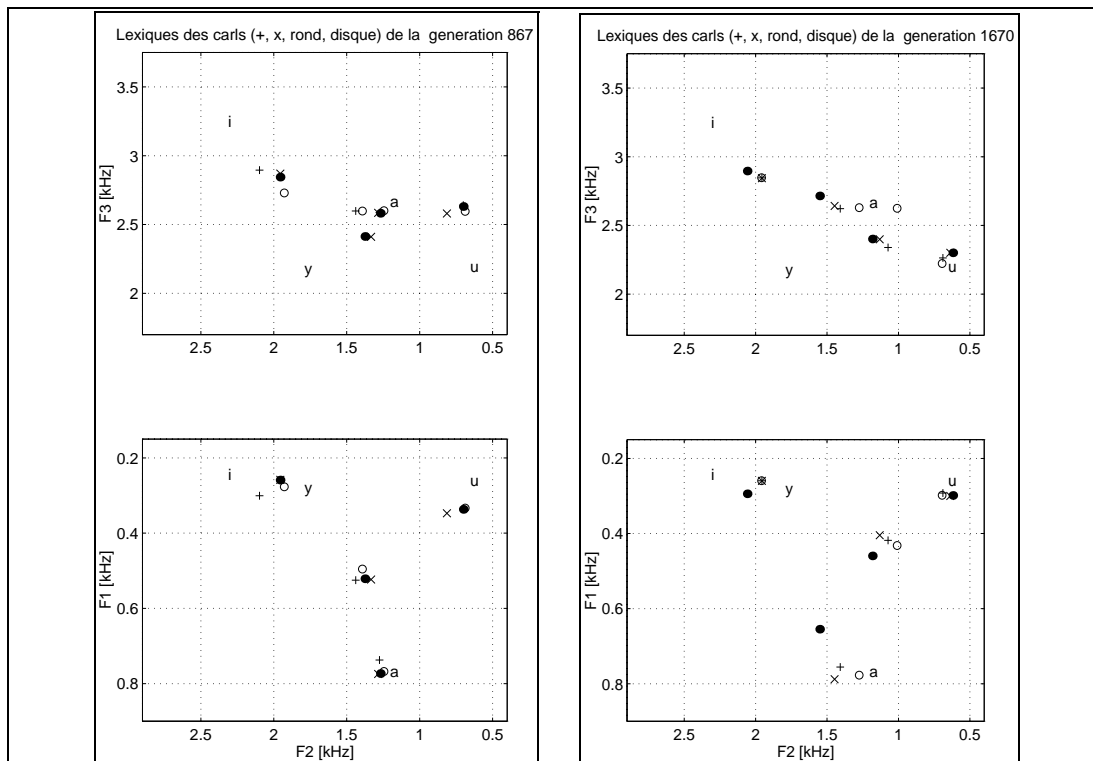
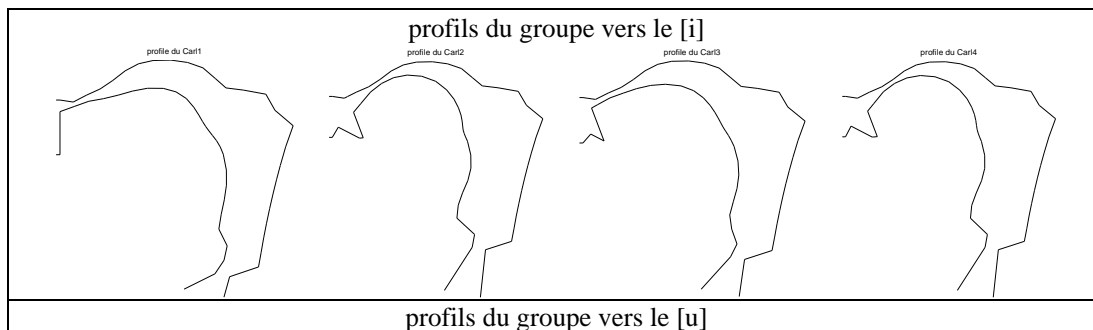


Figure 10: lexiques confondus de la population à la génération 867 à gauche. A droite ceux de la génération 1670. (Dans F1 F2 il y a superposition de quelques points).

Ce dernier est proche d'un des systèmes des langues à quatre voyelles. Les coupes sagittales articulaires de ces deux *carls* (cf. Fig. 13) montrent qu'ils ont choisi une stratégie articuloire différente. Les lexiques des deux autres *carls* sont inadéquats sur deux voyelles, elles n'ont pas été émises par des locuteurs lors de la phase d'apprentissage qui est courte.



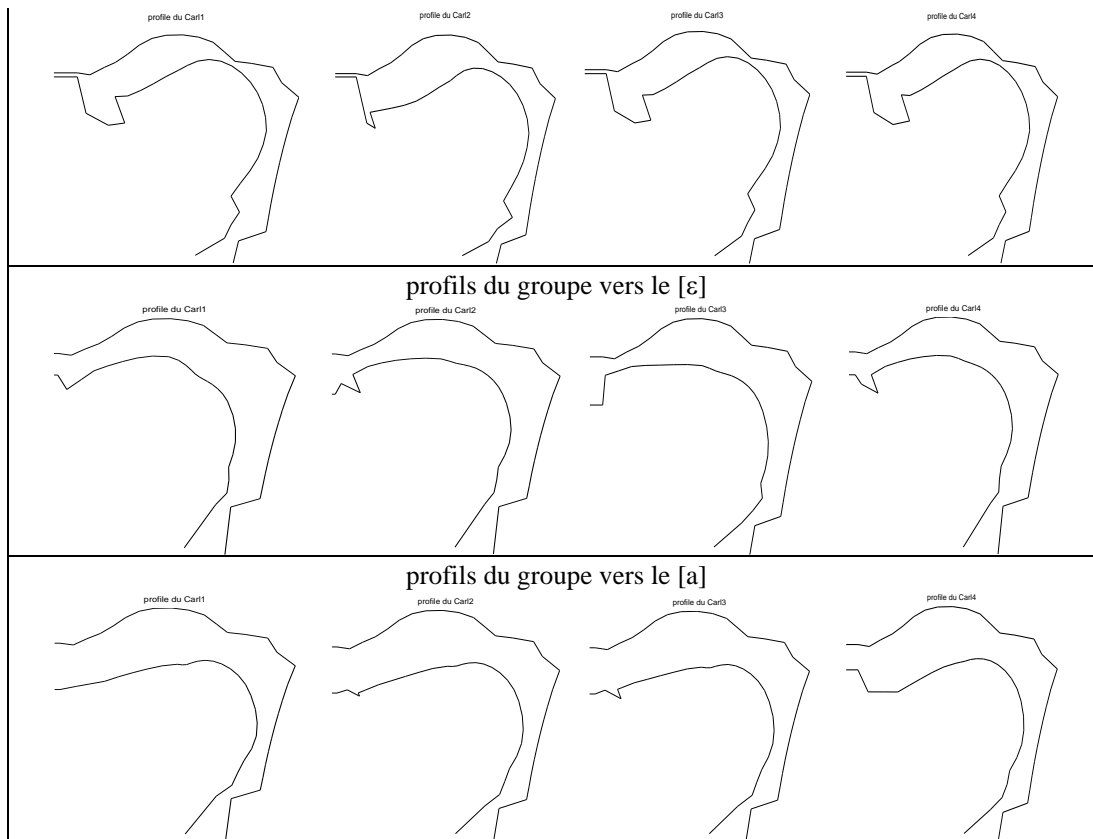


Figure 11 : coupes articulatoires sagittales des *carls* à la génération 867.

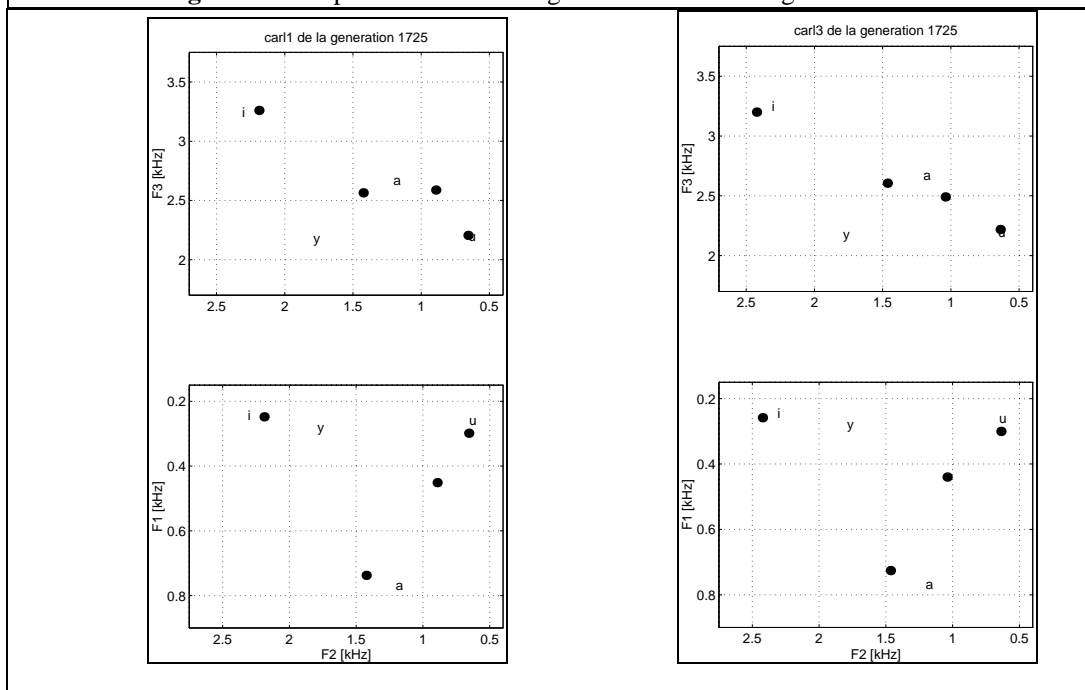


Figure 12 : lexiques du *carl* 1 (à gauche) et du *carl* 3 (à droite) de la génération 1725.

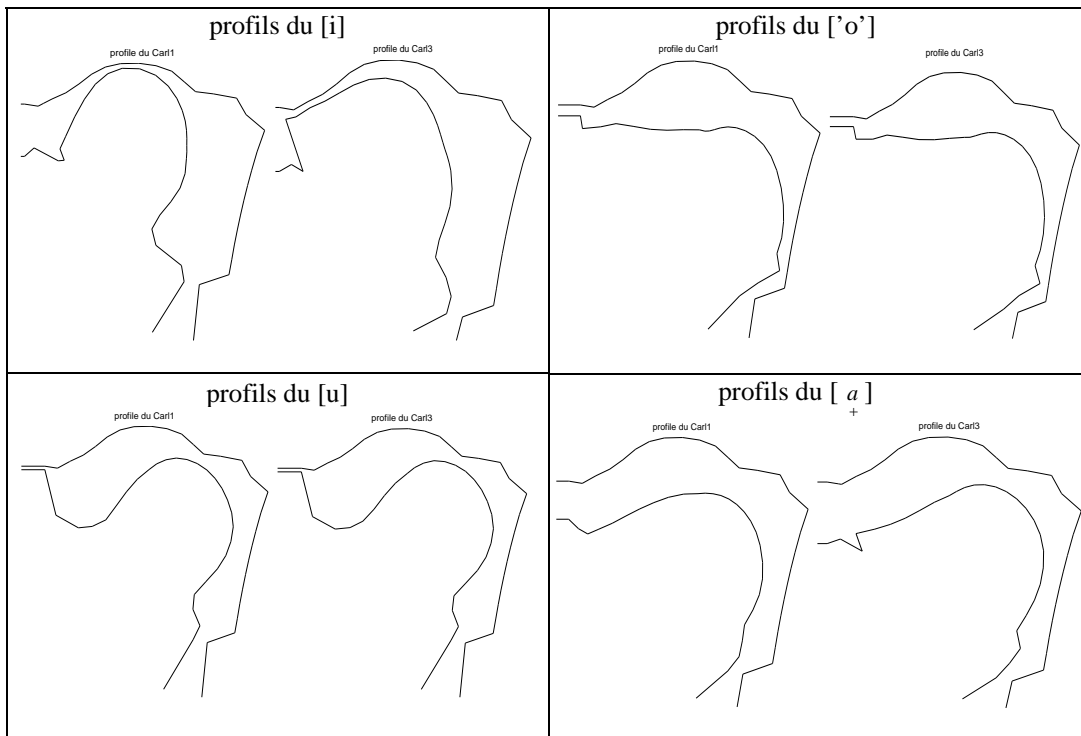


Figure 13 : coupes articutoires sagittales des *carls* 1 et 3 à la génération 1725 groupées par proximité acoustique.

6. Discussion

Nous avons montré que sans contrainte globale, uniquement par effet de coopération, des systèmes vocaliques propres à la communication parlée ont pu émerger de l'interaction entre agents sensori-moteurs. Ce sont les systèmes utilisés dans les langues du monde que l'on retrouve. Les systèmes les plus adéquats à la communication parlée émergent aux générations qui ont la fitness moyenne la plus élevée. Les produits acoustiques restent confinés dans un espace vocalique par le seul fait de la physiologie du système articuloire. De plus, chaque *carl* a élaboré sa propre stratégie articuloire pour coopérer, même si le produit acoustique est identique.

Nos résultats sont cohérents avec les statistiques sur les langues du monde (cf. Fig. 2). L'efficacité perceptive et l'économie motrice, sont les principes qui semblent régir la formation des systèmes linguistiques (Schwartz, 1995). Notre étude structure une première approche de la sorte dans le domaine de la phonétique.

Une fitness plus lisse peut être obtenue avec des populations de plus grande taille. Les fluctuations de la fitness sont dues au faible nombre de cycles d'apprentissage par rapport au nombre de voyelles, ainsi qu'à la trop petite taille de la population ce qui nuit à l'équiprobabilité des choix des voyelles lors des interactions entre les *carls*. Notre algorithme explore de la sorte l'espace des solutions, donc il n'est pas étonnant de tomber de temps en temps sur des configurations intéressantes. De récentes simulations montrent en effet que l'évolution de la fitness est similaire à celle de la simulation acoustique seule (cf. Fig. 7), lorsque l'on prend une population de 16 *carls*, pour quatre cycles d'apprentissage. Le temps des calculs prohibitifs pour l'instant des fonctions articuloires, est le frein à des simulations à plus grande échelle.

Un des résultats intéressants malgré l'effet de bord du aux valeurs des paramètres qui sont trop petits, est qu'il peut apparaître différents systèmes au cours d'une même simulation, séparés par des chutes de la fonction fitness, ce qui pourrait constituer un germe d'explication pour les phénomènes de changement phonétique.

7. Perspectives

L'éventail de recherche en phonétique qui est ouvert est vaste et novateur : l'émergence des systèmes vocaliques et syllabiques jusqu'à l'évolution phonétique des langues en recherchant des trajectoires optimales dans les phénomènes de coarticulation. L'évolution en continu dans les substrats acoustiques et articulatoires offrent à nos simulations un espace de recherche très fin et inédit. Nous présentons dans la suite deux possibilités d'extension de ce travail à travers la théorie des niches écologiques, et la dynamique des changements phonétiques.

7.1. Des niches écologiques pour la formation des dialectes

Comme Lindblom et Engstrand (1990) le proposent, une quantification de facteurs socioculturels serait utile pour améliorer les prédictions : la formation et l'évolution des systèmes de sons seraient une combinaison de paramètres sociaux, perceptifs et articulatoires. Nous proposons un modèle de l'émergence de dialectes dans le monde des *carls* pour donner une dimension supplémentaire à notre simulation, propre à l'étude d'apparition de dialectes, dont nous montrons la faisabilité par AGORA à travers la théorie des niches écologiques (Goldberg, 1994). Grosso (1985) a démontré que l'usage d'une géographie combinée à l'exploration génétique peut être un moyen utile d'assistance à la formation de sous populations variées, chaque sous population appartenant à sa niche écologique. Considérons qu'à un instant donné différents systèmes vocaliques sont en compétition, phénomène que nous avons observé dans nos simulations. Les niches écologiques permettent alors le développement des deux systèmes, deux dialectes, qui tendent vers leur propre optimum, grâce à des probabilités d'échanges qui sont fonction des positions géographiques des deux *carls* communiquant.

7.2. Vers un modèle de la théorie Hypo/Hyper : dynamique et étude des changements phonétiques

La théorie motrice (Liberman et al., 1985) et la théorie Quantique (Stevens, 1989), postulent l'existence des invariants physiques en parole soit articulatoires pour l'une, soit acoustico-articulatoires pour l'autre. Cette invariance n'est pas de nature phonétique pure, mais peut être définie au niveau de la compréhension de l'auditeur : « *The invariance problem is not a phonetic issue at all for ultimately invariance can be defined only at a level of listener comprehension* » (Lindblom, 1987). Le locuteur adopterait une stratégie de Variabilité Adaptative résolvant un compromis entre la demande de compréhension et l'effort articulatoire pour réaliser la tâche. Cette adaptation du locuteur permet d'expliquer pourquoi les paramètres physiques décrivant la parole peuvent être variables pour un message linguistique invariant, ce qui est à l'origine d'un grand nombre de phénomènes de variabilité en parole. La théorie Hypo/Hyper (H&H) (Lindblom, 1995), propose aussi des explications pour les changements phonétiques. L'attention de l'auditeur peut être mobilisée par la prononciation d'un mot, autant que par son sens : « [...], on peut s'imaginer qu'une sélection, ou un rejet, d'une innovation par d'autres locuteurs est un résultat d'une évaluation qui se fait, le plus souvent, d'une manière très implicite » (Lindblom 1995). Lindblom propose un schéma (Fig. 14) de source des adaptations. La structure H&H se situerait au niveau phonologique, pour la sélection et l'adaptation.

Les variations phonétiques proviendraient des productions entre « SPEAKERS » et « LISTENERS ». Nous pourrions simuler des changements phonétiques (Labov, 1994) par AGORA

en concevant des séquences de voyelles et en étudiant les phénomènes H&H. Les transitions vocaliques seraient des indissociables dans les lexiques des *carls*. Elles ne seraient pas stockées en tant que telles parce qu'elles sont susceptibles de changer selon le degré H&H. Elles seraient plutôt définies par des cibles. Elles posséderaient un caractère non-segmental du fait que les cibles n'appartiennent pas à un alphabet limité (en nombre), donné préalablement. Notre algorithme, modulo quelques modifications ferait émerger des syllabes qui maximisent à l'intérieur d'elles mêmes leur contraste perceptif entre le départ et l'aboutissement du geste, et qui minimisent leur contorsion articulatoire. Le coût articulatoire intra-syllabique serait obtenu en calculant la distance entre les paramètres articulatoires de la consonne et ceux de la voyelle; il agirait au niveau de l'adaptation, et notre algorithme de coopération agirait au niveau de la source de variation phonétique.

A plus long terme une telle étude nous amènera à mieux comprendre comment sont optimisés les parcours et négociés les objectifs de la communication parlée, ce qui est un atout pour l'élaboration de systèmes de reconnaissance automatique et de synthèse de la parole.

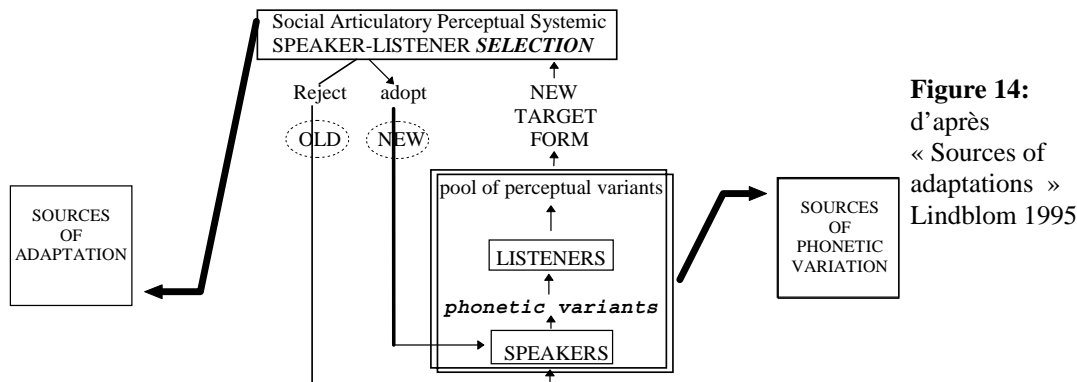


Figure 14:
d'après
« Sources of adaptations »
Lindblom 1995

Remerciement

Nous remercions vivement Arturo Galvan pour l'implémentation des fonctions du modèle articulatoire.

Références

- Axelrod, R. (1984). *The evolution of coopération*. New York: Basic Books.
- Baker, J.E. (1985). Adaptive selection methods for genetic algorithms. *Proceedings of an international Conference on Genetic Algorithms and their applications*, pp 101-111.
- Boë, L.-J., Schwartz, J.L. & Vallée, N. (1994). The prediction of vowels systems: perceptual contrast and stability. In E. Keller (Ed.) *Fundamentals of speech synthesis, and speech recognition*, pp 185-212.
- Bosh, L.T. (1986). Architecture of vowels systems. *Proceedings from the institute of phonetic sciences of the University of Amsterdam*. The University of Amsterdam. 55-72
- Butcher, A. (1994) On the phonetics of small vowels systems: evidence from australian languages. *Proceedings of the fifth Australian International Conference on Speech Science and Technology*. Volume 1. The University of W. Australia.
- Caraty, M. (1994). *Introduction à la parole*. Laforia, Université Pierre et Marie Curie, Paris VI.
- Goldberg, D. (1994). *Algorithmes génétiques, exploration, optimisation et apprentissage automatique*. Addison-Wesley.
- Grosso, P.B. (1985). *Computer simulation of genetic adaptation: Parallel subcomponent interaction in a multilocus model*. Doctoral dissertation, University of michigan.
- Heudin, J.C. (1994). *La vie artificielle*. Paris. Hermès.
- Kugler, P. & Turvey, M.T. (1988) Self organization, Flow fields and information. *Human Movement Science*, 7,97-129.
- Laboissière, R. (1992). *Préliminaires pour une robotique de la communication parlée: inversion et contrôle d'un modèle articulatoire du conduit vocal*. Thèse de docteur de l'Institut National Polytechnique de Grenoble.
- Labov, W. (1994). *Principles of linguistic change*. Blackwell. Cambridge.

- Lieberman, A.M. & Mattingly, I.G. (1985). The motor theory of speech perception revisited. *Cognition*, 21:1-36.
- Liljencrants & Lindblom, B. (1972). Numerical Simulation of Vowel quality systems: The role of perceptual contrast. *Language* 48, 839-862.
- Lindblom, B. (1987). Adaptive variability and absolute constancy in speech signals: two themes in the quest for phonetic invariance. *Proceedings of the Xith International Congress of Phonetic Sciences*, Tallin, Estonia, Volume 3, 9-18.
- Lindblom, B. (1990). Explaining Phonetic Variation. A Sketch of the H&H Theory. *Speech Production and Speech Modelling*, Hardcastle W.J. & Marchal A. (Eds.), Academic Publishers, pp 55-68.
- Lindblom, B. (1995). Approche intégrée de la production et de la perception. In Méloni H. (Ed.) *Fondements et perspectives en traitement automatique de la parole*. Actes de l'école thématique organisée par le GDR-PRC Communication Homme_Machine. Université d'Avignon et Pays de Vaucluse. pp 1-8.
- Lindblom, B. & Engstrand, Q. (1990). In What sense is speech Quantal. *Phon. Exp. Res.Inst. Ling. Univ. Stockholm* 8, 1-20.
- Lindblom, B., Guion, S., Hura, S., Moon, S.J. & Willerman, R. (1995, à paraître). Is sound change adaptive? Department of Linguistics, University of Texas at Austin.
- Lindblom, B., MacNeilage, P. & Studdert-Kennedy, M. (1984). Self-organizing processes and the explanation of phonological universals. In *Explanation of Language Universals*. Mouton, pp 182-203.
- MacLennan, B.J. & Burghardt, G.M. (1994). Synthetic Ethology and the Evolution of Cooperative Communication. In *Adaptive Behavior*. Cambridge, MA: MIT Press. Vol 2, N°2, 161-188
- Maeda, S. (1988). Improved articulatory model. *Journal of the acoustical society of america*. Vol 2, pp 146.
- Meyer, J.-A. & Wilson, S.W. (1991). *Form animals to animats*. Cambridge, MA: MIT Press.
- Payan, Y. (1992) *Normalisation de l'espace de réalisation acoustique des locuteurs: intégration de contraintes de production*. D.E.A. Signal Image Parole, Institut National Polytechnique de Grenoble.
- Robert-Ribes, J. (1995). *Modèles d'intégration audiovisuelle de signaux linguistiques*. Thèse de docteur de l'INPG
- Schroeder, M.R., Atal, B.S. & Hall, J.L. (1979). Objective measure of certain speech signal degradation based on masking properties of human auditory perception. In B. Lindblom and S. Ohman (Eds.), *Frontiers of Speech Communication Research* (pp.217-229). London: Academic Press.
- Schwartz, J.L., Boë, L.-J., Vallée, N. & Abry, C. (1995). The dispersion - Focalisation Theory of vowel systems, submitted to journal of Phonetics.
- Schwartz, J.L. (1995). Perception de la parole : des représentations sensori-motrices à l'émergence de systèmes linguistiques. In *Fondements et perspectives en traitement automatique de la parole*, Méloni H.(Ed.),pp9-21
- Stevens, K.N. (1989). On the quantal nature of speech. *Journal of phonetics*, 17(1/2):3-45.
- UPSID (1981). UCLA Phonological Segment Inventory Database. UCLA WPP 53, University of California, LA.
- Vallée, N, L.-J., Boë, Chaize, M., Dequier, P., Dumay, A., Grimont, E., Perrier, P., Payan, Y. (1991). *Codebook and Vowel Prototypes of the Vocal Tract Plant*. In Speech Maps (Esprit/BR N°6975), Deliverable 17, pp 12-29.
- Vallée, N. (1994). *Systèmes vocaliques: de la typologie aux prédictions*. Thèse de docteur de l'INPG
- Varela, F.J. & Bourgine, P. (1992). *Toward a practice of autonomous systems: Proceedings of the first European Conference on Artificial Life*. Cambridge, MA: MIT Press.

Une *société* de mots : le texte

par

un *collectif* d'auteurs :

E. Andreewsky¹, D. Bourcier², C. Parisse¹ et H. Péraïta³

¹INSERM-TLNP, La Salpêtrière, 47, Bd. de l'Hôpital, 75013 Paris-

E-mail : andreews@idf.ext.jussieu.fr - parisse@idf.ext.jussieu.fr

²IDL-CNRS, 27, Rue Damesme, Paris 13^e - E-mail : bourcier@idl.msh-paris.fr

³UNED, Madrid

Résumé

*Nous proposons quelques jalons visant à introduire une discussion générale sur des “formes” aussi complexes qu’usuelles, à savoir le **texte** et la **société**. Dans leurs dimensions physiques ou biologiques, ces formes regroupent des ensembles d’éléments (respectivement, mots ou individus), et sont qualifiées de simplement « collectives ». Mais dès que l’on aborde les aspects plus complexes de leurs éléments (identité, liaisons, dimension cognitive, etc.), il convient de considérer ces formes collectives comme des **systèmes**. Les **analogies** fonctionnelles qu’on observe dans les deux cas devraient contribuer à l’intelligibilité du “collectif au social”, thème général des journées de Rochebrune 1996.*

mots clés : système, texte, société, complexité, analogie

1. - Introduction

Dans toutes les cultures, les conceptions respectives de l’homme, de la nature et du langage sont, d’une manière ou d’une autre, *reliées*. Signalons par exemple que selon Confucius (trad. 1987) “qui ne connaît le sens des mots ne peut connaître les hommes”, ou que selon Schrödinger (1990) “la manière dont l’individu conçoit son identité indique celle dont il conçoit le monde”.

Le statut des relations individu-société est une question majeure en sociologie. Dans le domaine de la linguistique, il en est de même pour les relations mot-texte. Dans les deux cas, ce statut est en général “réduit” à celui d’une relation de type *élément-collection* ; la *complexité* de ce

statut, qui est alors laissée pour compte, invite cependant à une approche *systemique*. Nous présenterons des exemples d'une telle approche qui font apparaître un certain nombre d'*analogies* entre les relations individu-société et mot-texte - contribuant ainsi à l'intelligibilité du thème général "du collectif au social".

2. "Collection" d'individus ou société ?

Considérer les éléments - mots ou individus - comme des "objets indépendants" de ce qui les entourent, et notamment du texte ou du social, traduit une approche réductionniste qui permet difficilement d'expliquer un certain nombre de phénomènes.

Il est intéressant de remarquer qu'en sociologie, si on essaye de concevoir l'*indépendance* individuelle, on se heurte à des problèmes similaires à ceux des *dépendances contextuelles* qui sont au coeur du langage. Les arguments à l'encontre de la sémantique "context free" (cf. Andreewsky, 1991, Shanon, 1993) peuvent être utilisés en effet pratiquement mot à mot pour mettre en évidence l'interdépendance de cette indépendance ... En effet, dans les deux cas, si mots ou individus sont considérés comme des éléments indépendants, on ne peut pas expliquer (même en première approximation) les phénomènes qui ancrent mots ou individus dans le langage ou la société ; comme le soulignait en effet déjà Mead (1934), "si les mots *prennent un sens* à partir de leur présence avec d'autres mots, le *moi* ne se met à *exister* qu'à partir des *interactions interindividuelles*".

2. 1. - Les problèmes liés aux approches réductionnistes

La signification du mot est loin d'être figée (ni donc *stockée et retrouvée*) ; elle est *générée* en texte et en contexte, c'est à dire, comme dans la citation précédente, "à partir de la présence du mot avec d'autres mots". Cette hypothèse de générativité est nécessaire pour rendre compte de la *dynamique des significations lexicales*. En effet, dans la mesure où même la signification du terme le plus simple est potentiellement affectée à l'infini par *tout* ce que le locuteur connaît¹, seul un processus dynamique comme celui de générativité contextuelle peut en rendre compte.

La psycholinguistique expérimentale, qui cherche à construire des expériences comparables à celles des sciences dures comme la physique, est confrontée à ces phénomènes complexes, inextricablement liés à la nature du langage. Mais il est souvent très difficile d'accommoder ces phénomènes en laboratoire. D'où les efforts de cette discipline pour *réduire* ces phénomènes, dans

¹ On peut en juger sur l'exemple d'une exclamation aussi banale que *il n'y a plus d'eau !* Dans le contexte d'un supermarché, elle signale que le rayon de bouteilles d'eau minérale est actuellement vide ; l'item lexical *eau* a alors pour signification *bouteilles*. Si on se trouve "perdu dans le désert", la signification de cet item devient quelque chose comme : *boisson précieuse indispensable à la survie*, et l'exclamation "il n'y a plus d'eau !" est un cri de désespoir ; elle se métamorphose en cri de joie, si on est en train d'écoper sans trêve un frêle esquif au cours d'une régata - la signification de l'item *eau* se transformant alors en : *lest saumâtre, indésirable, ralentissant le bateau*. Elle devient quelque chose comme : *notre concept traditionnel d'H₂O*, avec "il n'y a plus d'eau !" pour commenter des travaux sur la "mémoire de l'eau" susceptibles d'entraîner des remaniements de nos conceptualisations de ce liquide ... Ou comme : *eau lourde*, s'il s'agit d'un drame dans le fonctionnement d'une usine atomique", ou encore comme : *bouillon de légume*, si le drame est ménager, et que la soupe est en train de se carboniser ...

les expériences, à des *éléments simples*. Le résultat de cette attitude est, aujourd'hui plus que jamais, conforme à ce qu'était pour L. S. Vygotsky (1982) la psychologie des années 30 : *un panorama atomiste grandiose de l'esprit humain désarticulé*.

On peut en effet constater que le mot, dans la littérature psycholinguistique, est morcelé en éléments (psycho) linguistiques de base, atomisé en niveaux de représentation différents (sons, phonèmes, morphèmes, syllabes, lexèmes, traits sémantiques, modèles orthographiques, etc.), réduit à ne jamais être considéré dans sa complexité. De plus, seuls les composants les plus physiques de ce mot (dans les modalités visuelle et/ou auditive) et la structure syntaxique, semblent mériter un traitement détaillé sur les plans théoriques et méthodologiques (il s'agit en fait pratiquement des seules approches considérées comme scientifiquement "respectables" au niveau académique). Le "reste", c'est à dire ce qui concerne le conceptuel ou la signification, ainsi que - encore que en moindre mesure - les modèles auditifs, les morphèmes, lexèmes et plus généralement les règles de sous-catégorisation, est ignoré - probablement parce qu'il risque de contaminer la clarté et la précision requises dans toute expérience scientifique ...

2. 2. - Approches systémiques

Le statut des relations individus-société est le débat fondamental de la sociologie, et a déterminé les méthodologies de ce domaine. D'un côté, avec Durkheim (1894), la société est un "objet" *sui generis* et non une somme d'individus. On étudie les "faits sociaux" comme externes aux individus, et on les étudie dans leur globalité. De l'autre, avec Weber (1922), la société est la "somme" des actions des individus ; elle est la simple reproduction macroscopique de toutes les individualités.

Pour les deux conceptions précédentes, l'individuel et le social doivent être compris comme des phénomènes *indépendants*, et l'étude de l'un ou de l'autre est privilégiée selon le cas. Mais soulignons que certaines théories du social (comme la théorie de la specularité, Vullierme, 1994) tentent d'articuler ces phénomènes en termes *d'interdépendance*. Ces théories impliquent des processus dynamiques complexes qui, à la fois, "font *émerger* la société des comportements des individus, et font déterminer à ces derniers leur propre comportement sur la base du *modèle* qu'ils se forgent de leur contexte social". Ce modèle individuel (baptisé par J.-L. Vullierme de *spéculaire*), devient ainsi la source permanente de la *dynamique sociale, qui fait de la société toute autre chose qu'une collection d'individus*.

3. - Analogies mot-individu / texte-société : quelques exemples

On vient de proposer, avec les concepts systémiques *de processus dynamique complexe* et *d'émergence*, des *analogies fonctionnelles* potentielles entre les objets et les méthodes de la *linguistique* et de la *sociologie*. Nous allons présenter des exemples permettant d'illustrer certaines de ces analogies.

3. 1. - Une dynamique commune mot-individu : polysémie et schizophrénie

Pour H. von Foerster (1974), l'« identité personnelle », le « langage » ou la « communication » - concepts qu'il désigne sous le nom de « registres » - sont des représentations qui *s'engendrent mutuellement*. D'où notamment des liens génératifs entre *langage et représentation de soi*. Dans la mesure où l'identité personnelle de l'individu se construit en termes de références croisées (à l'aide de processus récursifs), les relations liant langage et connaissance seront aussi articulées sur cette identité. Dans ces conditions, qu'advient-il si l'expression autoréférente récursive qui définit l'identité de l'individu se révèle non stable ? Plusieurs cas peuvent en effet être envisagés en fonction du comportement d'une telle expression (X. Fargeas et E. Andreewsky, 1994) :

- Si la trajectoire *converge* vers un attracteur, l'individu va générer son identité (et donc aussi ses interprétations du monde) avec une certaine rigidité.

- Si cette trajectoire *oscille* autour de cet attracteur, le degré de liberté représenté par ces oscillations évitera les rigidités précédentes ; par ailleurs, la cohérence des interprétations de l'individu, en référence à l'unité de sa propre identité, sera assurée.

- Si la trajectoire oscille maintenant autour de deux attracteurs (ou plus), les extrapolations en fonction de l'expérience préalable seront difficiles, et les discontinuités de l'identité de l'individu se répercuteront sur ses interprétations, et donc sur son comportement. Ces discontinuités ont des conséquences théoriques permettant notamment de sous-tendre un certain nombre de caractéristiques cliniques des *patients schizophrènes*².

Sans entrer dans plus de détails sur les formalisations proposées, on remarquera que la théorie de la communication de von Foerster, base des formalismes précédents, est associée avec une *théorie du langage* en termes d'*émergence*. Contrairement à la majorité des théories de la communication (qualifiées par l'auteur de *technologies* de la communication), cette théorie ne présuppose pas l'existence de symboles primaires, autrement dit, d'éléments de communication - mots ou énoncés - pour lesquels la communication serait considérée comme acquise. Dans cette approche, la signification des mots est générée à la fois en termes de référence croisée avec celle des autres mots (et leurs références respectives), et des caractéristiques du contexte. Cette signification, par conséquent, a des dimensions autoréférente et récursive, assez analogues à ce que nous avons décrit plus haut, concernant l'identité personnelle. Les différentes trajectoires des systèmes dynamiques représentant cette identité sont alors susceptibles de modéliser également certains phénomènes du langage - comme notamment les différences de fonctionnement entre noms propres, mots « pleins » et polysémies³.

² Il s'agit des symptômes bleuleriens qui caractérisent la schizophrénie, de l'aspect développemental des dissociations des malades, et de tout ce que la théorie du « double bind » a tenté d'expliquer. Mais, à l'inverse de cette dernière théorie qui postule des oscillations exogènes (induites socialement par la famille du patient), pour expliquer les phénomènes pathologiques observés, avec l'hypothèse proposée, ces oscillations ont une origine endogène (les attracteurs liés aux systèmes dynamiques des « registres » individuels), et des dimensions aussi bien biologiques que culturelles ou sociales.

³Citons les trois cas :

1er cas : stabilité, convergence vers un attracteur. Ceci représente la trajectoire de la signification des noms propres, qui se fige dès qu'on connaît la référence de ce nom, et ne dépend plus d'autres facteurs, notamment du contexte.

Il n'est pas inintéressant de noter que les *discontinuités* de l'identité personnelle des patients *schizophrènes* semblent se répercuter spécifiquement, *en matière de langage*, dans le traitement des *polysémies* par ces patients, dans la mesure où ce traitement est pratiquement leur seule activité psycholinguistique qu'on a réussi à objectiver expérimentalement comme différente de la norme.

3. 2 - Complexités “récurives” : société et texte

Un des traits communs le plus important des *relations* qui lient *individu* et *société* d'une part, *mot* et *texte* d'autre part, est le rapport de *complexité* entre les termes de ces relations. Dans une construction classique où les éléments n'ont ni vie propre, ni autonomie⁴, ils restent simples par rapport à la complexité de l'ensemble qu'ils constituent. Dans de tels cas (et par exemple pour la plupart des artefacts modernes utilisés quotidiennement), l'ensemble est nécessairement composé de beaucoup d'éléments - et cette quantité crée sa complexité, ou plutôt sa complication (Le Moigne, 1990). A l'opposé, un texte, ou une société, peut exister avec quelques éléments seulement. Un simple titre, ou un slogan, présente complexité et richesse aussi bien qu'un article, une nouvelle ou un roman. Certes, il y a de grandes différences entre une phrase courte et une encyclopédie, mais dans les deux cas, un message est délivré, des interprétations sont suggérées - qui peut-être sont susceptibles de changer l'avenir de l'humanité ...

La complexité des sociétés et celle des textes, de faible taille, n'est pas surprenante. En effet, les “éléments” des textes - comme ceux des sociétés - sont loin d'être simples ou “élémentaires” ! Ils sont au contraire très complexes et leur combinaison n'est pas une simple concaténation. Les mots sont des *systèmes* possédant un grand nombre de propriétés qui, pour la plupart, reflètent celles des systèmes formés par les textes où ces mots figurent. En lisant un texte, le lecteur, dès le premier mot, crée une certaine interprétation. A chaque nouveau mot, celle-ci évolue, se transforme. A tout moment, si la lecture est arrêtée, le lecteur peut fournir ses idées sur ce qu'il a lu, ses analyses, faire un résumé... Il peut caractériser le texte. Les règles de composition gouvernant l'enchaînement des mots sont multiformes et le “message” du texte émerge de façon non-prédictible (si, souvent, l'ajout de plusieurs mots ou de plusieurs phrases ne fait pas beaucoup évoluer le sens d'un texte donné, un mot spécifique peut tout à coup faire basculer la signification de tout le texte).

2° cas : continuité, cohérence - oscillations autour d'un attracteur. Ces oscillations aménagent le degré de liberté nécessaire pour générer la signification des mots “pleins” usuels à chacune de leur occurrence, en fonction du contexte de cette occurrence. L'espace des significations potentielles est déterminé par l'amplitude des oscillations.

3° cas : discontinuités, oscillations autour de deux ou de plusieurs attracteurs. Ces discontinuités représentent celles des polysémies, c'est à dire les discontinuités (très courantes dans les langues naturelles) des significations d'un mot donné, à chacune de ses occurrences.

⁴ Comme les individus, les mots naissent, vivent et meurent. Ils ont des ancêtres (et même des ancêtres communs). Leur date de naissance n'est pas toujours connue avec précision et leur identité n'est reconnue qu'à l'âge adulte, même si leurs parents adoptifs (eux, vous ou moi) les utilisent dès leur plus jeune âge.

Plus sérieusement, le mot présente des caractéristiques autopoïétiques. Son existence suppose qu'il se distingue (pour les utilisateurs de la langue) des autres mots. Il présente une dynamique propre qui gravite autour d'un attracteur sémantique spécifique à l'époque, à la culture, au lieu, à la société. Par ailleurs, il interagit avec les autres mots. Il peut prendre leur place, tisser des relations privilégiées. Enfin, il disparaît lorsqu'il n'est plus utilisé, ce qui est souvent lié à ce qu'il ne se distingue plus aussi clairement lors de son usage (c.a.d. sa dynamique propre et donc son identité disparaît).

La manière dont les mots se combinent traduisent leur complexité, qui elle-même reflète celle des textes où ils figurent. On a affaire, en quelque sorte, à des *systèmes "fractals"* en ce sens que même en focalisant sur une partie ou un élément d'un texte donné, on retrouve une complexité proche de celle de ce texte. C'est là que se trouve la *différence majeure entre collection et société*. Il n'y a pas de rupture de la complexité entre tout et partie d'un texte, alors qu'il y en aurait une s'il ne s'agissait que d'une simple collection. On peut faire des remarques analogues sur les individus et les sociétés. Si ajouter ou soustraire un individu modifie la taille physique de la collection (de la *collectivité*), cela ne passionne personne. C'est le *sens*, c'est à dire la finalité, de cette "collectivité" qui est intéressant. Il peut aussi bien rester inchangé par l'ajout d'un individu, comme être bouleversé sur de nombreux aspects, bien au-delà des modifications physiques induites. C'est ce qui fait de la collectivité une *société*. Il s'agit d'une analogie forte entre les fonctionnements *mot-texte* et *individu-société*.

3.3. - Régulation systémique et décision individuelle - l'exemple des textes de jurisprudence.

Tout texte est un *système*. Il n'est pas une 'somme sémantique' de ses éléments lexicaux car il est soumis à une dynamique interne qui lui permet d'orienter le sens, voire de suggérer des *rétro-interprétations* (Bourcier & Bruxelles, 1995). Le texte peut s'auto-représenter et fournir des indications sur la façon dont il convient de le comprendre. Ces propriétés remarquables lui permettent de modifier l'environnement textuel où il est produit et d'interférer avec les conditions mêmes de sa production et de son interprétation.

Ceci est particulièrement clair quand il s'agit de textes juridiques. En effet, le droit n'est pas une collection de lois et de décisions individuelles, mais un système auto-référentiel qui a sa propre dynamique interne. La dynamique des rapports entre texte et interprétation invite à introduire le concept d'*émergence* : le droit serait déterminé par des phénomènes qui permettraient à des processus non définis *a priori* de se co-déterminer dans leur fonctionnement (rôle des différents acteurs juridiques dans l'élaboration du droit). L'opposition entre la conception *positiviste* du droit (le droit est une collection de textes et de règles) et celle du droit comme *émergeant* de décisions individuelles et d'interactions avec la société, a fait l'objet de débats récurrents en théorie du droit. Cette opposition est analogue à celle qui a été développée plus haut, en matière de langage.

Examinons plus particulièrement de quelle façon se forme la jurisprudence ; une définition classique de la jurisprudence fournit une première réponse : la jurisprudence n'est pas l'ensemble des décisions rendues par les juges mais « l'interprétation de la règle de droit, ou la solution suggérée par un ensemble suffisamment concordant de décisions rendues par les juridictions sur une question de droit » (Lexique des termes juridiques, Dalloz (1991)).

La jurisprudence n'est donc pas seulement une source de droit mais une « société de textes » d'où l'on peut "dégager des réseaux de sens et des contraintes logiques". Autrement dit, les textes évoluent par eux-mêmes et se complexifient au cours d'une série d'interactions, et en fonction de leurs lois propres d'organisation.

Pour Luhmann (1985), le système juridique est une auto-propagation circulaire : les décisions créent des normes et les normes, des décisions. Des actions dérivent de règles législatives mais aussi des interactions, par équilibrations successives et auto-construction de représentations. Autrement dit, coopèrent dans le système à la fois des règles issues de textes mais aussi des règles non écrites issues des pratiques. “L’esprit ne fabrique point tant de règles qu’il ne se compose de règles pour l’action, c’est-à-dire d’un complexe de règles qu’il n’a pas faites mais qui ont fini par gouverner l’action des individus parce que, lorsqu’elles s’appliquaient, leurs actions s’avéraient plus efficaces” (Hayek, 1985). Le sens des textes normatifs se construit à chaque interaction avec le décideur lors de l’exercice de son pouvoir d’interprétation⁵. Mais le juge interprète les faits et les normes de façon à rendre son activité de jugement plus *efficace* par rapport à son environnement et à l’ensemble des décisions précédentes

Ce processus d’adaptation continue produit ce qu’on appelle la jurisprudence. La jurisprudence n’est pas seulement une somme de cas, elle génère son propre système de normes, repérable à travers les régularités qui *émergent* des décisions individuelles.

4. - Conclusion

De nombreux arguments mettent en évidence des points communs entre les relations mot-texte et individu-société. Ces points communs expriment évidemment la similitude entre deux *systèmes de systèmes* extrêmement complexes, où chaque élément peut présenter une interdépendance constitutive avec tous les autres.

Les analogies que nous avons repérées entre ces deux systèmes (et systèmes de systèmes) sont peut-être dues au hasard et/ou à notre incapacité, en l’état actuel de nos connaissances, à mieux caractériser les différences entre systèmes complexes.

Toutefois, il reste possible que ces analogies ne soient pas vraiment dues au hasard. Comme l’indiquent Confucius et Schrödinger (cités en introduction), les conceptions de l’homme et du langage sont liées, dans toutes les cultures. Aucune société ne peut exister sans communications entre ses membres, donc sans langage. Et aucune langue n’a jamais préexisté à une société. Cette indissociabilité *constitutive* induit peut-être une communauté structurelle forte et profonde, ce que n’a fait que suggérer cet article.

⁵ Pour représenter ce phénomène d’émergence de régularités sociales (règles) à partir de décisions individuelles, nous avons utilisé l’approche connexionniste. Le modèle utilisé avait pour objectif de monter un protocole très limité sur une partie de la jurisprudence du Conseil d’état (Bochereau, Bourcier & Bourguine, 1991). Ce modèle se proposait de voir si l’on pouvait réduire la jurisprudence à une pratique d’expériences successives fondées sur des exemples de conflits résolus. Dans nos résultats, on a pu retrouver les règles positives applicables, qui étaient représentées dans les textes de lois par une collection de mots, mais aussi des *règles artificielles inconnues*, localisées dans la structure du réseau et le poids des connexions. Le contexte des lois est devenu l’ensemble du réseau.

Références

Andreewsky, E. (ed.) (1991). *Systémique et cognition*, Paris, Dunod.

Confucius (1987). *Entretiens familiaux*, XX. 3, trad. P. Ryckmans, Paris.

Bochereau, L., Bourcier, D. & Bourguine, P. (1991). Extracting legal knowledge by means of a multilayer neural network, *Proceedings of the 3rd ICAIL*, Oxford, ACM, p 288.

Bourcier, D. & Bruxelles, S. (1995). *Une approche sémantique de l'argumentation juridique, Dire et c'est-à-dire*, Vol. 45, N°1, 34-57.

Durkheim, E. (1988 - 1^{ère} éd. : 1894). *Les règles de la méthode sociologique*, Paris, Flammarion,

Fargeas, X. & Andreewsky, E., (1994). Schizophrenic Troubles of Personal Identity : A Cognitive Model. In Andreewsky, E. (ed.), *Problems of Systems Identity in the Natural, Human and Social Sciences, World Futures*, Vol 42, N° 1-2, 119-124.

Foerster, von H. (1974). Notes pour une épistémologie des objets vivants. In Morin, E. & Piattelli-Palmarini, M. (Eds.) : *L'Unité de l'Homme, 2. Le cerveau humain*, Paris, Seuil, 187-205.

Hayek, F. (1985). *Droit, législation et liberté, 1 : Règles et ordre*, Paris, Puf

Le Moigne, J.-L. (1990). *La modélisation des systèmes complexes*, Dunod, Paris.

Luhmann, A. (1985). *A Sociological Theory of Law*, Routledge.

Mead, G. H. (1934). *Mind, Self and Society*, Chicago, Univ. Chicago Press.

Schrödinger, E. (1990). *L'esprit et la matière*, Paris, Seuil.

Shanon, B. (1993). *The Representational and the Presentational*, London, Harvester Wheatsheaf.

Vullierme, J.-L. (1994). Alter Egos: Notes on the Basic Processes of Specular Identification. In Andreewsky, E. (ed.). *Problems of System Identity in the Natural, Human and Social Sciences, World Futures*, Vol. 42, N° 1-2, 125-131.

Vygotski, L. (1985). *Pensée et langage*, trad. F. Sève, 1985, Paris, Messidor.

Weber, M. (1922). *Essais sur la théorie de la science*, Paris, Agora, Presses-Pocket

Une analyse de conversation au service de la conception d'un système de production automatisé

James De Almeida
Groupe de Recherches sur les Communications
Université de Nancy 2, BP 33-97
54015 Nancy Cedex

Résumé

Cet article présente une recherche menée en interdisciplinarité entre psychologues, éthologues et automaticiens. Cette recherche vise la mise en place d'un système multi-agents qui fonctionnerait d'une manière similaire à celle d'un groupe d'humains qui résout un problème d'agencement en temps réel en usant du langage ; c'est-à-dire dans le cours même de son activité, sans planification et sans recherche systématique d'optimalité. Ainsi, à partir d'un modèle développé en éthologie, avec ses notions d'émergence et d'auto-organisation et d'une technique d'analyse des interactions langagières développée au Groupe de Recherche sur les Communications, nous comptons systématiser la dynamique socio-cognitive de la résolution interlocutoire d'un problème d'usinage par un multi-robots fonctionnant en parallèle. Pour ce faire, nous avons mis en place une méthodologie ascendante alliant un jeu de rôle, où des sujets humains ont pris la place d'agents robots, et une analyse de conversation de l'enregistrement (audio-vidéo) de l'interaction.

mots clés : analyse interlocutoire, système multi-agents, résolution de problème, cognition distribuée, auto-organisation .

1. Théorie des actes de langage et formalisation des messages dans les systèmes artificiels

1.1. Introduction

Si la recherche en Intelligence Artificielle (IA) classique s'est longtemps consacrée à la modélisation du cognitif, la recherche en Intelligence Artificielle Distribuée (IAD) marque une tentative d'intégrer du *social* à l'élaboration cognitive. Cette tentative en IAD est bien sûr suscitée par le projet distributionniste. Distribuer l'intelligence revient à envisager un Système Multi-Agents (SMA) où les agents sont spécialisés dans le traitement d'un certain type de connaissances. Ce sont des agents autonomes qui doivent néanmoins coopérer, interagir, communiquer pour mettre en commun leurs productions individuelles. Cette mise en commun est fondamentale. Tout d'abord, elle explique que l'on qualifie ce champ de recherche, d' "Intelligence Artificielle Distribuée". Pour nous, en nous inspirant de Cicourel (1994), "Intelligence Distribuée" signifie que les cognitions spécifiques accomplies conjointement par les différents agents d'un système, dans le cadre de la résolution collective d'un problème, appartiennent à différents registres (logique, mathématique, sens commun, sémantique...).

1.2. Reconnaissance d'intention et constructibilité des actes de langages

Dans le paradigme classique en IAD, le pouvoir est décentralisé et le fonctionnement du système émane de décisions locales prises par les différents agents du SMA. Le type de

communication proposé ne reflète guère la dynamique des interactions humaines en langage naturel. Il n'y pas véritablement communication inter-agents, mais plutôt transmission d'informations d'un agent à l'autre. Plus précisément la communication entre les agents repose sur la reconnaissance, par l'agent récepteur d'un message et de l'intention de l'émetteur à partir du message reçu. Cette reconnaissance permet à l'agent récepteur d'activer un protocole ou un plan local lui permettant de réagir au message et d'envoyer à un (ou plusieurs) agent(s) spécifique(s), un nouveau message. La reconnaissance d'intentions est donc essentielle. Elle suppose que les messages soient suffisamment formalisés pour permettre ce type de reconnaissance. Cette formalisation vise à ce que les messages émis soient suffisamment structurés, explicites pour qu'ils soient envoyés au "bon" destinataire et correctement interprétés par ce dernier. Un message est en quelque sorte assimilable à la formalisation d'un acte de langage, dans le sens où le contenu propositionnel est représenté ainsi que certains composants de la force illocutoire de l'acte comme le degré de force et de puissance - rendant compte de l'importance du message - et la direction d'ajustement - spécifiant le destinataire du message. Toutefois, nous constatons que la force est une entité dissociée du contenu propositionnel de l'acte. Ce qui signifie que force et contenu propositionnel sont indépendants et ainsi, n'exercent aucune influence l'un sur l'autre. Ce constat révèle une mauvaise appropriation de la théorie des actes de langage (Searle et Vanderveken, 1985 ; Vanderveken, 1988). dans celle-ci, l'acte de langage est un contenu propositionnel sur lequel s'exerce une force illocutoire (Vanderveken, 1990). et non pas - comme semble le penser les chercheurs en IAD - un contenu propositionnel auquel viendrait se surajouter une force illocutoire. D'ailleurs un composant de la force illocutoire à savoir les conditions de contenu propositionnel traduit l'interdépendance entre contenu propositionnel (P) et force illocutoire (F) : par définition les conditions de contenu propositionnel sont des contraintes exercée par la force sur le contenu propositionnel de l'acte. En fait, lors de l'accomplissement d'un acte de langage *in situ*, la force et le contenu propositionnel apparaissent dans une *Gestalt* non réductible à leur somme.

Nous assistons aussi à une centration sur l'acte de langage lui-même (le message), ce qui traduit une inscription à la version monologique de la théorie des actes de langage. Pour eux la communication consiste à reconnaître dans le message l'intention de l'émetteur de façon à la satisfaire. Cette reconnaissance suppose l'existence d'un savoir mutuel (connaissances communes prédéfinies).

Par exemple, chez Cohen et Levesque (1990), ce sont des plans locaux qui sont mutuels. C'est sur cette base que les intentions sont reconnues. Quoi qu'il en soit, tous les auteurs s'engagent à écrire que la communication repose sur la reconnaissance, par le récepteur, de l'intention de l'émetteur (Brassac, De Almeida, Grégori, Saint-Dizier, 1996).

Pourtant et depuis plusieurs années, les conversations naturelles, analysées sous un angle dialogique, révèlent que l'intention d'un locuteur n'est pas donnée via la profération d'un énoncé en contexte : l'intention se négocie, se co-construit au gré de l'interaction.

Prenons un exemple :

t1 L1	A1: Où est mon journal ?	(L = locuteur)
t2 L2	A2: Bon je vais te le chercher	(A = acte de langage)
t3 L1	A3: Non, je veux juste savoir où il est !	

A2 montre que L2 a interprété A1 comme un ordre et non comme une requête, puisque L2 en t2 s'engage à aller chercher le journal de L1. A3 révèle que l'intention de L1 en t1 était que L2 satisfasse la requête littéralement exprimée au moyen de A1 et non l'ordre qui en dérive.

Si l'intention se négocie, c'est parce que nous, humains, ne disposons pas d'un savoir mutuel prédéterminé qui ferait que l'on se comprendrait parfaitement (comme les SMA). Certes au préalable d'une interaction, nous avons des suppositions sur un univers de référence commun ; mais ce ne sont que des suppositions, qui toutefois deviendront des croyances partagées si elles sont négociées dans l'interlocution. Dans l'exemple, trois tours de parole ont permis à L1 et L2 de fixer

de l'intercompréhension à propos de l'intention de L1 en t1. Nous voyons ici l'intérêt du principe de séquentialité qui gouverne les conversations. Si nous parlons en alternance, c'est pour créer de l'inter-compréhension et ainsi un espace d'intersubjectivité.

Nous soutenons que l'élaboration d'un espace intersubjectif est nécessaire à toutes activités véritablement sociales. Mais la création d'un espace intersubjectif suppose une volonté d'intercompréhension des agents, or en IAD les processus d'intercompréhension s'avèrent inutiles puisque les agents disposent d'un savoir mutuel et sont d'emblée sincères et coopératifs ; ce qui signifie qu'ils n'ont pas besoin de négocier pour se comprendre parfaitement ; c'est pourquoi d'ailleurs dans les SMA, les agents qui envoient des messages à d'autres agents, n'ont jamais de retour sur la façon dont les agents récepteurs ont perçu et traité leur message.

Pour notre part, nous pensons que si il n'y a pas construction collective d'intercompréhension et donc pas élaboration d'intersubjectivité, alors il ne peut y avoir véritablement coconstruction interactive d'une solution.

1.3. Les SMA et la gestion des relations sociales

Si un des projets en IAD est de concevoir des SMA simulant la dynamique des interactions humaines dans le cadre de résolution collective de problème, alors nous pensons qu'il est nécessaire de renoncer au savoir mutuel et donc à la reconnaissance d'intentions, pour s'intéresser aux processus sous-jacents à la dynamique conversationnelle. Ces processus relèvent du cognitif et du social. Nous soutenons que le social via la gestion relationnelle occupe une part active et importante dans l'élaboration conjointe des cognitions. La gestion relationnelle peut s'envisager à la fois comme une ressource servant à formater les interactions et comme le résultat d'un travail interactionnel qui permet à des cognitions spécifiques de s'accomplir (Perret-Clermont, Schubauer Léoni, Grossen, 1991 ; Trognon, 1994). En IAD, l'intégration du social aux SMA s'est réduite à doter les agents, de rôles n'évoluant pas en cours d'interaction. Nous pensons que la gestion d'un rapport de places (Goffman, 1974), permettant d'acquérir ou de perdre du pouvoir, est un moteur dans la dynamique des interactions humaines. Précisons que dans cette perspective les agents doivent pouvoir ne pas être toujours sincères.

Pour conclure, nous pensons que pour parvenir à des SMA où les agents, véritablement, élaborent conjointement une solution à un problème, les chercheurs en IAD doivent se résoudre à acquérir une vision dialogique de la conversation.

Pour parvenir à simuler la dynamique des enchaînements conversationnels, la première étape sera de comprendre cette dynamique. Un modèle descriptif et explicatif (la logique interlocutoire) à même d'élucider cette dynamique conversationnelle est actuellement en cours de développement à Nancy 2, et devrait prochainement offrir tous les critères à une implémentation (Trognon et Brassac, 1992 ; Brassac et Trognon, 1994 ; Brassac, 1992 ; Trognon et Ghiglione, 1993). Mais pour l'heure, le projet consiste uniquement en une exploitation de cette logique comme moyen pour mettre en évidence les mécanismes intrinsèques de la résolution du problème SHIVA.

2) Le projet SHIVA : vers une réalisation interdisciplinaire

2.1. Intelligence artificielle distribuée et interactions sociales

La démarche habituellement utilisée en intelligence artificielle classique, dans le cadre de la conception de systèmes informatiques interactifs, est descendante. Elle consiste à postuler sur la modélisation d'une cognition censée produire l'activité à simuler. Néanmoins, comme nous l'avons illustré précédemment, les chercheurs intéressés par la communication et plus particulièrement pour

certains par la théorie des actes de langage (Cohen et Levesque, 1990) et parfois plus précisément pour son aspect pragmatique (Perrault, 1990) semblent vouloir s'ouvrir à l'interactionnisme. C'est dans cette ouverture que s'inscrit notre travail.

2.2. Cognition distribuée et interactions sociales

Nous soutenons une perspective constructiviste (Varéla, 1989 ; Stewart, 1993 ; Trognon, 1992, 1994 ; Lenay, 1994) selon laquelle l'activité humaine n'est pas réductible au résultat d'opérations formelles sur des représentations mentales d'un monde existant objectivement, mais résulte d'un travail co-construit par les interactants dans un cadre social qui oriente leurs productions et en émerge (Beauvois, Monteil et Trognon, 1991).

Nous n'envisageons l'activité à simuler que dans le *hic et nunc* de son accomplissement.

A partir de ces deux revendications, nous proposons le passage d'une conception centrée sur l'agent à une conception socialement fondée (Winograd et Florès, 1989). Ainsi, plutôt que de postuler sur la modélisation de la cognition de l'agent ou des agents informatique(s), nous proposons un travail centré sur l'accomplissement de l'activité à simuler - c'est-à-dire au niveau de ce qui s'élabore dans l'interlocution. La démarche de conception proposée est donc ascendante : c'est à partir des analyses inductive et interlocutoire d'un corpus conversationnel où l'activité à simuler s'est produite que l'organisation fonctionnelle du système doit être définie. L'avantage de cette démarche est qu'elle permet de systématiser l'accomplissement simultané de processus socio-cognitifs variés qu'une approche centrée sur la modélisation de la cognition individuelle n'est pas à même d'envisager.

2.3. Analyse d'une interaction socio-cognitivement jouée

L'application (SHIVA, une machine-outil à architecture parallèle) sur laquelle nous travaillons (De Almeida, 1995) avec des automaticiens (Patriitti, Garro et Charpentier, 95) et des éthologues (Desor, Toniolo, Krafft et Dickes, 92) devrait permettre à partir de ce type d'analyses, de construire un système informatique gérant l'interaction d'une façon similaire à celle d'un groupe d'humains. La méthodologie que nous proposons, à l'instar de la recherche en éthologie, consiste en partant du système en fonctionnement, à décrire par son analyse, comment les propriétés de coordination émergent en temps réel.

Deux méthodes sont envisagées : la première consiste à étudier la progression des différents paramètres à l'aide d'analyses multi-variées ; la seconde, à rendre compte des différents problèmes relatifs à la coordination des agents en décrivant comment ils sont négociés, comment les décisions sont prises, contredites, rectifiées... bref comment représentations et actions s'articulent dans une élaboration collective de cognitions.

Pour décrire et analyser un tel système - qui n'existe pas dans la nature - avant de le concevoir, nous avons mis en place un protocole expérimental¹. Ce dernier consiste, à travers une situation quasi naturelle, proche du "jeu de rôle", à considérer le SMA comme un système groupal de coordination d'actions, *comme* un système *intentionnel* (Dennett, 1990) : dans une telle situation, les agents (pièces et outils) étant remplacés par de véritables systèmes cognitifs (sujets humains) les processus de décision, de négociation, d'explicitation ..., peuvent être étudiés et formalisés en terme d'heuristiques grâce à la logique interlocutoire.

¹ en collaboration avec le Centre de Recherche en Automatique de Nancy (CRAN - CNRS/URA 821-Université de Nancy 1) et le Laboratoire de biologie du comportement et de la physiologie (LBCP-CNRS/URA 1293 - Université de Nancy 1)

2.3.1. *Coordination discursive vs coordination actionnelle*

Il faut distinguer la coordination discursive qui permet aux sujets de résoudre le problème, de la coordination actionnelle nécessaire à l'accomplissement physique de l'activité de SHIVA. En effet, nombre d'éléments participant de la coordination discursive ne sont pas nécessaires au fonctionnement du système. La coordination discursive n'est que le moyen permettant de faire émerger la coordination actionnelle.

Dés lors que l'organisation fonctionnelle a émergé (et s'est stabilisée) et qu'elle a été extraite de l'interaction grâce à la logique interlocutoire, la dynamique ayant permis son émergence devient inutile pour l'implémentation du SMA.

2.3.2. *Coordination cognitive vs coordination réactive*

En ce qui concerne la gestion des intentions en vue d'une coordination plus efficace, il est évident que nos sujets en ont usée et qu'elle prend une place prépondérante dans notre analyse. Toutefois, au regard de la solution qu'il ont mis en place et de sa dynamique, nous ne pouvons affirmer qu'elle sera nécessaire à l'implémentation. En nous appuyant sur les propos de Ferber (1994), de Lenay (1994) et de Lestel, Grison et Drogoul (1994), nous pouvons nous attendre à ce qu'une implémentation sur un mode réactif soit à même de rester proche de la solution mise en place et semblable au niveau des performances.

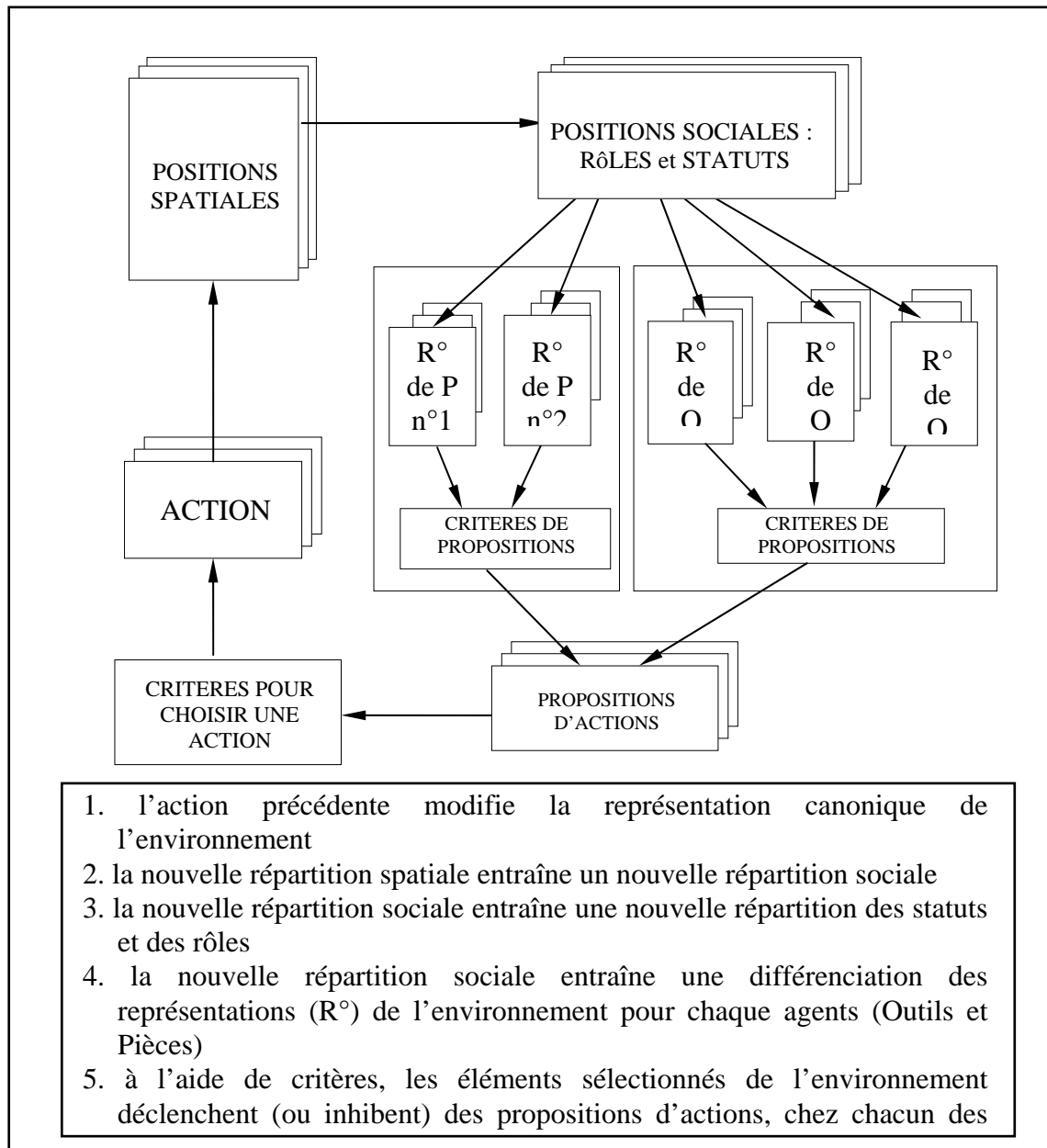
2.4. **Quelques résultats**

L'analyse du corpus n'étant pas à ce jour terminée, nous ne pouvons fournir que quelques résultats partiels, mais qui, et malgré les futurs ajustements, sont pour la plupart tenus pour acquis.

En premier lieu, il convient de signaler que la solution co-construite par les interactants est une solution qui repose sur une procédure d'ajustement à la situation locale (spatial, temporel et relationnel). Elle se réalise pas à pas et ne fait appel à des investigations prédictives que pour des évaluations à court terme ; ces tentatives de planification de l'action étaient limitées par l'expérimentation (pas moyen d'écrire), elles n'ont pas jouées le rôle de programme d'actions (par abandon, oubli, erreur...) au cours de l'interaction, puisqu'elles ont été préférentiellement remplacées par des heuristiques faibles en coût cognitif et aisées à communiquer aux partenaires.

Ensuite nous avons mis en évidence qu'une sociogenèse avait réparti les fonctionnalités sur les différents agents, réduisant de façon importante les possibilités d'actions à chaque instant. Mais surtout nous avons remarqué que cette répartition était fluctuante. En effet, les variations d'appariements (agents ; fonction) sont directement déterminées par les positions relatives de chacun des agents dans l'environnement SHIVA.

Aussi, nous avons tenté à partir de ces premiers résultats de schématiser l'interaction à travers un modèle du processus de résolution construit comme une boucle de perception-action.



-Figure 1-
 SOLUTION SOCIO-COGNITIVEMENT
 CO-CONSTRUITE PAR LES INTERACTANTS
 DU JEU DE RÔLES

3. Discussion

La méthodologie ascendante, en ce qu'elle concilie l'utilisation d'un véritable groupe social à travers le jeu de rôles et l'extraction du processus de résolution grâce à l'analyse d'une interaction conversationnelle, nous permet à la fois d'éviter les écueils d'une implémentation hasardeuse d'agents communicants et de ne pas passer à côté des phénomènes essentiels que supporte une authentique relation sociale, si rationnelle et contrôlée qu'elle se veuille.

Plutôt que de chercher à systématiser une interaction langagière artificielle, nous avons donc choisi, à partir de l'analyse d'une interaction réelle, de n'implémenter que le résultat (l'organisation fonctionnelle telle qu'elle s'est stabilisée) que cette dernière a permis de faire émerger.

Cette recherche n'ayant pas encore atteint le stade de l'implémentation, nous retiendrons qu'elle offre une perspective de conception innovante et plus à même de restituer les multiples phénomènes intriqués de l'interaction langagière. Néanmoins le pari que nous faisons concernant l'utilisation et la validité pratique en IAD de la logique interlocutoire semble être lancé sur une voie riche de perspectives que nous comptons rapidement exploiter.

Quant à une implémentation intégrale de la logique interlocutoire, pour des agents conversationnels véritablement interactifs, l'avenir reste prometteur mais nécessitera de nombreux travaux et investigations d'autant plus fastidieux qu'il nous faudra en passer systématiquement par des techniques de simulation et d'analyses.

4. Conclusion

Définir l'organisation fonctionnelle d'un SMA, dans une démarche de conception ascendante, pour un système multi-robots autonome voué à l'industrie de production automatisée, à l'aide d'une technique d'analyse des conversations utilisée en psychologie sociale, à l'aide d'un modèle pragmatique de la gestion des interactions socio-cognitives et à l'aide d'une méthodologie empruntant au travail de terrain des éthologues, et surtout en n'omettant pas de respecter les contraintes de l'automatisme et de la production industrielle, reste un objectif ambitieux. En effet de nombreuses embûches sont sur notre chemin : trouver un langage partagé correspondant à des concepts communs, s'approprier les perspectives de l'autre sans (trop) les dénaturer... Ces problèmes liés à l'interdisciplinarité - qui selon nous ne sont pas véritablement des problèmes puisqu'en les dépassant chaque jour, nous progressons dans notre projet - sont du même ordre que ceux que nous envisageons intégrer au fonctionnement de notre système. Ainsi, c'est dans une sorte de va-et-vient entre la conception (action) et la conception (objet), que s'élabore collectivement et donc interactivement le déroulement et le produit de notre collaboration.

Références

Beauvois, J.-L., Monteil, J.-M., Trognon, A. (1991). Quelles conduites, Quelles cognition ? Repères conceptuels. In J.-L. Beauvois, R.V. Joule et J.-M. Monteil (eds) *Perspectives cognitives et conduites sociales*, 3. Fribourg : Delval. 203-287.

Brassac, C.(1992). S'engager en conversation. Thèse de doctorat en psychologie, Université de Nancy 2.

Brassac, C., De Almeida, J., Grégori, N., Saint-Dizier, V. (1996). "La théorie des actes de langage en Intelligence Artificielle Distribuée : utilisations et limites". *Actes des quatrièmes journées francophones sur l'Intelligence Artificielle Distribuée et les Systèmes Multi-Agents (JFIADSMA'96)*, Port-Camargue, 1-2-3 avril 1996. (communication acceptée).

- Brassac, C., Trognon, A. (1994). Speech Act Theory and Cognitive Psychology. In D. Vanderveken & S. Kubo (eds) *Essays in Speech Act Theory*. Amsterdam/Philadelphia : J. Benjamins Publishing Company. pp.
- Cicourel, A.V. (1994). La connaissance distribuée dans le diagnostic médical. *Sociologie du travail*, 4, 427-449.
- Cohen, P.R., Levesque, H.J. (1990). Persistence, intention and commitment. In: P. Cohen, J. Morgan and M. Pollack, eds., *Intentions in communication*. Cambridge: MIT Press, 33-69.
- De Almeida, J. (1995). Emergence de l'organisation fonctionnelle d'un système multi-agents. *Actes du Congrès annuel de la Société Française de Psychologie*; 11-13 mai 1995 Toulouse.
- Desor, D., Toniolo, A.M., Krafft, B., Dickes, P. (1992). Cognition sociale chez l'animal à propos du comportement d'incitation chez le rat. *Actes du cinquième colloque de l'ARC, percevoir, raisonner, agir, articulation de modèles cognitifs* ; 24-26 mars 1992 nancy, 233-248.
- Dennett, D. (1990). *La stratégie de l'interprète : le sens commun et l'univers quotidien*. Gallimard.
- Ferber, J. (1994). Coopération réactive et émergence. *Intellectica*, 19, 19-52.
- Goffman, E. (1974). *Les rites d'interaction*. Paris : Editions de Minuit.
- Lenay, C. (1994). Organisation émergente dans les populations : biologie, éthologie, systèmes artificiels. *Intellectica*, 19, 9-17.
- Lestel, D., Grison, B., Drogoul, A. (1994). Les agents réactifs et le vivants dans une perspective d'évolution coopérative. *Intellectica*, 19, 73-90.
- Patriti, V., Garro, O., Charpentier, P. (1995). Application des SMA à la fabrication : cas de SHIVA : machine à architecture parallèle. *Actes des troisièmes journées francophones sur l'IAD et les SMA* ; 15-16-17 mars 1995 Baldoph - Savoie - France, 67-78.
- Perrault, D. (1990). An application of default logic to speech act theory. In: P. Cohen, J. Morgan and M. Pollack, eds., *Intentions in communication*. Cambridge: MIT Press, 161-185.
- Perret-Clermont, A.N., Schubauer-Leoni, M.L., Grossen, M. (1991). Interactions sociales dans le développement cognitif : nouvelles directions de recherche. *Cahiers de Psychologie*. Université de Neuchâtel.
- Searle, J.R., Vanderveken, D. (1985). *Foundations of illocutionary logic*.
- Stewart, J. (1993). Au delà de l'inné et de l'acquis. *Intellectica*, 16, 151-174.
- Suchman, L.A. (1987). *Plans and Situated Actions-The problem of Human/Machine Communication*. Cambridge : Cambridge University Press.
- Trognon, A., Brassac, C. (1992). L'enchaînement conversationnel, *Cahiers de linguistique française*, 13, 67-108.
- Trognon, A., Ghiglione, R. (1993). *Où va la pragmatique ?* Grenoble : PUG.
- Trognon, A. (1994). Théories et modèles de la construction interactive. In A. Trognon, U. Dausendschön-Gay, U. Krafft, C. Riboni *La construction interactive du quotidien*. PUNancy. 7-52.
- Vanderveken, D. (1988). *Les actes de discours*. Bruxelles : Pierre Mardaga.
- Vanderveken, D. (1990). On the unification of speech acts theory and formal semantics. In P. Cohen, J. Morgan & M. Pollack eds, *Intentions in communication*. Cambridge: MIT Press, 195-120.
- Varela, F. (1989). *Autonomie et connaissance*. Paris : Seuil.
- Winograd, T., Flores, F. (1989). *L'intelligence artificielle en question*. Paris : PUF.

De l'apport potentiel de la sociologie pour l'ingénierie des systèmes sociaux artificiels

Nils FERRAND

Laboratoire LEIBNIZ - IMAG
46 Av. Félix Viallet - 38031 GRENOBLE CEDEX
e-mail : nils.ferrand@imag.fr

Résumé

Dans cet article nous nous efforçons de montrer que la sociologie peut et doit nous aider dans l'ingénierie des systèmes sociaux artificiels. Dans une première partie nous détaillons nos motivations en illustrant que les travaux sur les systèmes collectifs n'échappent pas à des références sociologiques implicites, et qu'il est donc utile d'en expliciter les concepts. Nous avançons aussi que la sociologie est une discipline déjà ancienne et constituée depuis son origine comme une science de la complexité, et qu'elle peut donc inspirer les travaux actuels relatifs aux systèmes collectifs.

La plus grande part de l'article consiste en un ensemble de rappel de notions sociologiques relatives à l'action sociale. Ces rappels sont assortis de commentaires portant sur la validité et l'utilité de ces concepts pour l'ingénierie des systèmes sociaux artificiels. Enfin nous posons la question de l'intérêt de prendre les sociétés humaines comme modèle pour l'ingénieur. Nous montrons que selon les situations (simulation, résolution de problèmes, agents autonomes en environnement ouvert) la place des sciences sociales n'est pas identique. Et nous nous interrogeons sur la réductibilité des sociétés artificielles à un cadre autre que celui des sociétés humaines.

mots clés : sociologie, systèmes multi-agents, systèmes sociaux artificiels, analogie structurelle

1. Introduction

Le thème des journées de Rochebrune 1996 : “Du collectif au social”, ainsi que les questions afférentes relatives à la complexité des individus dans les organisations collectives, ne pouvait que résonner avec des préoccupations antérieures relatives à la construction de systèmes multi-agents à des fins de résolution de problème ou de simulation (Demazeau, 1993 ; Ferrand, 1995). En effet, nous avons illustré que de nombreux problèmes réputés complexes (problèmes multi-critères spatialisés avec contraintes locales) pouvaient être résolus au moyen de systèmes constitués d'agents très simples (“réactifs”). Cependant que, à l'autre extrême, certains collègues travaillaient dans la perspective de l'intégration d'agents symboliques de haut niveau, soit par les interactions (Demazeau, 1995), soit par les organisations (Baeijs & Demazeau, 1995). Or dans ce dernier cas, les références sociologiques sont apparues incontournables, avec la perspective de poursuivre une analyse conceptuelle initiée dans (PLEIAD, 1992).

Ainsi, prenant pour acquis que les systèmes réactifs (individus simples) ont des potentialités collectives très importantes, l'objectif de cette contribution est de caractériser et de discuter l'intérêt pour l'ingénieur de systèmes fondés sur un modèle social humain. A cette fin, nous avons donc pu concrétiser cette réflexion sur l'apport potentiel de la sociologie à l'ingénierie des systèmes collectifs artificiels.

Cet article a une vocation essentiellement incitatrice : il cherche à promouvoir une lecture des sciences sociales, qui puisse permettre des discussions fructueuses entre les différentes communautés. Car la progression des ingénieurs dans la complexité des structures et systèmes symboliques les conduit à rejoindre peu à peu les sociologues dans leur champ d'incertitudes.

2. Pour une lecture des sciences sociales

La motivation qui supporte actuellement notre intérêt pour les processus collectifs (vocabulaire générique choisi ici temporairement) est d'ordre pragmatique et constructive. En effet on peut considérer que la plupart des phénomènes observés dans notre environnement résultent, selon un terme désormais consensuel, de "l'émergence" d'organisations, de formes structurées, ou même d'intelligence, à partir des dynamiques enchevêtrées des objets ou entités impliqués (Casti, 1994 ; Ruelle, 1991). De plus, les capacités computationnelles symboliques ou non, la robustesse, et plus généralement l'autonomie de ces systèmes naturels "collectifs", en font des modèles aptes à susciter l'attention d'observateurs et gestionnaires de systèmes complexes. Dès lors, étant admis que le "collectif" est un objet digne d'étude, et qu'il offre à l'ingénieur des perspectives quant à la résolution de ses problèmes, on se doit de l'aborder aussi largement et méthodiquement que possible. Or, dans les communautés de chercheurs - ingénieurs intéressés par cette ouverture, si le concept de "collectif" a été déjà traité par de multiples machineries disciplinaires (éthologie, informatique, biologie, physique, etc.) (Ref collectif), il n'a paradoxalement pas été lu transversalement et systématiquement par la discipline fondatrice du concept, à savoir la sociologie. Si on se réfère à la très importante littérature relative aux systèmes multi-agents dits "cognitifs" (Demazeau & Muller, 1991 ; Ferber, 1995), à leurs ancêtres, ou apparentés (cf. la vie artificielle (Heudin, 1994)), c'est à dire à la plupart des systèmes collectifs artificiels, les analogies sociales ou psycho-sociales constituent un outillage permanent, qui articule un jargon structuré (PLEIAD, 1992) appelant inévitablement la construction de représentations mentales "socialisant" les processus collectifs considérés. Or si cette idéalisation facilite la compréhension, elle entretient d'une part largement la méfiance de communautés scientifiques "dures" envers ces travaux (perçus...) forcément "ramollis", puisque assimilables au social, et d'autre part elle inféode toute la pratique interne de l'ingénierie des systèmes collectifs à un ensemble d'idées reçues, d'images introspectives, ou de projections sociales, rarement explicitées.

Afin de résorber cette "pollution" rhétorique, on peut soit la rejeter, soit la digérer. La rejeter, c'est décider d'écarter systématiquement du vocabulaire tout concept susceptible d'une interprétation ou d'une résonance dans le champ du social, ou bien de reconstruire intégralement un vocabulaire acceptant éventuellement ces concepts polysémiques, mais toujours avec une définition référente au lexique. Cette dernière solution est cependant bien risquée, car la tentation analogique est pernicieuse, et les précautions méthodologiques imposées par un tel nettoyage sémantique et conceptuel paraissent devoir entraîner une refondation *ab-nihilo* des travaux (au moins des écrits) relatifs au "collectif" (par quels termes et acceptions remplacer ce titre lui-même empreint de multiples images culturelles ?). Quant à la solution de l'exclusion stricte des concepts ambigus (*collectif, organisation, relation, social, coopération, etc.*), elle réduit drastiquement les possibilités d'exploration des systèmes complexes considérés, en interdisant l'accès par l'outillage analogique, alors qu'il s'agit là d'un puissant "macroscope".

Il reste donc l'option mesurée, la "digestion", dont les ressorts sont de deux ordres : il faut d'une part chercher à préserver les énoncés de modèles des perturbations subjectives issues de la migration conceptuelle, et d'autre part aborder explicitement la référence sociale en réarticulant les propositions et modèles utilisées comme des transpositions affirmées du savoir sur le social. Ces perspectives supposent respectivement de savoir identifier et organiser les concepts "du social", de même que leur représentation par les scientifiques "passeurs de frontières", et de tout simplement accepter comme référence structurante de ces concepts les disciplines qui les ont fondées et diffusées dans le sens commun : sociologie, psycho-sociologie, ethnographie, sciences politiques, et partiellement économie.

Dans les communautés de l'informatique (dont l'Intelligence Artificielle, la Vie Artificielle, etc.) ou celles des mathématiques ou de la physique, l'image des sciences sociales est en général très dégradée. Par conséquent cela paraît être une gageure que de demander d'accepter ces sciences réputées infondées, comme des références cohérentes pouvant structurer un raisonnement et une

construction. Il ne nous appartient pas ici de reproduire, affaiblies, les controverses épistémologiques relatives à la qualité respective des sciences. Cependant, en préambule, nous souhaitons faire deux remarques : premièrement, les sciences sociales sont beaucoup plus anciennes que l'informatique en particulier, et pendant des siècles elles ont coévolué avec les sciences "dures" dans l'esprit de chercheurs encyclopédiques ouverts aux *mystères* en général, et souvent acteurs engagés de leur temps (comme Auguste Comte, Wilfredo Pareto, ou plus encore Karl Marx). La prudence s'impose donc quand on considère avec légèreté des travaux qui ont tant participé à notre histoire et ont ainsi défini la matrice sociale des découvertes scientifiques. L'esprit des Lumières, les révolutions industrielles, les grands cadres idéologiques du XX^{ème} siècle, ont fait naître des révolutions techniques, sont issus d'autres révolutions, et ils sont indissociables des courants dans les sciences sociales de leur époque. L'existence même des sciences sociales dès la fin du XVIII^{ème} siècle a certainement marqué le monde, et en particulier conduit aux premières tentatives de lectures rationalisées des systèmes sociaux (Comte, puis Durkheim). Ainsi, la jeune "science du collectif" qui nous intéresse ici se doit de considérer ces ancêtres et ascendants comme condition de sa réalisation.

Notre deuxième remarque est plus fondamentale. Penser le collectif est un des avatars multiples des sciences de la complexité, telles qu'elles ont été instituées depuis une vingtaine d'années. Et la science moderne se retrouve largement démunie d'outils aussi rationalisés et étalonnés qu'elle le souhaiterait pour aborder la complexité des phénomènes réels, qu'ils se situent dans la sphère du physique, du biologique, ou du cognitif. Les approches formalisées, axiomatisées, et empiriquement vérifiées sont minoritaires, et ne produisent généralement de résultats que sur une fraction du réel, ou sur des modèles idéaux largement réducteurs. En particulier les travaux en physique statistique, ou en physique "non-linéaire" et son pendant mathématique, qui constituent un réservoir théorique incontestable, s'ils abordent effectivement l'analyse de systèmes complexes, n'en apportent pas pour autant de réponse immédiate aux interrogations pragmatiques et ingénieristes que posent la compréhension et la régulation (gestion) de systèmes "cibles" réels, tels la cognition humaine et son support biologique, les écosystèmes anthropisés, les grands systèmes d'information ouverts, ou encore les organisations sociales et économiques. Par ailleurs ces systèmes ne sont pas accessibles à l'expérimentation, que ce soit de par leur taille, leur intangibilité "éthique", ou leur résistance structurelle à la perturbation inquisitrice. Tout au plus peuvent-ils être observés. Instituant le complexe comme objet d'étude non réductible, on se trouve ainsi rapidement dépourvu de perspectives méthodologiques scientifiquement fondées (c'est à dire testables, réfutables, au sens de Karl Popper). La recherche oscille alors entre deux pratiques : la confrontation dialectique (fondation discursive d'une rationalité intersubjective, construite incrémentalement par modification et enrichissement de concepts et validation consensuelle) et la modélisation / simulation, rendue opératoire par l'utilisation des ordinateurs. Mais dans l'un et l'autre cas, les sciences du complexe, dès lors qu'elles abordent le réel, perdent leur légitimité rationaliste, et subissent une dévaluation théorique au regard de leurs disciplines parentes, réputées "dures". Parvenu à ce point de l'analyse, il apparaît que ces recherches, issues essentiellement de sciences dures, se retrouvent dans la situation et avec les mêmes difficultés que les sciences sociales. Ceci pour la simple raison que, ayant historiquement fait le détour empirique par des objets réductibles et manipulables, elles ont été contraintes, sans doute après la théorie de la relativité, la mécanique quantique, et la physique non-linéaire, d'accepter l'imprévu, la subjectivité, l'imprédictibilité, ou des contingences insoupçonnables, comme parties intégrantes des phénomènes. Or, de leur côté, les sciences sociales n'ont jamais pu se départir de ces contraintes, et ont donc évolué en permanence avec un objet complexe. On peut donc considérer qu'il y a comme une leçon de leur part, qui énoncerait : "Vous avez rationalisé ? Eh bien discourez maintenant !". Et il paraît raisonnable de s'intéresser à ces sciences naturellement complexes, pour en observer les méthodologies et hypothèses structurelles. En effet, les résultats de 150 ans de travaux en sciences sociales peuvent certainement être utiles à d'autres domaines. Il reste cependant que confrontées à un objet fluctuant ne permettant pas la

systématique poppérienne, les sciences sociales ont certainement engendré des courants "folkloriques" improductifs (mais très démonstratifs), qu'il n'appartient sans doute pas à des "extérieurs" de juger tels, mais dont il faut être prévenu.

Si l'on admet l'intérêt d'une approche "sérieuse" des sciences sociales, il reste à s'ouvrir et accepter de travailler pour connaître et comprendre. En effet, il faut s'élever contre l'illusion qui veut que les objets considérés nous étant quotidiens et donc connus, point n'est besoin d'une référence à leur étude disciplinaire. En effet, comme le fait remarquer Paul Lazarsfeld (cité dans Bourdieu *et al.*, 1968), sens commun et faits sociaux sont souvent contradictoires. Et par conséquent, seule une recherche approfondie, maîtrisant autant que possible la subjectivité de l'observateur, peut nous renseigner sur la réalité. Notre objet n'étant pas directement le social, si nous souhaitons l'utiliser comme archétype, il faut en chercher les analyses à leur source. Et s'appropriant les concepts, on veillera à les intégrer dans leur totalité, structurés dès leurs fondements, et cohérents avec un cadre méthodologique nécessairement limitant quant à leur validité. Ce faisant on évite de n'en extraire qu'une version affaiblie et superficielle, biais dans lequel tombe bien souvent l'analogie ou la migration conceptuelle (penser aux compromissions dans lesquels le chaos déterministe ou la thermodynamique ont été impliqués). En bref, pour parler du social et en faire un modèle, il faut lire, utiliser et référencer les sciences sociales. L'investissement temporel requis doit alors être regardé comme un gain de clarification et d'objectivation. Dans cette perspective, nous ne cherchons pas ici à faire ce travail par rapport à une question particulière qui nous serait posée en tant qu'ingénieur, mais plutôt à illustrer à partir de quelques concepts clés communs aux sciences sociales et à "la" science des systèmes collectifs, quels gains on peut attendre d'un tel investissement.

Avant de poursuivre, remarquons aussi que si nous avançons en direction des sciences sociales, le contraire doit certainement être recherché aussi. La valeur structurelle des réflexions engagées sur diverses formes de systèmes collectifs, ainsi que l'utilisation d'un outillage avancé et réfléchi de simulation, doivent intéresser en retour les sciences sociales. Il doit donc s'agir d'un dialogue réciproque, dans lequel chacun doit d'une part apporter ses outils et concepts, et d'autre part poser ses questions, auquel l'autre pourrait aider à répondre.

Notre travail étant inscrit dans une démarche d'ingénieurs, l'objectif de cette analyse est donc d'évaluer l'apport potentiel des sciences sociales pour la modélisation, la construction et le contrôle de systèmes collectifs artificiels aptes à résoudre les problèmes concrets qui nous sont posés. Pour cela, nous avons choisi de nous focaliser sur la sociologie, qui nous paraît contenir et manipuler la plupart des concepts nous intéressant. On notera à cet égard que la nature des questions que posent l'articulation collectif / social nous conduit souvent à louvoyer entre sociologie et psychologie sociale. Mais les références que nous avons retenues nous situent au moins formellement dans le domaine de la sociologie. La méthodologie suivie ici va donc consister à choisir un certain nombre de concepts clés, à en donner ou rappeler quelques éléments de définition interne à la sociologie, puis à montrer comment on peut envisager de transposer ces hypothèses sociologiques comme constituants d'une construction raisonnée de systèmes collectifs artificiels. C'est encore une fois à dessein ici que nous choisissons le qualificatif "collectif artificiel" ; dans notre présentation, ce choix ne se voudra pas porteur de sens en tant que tel, mais devrait permettre de bien distinguer ce qui est relatif à la sociologie, et donc aux "collectifs humains", de ce qui est artificiel. Pour les collectifs artificiels, on appellera *agent* les entités.

Dans la suite on a distingué dans chacune des parties le rappel des notions et concepts de sociologie, et *en italique*, l'analyse que l'on peut en faire dans la perspective de l'ingénierie des systèmes sociaux artificiels.

3. Pourquoi le social ? Définition et conditions d'existence

Dans cette partie nous nous efforcerons d'expliquer comment les sociétés humaines se construisent et se maintiennent à partir d'un collectif d'individus.

3.1. Fragments d'une histoire de la sociologie

Historiquement on fait généralement remonter l'institution de la sociologie à Auguste Comte (1798-1857), qui a forgé le terme, et, le premier, a exprimé la nécessité de procéder à une analyse méthodique et rationalisée des faits sociaux et politiques, poursuivant ainsi son projet d'unification du savoir sur une base scientifique. Cette approche positiviste n'a en réalité fait que cristalliser des réflexions philosophiques et historiques antérieures, relatives à l'organisation des sociétés et à leur évolution. En particulier on considère souvent que les fondements d'une sociologie raisonnée sont à rechercher dans Montesquieu (1689-1755) et sa théorie du gouvernement. En montrant que des principes d'action individuelle peuvent structurer une organisation sociale, il amorce le débat relatif à la permanence de certains faits sociaux, et à leur origine. Nous ne nous référerons ici que marginalement à ses travaux, car ils transparaissent dans les théories ultérieures.

C'est dans les révolutions politiques et industrielles du XIX^{ème} siècle que vont se poursuivre les fondements de la sociologie avec Comte, Marx (1818-1883) et Tocqueville (1805-1859). Ceux-ci vont s'opposer largement sur le déterminisme des faits sociaux et leur articulation dans l'histoire. La perspective politique de Montesquieu et Tocqueville laisse une plus large part à la liberté individuelle que les approches structurelles de Comte et Marx, qui insistent sur la stabilité et le persistance de phénomènes sociaux dans les sociétés industrielles. Reprenant la lignée de Comte et Marx, Emile Durkheim (1858-1917) en France et Max Weber (1864-1920) en Allemagne vont achever de structurer les deux grands courants classiques de la sociologie selon des approches respectivement "positive" et "compréhensive" que nous détaillons plus loin. L'oeuvre de Wifredo Pareto (1848-1923) relative essentiellement à la rationalité des comportements humains va introduire (ou reprendre si on admet que Montesquieu en a été l'instigateur) l'idée d'une société organisant l'irrationalité des intentions et des actions des individus.

Enfin tout au long du XX^{ème} siècle, de nombreux courants attachés à des personnalités spécifiques et à des écoles de pensée (souvent politisées) vont contribuer à enrichir les acquis "classiques", et vont en particulier produire une école américaine forte, avec George Mead (construction sociale de la personnalité psychique), Kurt Lewin (théorie du champ social), et Talcott Parsons (théorie de l'action). En France, plus récemment (à partir de 1960), nous retiendrons essentiellement ici deux courants attachés respectivement à Alain Touraine (sociologie des mouvements sociaux) et à Crozier / Friedberg (approche stratégique de l'acteur social).

Cette histoire succincte montre que la sociologie a environ deux siècles d'existence structurée. Et que, comme on le verra par la suite, se sont grossièrement affrontés les tenants d'une part de la rationalité positiviste, davantage intéressés par les phénomènes globaux comme objets cohérents d'étude, et d'autre part les défenseurs d'une approche "locale", compréhensive (au sens de la capacité du sociologue de comprendre de l'intérieur les ressorts de l'action sociale), et interactionniste (l'interaction entre individus comme déterminant de la structure sociale). Il est donc frappant de constater que l'on retrouve ici transposé le débat actuel relatif au système collectif et à leur appréhension respectivement par les approches "top-down" et "bottom-up". On notera qu'en France le courant positiviste de Comte puis Durkheim a très longtemps été dominant.

Par ailleurs l'analyse des faits sociaux n'a pas été l'apanage des seuls "littéraires humanistes" ; ainsi, Comte, Pareto, ou plus tard de nombreux américains comme Kurt Lewin et Herbert Simon, ont fondé leurs travaux sur un corpus de mathématique ou de physique, en appliquant une perspective scientifique rationalisante à l'étude des sociétés.

3.2. L'action sociale

Ce que la sociologie appelle "l'action sociale" va nous intéresser au premier plan, car ce terme recouvre *l'ensemble de l'action humaine dans et relativement aux différents cadres sociaux*.

Niveaux d'analyse

L'action élémentaire est *l'interaction* entre deux individus. L'étude spécifique de la structuration et du comportement d'une personne relève de la psychologie que nous n'abordons pas ici. La combinaison et la structuration des interactions produit des systèmes sociaux intermédiaires comme la famille ou d'autres groupes restreints, puis en poursuivant le regroupement, on atteint les sociétés globales, cohérentes et suffisantes. Il est acquis désormais (cf. Gurvitch, cité dans (Rocher, 1968)) que la séparation des différentes échelles n'a pas de sens ; ainsi l'analyse *microsociologique* (interactions unitaires) et l'observation *macrosociologique* sont inséparables.

Dans la sociologie moderne, on se réfère aussi aux *réseaux sociaux*, des systèmes transversaux, structurés comme des systèmes d'interaction plus étendus mais non institués symboliquement, et qui constituent ainsi des maillons cohésifs de la société.

Définition

Par **Action Sociale**, nous entendons *toute manière de penser, de sentir et d'agir dont l'orientation est structurée suivant des modèles qui sont collectifs, c'est à dire. qui sont partagés par les membres d'une collectivité quelconque de personnes* (Rocher, 1968). On retiendra ici que aussi bien la *manière de penser*, que la *manière d'agir*, et même la *manière de sentir* sont structurés socialement. En particulier la perception et la représentation que l'on a d'autrui déterminent les processus d'action collective.

Dans une perspective artificielle, l'action sociale des agents recouvre ainsi leurs attributs et comportements qui directement (explicitement) ou indirectement (après analyse et transcription dans la structure des contraintes sociales) prennent en compte l'existence et le comportement des autres agents. Il s'agit ici de modèles collectifs "par construction", et non par "observation".

Il existe deux approches principales de l'action sociale, qui ont trait au niveau d'analyse retenu. Tout d'abord la définition subjective de Weber aborde l'action sociale par l'individu et ses interactions. L'idée principale est que l'action est sociale quand elle tient compte des autres, avec une triple condition :

- elle tient compte du comportement des autres et de leur présence et existence ;
- elle a une valeur signifiante ou symbolique pour les autres, et respectivement l'action affiche symboliquement la reconnaissance des attentes perçues des autres (avec tous les malentendus possibles) ;
- elle est modifiée en fonction de l'action des autres et de soi-même.

Pour un agent, la première condition signifie la dépendance (statique) de son état à l'état des autres, la deuxième qu'il existe un mécanisme d'interprétation symbolique de l'action des autres, et un processus d'élaboration d'une réponse effective, enfin la troisième implique que le comportement même de l'agent (sa dynamique, sa stratégie comportementale) dépend de l'état des autres et de lui-même.

Par ailleurs, dans la définition objective de Durkheim, l'action sociale consiste en "des manières d'agir, de penser, et de sentir, *extérieures* à l'individu, et qui sont douées d'un pouvoir de coercition en vertu duquel elles s'imposent à lui" (Durkheim, 1904, cité dans [Rocher, 1968]). L'exteriorité des modes d'action sociale situe cette analyse comme davantage globale et structurelle. Durkheim introduit ainsi le concept de *conscience collective*, qui vient s'opposer à la *conscience individuelle*. La force de coercition des modes collectifs d'action est variable selon les sociétés ; et

détermine le processus d'intériorisation de la conscience collective en chacun. La contrainte est ainsi remplacée par l'habitude et par la conscience morale. Dans la perspective durkheimienne, l'action individuelle peut donc être influencée par le collectif sans qu'il y ait interaction effective, tout en restant cependant sociale.

Il faut rappeler ici que l'objectif premier de la sociologie est l'analyse et la compréhension de l'action sociale, et non sa construction. On peut donc s'interroger sur la pertinence des distinctions opérées entre les deux niveaux d'explication de Weber et Durkheim, pour l'ingénierie des systèmes collectifs artificiels. En particulier le développement de la "conscience d'un agent" n'a à priori pas de sens dans la mesure où leur construction doit être achevée pour envisager l'utilisation effective du système collectif artificiel. Mais cette hypothèse de pré-construction n'est pas indispensable ; en effet on peut tout à fait envisager un processus d'adaptation par apprentissage, et donc effectivement d'intégration du fait social.

Par ailleurs les deux analyses ont bien une transcription possible : soit ("subjectivement") on ne spécifie pour les agents que des modes d'interaction sociale qui vérifient tout ou partie des conditions weberiennes, et on va alors observer ou non l'émergence de processus collectifs particuliers, et l'obtention d'une réponse fonctionnelle à certain état de l'environnement ; soit ("objectivement") on construit à priori un répertoire d'actions et de connaissances (une conscience collective artificielle), un ensemble de règles parmi lesquels les agents vont puiser pour déterminer leur comportement. Le travail de construction de cette "conscience collective artificielle" suppose que l'on ait analysé auparavant les différentes configurations collectives possibles afin de toujours pouvoir orienter l'action. En pratique il semble que les deux approches doivent être concomitantes et qu'il est souvent difficile de tout spécifier individuellement, sans en passer par des déterminants collectifs donnés à priori. Le cadre de l'agent autonome (p.e. un "agent WWW") nous limite cependant souvent au local, par méconnaissance exhaustive des situations et autres agents rencontrables (on touche ici pour la première fois au problème des systèmes collectifs artificiels en environnement ouvert).

3.3. Les normes dans l'action sociale

Par nature, l'humain est être de pulsion, désir ; et ce qui va mettre la conduite des individus en conformité avec des manières collectives d'agir, de penser et de sentir, ce sont *les normes*. Elles vont leur donner un cadre comportemental commun. On peut donc lire la signification de la conduite des individus à partir de règles ou *normes de communication et d'action*, sur lesquelles il existe un consensus accepté par les acteurs. On notera que *les normes se réfère bien à des actions et comportements concrets et observables*. Elles n'ont pas en tant que tel d'expression symbolique (la sentence "tu ne tueras point" n'est pas une norme ; l'absence de meurtre en est une).

Les normes assurent la régulation fonctionnelle du collectif. Elles en constituent le "code de fonctionnement". Par exemple "les agents laissent toujours passer plus gros qu'eux", ou en terme de communication, "ils respectent tel protocole". Du point de vue constructif il est difficile de distinguer clairement les normes imposées par le concepteur des simples limitations fonctionnelles non évoquées en sociologie, comme celles relevant du principe de réalité : "on ne peut pas voler", "on peut mourir", etc. On peut admettre que les normes sont des limitations imposées à des fonctions disponibles (les agents se déplacent : fonction, verticalement : norme). Mais ceci est très artificiel car inversement on peut admettre que certaines contraintes fonctionnelles humaines résultent de normes.

D'autre part le caractère non symbolique des normes peut être traduit par ce qu'on qualifie parfois de registres de comportements des agents, c'est à dire des ensembles d'actions accessibles en fonction directement de percepts (registres sensori-moteurs).

L'interaction est ainsi structurée à partir de normes collectives, mais celles-ci peuvent relever de différentes "sous-sociétés" ou "tribus". L'individu est donc le "produit des tribus" (Rocher, 1968), au sens que sa conscience individuelle est forgée par la combinaison des différents niveaux de conscience collective qu'il intègre.

Si on définit différentes "classes d'agents" par des registres d'action, un agent particulier peut éventuellement être construit par combinaison de classes.

D'autre part les normes sont fondamentales pour assurer la deuxième condition weberienne : la conduite humaine n'est significative et cohérente aux yeux d'autrui que par référence à une structure de règles ou normes collectives. Elle s'inscrit dans un cadre reconnu qui la rend sociale. Et cette cohérence permet à tout individu de prévoir les comportements d'autrui, et donc d'élaborer des stratégies sociales. La prospective des actions et réactions des autres obéit ainsi à un ensemble de *modèles culturels* ("cultural patterns").

Le respect des codes décrits par les normes est certes une garantie de bon fonctionnement. La référence au code de la route est édifiante : en fonction de ce qui est supposé respecté par autrui, chacun construit une stratégie locale. La confrontation des modèles culturels français et anglais de code de la route est évidemment désastreuse.

Le rôle social

Le répertoire de normes utilisés par les différents acteurs n'est pas identique. Même si il y a cohérence des différents systèmes, l'action des sujets qui occupent une position ou une fonction particulière dans un groupe ou une collectivité est soumise à un ensemble spécifique de normes, qui définissent ainsi un *rôle social*. Celui-ci est modulé selon les personnalités, mais il garantit certaines stabilités comportementales utiles à l'action sociale. Le *rôle* doit être bien différencié de la *fonction* : le *rôle* est l'ensemble des manières d'agir qui, dans une société donnée, sont censées caractériser la conduite des personnes dans l'exercice d'une *fonction* particulière (cette définition du *rôle* n'est pas très naturelle).

Pour assurer telle fonction (définie par l'organisation, cf. ailleurs), un agent doit respecter telles normes, qui précisent son rôle : la fonction de balayeur suppose qu'en présence de déchets l'agent les balaye...

Etant donné la multiplicité des fonctions que remplit un individu particulier au sein d'un collectif, il est donc amené à "jouer" différents rôles sociaux. Ce qui a conduit (dès Weber) à la notion d'*acteur social* (reprise et enrichie par Touraine et Crozier), qui dispose, de même qu'au théâtre, d'une capacité de modulation de son jeu. Cependant la complémentarité et l'interdépendance des rôles au sein d'une société cohérente contraignent leurs variations. De manière générale, il est souhaitable (mais difficile) de distinguer les rôles pour bien comprendre le fonctionnement social.

La modulation du jeu de l'agent est l'élément de fluctuation qui doit permettre une plus grande souplesse dans les comportements collectifs. Ainsi, toujours dans l'exemple du code de la route, le comportement face à un feu orange peut fluctuer depuis l'arrêt jusqu'à l'accélération, en fonction des capacités du véhicule (paramètre intrinsèque), ou de la présence d'un policier (paramètre extrinsèque).

Sanctions

Les normes sont fonctionnellement justifiées au titre du collectif social, mais cela ne suffit pas à les imposer aux individus, qui peuvent vivre avec des systèmes d'action "anormaux". Il existe donc un mécanisme de *sanction* : la conformité est récompensée, et l'insoumission punie. On distingue classiquement (Rocher, 1968) quatre types de sanctions (négatives ou positives):

- *physiques*, qui instituent la violence sociale,

- *économiques*, par contrôle social des biens et richesses de l'acteur,
- *surnaturelles*, par imposition d'une logique symbolique magique ou surnaturelle,
- *sociales*, par modification des processus d'interaction et de la place sociale,

L'ensemble des sanctions assurent la cohésion et le fonctionnement de la société, et permettent le *contrôle social* (mise en conformité des conduites aux modèles établis).

Par construction l'agent est conforme. Cependant si on accepte les fluctuations ou l'adaptation, à des fins d'amélioration du système, il faut pouvoir sélectionner. Dans les systèmes évolutionnaires artificiels on définit ainsi la "fitness" et un paramètre énergétique qui éliminent peu à peu les moins adaptés. C'est un exemple de sanction artificielle de type "économique". Si on reprend les différents types de sanction, on observe que pour permettre une sanction physique, il faudrait concevoir une forme d'intégrité "corporelle" de l'agent qui pourrait être dérangée, l'extrême étant la disparition (mort). Imaginer une sanction surnaturelle nous entraînerait trop loin dans les fondements sociaux du surnaturel. Par contre la sanction sociale est fondamentale car elle recouvre tous les processus d'avancement hiérarchique, d'éviction, et de "boycott" social.

En général les sanctions vont s'appliquer sans qu'il y ait eu "crainte" auparavant puisqu'un tel niveau de représentation n'est en général pas fourni à l'agent.

Une théorie récente (Coleman, 1990, *Foundation of Social Theory*, Chicago), d'orientation apparemment "utilitariste", pose différemment le problème de la norme et de la sanction. Elle énonce que les normes n'ont de réalité qu'au travers d'un ensemble d'acteurs les supportant, et qu'en particulier les normes ne s'imposent pas à priori aux acteurs. Il faut pour que le contrôle social s'effectue que des acteurs "s'en chargent", et en aient les moyens. Car en effet l'imposition de normes n'est pas gratuite. Comme toute activité, elle a un coût. Dès lors il faut que le collectif institue un ou des acteurs comme régulateurs et leur attribue une part des ressources collectives à cet effet. Cette théorie est justifiée par l'idée que si le collectif fonctionne bien, alors la part redistribuable est augmentée ; mais qu'en général chacun n'a pas intérêt à collaborer (problème standard du compromis vs. consensus en terme d'utilités). Ce mécanisme de sanction est très élaboré car il suppose déjà constituée une conscience collective, ainsi qu'une capacité de représentation sociale. Par contre on peut s'interroger sur la capacité d'une telle société à innover, au moins formellement.

Socialisation normative

Que ce soit dans l'éducation de l'enfant ou l'intégration d'un "étranger" dans une société, il doit y avoir acceptation et intériorisation des normes sociales. Cette *socialisation* vise le développement de dispositions ou tendances conformistes. Pour ce faire, les désirs, attentes ou besoin subissent par le jeu des sanctions un processus de sélection, qui institue "le bien et le mal". Les normes sont ainsi toujours relatives à une société particulière, et elles n'ont pas de valeur absolue.

Le processus de socialisation illustre l'imbrication des perspectives de Durkheim et Weber. En effet, partant de ce qui est collectivement accepté, chaque individu se construit en intégrant et faisant sien des rôles particuliers, qui structurent intégralement son être. L'intériorisation due à la socialisation conduit même, comme Freud l'a montré, à produire une image sociale autocontraignante. La conformité est alors désirée.

Dans les systèmes sociaux artificiels clos, la socialisation n'a de sens que relativement à l'arrivée d'un agent dans un groupe préétabli, ou à sa migration. Par contre, dans les systèmes ouverts (du type agents sur réseaux télématique), la socialisation est fondamentale car elle permet l'intégration des "nouveaux" au système préexistant, sans connaissance à priori du contenu.

Ordre social et liberté des acteurs

L'ordre du social ne reflète que le processus de régulation permettant le fonctionnement, et non un quelconque contrat explicite ou "ordre naturel". Le déterminisme social qui conduit à la standardisation des conduites individuelles laisse cependant une place à la spontanéité (jusqu'à même affronter l'ordre social). Comme on l'a énoncé précédemment, il existe une échelle "de rigidité" des modèles allant du plus impératif au plus facultatif, de l'interdiction à la préférence ; par ailleurs les modèles sont plus ou moins réalistes, certains relevant de "l'idéal". Rappelons aussi qu'il existe un ensemble de choix possible de modèles : il y a donc toujours une part de décision, en particulier entre des modèles divergents. La liberté de l'acteur réside alors dans ce choix (*variance*).

Le recours à des modèles marginaux ou interdits reste possible, il y a alors *déviance*. Ce choix peut être source de changement social lorsqu'il y a reconnaissance par un groupe marginal. La coévolution des différents systèmes de normes est une des lecture possible de l'histoire sociale.

Le déviance d'un agent est un concept séduisant, mais difficile à mettre en oeuvre sans utiliser des processus aléatoires. On peut aussi introduire une proportion de déviants et observer si ils persistent et quel effet ils produisent.

3.4. Eléments symboliques de l'action sociale

Parmi les problèmes que posent la distinction entre "collectif et social", la place du symbolisme nous paraît déterminante car il est le propre de l'homme, et donc on en vient toujours à s'interroger sur les propriétés sociale qu'un système collectif non symbolique peut et ne peut pas présenter.

Valeurs

Les valeurs d'une personne ou d'un collectif sont les manières d'être ou d'agir reconnues comme idéales (références positives), et qui donc rendent désirables ou estimables les êtres ou les conduites auxquelles elles sont attribuées. La valeur est dans l'ordre idéal, mais elle se concrétise par des conduites ou des objets concrets. Les valeurs permettent le jugement ("de valeur") et le choix des conduites. Les modèles (normes) se justifient à l'intérieur d'un système de valeur particulier. Cependant on observe dans les processus historiques des déviations relatives aux valeurs, qui peuvent conduire à des modèles perçus comme inconsistants.

Les valeurs donnent des critères d'évaluation, qui permettent de choisir des actions. En l'absence de valeurs, le comportement est indéfiniment réactif. Et tous les choix de modèles précités sont impossibles. Sans valeurs l'agent n'est pas libre, au sens qu'il ne dispose pas de l'outillage symbolique lui permettant d'élaborer un raisonnement spéculatif et le jugement afférent. Par contre dans le cadre artificiel les valeurs ne sont pas indispensables pour obtenir une société structurée, puisque le concepteur peut traduire des systèmes implicites de valeur dans les modèles qu'il impose. Cependant le système collectif artificiel ainsi formé vit avec ces valeurs sans les reconnaître (de même que beaucoup de sociétés humaines n'explicitent pas certaines valeurs).

De même que les normes, il y a relativité des valeurs vis à vis d'une société et d'un temps. Et la place des différentes valeurs à l'intérieur d'une société n'est pas immuable. La *hiérarchie des valeurs* ainsi constituée pose des *valeurs dominantes* et des *valeurs variantes*, dont les places respectives peuvent s'inverser à l'occasion d'un changement social.

Par ailleurs, les valeurs sont acceptées par un processus plus émotionnel ou intuitif que raisonné. Dès lors le recours individuel ou collectif à des systèmes de valeurs contradictoires est possible.

Selon Talcott Parsons, le choix d'un système de valeurs ("pattern variables"), traduit cinq alternatives sociales :

- 1) *affectivité* ("le flot des sentiments") vs. *neutralité affective* (contrôle et rationalisation),
- 2) *universalisme* (unicité du jugement relatif aux acteurs et faits sociaux) vs. *particularisme* (usage de normes particulières spécifiques et contextuelles),
- 3) *option de l'être* (évaluation d'autrui en fonction de son être propre) vs. *option de l'action* (jugement en fonction des actions),
- 4) *globalisme* (autrui vu comme un tout cohérent) vs. *spécificité* (autrui réduit à une dimension de son être ou action),
- 5) *égocentrisme* (selon des buts personnels) vs. "*sociocentrisme*" (selon des choix et buts collectifs).

L'alternative 1) a peu de traduction artificielle du fait de l'absence de "sentiment". Le particularisme peut se traduire par une priorité accordée à certains autres agents. L'option de l'être ne considérera que les attributs d'autrui, alors que l'option de l'action verra son état. La spécificité s'intéressera à une fonction particulière, alors que le globalisme retiendra un agent donné. Enfin le dilemme le plus important est celui de l'égo vs. socio centrisme car il se réfère aux conflits entre intérêts locaux et globaux.

Les choix ("options") de valeur peut être fait soit par la communauté soit par un individu. Les options individuelles sont cependant conditionnées par les options collectives. Toute option individuelle non collective est facteur d'innovation sociale.

Fonction sociale des valeurs

Les valeurs ont différentes fonctions sociales :

- la *mise en cohérence* des modèles (normes) : les valeurs permettent d'expliquer (en partie seulement du fait de leur caractère irraisonné, et donc parfois contradictoire) la coexistence de modèles culturels apparemment discordants ;
- les valeurs permettent l'*unité psychique des acteurs* en rendant cohérent et structuré leur perception de soi et du monde, ainsi que leur processus de décision (motivation unifié) ;
- enfin et surtout elles assurent la *cohésion sociale dans le champ symbolique* : l'intégration sociale passe par l'adhésion aux valeurs communes.

La cohésion du champ symbolique d'un système collectif artificiel renvoie à la définition de standard de communication (choix de langages et protocoles) au niveau le plus technique, mais aussi surtout à des registres d'évaluation.

Rôle des symboles

Pour assurer leur fonction sociale, les valeurs doivent être affirmées et donc visibles. Or elles sont idéales, et ne sont représentées que par certains modèles "de référence". Si on retient qu'*un symbole est une chose qui tient la place d'autre chose*, alors on conçoit que l'univers des valeurs et des modèles est un vaste ensemble symbolique dans lequel baigne l'action sociale. En particulier, les symboles vont rendre réels et accessibles les valeurs, et vont donc permettre leur explicitation ou leur réfutation. Ainsi le symbolisme va faciliter les changements de valeurs et donc l'adaptation de la société aux contraintes du milieu. L'extension de sa capacité symbolique a ainsi permis à l'homme de maîtriser le monde en se socialisant, mais inversement la socialisation a facilité et renforcé l'activité symbolique.

On est ici au coeur du débat relatif à la nécessité d'un systèmes symbolique dans les collectifs artificiels. Si on fait l'hypothèse d'agents non "symbolistes", et en prenant le contre-pied de la proposition précédente, on a absence d'affichage des valeurs. Dès lors il ne peut y avoir de changement intentionnel dans le système de valeur. Par ailleurs les valeurs ne peuvent exister qu'implicitement, telles que le concepteur les a choisies, et transcrites dans des modèles culturels (systèmes de normes) artificiels. L'adaptation de la société aux contraintes du milieu est limitée alors à ses évolutions sélectives. L'identification d'une

structure complexe, imbriquée, de sous groupes identifiés par des modèles culturels cohérents dépend entièrement de l'observateur extérieur, en l'absence de système symbolique interne. Il ne peut donc y avoir changement social dans un tel collectif, et la cohérence globale reposera sur la continuité du milieu environnant.

Au contraire de ce qui est énoncé pour les sociétés humaines, les modèles culturels artificiels ne sont pas forcément symboliques. Car en tant que systèmes cohérents de normes ils constituent essentiellement des répertoires de comportements réductibles à de simples processus sensori-moteurs conditionnés par le milieu.

La fonction sociale des symboles est de deux types :

- 1) *communication* sous différentes formes et langages de différents concepts ;
- 2) *participation des acteurs au collectif* au travers de symboles de solidarité, d'organisation hiérarchique (le symbole représente une échelle sociale exogène), de l'histoire collective (mémoire collective structurante), ou encore religieux ou magiques.

La fonction symbolique de communication est prépondérante pour les systèmes collectifs artificiels. Elle permet d'opérer une distinction fonctionnelle entre les actions et interactions avec les objets de l'environnement et avec les autres agents. Car en effet, en l'absence de symbole, les interactions sont limitées à des processus passifs où les agents se perçoivent mutuellement. L'action éventuelle d'un agent sur un autre ne résulte que de l'application d'une norme, sans référence à un quelconque message signifié par cette action. Dès lors il n'y a plus lieu de distinguer la société de l'environnement, puisque les processus sont strictement de même nature. Autrui perd son statut de "semblable".

Cette analyse est trop extrême, car le concepteur / observateur est toujours susceptible de distinguer les normes relatives à l'environnement de celles relatives au corps social. Hormis pour des besoins analytiques formels, on peut donc conserver, même pour des systèmes très "réactifs", deux catégories fonctionnelles séparées.

Les symboles restent toujours ambigus, car il n'y a pas vérification et réactualisation perpétuelle des significations. Or les symboles remplacent la réalité, et donc la société peut fonctionner avec une structure symbolique distordue d'avec son environnement réel ; jusqu'à ce qu'elle en subisse les conséquences.

Dans le cas des systèmes collectifs artificiels ouverts, la construction symbolique doit en permanence être adaptée, puisque l'environnement n'est pas clos (condition d'un système symbolique statique). On peut ici faire référence aux systèmes de révision de croyance en Intelligence Artificielle.

Pour résumer cette partie, on voit que l'action sociale a trois fondements : les normes, les valeurs (idéaux) et les symboles (qui permettent l'actualisation des valeurs).

Quels que soient les systèmes collectifs artificiels, les normes sont toujours présentes, alors que les valeurs et les symboles peuvent n'exister que dans l'esprit du concepteur qui les transcrit dans des structures et processus d'interaction.

3.5. La culture

La culture est le niveau le plus global d'analyse de l'action sociale. Nous retiendrons que *la culture est l'ensemble lié de manières de penser, de sentir et d'agir plus ou moins formalisées, qui étant apprises et partagées par une pluralité de personnes servent, d'une manière à la fois objective et symbolique à constituer ces personnes en une collectivité particulière et distincte.* On peut distinguer deux fonctions de la culture :

- 1) fonction sociale : réunir une pluralité de personnes en une collectivité spécifique. La culture est donc l'univers mental, moral et symbolique commun à une pluralité de personnes, grâce

auquel et à travers lequel des personnes peuvent communiquer, se reconnaître, se lier, et se sentir membres d'une même entité.

2) fonction psychique : la culture constitue le moule des personnalités ; elle offre un choix de modèles, de valeurs, de significations symboliques ; mais ce choix n'est pas illimité, il caractérise les options possibles. La culture confère ainsi une "forme" à la société.

Une culture artificielle est donc un cadre symbolique dans lequel inscrire les agents. Elle caractérise l'ensemble des attributs d'un système social et de ses individus. On peut admettre que tout système collectif artificiel apte au symbolisme a une culture.

Guy Rocher (1968) établit un parallèle intéressant entre la culture et l'instinct. Il montre ainsi que si l'instinct traduit "l'adaptation principale de l'animal à ses besoins et à son environnement", et donc lui permet de "répondre à la réalité", la culture, qui est propre à l'homme, remplit le même rôle pour l'homme dans son cadre social. Cette hypothèse implique en particulier que le recul de l'instinct chez l'homme correspond aux avancées de la culture, comme le montre aussi différentes analyses des évolutions de l'enfant.

La coexistence dans un système social artificiel de processus de type instinctifs (ou réactifs) et culturels (symboliques) traduit la distinction désormais explicite entre société et environnement.

Idéologie

L'idéologie est un système d'idées et de jugements, explicites et organisé, qui sert à décrire, expliquer, interpréter et justifier la situation d'un groupe ou d'une collectivité, et qui, s'inspirant de valeurs, propose une orientation précise à l'action historique de ce groupe ou de cette collectivité (Rocher, 1968). L'idéologie est donc cohérente, explicative, et oriente l'action. Elle constitue une représentation symbolique de la société et de ses valeurs par elle-même, et dans sa globalité. Elle affiche explicitement son intention de faire l'unanimité (alors que ceci n'est qu'un effet des systèmes de valeurs). Elle est le premier niveau d'expression symbolique de la société comme entité objective. Et donc elle est un prérequis pour l'élaboration d'une stratégie de changement social.

Une idéologie dans un système collectif artificiel est une construction symbolique qui a pour objet l'action sociale elle-même. Il s'agit donc de ce qu'on pourrait appeler un niveau "méta-social" qui n'est pas encore très répandu, à notre connaissance, parmi les agents artificiels. Cependant dans tous les systèmes collectifs supervisés (avec un contrôle centralisé) il existe au moins une idéologie implicite qui est celle utilisée par le superviseur.

Les idéologies sont souvent le fait d'élites qui ont la capacité d'analyser le social. Il peut coexister des idéologies différentes autour d'une même société.

Socialisation "culturelle"

Il s'agit ici des processus d'apprentissage et d'intériorisation des éléments socio-culturels du milieu, et de leur intégration à la personnalité, sous l'influence d'expériences et d'agents sociaux significatifs. Il y a respectivement acquisition symbolique de la culture, intégration à la personnalité, et donc tendance vers la conformité.

Les mécanismes de socialisation sont de nature psychique (et non biologiques), et ils s'inscrivent dans une évolution liée à l'histoire de chaque individu au sein de la collectivité. D'une part il y a *apprentissage* par essai, répétition, imitation, avec des récompenses et des punitions. Le milieu social transfère ainsi la culture instituée. D'autre part il y a *intériorisation d'autrui* par des processus spéculaires (j'imagine l'image que j'ai), ou le développement de ce que George Mead appelle la "notion d'autre généralisé", c'est à dire une image générique du collectif.

La perspective de construction d'un mécanisme de socialisation culturelle artificielle nous paraît bien utopique pour au moins deux raisons : tout d'abord la durée et la quantité

d'information digérée durant la socialisation humaine est bien supérieure à tout ce qu'on peut imaginer dans un système artificiel, et d'autre part les bases fonctionnelles physiques et psychiques de la socialisation sont d'une telle complexité que l'on imagine mal de les reproduire (encore faudrait-il les connaître...)

L'apprentissage est un processus artificiel déjà largement étudié. Articulé autour de normes, valeurs et sanctions, il peut justifier d'un nouveau regard. L'intériorisation d'autrui est aussi intégrée dans les modèles artificiels d'agents cognitifs, qui disposent de modèles des autres et de processus spéculaires d'élaboration de plans coopératifs.

3.6. Conclusion temporaire

Les théories sociologiques sont très nombreuses et très variées. Nous n'avons considéré ici que les théories classiques, c'est à dire anciennes. Nous souhaitons cependant citer au moins une fois une approche qui nous paraît fondamentale pour nos préoccupations d'ingénieurs, et sur laquelle nous pourrions revenir dans une suite à cet article. Il s'agit de l'approche "stratégique" de Crozier et Friedberg (1977) qui posent trois postulats :

1. les acteurs conservent toujours des objectifs propres, et ne sont donc jamais de simples rouages dans une grande machine sociale,
2. les acteurs se conservent toujours une marge de manoeuvre, une liberté individuelle, et en usent,
3. la rationalité des stratégies est limitée (cf. les travaux de Simon).

Sur cette base s'élabore une théorie compréhensive qui illustre en particulier comment le social est une combinaison de tensions contradictoires qui s'équilibrent, et donc qui instituent le conflit comme "moteur social".

Cette perspective déplace complètement l'analyse puisqu'elle conduit l'ingénieur à imaginer des types d'agents avec leurs stratégies globales, alors que précédemment on se référait des modes d'interactions. Elle est la transcription sociologique de l'approche "bottom-up" la plus radicale.

4. Les systèmes sociaux humains doivent-ils constituer un modèle pour l'ingénierie des systèmes collectifs artificiels?

Après avoir présenté quelques notions de sociologie, assorties de commentaires relatifs à leur applicabilité à l'ingénierie de systèmes collectifs artificiels, nous souhaitons aborder ici la question de la pertinence de cette référence sociologique, qu'elle soit implicite ou explicite. En effet, soit elle apparaît comme nécessaire ou incontournable, et le travail de lecture transversale systématique entamé ici s'impose ; soit une alternative "désocialisée" existe, et l'effort méthodologique doit alors porter sur le "nettoyage" conceptuel.

Nous allons donc détailler le cadre de l'ingénierie des systèmes collectifs artificiels en trois sous-classes d'épistémologie différente : la simulation de systèmes complexes, la résolution de problèmes, et le développement d'agents autonomes pour des systèmes ouverts distribués.

Simulation de systèmes complexes

La simulation de systèmes complexes à partir de systèmes collectifs artificiels repose sur un principe d'analogie structurelle consistant à transcrire dans des agents artificiels les objets du système source. En construisant ainsi un ensemble d'agents en interaction reproduisant individuellement la dynamique locale des entités systémiques, on obtient une image artificielle du système source, qui peut soit valider des hypothèses locales par observation des processus émergents, soit servir de laboratoire virtuel pour tester des hypothèses comportementales ou des registres d'action. La condition de réalisation et d'utilisation d'une telle simulation est que le support

matériel à l'implémentation des agents respecte le parallélisme intrinsèque (ou la séquentialité fine) des processus réels simulés.

Dans ce cadre, on peut admettre que les systèmes sociaux humains ne vont constituer un modèle que lorsque le système à simuler inclue une société humaine (proposition quasi-tautologique). Le recours à des modèles sociaux humains comme métaphore pour la lecture d'autres types de systèmes complexes ne devrait constituer qu'une médiation transitoire entre ces systèmes sources et leur simulation ; cependant, cette possibilité est souvent une réalité dans la mesure où les systèmes d'agents artificiels utilisés ont été fondés initialement sur des modèles explicites ou implicites de comportements humains. Si en revanche on cherche à simuler une société humaine, alors toutes les ressources de la sociologie (ou de la psychologie) peuvent être utiles pour déterminer les bons attributs et comportements à attacher aux agents artificiels. Dans une stricte perspective de modélisation / simulation, cela suppose que l'on dispose effectivement d'un modèle social transposable dans un système informatique. A cela deux perspectives : soit on ne reconnaît comme modèles *sui-generis* que des descriptions formalisées au moyen de langage axiomatisés autoréférents ("formels"), comme les mathématiques, et dont la transposition ne rencontre pas de difficultés techniques insurmontables. Dans ce cas, on s'expose à ne pouvoir rendre compte que de fractions limitées de la réalité sociale humaine, qui semble par nature difficilement réductible à des déterminismes stricts. Ainsi, par exemple, les théories économiques peuvent être supportées analytiquement, mais leur validité reste évidemment limitée (encore que les systèmes économiques puissent à l'inverse espérer l'adaptation des individus réels au modèle). L'alternative est d'accepter que des modèles sociaux soient issus de démarches descriptives, fondées sur le discours, et structurées autour de lexiques argumentés, et eux aussi largement autoréférents. Dès lors on sort du cadre poppérien, mais en retour on étend les modèles, en particulier vers les fluctuations du champ symbolique, ou la diversité des comportements humains. Or il ne semble pas que cette approche soit moins bonne en terme de vraisemblance et d'opérationalité des modèles. En effet, la validité d'un modèle n'est pas intrinsèque ; elle résulte du choix fait par l'observateur / modélisateur quant à la clôture du système considéré. Et en particulier l'universelle validité de ce modèle (comme constitution d'une réalité acceptée socialement, et donc opératoire) repose sur le consensus. Qu'il se fasse sur la clôture comme dans les modèles formels, ou sur l'idiolecte pour les modèles discursifs, on se trouve confronté à des enjeux subjectifs de même nature. On objectera cependant que la transposition informatique n'est, dans l'approche discursive, pas acquise. Cela tient au paradigme dominant actuel qui suppose une description globale et séquentielle des processus ("de haut en bas"). Si on reconnaît, comme les tendances actuelles de la simulation distribuée portent à le croire (Ferber, 1995b), que l'approche adéquate de la simulation des systèmes complexes est la construction "de bas en haut", par production de micro-modèles interdépendants au sein d'un système englobants, alors on ouvre la voie à l'implantation informatique des modèles discursifs, au moyen de systèmes symboliques structurés, traduisant directement les concepts et leur articulation. Et donc, il revient aux disciplines descriptives et explicatives, ici la sociologie, ou les sciences sociales en général, de renseigner sur les modèles consensuels (ou prospectifs) à considérer. La notion de règle symbolique, au sens des systèmes à base de connaissance, constitue alors le médium de leur implémentation, transposé dans des modèles distribués de type multi-agents. La conclusion ici est qu'il est possible de modéliser et simuler des sous-systèmes de sociétés humaines, formellement ou non. La validité de l'approche va alors dépendre des objectifs, comme on le voit ci-dessous.

Lorsque l'objectif de la simulation est la vérification d'hypothèses structurelles ou fonctionnelles relativement à certain objet connu à priori (en général un système ou un fait social observé globalement), il faut disposer de données permettant la comparaison et la validation. Les enquêtes et analyses sociologiques nous les fournissent. On construit alors un modèle d'individu, et on observe le résultat social émergent. Si il correspond aux données, le modèle local est validé ; sinon, le modèle local doit être ajusté.

Si le modèle social a été validé, il peut alors servir un autre objectif : la prédiction et le test d'hypothèse. On accède alors à ce type de laboratoire virtuel dont rêvent de pouvoir disposer les sociologues pour une pratique expérimentale impossible autrement. On peut s'interroger en retour sur les potentialités induites de contrôle du système réel, c'est à dire la société (ici le sous-système modélisé).

La sociologie peut aussi à l'inverse être utile en tant que référence méthodologique pour l'observation de la société artificielle : en effet les outils statistiques et les procédures d'enquête établies systématiquement en sociologie pourraient être utiles pour évaluer l'état des systèmes sociaux artificiels autrement que par la seule observation "gelstatique". L'absence de méthodologie claire en matière de simulation provient essentiellement de deux causes : la multiplicité des paramètres pas toujours indépendants et la difficulté à en mesurer les effets. Or précisément, la sociologie a, comme on l'a montré plus haut pour l'analyse de l'action sociale, cherché à discerner dans la complexité des sociétés quels en sont les déterminants structurels et fonctionnels. Par référence à la sociologie on peut donc envisager de "nettoyer" un certain nombre de modèles sociaux redondants. D'autre part les outils et les méthodes d'analyse factorielle manipulées en sociologie pourraient faciliter l'identification des dépendances dans les systèmes sociaux artificiels et donc permettre leur ajustement.

Résolution de problèmes

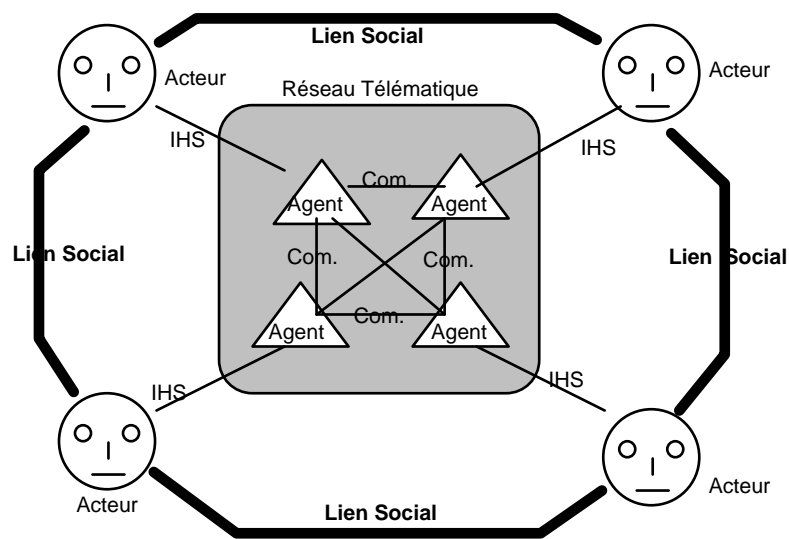
Dans le cas de la résolution de problèmes, l'idée est de construire un système collectif artificiel qui va pouvoir résoudre collectivement le problème posé. Dans ce cadre toutes les options sont ouvertes tant que l'on trouve effectivement une solution. En particulier de nombreux travaux ont montré qu'il n'était pas du tout nécessaire d'utiliser des agents très compliqués pour parvenir à cette fin (computation émergente ; Forrest, 1991). On peut alors renverser la question et s'interroger sur les problèmes qui ne sont solvables que par un collectif d'une complexité équivalente à une société humaine. Ce "méta-problème" reste ouvert (est-il solvable par un collectif simple ?). A cet égard on peut observer que la référence sociale humaine est complexe, car une bonne part du problème que résout implicitement le groupe est simplement de se maintenir. Si on se réfère à l'analyse stratégique (Crozier), le groupe humain est conflictuel, et donc il consomme des ressources pour se réguler. Dans le cadre artificiel, on peut imaginer d'optimiser le fonctionnement en réduisant les conflits. Mais il faut auparavant analyser le rôle fonctionnel des conflits : ils sont source de dynamique ; ils permettent l'innovation ; ils assurent le changement. Un collectif artificiel sans conflits peut apparaître trop statique et non adaptatif. En résumé, en résolution de problème la fin (la résolution) prime sur les moyens, et on peut être tentés, comme en résolution distribuée de problèmes, par des analogies sociales, mais le coût risque d'en être élevé si on néglige ou méconnaît le coût structurel, et donc corrélativement computationnel, d'auto-préservation de la société.

On notera enfin que la perspective de la résolution de problèmes doit sans doute être abordée par l'intermédiaire de la sociologie des organisations (systèmes structurés autour de buts donnés), non traitée ici (Bernoux, 1985).

Agents autonomes en environnement ouvert

Le développement des grands réseaux télématiques (type Internet) a conduit différents laboratoires à concevoir des agents "autonomes" évoluant sur ces réseaux et effectuant un certain nombre de tâches (Maes, 1994). Ces agents vivent dans un environnement constitué de serveurs d'informations, et ils communiquent et interagissent avec d'autres agents, et éventuellement des utilisateurs humains (dans le cas des agents assistants ou prospecteurs). On a donc le schéma de la figure suivante où trois types d'interactions existent : entre agents artificiels (communication), entre utilisateurs (lien social entre acteurs sociaux), et entre un agent et un acteur (IHS = Interface Homme-Système). Chacun de ces systèmes d'interaction peut relever d'une modélisation sociale.

Mais l'ingénieur n'intervient directement que dans la construction des agents eux-mêmes. Il ne "contrôle" évidemment pas les interactions sociales entre acteurs.



Les différentes interactions autour des agents autonomes sur les réseaux

Selon un modèle classique, un tel agent autonome doit avoir des compétences propres relatives au traitement des informations, des capacités relationnelles et coopératives vers ses semblables (accointances), et une capacité éventuelle d'interaction avec son environnement exogène, ici les acteurs. Sa conception peut donc faire appel à quatre sous-modèles, éventuellement sociaux :

1. le collectif des agents artificiels, ses règles, son organisation ;
2. l'IHS, en tant qu'interaction ouverte entre un agent et un acteur quelconque d'une société ;
3. le modèle de société humaine que les agents doivent utiliser pour planifier leur action (à noter que ce modèle est en fait "un modèle dans un modèle", puisqu'il est inscrit dans le modèle d'agent comme un modèle de référence) ;
4. le modèle d'interaction globale entre la société humaine et le collectif d'agents, en tant que possibilité de modifier indirectement le lien social par la mise en place d'un lien artificiel médiatisé par la relation "Acteur 1 - IHS 1 - Agent 1 - Communication - Agent 2 - IHS 2 - Acteur 2".

La problématique est globalement différente de celles de la résolution de problèmes car le milieu d'évolution de l'agent est largement ouvert et incertain. Il n'est évidemment pas contrôlé par le concepteur initial de l'agent. Les agents et contextes que l'on peut rencontrer ne sont pas connus a priori. Seul un ensemble de normes techniques (langages, protocoles de communication) sont fixées.

La nécessité ou l'utilité d'une référence sociale humaine diffère selon les sous-modèles considérés. Dans le cas du collectif artificiel (cas 1.), le recours systématique à un modèle de système social humain ne semble pas justifié, dans la mesure où il s'agit d'organiser un système de communication et de coopération entièrement artificiel, et donc non astreint aux contingences humaines ; dès lors une conception *ab nihilo* fondée sur une méthodologie d'analyse stricte est sans doute préférable.

Il reste que dans ce cas la référence sociale peut être utile si on considère l'enjeu de la constitution et la maintenance de relations entre agents très hétérogènes au sein d'un système qui doit rester malgré tout cohérent et fonctionnel. Ceci est particulièrement vrai pour les agents "libres", c'est à dire ne dépendant pas d'un utilisateur. Ils assurent des services généraux de collecte d'information, ou d'organisation des communications. Et ils peuvent alors former des sociétés artificielles à la condition qu'ils puissent communiquer entre eux et se reconnaître (Lashkari, 1994). La question que leur conception peut poser à la sociologie est d'une part de savoir quels sont les

conditions structurelles d'un fonctionnement social (quel modèle d'agent utiliser ?), et d'autre part comment garantir leur fonctionnement au gré des fluctuations multiples que peuvent subir les réseaux, avec toutes les possibilités de "déviante artificielle" que l'on peut imaginer (faut-il des agents policiers ?). Les concepts de normes, rôles, valeurs, culture, sanction, peuvent avoir un sens ; mais on est fondé à s'interroger sur leur validité stricte pour un système clos artificiel.

Le cas des agents assistants est différent car il s'agit de concevoir l'agent pour qu'il assure des fonctions particulières relativement à un utilisateur déterminé. Ils doivent le satisfaire, c'est à dire se conformer à ses valeurs et sa culture. Les sous-modèles concernés sont le 2 et le 3. Dans le cas de l'IHS, les contraintes de l'ergonomie cognitive font que l'on doit envisager une articulation entre la culture artificielle de l'agent (les modèles conceptuels qu'il utilise) et la culture réelle de l'acteur. La culture artificielle de l'agent peut être soit posée à priori par le concepteur, qui peut être différent de l'utilisateur futur, ce qui peut engendrer des discordances ; soit elle peut se construire par apprentissage au contact de l'utilisateur, et c'est la tendance que l'on observe dans de nombreux travaux ("Training a Personal Digital Assistant", p. 32, in (Maes, 1994)). Si par ailleurs l'agent doit organiser ou participer indirectement à des interactions entre acteurs, alors, afin de respecter les schémas classiques de communication "directe", il doit intégrer au niveau du sous-modèle 3 une représentation du système social réel. En particulier, dans une perspective de production, l'agent devra intégrer l'organisation et les positions respectives des agents avec lesquels il communique.

Le sous-modèle 4, suggéré par John Stewart, a déjà été considéré par ailleurs (Ferrand, 1995b ; Ferrand, 1996), et nous ne le détaillerons pas ici. On retiendra seulement qu'il supporte deux enjeux : d'une part les effets potentiels de l'instrumentation du lien social (Rheingold, 1993) et la restructuration induite sur les sociétés ou les organisations, et d'autre part l'inscription ("contextualisation") des systèmes dans un cadre social rendant l'information contingente.

Pour conclure sur cette question de la validité du modèle social humain pour les systèmes collectifs artificiels, nous nous référons à la conclusion ("The Natural and the Artificial") de (Chatila, 1994) qui énonce clairement que la nature résout des problèmes en adaptant des solutions existantes à de nouvelles situations, c'est à dire par *réaction*, alors que l'ingénieur humain peut *prévoir et planifier*. La référence aux sociétés humaines relève de la même remarque : il n'est sans doute pas indispensable de prendre la sociologie pour référence dans la conception des systèmes collectifs artificiels, si ceux-ci sont destinés à évoluer dans un milieu très différent du cadre d'évolution de l'espèce humaine. Cependant, si ils doivent interagir avec des humains, et assurer des fonctions relatives aux sociétés humaines (gestion d'agendas, aide à la négociation, médiation, etc.), cette référence s'impose.

5. Conclusion et Travaux ultérieurs

Le travail et la réflexion entamés ici ne constituent qu'une amorce de discussion. Il doit être discuté, commenté, amendé, et il faut y adjoindre de nombreux concepts sociologiques qui n'ont pas été inclus. En particulier la sociologie des organisations et la sociologie du changement n'ont été que très partiellement abordé. Nous souhaitons donc poursuivre afin de parvenir peut-être à distinguer clairement ce qui est social de ce qui ne l'est pas ; et à en inférer des principes de construction des systèmes collectifs artificiels qui ne s'inspirent des sociétés humaines qu'à bon escient, et qui pour tous les autres cas soient davantage fondés sur l'analyse des systèmes complexes non symboliques (physiques).

Pour préciser concrètement les points abordés dans cette article nous proposons aussi de développer un système collectif artificiel très simple qui permettrait d'illustrer les différents déterminants sociaux. Nous pensons pour cela poursuivre des travaux antérieurs sur un système de circulation d'agents dans un environnement structuré contraignant, A l'inverse, nous avons aussi entamé un travail de simulation sociale qui s'appuie intégralement sur un corpus sociologique portant sur la diffusion des comportements sexuels dans les réseaux sociaux. La simulation doit être

validée par application "d'enquêtes virtuelles" permettant la comparaison avec les données statistiques disponibles (échantillon de 5000 personnes). Elle doit permettre de tester un certain nombre d'hypothèses relationnelles.

Remerciements

Cette réflexion trouve certainement son origine intellectuelle (et intrinsèque) dans des discussions prolongées avec Alexis Ferrand (Institut de Sociologie, Lille 1) et Dan Ferrand-Bechmann (Université Paris - Saint-Denis). Qu'ils en soient remerciés, comme pour tout le reste.

Les relecteurs de ce texte ont contribué à son élaboration et sa maturation, et en particulier John Stewart qui a permis la réécriture de la partie 4.

Par ailleurs les apports multiples de la communauté des systèmes multi-agents, et en particulier du groupe MAGMA de Yves Demazeau, sont évidemment déterminants.

Enfin Paul Bourguin et Eric Dedieu, en des temps différents, y ont aussi contribué.

Références

- Aron, R., 1967, *Les étapes de la pensée sociologique*, Paris : Gallimard
- Baeijs, C., & Demazeau, Y., 1995, *Les organisations dans les systèmes multi-agents*, IVème journées du GDR- PRC IA, SMA, Toulouse
- Bernoux, P., 1985, *La sociologie des organisations*, Paris : Points - Seuil
- Boudon, R., & Bourricaud, F., 1982, *Dictionnaire critique de la sociologie*, Paris : PUF
- Bourdieu, P., Chamboredon, J.-C., et Passeron, J.-C., 1968, *Le métier de sociologue*, Paris : Mouton-Bordas.
- Casti, J.L., 1994, *Complexification - Explaining a paradoxical world through the science of surprise*. New York : HarperCollins Pub.
- Chatila, R., 1994, Control Architecture for Autonomous Mobile Robots, in *Proceedings of PERAC Conference 1994*, pp 254-265, Lausanne : IEEE Comp. Soc. Press.
- Crozier, M., & Friedberg, 1977, *L'acteur et le Système*, Paris : Seuil
- Demazeau, Y., 1993, *La plate-forme PACO et ses applications*, 2èmes journées nationales du PRCIA - SMA, Montpellier - PRCIA CNRS.
- Demazeau, Y., 1995, *From interactions to collective behaviour in agent-based systems*, European Conference on cognitive sciences, St-Malo.
- Demazeau, Y., Muller, J.-P., 1991, From Reactive to Intentional Agents, in *Decentralized Artificial Intelligence II*. Y. Demazeau & J.-P. Muller, eds, North-Holland, Amsterdam : Elsevier. pp. 3-10.
- Encyclopédie Universalis, 1985
- Ferrand, N., 1995,
- Ferrand, N., 1995b, Des systèmes de Traitement de l'information pour la gestion de l'environnement et du territoire : vers une éthique et un nouveau paradigme téléonomique d'intégration du contexte social et des milieux. Proposition de *contribution au colloque "Maîtrise du long terme et démocratie"*, CPVS, Paris : Ministère de l'équipement.
- Ferrand, N., 1996, Modelling and supporting Multi-Actor Spatial Planning using multi-agents systems, in *3rd. NCGIA Conference on environmental modelling and GIS*, Santa Fe : NCGIA
- Ferber, J., 1995, *Les Systèmes Multi-Agents - Vers une Intelligence Collective*. Paris : InterEditions
- Ferber, J., 1995b, Modélisation multi-agents : un outil d'aide à la simulation de phénomènes complexes, in *Lettre de l'Environnement - Revue du CNRS*, Paris : CNRS.
- Forrest, S., ed., 1991, *Emergent Computation*, MIT Press, Cambridge : Elsevier.
- Heudin, J.-C., 1994, *La Vie Artificielle*. Paris : Hermès.
- Lashkari, Y., & al., 1994, Collaborative Interface Agents, in *Proceedings of the 12th. national Conference on Artificial Intelligence*, vol. 1, pp. 444-449, Seattle : AAI Press.

- Maes, P., 1994, Agents that reduce work and information overload, in *Communications of the ACM*, vol. 37, n° 7, pp. 31-40.
- Mendras, H., 1979, *Eléments de Sociologie*, Paris : Colin
- Morin, E., 1977, *La Methode - 1. La Nature de la Nature* - Paris : Seuil
- PLEIAD, 1992, *Vers une taxinomie de vocabulaire pour les Systèmes Multi-Agents*. 1ères journées du PRCIA - SMA, Nancy : CNRS.
- Rheingold, H., 1993, *The virtual community - Homesteading on the electronic frontier*. Addison - Wesley.
- Rocher, G., 1968, *Introduction à la sociologie générale*, 3 vols., Paris : Points - Seuil
- Ruelle, D., 1991, *Hasard et Chaos*, Paris : Ed. Odile Jacob

Mécanismes empathiques et complexité sociale : autour de J.P. Dupuy

Guillaume Deffuant
Laboratoire d'ingénierie appliquée aux systèmes complexes
Cemagref
Parc de Tourvoie
92 160 Antony France

1 Introduction

Nous proposons d'aborder le thème "du collectif au social" par une revue (partielle) de plusieurs modèles de mécanismes d'interactions entre les individus censés rendre compte de comportements collectifs ou sociaux. Nous nous appuyerons particulièrement sur la présentation que fait J.P. Dupuy de ces mécanismes dans plusieurs de ses ouvrages [6][7]. Une telle revue, malgré son caractère partial et partiel nous paraît utile pour fournir quelques points de repères dans la tradition, permettant de donner du sens à cette distinction entre collectif et social.

Nous étudions les approches suivantes : l'individualisme méthodologique, les mécanismes de sympathie proposés par A. Smith, le couple amour propre / amour de soi proposé par Rousseau, les mécanismes du désir mimétique [9][10] et enfin les mécanismes spéculaires [14]. Nous ne pouvons qu'évoquer de manière assez caricaturale chacun de ces mécanismes, en nous attachant à préciser leurs points communs et leurs différences.

Pour ce faire, nous proposons trois points de vue essentiels qui sont liés à différentes visions de la complexité. En effet, il nous semble qu'une distinction féconde entre collectif et social pourrait se référer à la complexité des relations et des organisations globales observées. Nous commençons par décrire le point de vue de J.P. Dupuy qui fait principalement référence à la complexité au sens de Von Neuman et à la notion de hiérarchie enchevêtrée (Varela, Maturana [13]). Puis nous nous référons à C. Castoriadis [1] [2], selon lequel les relations sociales ne peuvent être comprises entièrement dans une logique de type mathématique habituelle (ensembliste identitaire). Il s'agit là d'une complexité présentant des points communs avec la conception proposée par E. Morin [12]. La compréhension de cette complexité passe par la reconnaissance des limites intrinsèques à une approche scientifique, à laquelle échappe une grande part de ce qui fait sens pour l'homme.

Enfin, nous considérons la complexité des mécanismes cognitifs individuels sous-jacents à chacun des mécanismes sociaux proposés. Quel niveau de simplification est-il acceptable ?

Ces points de vues nous amènent à prendre parti dans la discussion et à proposer des mécanismes simples, inspirés des exemples étudiés, et qui nous paraissent susceptibles de se plier à une formalisation. Nous garderons également les travaux de L. Dumont [3] [4] comme guide dans cette revue des interactions sociales. Ces travaux, notamment sur l'individualisme et l'émancipation de l'économie, nous paraissent en effet particulièrement pertinents pour le thème que nous abordons.

2 Quelques mécanismes d'interactions sociales

Nous décrivons rapidement les différents mécanismes d'interactions sociales que nous avons relevés dans les travaux de J.P. Dupuy[6] [7].

2.1 L'individualisme méthodologique

L'individualisme méthodologique est une figure familière de notre temps. Elle est issue des travaux fondateurs de l'économie moderne, dont les grands noms sont Walras et Pareto. L'homme y est vu séparé de ses semblables, confronté à des choix matériels par lesquels il doit optimiser son bien-être matériel. Son comportement est donc raisonnable, il consiste à choisir les actions qui pour lesquelles il anticipe les récompenses matérielles les plus importantes. Ces choix ne sont faits qu'en fonction des avantages que l'individu peut en tirer. Il s'agit donc d'une position d'égoïsme total, pour laquelle le regard des autres n'est d'aucune importance, il est même inexistant.

Ce comportement égoïste de l'individu permet une amélioration globale de la situation matérielle de tous, car l'accumulation de richesse des riches les amènent à dépenser de plus en plus, et donc à redistribuer leurs richesses. L'égoïsme des riches profite ainsi également aux pauvres par le mécanisme de la "main invisible". Cette thèse est proche de celle de la fable des abeilles de Mandeville : "vices privés, bénéfices publics", dont A. Smith aurait développé la théorie.

Nous verrons que cette interprétation de la pensée de A. Smith est fortement contestée par J.P. Dupuy.

2.2 Le désir mimétique

Le mécanisme de "désir mimétique" constitue l'élément majeur des thèses de R. Girard. Ce mécanisme serait une incontournable loi de l'être selon laquelle on ne peut désirer que ce que l'autre désire. L'autre est donc constamment un rival (car il devient un concurrent) et un modèle (car on imite son désir). R. Girard a longuement illustré et argumenté ses thèses

par l'interprétation de textes littéraires, ainsi que par de nombreux mythes [9][10]. Ces mécanismes lui ont également fourni des interprétations riches des périodes de crises, dans lesquelles les désirs tendent vers l'indifférencié, et où la cohésion sociale est menacée. Le phénomène du "bouc émissaire", permettant d'extérioriser le désir et de rendre sa cohésion à la société s'explique de manière particulièrement élégante dans ce cadre. Cette approche permet à R. Girard de déceler des caractères spécifiques de la culture judéo-chrétienne, qui avoue au grand jour l'innocence des victimes émissaires, contrairement aux autres cultures qui se fondent sur la croyance de leur culpabilité.

2.3 La spécularité

J.P. Dupuy signale l'importance des mécanismes spéculaires, notamment identifiés par J.M. Keynes dans ses modèles de la spéculation financière [11]. On peut également lier la spécularité à la théorie des jeux. Selon J.L. Vullierme [14], la spécularité correspond essentiellement à la représentation des représentations des autres, en particulier la représentation qu'ils ont de nous-mêmes. Il s'agit par exemple, dans la théorie des jeux, de tenir compte de l'anticipation de l'autre sur nos propres coups. Il est clair que ces mécanismes, comme le désir mimétique et l'empathie, peuvent provoquer des croisements et (potentiellement) des régressions à l'infini dans les interactions entre plusieurs personnes.

2.4 La sympathie (empathie) de A. Smith

Dans ce qu'il considère comme son ouvrage majeur, intitulé "théorie des sentiments moraux", A. Smith soutient que le mécanisme fondamental des comportements sociaux humains est la sympathie. Comme le précise bien J.P. Dupuy, la sympathie ne doit pas être confondue avec la bienveillance. Il s'agit plutôt de la capacité qu'ont les hommes à s'identifier par imagination à leurs semblables, pour vivre à leur place les situations auxquelles ils sont confrontés. Il est important de remarquer que cette identification n'est en général pas totale, c'est-à-dire que le sujet est à la fois lui-même et l'autre en même temps, et n'induit pas forcément une attitude bienveillante (cela peut même être souvent le contraire, nous y reviendrons). Pour éviter cette confusion avec la bienveillance, nous utiliserons le terme empathie pour désigner la sympathie dans le sens de A. Smith. Dupuy fait remarquer que cette conception garde un caractère moderne, puisque l'individu reste séparé de ses semblables, et ne dispose que de son imagination pour se mettre à leur place.

Les mécanismes d'empathie induisent des phénomènes aux dynamiques complexes. Les dynamiques croisées sont celles qui apparaissent le plus naturellement. Le plaisir n'est plus directement la jouissance procurée par

les objets, mais la satisfaction de ressentir la considération des autres, en raison de la possession de cet objet. Or le sentiment de la considération des autres est engendrée par des mécanismes croisés d'empathie. En effet, il s'agit du plaisir à s'identifier à celui qui prend plaisir à s'identifier à soi. Les rapports économiques ne se distingueraient donc pas des autres rapports humains. L'objectif reste l'obtention de la considération des autres, c'est à dire d'imaginer le plaisir qu'ils ont à s'identifier à soi.

J.P. Dupuy argumente ces propos en insistant sur le peu de cas que faisait A. Smith des biens matériels. Selon lui, le bonheur et le plaisir individuels peuvent être facilement comblés par un minimum de biens matériels, et leur accumulation n'est d'aucun intérêt. Leur valeur vient donc uniquement de la considération qu'ils apportent à leur possesseur. Le comportement de l'individu Smithsien a donc son origine dans une interaction avec ses semblables, qui façonne son interaction avec les objets. Nous nous trouvons donc en contradiction complète avec l'individualisme méthodologique (l'individualisme moderne décrit par L. Dumont), pour lequel la hiérarchie est inverse : ce sont les interactions des hommes avec les objets qui façonnent leur interactions entre eux.

2.5 L'amour de soi et l'amour propre (J.J. Rousseau)

Rousseau propose un double principe de fonctionnement là où Smith n'en a besoin que d'un. Le premier de ces deux mécanismes est l'amour de soi, qui guide totalement l'homme à l'état de nature. Cet homme à l'état de nature, figure théorique artificielle, a des points communs avec l'individu moderne au sens de Dumont : il est totalement séparé de ses semblables, et ne prend ses décisions qu'en fonction de ce qu'il estime bon pour lui-même.

Le deuxième principe de fonctionnement est l'amour propre qui tend à prendre le pas sur le premier dès que l'homme se trouve en société. L'amour propre se manifeste dès que l'individu se compare à ses semblables, et transforme autrui en concurrent et obstacle. Le désir de l'objet se polarise en sentiment violent envers cet obstacle, ce qui peut mener à des situations où l'on préfère détruire l'objet plutôt que de supporter la comparaison désavantageuse entre soi et l'autre.

L'objectif de Rousseau est de construire une organisation sociale tendant à un retour de la prédominance de l'amour de soi sur l'amour propre, afin d'éviter cette violence. Il proposera une solution avec le schéma du contrat social, qui dépasse le cadre de cet article.

Trois de ces mécanismes (empathie, specularité, désir mimétique) sont considérés par J.P. Dupuy comme équivalents à un certain niveau d'analyse, les deux autres étant à part. La section suivante explique cette classification, et les autres poursuivent l'analyse.

3 La complexité de Von Neuman et construction de hiérarchie enchevêtrée

Le premier point de vue sur la complexité que nous proposons est étroitement inspiré par J.P. Dupuy. Ce point de vue contient deux approches de la complexité : les hiérarchies enchevêtrées et la complexité au sens de Von Neuman. Cette distinction entre deux types de complexité nous permet déjà une première distinction parmi nos mécanismes d'interactions sociales, et de donner un premier sens à la distinction entre collectif et social.

3.1 Présentation des deux types de complexité

3.1.1 Complexité de Von Neuman

J.P. Dupuy insiste beaucoup, notamment dans "A l'origine des sciences cognitives" [5], sur la complexité au sens de Von Neuman. Cette idée de la complexité a été initialement proposée en évoquant la machine de Turing et le théorème de Gödel. Voilà en effet une machine dont les règles de fonctionnement sont extrêmement simples et faciles à décrire, alors que ses comportements possibles sont infiniment plus difficiles à décrire (puisque dans certains cas, il on ne peut prévoir si elle s'arrêtera ou non). A partir de cet exemple, Von Neuman propose une définition de la complexité dont l'esprit est le suivant : pour une machine non complexe, il est plus simple de décrire ce qu'elle produit que ses règles de fonctionnement, alors que pour une machine complexe, c'est l'inverse, les productions sont infiniment plus difficiles à décrire que ses principes de fonctionnement.

En considérant l'organisation sociale comme un système complexe dans ce sens, il ressort que cette organisation reste toujours en partie inaccessible aux individus qui la produisent. En effet, les règles d'interaction entre les individus peuvent être simples et accessibles à tous, sans que le résultat global de ces interactions le soit.

3.1.2 Hiérarchie enchevêtrée

Un autre point de vue adopté par J.P. Dupuy dans la caractérisation des mécanismes de fonctionnement sociaux est leur capacité à engendrer des hiérarchies enchevêtrées. De telles hiérarchies se caractérisent par les propriétés suivantes : le niveau individuel contribue à la constitution du niveau collectif, qui échappe dans une large mesure au niveau individuel et rétroagit sur les dynamiques individuelles. On assiste donc à une double indétermination, tant au niveau individuel qu'au niveau collectif, qui se codéterminent mutuellement dans une danse imprévisible. Ainsi, dans une telle perspective, les individus sociaux participent à l'institution des règles sociales et sont en même temps déterminés par elles. Leur comportement est donc initialement en

grande partie indéterminé. Ce comportement se détermine dans leur interaction avec le tout social qui leur préexiste et les dépasse, mais qu'ils contribuent à modifier, et qui les modifie en retour. Ce type de configuration, comportant une circularité intrinsèque est à rapprocher des systèmes autopoïétiques ou opérationnellement clos [13] et semble souvent considéré par Dupuy comme prototypique de la complexité.

Ces deux approches de la complexité semblent être parfois mêlées dans certains textes de Dupuy, alors qu'il nous semble important de bien les différencier.

Observons maintenant nos mécanismes d'interactions sociales à la lumière de ces concepts.

3.2 L'individualisme méthodologique : uniquement la complexité de Von Neuman

L'individualisme méthodologique se distingue des autres approches considérées. En effet, l'individualisme méthodologique peut engendrer une complexité de type Von-Neuman, et c'est là ce qui sauve cette tradition aux yeux de Dupuy (Castoriadis au contraire prononce une condamnation sans appel en qualifiant l'individualisme méthodologique d' "infra-débile"). Des règles individuelles simples, de type optimisation rationnelle des choix, peuvent donner des effets globaux inattendus, opaques aux individus. C'est ainsi que les égoïsmes individuels vont se transformer en une maximisation globale du bien-être (la main invisible). Ainsi la société est-elle le produit des actions des hommes, mais pas de leurs desseins. On obtient donc un schéma de fonctionnement proche de celui de la complexité de Von Neuman, pour lequel des règles de fonctionnement individuelles simples produisent des comportements globaux qui échappent à ces règles simples.

En revanche, le schéma de la complexité de type Varela s'applique mal. En effet, le comportement des individus est totalement spécifié initialement, et les équilibres globaux ne modifient que les objets disponibles pour leurs choix, pas la manière de faire leurs choix.

3.3 Les autres mécanismes : construction de hiérarchie enchevêtrée

Dupuy montre très bien que les mécanismes d'empathie, de mimétisme ou de specularité donnent lieu à des dynamiques morphogénétiques de hiérarchies enchevêtrées. En effet, dans les trois cas, on observe une indétermination initiale, à la fois du niveau global et du niveau individuel. Ces différents niveaux se cospécifient dans la dynamique, donnant lieu à différents états (stables ou instables) possibles (même si A. Smith n'avait pas vu cette richesse et cette indétermination). En effet, selon ces différents mécanismes d'interactions, les mécanismes de choix des individus sont largement indéterminés au départ, et se spécifient dans l'interaction avec les autres, par mimétisme du désir, par

identification partielle (empathie), ou en tenant compte de la représentation des représentations que se font les autres (spécularité). C'est ainsi qu'on peut retrouver des positions plus "sociologiques", dans lesquelles la société comme un tout détermine en grande partie les règles de comportement des individus, même si ce sont ces mêmes individus qui produisent cette société. On voit donc que de ce point de vue, les différences entre ces mécanismes peuvent être négligées, et c'est pourquoi J.P. Dupuy les traite parfois comme équivalents.

Le cas de Rousseau paraît moins clair, et Dupuy semble regretter une complication inopportune du double principe (amour de soi et amour propre), alors que les autres modèles n'en nécessitent qu'un seul. Le lien entre l'empathie de Smith et l'amour propre est suggéré, mais pas développé totalement.

3.4 Proposition d'un sens pour la distinction entre collectif et social

Les remarques précédentes nous amènent à proposer un premier sens à la distinction entre collectif et social : le collectif serait ce qui ressort de mécanismes complexes de type Von Neuman, le terme social ne serait acceptable que lorsque des hiérarchies enchevêtrées sont produites. Dans ce sens, la complexité du social est évidemment plus grande que celle du collectif. Cette distinction se justifie de notre point de vue car les interactions de type spéculaire, empathiques ou mimétiques nous paraissent rendre plus finement compte de la réalité des rapports humains, alors que le mécanisme d'interaction sociale de l'individualisme méthodologique est directement inspiré de la physique qui spécifie le comportement des objets par la minimisation d'une grandeur physique.

Cependant, ce premier sens permet de distinguer entre l'individualisme méthodologique et les autres mécanismes, mais pas entre ces derniers. Sont-ils totalement équivalents ? Il nous semble intéressant de poursuivre l'analyse, à la lumière d'autres approches de la complexité. Peut-être pourrions-nous préciser encore le sens d'une distinction entre collectif et social.

4 Mécanismes sociaux et logique ensembliste identitaire

Nous regardons maintenant nos mécanismes d'interactions sociales à la lumière de la logique ensembliste identitaire, afin de mieux appréhender leurs propriétés, et si possible de préciser notre distinction entre collectif et social.

4.1 Présentation de la logique ensembliste identitaire

La logique ensembliste identitaire (abrégée par le terme ensidique) est introduite par C. Castoriadis pour caractériser les raisonnements à l'oeuvre dans la théorie des ensembles, et par extension dans l'ensemble des mathématiques. Dans ce type de logique, les objets ont une identité déterminée, qui ne souffre aucun doute : un objet est soit égal à un autre, soit différent, aucune troisième voie n'est possible.

La logique ensembliste identitaire s'oppose à des logiques oniriques, typiques de l'inconscient freudien, dans lesquelles l'identité des objets n'est pas fixée, où deux objets peuvent à la fois être distincts et identiques, séparés et liés [1]. Cette distinction se rapproche de la conception de la complexité selon E. Morin [12], qui consiste à prendre acte des limites d'un discours scientifique qui occulte des pans entiers de notre vie psychique et sociale. Notons que cette distinction n'a rien à voir avec les distinctions entre discret et continu, déterministe ou probabiliste, qui pour Castoriadis prennent toutes place dans le cadre de la logique ensidique.

Il n'est pas étonnant que la logique ensidique ait conquis totalement les discours dominants des sociétés modernes. Ces discours dominants, centrés sur les activités marchandes et scientifiques, se caractérisent par des relations aux objets qui semblent mieux s'accorder des mutilations imposées par la logique ensidique.

Cependant, si nous suivons Castoriadis et Morin, les mécanismes sociaux, s'ils veulent rendre compte de la richesse des comportements humains, ne peuvent se cantonner à la logique ensidique. En effet, selon ces penseurs, les explications uniquement ensidiques du social sont obligatoirement mutilantes, car des parties essentielles de la psychologie et du social leur échappent. Il est donc particulièrement intéressant d'observer nos mécanismes à la lumière de cet éclairage.

4.2 Individualisme méthodologique

Pour l'individualisme méthodologique, la conclusion est immédiate : chaque individu est séparé des autres, et n'a d'interaction avec eux que par l'intermédiaire de signaux sur les choses, nous nous trouvons donc dans un discours en parfait accord avec la logique ensidique puisque les individus sont parfaitement séparés les uns des autres et que leur identité ne fait aucun doute. Cela n'est pas étonnant pour un discours qui est au coeur des sociétés marchandes modernes.

4.3 Spécularité

Le statut de la spécularité par rapport à la logique ensidique est plus difficile. Chacun se fait des représentations des autres et de leurs intérêts, incluant

éventuellement des représentations de soi-même. Ces représentations se distinguent clairement des individus ou des objets qu'elles représentent, et donnent lieu à des raisonnements permettant de maximiser un intérêt, comme on le constate dans la théorie des jeux. Le fait que des boucles infinies de spécularité puissent donner des figures originales ne contredit en rien la logique ensidique. Il est clair que les choix d'un individu vont dépendre des représentations qu'il se fait des autres et de leurs propres représentations (y compris et surtout de celles qu'ils se font de lui-même), mais on peut très bien considérer des individus rationnels, qui maximisent leur intérêt en fonction de ces représentations qui sont simplement un peu plus complexes que dans le cas de l'individualisme méthodologique.

Le point important ici est de savoir comment se définit le sujet qui maximise cet intérêt. S'il est défini en dehors des représentations, donc comme spectateur de ces représentations dans un théâtre cartésien, ce sujet est parfaitement défini et ne contredit en rien la logique ensidique (mais d'autres problèmes de type homuncule surgissent). On peut soupçonner que cette conception néocartésienne du sujet est sous-jacente dans les discours sur la spécularité.

4.4 Désir mimétique

Les difficultés sont du même ordre pour le désir mimétique. En effet, le mimétisme implique l'observation de l'autre et la reproduction de son attitude. Quelle est l'influence de cette reproduction sur l'identité des personnes. Qui est qui ? En fait, le mimétisme nous semble éluder cette question, en laissant dans la possibilité de rester en accord avec la logique ensidique. Le fait de copier le désir de l'autre implique que l'on se transforme au contact de l'autre, ce qui ne contredit pas la logique ensidique. Il y a simplement le postulat d'une très grande importance de l'influence de l'autre.

4.5 Empathie

l'empathie de Smith nous semble en revanche, à un certain niveau, en contradiction avec la logique ensidique. En effet, par définition de l'empathie, l'individu est à la fois lui-même et l'autre (ou ce qu'il imagine de l'autre), l'autre qui est aussi à la fois lui-même et l'autre, et ce éventuellement avec des régressions à l'infini. Ceci nous semble être une différence très importante par rapport au mimétisme et la spécularité, car l'empathie questionne l'identité du sujet au plus profond, en la posant comme multiple et indéterminée, alors que le mimétisme et la spécularité permettent encore de garder cette identité en conformité avec la logique ensidique.

4.6 Proposition pour la distinction entre collectif et social

Ce détour par la logique ensidique nous a permis de mieux comprendre les propriétés et les relations entre les différents mécanismes sélectionnés par notre premier critère de complexité (hiérarchie enchevêtrée). Les mécanismes empathiques, qui posent des sujets potentiellement multiples et indéterminés, contradictoires, se rapprochent plus de la complexité dont parlent Castoriadis et Morin en s'écartant nettement de la logique ensidique. De plus, ils nous paraissent être l'expression achevée des mécanismes mimétiques ou spéculaires qui ne font que frôler les limites de la logique ensidique. En effet, on peut considérer que le désir mimétique ou la spécularité sont des manifestations des mécanismes empathiques : la propension à l'identification à l'autre inclut la propension à s'approprier les désirs, les craintes, et les intentions de l'autre, qui deviennent des représentations dans un cadre représentationnaliste adopté par certains discours sur la spécularité.

Notre premier critère de complexité (hiérarchie enchevêtrée) nous a permis de faire une distinction entre l'individualisme méthodologique et les autres mécanismes. Le second critère de complexité (non conformité avec la logique ensidique) nous permet de distinguer les mécanismes empathiques des autres mécanismes, qui apparaissent comme des manifestations partielles de l'empathie.

Nous avons laissé de côté le couple amour de soi / amour propre qui mérite une attention particulière. Le fait d'avoir affaire à un double mécanisme introduit une difficulté nouvelle que nous traitons dans la section suivante. Dupuy nous semble passer un peu rapidement sur cette difficulté, en soulignant la plus grande simplicité des mécanismes d'empathie ou de mimétisme. Nous proposons au contraire de nous y attarder un peu.

5 Mécanisme unique ou multiple ?

5.1 De la nature bicamérale de l'esprit

Réduire les mécanismes sociaux à un mécanisme unique n'est-il pas excessivement réducteur ? Peut-on espérer qu'une réduction aussi drastique de la complexité psychologique soit acceptable ?

On peut en douter et préférer un mécanisme cognitif double, fondé d'une part sur les interactions sensori-motrices directes, et d'autre part sur les interactions sociales. En effet, de nombreuses situations romanesques sont issues d'une opposition entre les deux aspects. Citons par exemple le personnage de Hulot dans "la cousine Bette" de Balzac. Ce personnage ruine sa réputation et sa famille à cause de sa passion immodérée pour les jeunes filles, qui présente le facteur aggravant de s'accroître avec l'âge. On ne voit pas en quoi ce désir pourrait provenir d'une imitation du désir d'un autre. Il s'agit d'un désir primaire, direct, qui contredit les intérêts sociaux

du personnage. Le romanesque naît de cette tension entre les deux types de désir. Mais on doit à Proust, avec ses observations sur les mécanismes de la mémoire, la mise en évidence la plus complète de l'existence d'un fonctionnement cognitif directement sensori-moteur, qui court-circuite la diversité des moi sociaux. La "madeleine" en est l'exemple le plus célèbre. Nos analyses sont donc ici en opposition très claire avec celles de Girard, pour lequel ces aspects sensori-moteurs jouent un rôle négligeable. Cette position nous paraît cependant suffisamment solide pour être défendue.

En effet, nous la retrouvons également dans de nombreux travaux scientifiques ou philosophiques, que nous ne pouvons qu'évoquer brièvement. Elle peut se rapprocher de la distinction freudienne entre libido et identification (comme le signale Dupuy). Castoriadis la reprend à son compte dans sa distinction entre strates, et notamment entre la psyché et le social-historique. Plus récemment, à partir de considérations neuro-biologiques, Edelman [8] propose un modèle double de la conscience, mettant en scène d'une part une "conscience primaire" fondée sur les aires sensori-motrices du cerveau, et une "conscience d'ordre supérieur", centrée sur les aires du langage.

Cette nature "bicamérale" de l'esprit nous paraît donc être une idée forte, qui, dans une perspective que nous allons développer, n'est pas en contradiction avec le point de vue d'A. Smith, ni avec celui de J.P. Dupuy. En effet, chez ces deux auteurs, le mécanisme empathique ne fonctionne pas seul. Les figures d'empathies croisées que nous avons décrites sont en interaction avec des sentiments de désir (crainte), de plaisir (douleur) (plaisir empathique du plaisir de l'autre). Cette empathie est en général partielle, et des désirs ou craintes non empathiques peuvent rester influentes. D'où peuvent venir ces désirs ou craintes si ce n'est d'un mécanisme biologiquement contraint, qui initie les processus empathiques ? Les mécanismes cognitifs sociaux (fondés sur l'empathie selon notre analyse) peuvent transformer totalement ces désirs, les inhiber pour en introduire d'autres (imités), les sublimer en provoquant des régressions à l'infini comme nous le verrons au paragraphe suivant, etc... Mais il nous semble que l'empathie ne peut fonctionner sans un niveau de désir ou de répulsion directs (sensori-moteurs), qui sert d'"étayage" (selon une expression de Castoriadis) aux divers niveaux empathiques.

5.2 Amour de soi et désirs biologiques

On peut déceler de troublantes ressemblances entre l'homme à l'état de nature, dominé par l'amour de soi, et la conscience primaire d'Edelman. En effet cette conscience primaire est essentiellement commandée par des désirs biologiques, transmis génétiquement (principalement sexuels et alimentaires). La mémoire primaire s'organise en un réseau de mémorisations d'expériences plus ou moins agréables, qui vont se superposer et modifier la perception en associant des anticipations de plaisir ou de douleurs à diverses

classes de situations. Il s'agit d'un comportement maximisant le plaisir, cette fonction de plaisir étant en constante évolution.

La description de l'homme à l'état de nature proposée par Rousseau semble en concordance avec ce type de comportements. En effet, la principale caractéristique de l'homme à l'état de nature, est qu'il est totalement indépendant de ses semblables. On peut d'ailleurs voir des ressemblances importantes avec l'individu moderne, vu par l'individualisme méthodologique. De ce point de vue, l'individualisme moderne aurait une composante de régression vers la conscience primaire, c'est à dire vers une négation de ce que l'homme peut avoir d'humain. En effet, selon l'identification des circuits biologiques impliqués dans la conscience primaire que propose Edelman, tous les animaux qui possèdent un cortex doivent posséder une conscience primaire, et c'est le cas de tous les mammifères. La spécificité humaine serait due aux circuits neuronaux du langage. A considérer certains aspects de l'histoire de l'époque moderne il peut apparaître que cette régression n'est pas restée à l'état de figure théorique, mais a réellement été expérimentée à grande échelle.

5.3 Amour propre, self-love et empathie

Selon Edelman, la conscience d'ordre supérieur, liée aux aires du langage est spécifique de la cognition humaine. Il s'agit donc de la partie sociale de la cognition, qui selon notre analyse serait principalement due à des mécanismes empathiques. Les liens entre le langage et la capacité d'empathie sont de toute évidence très forts. Nous ne tenterons pas de développer ce point qui nous entraînerait trop loin. Il nous suffira de signaler que le langage permet de communiquer des signaux d'approbation ou de désapprobation sociales de manière très détaillée et précise. Si l'on considère que la dynamique de base de l'individu social reste celle de la maximisation de son plaisir social, il s'agit de la maximisation de l'approbation des personnes qui comptent pour lui. Pour obtenir cette maximisation, l'individu est obligé d'anticiper les réactions de ces personnes, afin de prévoir les comportements qui recueilleront leur approbation. C'est ici que se fait le lien avec l'empathie : l'anticipation des réactions des autres impose le développement de mécanismes empathiques.

Le lien avec l'amour propre Rousseauiste est-il possible ? Dupuy le suggère et un examen un peu détaillé le révèle pleinement. En effet, selon Rousseau, la principale caractéristique de l'amour propre est d'engendrer l'envie, et de préférer détruire l'objet de nos convoitises plutôt que de voir un autre s'en emparer. Il y a dans ce comportement autre chose que la volonté de contourner un obstacle, qui peut se rencontrer aussi bien dans l'état de nature. Cet obstacle est d'un type particulier. Supposons que A et B convoitent l'objet O. Si A se rend compte que B convoite l'objet O, et qu'il s'identifie partiellement à B, alors son désir se renforce de la modélisation du

désir de B (et symétriquement pour B). Ce mécanisme peut se poursuivre en une régression à l'infini qui exacerbe le désir et peut donc amener la violence.

Cependant, l'amour propre se caractérise plus par les comparaisons entre les semblables, qui provoquent l'envie, que par cette exacerbation du désir. De telles comparaisons provoquent la violence car elles induisent le désir de rabaisser l'autre pour que la comparaison soit moins désavantageuse. Cette possibilité de comparaison se trouve également dans les mécanismes empathiques. En effet, il suffit pour cela d'introduire le plaisir provoqué par l'empathie. Considérons, comme le fait A. Smith, que l'important dans les mécanismes sociaux est la recherche du plaisir provoqué par l'identification des autres à soi (le self-love). Dès lors, la comparaison entre le plaisir de s'identifier à l'autre et le plaisir que l'autre prend à s'identifier à soi apparaît dès la mise en place d'une opération empathique double, impliquée par le self-love d'A. Smith. C'est la différence entre les deux plaisirs qui exacerbe la concurrence.

On peut pousser plus loin l'interprétation : la création d'un "moi social" passe par le regard de l'autre. La jouissance sociale ne peut donc être que la jouissance empathique d'un autre (éventuellement imaginaire). Le fait que A possède plus que B, et donc ne puisse pas avoir de plaisir à s'identifier à B implique que B ne peut pas jouir socialement de ce qu'il possède. Il doit donc déposséder A de ses biens pour pouvoir jouir des siens. On peut supposer que la régression à l'infini de l'empathie répète cette frustration indéfiniment et la rend totalement insupportable.

On retrouve alors les principes proches de ceux du désir mimétique, pour lesquels le désir (et le plaisir) ne peuvent être que ceux de l'autre. Le désir mimétique, ainsi que le self-love et l'amour propre apparaissent donc comme des figures de l'empathie.

La distinction entre mécanisme simple (empathie, mimétisme) et mécanisme double (amour de soi - amour propre) n'est probablement qu'apparente. En effet, dans les mécanismes simples, le désir biologique est toujours sous-jacent même s'il est très souvent détourné et modifié par les mécanismes sociaux. Le projecteur de la complexité psychologique, et des substrats biologiques associés, nous a fourni un nouveau regard sur nos mécanismes d'interactions sociales, et nous a permis d'analyser le double mécanisme proposé par Rousseau. Le moment est venu de faire un point sur l'ensemble de notre démarche, afin d'en bien faire ressortir les points saillants.

6 Synthèse

Nous proposons un modèle synthétique issu de cette étude, en nous appuyant sur les hypothèses suivantes :

- Les mécanismes sociaux doivent permettre la morphogénèse de hiérarchies enchevêtrées, ce qui exclut l'individualisme méthodologique.
- Un modèle de la cognition sociale ne doit pas respecter la logique ensidique, qui à l'évidence n'est pas pertinente pour ce qui concerne les rapports entre les hommes. L'abandon de la logique ensidique ne signifie pas l'abandon de toute formalisation, mais le développement d'une formalisation différente, dans laquelle les objets peuvent être à la fois différents et semblables. La recherche d'une telle formalisation, si elle est possible, nous paraît d'un enjeu très important. Une piste à explorer est celle de la mise au point d'une théorie des ensembles qui peuvent se contenir eux-mêmes à différents niveaux (et même une infinité de fois).
- Les mécanismes de comportements sociaux prennent racine dans des comportements biologiques, qui ne peuvent donc être totalement négligés, même s'ils peuvent être rendus méconnaissables par les mécanismes sociaux.
- L'amour propre, le self-love, le désir mimétique et la spécularité peuvent être vus comme des figures de l'empathie.

La proposition d'un modèle bicaméral, comportant une strate de désir biologique sous-tendant des mécanismes empathiques nous paraît donc être une bonne synthèse de notre revue. Nous pouvons donner quelques pistes de son utilisation :

- l'un des fondements de l'organisation sociale est de permettre une régulation des régressions à l'infini du désir (ou de la répulsion), ou même de l'opération empathique carrée, qui engendrent la violence. Un des moyens de bloquer la régression est de forcer l'identification à personnage particulier qui ne cherche pas lui à s'identifier (et donc qui bloque la régression). On reconnaît là le fonctionnement classique des divinités archaïques. Dans de telles sociétés, les individus au sens moderne n'existent pas, puisqu'ils ne cherchent pas à se voir dans le regard des autres.
- L'individu ne vient à exister que lorsqu'il se voit dans le regard d'un autre. Alors se pose l'éternel problème de la permanence de cet autre (même imaginaire). L'individu s'identifie donc toujours à un autre que lui-même qui le regarde pour lui permettre de se voir. Il s'agit là d'une différence importante avec la spécularité, qui considère que des représentations de soi-même sont directement accessibles. Dans notre cadre, ce n'est pas le cas. L'identification à soi-même nécessite l'intermédiaire de l'identification à une autre (même imaginaire), qui

s'identifie à soi-même. C'est cet autre imaginaire, souvent négligé, qui va conditionner l'apparition du moi social et le spécifier.

- La construction systématique de cet autre imaginaire chargé de faire être l'individu est quelque chose de récent et de singulier dans l'histoire de l'humanité, correspondant à l'émergence de l'individu moderne. Il serait intéressant de reprendre les analyses de L. Dumont sur l'influence du christianisme dans ce domaine. Reprendre dans notre perspective les analyses de R. Girard sur la question pourrait également s'avérer fertile.

7 Conclusion

Notre parcours, provoqué initialement par la distinction entre collectif et social, nous a permis de proposer une interprétation inspirée par J.P. Dupuy de cette distinction en terme de hiérarchie enchevêtrée. En effet, ce type de hiérarchie nous paraît rendre compte de manière plus adéquate des interactions entre liberté individuelle et détermination sociale. Par opposition, les effets collectifs peuvent être complexes, mais sont issus de comportements totalement spécifiés des agents.

Nous avons poussé l'étude au delà, en introduisant des critères supplémentaires sur nos modèles sociaux, comme la transgression de la logique ensidique, et le respect d'une organisation bicamérale de la cognition. Nous proposons une synthèse à partir de ces critères, qui nous paraît être mieux placée pour prétendre à la production d'organisations sociales (par opposition à collectives) : la composition d'une strate de désirs biologiques et d'une propension à l'empathie.

Cette synthèse nous semble ouvrir des perspectives intéressantes. En effet, bien que transgressant la logique ensidique (en considérant des ensembles qui se contiennent eux-mêmes par exemple), la mise en place d'un formalisme rigoureux, en relaxant certaines contraintes logiques habituelles, ne semble pas impossible. De plus cette synthèse nous semble pouvoir rendre compte de manière plus satisfaisante d'un certain nombre de phénomènes sociaux que ne le feraient les autres mécanismes évoqués.

References

- [1] Cornelius Castoriadis. *L'institution imaginaire de la société*. Le Seuil, 1975.
- [2] Cornelius Castoriadis. *Domaines de l'homme*. Le Seuil, 1986.
- [3] Louis Dumont. *Homo Aequalis - Genèse et épanouissement de l'idéologie économique*. Gallimard, 1977.

- [4] Louis Dumont. *Essais sur l'individualisme*. Le Seuil, 1983.
- [5] Jean-Pierre Dupuy. *Aux origines des sciences cognitives*. La découverte, 1992.
- [6] Jean-Pierre Dupuy. *Le sacrifice et l'envie*. Calmann-Lévy, 1992.
- [7] Jean-Pierre Dupuy. *Logique des phénomènes collectifs*. Ellipse, 1992.
- [8] Gerald M. Edelman. *Biologie de la conscience*. Odile Jacobs, 1992.
- [9] René Girard. *Mensonge romantique et vérité romanesque*. Grasset, 1961.
- [10] René Girard. *Le bouc émissaire*. Grasset, 1982.
- [11] John Maynard Keynes. *The General Theory of Employment, Interest and Money*. Cambridge University Press, 1936.
- [12] Edgar Morin. *La complexité humaine*. Flammarion, 1994.
- [13] Francisco Varela. *Principles of biological autonomy*. North Holland, 1979.
- [14] Jean-Louis Vullierme. Spéculation morale et spécularité anthropologique. *Individu et Justice sociale - autour de John Rawls*, pages 145–157, 1988.

Complexité individuelle et complexité sociale : antagonismes et complémentarités.

Marie-Noëlle Sarget

Ecole des Hautes Etudes en Sciences Sociales
25 Avenue Bolviller, 91.800, Brunoy.

Résumé

Si la notion de "collectif" renvoie à quelque chose de quantitatif, celle de "société" renvoie aux relations existant entre ses membres, ce que les sociologues définissent comme du "lien social". Une organisation qualifiée de "sociale" suppose la complexité des individus qui la composent. Si elle est une source de richesse et de renouvellement pour la société, cette complexité introduit également une variété qui peut entrer en conflit avec les intérêts de la société ou les règles collectives, à court ou à long terme.

mots clés : complexité, collectif, communauté, société.

1. Remarques préalables à propos de la définition du collectif et du social

Vous proposez comme critère de distinction entre le collectif et le social la "capacité individuelle de se représenter la société, et d'agir délibérément afin de modifier l'organisation sociale". Il me semble qu'un tel critère n'a guère de sens pour distinguer une collectivité d'une société dans les sociétés humaines ; il paraît, cependant, relativement plus adéquat pour différencier les sociétés animales des sociétés humaines.

Quand on passe des sociétés animales aux sociétés humaines, le degré de complexité s'accroît nécessairement. mais la plus ou moins grande complexité n'est toujours que relative. D'où un problème de méthode dans cette réflexion : déterminer des *seuils* de passage du collectif au social n'est pas possible de manière universelle, il y a nécessairement des seuils différents d'une collectivité ou d'une société à l'autre, selon le type de société ou de collectivité dont il est question. Il serait évidemment souhaitable de pouvoir disposer d'une définition théorique universelle, mais il paraît extrêmement difficile de pouvoir y parvenir.

Par ailleurs, celle qui est proposée paraît très marquée par les concepts philosophiques occidentaux, dominés par l'action sur le monde et la société elle-même, à l'opposé d'une vision bouddhiste, ou animiste, par exemple. Cette définition ne nous paraît pas établir des conditions nécessaires à la constitution du social ; elle répond peut-être à un souci de différencier les sociétés humaines des sociétés animales, mais il faut être conscient que la même distinction peut être établie dans les sociétés humaines elles-mêmes : ainsi, les sociétés dites "primitives", que l'on peut qualifier de non prométhéennes, ne répondent guère à ces critères, car l'individu ne cherche pas à modifier la société, mais au contraire, en quelque sorte à s'y couler et à s'y identifier. Mieux vaudrait se référer, à notre avis, comme point de départ, au critère de la complexité de l'organisation, qui apparaît dans les questions qui nous ont été posées, sans entrer dans les modalités de ses implications. Et y ajouter la forme et le contenu quantitatif et qualitatif des interactions que les membres de la société en question ont entre eux, dont la nature doit être précisée.

En effet, si toute société est un collectif, tout collectif n'est pas nécessairement une société : par exemple, des personnes voyageant dans le même avion constituent un collectif, pas une société. Alors que le terme de collectif est défini en quelque sorte par un critère quantitatif - il s'oppose à "individuel", et réunit un ensemble d'individus dont les liens ne sont pas définis -, en sociologie, le

terme de social se réfère aux rapports des hommes entre eux dans la collectivité. On parle de "lien social" pour désigner les rapports entre les personnes, et les faits ou phénomènes sociaux se définissent comme "ceux qui résultent des relations réciproques entre les membres du groupe ou du tout organique que forme la société". Le social concerne le comportement de l'individu par rapport à ses semblables dans la société ainsi que des groupes et classes de la société entre eux. Une société est ainsi constituée de relations entre des personnes (ou des animaux) régies par une forme de division du travail, et qui mettent quelque chose en commun. La notion de social ou de société renvoie donc nécessairement à un système d'interactions organisé et structuré complexe, ce qui n'est pas obligatoirement le cas pour le collectif.

Rappelons encore que la sociologie reprend traditionnellement la distinction de Tönnies entre deux types de fondements de la société, la *Gemeinschaft* - ou communauté- , groupement spontané fondé sur une proximité d'origine naturelle (voisinage, consanguinité), qui concerne un type de lien social relativement traditionnel, et la *Gesellschaft* - ou association - groupement fondé sur l'adhésion volontaire et sur l'intérêt, qui caractérise surtout les relations sociales contemporaines. *Gemeinschaft* et *Gesellschaft* peuvent d'ailleurs se combiner de diverses manières, et à des degrés divers. De toutes manières, pour la sociologie, c'est la nature du lien social qui caractérise la société, indépendamment de la représentation que les individus pourraient en avoir. Ces liens peuvent être, par exemple, de type hiérarchique ou égalitaire, se caractériser par la participation ou l'exclusion, ou encore par la coopération, l'antagonisme ou la concurrence... Il me semble que, pour différencier valablement une collectivité d'une société, il faudrait rechercher des critères non pas forcément identiques, mais tout au moins de même type, en complément du critère de complexité, qui, seul, n'est pas assez précis.

Le passage du collectif au social suppose donc à la fois un accroissement du degré d'organisation, et le développement de liens entre les individus qui composent la collectivité. Mais il ne s'agit pas seulement de multiplier les liens : certains liens peuvent être plus ou moins vitaux pour la constitution de la société en question, en fonction, par exemple, de l'environnement : là encore, on retrouve le problème de la possibilité d'une définition universellement valable.

2. Complexité individuelle et complexité sociale

Je partirai maintenant , dans cette présentation, de la première question : "Une organisation collective suffisamment complexe pour être qualifiée de "sociale" peut-elle être composée d'individus très simples ?"

Je pense qu'une première réponse, négative, peut être trouvée dans l'anthropologie : même des sociétés dites "primitives" supposent des rôles sociaux multiples et différenciés. Et dans ces sociétés mêmes, où les normes sociales s'inscrivent puissamment jusque dans la chair des adolescents lors des cérémonies d'initiation, tendant ainsi à limiter les déviances, l'individu ne saurait être qualifié de "très simple", dans la mesure où il n'y a pas seulement reproduction à l'identique des pratiques et de rôles sociaux définis une fois pour toutes. Ce ne sont pas des sociétés "sans histoire", mais des sociétés où les rythmes de l'évolution historique et du changement social ont été jusque là relativement lents ; l'aspiration au changement peut également s'y manifester, et passer là comme ailleurs, par l'individu. Et Pierre Clastres nous rappelle, à partir de l'interprétation qu'il fait du chant des indiens Guayaki, que même dans les sociétés primitives, le désir d'échapper aux contraintes sociales existe : ils s'affirment par ce chant individualistes, refusant l'échange et le partage que la société leur impose pour le produit de la chasse et les femmes. Et dans nos propres sociétés, le fait qu'au sein d'une même famille, certains enfants se forment par identification et d'autres par opposition aux adultes témoigne également de la difficulté d'une reproduction sociale programmable.

Une seconde réponse, complémentaire de la première, et également négative, se trouve chez Yves Barel dans son ouvrage "*Le Paradoxe et le système*" : les rôles sociaux relativement simples que définissent les spécialistes des différentes sciences humaines ne sont que des procédés analytiques,

qui isolent les comportements des acteurs dans les différents systèmes politiques, économiques, sociaux, culturels, etc. Mais il va sans dire que *l'homo oeconomicus* est également père de famille, citoyen, participant à tel ou tel culte, etc. Autrement dit, la multiplicité des rôles sociaux fonde la complexité de la société, mais la simplicité relative de chacun de ces rôles isolés par l'analyse sociale ne doit pas faire oublier la complexité de l'individu concret qui joue tour à tour et parfois simultanément plusieurs de ces rôles : Barel insiste ainsi sur le fait qu'il ne peut y avoir d'éléments simples lorsqu'il est question de systèmes sociaux, car leur place y est tenue par des individus toujours complexes.

La complexité de l'individu reflète donc, en grande partie, celle de l'organisation collective. Elle lui est d'ailleurs largement transmise par cette organisation dans les sociétés humaines, à travers l'enseignement reçu dans la famille, le village, ou maintenant, l'école. L'éducation vise à transmettre les normes sociales, les valeurs, le savoir théorique et technique de la société à l'individu, qui les intériorise. Mais la complexité de l'individu dépasse, ou en tout cas, échappe en partie à l'organisation collective, et le tout est autre chose que la somme des parties, qui sont de leur côté, autre chose que le tout. Cette complexité est aussi variété, richesse, et peut permettre à l'organisation collective de se développer, de changer, car tout système clos tend vers l'entropie, et seule l'ouverture vers l'extérieur rend possible la néguentropie. Quand l'identité et la soumission aux normes sont trop fortes, la complexité de l'individu tend à se réduire, et peu de changement reste possible. Au contraire, dans les sociétés où la contrainte sociale est moins puissante, plus souple, un autre virtuel peut surgir de l'instinct, de la génétique, du naturel individuel, de l'art, de la création, ou de relations établies au niveau individuel ou collectif avec l'extérieur, par lesquelles peut transiter l'information : un tel contact avec l'altérité conditionne renouvellement et changement. L'autonomie des individus et des groupes dans la société est ainsi condition de la variété et de la vitalité de celle-ci, et condition de l'innovation.

Il peut cependant y avoir des oppositions entre l'individuel et le collectif liées à des effets de nombre, de seuil : ce qui est valable à l'échelle d'un individu ou d'un groupe restreint ne l'est pas toujours à l'échelle de toute une société, et il faut tenir compte de seuils quantitatifs. En outre, la viabilité d'une organisation collective implique non seulement la diversité, mais une complémentarité positive des fonctions individuelles. Par exemple, il est agréable pour un individu d'avoir une voiture, mais il risque de se produire une catastrophe écologique ou des bouchons bloquant toute circulation si tous l'utilisent au même moment. Il y a, également, un équilibre précaire et difficile à tenir entre la normalisation, et l'autonomie : pour que l'autonomie des individus enrichisse la société, le lien entre l'individu et cette dernière doit l'emporter. C'est, en quelque sorte, le sens du message freudien : pour que la vie sociale et la civilisation soient possibles, le "ça" doit rester subordonné au "moi" et au "surmoi" socialisés ; et l'individu doit contrôler ses pulsions agressives ou sexuelles au profit du respect des normes qui conditionnent la cohésion sociale. Si l'autonomie de l'individu l'emporte sur son lien à la société, son effet sur cette dernière peut être plus négatif que positif, et tendre vers sa moindre progression, voire sa désintégration : c'est tout le problème, par exemple, de l'exode des cerveaux des pays du tiers-monde vers les pays développés, le sens aussi des essais infructueux de l'U.R.S.S pour retenir ses savants et intellectuels qui tentaient de partir vers l'occident. Là, il y a contradiction entre la logique de l'intérêt individuel et la logique collective, et l'intérêt individuel peut pousser les membres de la communauté à partir pour valoriser leur capital intellectuel et scientifique dans une autre société, au lieu de participer au développement, souvent lent et difficile, de leur société d'origine.

En réalité, le problème de l'existence ou non d'une contradiction ou d'une opposition entre les principes de fonctionnement individuel et social fait l'objet depuis longtemps de divergences entre les spécialistes des sciences sociales.

Prenons, par exemple, le cas des économistes. Pour les économistes libéraux, et notamment, les utilitaristes, cette opposition n'existe guère : l'intérêt individuel et les intérêts collectifs vont dans le même sens, le travail individuel profite à l'ensemble de la société, le travail du plus riche enrichit

même le pauvre par le dynamisme qu'il entraîne et les emplois qu'il crée. De même, sur le plan politique, à peu près à la même époque, dans le contrat social, la volonté collective émane des choix individuels, et la majorité qui se dégage du vote de ces derniers est supposée représenter un optimum à la fois pour la société et pour les individus. Dans un cas comme dans l'autre, le conflit n'est pas un objet de réflexion, et l'accent est mis sur la cohérence entre l'individu et la société, réunis autour d'un objectif commun, le progrès.

A l'opposé, les penseurs socialistes et Karl Marx mettront l'accent sur le conflit entre les intérêts des prolétaires, d'une part, leur prise en charge économique par la main invisible du marché, et politique par les procédés de représentation de la démocratie "bourgeoise", d'autre part. Ils dénoncent l'intérêt bafoué du travailleur aliéné à la compartimentation des tâches, et le sacrifice des intérêts du prolétariat, mais aussi de l'ensemble de la société, à l'ensemble de la bourgeoisie. Pour Marx, seul le prolétariat peut prendre en charge l'intérêt collectif, et, surmontant la crise du capitalisme, la conduire vers une société d'abondance où l'aliénation aura disparu : le communisme. Si, dans la société capitaliste, il y a contradiction entre les principes de fonctionnement individuel avec ceux de la société prise dans son ensemble, elle est censée disparaître avec l'avènement d'une société sans classes : il n'y aura pas de contradiction possible entre la classe ouvrière et son émanation, l'Etat socialiste, et la grève sera interdite.

Ne pas concevoir ni accepter l'existence de contradictions entre les principes de fonctionnement des individus ou de certains groupes et ceux de la société globale, est ainsi une conception qui peut conduire tout droit au totalitarisme. Il en est de même lorsqu'il s'agit de minorités nationales d'origine ethnique ou religieuse, qui sont alors considérées comme un danger pour l'unité de la nation, porteuses de valeurs nocives, et responsables des dysfonctionnements divers de la société. La marginalité est rejetée au nom de la cohésion et de la sauvegarde de la collectivité, dont la diversité et donc, la complexité, peut être réduite par des moyens qui peuvent aller du renforcement du contrôle social, comme, par exemple, dans la société singapourienne, à l'épuration ethnique.

Entre une autonomie excessive par rapport à ce que peut supporter une société donnée, allant jusqu'au détachement de l'individu, et une conception de l'unité collective qui refuse toute autonomie à l'individu, il y a place pour une autre conception et une autre réalité de l'autonomie, où l'individu ou le groupe marginal enrichissent le corps social, qui les enrichit à leur tour, mais moins naïve que les conceptions libérales : l'autonomie des individus ou des groupes est un lieu de tensions entre l'individuel, le minoritaire, porteurs d'un ailleurs interne ou externe, et le collectif auxquels ils se rattachent. C'est en fonction de cet ailleurs que cet enrichissement même est possible, qu'il provienne des contacts avec d'autres sociétés, ou de la créativité originale d'un artiste non réductible au présent de sa société, de déviants, fous et marginaux, dont l'imaginaire en deçà ou au-delà de leur propre monde le déborde de toute manière: ils témoignent d'autres aspirations, d'autres possibles. L'individu relie la société à la fois au réel et à l'extérieur d'elle-même, au niveau qui est le sien, et l'enrichit de ces liens : par exemple, les contacts et la vision d'un français à l'étranger ne sont pas les mêmes que ceux des représentants de son gouvernement ; de retour en France, il diffuse sa propre vision, qui peut être différente de celle de son ambassade, et contribuera à modifier et enrichir celle-ci. De même, l'expérience concrète que fait un individu donné, par exemple du chômage ou du divorce, ne correspond pas toujours à ce qui est l'objet d'un large consensus dans sa société, ni à ce qu'exprime la loi, la règle collective, à ce sujet ; ce vécu concret alimente une somme d'expériences et d'idées qui peuvent conduire à modifier les normes qui ont jusque là régi les comportements individuels. Il y a donc interaction permanente entre les niveaux.

L'opposition ou la contradiction n'est souvent qu'un moment de ce processus : ainsi, l'artiste doit-il parfois se retirer dans son individualité pour cultiver sa marginalité et sa capacité de création, contre sa propre société, pour pouvoir, justement, lui donner quelque chose de neuf : Van Gogh ne serait pas Van Gogh s'il n'avait accepté la marginalité liée à sa conception nouvelle de la peinture, qui n'a été reconnue par l'ensemble de la société que plus tard, et qui l'a alors fécondée. La plupart des oppositions et contradictions disparaissent ainsi dans le temps d'une durée suffisamment longue

pour prendre en compte leur intégration, et l'individuel ou le particulier se fondent dans le social. De tels processus sont cependant toujours douloureux, à la fois pour l'individu ou le groupe social innovateur, comme pour la société concernée, car le changement ou le développement sont vécus le plus souvent comme un arrachement difficiles au passé, même s'ils s'accompagnent d'améliorations apparemment mesurables et tangibles.

BIBLIOGRAPHIE

Barel Yves .- *Le paradoxe et le système*, Grenoble, Presses universitaires, 1979.

Clastres Pierre .- *Recherches d'anthropologie politique*, Paris, Seuil, 1980.

F. Tönnies .- *Communauté et Société*, Paris, PUF, 1944.

L'intelligence collective des êtres humains est-elle négative ?

John Stewart ¹

¹ COSTECH, Université de Technologie de Compiègne, France.

E-mail: stewart@pasteur.fr

Résumé

L'intelligence collective des animaux, et prototypiquement celles des insectes dites "sociaux", est le résultat émergent des interactions entre des agents "réactifs" qui fonctionnent en suivant des règles purement locales, comme le montrent de nombreuses modélisations récentes. Autrement dit, la viabilité collective d'une fourmilière n'est en aucun cas le résultat d'une intention délibérée de la part des agents individuels. Cet article examine la possibilité que la viabilité à long terme (des millions d'années) de cette forme de vie collective soit due en partie précisément au fait que les individus ne pensent pas leur organisation collective. En ce qui concerne les êtres humains, le système d'économie de marché requiert et résulte d'une application de notre intelligence à notre organisation sociale. Il semblerait que ce système condamne l'humanité au progrès continu, et à une refonte permanente de nos relations sociales sous l'impulsion des innovations technologiques. Il n'est pas certain que notre intelligence collective, coupée des traditions qui font la "sagesse" et la viabilité des sociétés animales, puisse suivre indéfiniment.

mots clés : intelligence collective, société, économie, tradition, innovation.

1. L'intelligence collective des animaux.

Lorsqu'un observateur humain regarde une fourmilière, il est émerveillé (à juste titre) par l'organisation de la colonie. Les activités incessantes de milliers, voire de millions d'individus sont co-ordonnées de telle sorte que le nid est construit, entretenu et défendu contre les agressions d'intrus ; des pistes sont formées pour ramener de la nourriture par le chemin le plus court ; les oeufs sont soignés, les larves sont nourries, les cocons sont soignés, et les nouveau-nés se voient attribuer une place en fonction d'une division du travail dynamiquement adaptée aux besoins de la colonie. Plus profondément encore, chaque individu agit de telle sorte qu'il contribue à la viabilité d'un super-organisme - la colonie - dont il dépend lui-même pour sa propre survie. Puisque notre imaginaire est nourri par l'expérience de notre propre vie sociale, il est tout à fait normal que nous pensions immédiatement qu'une telle organisation doit provenir d'un centre de décision intelligent. Ainsi, quand il fut découvert que tous les oeufs sont pondus par une seule femelle, et que si cette femelle vient à mourir l'activité de toute la colonie se désagrège, il était tout à fait normal de nommer cet individu "reine" et de penser que toute cette organisation découlait de ses ordres, relayés sans doute par une structure administrative hiérarchique, et une disposition juridique et policière pour renforcer les comportements individuels nécessaires à la viabilité collective. Une grande partie de la fascination intellectuelle exercée par le champ nommé "intelligence collective" provient de la découverte qu'il n'en est rien. Des travaux comme ceux de Bonabeau et Théraulaz (1994) montrent que la construction du nid ou la formation d'une piste (pour ne prendre que ces exemples) est le résultat *émergent* d'un grand nombre d'interactions relativement simples et *purement locales*. "L'intelligence" en question n'est localisée nulle part ; elle est entièrement *distribuée* sur la totalité des entités qui entrent en interaction au cours du processus, qu'il s'agisse des différents agents, des phéromones qu'ils émettent et auxquels ils réagissent, ou du milieu physique lui-même. La "reine" fourmi n'a rien d'une souveraine omnisciente, elle n'a aucune vision stratégique lui permettant de planifier dans leurs moindres détails les oeuvres à entreprendre, les batailles à mener. Son cerveau minuscule, comme celui de chacun de ses congénères, n'a pas les

capacités de mémoire et de raisonnement nécessaires pour cela ; il est tout juste suffisant pour lui permettre de réagir *localement* de manière appropriée en fonction des stimuli *locaux*.

Ces travaux, qui réunissent des modélisations mathématiques avec simulation sur ordinateur et d'autre part des observations empiriques d'une grande finesse aussi bien au niveau des propriétés des composants du système (agents, phéromones, milieu) qu'au celui des propriétés émergentes, sont extrêmement convaincants. Néanmoins, dans un certain sens ils nous laissent sur notre faim. Plutôt qu'une pleine explication des phénomènes, ils semblent indiquer qu'en fin de compte il n'y avait rien à expliquer. En anglais, on dirait qu'ils n'expliquent pas, "they don't explain so much as explain away". On peut exprimer plus clairement ce "manque à gagner" en se référant aux modélisations et aux simulations. Pour que le phénomène émergent se produise, il faut qu'un assez grand nombre de paramètres et de fonctions régissant les interactions entre les composants prennent des valeurs et des formes assez particulières. Si on demande comment il se fait que, dans la nature, ces paramètres et ces fonctions aient effectivement exactement les valeurs et les formes requises, les travaux que j'ai cités n'apportent pas de réponse. Nous nous trouvons ici en présence de l'incapacité de la biologie contemporaine de penser sérieusement la causalité finale d'Aristote (Rosen 1991). La réponse habituelle est d'évoquer la "sélection naturelle" ; mais dans la mesure où celle-ci se réduit au jeu du hasard et de la nécessité, pour reprendre le célèbre formule de Monod, il s'agit encore d'une "explication" qui n'en est pas une.

Cette difficulté est en relation étroite avec la question de l'hérédité, et la relation entre la génétique (une sous-discipline) et la biologie en général. Le fait est que la biologie contemporaine ne possède pas une théorie satisfaisante de l'hérédité, à savoir le processus par lequel les générations successives parviennent à se reproduire et à se ressembler. Il est bien reconnu par des philosophes de la biologie (Lenay 1993, Guyon 1996) que la génétique mendélienne a opéré une disjonction entre hérédité et développement. En fait, et aussi surprenant que cela puisse paraître, la génétique n'est même pas une science de l'hérédité (Oyama 1985, Stewart 1993). La raison profonde en est que la génétique est épistémiquement aveugle sur tout ce qui est invariant au sein d'une espèce donnée. Un gène, par définition, est ce dont la *variation* donne lieu à une *variation* dans un phénotype. Sans variation, il n'y a aucun lien entre génotype et phénotype, et par conséquent aucun lien entre génétique et biologie. Par conséquent, la biologie contemporaine est affectée d'une immense tâche aveugle (qu'elle ne peut voir, pas plus que nous ne pouvons voir la tâche aveugle dans notre champ visuel) concernant les processus qui permettent à des formes vivantes de perdurer au fil des générations.

On pourrait objecter ici que chez bien des espèces d'insectes "sociaux", seule la future reine passe l'hiver : où se situe l'hérédité sinon dans les gènes ? Eh bien, très précisément, dans tout ce qui est invariant et que la génétique ne voit donc pas. Très fondamentalement, elle se situe en partie dans certaines propriétés de la matière. Par exemple, l'eau est un liquide ; la seule substance inorganique, à part le mercure, qui soit liquide aux températures ambiantes à la surface de la planète Terre. Ce fait est déjà remarquable : l'oxygène est plus léger que le soufre, et H₂S est un gaz, donc normalement H₂O devrait également être un gaz. En fait, du point de vue physico-chimique, l'eau est une substance très particulière qui donne lieu notamment à la polarité hydrophile-hydrophobe sans laquelle la morphogenèse du vivant n'aurait pas lieu (Saunders 1993). Plus spécifiquement, l'hérédité se situe également en partie dans la niche écologique de l'espèce en question, dans son "Umwelt" (von Uexküll 1966, Clément et al 1996). Et enfin, l'hérédité se situe également en partie dans la reine elle-même, un organisme vivant ayant une anatomie, une physiologie et un comportement qui sont loin de se réduire à ses gènes (Lewontin 1993). Ainsi, s'il n'y avait que des gènes, il n'y aurait pas de fourmis ; et si l'on élevait des larves de reine en complète isolation (avec ou sans gelée royale, mais sans eau et sans niche écologique), on n'obtiendrait certainement pas une ruche fonctionnelle.

Or il me semble que ce qui nous manque pour *expliquer* pleinement l'intelligence collective des fourmis se situe précisément au centre de cette tâche aveugle de la génétique. Pour nommer ce

manque (à défaut de le combler, car je n'ai aucune prétention d'apporter moi-même une solution satisfaisante au problème), je propose de parler d'un "respect de la tradition". Les colonies de fourmis parviennent à s'organiser parce que leurs ancêtres possédaient ce "secret", et elles ont respecté la tradition en se comportant individuellement et collectivement de la même manière. Les fourmis ne savent pas ce qu'elles font (cela, les travaux modernes sur l'intelligence collective le montrent à merveille), mais elles savent faire et elles maintiennent ce savoir-faire avec une belle fiabilité.

2. L'intelligence collective des êtres humains.

Le but de cette caractérisation, volontairement provocatrice, de l'intelligence collective des fourmis est de fournir un point de repère pour aborder le sujet des Journées de Rochebrune en 1996, à savoir l'intelligence collective des êtres humains. Pour aller droit au point de contraste, l'intelligence collective des êtres humains est "complexe", selon la définition proposée par Bonabeau & Théraulaz (1994) :

"L'intelligence collective... peut être qualifiée de "simple" dans la mesure où les agents cognitifs ne perçoivent pas leurs interactions sémiotiques en tant que telles. Ceci nous permet de définir, par contraste, une "intelligence collective complexe" : celle-ci est caractérisée justement par le fait que les agents peuvent *penser* intentionnellement leur propre sociabilité, et agir en conséquence. C'est probablement le propre des sociétés humaines, avec naissance de la politique et de la culture". (op. cit, p. 24).

Autrement dit, les êtres humains agissent délibérément en vue de modifier leur propre organisation collective - ou bien, le cas échéant, afin de ne pas la modifier en préservant "la Société contre l'Etat" (Clastres 1974). Comme proposition terminologique, on peut y voir une définition même de ce que l'on entend par "social".

Dans ce court article, je ne prendrai qu'un seul exemple pour illustrer les complexités qui peuvent en résulter, notamment concernant un thème cher aux historiens : les effets non voulus de l'action (Giddens 1976, 1984). Depuis quelques deux mille ans (c'est-à-dire depuis l'invention de la monnaie frappée chez les Grecs), les êtres humains ont délibérément mis en place une modalité radicalement inédite de régulation de leurs relations sociales : il s'agit de l'économie de marché. Cette mise en place fut longue et progressive ; elle était initialement réservée au seul commerce, car la mode de production chez les Grecs était esclavagiste. A la Renaissance, avec la naissance du capitalisme, le secteur productif entra à son tour dans l'économie de marché, mais partiellement seulement. Il y a seulement un siècle, la masse de la population dans tous les pays du monde vivaient à la campagne où la production était encore en grande partie une auto-production échappant au marché. Mais aujourd'hui les campagnes sont désertées et le secteur agricole est entré de plein pied dans une économie de marché internationale. Pour résumer, depuis l'effondrement du stalinisme dans les pays de l'Est, nous sommes actuellement dans une situation où (à une première approximation) la quasi-totalité de nos relations sociales sont déterminées par les "lois" du marché. Notre dépendance sur "l'économie" est cruellement illustrée par le phénomène du chômage, qui augmente de manière incontrôlable à chaque dysfonctionnement de l'économie.

Il est courant de dire que le système d'économie de marché, contrairement au système communiste, n'est pas un système "pensé". On peut alors se demander si les êtres humains ont vraiment *délibérément* mis en place un tel système ? Mais que l'idéologie néo-libérale soit ou non une "pensée", il est certain que la "construction" de l'Europe sur la mode de la "déréglementation", de la privatisation et du démantèlement des services publics, et plus généralement "le Nouvel Ordre Mondial" d'inspiration clairement économiste, sont des actes politiques délibérés - en toute connaissance de cause concernant leurs effets en terme de fracture sociale, d'inégalités Nord-Sud structurelles, et de saccage de l'éco-système.

Or le point essentiel que je souhaite souligner est ceci. Dans une économie de marché parvenue à maturité, comme la nôtre, *l'innovation permanente est une nécessité*. Dans la guerre économique de tous contre tous, ceux qui n'innovent pas sont condamnés à disparaître. Tous nos dirigeants et faiseurs d'opinion nous le répètent assez souvent pour qu'il soit inutile d'argumenter longuement sur ce point. Je veux seulement remarquer que ce faisant, nous prenons le contre-pied exact de l'intelligence collective des animaux - les fourmis, les mammifères, et même les sociétés "traditionnelles" humaines qui nous ont précédés. Nous nous trouvons dans une situation où nous sommes obligés de nous *interdire* de respecter les traditions. Et je pose la question : est-ce que nous savons réellement ce que nous faisons ?

Cette situation comporte une énorme ironie. Au niveau individuel, notre intelligence est incommensurablement plus grande que celle des animaux. Nous sommes capables d'imaginer des buts, et par la technologie de les réaliser, d'une manière inaccessible aux animaux. Et cette intelligence est "collective", jusqu'à un certain point : nous sommes capables de co-ordonner nos actions à une assez grande échelle pour placer un homme sur la lune, pour ne prendre que cet exemple-là. Mais une innovation de cet ordre n'est pas sans ses dangers. L'évolution biologique, qui à l'échelle de temps des millions des années a connu ses propres innovations, peut illustrer les enjeux. Quand, il y a quelques 1000 millions d'années, les premiers organismes multicellulaires ont fait leur apparition, tout un monde phénoménal venait à exister. Mais ces nouveaux êtres devaient affronter des problèmes qui étaient littéralement inexistantes pour leurs prédécesseurs unicellulaires, un exemple notable étant le cancer. Un organisme unicellulaire ne peut mourir d'un cancer, car sa croissance ne menace pas l'existence d'un super-organisme dont il dépend lui-même. Il n'en est plus de même dès lors que de nombreuses lignées cellulaires contribuent à la formation d'un seul organisme multicellulaire dont toutes les lignées dépendent. Ainsi, corrélativement avec l'apparition d'une nouvelle forme d'organisation, les nouveaux êtres qui la composent sont *condamnés à réussir* dans la résolution de problèmes inédits. Buss (1987) a bien souligné la complexité des processus en jeu.

Il est bien reconnu, depuis assez longtemps, que dans certaines conditions l'intelligence collective des êtres humains peut être bien en deçà de leurs intelligences individuelles. Cela est notamment le cas dans des phénomènes de "foule", étudiés par Le Bon, Tarde et Freud et magistralement synthétisés par Moscovici (1981). Mais ce genre de phénomène a généralement été considéré comme marginal, l'exception pathologique qui confirme la règle de l'ordre social. Nous avons moins souvent envisagé la possibilité - inquiétante, certes, mais la lucidité serait alors plus que jamais nécessaire - qu'avec la refonte permanente de nos relations sociales sous l'impulsion de l'innovation technologique qui découle des mécanismes d'une économie de marché, nous sommes pris à notre insu et indépendamment de notre volonté dans un processus *massif* de réduction à l'état de foule. On pensera ici au phénomène d'une société de consommation (individuelle), et tout particulièrement aux sondages d'opinion et aux mass-média régies par l'audimat, qui réduisent des citoyens à l'état d'une foule d'individus qui ne se parlent plus en termes de responsabilité politique et culturelle.

Résumons donc à propos de notre propre intelligence collective dans la situation actuelle. Il se peut que nous soyons *condamnés* à réussir la résolution d'un problème au moins aussi radicalement inédit que celui des organismes multicellulaires face au cancer : celui de gérer les conséquences d'une innovation permanente. Le défi est d'autant plus redoutable que nous nous sommes privés précisément du ressort de l'intelligence collective de nos ancêtres (humains et animaux), à savoir ce que j'ai appelé "le respect des traditions". Aussi immense que soit notre intelligence collective, il n'est pas sûr qu'elle soit égale à cette tâche qu'elle s'est créée de toutes pièces. C'est peut-être justement parce que l'on croit pouvoir se passer des traditions, entre autres la dimension morale, que la condition humaine devient de plus en plus problématique. Comme le dit Thuillier (1995), nous avons perdu le sens de la poésie, et il se peut que notre société occidentale soit déjà morte dans l'âme, prélude à une "implosion" matérielle qui nous menace sans qu'on la voit.

C'est pourquoi que je me demande si, en fin de compte, notre intelligence collective n'est pas globalement négative.

Remerciements

Je tiens à remercier les deux relecteurs pour leurs remarques utiles et constructives. Je me suis permis d'incorporer certains de leurs commentaires dans la version révisée de ce texte. Je remercie également Véronique Havelange pour des discussions enrichissantes au cours desquelles elle m'a fait bénéficier de son immense connaissance de la littérature sociologique.

Références

- Bonabeau E. & Théraulaz G. Eds, (1994). *Intelligence collective*. Hermes, Paris.
- Buss L.W. (1987). *The Evolution of Individuality*. Princeton University Press.
- Clastres P. (1974). *La société contre l'état*. Editions de Minuit, Paris.
- Clément P., Scheps R. & Stewart J. (1996). Une interprétation biologique de l'interprétation. Colloque de Cerisy, *Herméneutique, Sciences, Textes*. Editions du Seuil, Paris.
- Giddens A. (1976). *New rules of sociological method*. Hutchinson, London.
- Giddens A. (1984). *The constitution of society*. Polity Press, Cambridge.
- Guyon J.P. (1996). Le temps des gènes. In Couloubartsis L. Ed, *Actes du Colloque "Figures du temps"*, Université Libre de Bruxelles.
- Lenay C. (1993). Caractères adaptatifs et représentations symboliques. *Intellectica* **16**, 209-257.
- Lewontin R.C. (1993). Le rêve du génôme humain. *Ecologie Politique* **5**, 125-152.
- Moscovici S. (1981). *L'âge des foules*. Fayard, Paris.
- Oyama S. (1985). *The ontogeny of information: developmental systems and evolution*. Cambridge University Press, Cambridge.
- Rosen R. (1991). *Life Itself*. Columbia University Press, New York.
- Saunders P.T. (1993). L'évolution des formes biologiques. *Intellectica* **16**, 61-83.
- Stewart J. (1993). Au-delà de l'inné et de l'acquis. *Intellectica* **16**, 151-174.
- Thuillier P. (1995). *La grande implosion*. Fayard, Paris.
- von Uexküll J.V. (1966). *Mondes animaux et monde humain*. Gonthier, Paris.

Index des auteurs

<i>Evelyne Andreewsky</i>	127
<i>Danièle Bourcier</i>	127
<i>Christian Brassac</i>	85
<i>Michèle Courant</i>	9, 31
<i>James De Almeida</i>	135
<i>Fabienne de Gaulejac</i>	71
<i>Guillaume Deffuant</i>	163
<i>Ludovic Delépine</i>	21
<i>Didier Desor</i>	95
<i>Jean-Louis Dessalles</i>	49
<i>Nils Ferrand</i>	143
<i>Thierry Fierville</i>	95
<i>Alain Gallo</i>	71
<i>Hervé Glotin</i>	113
<i>Juan Gonzalez</i>	39
<i>Rafael Laboissiere</i>	113
<i>Patricio Lerena</i>	9
<i>Anne Nicolle</i>	21
<i>Christophe Parisse</i>	127
<i>Herminia Peraita</i>	127
<i>Antony Robert</i>	31
<i>Marie-Noelle Sarget</i>	179
<i>John Stewart</i>	85, 185
<i>Anne-Marie Toniolo</i>	95
<i>Laurent Toubiana</i>	105
<i>Jean-François Vibert</i>	105
<i>Tiziana Zalla</i>	61

Dépôt légal : 2ème trimestre 1996

Imprimé à l'Ecole Nationale Supérieure des Télécommunications -
Paris

ISSN : 0751-1353