

**Journées de Rochebrune : Rencontres interdisciplinaires sur les
systèmes complexes naturels et artificiels**

Edition 97 : du 2 au 9 Février 1997

***Invariance, Interaction, Référence :
L'identité en question***

Avec le soutien de :

ECAL

ENST

DGA (contrat DRET/DS/DR n°96-1189/A000)

Sous le patronage de :

l'Association pour la Recherche Cognitive

l'AFCET (Comité technique Vie Artificielle)

**Journées de Rochebrune : Rencontres interdisciplinaires sur les
systèmes complexes naturels et artificiels**

Edition 97 : du 2 au 9 Février 1997

***Invariance, Interaction, Référence :
L'identité en question***

L'une des caractéristiques des processus cognitifs est la capacité de reconnaître le même dans le différent, la permanence dans le changement. Quels sont les ressorts cognitifs donnant ces possibilités ? Quel est le rôle de l'interaction sensori-motrice, de la capacité de désignation (référence), de la dimension pragmatique ? Quelles sont leur importance ou leur utilité dans l'ensemble des mécanismes cognitifs ?

Ces questions se posent avec peut-être encore plus d'acuité pour le scientifique face à un système complexe. De tels systèmes sont généralement en perpétuelle transformation, ce qui peut les rendre totalement méconnaissables d'un instant à l'autre. Leur identification par l'observateur est donc au cœur de la démarche scientifique. Elle implique la définition des frontières et des dynamiques de transformation. Comment ces opérations s'effectuent-elles ? Sur quelles bases scientifiques ?

Peut-on se poser ces questions sans poser celle de l'identité de l'observateur ? Quel sens donner à l'identité humaine individuelle, collective, culturelle ? Peut-on les caractériser, en proposer des mécanismes de construction ou d'évolution ? Peut-on en retracer l'histoire précise dans des circonstances particulières ? Quelle interaction entre cette identité et la démarche cognitive d'identification par catégorisation et/ou de désignation ?

Comité d'organisation

Christian Brassac
Guillaume Deffuant
Christophe Parisse
Pierre Yves Raccach
Anne Rouille
John Stewart

Comité de Programme

Hugues Bersini
Paul Bourguin
Christian Brassac
Michele Courant
Guillaume Deffuant
Wolf Eberwein
Nils Ferrand
Anne Nicolle
Pierre Yves Raccach
Martine Timsit

Avec le soutien de :

ECAL
ENST
DGA (contrat DRET/DS/DR n°96-1189/A000)

Sous le patronage de :

l'Association pour la Recherche Cognitive
l'AFCET (Comité technique Vie Artificielle)

Programme

Lundi 3 février 1997

Conférence sollicitée :

F. de Gaulejac et A. Gallo p. 7
Conceptualisation et contextualisation : d'une épistémologie, l'autre.

Conférences ordinaires :

N. Grégori p. 15
Désigner et faire évoluer un objet en conception distribuée.

J. Stewart p. 27
L'identité en biologie.

P.Y. Raccach p. 39
L'identité des sciences cognitives : de l'épistémologie à la sémantique.

H. Glotin et P. Pinel p. 49
Dictyostelium discoideum : d'une identité individuelle à une identité collective.

Modérateur pour la journée : G. Deffuant

Agitateur pour la journée : H. Bersini

Mardi 4 février 1997

Conférence sollicitée :

J. Sallantin —
Une philosophie computationnelle de l'identification

Conférences ordinaires :

P. Besnard —
Sur la quantification et l'identité en logique formelle

G. Deffuant p. 69
De l'identité logique à l'identité empathique : un cheminement impossible ?

P. Benhamou p. 77
La voiture de mon beau-frère est-elle toujours la voiture de mon beau-frère ?

D. Dubois et W. Eberwein —
Contribution à la constitution de l'objectivité en sciences cognitives

Modérateur pour la journée : D. Snyers

Agitateur pour la journée : P.Y. Raccach

Mercredi 5 février 1997

Conférence sollicitée :

- M. Timsit et E. Andreewsky** p. 87
Dynamique identitaire et psychopathologie

Conférences ordinaires :

- E. Andreewsky** p. 93
Problèmes d'identité des systèmes

- J. Gonzalez** p. 99
"Qui suis-je?" "Simple: je suis le chef". Identité individuelle, perceptions et descriptions

- S. Herbin** p. 113
Critique de la reconnaissance comme identification

- P. Beust et A. Nicolle** p. 125
L'identité dans un modèle de la dépendance nominale

Modérateur pour la journée : D. Luzeaux

Agitateur pour la journée : N. Ferrand

Jeudi 6 février 1997

Conférence sollicitée :

- H. Bersini** —
L'identité incertaine des systèmes chaotiques

Conférences ordinaires :

- C. Lenay** p. 135
Invariants sensorimoteurs et Perception

- D. Bourcier** —
Réflexions sur les modes de construction de l'identité dans les systèmes juridiques à partir de l'exemple des "sans-papiers"

Modérateur pour la journée : J.M. Fouet

Agitateur pour la journée : J. Stewart

Vendredi 7 février 1997

Conférence sollicitée :

J. Ferber et J.P. Müller —
Invariant et identité dans les systèmes multi-agents

Conférences ordinaires :

D. Luzeaux p. 145
Apprentissage de la commande de robots : le rôle de la boucle sensori-motrice

J.M. Fouet p. 155
Problèmes liés à l'identité dans la construction de Systèmes à Base de Connaissances

R. Munos p. 161
Catégorisation adaptative de données sensori-motrices pour un système d'apprentissage par renforcement

V. Decugis et J. Ferber p. 173
Architecture multi-agents de Robots Mobiles et Autonomes

E. Dedieu et P. Bessière p. 185
L'identité comme fonction ou structure : une illustration dans le développement de robots

D. Snyers p. 191
Identification d'un opérateur pour le contrôle de processus dynamiques

S. Ploux p. 201
Construction et interaction de modalités perceptives - Définition d'un cadre théorique et propositions pour un atelier expérimental

Modérateur pour la journée : C. Parisse, H. Peraita
Agitateur pour la journée : M. Courant

Index des auteurs

E. Andreewsky	p. 87, p. 93
P. Benhamou	p. 77
H. Bersini	—
P. Besnard	—
P. Bessière	p. 185
P. Beust	p. 125
D. Bourcier	—
F. de Gaulejac	p. 7
V. Decugis	p. 173
E. Dedieu	p. 185
G. Deffuant	p. 69
D. Dubois	—
W. Eberwein	—
J. Ferber	p. 173
J.M. Fouet	p. 155
A. Gallo	p. 7
H. Glotin	p. 49
J. Gonzalez	p. 99
N. Grégori	p. 15
S. Herbin	p. 113
C. Lenay	p. 135
D. Luzeaux	p. 145
J.P. Müller	—
R. Munos	p. 161
A. Nicolle	p. 125
P. Pinel	p. 49
S. Ploux	p. 201
P.Y. Raccah	p. 39
J. Sallantin	—
D. Snyers	p. 191
J. Stewart	p. 27
M. Timsit	p. 87

CONCEPTUALISATION ET CONTEXTUALISATION : D'UNE EPISTEMOLOGIE, L'AUTRE

Fabienne de Gaulejac et Alain Gallo

Laboratoire de Neurobiologie et Comportement

Université Paul Sabatier

118, route de Narbonne

31062 Toulouse cedex, France

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

Reconnaître le même dans le différent : l'activité scientifique de conceptualisation et de décontextualisation

La conceptualisation, quelle que soit la conception que l'on retienne, semble pouvoir se définir comme un processus d'abstraction (Cassirer, 1977) composé de deux opérations essentielles : abstraire l'objet d'étude du contexte d'observation et abstraire le commun des caractéristiques des objets. Elle permet finalement de reconnaître comme identiques des éléments différents les uns des autres. Selon la thèse aristotélicienne, par exemple, la conceptualisation consiste en un processus de condensation de l' "insurmontable" diversité des données sensibles. De la comparaison des multiplicités données dans l'expérience sensible, la conceptualisation retient les aspects saillants possédés en commun par la majorité des objets, les traits essentiels qui définissent la convergence de ces objets, par exemple leur similitude physique. Dans cette psychologie de l'abstraction, c'est cette identité, maintenue envers et contre tous les changements affectant les contenus particuliers, qui constitue la forme spécifique du concept. On attend du concept scientifique qu'il substitue au caractère plurivalent du contenu perceptif une détermination rigoureuse et univoque.

La caractéristique essentielle du concept est donc sa fonction unitaire. Les objets sont subordonnés aux concepts. Il en résulte une stabilité des énoncés par delà la pluralité des phénomènes. Ainsi en éthologie, on forme l'item 'manger' en dégageant de la diversité des actes alimentaires observés en différents lieux et à des moments distincts chez des espèces diverses, c'est-à-dire en dégageant de leur contexte spatio-temporel, l'ensemble des indices apparentés. Après abstraction de ces traits dépouillés d'éléments contextuels hétérogènes et non essentiels, les comportements partageant cette même structure sont ainsi regroupés en classes. A la perception effectivement éprouvée est alors substituée une image globale revendiquant un certain degré de généralité. Ce que nous appelons comportement animal est donc, en ce sens, un concept qui subsume une collection variée d'activités diverses émises par des individus différents appartenant aux multiples espèces. Mais pour une éthologie zoologique, liée à la diversité de la systématique, cette définition est trop large car (a) les comportements des différentes espèces animales doivent être étudiés comme des réalités différentes les uns des autres et (b) les comportements doivent être distingués les uns des autres (comportement alimentaire, de reproduction, de communication...).

L'emboîtement relève aussi de la pensée classique : on s'élève d'un niveau conceptuel à l'autre (items alimentaires, comportement alimentaire, comportement) en ouvrant le champ de la réflexion à un domaine encore plus étendu, par l'élimination de tel ou tel indice retenu jusque là et dorénavant passé sous silence. Mais, quel que soit le niveau d'abstraction considéré, le comportement étudié

tombe sous un concept qui, soit gomme les particularités des observations (thèse empiriste), soit doit s'énoncer comme la règle de construction de ces particularités mêmes (thèse idéaliste). Pour cette dernière école, le concept véritable ne doit pas écarter les particularités qui spécifient les contenus qu'il subsume. Au contraire, il doit chercher à dévoiler (règles de construction) la nécessité de leur manifestation et de leur enchaînement. Dans les mathématiques par exemple, les cas particuliers se voient décerner le statut de phases entièrement déterminées au sein du processus général de variation. Le concept se caractérise, non par la généralité d'une image représentative, mais par la validité générale d'un principe sériel. Au lieu de prélever des parties quelconques sur la multiplicité offerte, il faut engendrer une relation univoque pour les éléments qui la composent. A l'universalité abstraite du concept (mise à l'écart de toutes les différences spécifiques) est ainsi opposée l'universalité concrète de la formule mathématique (concept accueillant en lui les traits particuliers de tous les éléments). Plus précisément, chaque fonction mathématique représente une loi universelle qui, grâce aux valeurs successives susceptibles d'être prises par la variable, sous-tend tous les cas particuliers pour lesquels elle vaut. Cependant, pour ce qui est des systèmes vivants auto-organisés, la loi n'est pas de même nature. En effet, la finalité apparente du vivant ne renvoie pas seulement à un principe prédéterminé d'autorégulation (l'autorégulation n'est jamais qu'une adaptation d'une loi à des conditions empiriques particulières) mais à une auto-organisation d'un système loin de l'équilibre (Lestage, 1994), qui échappe à une loi immanente responsable des régularités observées.

Un fonctionnement psychologique universel ?

Cette activité scientifique de conceptualisation vise à produire, sur le plan de la technologie, des artefacts, des machines, des robots se comportant de façon adaptée quel que soit le contexte. On conçoit alors que le modèle classique de rationalité scientifique puisse fasciner l'éthologiste. Dès lors il recherchera les lois générales du comportement des animaux telles qu'elles expliquent les activités observées quel que soit leur contexte. Dans cette optique dominante, à cause de l'opération épistémique de décontextualisation propre à la conceptualisation, les animaux héritent en définitive d'une conception des systèmes vivants comme êtres décontextualisés, décontextualisants et aptes à pratiquer la seconde opération consistant à reconnaître l'identité sous la diversité phénoménale (invariance). C'est ainsi que certains éthologistes finissent par se (nous) persuader que le concept existe chez le pigeon (Herrnstein, 1984). L'approche, en effet, implique que l'animal, en tant que système cognitif naturel, possède des représentations qui sont des états mentaux portant sur le monde et une possibilité de combiner ces représentations d'une façon qui évoque un calcul, celui-ci aboutissant à la résolution d'un problème (Andler, 1992). Pichot (1992) résume cette conception (qu'il critique) par la formule suivante : "*Penser, c'est résoudre un problème posé ; résoudre un problème, c'est effectuer un calcul logique plus ou moins comparable à ceux qu'effectuent les ordinateurs. Penser, c'est donc calculer*". Ce calcul, ou traitement computationnel, se ramène à des procédures algorithmiques qui, par définition, permettent de résoudre un problème en un nombre fini d'étapes. Leur exécution ne demande pas d'interprétation immédiate : issu du formalisme mathématique, le cognitivisme doit beaucoup au concept de système syntaxique composé de symboles dénués de toute signification intrinsèque. Le système cognitif interagit avec la forme des symboles

qu'il manipule (avec des éléments physiques discontinus), et non avec leur sens. Ce n'est qu'ensuite que la dynamique de l'esprit attribue un relief interprétatif. Ainsi dans cette philosophie cognitive de l'esprit, la pensée serait structurée comme un langage technique composé de termes renvoyant à des concepts et non de mots (Racah, 1995) et la représentation serait envisagée sous une forme propositionnelle. Notons que si aucune interprétation n'est requise durant le traitement informationnel, ces calculs peuvent donc aussi bien être réalisés par une machine.

Mais ne peut-on imaginer que certains êtres vivants ne fonctionnent pas selon ce modèle, parce que, pour eux, les formes s'inféderaient au sens de telle sorte qu'il ne soit jamais assuré, contrairement à l'opinion des instinctivistes, qu'un stimulus constitue une détermination univoque. Chez l'animal par exemple, un objet, dans des contextes variés, peut avoir des significations totalement différentes : ce grand mâle chimpanzé, manifestant à l'égard d'un jeune les signes affectifs de l'épouillage, lorsqu'il est en colère, s'en saisit comme d'une branche ou d'une pierre et le fracasse violemment contre le sol. Il conviendrait alors de se demander comment faire fonctionner un autre être vivant que nous dans la limitation fondatrice ainsi exprimée : "*nous sommes les porte-paroles d'êtres qui ne parlent pas*" (Latour, 19XX) ; "*nous ne pouvons pas sortir de notre propre domaine cognitif*" (Maturana, 1987).

Une lecture des relations sociales des loups en situation de réaction à la nouveauté

Nous partirons des travaux de ces dernières années centrés sur un comportement important dans la vie des animaux, à savoir la réaction à la nouveauté. Le comportement des animaux vis à vis d'objets nouveaux introduits dans leur milieu se définit classiquement de façon unitaire comme une activité exploratoire dont la forme peut varier selon les individus ou le contexte : exploration olfactive ou tactile, manipulation... sous dépendance d'une fonction psychologique, la curiosité (Berlyne, 1960). Cette fonction psychologique est soustendue par l'activité de structures nerveuses connues (Vankov et al., 1995). Elle s'expliquerait, au niveau neurophysiologique, par l'existence postulée d'un mécanisme comparateur entre les informations entrantes et les informations mémorisées (Sokolov, 1963; Jensen, 1994).

Chez les animaux vivant en collectivité, c'est souvent en relation avec la réaction à la nouveauté que les rapports sociaux sont étudiés (Chamove, 1983 sur les macaques ; Nakanishi et al., 1992 chez les bovins...). Parmi les canidés, les loups sont connus pour présenter le plus haut degré d'organisation sociale (Sheldon, 1992). Il semble que leur structuration sociale renvoie à un système hiérarchique basé sur la dominance, via les statuts sociaux (Zimen, 1983 ; Fox, 1987). A l'état sauvage comme en captivité, lorsqu'une meute capture une proie ou est artificiellement nourrie, la position sociale détermine l'ordre des visites à la nourriture (Fox, 1987; Vila et al. 1990). Les membres dominants du groupe, souvent le couple reproducteur, mangent les premiers et davantage que les autres. Suivent les animaux de rangs intermédiaires, puis les subordonnés, les juvéniles et enfin les individus de bas rang (omega), écartés de la meute.

Bien que le concept de statut social soit largement utilisé dans la littérature éthologique, et particulièrement chez les loups (Schenkel, 1947, 1967 ; Zimen, 1983 ; Fentress et al. 1987 ; Sheldon, 1992), il semble cependant quelque peu problématique. Par exemple, lorsque la situation alimentaire

est modifiée, en raison de l'introduction d'aliments nouveaux, de la modification du système distributeur de nourriture... (Menzel et Menzel, 1979 ; Katzir, 1983 ; Chamove, 1983 ; Brennan et Anderson, 1988) ou en raison du placement des animaux dans un nouveau lieu (Katzir, 1982 ; Fox, 1987), le classement initial n'est que rarement observé : les dominants participent exceptionnellement à l'exploration et l'investigation du nouveau. Selon Chamove (1983), ces observations tendent à montrer que l'ordre d'accès à l'événement nouveau dépendrait davantage de rôles d'explorateurs ou initiateurs qui seraient dévolus à certains individus, que du statut social des membres du groupe.

Si l'ensemble de ces travaux s'accorde pour remettre en cause le concept de statut social et lui substituer celui de rôle social, ils considèrent, en revanche, comme un *a priori* le concept de dominance. Tout comme celui d'agression auquel il est souvent associé, ce concept est clairement inférentiel, ce qui revient à dire que ni l'agression, ni la dominance ne sont observables. Il s'agit de concepts qui ont en quelque sorte valeur explicative. Ils regroupent, en un ensemble supposé cohérent, différents comportements (par exemple, attaquer, menacer, mordre...) et en excluent d'autres.

Tous les protocoles ci-dessus mentionnés présentent une particularité commune, celle de disposer les objets, dispositifs... nouveaux aux mêmes emplacements au cours des différentes sessions expérimentales. Sans doute cette procédure est-elle dictée par le souci scientifique de maîtrise des variables parasites. La chose est à ce point entendue que souvent, leur position n'est pas indiquée. Brennan et Anderson (1988), Fornasieri et al. (1990), par exemple, ne précisent pas la position des piles de nourriture ou des boîtes renfermant de la nourriture dans leurs études sur la compétition alimentaire en situation d'acquisition nouvelle de nourriture. Ainsi, est-il implicite que, chez les animaux vivant habituellement sur le lieu des observations/expérimentations, la localisation spatiale des conduites exploratoires puis alimentaires, et plus généralement du 'comportement', importe peu dans leurs réactions à la nouveauté et leurs interactions sociales concomitantes. Or, Brennan et Anderson (1988), après Richards (1974), observent que tous les animaux ne consomment pas la nourriture à l'endroit où ils la trouvent. On sait, d'autre part, que nombre d'espèces organisent spatio-temporellement leurs conduites. On reconnaît en œuvre les opérations cognitives de conceptualisation et de décontextualisation initiales puis de l'instauration d'un lien permettant de prendre en compte la relation entre comportement et référentiel spatio-temporel, entre activité et contexte.

Nous avons étudié le comportement d'une meute de huit loups européens confrontée à trois situations de changement de place de la nourriture. La meute étant nourrie en un lieu toujours identique, nous avons quantifié, trois séances de nourrissage durant, les rapports de dominance habituels à l'aide des indices suivants : ordre d'arrivée dans la zone de dépôt, fréquence des présences dans cette zone, ordre de consommation, fréquences des consommations puis établi des classements en fonction de ces indices. Les trois séances de nourrissage suivantes, nous avons fait déposer la nourriture dans une zone inoccupée par l'ensemble de la meute, à l'exception d'une louve omega, exilée à cet endroit. Les trois séances alimentaires suivantes, elle a été déposée dans la zone de repos des adultes et enfin, les trois fois suivantes, simultanément dans les zones alimentaires des adultes et des juvéniles. Nos données indiquent clairement qu'il existe des rapports de dominance différents selon l'emplacement de la nourriture. Dans toutes ces situations de changement de lieu de dépôt de la

nourriture, les animaux les plus favorisés sont ceux qui occupent habituellement les zones de dépôt ou les plus proches abords. Ainsi, quand la nourriture est déposée à l'emplacement habituel, les sub-adultes, dont la louve omega, sont les plus présents au point d'alimentation. Après le premier changement de place de nourriture, c'est-à-dire lorsqu'elle est déposée dans la zone de refuge de la louve omega, celle-ci conserve son rang de leader. Ces premières données s'accordent mal avec la littérature. A cela deux raisons : (a) elles placent en haut de l'échelle sociale un individu chassé de la meute par les reproducteurs ; (b) après modification du lieu de dépôt de la nourriture, elles ne montrent pas de changement dans la hiérarchie de dominance. Mais au cours des changements ultérieurs les sub-adultes sont très peu présents. Ils semblent être relayés par les adultes lors du second changement de place de la nourriture puis par les adultes et les juvéniles lors de la dernière situation expérimentale. Ce phénomène est moins flagrant pour ce qui est de la consommation, tous les individus parvenant à se nourrir quelle que soit la situation, mais la tendance va tout de même dans le même sens. On observe donc autant de 'hiérarchies' que d'investissements différents des lieux de dépôt de la nourriture. Qui plus est, on montre que si la nourriture est déposée dans des zones caractérisées par des activités sans rapport avec l'alimentation, elle est d'abord la cible de ces activités (elle est foulée dans des zones de passage, enterrée dans les zones post-alimentaires, elle sert de support pour se coucher dans les zones de repos...) avant d'être consommée. La dominance semble donc relative, c'est-à-dire fonction d'une forme de priorité spatiale. On proposera que la relation de l'animal au contexte ne fait pas varier, comme l'affirme la pensée classique, son comportement mais le détermine. On distinguera cette formule de la thèse environnementaliste (Skinner, 1971), selon laquelle le milieu (et non pas la relation sujet-contexte) a un double rôle de stimulation et de sélection.

Contextualisation du comportement et conceptualisation du contexte

Ce dernier énoncé instaure un nouveau type de relation entre l'animal, le comportement et le contexte. Il convenait de régler, en effet, le statut du contexte spatio-temporel d'apparition ou de réalisation du comportement étudié. En éthologie, ce contexte est le plus souvent considéré comme une des déterminations (exogènes) de la variabilité des données sensibles. L'éco-éthologie classique montre ce lien entre le comportement (comportement alimentaire, par exemple) sous ses formes apparentes diversifiées et les différents milieux où ces observations sont réalisées (topographie, superficie du milieu...). Dans ce cadre, le concept recouvre des réalités (a) dont la diversité n'est qu'accessoire car due à des différences contextuelles si l'on s'intéresse au comportement comme révélateur de processus psychologiques fondamentaux ou comme l'expression de mécanismes physiologiques généraux ; (b) dont la diversité joue un rôle important si l'on s'intéresse aux causes ultimes (théorie évolutionniste) comme cible de la sélection naturelle. Assurément, des facteurs endogènes (détermination génétique, état psychophysiologique...) interviennent et co-agissent avec les facteurs exogènes. Mais, dans tous les cas, la réalité comportementale est décrite comme unitaire (tombant sous un concept), c'est-à-dire rebelle à toute division par des événements que l'on peut globalement définir comme contextuels. Quand l'étude porte sur l'organisation spatio-temporelle des activités, elle consiste à établir des relations non pas entre l'animal et le contexte, mais entre les séquences comportementales de cet animal et le contexte de leur manifestation. Si l'on y pense, cette

conception du rôle des événements contextuels, générale en éthologie, renvoie peut être davantage au sens commun qu'à l'exercice d'une activité réflexive de type scientifique - auquel cas, le comportement ne serait encore qu'une notion et non pas un concept. On reconnaîtra, ici, la distinction chère à Bachelard entre la pensée du sens commun utilisant des notions polysémiques et contaminées par des images et la rationalité scientifique d'un langage discursif où les concepts, constructs monosémiques, se sont dégagés de leur gangue originale.

On renverra, par exemple, à l'histoire des travaux sur la chute des corps en physique pour étayer l'idée que le contexte ne se traite pas comme un facteur de diversité, ni comme le substrat d'une séquence de phénomènes. En ce sens, on rejoint la nécessité, soulignée par Cassirer (1977), de reprendre le schéma couramment admis pour la conceptualisation et d'en remanier jusqu'à la formulation, car il amalgame sans nuances et met sur un même plan les propriétés qui expriment le contenu même des choses (possession d'un même attribut, par exemple) et les aspects purement relationnels. Sous une nouvelle forme, la théorie de la conceptualisation doit donc contenir des traits qui exigent d'être fondés sur d'autres bases. Car la notion de substance (réalité du sujet concret par opposition à la réalité de la synthèse conceptuelle) a préséance logique dans la thèse classique. Seules les substances dont l'existence est donnée peuvent servir de référence et de support à la diversité des déterminations de l'être. Seul un donné présentant toutes les garanties de consistance objective et imposant sa présence immédiate peut offrir aux variétés logico-grammaticales qui spécifient l'être le point d'appui et le soutien dont elles ont besoin. Déterminations spatiales et temporelles n'ont pas de subsistance en soi et pour soi ; elles ne valent qu'en leur qualité de propriétés. Par exemple, si l'on considère que, pour un animal, le lieu où il s'alimente ne fait pas partie (effet de contexte) du comportement alimentaire, on construira des répertoires comportementaux composés d'items désignant des actions 'pures', indépendantes de leur contexte spatio-temporel (flairer, manger, dormir, se toiletter...). Or cette division classique entre substance et arrière-plan occulte la réalité du comportement si le fait de s'alimenter ici et là n'a pas la même signification pour l'animal. Prendre en compte cette dernière hypothèse oblige à réviser les unités de description et à remplacer systématiquement les verbes par des locutions adverbiales reliant verbe et adverbe, c'est-à-dire un comportement et son contexte (de Gaulejac & Gallo, 1996). Cependant, la multiplication des locutions adverbiales ne constitue qu'une étape descriptive ("monde de description de l'observateur", Maturana, 1987) bien qu'essentielle. Le foisonnement qui en résulte interdit qu'on s'y soumette. Une autre étape reste à franchir, liant de telle manière l'animal et le contexte spatio-temporel, que les séquences comportementales observées n'en soient plus qu'une conséquence. L'objectif serait d'énoncer la loi qui régit la relation de l'animal au contexte spatio-temporel (loi d'appropriation) et non pas les règles qui régissent la succession des événements (comportements, items). La prise en compte du contexte dans sa relation à l'être que l'on observe constituerait donc le fondement de l'analyse comportementale ultérieure. Dans cette conception, l'animal prend plutôt statut de système autonome (Varela, 1993) que de système à input. Il est vraisemblable que cette problématique soit aussi au cœur de la psychologie humaine : la contextualisation des comportements permet de décaler la causalité propre à une psychologie des fonctions vers une "psychologie" où les déterminants seraient à rechercher dans un contexte devenu essentiel, celui de la parole des autres.

Conclusion

On ne peut sortir de notre domaine cognitif. Pour autant, cela ne nous enferme pas au point de ne nous laisser aucun choix heuristique. La contrainte d'une pratique des mêmes opérations cognitives (conceptualisation), que l'on traite ou non d'un autre être que nous, ne nous conduit pas nécessairement à conserver éternellement les cibles initiales. La subversion épistémologique, ici, si elle existe, aura consisté à intervertir ces cibles : à la conceptualisation du comportement, on vise à substituer la conceptualisation du contexte.

Références

- Andler D. 1992. Introduction aux sciences cognitives. Paris : Folio Essais.
- Berlyne D.E. 1960. Conflict, arousal and curiosity. London : McGraw-Hill.
- Brennan J. & Anderson J.R. 1988. Varying responses to feeding competition in a group of rhesus monkeys (Macaca mulatta). Primates, 29, 3, 353-360.
- Cassirer E. 1977. Substance et fonction : éléments pour une théorie du concept. Paris : éditions de Minuit.
- Chamove A.S. 1983. Role or dominance in Macaque response to novel objects. Motivation and Emotion, 7, 2, 213-228.
- de Gaulejac F. & Gallo A. 1996. La description en éthologie : du mot à la locution. Communication au troisième Congrès Européen de Systémique. Rome, Octobre 1996.
- Fentress, J.C., Ryon, J., McLeod, P.J., & Havkin, G.Z. (1987). A multidimensional approach to agonistic behavior in wolves. In Man and Wolf, H. Frank (Ed.), pp. 253-273, Dr W.J. Publishers, Dordrecht, The Netherlands.
- Fornasieri, I., Anderson, J.R., & Roeder, J.J. (1990). Responses to a novel food acquisition task in three species of lemurs. Behavioural Processes, 21, 143-156.
- Fox M.W. 1987. Behaviour of wolves, dogs and related canids. Malabar: Krieger Publishing Company.
- Herrnstein R.J. 1984. Objects, categories and discriminative stimuli. In Animal Cognition, H.L. Roitblat, T.G. Bever, H.S. Terrace (Eds), Hillsdale, Lawrence Erlbaum Associates.
- Jensen P. 1994. Fighting between unacquainted pigs - effects of age and of individual reaction pattern. Applied Animal Behaviour Science, 41, 37-52.
- Katzir G. 1982. Relationships between social structure and response to novelty in captive jackdaws, Corvus monedula L., I. Response to novel space. Behaviour, 81, 231-263.
- Katzir G. 1983. Relationships between social structure and response to novelty in captive jackdaws, Corvus monedula L., II. Response to novel palatable food. Behaviour, 87, 183-208.
- Latour B. (19XX). Sens et place des connaissances dans la société.
- Lestage P. 1994. Complexité et structures cérébrales : Piaget et Prigogine. In Les défis de la complexité, D. de Béchillon (Ed.), Paris : l'Harmattan.

- Menzel E.W. & Menzel C.R. 1979. Cognitive, developmental and social aspects of responsiveness to novel objects in a family group of Marmosets (Saguinus fuscicollis). Behaviour, 70, 3-4, 251-279.
- Maturana H. (1987) Communication et Représentation, in J. Piaget, P. Mounoud, & J.P. Bronckart (Eds), Psychologie (pp. 1442-1475), Paris : Encyclopédie de la Pléiade, Gallimard.
- Nakanishi Y. Mutoh Y. & Umetsu R. 1992. Social relationship and spatial distribution in a small herd of Japanese black cattle in a dry-lot. AJAS, 5, 1, 183-188.
- Pichot A. 1980. Eléments pour une théorie de la biologie. Paris : Maloine.
- Racah P.Y (sous presse). Argumentation and knowledge : from words to terms. Actes du Colloque Languages of Science. Bologna; Octobre 1995.
- Richards S.M. 1974. The concept of dominance and methods of assessment. Animal Behaviour, 22, 914-930.
- Sheldon J.W. 1992. Wild dogs: the natural history of the nondomestic canidae. San Diego: Academic Press, Inc, Harcourt Brace Jovanovich, Publishers.
- Schenkel R. 1947. Ausdrucks-Studien an Wölfen. Behaviour, 1, 81-129.
- Schenkel R. 1967. Submission: its features and functions in the wolf and dog. Am. Zool., 7, 319-329.
- Skinner B.F. 1972. L'analyse expérimentale du comportement. Bruxelles : C. Dessart.
- Sokolov E. 1963. Perception and the conditioned reflex. New York: Pergamon.
- Vankov A. Hervé-Minvielle A. & Sara S.J. 1995. Response to novelty and its rapid habituation in locus coeruleus neurons of the freely exploring rat. European Journal of Neuroscience, 7, 1180-1187.
- Varela F.J. Thompson E. & Rosch E. 1993. L'inscription corporelle de l'esprit. Sciences cognitives et expérience humaine. Paris : Seuil.
- Vila C. Urios V. & Castroviejo J. 1990. Possible not male dominated interaction among Iberian wolves (Canis lupus). Mammalia, 54, 2, 312-314.
- Zimen E. 1983. Social dynamics of the wolf pack. In The wild canids, M.W. Fox (Ed.), pp. 336-362, Malabar: Krieger Publishing Company.

Désigner et faire évoluer un objet en conception distribuée

Nicolas GREGORI

Groupe de Recherches sur les Communications
Université de Nancy 2
BP 33-97 - 54015 NANCY CEDEX
gregori@clsh.u-nancy.fr

Résumé : Au cours de ce travail nous allons nous intéresser à une expérience de conception distribuée. En tant que psychologue social spécialisé dans l'étude des interactions, nous observerons le processus d'intercompréhension interhumaine qui préside à cette activité collective. Nous verrons comment les acteurs font pour désigner l'objet qu'ils ont à concevoir et comment ils font pour s'entendre et progresser dans le processus de conception alors qu'ils ont des représentations différentes de cet objet.

Mots-clé : conception distribuée, intercompréhension, espace intersubjectif, objet intermédiaire.

INTRODUCTION

Il est des situations où l'on s'entretient d'objets encore indéfinis, incertains. Tel est notamment le cas de l'activité de conception puisque l'objet fini n'a d'existence effective qu'une fois cette activité achevée. Tant que dure le processus de conception, les concepteurs s'entretiennent à propos de quelque chose d'instable mais dont les fonctions générales sont relativement invariantes. Les acteurs ont alors à réaliser un travail de co-référenciation fait de multiples négociations qui apparaissent et se règlent dans l'interaction. L'objet en gestation passe par différents stades et l'on peut suivre son ontogenèse en pistant aussi bien les traces laissées sur des supports physiques -les notes et dessins réalisés sur papier par exemple- comme le font Blanco *et al.* (1996), que les traces langagières des concepteurs qui surgissent au cours de leurs interactions. C'est ce dernier travail qui nous occupe plus particulièrement, c'est-à-dire l'analyse pragmatico-linguistique des dires des participants, même si notre collaboration avec le groupe de chercheurs cité plus haut nous amène également à considérer les "objets intermédiaires", nom donné aux traces qu'ils étudient.

La situation dite de « Conception Distribuée » qui fait l'objet de notre travail prend place dans un paradigme apparu il y a une dizaine d'années aux Etats-Unis et dénommé « Ingénierie Concourante ». Il s'agit de prendre le contre-pied des méthodes classiques de conception -lesquelles suivent un déroulement séquentiel : un secteur d'activité agit, puis un autre, puis un autre, etc.- pour développer un modèle qui soit, lui, parallèle ou en réseau (Garro, 1996).

En conception, on est donc amené à discuter, et mieux même, à s'entendre à propos d'un objet structurellement et fonctionnellement instable, dont la dénomination évolue au cours du temps de manière plus ou moins subreptice. Malgré cela, les concepteurs ont pour obligation de se comprendre les uns les autres sous peine d'échec. Comment cela est-il possible ? Comment cela se produit-il ? C'est à ces questions que nous, psychologue spécialiste des interactions, allons essayer de répondre dans cet article.

Pour le faire, nous procéderons de la sorte. Dans un premier temps, nous dirons quelques mots sur le modèle de conception distribuée que nous défendons et sur la méthode d'analyse que nous mettons en œuvre pour montrer comment opèrent les mécanismes d'intercompréhension entre les individus aux prises. La seconde partie sera l'occasion de nous intéresser à la façon qu'ont eu les concepteurs de s'entendre -ou se "mésestimer"- pour développer cet objet qu'ils appellent « plots » et comment celui-ci a acquis structure et forme.

1. ANALYSE D'UNE SITUATION DE CONCEPTION

1.1. Le modèle de « Conception Distribuée »

Ce n'est que récemment, une cinquantaine d'années, que des méthodes de conception ont été élaborées (Garro, 1996). Cette activité est alors modélisée de façon séquentielle. Un Bureau d'Etudes analyse la fonctionnalité du système à concevoir, transmet le fruit de sa réflexion au Bureau des Méthodes qui produit une analyse structurelle et envoie à son tour le résultat de son travail à l'atelier de fabrication qui va réaliser le produit. Voilà grossièrement résumée la chaîne de conception d'un produit. Dans ce modèle on peut observer quelques rétroactions depuis un secteur d'activité vers un autre, mais ces remarques, en tout état de cause, ne peuvent être formulées qu'une fois réalisé le travail de l'entité antérieure.

Imaginer une ingénierie concourante, c'est intégrer de l'interactivité au sein du processus de conception en faisant en sorte que les experts des différents secteurs d'activités puissent nourrir le processus de leurs compétences à tout moment. Il s'agit pour les défenseurs de ce paradigme de déployer des réseaux de communication entre les différents acteurs responsables de la conception d'un produit. Donner une telle orientation aux nouvelles méthodes de conception, c'est reconsidérer la complexité des objets à concevoir, lesquels objets allient souvent diverses technologies : informatique, électronique, mécanique, etc. C'est aussi permettre à la créativité de s'exprimer tout au long du processus de conception (Charpentier *et al.*, 1995).

C'est dans ce nouveau paradigme de recherches que s'inscrit le projet « Conception Distribuée ». La distribution autour de l'activité de conception est rendue par la coopération d'acteurs spécialisés dans les secteurs fonctionnel, structurel et de fabrication et qui interviennent chacun dès le début et tout au long de l'expérience de conception. Les trois modules ont chacun des objectifs locaux mais, reliés, ils forment un système ayant pour objectif global de concevoir un dispositif mécanique de mise en position et de maintien de pièces planes en bois pendant leur usinage (un contournage). La séance s'est déroulée sur deux demi-journées de trois heures chacune, temps imparti par avance, et a été enregistrée en vidéo. Précisons dès maintenant que notre analyse ne porte que sur la première demi-journée.

1.2. Analyser l'intercompréhension interhumaine

Au cours de cette partie, nous allons présenter très succinctement la méthodologie que nous suivons pour analyser la situation de conception. Il s'agit d'une méthodologie pragmatique-

linguistique liant les principes empiriques de l'analyse conversationnelle des ethnométhodologues à ceux formels de la théorie des actes de langage dans les versions axiomatisées de leurs conditions de succès (Searle et Vanderveken, 1985) et de satisfaction (Vanderveken, 1988). Autrement dit, nous observons *in situ*, comme le prône l'ethnométhodologie, les réalisations illocutoires des interactants aux prises. En réagissant les uns aux autres, ceux-ci rendent manifestes leurs interprétations réciproques, ils disent la compréhension qu'ils ont de ce qui s'échange : connaissances, intentions, représentations, états-mentaux, etc.

Nous rendons compte de cette activité en analysant l'articulation entre réussite et satisfaction des actes de langage ce qui nous permet de mettre en lumière les états du monde qui circulent, ainsi que les rapports sociaux qui s'instaurent, dans l'espace intersubjectif créé par les acteurs en présence. Notre travail est un travail de description de cet espace intersubjectif rendu observable et intelligible par les accomplissements successifs des interlocuteurs. Cette méthodologie pragmatique, ainsi que son application en psychologie sociale, est longuement décrite dans Trognon et Brassac (1992), Trognon et Ghiglione (1993), Trognon et Larrue (1994) et Brassac et Stewart (1996). Pour son application en conception distribuée, voir Grégori (1996).

2. DE L'IN-FORMATION A L'EN-FORMATION DES PLOTS

« La fabrication des faits et des machines est un processus collectif » affirme Latour (1989, p. 79). L'objectivité est une intersubjectivité renchérit Vinck (1995 p. 90). Il faut donc s'intéresser aux conditions de production des faits et des objets scientifiques et non pas seulement aux résultats auxquels on aboutit. Nous avons envie de reprendre cette revendication à notre compte en l'appliquant à l'activité de conception, convaincu que nous sommes que c'est dans les formulations et reformulations successives, dans l'émergence des malentendus, dans les modifications continues que se forme l'objet de conception. Si l'on considère le modèle de conception distribuée, alors il faut accepter que l'évolution de cet objet n'est pas le fruit d'une succession de transmissions de l'objet d'un acteur vers un autre mais qu'elle est le reflet d'une composition des acteurs. On décrira ce processus en contant l'histoire d'un objet qui surgit d'une sorte de magma primitif, d'un bouillon originel où les significations sont vagues et se recoupent et qui, en grandissant, se précise jusqu'à devenir autonome, c'est-à-dire structurellement déterminé.

2.1. L'émergence des plots du « bouillon originel »

Rappelons la tâche des trois modules. Il leur faut concevoir un objet mécanique permettant de mettre en position et de maintenir des pièces planes en bois dont il s'agit d'usiner les contours. Cet objet prend place sur une machine spécialisée : une défonceuse. Nous nous intéresserons au cours de cette partie à la première phase de conception, d'une durée de

cinquante minutes environ, au cours de laquelle les modules découvrent la commande du client puis discutent des fonctions à remplir par l'objet¹.

Dès le document fourni par le client en guise d'ébauche du cahier des charges fonctionnel et dès les premiers mots proférés par le module fonctionnel décrivant ce document, le terme « plots » apparaît pour désigner l'objet à concevoir.

[1] FCT1² on va essayer de concevoir globalement euh des:³ des plots ou des montages d'usinage pour une machine spécialisée dans l'usinage euh des pièces en bois [...]
[1ère mn]

En fait, dans [1], on voit apparaître deux des phénomènes qui nous occupent dans cette partie. Premièrement, comme nous l'avons déjà annoncé, c'est le terme 'plots' qui est d'abord utilisé pour désigner l'objet mécanique à concevoir. Or, ce qui sera conçu au final, ce sont également des plots. La référence aura bien sûr évolué, mais le terme aura subsisté. 'Plot' est donc une dénomination-clé dans cette situation. Deuxièmement, on observe qu'immédiatement après le profération du terme plots, FCT1 opère une rectification ou propose une alternative, l'analyse du "ou" ne permettant pas de décider à cet instant (est-ce un « ou plutôt » ou alors un « ou bien »?). Toujours est-il que le terme 'montages d'usinage' entre en scène et qu'il se trouve juxtaposé à 'plots'. La suite de cette partie introductive, au cours de laquelle le module fonctionnel présente la commande du client à ses partenaires, nous incite à penser que pour le moment 'plots', 'montages d'usinage' et 'système', autre terme utilisé, réfèrent au même objet et sont substituables l'un à l'autre comme le montre [2].

[2] FCT1 sur quoi on fixe nos plots? ben ben y a y a la structure de la table donc euh y faut penser à un système qui viendrait se placer sur la structure de la table
STR c'est quoi la structure de la table ?
FCT1 euh c'est c'est une structure en en tôle euh [...].
[13ème mn]

Sans que cela pose problème, ni ne soulève de remarque, FCT1 glisse d'un terme à l'autre en [2], signifiant implicitement par là leur synonymie. En effet, ce n'est pas à propos de ce glissement qu'intervient STR quand il met en cause la satisfaisabilité de l'acte de FCT1, c'est un autre élément du contenu propositionnel de cet acte (la structure de la table) qu'il interroge. Il est cependant de rares cas [3] où FCT1 différencie 'plot' et 'système', ce dernier terme étant annoncé comme neutre dans la désignation de l'objet à concevoir tandis que le premier représenterait un type particulier de technologie pour solutionner le problème. Malgré cette précision, on observe que, dans la suite immédiate de l'interaction, 'plots' réapparaît utilisé par STR et "accepté" par FCT1.

¹ Par commodité et pour ne pas créer d'ambiguïté entre notre façon de désigner le dispositif à concevoir et celle pratiquée par les concepteurs, nous utiliserons systématiquement le terme « objet » en référence à ce dispositif.

² Le module fonctionnel est interprété par deux acteurs notés FCT1 et FCT2. De même, nous utiliserons l'abréviation STR pour désigner le module structurel et FAB pour celui de fabrication.

³ Par convention de transcription des discours oraux, les deux points [:] accolés à une syllabe indiquent que le locuteur a la voix qui traîne sur celle-ci.

- [3] FAB on cherche à faire un seul plot là ou plusieurs types de plot ?
 FCT1 ah ben enfin là on on s'est un peu axé déjà par par la solution mais c'est un système de maintien qui globalement peut tenir 5 à 6 pièces en parallèle sur la machine
 FAB oui
 STR y en a plusieurs euh d'autant plus que les surfaces euh [*de pièces*] vont de 50 par 50 euh
 FCT1 si on arrive à trouver un système qui peut tenir dans un coin une pièce et pas dans l'autre, une pièce quelles que soient les dimensions
 STR bon dans les contraintes liées à l'environnement tu nous parles des 277 millimètres maxi euh du côté de la broche et tu nous dis que ta hauteur de pièce est de 170 millimètres
 FCT1 ouais tout à fait
 STR le plot, y fait combien là-dedans ?
 FCT1 le plot est pas prévu [...].
 [8ème mn]

En réagissant à la requête de FAB, FCT1 en livre les conditions de satisfaction. Dans un premier temps, il indique que formulée telle qu'elle l'est, la requête de FAB est biaisée. Et ceci parce qu'elle présuppose un certain état de choses : l'objet à concevoir est un plot. C'est donc la réussite de la requête qui est menacée ici. Cependant, FCT1, dans un deuxième temps, satisfait quand-même la requête de FAB en substituant 'plot' par 'système de maintien' et en précisant qu'il s'agit d'un seul système pour plusieurs pièces. Ainsi, en menaçant un instant la réussite de la requête de FAB du fait d'un présupposé maladroit mais en satisfaisant tout de même cette requête en en précisant les conditions de satisfaction -ce qui rend la requête réussie-, FCT1 a-t-il indiqué le fait que 'système de maintien' lui paraît être un terme neutre, contrairement à 'plot' qui, lui, réfère à un type de solution technologique particulière. Par la suite, FAB se contente d'acquiescer, ce qui ne nous permet pas de savoir s'il est engagé sur la ratification de l'explication fournie par FCT1, STR réagit d'une façon ambiguë sur laquelle nous allons revenir et FCT1 précise le rapport entre 'système' et pièces, répétant alors de façon implicite la distinction à faire entre 'plot' et 'système'.

Cependant, la dénomination 'plot' réapparaît immédiatement sans plus faire l'objet d'une controverse. Des plots, un système, un montage, quel que soit le terme utilisé, la conception de l'objet progresse. Les acteurs s'entretiennent du nombre de pièces à placer sur la table, de la hauteur de l'objet, de la structure sur laquelle il faudra le fixer. Ils discutent de son aspect modulaire, du système de dépression (c'est-à-dire du système de maintien de l'objet sur la table et des pièces sur l'objet). Ils en détaillent les fonctions. Et pourtant, même s'ils réussissent cela, on peut se demander s'ils parlent bien du même objet, s'ils partagent une même représentation.

On a dit que les trois termes fonctionnaient la plupart du temps de façon synonymique, voir [2] et [3], et que l'on relevait parfois des distinctions, voir [3]. Mais il est un autre phénomène important à mettre en lumière, qui apparaît dans [3] mais est plus évident dans [4].

- [4] STR donc là c'est figé: ça serait des systèmes qui seraient modulaires
 FCT1 euh enfin j'ai rien dit quant à la mod' j'ai rien dit quant à la modularité du système pour l'instant
 STR bon en tout cas ça serait des systèmes euh séparés, enfin y aurait euh
 FCT1 enfin: enfin tout dépend comment comment on va le concevoir mais pour l'instant euh rien ne dit qu'y peut être séparé
 STR d'accord.
 [19ème mn]

Ce que nous tenons à souligner ici, c'est la façon qu'ont STR et FCT1 de désigner l'objet à concevoir. En [3], STR parle de *plusieurs* 'plots' ou 'systèmes'. Son intervention répond à celle de FAB tout en intégrant la réaction de FCT1, d'où la difficulté de décider à quoi réfère « en ». Difficulté et même impossibilité puisque STR utilise, immédiatement après, le terme 'plots' malgré la remarque de FCT1. Poursuivons. FAB demandait s'il fallait concevoir un ou plusieurs types de 'plot'. STR répond plusieurs, or le voilà maintenant qui parle *du* 'plot'. On glisse donc de 'système' à 'plots', du pluriel au singulier, ce qui apparaît nettement dans [4]. Ici, STR désigne l'objet par le terme 'systèmes' au pluriel tandis que FCT1 l'utilise au singulier. L'aspect modulaire ou séparé renvoie aux « plusieurs plots/systèmes » indiqués en [3] par STR, alors que FCT1 préfère ne pas s'engager sur cet aspect de la conception, qui ne peut actuellement faire l'objet d'une décision, en ne satisfaisant pas par deux fois les tentatives d'assertion de STR. Malgré tout, ceci nous renseigne sur les représentations divergentes des acteurs. Pour l'un, STR, l'objet semble se fractionner en plusieurs objets, pour l'autre, FCT1, l'objet est une sorte de gros objet structurellement indéfini, et qu'il faut sûrement se garder de préciser trop. Toutefois, le déroulement de l'interaction démontre que les acteurs sont performants dans leur activité de conception puisqu'ils arrivent à un accord. Loin d'être un frein (le fait que l'un parle *des* systèmes et l'autre d'*un* système n'est pas discuté), cette apparente mésentente serait plutôt un moteur pour l'action des concepteurs. Après tout, en ne relevant pas la différence de représentation, ceux-ci progressent dans le processus en s'entretenant d'un objet à la fois singulier et pluriel, ce qui n'est certainement pas un désavantage en terme de potentiel de créativité.

Durant la première demi-heure d'où sont extraites les séquences [1], [2], [3] et [4], les acteurs s'approprient donc la tâche à réaliser et construisent les premières données techniques. Le vocabulaire qu'ils utilisent les amène à parler indifféremment, ou presque, de 'plot(s)', de 'système(s)' ou de 'montage(s)'. Et ceci alors que le terme plots renvoie déjà à une certaine représentation de l'objet à concevoir, comme le signale FCT1 en [3]. Pour expliquer cela, nous faisons l'hypothèse que les plots n'ont pas encore été structurellement définis, ni dans le discours, ni sur le papier. Nous verrons par la suite que cette hypothèse semble confirmée par le fait qu'une esquisse sur papier de la forme et du fonctionnement général des plots suffit à créer une irréversibilité telle que 'plots' ne sera plus équivalent à 'montage' ou 'système'.

Vers la trentième minute, nous en sommes rendus au point suivant. L'objet à concevoir n'a pas encore de forme, il est in-formé. Une nouvelle phase est alors entamée par les acteurs, principalement FCT1 et FCT2, au cours de laquelle il est discuté des fonctions précises à remplir par l'objet. Celui-ci est alors presque exclusivement désigné par 'système'. On dénombre en effet 32 occurrences de ce mot pour désigner l'objet contre 7 pour 'montage' et une seule pour 'plots'. C'est là un effet massif - dans la première phase l'objet est désigné 23 fois par le terme 'système', 20 fois par 'montage' et 20 fois par 'plots' - que l'on peut expliquer si l'on se souvient qu'en [3], FCT1 avait utilisé ce terme système pour neutraliser son discours en termes

d'options technologiques pointant derrière l'utilisation de 'plots' par FAB. Autrement dit, nous sommes là dans une phase qui réclame précision et bonne entente. 'Système', terme plus neutre que 'plots' du point de vue des options technologiques, est alors le terme utilisé par les concepteurs, comme le montre la figure 1 qui représente un des objets intermédiaires -l'objet 2⁴- réalisés en situation par les acteurs et analysés dans Blanco *et al.* (1996).

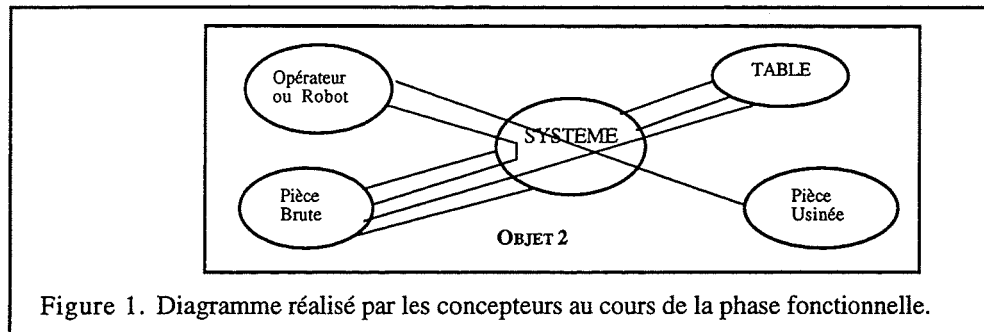


Figure 1. Diagramme réalisé par les concepteurs au cours de la phase fonctionnelle.

Nous arrivons maintenant au terme de cette première partie, terme marqué par le fait que STR, à la suite du travail fonctionnel, propose trois types de solutions technologiques. Pour résumer ce qui s'est passé au cours de ces cinquante premières minutes, nous dirons que les plots, en tant qu'objet à concevoir, ont commencé d'exister bien qu'ils ne soient encore ni structurés ni consensuels. La désignation 'plots' est toujours indistincte des termes 'système' ou 'montage'. Indistincte, pas tout à fait, car en quelques endroits son utilisation, ou sa non utilisation, lui a conféré une identité en devenir et qui va s'affirmer maintenant.

2.2. Evolution des plots vers l'autonomie

Cette deuxième partie ne fera pas l'objet d'un développement aussi long que la première. Ce qu'il nous importe de montrer, c'est i) comment, à partir du stade précédent, les plots encore informés acquièrent une structure propre leur donnant une autonomie, et ii) ce que cette évolution a de distribué sur les concepteurs, autrement dit en quoi cette autonomie est le produit d'une enformation, c'est-à-dire d'une mise en forme dans un espace intersubjectif.

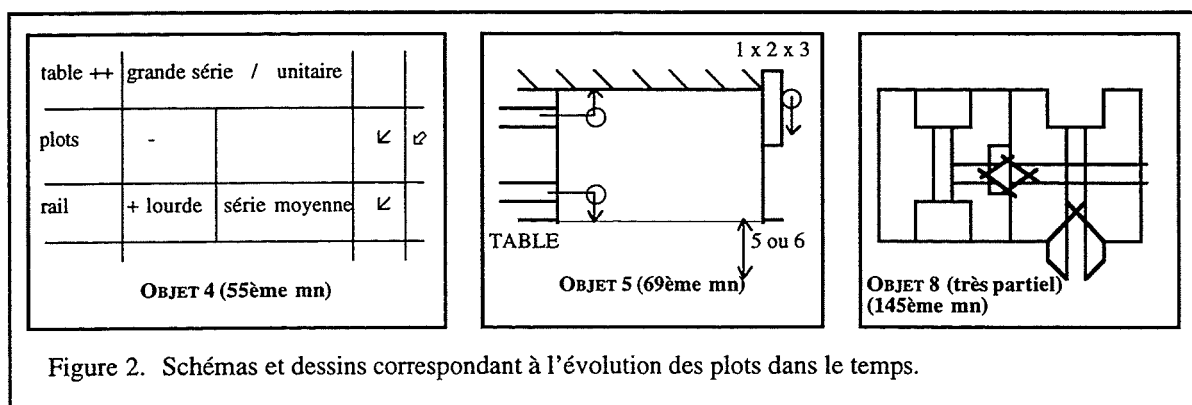


Figure 2. Schémas et dessins correspondant à l'évolution des plots dans le temps.

⁴ Il ne s'agit pas du véritable objet 2, réalisé à la main par les acteurs et comportant des annotations que nous n'avons pas reproduites ici, mais d'une représentation simplifiée. Il en sera de même pour les objets 4, 5 et 8 de la figure 2.

En deux temps, que Blanco *et al.* notent les objets 4 et 5 (figure 2), la référence 'plots' évolue et acquiert une première représentation physique. Ceci est capital car c'est la marque d'une articulation dans le processus de conception. L'objet, en se matérialisant sous l'action de ses concepteurs, reflète d'abord un choix à faire entre trois solutions possibles (objet 4) puis prend forme (objet 5) créant alors une sorte d'irréversibilité.

Dans un premier temps, STR propose trois solutions possibles. La première appelée « table » consiste en une amélioration de la table actuelle, la deuxième appelée « plots » où il s'agit de disposer des plots n'importe où sur la table afin de pouvoir mieux disposer de l'espace et la troisième appelée « rail » où il s'agit de disposer des plots sur des guides de translation, des rails. Ici, nous sommes encore dans une phase intermédiaire entre la partie fonctionnelle qui a précédé et la partie structurelle qui s'annonce. Cependant, en discutant pour la première fois de solutions possibles, les concepteurs opèrent une rupture dans le processus de conception puisque, désormais, le terme 'plots' renvoie à quelque chose de tangible. Ceux-ci réfèrent en effet à la fois à un type spécifique de solution et sont aussi des éléments intégrant les trois solutions. Ils peuvent être des 'plots fixes' dans le cas de la solution « table » ou des 'plots mobiles' pour les solutions « plots » et « rail ». Cette double référence est la source d'une ambiguïté pour FAB, comme il l'exprime dans [5]. FCT1 y apporte alors une réponse.

- [5] FAB bon j'ai pas bien compris alors c'était une table améliorée parce que pour moi euh mettre des plots partout c'est la solution à tout
 FCT1 ben c'est des plots mais fixes, y en a un c'est des plots fixes et l'autre c'est des plots mobiles en gros quoi
 [65ème mn]

Par la suite, on observe à nouveau une évolution de la référence plots dans le processus de conception (objet 5). C'est STR qui réalise ce dessin. Nous avons montré ailleurs (Grosjean *et al.*, 1996) le rôle essentiel de ce moment. Trois solutions ont été envisagées et la discussion qui s'en est suivie n'a pas permis de faire un choix explicite même si une préférence pour la solution « plots » semblait de dégager. En dessinant un plot tout en attirant l'attention de ses partenaires par l'intervention « *bon maintenant imaginez que j'ai(e) un plot* », STR s'engage, en même temps que les autres se laissent engager, plus avant dans le processus de conception des plots. Le principe général en est alors défini avec une aspiration dirigée vers la table et une autre vers les pièces et la mise en position des pièces est réalisable au moyen de butées rétractables (une, deux ou trois). Après avoir été définis comme 'plots fixes' vs 'plots mobiles', à présent les plots sont équipés ou non de butées. Les plots sont ainsi des objets munis d'un certain mécanisme et forment un système général de mise en position et de maintien des pièces à usiner. Les termes 'plots' et 'système' ne sont définitivement plus équivalents dans leur capacité à désigner l'objet à concevoir puisque cet objet devient un système composé de plots. En acquérant un mode de fonctionnement, les 'plots' prennent de l'autonomie vis-à-vis des entités 'système' et 'montage'. Les acteurs n'utiliseront dorénavant plus 'plots' et 'système' ou 'montage' dans le même registre. Et cela sans s'être mis d'accord pour le faire. C'est l'évolution technique des plots réalisée dans l'interaction qui incite à le faire.

Les plots, cependant, ne sont pas encore tout à fait définis. Un choix se pose entre réunir les trois butées nécessaires à la mise en position des pièces sur un même plot ou assembler plusieurs plots dont trois sont équipés de butées. 'Plots', 'butées', 'plots avec butées', 'plots sans butée', la difficulté de se représenter tout cela et la confusion s'installent. Si bien que le besoin de donner une définition au terme 'plots' -et ceci pour la première fois après plus d'une heure et quart de discussion- se fait sentir [6].

- [6] FCT2 t'as des plots qui sont uniquement des plots et des plots qui servent à faire des butées
 FCT1 alors ouais alors est-ce que t'as des plots, 'fin qui n'sont pas des plots mais qui sont des butées et qui peuv' être indépendamment, indépendants des plots parce que tu: tu peux promener tes:
 FAB eh: tu nous la refais celle-là!
 FCT2 doucement
 FCT1 non mais: ton plot, alors quand j'entends plot ça veut dire qu'on peut mettre la pièce dessus
 STR ouais
 FCT1 ok? t'entends bien exactement la même chose?
 STR mm-mmh ouais
 FCT1 c'est pas un plot qui permettrait pas d'mettre euh qui permettrait pas d'mettre la: la pièce dessus quoi.
 [76ème mn]

Voyons comment apparaît le besoin de définir 'plot'. FCT2 vient de préciser que certains plots servent également de butées et d'autres non. FCT1 se sert de cette assertion (*alors*) comme prémisses à une question portant sur une éventuelle indépendance des butées vis-à-vis des plots. Cette requête d'information est alors suspendue par FAB et FCT2 qui en mettent l'intelligibilité en cause. Cette question n'est dès lors pas réussie. Comme cela est l'usage quand l'inintelligibilité d'un acte est déclarée par l'interlocuteur, le locuteur réitère, reformule cet acte (Goffman, 1973). C'est ce que commence de réaliser FCT1 avant d'entamer un processus définitoire du terme plot. Ce faisant, il initie un échange subordonné -qui court sur cinq tours de parole- au cours duquel FCT1 et STR vont assurer leur bonne entente et dont la fonction est d'assurer la bonne complétude interactionnelle de l'acte en jeu pour tenter d'en assurer la réussite conversationnelle (Roulet *et al.*, 1985), en l'occurrence désambiguïser au maximum les termes utilisés.

Cette précision faite, le processus de conception poursuit son cours. Et le choix concernant plots et butées est arrêté plus loin par l'intermédiaire de l'objet 8 (figure 2). Nous n'avons représenté que très partiellement cet objet. De nombreux autres dessins y figurent et montrent la progression vers celui que nous avons figuré. Blanco *et al.* font une analyse très détaillée de cet objet car il représente une rupture dans le processus de conception. Il est un intermédiaire entre l'objet 5 qui était un schéma fonctionnel et les objets suivants qui sont, eux, des dessins techniques.

Les plots sont donc maintenant parés d'un système de fixation à la table et de fixation entre eux. C'est en effet l'option assemblage de plots qui a été choisie. L'objet 8 représente d'ailleurs deux plots accolés l'un à l'autre. Par la suite, la technologie des plots s'affine en même temps que les dessins se font plus précis. D'abord convoqués en tant que système indéfini durant la première phase, puis en tant que type spécifique de solution, les plots ont acquis une structure telle qu'ils sont maintenant parfaitement distincts d'autres entités telles que système, montage ou même butées, entité apparue au cours du processus de conception. Mieux, les plots sont divers,

comportant ou non une butée. Cette évolution s'est déroulée, comme on l'a vu, dans et au moyen de l'interaction. En suivant les traces physiques et langagières qui jalonnent la formation des plots, nous nous sommes appliqué à décrire l'espace intersubjectif qui a étayé cette évolution. Nous avons suivi un processus qui nous a mené de l'in-formation des plots à leur en-formation.

3. DISCUSSION

Nous venons de voir comment un objet, en conception distribuée, est passé du stade de l'in-formation à celui de l'en-formation. C'est-à-dire comment cet objet est passé d'un stade où il n'a ni forme, ni fonction, ni structure à un stade où il devient technologiquement défini, cette mise en forme ayant été produite par l'action conjointe de différents acteurs.

Blanco *et al.* étudient cette dynamique de changement en observant la production d'objets intermédiaires. Nous le faisons quant à nous à travers la production langagière des interactants. Ainsi, nous voyons comment les références se construisent et évoluent. L'une et l'autre façon de faire ne sont pas si éloignée. Il s'agit dans les deux cas de montrer le poids de la situation dans le processus de conception. Et ce processus est alimenté par des mécanismes tels que la prise de décision (voir Grosjean *et al.*, 1996) ou encore la désignation des objets dont il a été question dans notre développement. En montrant l'espace duquel le terme 'plots' a émergé, en soulignant sa proximité sémantique initiale avec les termes 'système' et 'montage', en nous attardant sur le processus qui lui a permis de devenir une entité propre, nous avons voulu mettre en valeur le poids de l'interaction dans l'activité de conception distribuée.

On comprend alors pourquoi il n'est pas nécessaire que les acteurs se comprennent toujours exactement lorsqu'ils discutent entre eux. En remontant le fil des trois premières heures de conception de l'objet mécanique, on s'aperçoit que chaque stade est supporté par le précédent.

Ainsi, en ne cherchant pas à définir scrupuleusement les espaces sémantiques des termes utilisés dès le début de l'interaction, les concepteurs ont-ils pu lancer le processus sans chercher à savoir quelle orientation technologique était sous-jacente aux mots proférés par eux-mêmes et leurs partenaires. En parlant indifféremment de l'objet au singulier ou au pluriel, ils ne se sont pas confrontés trop tôt à la répercussion que cela pouvait avoir en terme de conception ni au choix qu'il aurait alors fallu faire. Ainsi encore, c'est en appui sur cette phase d'indifférenciation sémantique ayant permis aux acteurs de discuter sans arrière-pensées de quelques généralités et de quelques spécificités de l'objet à concevoir que ceux-ci ont pu proposer et faire évoluer des solutions technologiques précises. En faisant comme s'ils se comprenaient, excepté bien sûr lorsqu'ils expriment un malentendu et en discutent, c'est-à-dire en faisant comme s'ils partageaient la même représentation de objet, les acteurs ont donc fait l'économie de métadiscours qui auraient pu les emmener loin du résultat escompté et nuire au processus de conception plutôt que lui bénéficier.

Toutefois, nous n'affirmons pas que ce comportement des concepteurs est intentionnel. Il n'a pas été décidé de ne pas préciser les termes utilisés non plus qu'il n'a été choisi d'évoluer d'abord dans une période d'indifférenciation sémantique afin de préserver le potentiel de créativité du groupe. Simplement, on observe que cette façon de faire s'est révélée efficace, ce qui est rendu par l'évolution subreptice du référent 'plots'. Sans que cela soit exprimé explicitement, l'usage du terme 'plots' n'a soudainement plus désigné le système mécanique à concevoir mais des éléments techniques de ce système.

CONCLUSION

Faire de la conception distribuée, c'est faire collaborer dans un même espace des individus aux compétences et aux intérêts divers. Or qui dit compétences et intérêts divers dit représentations diverses tant il est vrai que s'il est couramment admis que nos représentations guident nos actions sur le monde, il n'en est pas moins vrai que nos actions sur le monde guident elles-mêmes nos représentations. On conviendra alors qu'il serait illusoire d'espérer que des ingénieurs de spécialités différentes puissent faire strictement coïncider leur façon de penser et de représenter, fût-ce à propos d'un objet mécanique.

Cependant, ces individus, en tant qu'êtres humains, ont la capacité à faire se superposer des champs sémantiques différents, ce qui fait que le paradigme de conception distribuée se révèle efficace. Notre analyse du processus d'intercompréhension qui sous-tend cette activité de conception nous permet en effet de décrire la genèse et l'évolution d'un objet particulier : des plots. Nous avons vu d'une part comment les concepteurs sont d'abord restés dans un espace de significations très vague pour le restreindre petit à petit et d'autre part comment cet objet, tout en conservant la dénomination de plots, s'est transformé de façon irréversible. La forme des plots telle qu'elle apparaît à la fin du processus que nous avons analysé est donc le résultat d'une sorte de sédimentation des différentes phases qu'ils ont traversé, qu'on les appelle objets intermédiaires ou objets de discours apparaissant dans les mondes créés par les négociations successives. Au moyen d'une analyse pragmatique empruntant à l'ethnométhodologie, c'est-à-dire plaçant l'observateur au cœur du dispositif qu'il observe et non à l'extérieur, nous espérons avoir montré comment s'est construite l'identité d'un objet particulier, les 'plots', objet dont la représentation fut d'abord fluctuante même si les fonctions qui le définissent sont, elles, en quelque sorte invariantes, et en quoi ce travail de co-référenciation dépend de l'interaction langagière qui s'est instaurée entre les acteurs de la conception.

BIBLIOGRAPHIE

- BLANCO E., GARRO O., BRISSAUD D., JEANTET A. (1996). Intermediary objects in the context of distributed design. *Actes IEEE CESA '96*, Villeneuve d'Ascq, 9-11 juillet 1996.
- BRASSAC Ch. et STEWART J. (1996). Le sens dans les processus interlocutoires, un observé ou un co-construit ? *Cinquièmes Journées de Rochebrune "Du social au collectif"*, 29 janvier-3 février 1996.

- CHARPENTIER P., GARRO O., TONIOLO A.-M. et BRISSAUD D. (1995). Vers l'émergence d'un produit en conception distribuée. *Actes du Quatrième Colloque sur la Conception Mécanique Intégrée*, La Plagne, France, 3/5 avril 1995, pp. 111-118.
- CONEIN B. et JACOPIN E. (1994). Action située et cognition, le savoir en place. *Sociologie du travail*, XXXVI 4/94, 475-500.
- CICOUREL A. (1994). La connaissance distribuée dans le diagnostic médical. *Sociologie du travail*, XXXVI 4/94, 427-449.
- GARRO O. (1996). Conception distribuée. De l'industrie... à l'industrie, in J.P. Müller et J. Quinqueton (eds.) *IA distribuée et systèmes multi-agents (JFIADSMA'96)*. Paris : Hermès, pp. 7-20.
- GOFFMAN E. (1973). *La mise en scène de la vie quotidienne*. Paris : Minuit.
- GREGORI N. (1996). Cognitiones distribuées en activité de groupe, la distribution au cœur de l'inter-compréhension. *Actes du deuxième Colloque Jeunes Chercheurs en Sciences Cognitives*. Presqu'île de Giens, 5-7 juin 1996, pp. 97-108.
- GROSJEAN S., GREGORI N. et BRASSAC Ch. (1996). La conception distribuée d'un système de production, une coopération en temps réel. *Actes de la Conférence internationale sur l'Apprentissage Personne Système (CAPS'96)*. Caen, 8-9 juillet 1996, pp. 32-43.
- LATOUR B. (1989). *La science en action*. Paris : La découverte.
- LIVET P. (1994). *La communauté virtuelle. Action et communication*. Combas : L'éclat.
- ROULET E. (éds.) (1985). *L'articulation du discours en français contemporain*. Berne : P. Lang.
- SEARLE J. et VANDERVEKEN D. (1985). *Foundations of illocutionary logic*. Cambridge : CUP.
- SUCHMAN L. (1987). *Plans and situated actions. The problem of Human/Machine communication*. Cambridge : CUP.
- TROGNON A. et BRASSAC Ch. (1992). L'enchaînement conversationnel. *Cahiers de Linguistique Française*, 13, 76-107.
- TROGNON A. et GHIGLIONE R. (1993). *Où va la pragmatique ? De la pragmatique à la psychologie sociale*. Grenoble : PUG.
- TROGNON A. et LARRUE J. (1994). *La pragmatique du discours politique*. Paris : A. Colin.
- VANDERVEKEN D. (1988). *Les actes de discours*. Liège : Mardaga.
- VINCK D. (1995). *Sociologie des sciences*. Paris : Armand Colin.

L'identité en biologie

John Stewart

PHITECO, TSH, Université de Technologie de Compiègne

Intuitivement, les organismes vivants - des animaux, des plantes ou même des microbes - sont le prototype même des entités ayant une identité. Cependant, à l'heure actuelle, il n'existe aucune théorie scientifique pleinement constituée qui parvient à traduire cette intuition en termes précis et explicites. Le but de ce texte est de fournir quelques repères pour identifier la nature des difficultés, dans l'espoir de contribuer à les surmonter.

La première difficulté provient du fait que la biologie orthodoxe - la synthèse néo-darwinienne dont l'aboutissement contemporain est la génétique moléculaire - ne fournit aucune définition de ce qu'est un organisme vivant en tant que tel. Cette biologie possède certes un objet scientifique central, théoriquement et empiriquement bien fondé ; mais cet objet n'est pas "la vie", c'est "le gène". Il est vrai qu'une séquence de nucléotides dans l'ADN - par exemple, celle qui définit le "complexe majeur d'histocompatibilité" des vertébrés - constitue une "empreinte génétique" spécifique à chaque individu. Mais les séquences d'ADN, y compris le CMH, persistent bien après la mort de l'individu - c'est ainsi qu'on a pu les identifier chez des pharaons mommifiés, et même chez des mammoths préservés dans les glaces polaires. Par conséquent, je considère que la biologie moléculaire ne permet pas de cerner ce qui constitue l'identité d'un organisme *vivant*.

Le deuxième repère est la définition de la vie comme autopoïèse, due à Maturana et Varela. Je cite (Varela 1989) :

"Un système autopoïétique est organisé comme un réseau de processus de production de composants qui (a) régénèrent continuellement par leurs transformations et leurs interactions le réseau qui les a produits, et qui (b) constituent le système en tant qu'unité concrète dans l'espace où il existe, en spécifiant le domaine topologique où il se réalise comme réseau.

Il s'ensuit qu'une machine autopoïétique engendre et spécifie continuellement sa propre organisation. Elle accomplit ce processus incessant de remplacement de ses composants, parce qu'elle est continuellement soumise à des perturbations externes, et constamment forcée de compenser ces perturbations. Ainsi, une machine autopoïétique est un système ... à relations stables dont l'invariant fondamental est sa propre organisation (le réseau de relations qui la définit)."

Il y a beaucoup ici pour nous intéresser - en particulier, concernant le thème de ces Journées, une définition de l'identité qui renvoie à une *invariance* fondamentale en termes d'organisation. Il convient d'introduire également ici la notion de "clôture opérationnelle" (Varela 1989) :

"Nous dirons d'un système autonome qu'il est opérationnellement clos si son organisation est caractérisée par des processus :

- a) dépendant récursivement les uns des autres pour la génération et la réalisation des processus eux-mêmes, et
- b) constituant le système comme une unité reconnaissable dans l'espace (le domaine) où les processus existent."

Voilà qui semble être clair. Où est le problème ?

La difficulté provient du fait qu'il n'existe pas des modalités d'interprétation suffisamment bien établies au sein d'une communauté scientifique pour qu'il existe des réponses consensuelles sur l'application de cette définition à des cas concrets. Par exemple, quelles sont les entités autopoïétiques dans la liste suivante (en vrac) : une étoile, un atome, Gaia, une rivière, une colonie de fourmis, une université, la nation française, le système immunitaire..... ? Les avis sont divergents.

La difficulté provient en grande partie de la notion essentielle mais difficile de "clôture opérationnelle". En mathématique ensembliste, la notion de "clôture opérationnelle" est définie pour des sous-ensembles ; si l'opération n'est pas définie dans le sur-ensemble, la clôture est tautologique. Ceci illustre déjà les difficultés de la notion en biologie, et suggère que le formalisme de la mathématique ensembliste pourrait ne pas être adéquate pour exprimer un modèle du vivant.

Par ailleurs, il ne faut en aucun cas confondre la "clôture opérationnelle" varélienne avec une "clôture" matérielle ou thermodynamique. Ce point mérite explicitation. On sait que des organismes vivants - exemples prototypiques des systèmes autopoïétiques - sont nécessairement des structures dissipatives, thermodynamiquement *ouvertes* sur un flux d'énergie. En effet, la physique classique nous enseigne que dans un système thermodynamiquement clos, l'entropie ne peut qu'augmenter ; autrement dit, le système ne peut que tendre vers un équilibre stable où plus rien ne se passe. Or, même si nous ne savons pas très bien ce que sont les organismes vivants, nous savons au moins qu'ils ne sont pas des systèmes stables où rien ne se passe. La conclusion - qu'ils doivent être des systèmes thermodynamiquement ouverts - s'ensuit tout simplement. Il n'est peut-être pas inutile de noter que cette conclusion, qualitative et d'une certaine manière négative, correspond à une auto-limitation de la portée de la notion d'entropie. En effet, l'entropie est une quantité qui n'augmente de façon monotone que sur dans des systèmes clos ; dès qu'un système est ouvert, l'entropie n'est plus une fonction Lyapunov et perd donc l'essentiel de sa pertinence scientifique (Rosen 1991). Autrement dit, l'entropie est

une notion rigoureusement inutilisable pour conceptualiser les organismes vivants en tant que tels. Cette précision est importante car on a beaucoup glosé sur l'équivalence formelle entre "néguentropie" (l'opposé de l'entropie, notamment repris par la thermodynamique statistique) et la définition shannonienne d' "information" : la formule $H = -\sum f_i \cdot \log(1/f_i)$ est en effet rigoureusement analogue dans les deux cas. Ainsi on aboutit à des formulations du type "les organismes vivants se nourrissent d'information". Je considère que ce genre d'expression est un non-sens total. "L'information" n'est en aucun cas une mesure de l'organisation du vivant, laquelle n'est d'ailleurs pas quantifiable. Ceci est cohérent avec la notion de "clôture opérationnelle", qui signifie que les interactions des organismes vivants avec leur milieu constitue des *perturbations* - dont les significations sont déterminées par l'organisation propre de l'organisme - et non pas une source extérieure des formes dont l'organisme aurait besoin pour "s'in-former". (Stewart 1993).

Les malentendus engendrés par la confusion entre "clôture opérationnelle" et "clôture" thermodynamique, matérielle (ou autre, solipsisme compris) sont tels que l'on peut être tenté de ne plus employer le terme de "clôture". Mais ce serait une fausse solution de facilité, qui équivaldrait à jeter le bébé avec l'eau du bain. En effet, c'est bien la clôture (opérationnelle, certes, et non thermodynamique) qui permet de distinguer ce qui appartient à un système autopoïétique et ce qui n'y appartient pas ; autrement dit de définir ce qui constitue l'identité d'un système autopoïétique. Mais alors, qu'est-ce que cela veut dire exactement une clôture "*opérationnelle*" ? Il faut être lucide et honnête : si l'on a pas un critère permettant de distinguer une "opération" (comprise dans une circularité identifiée) et une "perturbation" (interaction avec ce qui sera donc "extérieur"), le discours n'aura pas d'autre cohérence que celle de la tautologie.

Mon troisième - et dernier - repère sera la définition proposée par Rosen (1991) : "un système matériel est un organisme si, et seulement si, il est clos sous causalité efficiente". A mon avis cette définition du vivant est profondément convergente avec celle de l'autopoïèse (bien qu'il convient de signaler que les protagonistes principaux ne se reconnaissent pas mutuellement). Nous retrouvons la même intuition qu'il faut définir le vivant en termes d'une invariance *organisationnelle*, et la même idée fondamentale d'une *clôture* constitutive. Si nous possédions les moyens d'une interprétation consensuelle de cette circularité, nous tiendrions une voie intéressante pour rendre pleinement scientifique une notion biologique de "l'identité". Malheureusement, dans des cas empiriques, nous ne sommes pas plus capables de parvenir à un consensus sur la "clôture sous causalité efficiente" que sur la "clôture opérationnelle".

Pour ne pas en rester là, je propose d'alimenter la recherche d'une solution (que je crois possible) à ces difficultés par une présentation plus détaillée de la théorie de Rosen telle que j'ai pu la comprendre moi-même. Rosen emploie le formalisme suivant : il dénote un *composant* élémentaire d'un système par la formule

$$f : (A \rightarrow B)$$

Mathématiquement, "f" est l'application de l'ensemble "A" vers l'ensemble "B", à ceci près que Rosen emploie la "Théorie de Catégories", branche controversée des mathématiques qui permet à des entités d'être *à la fois* des applications et des arguments. Selon l'interprétation générique de ce formalisme proposé par Rosen, "B" est l'entité produite par le composant en question, "f" est sa "cause efficiente" et "A" est sa "cause matérielle".

Toute la physique classique peut alors être caractérisée, génériquement, comme la science des "systèmes dynamiques déterminés par leur état" (ou "SDDE"). Dans ce cas, "A" est l'état du système à temps "t", "B" est l'état du système à temps "t + dt", et "f" est un endomorphisme qui correspond à la loi dynamique qui régit (en droit) toute l'évolution temporelle du système. Dans la physique des SDDE, il n'y a pas (et ne saurait y avoir) de réponse à la question "pourquoi f?" - la loi de la gravitation universelle, ou les lois de Maxwell, sont simplement des "lois de la nature" un point c'est tout. Mais en biologie - et c'est ici que Rosen établit une différence fondamentale entre biologie et physique - les organismes sont caractérisés par leur "clôture sous causalité efficiente", donc il doit y avoir une réponse à la question "pourquoi f?" Quelle sera alors la forme minimale - et générique - des organismes vivants ainsi définis ?

Admettons que la "cause matérielle" de f soit B (comme le dit Rosen, si on cherche une configuration minimale, avec une concentration maximale des implications causales, "pourquoi pas" ?) Il nous faut alors une cause efficiente pour la transformation (B → f) ; appelons-la "phi" :

$$\phi : (B \rightarrow f)$$

Il nous faut alors une cause matérielle pour phi ; en économisant au maximum, comme précédemment, on peut supposer que f fera l'affaire (mais noter que f doit alors être à la fois une application et, maintenant, un ensemble). Mais il semble que nous soyons pris dans une fuite en avant sans fin, car il nous faut maintenant une cause efficiente pour effectuer la transformation (f → phi) ; appelons-la "beta" :

$$\beta : (f \rightarrow \phi)$$

C'est à ce moment, sous la menace d'une régression à l'infini, que Rosen a trouvé une solution (et par construction, elle est bien la plus simple possible) : rien n'interdit de supposer que beta et B soient identiques! Evidemment, il faut admettre que ce composant clé soit à la fois un ensemble (B) et une application (beta), mais nous avons

déjà employé cette ressource de la Théorie des Catégories en ce qui concerne f . Si nous écrivons :

$Beta = (beta = B)$,

nous avons donc :

$Beta : (f \rightarrow \phi)$

Le théorème fondamental de Rosen peut donc s'énoncer comme suit : le modèle relationnel minimal qui possède la propriété de clôture sous causalité efficiente comporte trois composants : f est la cause efficiente de $Beta$, qui est la cause efficiente de ϕ , qui est la cause efficiente de f Rosen propose d'appeler "biologie" la science de tous les modèles ayant la propriété d'une clôture sous causalité efficiente ; pour marquer la différence avec la biologie néo-darwinienne, j'appellerai cette nouvelle science par le terme "biologie relationnelle".

Il faut cependant souligner que le théorème de Rosen marque non pas un point d'aboutissement, mais un point de départ. Ce qui manque, pour le moment, ce sont les procédures de mesure et d'interprétation qui permettrait de mettre ce genre de modèle à l'épreuve des observations empiriques. Il faut en effet dire quelques mots sur la nature *sémantique* des trois composants canoniques d'un modèle relationnel exhibant une clôture sous causalité efficiente. Que peuvent représenter les entités f , ϕ et $Beta$?

La première chose à noter est qu'en raison de la causalité circulaire, qui caractérise *par construction* le système en question, la structure des implications causales de ce modèle est à la fois *élargie* et *intensifiée* au-delà de ce qui est pensable dans la physique des SDDE. En particulier, il est maintenant possible de donner un contenu scientifique aux notions de cause finale et de fonction. Si nous posons par exemple la question "pourquoi $Beta$?" il est possible, comme pour les SDDE, de donner des réponses :

- parce que f (cause efficiente) ; et
- parce que A (cause matérielle).

Mais en plus, à la différence de la physique des SDDE, $Beta$ possède également des "raisons d'être" qui provient de ses *conséquences*. En effet, on peut aussi répondre à la question "pourquoi $Beta$?" en disant :

- parce qu'il est la cause (matérielle) de f (et sans f , $Beta$ lui-même n'existerait pas) ; et
- parce qu'il est la cause (efficiente) de ϕ (et sans ϕ , f et donc $Beta$ lui-même n'existeraient pas).

Autrement dit, les composants d'un modèle relationnel possèdent des *fonctions*, non pas au sens mathématique (des applications), mais au sens biologique du terme. Par exemple, la *fonction* du cœur est de pomper du sang, et ceci est nécessaire pour que le cœur lui-même soit alimenté en glucose et en oxygène (et plus généralement, pour que l'organisme dont il est un composant maintienne son autopoïèse). L'identification

sémantique des composants d'un modèle relationnel est difficile tout d'abord car jusqu'ici les biologistes ont intériorisé l'interdit d'employer des concepts de cause finale et de fonction. Par conséquent, ils emploient ces concepts subréptiquement mais sans imaginer qu'ils puissent avoir à le faire avec rigueur et précision. Par ailleurs, la quasi-totalité des observations en biologie jusqu'ici sont exprimés en termes de *structures matérielles*, susceptibles de servir comme bases empiriques d'un modèle du type SDDE. Or, pour des raisons fondamentales que nous allons examiner, un modèle relationnel et un modèle SDDE sont *nécessairement incommensurables*. Autrement dit, on sait d'avance que dans le cas des organismes vivants, il sera impossible d'observer empiriquement les composants fonctionnels d'un modèle relationnel directement en termes de la grande masse d'observations accumulées jusqu'ici par la biologie. Regardons de plus près pourquoi il en est ainsi en revenant à la question de la relation entre cette nouvelle biologie et la physique des SDDE.

La physique des SDDE tend à se considérer comme étant en droit "la science de tout ce qui existe"; et cela se comprend, car dans ses propres termes il donne lieu à un déterminisme total, comme l'avait si bien vu Laplace. Il s'ensuit comme corollaire que la biologie ne serait qu'un cas particulier de la physique. Rosen renverse radicalement cette perspective. La structure des implications causales qui définit la classe de modèles que j'ai appelé "biologie relationnelle" est beaucoup plus riche - et donc beaucoup plus *générale* - que celle de la physique classique. Par conséquent, c'est la physique qui est un cas particulier de la biologie, en non l'inverse. Certes, en raison des contraintes supplémentaires et la restriction conduisant à une extrême économie des implications causales, la physique classique possède des propriétés particulières qui n'existent pas dans la "biologie" en général. Tout comme la restriction aux nombres entiers constitue une arithmétique ayant des théorèmes sans contrepartie dans les mathématiques plus générales du réel continu, la physique possède une théorie - celle des SDDE - qui n'existent pas dans la science des modèles relationnels en général.

Quelle est donc la relation entre ces deux classes de modèles, celle (particulière) des SDDE, et celle (plus générale) d'une biologie relationnelle ? Il est clair, je pense, qu'ils sont *incommensurables*. La physique des SDDE *ne peut* penser la clôture sous causalité efficiente. Mais réciproquement, une biologie relationnelle ayant comme objet central les systèmes clos sous causalité efficiente *ne peut* avoir recours, pour penser les organismes vivants en tant que tels, à la théorie des SDDE. Evidemment, face à un système naturel que l'observateur scientifique considère comme "vivant", soit sur la base du sens commun, soit parce qu'il existe un modèle relationnel, rien ne peut empêcher un "bio-physicien" de construire un modèle SDDE. Au contraire, il s'agit de la démarche habituelle de modélisation "biologique", qui nous renseigne de façon irremplaceable sur le fonctionnement dynamique du système. Par conséquent, une biologie relationnelle

comportera deux types de modèle - SDDE et relationnellement clos sous causalité efficiente - qui sont incommensurables entre eux. On peut distinguer trois catégories de "complexité" : (i) une complexité "objective", caractérisée par la nature du modèle d'un système naturel ; (ii) une complexité "épistémique" caractérisée par le fait qu'il existe une pluralité de modèles mutuellement incompatibles ; et (iii) complexité "herméneutique", où le sujet de la connaissance s'inclut lui-même dans un schéma d'interprétation circulaire (Stewart 1996). Dans cette terminologie, une biologie relationnelle (si on parvient à la constituer) possédera nécessairement une complexité "épistémique" - et, si elle réussit à prendre en compte le fait que les biologistes sont eux-mêmes des organismes vivants dont le répertoire cognitif est constitué par une clôture, une complexité "herméneutique".

En raison de ces complexités, on peut donc comprendre que l'identification sémantique des trois composants f , ϕ et β , essentiels pour un organisme vivant minimal (et donc pour tout organisme vivant) sera une tâche difficile. Néanmoins, on ne peut totalement s'y dérober, sous peine de frapper toute l'entreprise d'une biologie relationnelle de nullité. Voici donc, sous toute réserve, ce qu'il me semble possible d'en dire dans l'état actuel de la question.

Nous avons déjà vu qu'une caractéristique essentielle des organismes vivants est d'être branchés sur un flux d'énergie (et de matière), et d'organiser spontanément (typiquement par des formes circulaires, comme les cellules de Bénard¹) la dissipation de ce flux. Dans le cas des organismes vivants sur la planète Terre, la source de ce flux est (directement ou indirectement) l'énergie solaire : le "flux" est donc photonic, transformé par photosynthèse en un flux électronique, et organisé moyennant des transformations (bio)chimiques. Mae Wan Ho (1993) a décrit avec beaucoup de poésie comment les organismes vivants "captent des électrons en chute libre", et par la suite assurent leur circulation à travers la cellule où ils assurent l'animation des processus vitaux avec des pertes d'énergie (production de chaleur) remarquablement minimisées. Ce n'est sûrement pas un accident si encore aujourd'hui le "centre" du métabolisme de tous les organismes vivants est le "cycle de Krebs", qui possède la forme circulaire d'une structure dissipative biochimique.

Il est donc assez convaincant de suivre Rosen (1991) et d'identifier le premier composant, " f ", avec la fonction du *métabolisme*, ou encore avec celle de la *physiologie*. Plus précisément, " f " est la *cause efficiente* des flux métaboliques. En pensant aux

¹ Si une liquide contenue entre deux plaques de verre horizontales est chauffée doucement par une source placée en-dessous des plaques, la chaleur est d'abord dissipée par diffusion. Cependant, si le taux de chauffage est augmenté, au-dessus d'un certain seuil la chaleur est dissipée par des nombreux petits courants de convection qui "s'auto-organisent" dans une configuration hexagonale (vue par-dessus).

organismes vivants actuels, il est tentant d'identifier "f" avec des enzymes, qui contribuent très certainement à l'organisation et à l'effectuation du métabolisme. Mais on court ici le risque de nous égarer, en faisant une identification réifiante et faussement simpliste entre un composant d'un modèle relationnel et des structures matérielles - alors que nous savons d'avance que cela n'est pas possible. On peut s'apercevoir de l'erreur si nous revenons un instant aux cellules de Bénard. Ici, "f" est le dispositif - les plaques de verre, mais aussi le fait qu'elles soient horizontales, qu'elles referment le liquide, et que l'ensemble se trouve juste au-dessus d'une source de chaleur suffisamment forte pour déclencher les courants de convection (mais pas trop forte, ce qui ferait bouillir l'eau...!) De plus, en raison de leur (auto)-organisation circulaire, les courants sont en plus en partie de leur propre cause efficiente - et donc participent eux-mêmes à "f". Autrement dit, l'interprétation de "f" fait intervenir tout un ensemble de structures *et leur organisation, notamment spatiale*. Revenant aux organismes vivants à la lumière de cet avertissement, il apparaît qu'un élément capital de "f" est la membrane cellulaire, qui évite la perte (par diffusion) des éléments du cycle métabolique, et peut-être aussi l'organisation du cytoplasme qui rend efficace le transfert de l'énergie le long de la chaîne électronique. Les enzymes y participent chez les organismes contemporains, mais "f" ne peut s'y réduire.

Qu'en est-il donc du deuxième composant, "phi" ? Sa fonction est de produire le composant "f". En effet, si "f" - l'ensemble du dispositif qui donne lieu au flux métabolique - n'est pas entretenu, il finira par tomber en panne et le processus s'arrêtera. Chez les machines non-vivantes - comme une voiture, ou encore comme les cellules de Bénard entre les plaques de verre - l'entretien est une opération distincte de la première fabrication. C'est sans doute pour cette raison que Rosen (1972, 1991) donne au composant "phi" le nom de "réparation". Cependant, chez les organismes vivants (qui nous intéressent en premier chef), le cas est un peu différent. Ici, il se trouve que la "réparation" des composants est un processus permanent qui, en fin de compte, n'est pas séparable de leur fabrication. Pour prendre la métaphore d'une vieille voiture (version moderne du bateau de la mythologie grec), il faut imaginer qu'avec le temps on en vient à changer toutes les pièces - les pneus, les roues, la carrosserie, le moteur - de sorte qu'à la fin il n'y a aucun élément matériel en commun avec la voiture originale. Pourtant, l'intuition nous dit qu'il s'agit toujours de la "même" voiture : en tout cas, il n'y a eu aucune solution de continuité, ni dans son organisation ni dans sa fonctionnalité. On sait que les organismes vivants sont comme cette vieille voiture : les atomes et les molécules matérielles qui les composent se renouvellent en permanence, alors que leur organisation essentielle reste invariante. Autrement dit, chez les organismes vivants "réparation" et "fabrication" coïncident. Ainsi, se restreindre au terme "réparation" pour interpréter phi est discutable. En effet, si la fonction "phi" peut parfois prendre les allures d'une "réparation", il s'agit peut-être plutôt d'une *conséquence secondaire* de l'organisation

autopoïétique dans son ensemble. Si les perturbations provenant des interactions entre l'organisme et son milieu n'infligent pas des "dommages", le composant "phi" qui régénère en permanence les causes efficientes du métabolisme n'apparaîtra pas comme une "réparation". Pour ces raisons, je propose de donner au composant "phi" le nom plus général de "production" ou (en anticipant sur la suite) "poïèse".

Rappelons - pour mieux l'écarter - la tentation à identifier "f" purement et simplement avec des enzymes. Dans cette optique réifiante, "phi" chez les organismes vivant actuellement sur la planète Terre pourrait être identifié avec les acides nucléiques - l'ADN et les ARNt et ARNm - qui sont, en effet, nécessaires pour la synthèse des protéines que sont les enzymes. Cette interprétation n'est peut-être même pas totalement fausse ; mais elle est peu rigoureuse, car dans le modèle de Rosen "phi" correspond à une *fonction*, l'ensemble du dispositif qui est nécessaire à la production permanente de l'ensemble de "f". Reprenant l'exemple des cellules de Bénard, "phi" est le processus de production de "f", c'est-à-dire la fabrication du dispositif comprenant les plaques de verre horizontales enfermant le liquide, l'ensemble étant placé au-dessus d'une source de chaleur. Dans ce cas - comme chez toutes les machines - le processus "phi" (qu'il s'agit de la réparation ou de la fabrication initiale) est réalisé par des êtres humains. Autrement dit, "phi" est un processus *hétéronome* tout à fait distinct du *fonctionnement* de la machine elle-même. L'usine où une voiture est fabriquée, et/ou l'atelier où elle est entretenue et réparée, n'ont rien à voir avec les routes sur lesquelles elle roule. Comme nous allons voir, il n'en est pas de même pour les organismes vivants.

Qu'en est-il, enfin, du troisième composant dans le modèle de Rosen ? Il s'agit tout d'abord du composant "beta" qui (par définition) est l'ensemble du dispositif qui assure la production de "phi" (qui produit "f" qui produit les flux métaboliques....). On se croit parti dans une régression sans fin : on prévoit déjà un "gamma" pour produire "beta", un "delta" pour produire "gamma", etc. et en effet, dans le cas des machines, la mise en place de l'usine (avec des ouvriers, un approvisionnement en matières premières etc) est distinct du fonctionnement de l'usine pour fabriquer la machine, laquelle fabrication est distinct du fonctionnement de la machine pour Mais dans le cas des organismes vivants, cette régression indéfinie n'a pas lieu, car la série *boucle sur elle-même* : le composant "beta" *est identique à* "B", le produit des processus métaboliques. Cette identité remarquable, qui est au cœur même de l'organisation du vivant en tant que tel, peut être désigné par la formule "beta = B". Notationnellement, on peut écrire "Beta" pour désigner le cas où, précisément, beta et B se recouvrent : "Beta = (beta = B)". Ainsi, "f" produit "Beta" qui produit "phi" qui produit "f"..... Autrement dit, chez les organismes vivants, les processus de réparation/fabrication/production, désignés par "phi", sont tirés à l'intérieur même du système où les processus métaboliques "f" sont à

l'oeuvre. Le composant "Beta" est donc ce qui permet la clôture sous causalité efficiente, et confère à l'organisme dans son ensemble son caractère *auto-poïétique*. Cependant, son interprétation en termes de structures matérielles est extrêmement délicat.

De même qu'il était tentant d'identifier "f" avec des enzymes, et "phi" avec les acides nucléiques, on peut songer à interpréter "beta = B" par le fait que *dans le contexte métabolique fourni par la cellule*, les acides nucléiques peuvent se répliquer. C'est peut-être pour cette raison que Rosen (1972, 1991) propose le terme "réplication" pour désigner ce troisième composant. Mais, comme pour le terme "réparation", cette désignation est discutable. Si la fonction "Beta" prend parfois - souvent, même - les allures d'une "réplication", il s'agit peut-être plutôt d'une autre *conséquence secondaire* de l'organisation autopoïétique dans son ensemble. D'une part, à *condition* qu'un organisme soit autopoïétique, il lui suffira de croître et puis de casser convenablement en deux parties pour se reproduire. D'autre part, si l'organisme est mortelle mais de fait ne meurt pas (comme "Gaia", la planète "vivante" selon Lovelock (ref)), "Beta" (généré par le premier composant "f" et qui re-génère en permanence le deuxième composant "phi") sera mal caractérisé par le terme "réplication" qui suggère une malencontreuse identification avec la reproduction des organismes. Cette objection rejoint celui de Maturana à l'encontre de la "reproduction" comme caractéristique essentielle des organismes vivants.

De toute façon, le terme "réplication" possède l'inconvénient de connoter les acides nucléiques. Encore une fois, il faut bien se garder de l'erreur réifiante qui confond un composant d'un modèle relationnel (défini en principe par sa fonction, ce dont il est la cause efficiente) avec des structures matérielles particulières qui sont, au mieux, une instantiation contingente. Même si les acides nucléiques sont des éléments nécessaires à tous les organismes vivants actuellement connus sur la Terre, l'équation "ADN = secret de la vie" est trompeuse. Le composant "beta = B" est avant tout un concept abstrait, signifiant que les produits du métabolisme (B) concourent en permanence à assurer le maintien dynamique des processus poïétiques "phi".

Ainsi, il n'est pas aisé de trouver un terme adéquat pour éviter les réifications et interpréter convenablement ce composant "Beta". Je propose, provisoirement, de le nommer par la préfix "auto-" ; cela a au moins le mérite de souligner son caractère profondément relationnel. En effet, "Beta" ne saurait exister en tant que tel tout seul, car sa définition même réside dans le *recouvrement* entre "B" et "beta".

En conclusion, je peux résumer ces tentatives interprétatives par la thèse suivante. L'identité d'un organisme vivant provient d'un invariant de son organisation, à savoir la circularité entre trois composants fonctionnels essentiels :

i) Son *métabolisme* ou *physiologie*, le fait que l'organisme est le siège d'un ensemble de processus dynamiques ;

ii) Sa *poièse*, le fait qu'il comporte un dispositif qui *fabrique* les flux métaboliques ;

iii) Sa caractéristique *auto-* poiétique, le fait que ce soient ses propres processus métaboliques qui fabriquent son dispositif poiétique (qui fabrique le métabolisme qui fabrique.... etc, recursivement).

Il reste à voir - pratiquement, ensemble - si cette élaboration contribue à rendre possible une discussion *empirique* sur le statut autopoïétique (ou non) des entités divers, et par conséquent sur leur identité.

Références.

Ho M.W. (1993). *The Rainbow and the Worm*. World Scientific, Singapore.

Lovelock J. (1988). *The Ages of Gaia*. Norton, New York.

Rosen R. (1972). Some relational cell models: the metabolism-repair systems. In: Rosen R. Ed., *Foundations of Mathematical Biology*, Academic Press, N.Y., pp 217-253.

Rosen R. (1991). *Life Itself*. Columbia University Press, Columbia.

Stewart J. (1993). Au delà de l'inné et de l'acquis. *Intellectica* 16, 151-174.

Stewart J. (1996). Towards an epistemology of complexity. *Actes du Symposium Echo*, Amiens.

Varela F.J. (1989). *Autonomie et Connaissance*. Editions du Seuil, Paris.

L'IDENTITÉ DES SCIENCES COGNITIVES : DE L'ÉPISTÉMOLOGIE À LA SÉMANTIQUE (ET RETOUR...)

Pierre-Yves Raccah
CNRS-idl, Paris

Position de la question

Les sciences cognitives, en raison de leur caractère interdisciplinaire, rassemblent des chercheurs issus de nombreuses disciplines (de la philosophie à la neurobiologie, et même jusqu'à la physique atomique) pour étudier, selon des éclairages divers, des phénomènes considérés –et c'est là ce qui caractérise les sciences cognitives– comme formant une certaine unité. Ce rapprochement de chercheurs de différents horizons provoque, dans chaque discipline, des situations de choix: se rapprocher de l'étude de la cognition vs refuser de s'en rapprocher (ce qui revient, de fait, à s'en éloigner). Je développe, au premier paragraphe, une manière de reconstruire, d'un point de vue théorique, l'unité phénoménologique à la base de la problématique cognitive. Au deuxième paragraphe, j'analyse, de manière abstraite, les choix qui me semblent marquer, dans les différentes disciplines, l'identité disciplinaire dessinée par la problématique cognitive. Au troisième paragraphe, je propose une analogie entre *chercheur* et *interprète*, analogie qui découle d'une identification entre *sens* et *connaissance*, identification qui s'impose lorsqu'on prend au sérieux l'unité phénoménologique telle qu'elle est proposée au premier paragraphe.

1. L'identité cognitive

Plusieurs disciplines, pour des raisons très diverses, ont suivi un chemin qui les conduit à s'intéresser à la connaissance et à son organisation –ou sa *gestion*– chez l'homme. Ces problématiques complémentaires –plus que communes– constituent le champ d'investigation des sciences de la cognition.

Le rôle de la recherche cognitive, indépendamment de ses motivations plus profondes, relevant de la philosophie de l'esprit, est de mettre en lumière les différents modes de *gestion des connaissances*: modes de représentation, de production, de transmission, d'acquisition de ces connaissances. Les différentes disciplines qui participent à cette nouvelle entreprise (psychologie, neurologie, sciences du langage, logique, informatique, ergonomie, ...) concernent toutes, chacune de son point de vue, certains aspects de la gestion des connaissances.

Elles contribuent à la recherche cognitive en apportant des théories et des modèles grâce auxquels on espère entrevoir, au delà du domaine spécifique pour lesquels ils sont construits, des constantes concernant la gestion des connaissances.

Il ne s'agit pas de transplanter, tels quels, des modèles propres à une discipline, dans une autre discipline. Ainsi, il n'est pas question de reproduire, sous forme d'un programme informatique, les mécanismes neurologiques que l'on aura pu mettre en évidence chez des sujets humains, par exemple, lorsqu'ils font une analogie, mais, si tant est que de tels mécanismes puissent être mis en évidence, de les comparer aux modèles psychologiques proposés, aux modèles linguistiques de l'expression de l'analogie, etc. afin d'en tirer des conséquences sur ce mode de production de connaissances. De même, il n'est pas question de comparer le cerveau à un ordinateur, si sophistiqué fût-il, en tentant de rendre compte de processus neurologique au moyen de modèles informatiques.

La recherche cognitive étudie un domaine plus abstrait que la psychologie, la neurologie et les autres disciplines ayant trait, de près ou de loin, à la gestion des connaissances, mais peut-être plus finalisé: il s'agit d'"inventer" des structures ayant des fonctionnalités précises, compatibles avec (ou mieux, rendant compte de) les processus naturels liés à la gestion des connaissances, et susceptibles d'être implantées, ne serait-ce que pour tester leur efficacité, dans des architectures artificielles. De ce point de vue, qui n'est pas le seul ni le plus accepté, les sciences de la cognition peuvent être considérées comme les bases théoriques de l'intelligence artificielle.

J'ai appelé *hypothèse de l'abstraction cognitive*, cette hypothèse selon laquelle les sciences cognitives s'intéressent à ces concepts abstraits, dont on trouve des exemples dans les différentes disciplines concernées par la gestion des connaissances. Voici une formulation de cette hypothèse.

HYPOTHÈSE DE L'ABSTRACTION COGNITIVE (vue par un linguiste...)

La langue est un produit de l'esprit humain et on peut penser que ses structures sont des manifestations de structures plus abstraites de cet esprit, structures dont d'autres manifestations, distinctes des premières mais néanmoins analogues, se retrouvent dans l'*usage* de la langue et, plus généralement, dans la manière dont l'esprit *gère* les connaissances. Ainsi, les structures de l'expression linguistique (considérée comme un produit, et non comme une production) sont la trace, dans le domaine de la langue, de structures plus abstraites (appelons-les *cognitives*), dont d'autres traces peuvent être trouvées dans d'autres domaines, par exemple, celui de l'esprit ou celui de l'intelligence artificielle.

Cette hypothèse, qui justifie les analogies possibles entre les modèles des différents domaines concernés par la gestion des connaissances, tout en en fixant les limites, crée, de fait, un nouveau corps de concepts qui généralisent ceux des domaines initiaux (comme le concept de *champ* généralise des concepts du magnétisme, de l'électricité et de la gravitation). Et ce sont précisément ces nouveaux concepts qui constituent l'objet des sciences cognitives.

Ce positionnement de la linguistique par rapport aux sciences de la cognition, s'il s'accorde assez bien avec la première des hypothèses que Michael Dummett considère comme caractéristiques du « tournant linguistique » sur lequel est fondée la philosophie analytique, n'implique nullement la seconde :

« Ce qui distingue la philosophie analytique en ses divers aspects d'autres courants philosophiques, c'est en premier lieu la conviction qu'une analyse philosophique du langage peut conduire à une explication philosophique de la pensée, et en second lieu la conviction que c'est la seule façon de parvenir à une explication globale. »¹

La notion d'*abstraction cognitive* exclut, au contraire, que l'on examine les relations entre langue et pensée en termes de détermination de l'une par l'autre. C'est cette « troisième voie » qui, en attribuant de manière parcimonieuse les déterminations causales, constitue l'originalité de la conception et en fait la source de sa richesse. Ainsi, par exemple, la thèse frégréenne de l'*objectivité du sens*, sur laquelle, selon Dummett, s'appuie tout l'édifice de la philosophie analytique n'est pas le seul rempart contre le psychologisme, fonction qui constitue pourtant sa motivation principale.

2. Frontières, identité et intégration

Les questions de fonctions cognitives, telles qu'elles ont été évoquées au paragraphe précédent, ne sont pas les seules abordées par la recherche cognitive; la façon même de concevoir les problèmes est révélatrice de certains choix, de certains courants dans les sciences cognitives. Ainsi, la manière dont j'ai présenté la problématique de la cognition suppose un intérêt pour la *simulation* du raisonnement et, plus généralement, du *comportement global*, approche qui a parfois été qualifiée de *symbolique*. Mais d'autres courants de la recherche cognitive s'intéressent aux étapes les plus élémentaires des activités humaines; elles ne cherchent pas seulement à simuler le comportement global, mais surtout à reproduire les *micro-événements* intervenant dans les processus humains. Cette approche, qui a parfois été appelée *sub-symbolique*, est représentée par certains des différents courants connexionnistes. Elle implique des recherches biologiques, y compris sur des formes de vies dites inférieures, que l'approche symbolique ne justifie pas; et, en amont, on conçoit que des modèles de phénomènes encore plus élémentaires, relevant des sciences de la nature, puissent contribuer à servir ces objectifs. Il ne s'agit pas nécessairement de démarches réductionnistes: l'étude de l'*émergence* d'un phénomène relevant d'un

¹ Dummett (1988).

niveau de complexité donné, à partir d'autres phénomènes, de niveaux de complexité inférieurs, ne vise pas nécessairement à nier la pertinence d'une problématisation à ce premier niveau. Enfin, très liée à l'approche sub-symbolique même si, par certains côtés, ses objectifs sont plus "globalisants" que l'approche *symbolique*, est apparue, depuis peu, une problématisation *sociale* de la gestion des connaissances: il s'agit, là aussi, de comprendre une émergence, mais cette fois, le niveau d'organisation à expliquer est plus élevé que l'individu. Pour la *cognition sociale*, le danger de réductionnisme vient du *symboliste* qui ne reconnaît pas la dimension sociale de la gestion des connaissances, et non pas du biologiste, lequel, qu'il l'admette ou non, étudie des collectivités.

De la façon dont j'ai présenté ces trois grands courants, on comprendra que je les considère comme complémentaires...

Ainsi, dans chacune des disciplines susceptibles de contribuer à la recherche cognitive le choix, pour un chercheur, d'adopter ou non une problématique cognitive me semble être déterminé par la position qu'il adopte vis-à-vis de deux paramètres non-indépendants (mais que j'exposerai séparément, par souci de clarté)²:

- la recherche -ou non- d'un principe intégrateur subsumant différents points de vue sur le même objet, objet qui touche, d'une manière ou d'une autre, à la gestion des connaissances, qu'il s'agisse d'approches empiriques ou d'approches exclusivement formelles.

Ce critère est souvent appelé "interdisciplinarité": l'interdisciplinarité qu'implique la problématique cognitive est donc *intégratrice* (cf. l'abstraction cognitive).

Exemples:

(i) Un linguiste peut chercher à décrire les structures d'une (ou de plusieurs) langue(s) sans s'intéresser (professionnellement) aux rapports entre ces structures et celles de l'esprit des personnes qui parlent cette (ou ces) langue(s). En ce cas, il n'a pas de raison de s'intéresser à la cognition.

(ii) D'une manière analogue, un psychologue peut chercher à décrire des structures ou des processus mentaux sans les concevoir comme intervenant dans un système de gestion de connaissances: toute psychologie n'est pas nécessairement cognitive.

(iii) Un logicien travaillant à des systèmes de logique formelle peut concevoir son travail comme destiné à pouvoir être utilisé pour rendre compte, par

² Je ne prétends pas limiter la caractérisation à ces deux paramètres: ce sont ceux qui m'ont paru les plus saillants, sans qu'ils soient, pour autant, uniques. D'autre part, une réponse positive à l'une de ces deux questions me semble suffisante pour justifier le caractère cognitif de la démarche du chercheur.

exemple, des structures mentales des êtres humains ou des processus inférentiels des êtres animés: en ce cas, il participe de la problématique cognitive.

- la recherche –ou non– d'une sorte de justification de l'émergence d'un niveau d'organisation plus abstrait à partir du niveau établi par la discipline d'origine, ou, à l'inverse, la recherche –ou non– à partir de niveaux plus élémentaires d'une justification du niveau établi par la discipline d'origine (ce paramètre ne concerne que les disciplines empiriques).

Exemples:

Un neurobiologiste qui étudie la physiologie des cellules du cerveau peut avoir comme seul objectif de décrire les processus physiologiques concernant le cerveau: en ce cas, son activité ne concerne pas les sciences de la cognition; mais sa recherche peut viser à rendre compte de l'émergence de structures psychologiques (représentations, inférences, affects, etc.)³: en ce cas, il contribue directement aux sciences cognitives.

De même, un biochimiste qui étudie les interactions moléculaires dans les échanges entre cellules peut, ou non, considérer son travail comme une contribution aux sciences de la cognition, selon qu'il s'intéresse, ou non, à fonder, au niveau de la biochimie, les concepts de la biologie cellulaire, conçue, elle-même, comme fondant des concepts relevant d'un niveau plus abstrait d'organisation, élément d'une chaîne aboutissant à la gestion des connaissances.

Ces deux paramètres, que j'appellerai, plus par convention que par nécessité, *principes d'intégration horizontale* (pour le premier) et *verticale* (pour le deuxième) ne sont pas toujours explicités dans les différentes démarches des acteurs du domaine cognitif. D'autre part, si le premier principe est, en général, facilement reconnaissable par ces acteurs, comme sous-tendant leur démarche, il n'en va pas de même pour le deuxième principe, qui requiert une réflexion épistémologique plus poussée.

Les nouvelles frontières disciplinaires dessinées par la problématique cognitive sont donc de nature épistémologique: elles concernent la conception de l'homme et la conception de la science. Le *cognitologue* considère l'homme comme une unité phénoménologique pouvant être étudiée selon plusieurs points de vue. Chaque point de vue peut relever d'une discipline particulière, dotée d'une méthodologie propre, mais l'étude de l'homme est soumise aux deux principes d'intégration.

³ Insistons sur le fait que "rendre compte de l'émergence" ne signifie pas *réduire*: dans le premier cas, ce qui *émerge* est considéré comme un *fait*, dont on décrirait la construction, tandis que, dans le deuxième cas, ce que l'on *réduit*, est considéré comme une *apparence* (éventuellement trompeuse) n'ayant pas de pertinence théorique.

Pour les chercheurs des disciplines formelles (logique, mathématique, informatique) la démarcation entre *cognitologues* et *non-cognitologues* sépare, d'un côté, ceux qui cherchent à fournir des modèles utilisables par (ou inspirés par) les approches relevant des sciences de l'homme et de la société, et, de l'autre côté, ceux qui ne s'embarrassent pas de telles exigences: on a souvent entendu dire, à juste titre⁴, pour justifier des modèles ou des programmes informatiques peu conforme à l'idée que l'on peut se faire de ce qui se passe chez l'homme, qu'il n'est pas nécessaire que les avions aient des plumes pour qu'ils puissent voler.

Pour les chercheurs des disciplines empiriques (de la physique et la biochimie à la linguistique et la psychologie), la ligne de démarcation passe entre ceux qui tentent de construire des théories disciplinaires dont l'intégration rende compte de l'émergence des structures et des processus de gestion des connaissances, et ceux qui limitent leur horizon meta-théorique en deçà de cette émergence.

A l'intérieur de cette catégorie des disciplines empirique, une autre catégorisation peut présenter un certain intérêt: du point de vue des sciences humaines et des sciences sociales, le camp des *cognitologues* valorise l'intégration verticale "vers le bas" (vers les niveaux d'organisation inférieurs), tandis que, du point de vue des sciences de la nature, le camp des *cognitologues* valorise l'intégration verticale "vers le haut". Ainsi, un chercheur en cognition sociale tentera de rendre compte de l'émergence de phénomènes de gestion sociale des connaissances, à partir de phénomènes, supposés plus élémentaires, concernant les acteurs de l'organisation sociale, tandis qu'un chercheur en neurobiologie cognitive tentera de décrire les phénomènes neurobiologiques de telle manière que cette description rende compte de l'émergence de phénomènes, supposés plus complexes, concernant la gestion des connaissances chez l'individu dans lequel ces phénomènes neurobiologiques siègent.

Enfin, les rapports entre les disciplines formelles et les disciplines empiriques sont marqués par le fait que les premières tendent à utiliser les secondes comme sources de modèles théoriques, tandis que les secondes tendent à utiliser les premières comme terrains d'expérimentation⁵.

3. Intégration et langue

Ainsi, comme nous venons de le voir, les migrations de concepts, d'une discipline à l'autre, sont soumises à de fortes contraintes: le même objet d'expérience ne produit pas le même objet d'étude, selon que l'on s'y intéresse d'un point de vue ou d'un autre; le même objet d'étude (ou *d'observation*) ne constitue pas le même objet scientifique selon qu'on l'aborde dans le cadre d'une discipline ou d'une autre. Nous avons vu les difficultés que ce «phénomène épistémologique» ajoute à la constitution du champ disciplinaire

⁴ Encore que, la question ne devrait pas être réglée trop vite...

⁵ Parmi d'autres.

des sciences cognitives (comme cela est d'ailleurs le cas pour toute problématique interdisciplinaire). Nous nous proposons maintenant de construire et d'étudier une double analogie entre *langue* et *connaissance* qui permettrait de considérer ce « phénomène épistémologique » comme un cas particulier d'un phénomène plus général, dont une autre manifestation se retrouve dans les rapports entre *interprétation*, *sens* et *signification*. Pour justifier cette double analogie, il ne suffira pas, bien sûr, de montrer que les relations entre ces trois derniers concepts sont semblables à celles qui tiennent entre *objet d'expérience*, *objet d'étude*, et *objet scientifique* : il faudra aussi convaincre que cette similitude n'est pas fortuite. C'est ce que j'essayerai de faire en illustrant l'intérêt heuristique de cette double analogie.

Ceux qui s'y sont intéressés se souviendront⁶ de la distinction entre *phrase*, *énoncé-type*, et *énoncé-occurrence* : la première est un objet théorique, construit par le linguiste et, on y reviendra, accessible aux sujets parlants ; le dernier est le produit de l'événement historique que constitue l'énonciation (non répétable, matériel, objet d'expérience), entièrement déterminé par la donnée de la phrase et de la situation concrète de l'énonciation ; quant à l'*énoncé-type*, l'objet d'étude de la sémantique, il peut être considéré comme une classe d'équivalence d'énoncés occurrences, *modulo* une relation d'équivalence sur les situations concrètes. Cette relation d'équivalence rend compte de la catégorisation que les sujets parlants font des situations d'énonciation lors du processus de compréhension.

Aux *énoncés-occurrences* sont associés des *interprétations*, aux *énoncés-types*, des *sens*, et aux *phrases*, des *significations*. Une même signification est donc susceptible de donner lieu à plusieurs sens, et un même sens, de donner lieu à plusieurs interprétations.

Ainsi, l'interprétation des énoncés occurrences dépend de la situation précise dans laquelle ils sont proférés et de la situation précise dans laquelle ils sont interprétés ; en particulier, l'interprétation dépend des objectifs et des intérêts de l'interprétant. De même, les objets d'expérience sont perçus par les observateurs de telle manière que la représentation qu'ils s'en font est co-déterminée par la situation précise dans laquelle l'objet est observée : la perception de l'objet d'expérience dépend, comme pour l'interprétation des énoncés occurrences, des objectifs et des intérêts des observateurs. L'énoncé occurrence n'est-il pas d'ailleurs lui-même un objet d'expérience ?

En continuant l'analogie, on constate que, de même que le sens d'un énoncé-type dépend du type de situation dans lequel il est employé et repose donc sur une certaine catégorisation du réel, les caractéristiques d'un objet d'étude dépendent de la manière dont l'observateur découpe l'observable.

⁶ Cette distinction a été exposée en détail ici-même en 1996.

Enfin, de même que la signification d'une phrase est abstraite des situations et même des types de situations dans lesquelles cette phrase peut être énoncée, de même, l'objet scientifique est conçu pour transcender cette expérience. Cette transcendance, loin d'être donnée *a priori*, nécessite, comme cela a été montré en détail dans Bailly *et al.* (1995), un travail de décontextualisation qui ne peut porter ses fruits que dans un type particulier de situation.

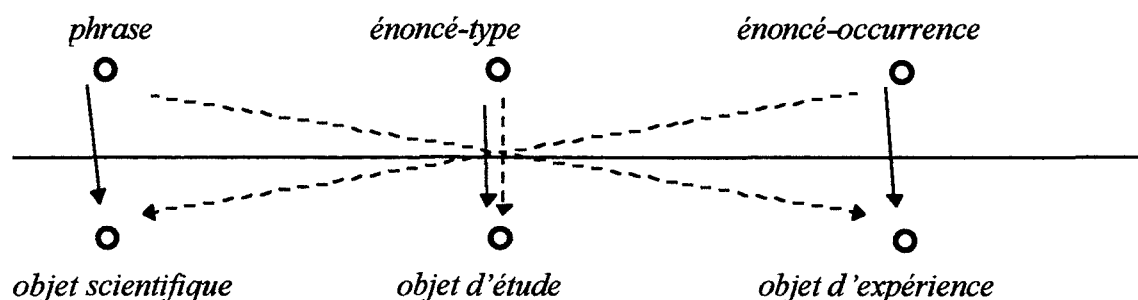
Ainsi, en suivant cette analogie, on est amené à concevoir les occurrences d'énoncés comme des *objets d'expérience*, les énoncés-types comme des *objets d'étude* (des *observables*), et les phrases comme des *objets scientifiques*. Une des particularités caractéristique des langues est que, contrairement à ce qui se passe pour, par exemple, les disciplines relevant des sciences de la matière, la reconnaissance, par un sujet quelconque, d'une occurrence d'énoncé *en tant qu'occurrence d'un énoncé* implique que ce sujet la conçoive comme une manifestation d'une phrase ; en ce qui concerne la matière, en revanche, les sujets peuvent reconnaître un objet d'expérience sans pour autant le concevoir comme un objet scientifique.

On pourrait dire, ainsi, qu'une des caractéristiques des langues est que les sujets qui les manipulent sont, *par le fait même qu'ils les manipulent*, des théoriciens.

Il y a cependant une autre analogie possible. Nous avons vu qu'un même objet d'expérience peut donner lieu à plusieurs objets d'étude et qu'un même objet d'étude peut donner lieu à plusieurs objets scientifiques. Il en va de même du côté de la langue : une même phrase peut donner lieu à plusieurs énoncés-types et un même énoncé-type peut donner lieu à plusieurs énoncés-occurrences.

Remarque :

Le lecteur a sans doute vu que cette analogie selon les potentialités d'actualisation est à contresens par rapport à l'analogie précédente, selon le degré de théorisation. On a ainsi le schéma suivant :



Ce fait peut paraître choquant de prime abord (remarque d'apparent bon sens : « il faudrait savoir : ça va dans un sens ou dans l'autre ? »). Mais, à y réfléchir, cette sensation qu'il y aurait contradiction ne pourrait s'expliquer que par la croyance – peu avouable selon nous... – selon laquelle si l'on peut ordonner des concepts selon une certaine relation, c'est parce qu'ils sont (ontologiquement) dans cet ordre. Sans cette croyance mystique ou, tout au moins naïvement objectiviste, le fait que deux séries de

concepts soient associées terme à terme selon une certaine relation ne renseigne pas sur les possibilités d'association selon une autre relation.

En fait, ce que cette seconde analogie donne à voir c'est que les déterminations idéologiques et culturelles fonctionnent, avec les situations d'interprétation des énoncés des phrases, comme des cadres théoriques pour l'interprétation. L'interprétation des énoncés serait ainsi soumise à ce qu'on pourrait appeler, en suivant Johnson et Lakoff⁷ des « théories populaires », au regard desquelles les sens constitueraient des *objets théoriques* et les occurrences d'énoncés des *objets d'expérience*.

Ainsi, selon la deuxième analogie, l'interprète serait un observateur théoricien qui construirait un objet théorique (le sens d'un énoncé) en s'appuyant sur un cadre théorique (caractérisé par la culture de l'interprète et la situation d'interprétation), à partir d'observations (les énoncés-types), elles-mêmes construites à partir d'objets d'expérience (les occurrences d'énoncé). Mais en tenant compte de la première analogie, on considérera le scientifique comme un interprète qui construit un sens (l'objet théorique) en s'appuyant sur sa culture (son cadre théorique), à partir d'énoncés-types (ses objets d'observation), eux-mêmes construits à partir d'occurrences d'énoncés (les objets d'expérience).

RÉFÉRENCES

Bailly, Francis, Lorenza Mondada et Pierre-Yves Raccach (1995).

Connaissances scientifiques et techniques, décontextualisation et re-contextualisation. Actes de la 5^{ème} École d'Été de l'Association pour la Recherche Cognitive : *Les rôles de la situation et du contexte dans la cognition*, vol. B. Château de Bonas, Juillet 1995

Dummett, Michael (1988). *Ursprünge der analytischen Philosophie*. Traduction par Marie-Anne Lescourret : *Les origines de la philosophie analytique*. Paris, 1991, Gallimard.

Johnson, Mark et George Lakoff (1980). *Metaphors we live by*. University of Chicago Press.

Raccach, Pierre-Yves (1995). Argumentation and knowledge : from words to terms. À paraître dans les actes du congrès *Languages of Sciences*, Bologna, Octobre.

Raccach, Pierre-Yves (1996). Lexique, conscience et société. Exposé aux *Journées de Rochebrune*, janvier-février.

⁷ « Folk theories ». Johnson & Lakoff 1980.

Dictyostelium discoideum, passage d'une identité
individuelle à une identité collective.

Glottin Hervé - IDIAP - glottin@idiap.ch
Pinel Philippe - ENST / Université Paris VI
Jean-Luc Cochard - IDIAP - cochard@idiap.ch
Rafael Laboissière - ICP - rafael@icp.grenet.fr

Communication aux journées de Rochebrune 1997
Rencontres interdisciplinaires sur les systèmes complexes naturels et artificiels
"Invariance, Interaction, Référence : l'identité en question"

Résumé

Parmi les systèmes naturels qui ont la capacité à partir d'interactions individuelles de faire émerger des processus collectifs sophistiqués, le *dictyostelium discoideum* est un exemple saisissant. En effet son cycle de vie est marqué par l'agrégation spontanée, d'individus de type amibe, en un organisme pluricellulaire reproducteur. Sur ce support biologique nous traitons de la question de l'identité et d'interaction environnementale. Puis nous abordons la question de la cognition et de l'identité sous l'angle constructiviste. Enfin sur des considérations de morphologiques et de fonctionnalisme, nous tenterons de donner des définitions de l'identité, qui seront analysées dans le cadre du *dictyostelium discoideum*. Nous ferons l'usage d'expériences mentales, d'intuitions contrôlées par la géométrie, pour l'approche de certaines définitions de l'identité.

Mots clefs :

identité, interaction, référence, constructivisme, invariance, autopoïèse, environnement, émergence, langage, morphogenèse

1 Introduction

La notion d'identité se retrouve à différents niveaux d'un même organisme. Ces niveaux sont imbriqués, et indépendants car ils mettent en jeu des processus de reconnaissance différents : identité génétique, cellulaire, immunitaire, psychique, sociale...

Beaucoup de systèmes naturels (Crutchfield & Mitchell, 1994. et Bonabeau & al., 1995), ont la capacité à partir d'actions individuelles en interaction, via un système de communication, de faire émerger des processus collectifs sophistiqués. Par exemple citons l'activité dans les nids d'insectes sociaux, ou le processus sensoriel parallèle et distribué dans la population corticale de neurones. Mais un exemple saisissant est l'agrégation spontanée suivant des contraintes environnementales, d'individus unicellulaires de type amibe en un organisme pluricellulaire reproducteur, durant le cycle de vie du *dictyostelium discoideum*, marquant le passage du stade unicellulaire à celui de la vie pluricellulaire.

Nous traitons à travers ce dernier exemple la question de l'identité.

Tout d'abord nous donnerons la description du cycle de vie du *dictyostelium*, base de notre étude.

Puis nous situons notre problématique sur plusieurs niveaux. Le premier est l'étude du passage concret d'un groupe d'individus unicellulaires, autonomes et identiques, à un organisme pluricellulaire avec différenciations cellulaires. Nous le doublons d'une réflexion sur le changement d'identité qui en découle vis à vis de l'environnement.

Le second niveau de notre problématique présente le cas du *Dictyostelium* et de l'identité dans l'optique du courant constructiviste.

Enfin sur des considérations morphologiques et de fonctionnalisme, nous tenterons de donner des définitions de l'identité, qui seront analysées dans le cadre du *dictyostelium dicoideum*.

Nous soulignons le caractère "arbitraire" des différents niveaux d'identités, et souligner que chacun d'eux a une sorte d'autonomie, une intégration des niveaux inférieurs à lui, une sorte de "loi de composition interne" (les règles de la biochimie ne sont pas celles de la psychologie...)

Plusieurs niveaux d'identité seront étudiés. En particulier, nous pouvons décrire deux classes principales. L'identité sera abordée comme étant le soi (par opposition au non soi), mais aussi comme étant l'invariance de l'être(l'identique par opposition au différent).

2 Description du Dictyostelium

2.1 Le cycle de vie du Dictyostelium

Le *Dictyostelium* est un genre d'amibe (myxomycète), étudié depuis longtemps car facilement maintenu en culture. Ils croissent sur le sol et se nourrissent de bactéries et d'autres matières organiques. Le cycle de vie du *Dictyostelium* dans un environnement stable et favorable est celui d'un unicellulaire classique, c'est-à-dire avec une croissance par division cellulaire. Il exhibe des fusions d'individus, des phases haploïdes et diploïdes.

En revanche, si le milieu devient défavorable (carence alimentaire), le cycle de différenciation s'amorce. Les cellules génèrent de l'ATP, et le chemiotactisme vers les centres d'agrégation commence à se manifester 6 h après le début de la diffé-

renciation. Ce cycle est illustré par la figure 1. Il y a convergence des individus en pôles comportant alors chacun en fin d'agrégation (après environ 10 heures) près de 10^4 individus. Les cellules ne fusionnent pas les unes aux autres mais conservent leur individualité. Environ deux heures plus tard, un bourgeon se forme, il contient les cellules qui deviendront des spores. Dans certaines conditions il y a une migration de cet ultra-organisme que nous nommerons limace sur une distance de 0.5 cm à 6 cm (Raper, 1976) vers la lumière ou la chaleur. Cette migration peut durer plusieurs jours (Darnell et al.). Puis l'étape finale de la différenciation est un allongement de la tige, et la remontée des spores vers le haut. Simultanément les spores sécrètent des substances qui forment un revêtement rigide autour de chacune d'elles. Les spores perdent de l'eau et leur activité métabolique devient nulle. Le corps végétatif mature contenant environ 70 000 spores, est supporté par un pédoncule constitué d'à peu près 30 000 cellules. Les spores matures sont beaucoup plus résistantes à la déshydratation et aux rayons ultra violets que les autres cellules. Si il y a apport adéquat d'aliments, un spore germe, son revêtement se rompt et une amibe viable est libérée. Les cellules similaires initialement se différencient en deux

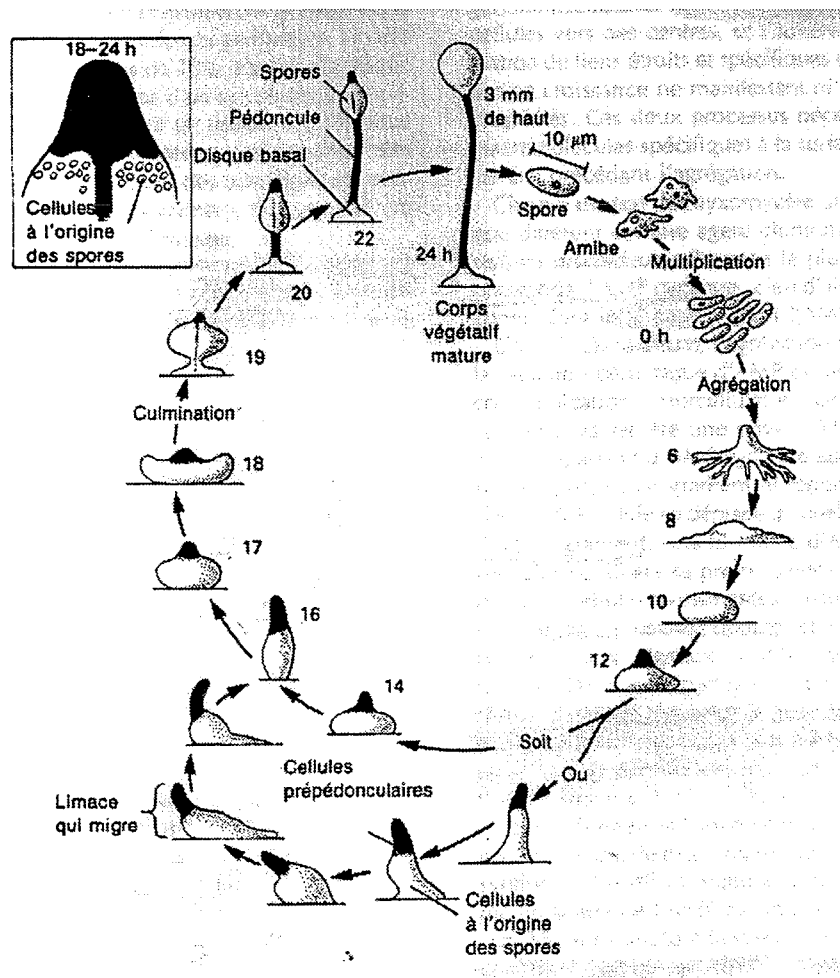


FIG. 1 - Le cycle de différenciation du *Dictyostelium*, d'après Darnell et al. Les trois grandes phases sont visibles: de la spore à 0 h, le stade unicellulaire, puis jusqu'à 10 h la phase d'agrégation, et enfin le stade pluricellulaire

sous populations de cellules, l'une qui compose le corps d'un être pluricellulaire, et

l'autre qui compose sa partie germinale. On a alors un véritable organisme pluricellulaire, possédant un support (cellules qui mourront avec ce "méta-individu"), et un appareil fructifère contenant les spores. Ces derniers, surélevés grâce au support de quelques millimètres de hauteur, pourront être essaimés vers d'autres lieux, peut-être plus favorables, et par division cellulaire, reformer des *Dictyostelium* au stade unicellulaire.

2.2 Interactions cellulaires

La mise en route du processus de regroupement décrit ci-dessus commence par l'activation de cellules dites "pionniers" (Nicolis & Prigogine, 1989), qui seront les premières à émettre un message aux cellules. Les interactions se font par le biais d'interagir avec les autres. C'est un mode de communication chimique, consistant en la sécrétion d'AMPc (Adénosine monophosphate cyclique). Ce composant va être présent dans chaque cellule, et sera à l'origine d'une structuration de la population dans le temps et l'espace.

Tout d'abord, chaque cellule qui capte l'AMPc de ses voisines va à son tour sécréter de l'AMPc. Il y a donc une amplification du message chimique autour du pôle des cellules "pionniers". Il y a donc formation d'un gradient qui va attirer et orienter tous les individus (chimiotaxie) comme l'illustre la figure 2

Enfin, ce regroupement va être structuré dans le temps, car il y a émission périodique d'AMPc au niveau cellulaire conséquence de phénomènes d'inhibition intracellulaire que nous ne détaillons pas ici (processus enzymatique). Il va donc y avoir des "vagues" d'AMPc à travers la population, qui vont, entre autre, réguler la progression de chaque cellule, puis leur différenciation (Devreotes, 1989).¹.

De plus il existe un système de communication chimique parallèle, visant à ne faire s'agréger que des cellules de la même espèce, c'est le rôle des CAM (specific Cell Adhesion Molecules).

2.3 Différenciation

Les propriétés de chaque cellule découlent de l'action de l'AMPc qui agit sur des effecteurs qui vont produire des messagers régulant l'expression des gènes. La concentration d'AMPc est donc à l'origine d'une différenciation cellulaire. Si elle est faible ou forte, si elle est maintenue constante ou appliquée périodiquement, ses effets inhiberont ou stimuleront les étapes du développement de l'organisme pluricellulaire.

3 L'identité à travers les trois grandes étapes du cycle du *Dictyostelium*

Le cycle de vie du dictyostelium peut se décomposer en trois grandes étapes (cf. Figure 1) : le stade unicellulaire, l'étape d'agrégation, et l'état pluricellulaire. A travers ces étapes nous allons développer la question de l'identité sous un premier

1. "Il est tentant d'imaginer que ces gradients fournissent aux tissus un "système de coordonnées" qui communique aux cellules individuelles de l'information positionnelle grâce à laquelle ces dernières peuvent reconnaître leur propre position vis-à-vis de celle de leurs partenaires." (Nicolis & Prigogine, 1989. chap 1.7).

aspect, certes arbitraire, qui est celui qui repose sur le caractère de ce qui est un, ou encore du soi, vis-à-vis du non soi.

3.1 L'état individuel:le stade unicellulaire

Dans le stade unicellulaire, chaque *Dictyostelium* constitue en lui même un individu autonome vis-à-vis des autres, autoreproductible, possédant donc une identité cellulaire et génétique. L'individu se limite à sa paroi cellulaire: autrement dit, cette paroi est l'interface entre l'environnement proximal et le génome de la cellule. Nous arrivons à l'idée que cet état unicellulaire est un état stable de l'identité. Un être unicellulaire se divise ou fusionne avec un autre pour faire d'autre unicellulaire, qui naîtront unicellulaire et qui finiront unicellulaires. Nous verrons plus loin état instable.

3.2 L'état agrégé

L'identité individuelle a laissé place à un niveau plus complexe. On peut dire que l'identité cellulaire est perdue car les cellules se sont engagées dans un cycle instable; elles forment un corps complexes, qui est une émergence, mais en même temps qui est une étape; ce corps n'est pas viable; on peut le voir non pas comme une "naissance", mais comme un appel suicidaire et ultime à la perpétration de l'espèce, une tentative qui aboutira à la mort de ce corps, et à un retour à l'être unicellulaire qui est un point d'équilibre.

Dans cet être pluricellulaire consistant en la structuration de la population, et en un changement de stratégie reproductrice, l'identité génétique persiste, car il n'y a pas formation de structure hybride entre plusieurs espèces de *Dictyostelium*. Cependant, l'identité cellulaire est perdue, les cellules sont différenciées, et donc elles sont dans une relation de dépendance. Elles ont perdu leur autonomie et une grande partie d'elles (le support est constitué de cellules mortes), n'est plus capable de s'autoreproduire. Le rôle d'interface entre environnement et génome est totalement changé: il est plus complexe. Autrement dit, le "soi" se concentre essentiellement autour d'un appareil fructifère, et l'interface consiste en une population de cellules différenciées contrairement à l'état décrit dans le paragraphe précédent.

Il faut cependant noter là un point fondamental. Cette structuration d'une population d'individus en un méta-organisme, ce "bouleversement" cognitif qui étend les limites de l'individu d'une cellule à une collectivité, n'existe qu'en tant que réaction à une contrainte de l'environnement. C'est un processus dynamique qui souligne particulièrement bien l'évolution et l'étroitesse du lien qui existe entre identité et environnement (Stewart, 1991).

3.3 L'agrégation

On a vu dans la première partie les mécanismes qui interviennent dans la structuration et la différenciation d'une population de *Dictyostelium*. Ce sont des systèmes stimuli/réactions simples, qui permettent l'interaction avec les génomes de la cellule et les génomes des autres cellules, via des systèmes protéiques (récepteurs, effecteurs...).

Il y a donc deux modes d'expression génomiques:

—dans la phase unicellulaire stable, il y a interaction du génome directement avec l'environnement intracellulaire, et indirectement avec l'environnement proximal extracellulaire.

—dans la phase d'agrégation, il y a interaction entre les différents génomes via le milieu extérieur. Cette interaction va aboutir à l'émergence d'un nouvel ordre pluricellulaire. D'une certaine façon, nous pourrions dire que le "soi" démultiplié et en interaction avec lui-même permet de passer à un niveau d'identité différent. Ces deux niveaux d'identité sont imbriqués, et portant chacun en eux leur propre définition du "soi" et leur propre stratégie d'autoreproduction. Car sinon comment l'identité génétique présiderait au changement de l'identité du soi? C'est bien le même mécanisme qui est à la base du changement, et non pas prise de pouvoir d'un mécanisme sur un autre.

4 Les diverses définition de l'identité dans le cadre du dictyostelium : identification de l'organisme par catégorisation.

4.1 Aspect construtiviste

L'aspect cognitif est représenté ici par des caractères tels que l'adaptation, la reconnaissance, la communication... mais il n'est bien évidemment pas symbolique. Tout comme elle le serait au niveau du système immunitaire, la cognition est non consciente, et découle de phénomènes naturels (Stewart, 1991).

Cependant, elle est intéressante dans le sens où elle nous permet de nous interroger sur l'évolution de l'identité. A chaque fois que le milieu environnant est défavorable, un processus est enclenché, faisant passer *Dictyostelium* d'un état unicellulaire à un état pluricellulaire, sous des contraintes externes, pour retourner ensuite au stade unicellulaire. Phylogenèse et ontogenèse sont donc ici étroitement imbriquées, si on laisse à l'état pluricellulaire la part ontogénétique². En réalité, on peut dire que les deux niveaux d'identité (individuelle et collective) sont omniprésents, se cédant la place à tour de rôle. En effet à l'état collectif succède aussi l'état individuel après l'éclosion des spores libérés par la limace.

La cognition se révèle donc être une capacité réactive; c'est l'environnement qui crée l'individu pluricellulaire. Il est issu de phénomènes physiques et chimiques, tels que:

- chimiotaxie et amplification du signal chimique,
- gradient et processus d'agrégation,
- multiplication de l'identité cellulaire.

Il faut souligner le lien qui existe entre multiplication de l'identité cellulaire et communication chimique. Il y a existence d'un langage commun du fait d'un même système stimulus/réponse partagé par toutes les cellules. En quelque sorte, c'est l'identique qui se répond à lui même.

Les *dictyostelium* ne partagent pas le même génome. Pourtant, ils ont le même type de rapport entre eux. Peut importe le génome entier. Ce qui compte, ce ne sont que quelques gènes responsables des communications (gènes du récepteur...), que chaque *dictyostelium* doit partager pour pouvoir agir comme des agents coopérant.

4.2 Questions sur la reproduction et l'identité

Parallèlement à ces deux niveaux d'identité décrits précédemment, on peut juxtaposer deux niveaux de reproduction: par division cellulaire, et par formation de

2. "Il est à noter que l'ontogenèse, intercalée entre évolution et autopoïèse, fait son apparition seulement avec les premiers organismes pluricellulaires; elle se complexifie progressivement tout au long de l'évolution ultérieure, et contribue certainement à l'accélération du rythme de celle-ci." (Stewart, 1991, p 10)

spores. Dans les deux cas, le rôle de la cellule n'est pas le même. Son rôle dans l'identité qu'elle représente non plus. La cellule ne vit plus pour elle-même, elle peut être sacrifiée pour une structure plus complexe qui porte le génome.

L'individu se réalise donc en tant que porteur optimal de son génome, qui est lui-même l'organisateur de cette structure autopoïétique, quitte à y perdre son autonomie.

Ce comportement "social" de coopération et la sporulation ont dû coévoluer. Cette capacité à se reproduire va définir les états stables/instables, et va donc décrire la stratégie trouvée comme réponse à l'environnement.

4.3 Identité et interactions

L'aspect intéressant à envisager, toujours dans l'optique de mieux définir ce que peut être l'identité, est de la voir comme le lien entre un individu (qui se définit par la présence de l'identité elle-même) et son environnement. Nous allons détailler cette idée dans les deux cas du *Dictyostelium*.

Dans le cadre unicellulaire, le sujet est la cellule, son environnement est le paysage physicochimique qui l'entoure. Plus généralement, on peut dire que le reste de la population de *Dictyostelium* fait partie de l'environnement d'un seul *Dictyostelium*, donc de son non-moi. En revanche, la cellule unique forme aussi un espace, composé d'éléments en interaction (protéine, gènes...), qui est son propre "soi". Cet espace d'interaction entre compartiments cellulaires définit l'identité du *Dictyostelium* unicellulaire.

Dans le cadre pluricellulaire, il y a communication, interaction, et dépendance des cellules les unes envers les autres. Elles sont donc toutes rassemblées en une activité partagée, de reconnaissance du soi, face au milieu extérieur reconnu comme non-soi. En revanche, au niveau moléculaire l'identité du *Dictyostelium* est perdue.

De ces remarques, on peut définir ici l'identité comme étant les frontières d'un espace relationnel, ou encore comme une somme d'objets partageant un même espace d'interactions.

La remarque précédente permet de dire aussi que l'identité est un espace de "communication partagé", que cette caractéristique permettrait d'associer une identité à une région de l'espace.

4.4 Définition des interactions

Dans la suite, nous allons développer des aspects propres à la théorie des catastrophes de R. Thom, ce qui demande quelques définitions que nous donnons ici.

Considérons les contraintes qui pèsent à la fois sur l'organisation spatio-temporelle de l'être vivant et sur sa régulation physiologique comme le fait René Thom dans l'apologie du Logos. Toutes ces contraintes sont les conséquences d'un seul principe : le principe de localité qui s'exprime ainsi : "tout processus interne à un organisme doit pouvoir se décrire comme résultant de déterminismes locaux" (Thom page 144). Ce principe se présente sous deux formes.

La première est la forme continue : les contraintes portent sur la propagation spatio-temporelle de champs (diffusions moléculaire, champ électromagnétique). Très souvent ces champs ont une signification biologique à valeur régulatoire, que René Thom appelle "prégnances".

La seconde forme est discrète : il s'agit d'interactions entre entités locales contiguës (comme des cellules par exemple), séparées par une interface (membrane ou paroi).

René Thom alors parle de formes individuées douées de "saillance". Les déterminismes locaux sont alors de deux types:

- Une prégnance investit une forme saillante compétente pour ce champs et y introduit des comportements nouveaux. Par exemple un champ de potentiel sur un tissu va y créer par ramification (ici amplification de diffusion de l'AMPc), des différenciations qui donneront naissance à des frontières entre régimes locaux, d'où l'apparition de formes saillantes nouvelles. - Une saillance, par une modification de sa forme ou de sa position, produit une modification topologique importante (scission, ramification, concentration) dans le flux énergétique d'une prégnance.

La propagation par contact des champs dans l'espace temps dans l'espace substrat est la communication proprement dite. C'est par onde, que l'on nomme prégnance, qui ricochent sur des bulles qui elles sont saillantes, individualisables et qui sont les *dictyostelium* proprement dits que les *Dictyostelium* se déterminent aussi un territoire beaucoup plus vaste que leur espace interne, ils se déploient par une activité extra corporelle qui est leur langage. Est ce que l'identité prend aussi un sens dans une forme prégnante? La forme est une identité visible, la force, i.e. le langage est une entité invisible qui produit des effets visibles. Elle est liée à l'irréversibilité du temps.

4.5 Perception, reconnaissance des formes et de l'identité:morphogenèse et naissance d'une identité, ou comment rompre son identité en se déformant.

Donner une identité à un objet c'est en faire un individu. Or ce qui caractérise l'individu c'est qu'il est limité par une paroi. D'un côté de cette paroi c'est l'intérieur de l'individu, de l'autre côté c'est le monde extérieur.

Le problème de l'identité peut être rattaché à celui de l'unité d'une forme. La conductivité topologique est une condition nécessaire mais la distinction intérieur-extérieur est fondamentale pour définir des formes d'être vivant. Il est alors intéressant de définir l'identité d'un être vivant par rapport à sa forme, chaque être vivant a une forme. Posons donc cette définition de l'identité: l'identité d'un être vivant est donnée par sa forme. Le passage "continu" d'une forme à une autre n'interfère pas dans l'identité de cet être. Par continu nous entendons que la déformation est à chaque instant "réalisable mentalement" par celui qui s'interroge sur l'identité de l'être observé. "Réalisable mentalement" signifie que l'observateur associe l'objet observé à celui identifié précédemment.

Mais qu'est ce qu'une forme au juste? La forme se définit en se séparant de son fond, à savoir qu'elle présuppose toujours une discontinuité qualitative (R. Thom p183) qui sépare les points de l'espace contenus dans la forme de ceux qui n'y sont pas.

On va supposer maintenant que l'on a affaire à un milieu naturel occupant un domaine D de l'espace. Mais alors, l'espace où vivent les *Dictyostelium* isolés est une forme percevable par un observateur, tout comme l'est l'amas agrégé de *Dictyostelium*. Si l'on se fixe comme définition une forme pour une identité, nous voici en possession de deux formes. De plus, l'une se déformerait continûment pour former la seconde (l'amas), mais ne serait ce pas alors une seule forme, une seule identité? Il nous faut trouver une rupture de continuité dans cette déformation si l'on veut parler dans ce cadre de définition de l'identité, d'un passage d'une identité à une autre. En fait dès que l'amas est formé entièrement, on peut parler d'une nouvelle forme: dans l'espace D de départ dans lequel on avait les *dictyostelium*, la première forme donc, en dehors de laquelle il n'y avait pas de *dictyostelium* (imaginons un espace vide, dépourvu de vie par exemple). Dès que l'amas est formé, et pas avant,

on peut définir une nouvelle forme, fermée, en dehors de laquelle les *dictyostelium* restants dans le domaine D ne signifient rien, même dans l'espace intérieur que décrivait la première forme. La première forme est celle qu'adopterait un gaz dans un récipient : elle occupe tout l'espace vital, cette forme adopte la forme de l'espace vital. La seconde forme a plutôt le comportement d'une goutte d'eau, et sa forme est régie par certaines contraintes supplémentaires à celles des limites du milieu vital (cohésion des cellules). On peut parler alors d'un changement de phase entre les deux formes, condensation d'une forme gazeuse à une forme liquide. La transition entre ces deux phases d'état d'organisme n'est pas réalisable mentalement. Il y a une discontinuité dans le changement de forme, ce qui du point de vue de cette définition de l'identité assure l'existence de deux identités distinctes.

4.5.1 Vers une définition proportionaliste de l'identité

Si comme nous l'avons proposé, l'identité d'une entité peut être rattachée au niveau de sa frontière avec son environnement, alors nous pouvons aborder l'identité comme étant un rapport entre l'entité et la "non entité". Si nous cherchons à regrouper nos considérations sur le langage et l'individu l'émettant, dans une proposition de définition de l'identité, nous pouvons poser des relations entre plusieurs termes. Les termes en questions sont le langage, le corps (l'individu, l'objet) et le monde (espace). Les différents niveaux d'identité qu'induisent ces éléments en interactions peuvent en fait être différents niveaux.

Un langage n'est pas localisé dans l'espace, il règne dans un espace ouvert. Mais un individu est localisé par son corps, il règne dans un espace fermé connexe. Pourtant l'identité attachée à une entité ou bien l'information qu'elle engendre sont des objets de pensée identique. Nous avançons alors les relations suivantes :

Identité = individu / corps (espace fermé) = langage / monde atteint (espace ouvert) = Objet / espace où règne l'objet

4.6 L'observateur et identité de l'observé

Le niveau de connaissances de l'observateur et l'interaction de ses croyances avec sa perception du monde (la perception est-elle influencée par ce que nous croyons?) sont des éléments fondamentaux dans le processus d'identification par cet observateur d'une entité.

L'identité qu'un observateur associe au *Dictyostelium* est vouée au concept de l'identité à priori de cet observateur. L'identité est définie par l'autre, donc dépendante de la perception de "l'autre". Deux cas se présentent : a/ Si l'observateur recherche l'identité dans les caractéristiques du génome, chaque *Dictyostelium* a une identité différente. Mais cette identité est elle même conservée dans la limace, on retrouve le même génome dans chaque *Dictyostelium* la composant, donc la même identité. C'est une observation de l'état dynamique interne du *dictyostelium*.

b/ Si l'observateur cherche l'identité à une échelle de forme supérieure, il est les contours de l'organisme, alors chaque *Dictyostelium* est une identité différente également, mais ces identités se fondent en une dans la limace. C'est une perception d'objet dans un milieu donné.

Dans les deux cas l'observateur peut décrire le phénomène d'agrégation comme une coordination des organismes orchestrée par l'échange de signaux.

Dans le cas (a) il observe une communication de *dictyostelium* avec eux mêmes depuis le mécanisme d'agrégation jusqu'à la différenciation de la limace.

Dans le cas (b) il observe la communication d'un langage externe lors de la phase d'agrégation, et d'un langage interne à la limace qui coordonne au sein d'un même

individu les opérations de différenciations .

La représentation que l'observateur se fait de ce langage est donc dépendante de l'identité qu'il donne au *dictyostelium*. Réciproquement le langage offre également une caractéristique de l'identité L'identité qu'on associe à une chose dépend de notre niveau de connaissance sur cette chose (la connaissance scientifique a tendance à diviser un corps en fonctions, sous partie interne.... à décomposer.) Un observateur généticien pourrait reconnaître toute forme de dictyostelium à partir de son génome que ce soit, sous forme d'unicellaire ou de pluricellulaire. Il se base sur une identité continue dans le temps et l'espace. Une connaissance locale lui suffit. Il a définit identité à partir du soi génétique. Un observateur naïf ne peut voir dans les stades unicellaire et pluricellulaire une même espèce qu'en les observant de façon continue. Il lui faut relier par une relation causale unicellulaires et pluricellulaires. Il lui faut une vision globale. Il se sert de l'identité au sens de similarité.

4.7 Différenciation et identité

4.7.1 Quelques notions sur la différenciation cellulaire

L'être agrégé dictyostelium est formé par des cellules indépendantes, c'est ce qui est à opposer avec un embryon est en effet dans un embryon une masse de cellules hautement ordonnée qui provient de la division d'une seule cellule : l'oeuf. Les fonctions et les organes eux mêmes se construisent progressivement chez l'embryon. Les "décisions" qui permettent à des cellules embryonnaires d'adopter diverses voies de développement relèvent du processus de différenciation cellulaire. Une cellule est dite déterminée lorsqu'elle a été informée ou qu'elle a décidé elle-même de son futur rôle de cellule spécifique. La différenciation d'une cellule déterminée est due à des gènes exprimés de manière sélective et dont les produits permettent à la cellule d'exprimer un phénotype distinct. L'engagement est la décision de débiter la différenciation.

La différenciation n'est pas un phénomène spécifique des organismes pluricellulaires. De nombreux micro-organismes unicellulaires comme le *Dictyostelium* peuvent subir des changements physiques modifiant leur physiologie sous la contrainte de modification de l'environnement, comme lors du processus de sporulation. Dans le cas du *Dictyostelium*, la sporulation nécessite une différenciation.

4.7.2 Protozoaires et pluricellulaires

Une définition commune chez les biologistes pour définir un groupement de cellule comme restant plusieurs unicellulaires groupés ou bien comme formant un être pluricellulaire repose sur la dépendance des cellules de ce groupe. Ainsi dans un regroupement cellulaire, il faut faire la distinction entre colonie cellulaire qui est formée de plusieurs êtres unicellulaires, et entre être pluricellulaire qui est formé de plusieurs cellules, mais qui est une entité à lui seul.

Ce sont la nature des relations entre les cellules d'un groupe qui donne la frontière entre unicellulaire et pluricellulaire. Si une cellule dans un groupe assume tous ses besoins physiologiques (respiration, nutrition, reproduction ...), alors le groupe en question formé de telles cellules est appelé une colonie. En revanche s'il y a différenciation de la cellule avec spécialisation, et donc dépendance des cellules entre elles, alors le groupe de cellule est un pluricellulaire. Il découle de ces définitions que dans une colonie chaque cellule vit pour soi, en être autonome. Il en est tout autrement dans un pluricellulaire où chaque cellule n'existe que pour le tout.

4.7.3 Lorsque le *dictyostelium* se différencie

Dans la première phase, celle de l'agrégation, il est à noter qu'un *dictyostelium* ne se déplace pas vers un autre *dictyostelium*, mais que dans le sens du gradient le plus fort du champ d'AMPc. Ainsi, ce n'est pas une reconnaissance d'un *dictyostelium* voisin, mais uniquement une réaction locale, indépendante de l'identité de celui qui créa ce champ (ça peut être un humain). Le *dictyostelium* ne répond pas à une identité, car il ne peut reconnaître l'origine de ce champ. Ce n'est qu'une fois agrégé, que la saillance se crée et qu'il y a reconnaissance et donc naissance du phénomène de devenir autre, différent : il y a différenciation. Une identité rencontre la même identité. A priori elle était identique lorsque l'on regardait les deux *dictyostelium* au niveau dispersé. Elles entrent en saillance, et alors se déterminent une identité différente, pour se reconnaître ? pour se compléter, peu être simplement pour se coordonner, il faut qu'elles soient différentes, car doivent se reconnaître, et ainsi permettre la construction d'une topologie interne, l'unification des déplacements etc...

L'agrégat est solidement pluricellulaire lié par des protéines que synthétisent les cellules. Le rôle des CAM (cf. section 2.2) est alors déterminant. L'identité de l'espèce est définie par la reconnaissance implicite lors de l'adhésion des cellules gérée par les CAM.

Peu après la formation des agrégats, une pointe apparaît sur la limace (cf. Figure 1). Dès ce moment la limace devient mobile et ses cellules sont déterminées : chacune d'entre elle est maintenant un précurseur de l'un des deux types cellulaires de la fructification : les cellules des spores et celles du pédoncule. Ce phénomène réversible est démontré par l'expérience de la figure 4

Les résultats montrent en effet que dans l'agrégat, au moins quelques cellules sont engagées dans la formation de spores au dans celle du pédoncule. Cependant certaines des cellules du pré-pédoncule peuvent se redifférencier en spores et certaines cellules de la pré-spore, dans la partie arrière de la limace peuvent se redifférencier en cellules pédonculaires. Comment expliquer qu'une cellule choisisse de se différencier en spore ou en pédoncule ? Deux modèles peuvent servir à comprendre ce changement de fonction (Darnel et al.). Dans le premier les cellules sont déjà déterminées à devenir une spore ou un pédoncule avant de former l'agrégat pluricellulaire, et dans celui-ci elles migrent sélectivement. Dans le second, toutes les cellules dans l'agrégat sont toutes initialement semblables, mais un gradient de substance est établi d'une extrémité à l'autre de la limace. Le devenir des cellules serait alors dépendant de ce gradient.

Est-ce l'environnement de la cellule qui détermine son identité ? Ou bien des cellules déjà déterminées sont-elles responsables des caractéristiques de l'environnement cellulaire ? La solution à ce problème ne pourra être fournie que par l'identification des gènes responsables de l'existence de deux états cellulaires différents et par l'analyse du mécanisme qui déclenche l'expression de ces gènes. Aucun problème de cet ordre n'a encore été résolu.

Mais en ce qui nous concerne nous pouvons donc conclure que les *dictyostelium* unicellulaires sont tous identiques, dispersés, mais qu'ils ont une identité propre, activité physiologique indépendante, etc... Une fois agrégé, ils coordonnent leur états internes, leur physiologie, et alors n'existent plus en tant qu'identités propres, mais forment à eux tous une seule identité.

4.8 Identité et fonctionnalisme

La capacité du *dictyostelium* à réaliser des tâches spécialisées grâce à une différenciation, démontre comment une coopération cellulaire permet de faire varier la

structure et la fonction d'un organisme. Le groupement de cellule devient organisme pluricellulaire quand on commence à y voir apparaître une structure décomposable en organes, en fonctionnalités, et que le niveau cellulaire n'est plus suffisant pour décrire la structure. L'identité (ou niveau d'identité pertinent) pourrait donc être définie comme ce qui permet de modéliser, de caractériser le plus simplement possible le système étudié (l'organisme..).

Autrement dit, dans l'organisme pluricellulaire, la cellule devient inférieure au niveau "utile" de description. Identifier par sa fonctionnalité un organisme, ou une partie de celui-ci serait une autre définition de l'identité.

Comme le pense R.Thom (p 260), " tout comme pour la scientificité dans les activités cognitives, une telle frontière n'existe pas, mais qu'il peut y avoir un "gradient" relativement bien défini dans la valeur régulatoire des activités organiques.". De plus, on ne peut parler d'une augmentation de complexité fonctionnelle chez le *Dictyostelium* de l'état unicellulaire à son état pluricellulaire. En effet si la complexité structurale, morphologique s'accroît car des fonctions deviennent visibles comme la motricité, la complexité fonctionnelle, elle, varie assez peu. En effet, l'amibe unicellulaire est capable de tropismes, elle est douée de sensibilité différentielle, tout comme la limace, qui se dirige vers la lumière ou la chaleur. De ces remarques nous pouvons déduire qu'une définition de l'identité relative aux fonctionnalités de cette entité, donne lieu à une constance identitaire du *Dictyostelium* lors de sa différenciation.

Le *Dictyostelium* après agrégation, regroupe les caractéristiques des systèmes pluricellulaires : le mouvement, les interactions (phototropisme), et différenciation cellulaire. Ses capacités sont augmentées (déplacement, tropisme) mais leur nature est la même. Ainsi peut-on parler de ce point de vue, d'une constance de l'identité.

4.9 Changement d'identité : schémas de transition identitaire

Le changement de la qualité des *Dictyostelium* isolés à $t=A$ en limace à $t=B$ peut poser le problème de la qualité du *dictyostelium* en un instant $t=C$ compris entre A et B . Quelle serait alors l'identité du *dictyostelium* en C ? A la fois celle des instants en A et en B ? Cette absurdité est levée en affirmant que l'identité est constante ou bien en affirmant que tout est changement. On peut donc adopter deux attitudes.

a/ On peut postuler que le changement est brusque et ne remplit aucun intervalle temporel.

b/ Ou bien on admet un temps indéfiniment divisible et on suppose que le changement consiste en une continuité d'états dont la différence est perceptible. L'identité est alors une qualité continuellement mouvante. C'est un état associé à un instant et non une définition d'une qualité intemporelle.

Des schémas de transitions identitaires pourraient être dessinés. Pour le cas a/, il s'agirait de deux plateaux parallèles pour deux types d'identités distinctes (unicellulaire et pluricellulaire), avec une transition orthogonale, instantanée reliant les deux plateaux. Dans le cas b/, il s'agirait d'une courbe sigmoïde qui serait la transition entre les deux plateaux, connotant l'effet de coopération de chaque unicellulaire pour l'émergence de l'être pluricellulaire. Ou bien même encore, si nous plaçons au niveau de l'identité de chaque être et non plus aux niveaux des différents types d'identité présents, alors nous pourrions concevoir un assemblage de sigmoïdes

5 Conclusion

Nous avons abordé deux aspects majeurs de l'identité. Celle au sens du soi, caractère de ce qui est un, et celle au sens de l'invariance de l'être, caractère de deux objets de pensée identique.

Nous avons montré dans cet exemple du *dictyostelium* le passage d'une identité individuelle, unicellulaire, à une identité collective. L'identité individuelle peut se perdre par interaction intercellulaire et l'émergence d'une relation de dépendance dans un organisme pluricellulaire possédant une identité collective. Autrement dit, le "soi" se concentre essentiellement autour d'un appareil fructifère qui alors est le porteur optimal du génome. Les échanges d'informations et les différenciations morphologiques entre cellules permettent de considérer les formes les plus évoluées non plus comme des colonies mais comme de véritables formes pluricellulaires.

Ce n'est plus être composé de qui donne une relation d'identité, mais bien plutôt le maintien des fonctions de l'organisme. Peut importe le nombre de cellules qui le composeraient, mais bien plus ce sont ces organes qui signifient son identité. Ainsi de ce point de vue l'identité du *dictyostelium* est constante. Ce point de vue serait peut être conservé si l'on parlait de la chenille et du papillon ...

Quand l'observateur, comment celui qui tenterait de rechercher des correspondances entre plusieurs définitions pourrait être cohérent avec lui même? Le rôle de l'observateur est déterminant. L'identité est dans l'œil de l'observateur.

De cette belle rythmique interne qui coordonne, organise ces îlots unicellulaires en un accord final qui est la naissance de l'individu limace, ne pourrait-on pas généraliser que toute identité émerge de coordination spatio-temporelle?

Le bouleversement cognitif engendré n'existe qu'en tant que réaction aux cellules voisines de référence, à la traduction d'un langage chimique. C'est une dynamique qui souligne le lien entre identité et environnement. De telles ruptures de symétrie conduites par des transitions dues à des réactions chimiques, constituent l'un des traits dominants de la vie, "susceptible de se ramener à des phénomènes d'auto-organisation physicochimique" (Nicolis, Prigogine, 1989. chap 1.7).

Pourquoi parler d'identité collective? Parcequ'un ensemble de *dictyostelium* différencies, ou rien d'autre qu'un nombre limité de gènes les relie génétiquement (seul leur langage est commun) est bel et bien une coopération intra-spécifique. C'est une identité collective, car ce sont plusieurs cellules à partager un groupe à qui on associe une identité, mais elles gardent cependant leur histoire individuelle, leur généalogie propre.

Remerciements

Nous remercions Anne Laure Monnier pour nous avoir initié à ses recherches sur le génome du *Dictyostelium*, et à Franck Lenormand pour nous avoir donné le goût au dictyo avec ses photos d'artiste. Nous remercions également John Stewart pour ses passionnantes discussions "constructivistes".
Merci à Erika Valencia, et Alvaro Cassinelli, pour leur relecture attentive et leurs arguments intéressants.

Bibliographie

Eric Bonabeau, Jean-Louis Dessalles et Alain Grumbach. (1995). *Characterizing emergent phenomena (1): a critical review*. Revue internationale de systématique. Vol 9, n 3, pp 327 à 346.

J.P. Crutchfield, Melanie Mitchell.(1994) The evolution of emergent computation. *Proceedings of the National Academy of Science*.SFI Technical report 03-012.

J.Darnell et al.(1988). *La cellule-biologie moléculaire* Decarie Editeur.

Richard Dawkins. (1976) *The selfish gene*.Oxford university press. Traduction française (1990) *Le gene egoiste*, ed.Odile Jacob.

P. Devreotes. (1989). *Dictyostelium discoideum* : a model for cell-cell interactions in development. *Science* , 245:1054.

Franck Lenormand. (1995). *Etude de la mort cellulaire programmee chez Dictyostelium Discoideum*. Rapport de stage de magistère ENS Ulm.

G. Nicolis, I. Prigogine. (1989). *Exploring complexity: an introduction* . R. Piper, GmbH&Co. KG Verlag.

K. B. Raper.(1984) *The Dictyostelids* . Princeton University Press.

John Stewart. (1991).*Epistémologie des sciences de la cognition et du langage* . Ecrits pour école d'été. Paris.

René Thom. (1990).*Apologie du logos*.Collection Histoire et Philosophie des Sciences. Hachette.

6 Annexe

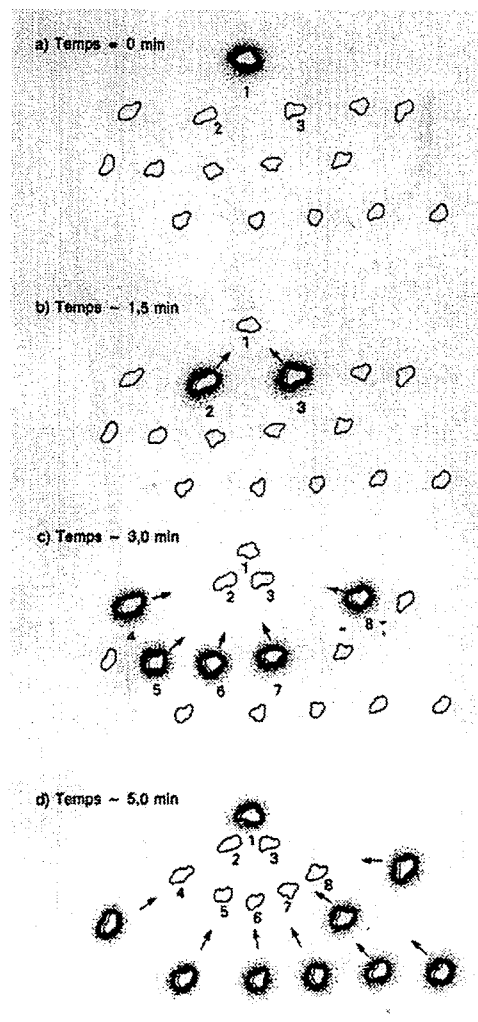


FIG. 2 - Communication entre les cellules de *Dictyostelium* à l'aide d'AMP cyclique, et convergence, d'après Darnell et al. a/ A $t=0$ une amibe libère une salve d'AMPc qui diffuse. b/ En $t=1.5$ min deux cellules adjacentes se déplacent de quelques micromètres vers la source d'AMPc. Elles libèrent leur propre salve d'AMPc et deviennent insensibles aux autres stimuli. Pendant ce temps l'AMPc libéré par la cellule en a/ est détruit par des enzymes. c/ En $t=3$ min, les cellules 4 et 5 se déplacent vers la cellule 2 et libèrent de l'AMPc et deviennent insensibles à l'AMPc, de même pour les 6, 7 et 8, vers la cellule 3. d/ $t=5$ min, d'autres cellules se déplacent, pendant ce temps la période réfractaire de la cellule 1 se termine et celle-ci synthétise et libère une nouvelle salve d'AMPc. Le cycle se répète.

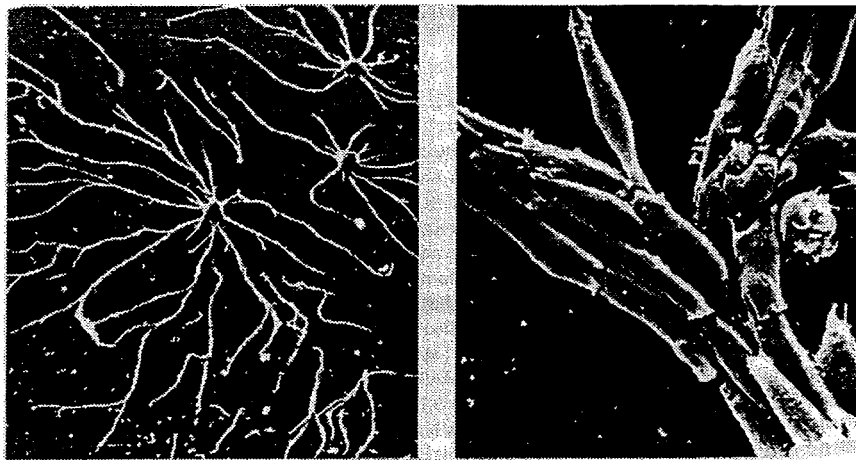


FIG. 3 – Micrographies électroniques à balayage d'un *Dictyostelium Discoideum* en cours de d'agrégation. A gauche un agrégat de *Dictyostelium Discoideum* en formation, grossissement $\times 50$. A droite une micrographie à plus fort grossissement ($\times 2000$) qui montre les cellules polarisées, reliées tête à queue et côte côte, formant un courant dirigé vers le centre de l'agrégation, d'après Darnell et al.

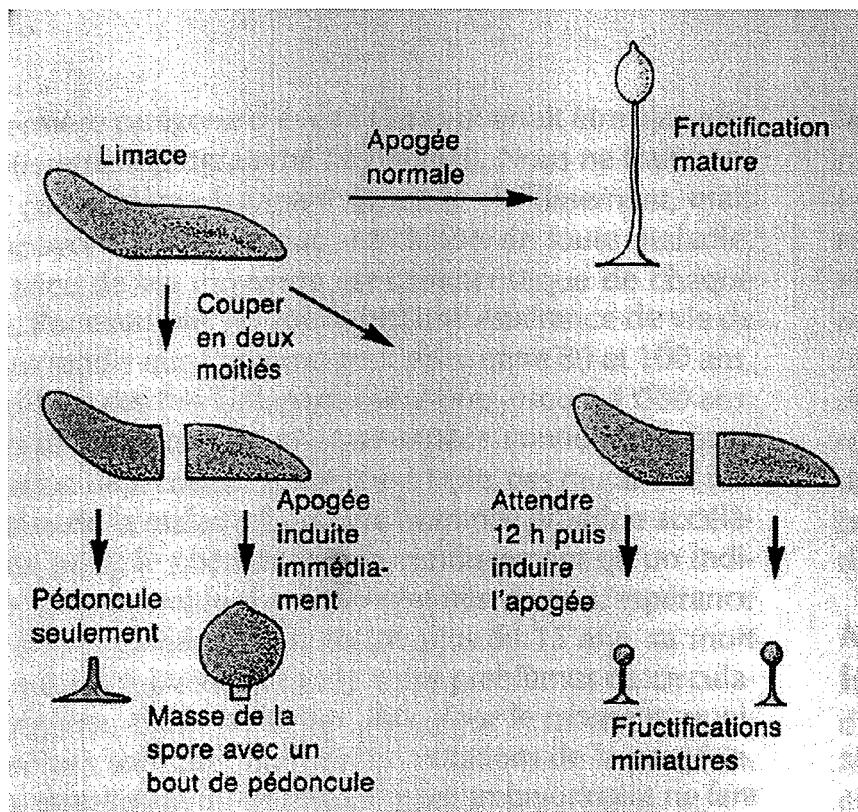


FIG. 4 – L'engagement des cellules de dictyostelium, d'après Darnell et al. Si une limace est sectionnée en deux parties et que l'apogée est induite immédiatement par un éclairage, la pointe se différencie en un pédoncule tandis que la portion arrière ne forme que des spores. Mais si l'apogée est différée de 12 h, chaque moitié se différencie en une parfaite fructification miniature et proportionnée.

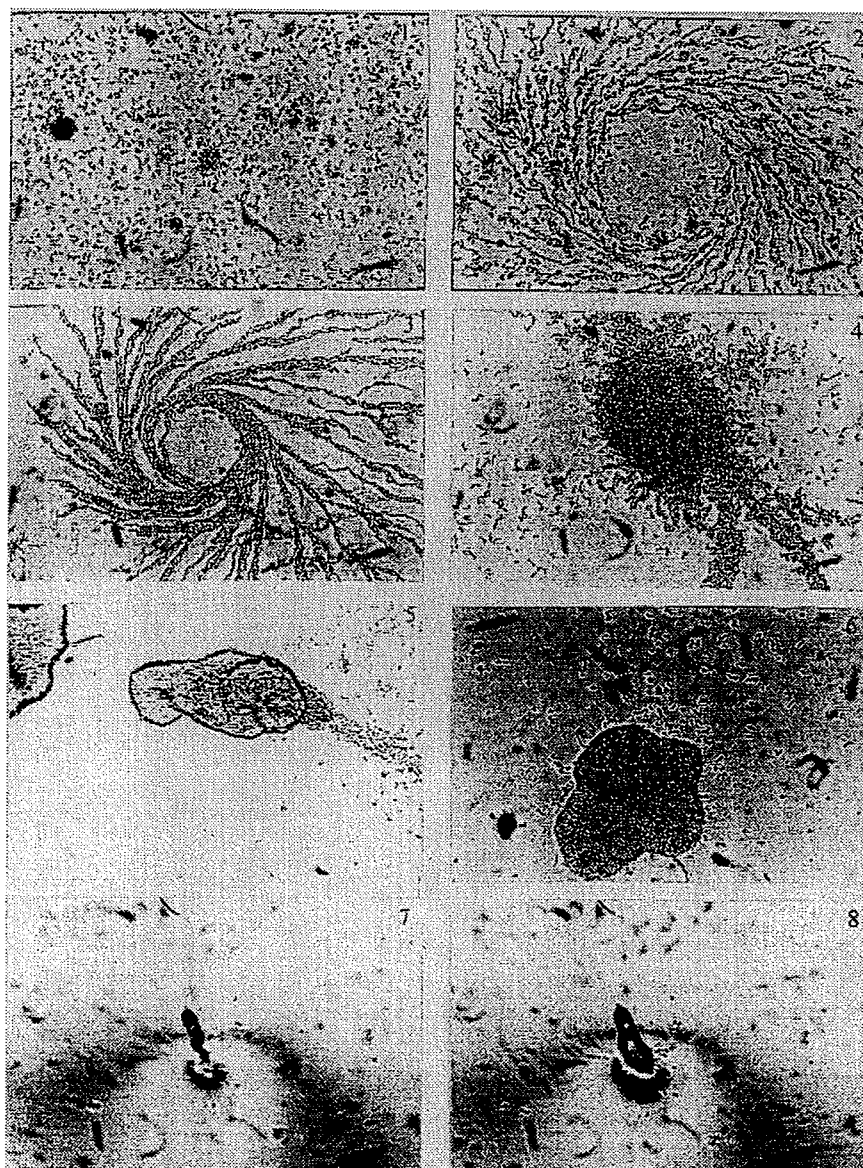


FIG. 5 – Annexe A : cinétique de différenciation et morphogénèse de *Dictyostelium Discoideum*, de 0 h à 16 h ,photos reproduites avec la permission de F. Lenormand. numérotation de gauche à droite, de bas en haut, grossissement $\times 10$; (1) Début de la phase de chimiotaxie; (2) 4h: Formation de spirales; (3) 6h: Spirales complètes; (4) 8h: Agglutinement; (5) 10h: Phase de mouvements amaeboïdes; (6) 12h: Arrêt; (7) 14h: Début de la morphogénèse; (8) 16h: Formation de la tête du sporocarpe

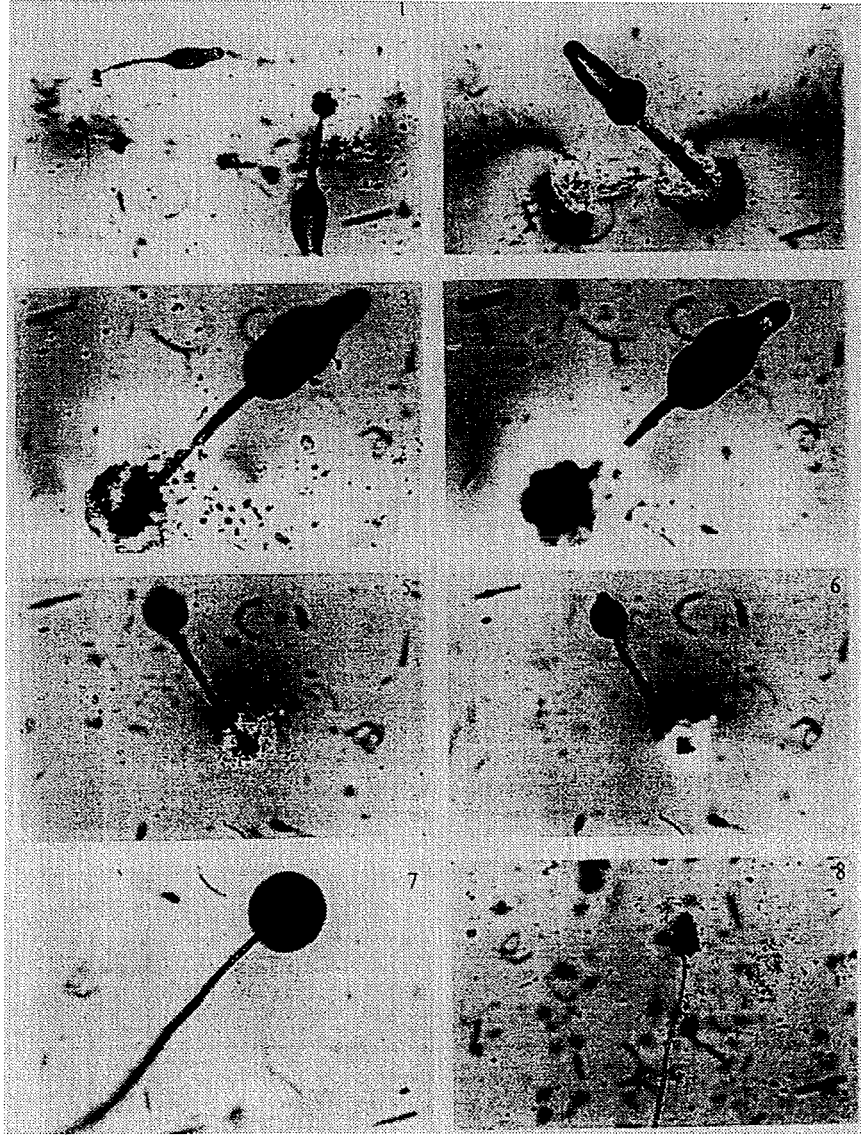


FIG. 6 – Annexe B: cinétique de différenciation et morphogenèse de *Dictyostelium Discoideum*, de 18 h à 24 h ,photos reproduites avec la permission de F. Lenormand, numérotation de gauche à droite, de bas en haut, grossissement $\times 10$.; (1,2) 18h: Développement des sorocarpes; (3,4,5,6) 20h: croissance de la tige; (7) 22 h: Arrêt de la croissance; (8) 24h: Libération des spores

Aller et retour de l'identité logique à l'identité empathique. Esquisse d'un cheminement possible

Guillaume Deffuant

Laboratoire d'Ingénierie des Systèmes Complexes
Cemagref
Parc de Tourvoie
92 185 Antony Cedex

Ce texte est un résumé étendu dont la plupart des points nécessiteraient de plus amples développements. Le mouvement général de l'argumentation est le suivant : partant de l'identité au sens mathématique, nous examinons tout d'abord son extension à l'identité par stabilité des formes proposée par la morphodynamique. Nous nous interrogeons sur l'applicabilité de cette extension concernant l'identité sociale des personnes. Nous proposons quelques directions nouvelles sur ce dernier point, en nous fondant sur des hypothèses fortes à propos du fonctionnement cognitif. Ces hypothèses nous donnent également une nouvelle perspective sur l'émergence de l'identité mathématique à partir des mécanismes de l'identité sociale.

1. Introduction

Nous nous intéressons particulièrement à deux aspects de l'identité intervenant dans les débats des sciences de la cognition. Tout d'abord, l'identité est une notion centrale dans l'activité de catégorisation, qui est la capacité de distinguer le même à l'intérieur de la diversité. Cette capacité, qui semble tout à fait fondamentale dans la cognition, fait l'objet de nombreux débats, qui sont loin d'être tous tranchés : quels sont les mécanismes permettant de catégoriser ? existe-t-il des catégories innées ? comment se forment et évoluent les catégories ?

Toutes ces questions sont évidemment très importantes et font l'objet de nombreuses recherches. Cependant, elles sont loin d'épuiser tous les aspects que recouvre la notion d'identité. En effet, l'identité d'un individu, perçue par l'individu lui-même ou par les autres, est également au cœur de nombreux débats au sein des sciences de la cognition et de la philosophie de l'esprit. Ces débats se trouvent en résonance avec les questionnements sur la conscience, la conscience de soi, les attitudes propositionnelles, etc..

Nous proposons de revenir sur quelques points forts de ces débats, en tentant de montrer les liens étroits entre ces deux acceptions de l'identité. Le texte est organisé comme suit : tout d'abord nous revenons au sens logico-mathématique de l'identité, ensuite nous décrivons rapidement l'extension de ce sens donnée par l'approche morphodynamique de Petitot (Petitot 92), puis nous rappelons quelques éléments du débat de la philosophie de l'esprit concernant l'identité des personnes, enfin, nous proposons des hypothèses sur le fonctionnement cognitif, inspirées des travaux d'Edelman (Edelman 92), et Dupuy (Dupuy 92), nous permettant de donner une vision différente de cette identité.

2. L'identité de la logique ensembliste identitaire ou l'identité logique

Considérons une première manière de concevoir l'identité, qui peut apparaître dans un premier temps comme la plus simple : il s'agit de l'identité au sens mathématique, au sens de la logique ensembliste identitaire de Castoriadis. Cette logique est fondée par la définition des ensembles dues à Cantor que commente assez longuement Castoriadis. Dans ce cadre, les objets du monde sont définis et sont soit égaux, soit distincts.

Aucune autre solution n'est possible. Il s'agit de la base des mathématiques. Remettre en cause cette notion d'identité serait mettre en péril tout l'édifice.

Cette approche de l'identité convient parfaitement à des objets mathématiques abstraits, tels que les nombres par exemple. Elle convient parfaitement également à l'informatique où l'identité des objets est donnée de manière univoque par leur adressage ou leur valeur. Elle convient plus généralement aux objets syntaxiques des langages formels.

Cette première facette de l'identité a une importance très grande dans les discours scientifiques notamment. En effet, il s'agit probablement de la notion d'identité ultime, que des discours réductionnistes considéreront comme la seule valide et légitime.

Or, nous y reviendrons, il apparaît que cette conception de l'identité n'a pas de sens hors du monde abstrait des mathématiques ou de la logique, et que son utilisation abusive dans des mondes où elle ne s'applique pas est très lourde de conséquences. Nous allons tenter de le montrer, et de proposer quelques alternatives.

3. Identité continuité des objets

Les propos de Jean Petitot sur le problème de la perception des objets et scènes de la vie courante est intimement lié à cette question de la compréhension de l'identité. En effet, Petitot constate le manque de fondation scientifique à notre expérience de tous les jours, notre perception des objets par exemple. Selon lui, les fondations scientifiques de notre époque, se trouvent dans la physique et dans la syntaxe formelle. Ailleurs, il n'est pas de fondement scientifique réellement solide. On peut remarquer que ces deux domaines sont ceux pour lesquels l'identité de type « ensembliste identitaire » est applicable dans une large mesure (encore que cela puisse être discuté dans le cas de la physique).

Nous considérons donc que le travail de Petitot peut être présenté comme une tentative de fonder mathématiquement un élargissement de la notion d'identité, qui permettrait de donner un fondement objectif à nos perceptions des objets usuels. On voit l'ambition extrême de ce projet.

Les moyens mis en œuvre sont à la mesure de cette ambition. En effet, Petitot appelle à la rescousse des mathématiques élaborées, fondées sur la théorie des catastrophes de R. Thom. L'idée centrale, empruntée à Husserl, est que l'identité est liée à la continuité : l'identité d'un objet ne change pas si son aspect se transforme continûment. Nous proposons d'appeler identité continuité cette approche de l'identité, puisqu'il elle se caractérise par une continuité de l'aspect de l'objet sous un ensemble de transformations.

Nous voici donc passés de la notion d'identité à celle de continuité, qui est elle-même fondée mathématiquement la notion de distance, elle même impossible à caractériser sans l'identité logique. Il s'agit donc d'une extension de l'identité logique.

Il faut souligner une grande parenté entre cette approche et le souci d'autres auteurs de caractériser la perception par un couplage sensori-moteur. En effet, l'objet se caractérise par ses aspects et les déformations associées, donc sur un ensemble de manipulations possibles de l'objet qui le change continûment. Cette vision de l'identité me paraît d'une très grande portée : l'identité se caractérise comme la zone de stabilité (i.e. de déformations continues) par rapport à un ensemble de degrés de liberté et de déformations possibles de l'objet considéré. Nous aurons l'occasion de retrouver cette vision par la suite.

4. Limites de la naturalisation de l'identité continuité

Une caractéristique très importante du travail de Petitot est la tentative de doter la théorie de la perception, et donc sa vision de l'identité des objets, d'une objectivité mathématique. Cette objectivité passe par une naturalisation de l'aspect d'un objet, c'est à dire de sa forme. Cette forme sera caractérisée par un ensemble de singularités objectives (à un certain niveau d'échelle). Les changements d'aspects correspondent à des changements dans le nombre ou les caractéristiques de ces singularités.

Il nous semble intéressant de discuter cette tentative. En effet, elle se fonde implicitement sur la distinction due à Thom entre saillance et prégnance. La saillance correspond aux caractéristiques mathématiques de la distribution des singularités de la forme, alors que la prégnance indique la signification associée à cette forme par l'organisme qui la perçoit.

Cette distinction est très lourde de sens, ainsi que d'implications philosophiques. En effet, elle est à rapprocher de la distinction moderne entre fait et valeur, que les anciens ne connaissaient pas. Jusqu'à l'époque moderne, il n'était pas venu à l'idée de distinguer un fait de sa signification. Cette dichotomie est donc très caractéristique de la modernité, et a été soulignée comme telle par H. Arendt (Arendt 61), et indissociable de notre débat sur l'identité. En effet, la naturalisation de la forme proposée par Petitot implique que l'on exclue le sens de la caractérisation de l'identité, car le sens, ou la prégnance sont attribués à l'objet par l'observateur au cours de son interaction avec l'objet, et se trouve donc incompatible avec la naturalisation.¹

On peut se demander jusqu'où cette exclusion du sens est praticable, et par là même, quelles sont les limites de la naturalisation proposée par Petitot.

D'une part, même en admettant l'hypothèse fondamentale que l'identité perçue des objets est fondée sur des singularités intrinsèques, les discontinuités de la prégnance peuvent se fixer sur des singularités mineures. La prégnance pourra donc changer radicalement sous un changement mineur de la disposition des singularités. On peut le constater dans des phénomènes de mode par exemple, où des différences subtiles dans la coupe d'un vêtement, l'ajout d'une marque, peuvent acquérir des significations subjectives énormes, tout en étant objectivement très proches.

D'autre part, même dans l'expérience courante de la perception, la naturalisation de la forme et de la stabilité sont loin d'être évidentes. Ainsi les expériences du phénomène phi par exemple mettent en évidence une perception de l'unité d'un objet liée à des constantes de temps internes: Ce n'est qu'en dessous d'un certain pas de temps auquel les lampes s'allument les unes après les autres que la perception d'un objet en mouvement s'établit. La proximité temporelle permet la reconstruction d'un mouvement. Ces faits plaident pour l'existence d'une constante interne permettant de décider de la continuité d'un mouvement. On s'aperçoit donc que la continuité n'est pas un phénomène intrinsèque, mais se trouve dans une interaction entre le sujet et l'objet, ce qui semble incompatible avec la tentative de naturalisation.

Au total, il nous semble que le modèle morphodynamique général garde sa pertinence pour nous aider à définir une extension de l'identité logique vers une identité continuité. Cependant, son utilisation pour une naturalisation de cette identité nous paraît plus que discutable, notamment en ce qui concerne la naturalisation des singularités. Nous préférons donc utiliser le modèle morphodynamique général, prenant en compte l'interaction entre le sujet et l'objet. Les discontinuités qualitatives sont définies par la dynamique particulière du sujet. Cette conclusion est un constat d'échec ou de modestie : il n'y a pas d'approche cohérente de l'identité sans des hypothèses minimales sur le sujet déterminant les zones de stabilité.

On mesure aisément les difficultés soulevées par cette conclusion qui suppose une théorie du fonctionnement cognitif pour aborder une théorie de l'identité, alors que l'identité est évidemment au cœur du fonctionnement cognitif... Cette enchevêtrement apparaît encore plus lorsqu'on aborde l'identité des personnes.

5. L'identité continuité en philosophie de l'esprit

L'identité des personnes a demandé de manière beaucoup plus pressante une extension de l'identité logique, qui est apparue immédiatement inadéquate. Locke par exemple propose de fonder l'identité sur la continuité des souvenirs. De manière générale, la question de la continuité apparaît naturellement pour la caractérisation de l'identité des personnes, alors que pour l'identité des objets, la tentative de Petitot apparaît comme relativement isolée.

Un des débats importants signalés par P. Engel (Engel 94) est celui entre une thèse matérialiste forte : la préservation de l'identité implique une continuité physique (les expériences de pensée effectuées sont des transplantations de cerveaux par exemple), ou les partisans d'une continuité psychologique seule. Là encore les raisonnements utilisés font appel à des expériences de pensée dans lesquelles les désintégrations et les réintégrations sur des planètes imaginaires vont bon train.

La plupart des arguments utilisés rappellent l'approche morphodynamique générale : il s'agit de soumettre l'individu à des expériences de pensée jusqu'à ce que l'on puisse définir des points de discontinuité qualitative, c'est à dire des déformations pour lesquelles l'identité change. P. Engel cite même un auteur qui présente le point commun avec Petitot de prétendre caractériser l'identité de manière objective, c'est à dire caractérisable à

¹ Un rapprochement entre cette distinction et la distinction également moderne entre syntaxe et sémantique serait probablement intéressant.

la troisième personne, ce qui tombe sous des objections encore plus difficiles que celles que l'on peut faire au travail de Petitot.

Engel prend la défense de l'approche générale utilisée, similaire à l'approche morphodynamique, avec des arguments convaincants, le principal étant que nous n'avons pas d'alternative. Simplement, comme nous l'avons remarqué, cette approche pose explicitement le rôle fondamental d'un observateur, qui nous permet de trancher sur la continuité ou pas de l'identité. Nous avons vu qu'il en va de même pour l'identité des objets, et cela nous paraît inévitable.

Dans le cas de la caractérisation de l'identité des personnes cependant, cette approche de l'identité implique des suppositions a priori (implicites en l'occurrence) sur le fonctionnement cognitif, qui font partie de cette subjectivité. Nous pensons que l'une de ces suppositions implicites est que l'identité des personnes est un objet quelconque, qui ne se distingue pas fondamentalement des autres objets, et peut être traité de la même manière. En effet, nous l'avons souligné, l'approche de la philosophie de l'esprit pour caractériser l'identité des personnes est similaire à l'approche morphodynamique utilisée par Petitot pour caractériser l'identité des formes.

Nous pensons qu'il s'agit d'une hypothèse très discutable, et nous aurions au contraire plutôt tendance à défendre une hypothèse différente, dans laquelle la relation aux personnes joue un rôle tout particulier dans la cognition, et doit être comprise totalement différemment des relations aux objets. Cette hypothèse oblige à reconsidérer l'espace des modifications possibles (contrôles) testant la continuité de l'identité.

6. Proposition d'une hypothèse schématique sur le fonctionnement cognitif

Nous avons soutenu que des hypothèse sous jacente sur le fonctionnement cognitif était toujours implicite dans les discours sur l'identité, même sur l'identité des objets, qui ne peuvent se départir d'un aspect subjectif. Pour faire des propositions, il nous est donc nécessaire de poser de telles hypothèses. Nous décrivons une proposition de fonctionnement cognitif (très schématisé), dont nous pensons qu'il pourra nous donner des indications intéressantes sur l'identité d'une personne.

Nous supposons très schématiquement que le fonctionnement cognitif humain se répartit en deux grandes parties (qui peuvent être vues comme une généralisation des propositions d'Edelman (Edelman 92)) :

- le système cognitif primaire, fondé sur l'histoire des interactions sensori-motrices et une grande part de déterminations biologiques initiales. Nous ferons l'hypothèse que l'élément de base du fonctionnement de ce système est l'anticipation en fonction de situations et d'actions. Aux motivations biologiques du moment, le système va choisir les actions qui lui paraissent de nature à satisfaire au mieux ces motivations biologiques, et les catégorisations induites sont liées à ces renforcements biologiques. On peut postuler, comme Edelman, l'existence d'une conscience primaire développée dans la tension du désir (ou de la répulsion) stimulé par des souvenirs, engendrant une impression de présent remémoré. Ce système cognitif primaire affecte les discontinuités perceptives en fonction de ses caractéristiques biologique et de l'histoire de son couplage avec l'environnement.
- le système cognitif d'ordre supérieur, ou socio-linguistique, fondé sur les interactions sociales empathiques ou mimétiques. Le développement de ce système s'explique par l'importance acquise par les interactions sociales qui définissent des enjeux très importants. Les anticipations des résultats de ces interactions sociales se font par empathie, ou par mimétisme : chacun doit se mettre à la place de l'autre afin d'anticiper ses réactions ou ses sentiments. Ici nous empruntons ces concepts à une longue tradition incluant Adam Smith, J.P. Dupuy, R. Girard...

On le voit, ces hypothèses sont très fortes, et nous ne pouvons donner ici l'argumentation qui s'imposerait pour les justifier. Nous proposons simplement de les étayer en observant les conséquences de telles hypothèses dans la compréhension de l'identité des personnes.

7. Dissolution l'identité personnelle par l'opérateur d'empathie

Une première conséquence de notre hypothèse de travail est que l'on peut considérer l'identité des personnes comme au moins double : elle comprend une composante issue du système cognitif primaire, et une seconde

issue du système cognitif d'ordre supérieur. Les deux sont probablement en interaction forte, mais présentent des aspects distincts également. Nous allons tenter de le mettre évidence.

Commençons par nous intéresser à la composante dite « primaire » de l'identité. Elle sera caractérisée par un noyau stable d'anticipations de désirs et de craintes, liées à des vécus marquants, ou à des déterminismes biologiques. On retrouve ici l'ensemble complexe de souvenirs, de désirs et de craintes, que la philosophie de l'esprit avait identifié comme le noyau psychologique caractéristique de l'identité. Les questions posées sur les opérations menant à des transformations permettant d'atteindre les limites de ce noyau restent entières.

Cependant, la question de l'identité sociale nous paraît assez différente. En effet, supposons que l'on en reste au même schéma : l'identité sociale est déterminée par l'ensemble des souvenirs, désirs et craintes issues d'opérations empathiques ou mimétiques. Alors, nous nous trouvons devant une difficulté importante, qui est au cœur des problèmes posés par ces opérateurs mimétiques et empathiques : où se trouve l'identité de la personne puisqu'elle emprunte ses souvenirs, désirs et craintes à la personne à laquelle elle s'identifie ? Le problème est d'autant plus épineux que la personne en question est probablement en train de faire de même, d'où une instabilité extraordinaire. Ainsi, l'identité empathique est-elle par définition en perpétuel changement. Le caractère essentiellement discontinu des désirs et admirations sociales est souligné par exemple par R. Girard, qui insiste sur l'instabilité des foules par rapport à leurs idoles, passant de l'adulation à la haine la plus radicale en un instant.

Utiliser la même approche pour l'identité sociale que pour l'identité primaire semble donc plutôt périlleux. En effet, les désirs et craintes, sont en grande partie dépendantes de l'empathie du moment (pour un interlocuteur, ou un personnage imaginaire, de roman, de film, de la mythologie...). A première vue, l'empathie dissout notre identité en la distribuant sur les personnages auxquels nous sommes confrontés et dont nous épousons en partie l'identité.

8. Construction d'une identité sociale fondée sur l'opérateur d'empathie

L'identité sociale est donc probablement à rechercher un peu plus loin. Nous proposons de la caractériser par *la manière* dont les mécanismes empathiques sont mis en œuvre. Il s'agit donc de percevoir de grands noyaux de stabilité dans la pratique courante que l'individu fait de ses capacités empathiques. Une des questions est par exemple la pratique de l'opérateur empathique au carré. L'individu se sert-il de l'autre comme d'un miroir (pour ressentir l'empathie de l'autre pour lui-même) ?

Une telle question nous permet de différencier l'identité de l'homme moderne, occidental (individualiste au sens de Dumont) de l'homme traditionnel. En effet, si l'on considère les descriptions de Mircea Eliade par exemple sur les schémas psychologiques des hommes traditionnels, on constate qu'ils disposent de modèles mythologiques pour toutes les actions importantes de la vie courante (Eliade 69). L'opérateur d'empathie n'est utilisé ici qu'une fois, sans retour sur eux même. La vie de l'homme traditionnel est donc une suite d'identifications complètes à des personnages mythologiques, dont il épouse les sentiments imaginaires. C'est ainsi qu'il rejette les événements qui n'entrent pas dans de tels schémas, ou bien les y fait entrer de force en les réinterprétant.

L'homme occidental se caractérise par une absence de personnage mythologique auquel s'identifier, et c'est sans doute là une des raisons profondes du mal de vivre des sociétés modernes (le château de Kafka est vide !). Il a de grandes difficultés à trouver des modèles disponibles dans lesquels il pourrait totalement s'oublier, ainsi que semble le faire l'homme traditionnel. Le schéma d'identification se complique alors, puisque l'identification n'étant pas complète, elle devient une sorte de miroir, dans laquelle on peut se regarder, et se regarder se regarder, et ainsi de suite.²

Dans ce cadre, l'identité due à la composition de l'opérateur empathique présente deux parties : l'une correspondant aux personnages externes (aux autres en général, réels ou imaginaires) et une image de soi que l'on perçoit par empathie de ces autres. C'est ainsi que peut se composer un type particulier d'image sociale de soi : l'image de soi renvoyée par les modèles ou les interlocuteurs, choisis délibérément ou imposés. Le moi

² Un romancier américain contemporain caricature la société américaine en décrivant des comportements d'identification totale à des schémas publicitaires ou à des personnages de feuilletons télévisés. Ces cas extrêmes sont dans notre perspective un curieux retour vers des identifications archaïques au sein même de la société moderne.

ainsi créé est donc curieusement vide, puisqu'il n'est que le regard des personnages en surplomb. Il correspond pourtant très probablement à l'essentiel de l'identité sociale telle qu'elle est perçue de l'intérieur.

Les deux composantes (le regard et l'image qu'il renvoie) sont évidemment en interaction avec le système cognitif dit primaire, avec lequel des contradictions sur les désirs ou les aversions peuvent s'établir. Ceci donne une identité dépendant des interactions entre ces trois composantes entre lesquelles des conflits peuvent être fréquents et violents. Les similitudes évidentes de ce système avec la psychanalyse freudienne ne sont sans doute pas fortuites.

9. Esquisse d'un chemin de retour

Notre parcours nous a conduit, par des extensions successives, de l'identité logico-mathématique à l'identité empathique. Ce chemin est celui du simple vers le complexe, dans le sens de la difficulté à caractériser l'identité. Cependant, selon notre postulat sur le fonctionnement cognitif, l'identité empathique est première, puisqu'elle est au centre des capacités cognitives d'ordre supérieur. L'identité logico-mathématique est donc issue d'un développement historique qui prend comme point de départ l'opérateur empathique. Nous tentons maintenant de donner des indications sur les mécanismes possibles de ce processus menant de l'identité empathique à l'identité logico-mathématique.

L'identité logico-mathématique est définie initialement par et pour une communauté scientifique, et plus largement, de manière universelle à tout homme capable de reproduire les procédures permettant de caractériser cette identité. On peut donc considérer que le scientifique se soumet par empathie au regard du scientifique universel qui critique son travail. En cela, il anticipe l'épreuve qu'il aura à subir pour convaincre la communauté, et s'identifie à ses contradicteurs potentiels. Ces contradicteurs sont censés respecter une règle du jeu, qui est d'utiliser des arguments acceptables par le contradicteur scientifique universel. La grande force de l'identité logico-mathématique repose donc sur son universalité, universalité qui peut s'interpréter comme empathie pour un scientifique quelconque ou universel, point de référence ultime. L'identité mathématique n'a de sens que parce que le mathématicien la considère acceptable pour le scientifique universel auquel il la soumet par la pensée.

Allons plus loin, et examinons les caractéristiques de ce scientifique universel, central selon nous à l'activité scientifique. Le point intéressant est que ce scientifique idéal se caractérise par l'absence de toute empathie qui influencerait son jugement et le détournant des procédures qu'il doit tester. Dans sa pratique scientifique, l'homme s'identifie à un être capable (en théorie) de refuser toute identification, et de n'utiliser que sa raison. Il s'agit finalement d'un ensemble de procédures testables et réfutables. C'est ici d'ailleurs que se noue la grande différence avec les principes divins : les dieux se posent comme références empathiques d'une communauté, mais ils sont capables eux aussi d'empathie (c'est peut-être cette différence qui a fait avorter les tentatives d'établissement d'un culte de la raison). Au contraire, le scientifique s'identifie (idéalement) à une procédure incapable d'empathie.

Le chemin proposé ici de l'identité empathique vers l'identité logico-mathématique passe donc par une utilisation spéciale de l'opérateur d'empathie, lui permettant en quelque sorte de se nier lui-même, ou de s'auto-détruire, puisqu'il s'agit d'une empathie pour un personnage irréel non empathique. Il est clair que si notre hypothèse sur le fonctionnement cognitif est correcte, cette négation est aussi la négation d'une part fondamentale de la cognition qui se trouve refoulée hors du champ de vision scientifique. Les conséquences de cette négation sur l'histoire de la modernité seraient probablement très intéressantes à étudier³.

10. Conclusion

Nous proposons dans ce papier une progression dans l'extension du sens de l'identité à partir de son sens logique mathématique. Une première extension, correspondant à l'approche morphodynamique, propose de caractériser l'identité des formes des objets à partir de la théorie des singularités. Elle repose sur la continuité de

³ Le Marxisme notamment, présente beaucoup de liens avec cette vision universaliste de l'homme, comme le souligne par exemple F. Furet. Or si notre raisonnement est juste, et si cette universalité reproduit le mouvement de l'universalité par identification à une référence niant l'empathie (ce qui reste à montrer), alors elle porte en elle une négation de l'homme et donc les germes du totalitarisme, presque autant que le fascisme qui effectue directement cette négation.

l'aspect de l'objet soumis à un ensemble de transformations. Elle donne un cadre très général dans lequel l'identification sensori-motrice des objets peut aisément prendre place.

Nous montrons ensuite que l'identité sociale ou l'identité des personnes est encore plus éloignée de l'identité logico-mathématique. En effet, la capacité d'empathie permet aux sujets de se mettre les uns à la place des autres, et de changer mentalement d'être constamment. Ceci oblige à aller chercher l'identité à un niveau plus élevé, dans la manière d'utiliser l'opérateur empathique par exemple. Nous identifions notamment les figures de l'empathie au carré présentant des rapprochements avec la psychanalyse qui, nous semble-t-il, mériteraient une plus grande attention.

Enfin, nous donnons quelques indications sur la constitution de l'identité logico-mathématique à partir des opérateurs d'empathie. Cette constitution passe par l'identification à un être universel réfractaire à toute empathie, et donc permet de fixer l'identité par rapport à cette référence ultime.

En conclusion, nous souhaitons insister sur l'importance de la distinction nécessaire entre les différents types d'identité. Utiliser une conception de type logico-mathématique pour caractériser l'identité des personnes est extrêmement dangereux car cela revient à nier la dimension essentiellement humaine qu'est la capacité d'empathie. On peut retrouver cette négation dans différentes idéologies, considérant l'identité de l'homme uniquement sous un angle biologique ou économique, et qui nient l'identité sociale au sens où nous l'avons décrite. Cette négation forcenée n'est pas tenable longtemps et des principes d'identification archaïques (au chef tout puissant par exemple) ne tardent pas à resurgir, entraînant des tensions énormes, lourdes de violences extrêmes, dont nous avons fait la terrible expérience au cours de ce siècle (L. Dumont propose un début d'argumentation pour une idée assez similaire dans (Dumont 83))

11. Bibliographie

Arendt H. *La condition de l'homme moderne*. Calman Lévy. 1961.

Castoriadis C. *Domaines de l'homme*. Le seuil, 1986.

Deffuant G. « Mécanismes empathiques et complexité sociale : autour de J.P. Dupuy ». *Actes des journées de Rochebrune* 1996.

Dumont L. *Essais sur l'individualisme*. Le Seuil, 1983.

Dupuy J.P. *Introduction aux sciences sociales*. Ellipses. 1992.

Edelman G. *Biologie de la conscience*. Odile Jacob. 1992.

Eliade M. *Le mythe de l'éternel retour*. Gallimard. 1969.

Engel P. *Introduction à la philosophie de l'esprit*. Editions la découverte. 1994.

Girard. R. *Mensonge romantique et vérité romanesque*. Grasset. 1981.

Petitot J. *Physique du sens*. Editions du CNRS. 1992

La voiture de mon beau-frère est-elle toujours la voiture de mon beau-frère?

Philippe BENHAMOU
COGNILOG
12, rue Georges Méliès
93100 MONTREUIL SOUS BOIS
email : benhamou@onera.fr

- Robert** Tu vois, ma Renault 21, et bien avec mon voisin mécano, on a tout changé, il n'y a plus une seule pièce d'origine, on a tout changé, le moteur, le carbu, le circuit d'eau, tiens la semaine dernière, il m'a encore changé l'embrayage.
- Philippe** Alors t'as une voiture neuve ?
- Robert** Ben oui puisque tout est neuf ou presque.
- Philippe** Donc t'as changé de voiture ?
- Robert** Ben non, c'est la *même*!
- Philippe** C'est pas la *même* !
- Robert** Ben si, si j'avais changé de voiture, j'aurais une nouvelle carte grise, petit malin! Et puis j'aurai deux voitures alors que j'en ai qu'une.
- Philippe** Moi, j'en n'ai pas.
- Robert** Si j'avais changé de voiture, je t'aurais donné ma vieille R21
- Philippe** Ou tu me l'aurais prêtée?
- Robert** Ben, oui
- Philippe** Dis, Robert...
- Robert** Quoi ?
- Philippe** Tu ne me prêteras pas ton ancienne/nouvelle voiture en février pour aller à Rochebrune ?
- Robert** T'as ton permis ? Je croyais qu'on t'avait retiré tous tes points pour excès de vitesse ?
- Philippe** Oui et non, je les ai regagnés dernièrement en prenant des heures de conduite
- Robert** Alors, non, je ne prête pas ma voiture à quelqu'un qui vient juste d'avoir son permis

1. Introduction

Aborder les problèmes soulevés par le concept d'identité, c'est se retrouver constamment face à un dilemme : soit le discours est contradictoire, paradoxal jusqu'à l'absurde, soit le discours reste élémentaire, trivial, jusqu'à devenir tautologique. En effet, le concept d'identité se définit grâce à lui-même, dans une auto-référence permanente qui entraîne tout chercheur dans sa spirale de contradiction et de tautologie. Dans ce papier, nous tenterons de distinguer les diverses formes du concept d'identité, à savoir l'identité entre deux objets et l'identité d'un même objet au cours du temps. Nous donnerons quelques éléments de réflexion tirés de mises en relation de domaines variés comme l'intelligence artificielle (IA), la systémique, la classification et la philosophie. Enfin nous proposerons une mesure de similarité entre objets d'un domaine pouvant conduire à une représentation graphique des relations de ressemblance

2. Le concept d'identité

On peut très schématiquement organiser cette problématique autour de deux problèmes [Ferret 96] :

- < **le problème de l'identité épistémique** : à savoir rechercher quels sont les critères, indices ou symptômes que nous construisons pour identifier les objets : les relations qu'une entité entretient avec elle-même et son environnement à un instant précis.
- < **le problème de l'identité dans le temps** : c'est-à-dire, la persistance de l'identité des objets à travers le temps : Les relations qu'une entité entretient avec elle-même à deux moments distincts de sa carrière.

2.1 L'identité en intelligence artificielle

Le problème de l'identité est très présent en IA bien qu'il soit peu abordé par les chercheurs eux même. Il présente deux facettes : soit on l'aborde du point de vue de l'identité des artefacts en général et à ce titre il n'y a pas lieu de distinguer les artefacts dit artificiellement intelligents d'autres artefacts; soit on l'aborde comme le problème de savoir si un système de traitement de l'information est un système intelligent. Dans ce dernier cas, nous nous trouvons face au problème de la définition même de l'intelligence et l'on sait l'importance de la notion d'identité dans le test de Turing puisqu'il s'agit de trouver qui est qui dans un dialogue où l'un des participants est un être humain et l'autre une machine.

Mais la notion d'identité a été projeté sur le devant de la scène par le passage de l'IA à l'Intelligence Artificielle Distribuée (IAD) impulsé par [Minsky 88]. En effet, cette rupture dans la démarche d'analyse, de conception et de développement de systèmes intelligents, en proposant le passage du *un* au *multiple* dans la vision même d'un problème et de sa résolution, a entraîné une réflexion sur l'identité du *un* par rapport au *multiple* et du *un* par rapport à lui-même au cours du temps.

Dans la pratique de la conception de systèmes informatiques distribués, l'identité spécifique se retrouve dans les choix conceptuels à effectuer lors du passage de la description globale d'un problème à résoudre à la description des entités (appelés *agents* dans la suite du papier) dont les interactions vont conduire à la résolution du problème. D'un point de vue méthodologique, le problème est le suivant : comment, à partir d'une spécification globale d'une tâche, spécifier les comportements individuels dont la co-action va permettre la résolution du problème global. La réponse se situe généralement dans la prise en compte d'organisations dynamiques (voir les aspects méthodologiques dans [Collinot & al. 96]). Il s'agit donc essentiellement d'un passage du global au local, ce découpage conduisant à des entités hétérogènes. La transition réciproque, le passage du local au global est assurée par les interactions (communication, coopération, processus homogénéisants).

Un collectif d'agents se construit avec des agents suffisamment proches pour pouvoir travailler ensemble et suffisamment différents pour passer de la concurrence (au sens des machines parallèles) à la coopération (le tout est supérieur à la somme de ses parties).

Un collectif se construit grâce à la coexistence de deux principes antagonistes:

- < Un collectif ne saurait exister sans un **principe intégrateur** [Barat & al. 92]: sémantique commune, zone de données communes et partagées pour les organisations à base de tableau noir; langage d'interaction commun pour les organisations à base d'envoi de messages (le succès de KQML [Finin & al. 93])

illustre cette nécessité) ; objectif commun pour les systèmes de résolution coopérative de problèmes...

- < Un collectif ne saurait exister sans un **principe de différenciation** au sein de ces éléments : La sémantique commune n'a d'intérêt que si elle permet des échanges d'informations hétérogènes, les Sources de Connaissances des systèmes à base de tableau noir sont hétérogènes. Elles traduisent des expertises différentes nécessaires à la résolution du problème.

Bref, on le voit, de façon consciente ou inconsciente, le concepteur de système distribué évolue en permanence dans ce double lien : il conçoit des entités à la fois semblables et différentes, proches et éloignées. Il est à la recherche d'un compromis permanent entre le semblable et le différent.

2.2 L'identité dans d'autres disciplines

Dans ce chapitre, nous présentons quelques aspects du concept d'identité que j'ai pu rencontrer au cours de mes lectures.

En systémique, le problème de l'identité se traduit par la mise en évidence d'un état d'équilibre entre l'identique et le semblable. Les sous-systèmes considérés devant être suffisamment proches pour former un système et suffisamment distants pour ne pas fusionner. En thérapie familiale systémique, l'identité de l'individu au sein de son système familial est vue comme un état d'équilibre entre l'identification totale au modèle proposé par le système familial (conformité) et le rejet total de ce modèle (marginalité) [Caillé 90].

Dès lors qu'il s'agit de l'identité d'êtres doués d'autonomie, il semble que l'identité soit comme un état de tension entre des opposés, des contradictions. On retrouve d'ailleurs à la base des travaux de Stéphane Lupasco ([Lupasco 87b] en particulier) la recherche d'une formalisation de cet état de tension entre les contraires. Cet état, que Lupasco a appelé *état T*, est cet état d'équilibre dynamique où tout basculement est possible, où potentiellement tout est présent, où tout peut s'actualiser. Ce sont les travaux de ce philosophe qui ont inspiré les recherches sur l'identité qui sont présentées dans ce papier (voir également [Benhamou & al. 94]).

A propos de l'identité des êtres vivants, citons également, le premier chapitre de [Monod 70] dans lequel l'auteur recherche des critères de distinction objectifs entre objets naturels et objets artificiels. Il recense trois propriétés : la téléonomie, la morphogenèse autonome et l'invariance reproductive. Il souligne enfin l'importance et la difficulté de trouver des critères objectifs pour définir les êtres vivants. Ainsi la première propriété : la téléonomie, c.-à-d. le fait que l'être vivant est un système qui réalise et poursuit un projet, n'est pas un critère objectif.

A propos du problème de l'identité à travers le temps, le lecteur trouvera beaucoup de pistes et un état de l'art complet dans [Ferret 96]. Le fil conducteur de son propos est le paradoxe du bateau de Thésé, ce bateau perpétuellement réparé dont les sophistes d'Athènes se demandaient, au fur et à mesure que les pièces en étaient modifiées ou remplacées, s'il s'agissait encore du même bateau.

Bien sûr tout ceci peut être vu comme un jeu de l'esprit. A quoi bon, en effet s'intéresser à l'identité du bateau de Thésé ou de la voiture de mon beau-frère ? Et pourtant, n'est-il pas troublant de savoir que notre identité demeure alors même que nos cellules sont constamment remplacées. Il en est de même avec certains groupes humains. L'esprit, le travail, l'identité d'un monastère demeure alors même que les

moins qui le constituent changent ; qu'ils disparaissent assassinés par la barbarie ne change rien à l'esprit qui demeure au sein de cette communauté, au contraire.

3. L'identité spécifique

Nous avons été confronté au problème de l'identité à propos d'un système d'aide à la décision pour les contrôleurs aériens s'occupant de planification de recherches d'avion ou de bateaux disparus (projet COMODD [Benhamou & al. 95]). Le problème est le suivant : à un instant donné, plusieurs cas de recherche sont en cours de résolution. Lorsqu'une information arrive dans un centre de recherche et sauvetage, il se peut que le contrôleur qui reçoit cette information n'arrive pas à la corréler avec un cas existant. Dans ce cas et devant l'urgence du traitement, il crée un nouveau cas et commence son traitement. Au fur et à mesure des traitements et de l'évolution des connaissances sur les différents cas en cours de traitement, deux cas peuvent se révéler être identiques et leur description respective devront être fusionnée. Les cas sont décrits à l'aide de fiches de renseignement (localisation, degré d'urgence, facteur temps, nom de l'appareil, etc.) qui évoluent au cours de la résolution du cas. Nous avons donc développé une distance pour mesurer l'identité des cas entre eux. A partir de cette distance, nous construisons une représentation graphique¹ qui évolue dans le temps en fonction des connaissances acquises sur les différents cas. Si la représentation graphique d'un premier cas se rapproche de la représentation graphique d'un second, le contrôleur aura intérêt à vérifier s'il ne s'agit pas du même cas.

3.1 Une mesure pour l'identité

Nous utiliserons la démarche adoptée dans les techniques de "clustering" (une description complète du clustering est donnée dans [Diday 76]) utilisées en reconnaissance des formes ou en classification. L'idée est de définir des classes d'objets à partir d'une mesure d'identité ou de dissemblance.

Soit $E = \{X_j\}_{j=1..m}$ l'ensemble de m objets du domaine et X_i l'objet i défini par n paramètres ou indices $(x_{ij})_{\text{pour } i=1..n}$

Les indices x_{ij} peuvent être des valeurs numériques interprétables dans un univers U , des propriétés (couleurs, forme, espace) ou des valeurs d'attributs.

De nombreuses mesures existent parmi lesquels citons, pour les paramètres numériques, la métrique de Chebychev : $d(X_i, X_q) = \max_j |x_{ij} - x_{qj}|$ ou la métrique de Camberra $d(X_i, X_q) = \sum_{j=1..n} |x_{ij} - x_{qj}| / |x_{ij} + x_{qj}|$

Pour les paramètres symboliques que nous utilisons en IA, nous proposons la métrique suivante à partir d'une description ensembliste des objets X_i :

$$\delta : E \times E \rightarrow [0,1] \text{ et } \delta(X_i, X_q) = \frac{\text{Card}(X_i - X_q) + \text{Card}(X_q - X_i)}{\text{Card}(X_i) + \text{Card}(X_q)}$$

où $X_i - X_q$ est la différence ensembliste de X_i et de X_q

L'interprétation de δ dans U est la suivante :

¹ A partir d'une analyse en composantes principales, on construit un repère orthonormé basé sur les vecteurs propres les plus représentatifs de la matrice des distances. La représentation finale est obtenue en projetant les objets de E dans ce repère.

$\delta(X_i, X_q)$ mesure le degré de dissemblance entre X_i et X_q .

Si $\delta(X,Y) = 0$ alors on dira que X et Y sont identiques,

Si $\delta(X,Y) = 1$ alors on dira que X et Y sont distincts.

δ vérifie les trois propriétés des distances (réflexivité, symétrie et l'inégalité triangulaire).

Mais définir une telle distance ne résout pas complètement le problème de l'identité spécifique. En effet, cette distance est conditionnée par :

1. un processus de choix : choisir les indices qui définissent les éléments à comparer (et on retrouve là le problème même de l'identité spécifique tel qu'il est énoncé dans l'introduction) : les indices traduisent-ils les fonctionnalités, la composition, la description ?
2. un processus de seuil : à partir de quand considère-t-on que la distance est proche de 0 ou de 1 ?
3. une identification des composants : le fait d'utiliser le même symbole permet d'identifier les composants deux à deux et de traduire ainsi une certaine proximité.

Ainsi cette distance présente en elle-même tout le paradoxe du concept d'identité puisqu'elle transpose le problème de l'identité relative de deux entités X et Y au problème de l'identité relative de leur fonctionnalité, de leurs composants ou de leur description. Il est clair que la définition de cette distance, tous les éléments sont mis à plat, sont considérés au même niveau.

Dans la suite et dans le contexte du projet Comodd, nous nous proposons d'affiner la définition de cette distance afin de tenir compte :

1. de la récursivité de la mesure :
 - (ainsi un cas fait référence à un avion qui est lui même décrit par un ensemble de caractéristiques parmi lesquels le pilote lui-même décrit par un ensemble de caractéristiques...
2. des valeurs numériques
 - (une mesure d'identité uniquement basée sur une similitude des objets décrivant les entités à comparer ne peut rendre compte des proximités numériques de certaines valeurs. Les proximités temporelles doivent être également prises en compte.
3. de mesures d'identité sur des structures :
 - (Aux cas de recherche et sauvetage sont associés des zones de recherches. Deux cas dont les zones de recherches sont proches présentent des similitudes ainsi le contrôleur en charge pourra par exemple organiser des recherches sur ces zones avec un seul avion de recherche.

Nous venons de voir comment mesurer le degré d'identité entre deux objets d'un point de vue statique. Nous allons maintenant nous intéresser à la deuxième problématique de l'identité, à savoir la permanence de l'identité dans le temps.

4. Le changement, continuité ou rupture?

Le changement affecte-t-il l'identité ? Entre cet enfant et l'adulte qu'il deviendra, l'identité demeure dans le changement. Entre ce vieillard mourant sur son lit d'hôpital et son cadavre quelques secondes plus tard, l'identité n'est plus la même. Ce cadavre n'est pas cet homme qui il y a quelques secondes était encore porteur de vie. Il y aurait-il alors différents types de changement ?

Humberto Maturana, lors de sa conférence invitée des journées francophones de l'IAD nous a donné sur cet aspect l'exemple suivant : "Son fils, ravi d'avoir une belle panoplie de bricoleur, a entamé un bon morceau de son bureau car il avait besoin d'un bout de bois. Humberto admet que son bureau a été sérieusement abîmé mais il le reconnaît encore comme étant son bureau. Et s'il dit à un tiers : pose ce livre sur mon bureau, c'est sans ambiguïté que cela sera fait. Le lendemain, pour expérimenter sa nouvelle scie, son fils coupe en deux le bureau de son père. Ce dernier se met en colère car son bureau a tout simplement disparu : le bureau de Humberto n'existe plus".

Pour Maturana, un système se caractérise par une **organisation** et une **structure**:

- < **L'organisation** est le système de relations entre les composants du système qui définit la classe de l'identité du système (l'identité spécifique). Dans le cas d'artefacts, cette organisation répond généralement à une ou plusieurs fonctions (un bureau se doit d'avoir un plateau, quatre pieds et au moins un tiroir).
- < **La structure** est constituée par les composants et les relations qui réalisent un système particulier, une instance d'une classe de système (c'est l'identité numérique pour reprendre la dénomination de [Ferret 96]).

Dans une démarche méthodologique de conception de systèmes, l'organisation correspond à la description fonctionnelle du système. La structure est ensuite construite à partir d'éléments, de composants mis en évidence lors de la phase fonctionnelle, et qui s'assemblent pour former une structure.

Cette distinction entre organisation et structure est le fil conducteur de la conception orientée objet: le concepteur définit des classes d'objets. Une classe est une description de l'organisation statique (les champs) et dynamique (les méthodes) d'un ensemble homogène d'objets. Une instance est une réalisation de la classe en terme de structure de données. Deux objets de la même classe sont relativement semblables (même organisation) et relativement différents au moins par les valeurs de leurs attributs (sinon pourquoi créer deux instances?). Les premières phases des méthodes de conception orientées objets ([Yourdon 89] par exemple) sont toujours des étapes de classement : il s'agit d'identifier les objets du domaine et de les réunir en classes en fonction de leur ressemblance structurelle ou comportementale. Les critères qui déterminent ces classes sont, dans ces méthodes, laissés au libre arbitre du concepteur. Le problème de l'identité spécifique se pose alors à lui. Notons que dans la majorité des cas, ce problème n'est pas très complexe à résoudre et que le bon sens suffit !

Les changements au cours du temps, qui n'affectent pas l'identité du système sont des changements où la structure change mais où l'organisation demeure. Dans le premier exemple, la structure du bureau a changé (un morceau de bois a été enlevé) mais l'organisation spatiale demeure (quatre pieds, un plateau, deux tiroirs,...). Dans le deuxième exemple, l'organisation a été modifiée (les quatre pieds ne sont plus solidaires), l'identité du bureau a été détruite.

Il existe donc deux types de changement que nous appellerons les changements de type 1 et les changements de type 2 (terminologie extraite de [Ferret 96]).

4.1 Les changements de type 1

Changement de structure et conservation de l'organisation. Ces changements conservent l'identité numérique.

Dans cette classe de changement, nous trouvons les changements de degré (du fœtus à l'adulte), les changements de lieu : Je suis le même ici ou ailleurs, les changements des

valeurs des propriétés (évolution d'un objet au sein d'une même classe) ou encore les altérations (exemple du bureau d'Humberto).

Ces changements illustrent le principe suivant : **"Pour changer, il faut être le même"**

4.2 Les changements de type 2

Changement de structure entraînant un changement d'organisation. Ces changements ne conservent pas l'identité numérique. On trouve ici les changements de nature (de vieillard à cadavre) et les changements de substance (génération ou corruption).

Ce type de changement illustre le principe : **"Lorsqu'on change, on n'est plus le même"**.

Ces changements concernent un seul système isolé et le problème d'identité se pose alors entre l'identité du système à l'instant T0 (avant le changement) et l'identité du système à l'instant T1 (après le changement). Les systèmes, qu'ils soient artefacts ou naturels, ne sont pas isolés. Ils sont regroupés en classes pour les artefacts (on parle par exemple des briquets, des voitures, des journaux) et de sociétés pour les systèmes vivants (le terme de société est pris ici au sens très large).

Les changements de type 1 sont de l'ordre de la continuité; les changements de type 2 sont de l'ordre de la rupture.

4.3 Problème de l'identité dans le changement

Nous ne nous intéressons ici qu'aux changements de type 1, c'est-à-dire aux changements où l'identité demeure.

Au cours de la résolution collective d'un ou plusieurs problèmes, les collectifs se créent, s'entretiennent et meurent en fonction des besoins. Les agents appartiennent donc à des collectifs qu'ils créent dynamiquement en fonction des configurations (voir un exemple dans [Collinot & al. 96]). Le problème se pose alors de l'identité d'un agent au cours du temps dans la mesure où son identité, nous l'avons vu, se définit par rapport aux autres agents d'un collectif.

4.3.1 Evoluteurs sur les agents

Soit X un agent (c.-à-d. une entité informatique appartenant à un système multi-agents et participant à un collectif afin de résoudre d'un problème).

On appelle **évoluteur** toute action f provoquée ou non par l'agent telle que

$$\delta(X, f(X)) > 0$$

On ne s'intéressera pas ici aux origines ni à la finalité des évoluteurs mais plutôt à leurs effets sur les agents et au sein des collectifs d'agents. A ce titre, on distingue deux classes d'évoluteurs.

4.3.1.1 les évoluteurs homogénéisants

Ces évoluteurs transforment les agents pour les rendre plus semblables les uns aux autres.

Soit X et Y deux agents appartenant au même collectif C . Un évoluteur f est un évoluteur homogénéisant ssi

$$\delta(X,Y) > \delta(f(X), Y)$$

Cette classe d'évoluteurs sera notée $>$ qui symbolise en quelque sorte le passage du mode binaire au mode unaire : de la différence à l'unité.

Nous noterons par $\mathbf{X>}$ l'application d'un évoluteur homogénéisant sur l'agent X .

4.3.1.2 Les évoluteurs hétérogénéisants

Ces évoluteurs transforment les agents pour les rendre plus distincts les uns des autres, pour accentuer leurs différences.

Soit X et Y deux agents appartenant au même collectif C . Un évoluteur f est un évoluteur hétérogénéisant ssi

$$\delta(X,Y) < \delta(f(X), Y)$$

Cette classe d'évoluteurs sera notée $<$ pour symboliser le passage du mode unaire au mode binaire : de l'unité à la différence.

Nous noterons $\mathbf{X<}$ l'application des évoluteurs $<$ sur l'agent X .

4.3.1.3 Définition d'un agent

Au sein des systèmes en équilibre, les évoluteurs sont organisés en couples (homogénéisant/hétérogénéisant). Ils constituent des couples antagonistes régis par les axiomes de la logique définie dans [Lupasco 87a]. Lupasco postule qu'il existe un état du système qui réalise l'équilibre dynamique ces deux antagonismes.

Un agent est une entité réelle ou abstraite capable de maintenir l'équilibre défini par :

$$X = \mathbf{X<X>}$$

Cet équilibre dynamique définit un point stable² appelé Identité de l'agent X .

5. Conclusion

Nous avons, dans cet article, présenté quelques éléments de réflexion pour définir le concept d'identité. Il est clair que ce concept est un concept dynamique. Le changement n'est qu'une facette de cette dynamique et comme le dit le philosophe :

² Je préfère le terme de stable plutôt que fixe. En effet, l'Identité d'un agent est soumise à des fluctuations, évolue au cours du temps mais reste relativement stable. Cette stabilité est justement la garantie de l'existence de l'agent.

Dans le monde de notre expérience, les contradictions souvent se complètent mutuellement et sont nécessaires à l'existence l'un de l'autre. Le changement n'est possible que s'il existe un stade à partir duquel changer, mais à son tour ce stade n'est qu'une étape où reprendre son souffle, une étape dans le passage continu du changement, ou encore une étape où le changement s'interrompt avant de passer à une autre étape dans son mouvement créateur. Derrière cette relation, il y a une dualité de l'état statique éternel et de l'éternel mouvement, et derrière cette dualité il y a quelque chose qui n'est ni état statique ni changement - et il est probable que c'est cela la vraie Réalité. [Aurobindo 59]

Derrière cette dualité, ni statique, ni mouvante, se trouve la voiture de mon beau-frère, qui est et n'est pas en même temps la voiture de mon beau-frère !

6. Remerciements

Je remercie Olivier Poirel et Jean Bourrely et pour leurs conseils et leur aide pour le développement du logiciel de classement.

7. Bibliographie

[Aurobindo 59] Shrî Aurobindo . The Hour of God, 1959

[Barat & al. 92] M.Barat et J. Erceau J. Le Principe Intégrateur pour concevoir des systèmes complexes multi-experts. Conception des systèmes du futur, Strasbourg, nov. 92.

[Benhamou & al. 93] P. Benhamou et J. Erceau Antagonisme et Identité, Second European Congress on System Science, AFCET, Prague, Oct. 93

[Benhamou & al. 94] P. Benhamou et J. Erceau Antagonism and Identity. Revue internationale de systémique WORLD FUTURE, Vol. 42 pp 49-57, 1994 (version anglaise de [Benhamou & al. 93])

[Benhamou & al.95] P. Benhamou, M. Barat et Luc Lamontagne : COMODD, Methods and Tools for Distributed Decision Making, *International Workshop on Decentralized Intelligent and Multi-Agent Systems*, Krakow, Poland, Nov. 95.

[Caillé 90] Ph.Caillé. L'Individu dans le système. Revue Internationale de Systémique, vol 4, n°2,1990.

[Collinot & al. 96] A. Collinot, A.Drogoul et Ph.Benhamou. Agent Oriented Design of a Soccer Robot Team ICMAS'96, déc. 96, Kyoto, Japon.

[Diday & al. 76]E.Diday et J.C.Simon. Clustering Analysis. In Digital Pattern Recognition, K.S.FU Editor. Springer-Verlag, Berlin, 1976

[Erceau & al. 91] J.Ecreau et J.Ferber. L'Intelligence Artificielle Distribuée. La Recherche 233, juin 1991, vol 22, pp 750-758.

[Ferret 96]Stéphane FERRET Le bateau de Thésée : Le problème de l'identité à travers le temps Les Éditions de Minuit, Paris, 1996.

[Finin & al. 93] T.Finin, J.Weber, G.Wiederhold, M.Genesereth, R.Fritzson, D.McKay, J.McGuire, R.Pelavin, S.Shapiro et C.Beck. Specification of the KQML Agent Communication Language plus example agent policies and architecture. DARPA Knowledge Sharing initiative External Interfaces Working Group, June, 93

[Lupasco 87a] S. Lupasco Le principe d'antagonisme et la logique de l'énergie Rocher, Paris 1951, réédité en 1987.

[Lupasco 87b] S. Lupasco L'Energie et la matière vivante Rocher, Paris, 1987.

[Matile & al. 87] L. Matile, P. Tassy et D. Goujet. "Introduction à la systématique zoologique" Biosystème 1. SFS Editions, Paris, 1987

[Minsky 88]. M. Minsky La société de l'esprit InterEditions, Paris 1988.

[Monod 70] J. Monod Le hasard et la nécessité, Essai sur la philosophie naturelle de la biologie moderne. Seuil, 1970.

[Yourdon 89] E. Yourdon. Moderne Structured Analysis. Englewood Cliff, New Jersey, Yourdon Press, 1989

IDENTITE ET PSYCHOPATHOLOGIE

M.Timsit-Berthier et E.Andreewsky
L.P.C.A. Liège et INSERM-TLNP Paris.

Les paradoxes de l'identité

Comment peut-on se construire une identité? Et, en particulier, comment peut-on y concilier la multiplicité et la singularité, la permanence et le changement? (Andreewsky 1994). Le but de ce travail est de tenter d'élargir notre réflexion en présentant quelques exemples de distorsions de l'identité telles qu'on peut les rencontrer en psychopathologie.

Il est habituel de déployer ce concept dans trois dimensions différentes:

-celle de l'identité corporelle, en relation avec un corps autonome, qui fonctionne comme un tout malgré la multiplicité de ses composants et qui se maintient dans le temps malgré la faible fiabilité de chacun de ses constituants. Ce corps "naturel" est soumis aux contraintes engendrées par son organisation structurale et par la temporalité rythmique des lois de la biologie.

-celle de l'identité sociale, "artificielle", en relation avec un rôle social, qui s'élabore dans un espace territorial par rapport à une communauté et à des traditions qui se maintiennent dans le temps à travers des contraintes représentées par des normes et des lois. Cette composante sociale peut s'exprimer à travers le concept de vie "publique" dans laquelle la distance aux autres est modulée en fonction de règles collectives et la temporalité est organisée de façon métrique, en fonction d'une référence stricte, stable, semblable pour tous. C'est le "temps réel".

-celle de l'identité personnelle, qui renvoie à la notion de "sujet" en relation avec la capacité de se représenter par rapport à l'autre et par rapport au monde et de construire une continuité temporelle entre passé, présent et avenir à partir d'"événements" tant corporels que sociaux auxquels on doit donner un sens. Les événements sont constitués comme tels par le sujet puisque c'est lui qui les nome, en fixe le début et la fin sans qu'aucun critère objectif ne permet cette segmentation. Cette composante subjective peut s'exprimer à travers le concept de vie "privée" dans laquelle la distance à soi est modulée de façon variable et la temporalité est organisée en fonction d'une référence floue, essentiellement instable avec des accélérations et des ralentissements qui se font en fonction de "l'histoire intérieure". C'est le "temps vécu".

Ces différentes composantes interfèrent entre elles, ce qui est à l'origine de plusieurs paradoxes, liés à la fois à des difficultés conceptuelles intrinsèques et aux mutations actuelles de notre société.

Ainsi, le rapport au corps pose problème, car d'une certaine manière, si l'homme "est son corps", se confond avec lui et peut se sentir forcé de le supporter comme une charge, il ne s'en contente pas. Il "a" aussi son corps comme un objet auquel il est extérieur qu'il peut explorer pour mieux en connaître les mécanismes et dont il peut disposer comme d'un outil. Ce nouveau savoir, qui, par certains cotés, augmente la distance au corps permet d'en diminuer les contraintes biologiques et d'en retarder la finitude. Ainsi, à l'occasion de déficiences, on peut en changer certaines pièces accessoires (hanche, genou artères) ou essentielles (foie, reins, coeur). On peut aussi en améliorer les capacités physiques, esthétiques, cognitives et sociales grâce à une panoplie de plus en plus diversifiée de produits et de techniques. Et parmi celles-ci, les drogues psychotropes prennent une place toujours grandissante, avec celles qui permettent de restaurer une sociabilité déficiente (les neuroleptiques), et celles qui augmentent l'estime de soi (les antidépresseurs). "Le mythe Prozac ouvre l'horizon identitaire" souligne Ehrenberg (1995) qui renvoie face à face "Sister Morphine face diabolique et Miss Prozac face angélique de la modification permanente de soi".

Ainsi, on assiste à une intrusion de plus en plus grande du social dans le corps biologique dont les contraintes reculent au fur et à mesure que la technique avance. Et le problème se pose alors de savoir jusqu'à quel point l'on peut (ou l'on doit?) mettre son corps à distance pour le maîtriser et le modifier pour parvenir à "être soi-même".

Le rapport soi et à autrui soulève également des questions difficiles car il s'agit d'assumer son rôle social sans pour autant se confondre avec lui. Par certains aspects, le rôle et les actions que la société nous assigne à travers notre sexe, notre âge, notre nom, notre état civil, notre type d'activité, sont indispensables à la construction de notre identité personnelle. Mais par d'autres aspects, cette identité sociale en nous infligeant des limitations et des contraintes peut nous apparaître comme une aliénation. Sartre a bien exprimé ce paradoxe en déclarant que "la tâche consiste à être ce qu'on n'est pas et à ne pas être ce qu'on est". Le problème consiste ainsi pour chacun de nous à mettre la "bonne distance" entre identité sociale et identité personnelle, entre temps réel et temps vécu. Mais, comment moduler cette distance lorsque c'est la société elle-même qui impose à tous l'affirmation de soi comme mode de réussite sociale, (Ehrenberg 1995) et la création de liens relationnels "lointains", médiatiques et électroniques comme mode de sociabilité (Virilio 1996)? Et le problème est bien de se demander comment l'on peut se retrouver "soi-même" dans une telle confusion entre vie publique et vie privée

A cet effet, la pathologie mentale qui nous apporte des exemples de distorsions des temporalités et des distances identitaires peut nous aider à élargir notre réflexion. Nous en retiendrons deux: celui de la Schizophrénie comme impossibilité de construire une identité et celui de la mélancolie comme incapacité à la faire évoluer.

Identité et Schizophrénie:

Il nous semble important de souligner, en préambule, que la Schizophrénie, en tant qu'entité nosologique, n'a été identifiée qu'au XIX^e siècle et que cette affection n'existe pas dans les sociétés traditionnelles. Ainsi, en milieu rural africain, la maladie mentale, perçue comme une menace pour l'intégrité du groupe, est prise en charge par la collectivité et ne présente ni la forme, ni l'évolution que nous lui connaissons en Occident.(Postel et Quételet, 1983).

La schizophrénie se caractérise essentiellement par un trouble de l'identité, dans sa dimension corporelle, personnelle et sociale. La maladie s'annonce souvent par des difficultés à constituer une subjectivité propre et une bonne "affirmation de soi". Et, si le sujet à personnalité "schizoïde" a souvent tendance à tenir les autres à distance et à se montrer rigide, enfermé dans son monde intérieur, c'est pour défendre sa faible autonomie et sa subjectivité déficiente. Dans la phase aiguë, les patients schizophrènes ont l'impression que leur corps est morcelé et fragmenté, qu'il n'y a pas de frontière entre eux et les autres, que l'on peut influencer leurs actes et leurs pensées et qu'ils sont aussi capables d'influencer les autres. Leur langage semble bizarre, ambigu, maniéré dans la mesure où ils omettent les traits contextuels qui le rendrait compréhensible. Lorsque ce vécu a été modifié par des drogues psychotropes qui diminuent le "syndrome d'influence" rétablissent la capacité de dialoguer et augmentent la sociabilité, ils ne parviennent, à long terme, que difficilement et sporadiquement à s'intégrer socialement et à exercer une activité professionnelle.

On peut attribuer la présence de ces troubles de l'identité à l'existence d'anomalies cérébrales structurales, portant en particulier sur les lobes frontaux, anomalies qui existeraient bien avant l'apparition de la maladie et ne se manifesteraient qu'à l'occasion de l'augmentation des demandes sociales, comme c'est le cas au début de l'âge adulte, âge d'apparition de cette affection (Kulynych et Weinberger, 1995). Cette interprétation a l'intérêt de prendre en considération les nouvelles données apportées par les neurosciences et de les articuler avec des données environnementales; mais elle ne permet de comprendre ni l'apparition tardive de la Schizophrénie dans l'histoire des maladies mentales, ni sa grande variabilité d'évolution en fonction des cultures (Sadoun R 1981), ni sa capacité d'être améliorée ou aggravée par le langage (Searles 1965)

Il nous semble plus intéressant d'adapter une position systémique qui pose d'emblée l'interdépendance des processus cérébraux, psychologiques et sociaux (Fargeas et Andreewsky 1991) et de tenter une interprétation à partir de certaines propositions théoriques de H. von Foerster, qui exprime la représentation de soi à l'aide d'une hypothèse "spéculaire". Un individu constituerait ainsi son identité à l'aide de processus autoréférent et récursif (en référence spéculaire croisée avec ses représentations d'autrui et ses représentations des représentations qu'autrui fait de lui). Ce serait précisément cette capacité de "spécularité" qui serait déficiente chez ces patients du fait de "discontinuités temporelles" qui les empêcheraient d'organiser de façon stable à la fois leur relation avec

leur corporéité dont ils ne parviennent pas à délimiter les frontières, leur relation avec les autres dont ils n'arrivent pas comprendre les intentions et à fortiori leurs rôle social.

Il est intéressant de rapprocher ces hypothèses de celles que Grivois (1995) a développé au sujet de la Psychose naissante. Le mimétisme moteur qu'il décrit, c'est à dire l'imitation involontaire du mouvement des autres qui bientôt s'emballe et s'inverse pour aboutir au phénomène de concernement puis de centralité, évoque lui aussi un dysfonctionnement de la temporalité avec incapacité à séquentialiser les causes et les effets et à intégrer des temporalités hétérogènes

Ces différents auteurs ne font pas d'hypothèses sur l'origine de ces distorsions temporelles mais les travaux récents de neurophysiologie et d'imagerie cérébrale ont mis en évidence le rôle des lobes frontaux à la fois dans l'organisation séquentielle des activités et dans la socialisation.(Posner 1994). Il n'en demeure pas moins que de tels dysfonctionnements ne sont associés à une impossibilité à construire ou à maintenir une identité personnelle que lorsqu'apparaît notre société moderne individualiste et que cette éventuelle "fragilité identitaire" d'origine neuronale peut être compensée par la force du lien social.

Identité et dépression

Il nous a semblé également intéressant d'approcher de façon analogue la distorsion des rapports établis entre les dimensions corporelle, sociale et personnelle de l'identité chez les patients présentant des troubles de l'humeur (mélancolie et dépression).

A la différence de la Schizophrénie, l'entité nosologique "maladie maniaque-dépressive" a été reconnue et décrite depuis l'antiquité (Postel et Quétel, 1983) mais sa fréquence semble actuellement augmentée et sa survenue est de plus en plus précoce dans la vie de l'individu.

Cette affection est caractérisée par la survenue épisodique d'accès de durée variable, entraînant une modification radicale de l'expérience personnelle. En période "normale", ces sujets vivent en parfaite harmonie avec la société et ce qui les caractérise, c'est leur hyperidentification à leur rôle social et leur horreur du changement (Tellenbach 1979). Leur vie privée se confond avec leur vie publique et leur temps vécu s'efface au profit du temps réel. Mais ce manque de flexibilité les rend vulnérables et cela d'autant plus que nous sommes maintenant dans une société de changement qui exige "un individu-trajectoire" qui doit évoluer en permanence (Ehrenberg 1995). Toute atteinte de leur identité sociale est vécue alors comme une atteinte de l'identité personnelle et risque de provoquer la mutation de l'humeur

Pendant les accès mélancoliques, même s'ils donnent à leur entourage l'impression d'avoir changé de personnalité, les patients ne développent pas de "troubles de l'identité" au sens où on l'entend pour les schizophrènes. Bien au contraire, ils ont l'impression d'être plus que jamais "eux-mêmes", enfermés dans leurs douleurs

physiques et morales. La tristesse semble constituer une cohésion exagérée de leur moi et assurer une unité inébranlable de leur identité. Seul, le suicide, stade ultime de la mélancolie, pourrait être considéré, au sens propre, comme une destruction de l'identité. Mais, en se tuant, le mélancolique tue également l'autre et arrête l'écoulement du temps. Au moment des accès, ces patients semblent avoir fixé, une fois pour toute, leur représentation du monde, des autres et d'eux mêmes. Le temps s'est figé, leurs pensées sont ralenties, stéréotypées, imperméables aux représentations des autres, toujours biaisées par une majoration péjorative des moindres événements. Leur parole est répétitive, monotone, empreinte d'idées générales et de stéréotypies sociales et peut se tarir dans le mutisme. Ils sont impuissants à agir, à entreprendre. Leurs souvenirs sont douloureux, leurs projets inexistant. Ils éprouvent de la haine envers les autres, les vivent comme hostiles et menaçants, se vivent eux même comme indignes et coupables.

Chez ces patients, la capacité de "spécularité" loin d'être déficiente semble exacerbée. Il semble exister une liaison exagérée entre la représentation qu'ils ont d'eux mêmes, celle qu'ils ont d'autrui et celle qu'ils se font de la représentation d'eux même par autrui. D'une certaine façon, ces patients n'ont plus à affronter les paradoxes de l'identité: un mélancolique est ce qu'il est et n'est pas ce qu'il n'est pas. Il abandonne totalement à autrui la construction de son identité.

L'approche neurobiologique de cette affection a mis en évidence une fragilité des horloges biologiques qui règlent l'alternance de période d'activité et de repos, au cours des jours, des mois et des années. Et il est intéressant de souligner que ces rythmes biologiques, relativement respectés par les sociétés traditionnelles, ne sont plus pris en considération dans les conditions de la vie moderne qui nous soumet à une temporalité métrique toujours plus rapide et plus précise. Par ailleurs, des travaux récents de neurophysiologie et d'imagerie cérébrale ont révélé, au cours de périodes pré-dépressives, l'existence d'une hyperactivité des régions frontales (Timsit-Berthier, 1990).

Identité et temporalité

L'intérêt de l'approche psychopathologique est de révéler que les modifications de l'identité s'accompagnent toujours de modifications conjointes de la représentation d'autrui, du langage, et de la capacité à remplir un rôle social. Elle insiste également sur le rôle de la temporalité dans cette construction. Les oscillations temporelles exagérées, qui ne permettent pas d'intégrer le temps réel au temps vécu, telles qu'elles apparaissent dans la Schizophrénie, empêche la construction ou le maintien du sentiment d'identité. Mais l'intolérance aux changements, la soumission exagérée au temps réel du mélancolique peut entraîner des troubles importants.

Ainsi, il existerait une certaine permanence de l'identité non pas parcequ'il y a une "structure" qui s'est élaborée dans un premier temps et qui demeure comme un échafaudage stable mais parcequ'il existe un procédé de fabrication, qui se définirait de

façon autoréférentielle et récursive. Pour décrire ce procédure de fabrication, Daniel Denett (1995) a utilisé la métaphore d'un "centre de gravité narrative" en développant l'idée que l'individu ne peut se construire qu'à l'intérieur du langage, d'une "toile de discours", qui est son moyen d'autoprotection, auto définition et auto contrôle et qui apparaît autant comme un produit neurobiologique que comme un projet social et éducatif.

Si l'approche psychopathologique confirme l'importance du dialogue et du langage intérieur pour construire et maintenir un sentiment d'identité, elle montre également que cette construction doit obéir à un certain "tempo", à une certaine rythmicité, acceptant alternances, desynchronisations et synchronisations entre temps réel et temps vécu. Et l'on peut se demander si les exigences temporelles toujours plus fortes de notre société, "la tyrannie du temps réel" que dénonce Virilio (1996) ne risquent pas de mettre en péril ce processus identitaire.

Bibliographie:

- Andreewsky E 1994 General introduction. *World Futures*, 42, p1-9
- Denett D.: Comment nous tissons notre moi . in *Mécanismes mentaux, mécanismes sociaux*; edt Henri Grivois et Jean-Rierre Dupuy . La découverte. Paris . 1995. p147-165
- Ehrenberg A. L'individu incertain. *Pluriel*. Calman Levy. Paris. 1995. 351p.
- Fargeas X. et Andreewsky E Une approche cognitive de la Schizophrénie. *L'évolution psychiatrique*, 56, 1991, p133-140.
- Grivois H. De l'individuel à l'universel: la centralité psychotique. in *Mécanismes mentaux, mécanismes sociaux*; edt Henri Grivois et Jean-Rierre Dupuy . La découverte. Paris . 1995. p 23-64
- Kimura Bin *Écrits de psychopathologie phénoménologique*. PUF, Paris. 1992, 198p.
- Kulynych J.J. et Weinberger D.R. Perspective neurodéveloppementale d l'étiologie de la schizophrénie ;in: *La Schizophrénie, recherches actuelles et perspectives*. Edt par Dalery J. et d'Amato T. Masson, Paris 1995, p273-283.
- Posner M.I and Rothbart M. Attentional Regulation: From Mechanism to Culture. in *International Perspectives on Psychological Science. Vol. 1 Leading Themes*. Edt P. Bertelson, P. Eelen and G. d'Ydewalle. 1994 p41-55.
- Postel J et Quételet C. *Nouvelle histoire de la Psychiatrie Privat*, Toulouse. 1983. 774p.
- Sadoun R. Comparaison transculturelle dans la Schizophrénie. in *Actualités de la Schizophrénie*; edt Pierre Pichot. PUF, Paris, 1981, p73-82.
- Tellenbach H. *La mélancolie*. PUF, Paris, 1979, 334p
- Timsit-Berthier: *Approche neurophysiologique des états dépressifs*. *Psychologie Médicale*, 1990, 22, 8:757-763
- Searles H. *L'effort pour rendre l'autre fou*. nrf Gallimard, 1965
- Virilio P. *Cybermonde, la politique du pire*. Textuel, Paris. 1996. 108p.

Problèmes d'identité des systèmes

Evelyne ANDREEWSKY

INSERM-TLNP

La Salpêtrière

47, Bd. de l'Hôpital, 75013 PARIS

andreews@idf.ext.jussieu.fr

L'*identité* est un concept que l'on retrouve dans des *domaines scientifiques* très différents. Cette *transversalité* reflète certaines *similitudes* entre les systèmes étudiés dans ces différents domaines, susceptibles d'enrichir nos approches de ces systèmes. Favorisant ainsi complémentarité et coopération entre domaines scientifiques, le concept d'identité semble tout particulièrement propre à être abordé dans un cadre multidisciplinaire tel que celui des Journées de Rochebrune¹.

* *
*

L'identité est une notion très *complexe*. Par exemple, l'*identité d'un système social* (comme une nation, ou une entreprise) peut être définie (Beer, 1993) comme *un ethos extrêmement complexe de self reconnaissance, produit par de multiples interactions - et notamment par les interactions entre le système observé et la multiplicité de ses environnements, qui constituent de fait les éléments principaux qui déterminent cet ethos et forgent cette identité*.

Quel que soit le type de système concerné, la complexité de son identité procède de celle des nombreuses dynamiques en présence :

- d'abord, *l'identité d'un système est elle-même un système*. L'*identité personnelle* en particulier, dans la mesure où elle procède de multiples composants intriqués (à la fois ethniques, sexuels, familiaux, culturels, biologiques, ...), en interaction avec l'environnement physique, biologique et social, est une *unité* qui constitue à l'évidence un *système*. Cette unité traduit la permanence des individus dans leur environnement fluctuant. Comme le dit Voltaire (*in* « Dictionnaire

¹ Le présent article reprend en grande partie l'introduction d'un numéro spécial de la revue « World Futures », Vol. 42, N° 1-2, 1994 : *Problems of Systems Identity in the Natural, Human, and Social Sciences*.

mots sont combinés, (3) les “mots pleins” des phrases (substantifs, verbes ou adjectifs) dénotent (dans le monde) objets, propriétés, relations, et ensembles des ces entités. Pour que la fonction qui définit la signification de la phrase (voir point 2), aie quelque intérêt, il convient de supposer que la signification des mots de cette phrase est stable, indépendamment du contexte dans lequel ces mots apparaissent ... (exception faite de quelques cas répertoriés évidents) ».

La stabilité et l'indépendance de la signification des mots sont ainsi des propriétés indispensables à la “simplicité organisée” des phrases et du langage - c'est ainsi que les approches linguistiques réductionnistes, qui sont ancrées dans ces propriétés, invitent à concevoir le monde sur le même registre.

Il est difficile, selon Winograd (1986) - et bien d'autres auteurs (cf. Fargeas & Andreewsky, *World Futures*, 1994) - d'adhérer à de telles approches. Que l'on postule stabilité et indépendance des identités - ou des significations - entraîne en effet des contradictions avec les faits scientifiques comme avec les évidences quotidiennes. Soulignons que certains arguments avancés pour récuser la simplicité organisée en matière de langage sont applicables aux approches réductionnistes en matière d'identité (ce qui traduit quelque similarité entre ce dernier concept et celui de signification des mots). Ainsi que le souligne Shanon (1993) : « les psychosociologues ont fait la même erreur que les psycholinguistes cognitifs (qui avaient tenté de caractériser les significations des mots en termes de *traits sémantiques*) en essayant de caractériser l'identité personnelle en termes de traits ... dans les deux cas, et de façon très similaire, les effets de contexte et ceux des variations contingentes font que de telles caractérisations ne sont rien d'autre que des entreprises sans espoir. ».

L'alternative de la « simplicité organisée » est la *complexité organisée*. Depuis longtemps, on a essayé de définir l'identité personnelle ou sociale dans des cadres pluridisciplinaires comme ceux de l'ethno ou de la socio-psychologie. Ces tentatives vont d'une caractérisation de l'identité en termes des seules interactions sociales, à des caractérisations purement biologiques, en passant par des conceptions spécifiquement culturelles. Par exemple, Mead (1934) définit toute identité personnelle en termes des autres identités personnelles, postulant par conséquent une construction du « moi » principalement ancrée dans les interactions sociales. Pour Touraine (1980), au contraire, l'identité personnelle, loin d'être un construit social, serait plutôt quelque chose de l'ordre d'une structure dynamique, émergent d'une « force » biologique, naturelle. Les données ethnologiques nous amènent souvent à réviser un certain nombre d'évidences : Levi-Strauss (1977) admet que les sociétés primitives ne conçoivent pas l'identité personnelle comme une « unité » stable (relevant docilement de telle description structurale ...), mais plutôt comme une combinaison instable de divers éléments. Dans le même ordre d'idée, et également à partir de données ethnologiques, Gadlin (1984), met en doute l'universalité de notre conception de l'identité personnelle, et suggère que ce concept est très lié à l'histoire et aux valeurs occidentales.

philosophique ») « le mot *identité* ne veut rien dire d'autre que *même chose* ; il pourrait être rendu par *mêmeté* ... c'est par conséquent la mémoire qui crée l'identité, la « *mêmeté* » d'une personne ».

- par ailleurs, l'identité d'un système donné présente des propriétés antagonistes très complexes ; il s'agit en effet d'une unité relevant à la fois d'une *multiplicité* - celle de toutes les unités similaires des différentes super-catégories auxquelles le système appartient (par exemple, mâle, VIP, musulman ...) - et d'une *singularité*, à savoir l'*unicité* de cette unité (par exemple, l'identité biologique). Ces propriétés antagonistes font du mot *identité* une polysémie pourvue de deux significations opposées, S1 et S2, donnant lieu à d'incessantes transmutations entre S1, *identique* (aux autres) et S2, *singulier* (significativement différent de tous les autres). Un équilibre dynamique (Benhamou & Erceau, World Futures, 1994) entre ces significations opposées peut définir un point stable représentant la signification « même chose » que Voltaire proclamait être la signification « vraie » du mot identité.

- Mais quelle que soit la définition adoptée pour l'identité personnelle, il convient de postuler l'unité des individus et la pérennité de cette unité, pour permettre, selon Macintyre (1985), l'émergence de la littérature et de la culture : *l'identité personnelle n'est rien d'autre que cette identité présupposée par l'unité des personnages qu'exige l'unité du roman ; sans elle, il n'y aurait en effet ni sujets, ni histoires à raconter à leur sujet ...*

L'identité est un concept très riche. Reprenons le cas de l'identité personnelle ; un très vaste système de valeur est associé à ce concept, dans la mesure où, si l'on en croit Schrödinger (1990), *la manière dont on conçoit l'identité personnelle indique la manière dont on conçoit le monde*. Il est clair que la plupart des approches classiques de l'identité personnelle et/ou sociale sont ancrées dans un système de valeurs traditionnelles. C'est notamment le cas des approches *réductionnistes* qui définissent l'identité en termes de *simplicité organisée*. La simplicité organisée constitue bien (Lazlo, 1972) un *concept réductionniste clé*. Il implique des entités *stables* et *indépendantes* sur lesquelles travailler, qu'il s'agit d'« organiser ». D'où la tendance à considérer les identités personnelles comme des unités *indépendantes*, des éléments à *organiser* en phénomènes psychologiques, culturels ou sociaux.

On peut mettre en évidence certains aspects communs des approches traditionnelles de l'essence de l'identité et des conceptions du « monde », à l'aide d'une petite incursion dans les théories (non moins traditionnelles) du langage. Classiquement, en matière de langage et de signification, en effet, on postule également des « entités stables et indépendantes » qu'on se propose d'organiser. Il s'agit de la signification des mots, telle que la conçoit le *rationalisme naïf* (dont relèvent, selon Terry Winograd, la plupart des théories sémantiques). En effet, dans le cadre du « rationalisme naïf », le langage réfère à la *réalité objective*, et la signification des mots dénote des parts définies de cette réalité. D'où la stabilité et l'indépendance de ces significations. Comme l'indiquent Winograd & Flores (1986) : « les approches traditionnelles du langage et de la signification sont fondées sur une hypothèse de *correspondance simple*, que l'on peut résumer comme suit : (1) les phrases disent des choses sur le monde, (2) ce que les phrases disent sur le monde est fonction des mots qu'elles contiennent et des structures dans lesquelles ces

La *doctrine de l'identité* des *Upanishads* (Schrödinger, 1990) s'éloigne considérablement des manières de penser et des conceptions occidentales de l'identité personnelle. Dans le cadre des *Upanishads*, tous les organismes vivants ne sont en effet rien d'autre qu'autant d'aspects d'un et d'un seul *Être* (Dieu, ou Brahmane, pour les *Upanishads*). Ces organismes, par conséquent fortement interdépendants, ont une diversité seulement apparente, dans la mesure où chacun d'eux constitue un des points de vue du même Être. Une des conséquences d'une telle doctrine est qu'elle détermine une grande solidarité entre les êtres vivants. Schrödinger s'est efforcé de mettre en évidence la cohérence logique et les valeurs éthiques du cadre des *Upanishads*, et de souligner l'intérêt de ce cadre (par rapport à nos évidences occidentales) pour élaborer une vision solide de notre monde. Il considère, comme nous l'avons indiqué plus haut, que *la manière dont on conçoit l'identité personnelle indique la manière dont on conçoit le monde*, et présente (sous une forme condensée - à savoir seulement sept mots, en anglais!) la raison pour laquelle notre identité personnelle n'a pas de place dans nos approches scientifiques du monde : *c'est parce qu'elle est elle-même l'image de ce monde*.

Comment caractériser un phénomène aussi complexe que l'identité des systèmes ? Un numéro spécial² de la revue *World Futures* (1994) s'est proposé l'exploration des problèmes d'identité de différents types de systèmes (des systèmes biologiques aux systèmes sociaux, en passant par les systèmes de connaissances). Des stratégies similaires pour caractériser l'identité des systèmes se retrouvent dans différentes approches ; par exemple, les troubles de l'identité sont invoqués pour essayer de mieux comprendre ce concept d'identité aussi bien en psychiatrie (troubles de l'identité personnelle) qu'en socio-politique (troubles de l'identité nationale) ou en épistémologie (troubles de l'identité des disciplines scientifiques).

² « *World Futures* », Vol. 42, N° 1-2, 1994, Special Issue, E. Andreewsky, (guest editor) : *Problems of Systems Identity in the Natural, Human, and Social Sciences*.

Table des matières :

- General Introduction, Evelyne Andreewsky, 1.
- The Characterisation of Systems Identity in the Physical and the Biological Sciences, Francis Bailly, 11.
- Definition and Identity of the Living Being, André Pichot, 21.
- A Cognitive View of the Immune System, Francisco J. Varela, 31.
- Biological Identity, Adaptation to Environment as a Model for Social Science, Emmanuel A. Nunez, 41.
- Antagonism and Identity, Philippe Benhamou & Jean Erceau, 49.
- Identity and Autonomy of Psychology in Cognitive Sciences: Some Remarks from Language Processing and -Knowledge representation, Danièle Dubois, 71.
- A « Child's » Identity Problem: The Complex Development of a Cognitive Science, Pierre-Yves Raccach, 79.
- Identity, Abduction and Autonomy, Jacques Miermont, 85.
- Autoreference, Autocreation and Subject in Psychiatry, Léon Cassiers, 97.
- The Identity of Schizophrenia: A Multilevel Systems Approach, Nicholas C. Paritsis, 107.
- Schizophrenic Troubles of Personal Identity: A Cognitive Model, Xavier Fargeas & Evelyne Andreewsky, 119.
- Alter Egos: Notes on the Basic Processes of Specular Identification, Jean-Louis Vullierme, 125.
- Which Scientificity for the Social Sciences?, Jacqueline Feldman, 133.
- Systemic Interpretation of Social Relations in Organizations, Gérard Donnadiou, 145.
- On the Archetypes in Information Systems Development: an Exploration, Wita Wojtkowski, 155.
- The End of History or the End of Democracy? National Identity and the Future of the Nation State, Wolf-Dieter Eberwein, 161.

Cette fin de siècle, avec les batailles terrifiantes brandissant la bannière de tel ou tel type d'identité - *ethnique* comme au Zaïre ou en ex-Yougoslavie, *religieuse* (à la mode des fondamentalistes de tous bords), ou encore *idéologique* (avec notamment le modèle Khmers rouges), reflète intensément les dynamiques étranges de l'identité sociale, et de ses troubles qui, comme l'indique Wolf-Dieter Eberwein (World Futures, 1994), ébranlent nos conceptions des fondements des mécanismes sociaux, de l'histoire, et de la démocratie.

Un autre phénomène non moins terrifiant - lui aussi spécifique de notre époque - est le sida. Tel le précédent, ce fléau ébranle nos conceptions des fondements - ceux des mécanismes biologiques - dans la mesure où ces conceptions ne permettent pas d'induire d'interprétation satisfaisante du comportement d'un virus comme le VIH. L'étude des interactions entre ce virus et le système immunitaire semble déclencher de véritables mutations dans les approches théoriques de ce système et de ses mécanismes identitaires. Le fonctionnement immunitaire au niveau cellulaire traduit une reconnaissance du « soi » biologique, forgé à un niveau très différent - où le système immunitaire apparaît comme une *unité* (que Francisco Varela (World Futures, 1994) qualifie de *cognitive*) dans un *environnement*. C'est à ce niveau qu'il est pertinent de concevoir la *constitution dynamique de l'identité biologique* - et les perturbations de cette dynamique.

Références

- Andreewsky E. (ed.), 1994, Problems of Systems Identity in the Natural, Human, and Social Sciences, *World Futures*, Vol. 42, N°1-2.
- Beer S., 1993, « World in Torment: A time Whose Ideas Must Come », *Kybernetes*, Vol. 22, N°6, *Special Issue dedicated to Stafford Beer*, 15-43.
- Gadlin H., 1984, « The Deconstruction of the Self », *Proceedings of the International Conference on Self and Identity*, Cardiff.
- Laszlo E., 1972, *The Systems View of the World*, New York : George Braziller.
- Levi-Strauss C., 1972, *L'identité*, Paris : Grasset.
- Macintyre A., 1985, *After Virtue, a Study in Moral Theory*, London : Duckworth.
- Mead G. H., 1934, *Mind, Self and Society*, Chicago: Univ. Press.
- Schrödinger E., 1990, *L'esprit et la matière*, Paris : Seuil.
- Shanon B., 1993, *The Representational and the Presentational*, New York: Harvester Wheatsheaf.
- Touraine A., 1980, « Les deux faces de l'identité », in Tap, P. (ed): *Identités collectives et changements sociaux*, Toulouse : Privat, 19-26.

"Qui suis-je?" "Simple: je suis le chef":
Identité individuelle, perception et descriptions
par Juan GONZALEZ

Dans ce papier j'aborde le thème de l'identité individuelle sous un angle philosophique. Je m'intéresse à la dynamique liant notre structure perceptive et cognitive à nos schémas conceptuels et, particulièrement, au rôle que les catégories de classification sociale jouent à l'intérieur de ces schémas pour conférer un sens à l'identité individuelle et aux descriptions s'y attachant. Je suggère que perception, identification et référence se co-déterminent réciproquement de manière à incorporer, dans une seule perspective, une vision significative et cohérente du monde : un monde que structurent l'environnement social et ses institutions fixant l'espace logique des descriptions qu'on fait de nous-mêmes et d'autrui. Notre identité individuelle et ses multiples descriptions se verraient ainsi largement déterminées par cet environnement, grâce auquel on devient quelqu'un...qui suit le chef.

I- Introduction

Platon, dans La République, conçoit le psychisme individuel comme étant composé de trois éléments. Dans le même temps, il décrit la société comme, elle aussi, composée d'éléments qui correspondent en nombre et en nature aux éléments du psychisme individuel¹. Platon établit donc un parallèle ontologique et architectonique entre la structure psychique individuelle et la structure sociale. Celles-ci partageraient, mutatis mutandis, une

¹ Ces éléments sont : l'appétit, l'ardeur et la raison. Ceux-ci correspondent, respectivement, à trois classes sociales : les commerçants, les soldats/fonctionnaires et les conseillers ; et à trois vertues : la sobriété, le courage et la sagesse. Platon, *La République*, IV, 434c-436c; 439d-445e

même identité formelle. Autrement dit, la conception de Platon postule un rapport identitaire *fractal* entre société et individu.

Aristote, pour sa part, dans Les Catégories, aborde la question de l'identité individuelle à partir d'un raisonnement sur les catégories logiques utilisées pour qualifier ou définir le sujet grammatical (i.e., le substantif —qui représente le sujet ontologique dans une proposition) et pour nous permettre de donner un sens univoque à nos énoncés². Ainsi, le prédicat 'est un homme' permet d'attribuer un sens univoque à la phrase 'Socrate est un homme' en vertu de la signification de (la catégorie) 'homme'. Ce prédicat, si l'on dit vrai, nous renseignerait sur une propriété ou caractéristique de Socrate au moyen de laquelle on identifierait le sujet. Aristote introduit donc d'emblée le problème de l'identité individuelle dans une dimension logico-linguistique.

Ce que je retiens pour mon propos de Platon —tout en récusant son essentialisme— c'est qu'il aborde l'identité individuelle comme étant une problématique psychosociale sous-tendue par un souci d'ordre politique et éthique³. D'un autre côté, ce que je relève d'Aristote c'est qu'il aborde cette question en tant que problématique psycholinguistique sous-tendue par un souci d'ordre logique et, si j'ose dire, cognitif.

Dans ce papier je ferai un bref écho à ces problématiques, car, bien que les contextes aient changé, elles demeurent d'une actualité remarquable pour toute discussion sur l'identité et les descriptions.

² Aristote, *Categoriae*, ch 1-5, dans Organon.

³ On trouve aussi chez Hume la comparaison entre le psychisme individuel et l'institution sociale: "...I cannot compare the soul more properly to any thing than to a republic or commonwealth, in which the several members are united by the reciprocal ties of government and subordination, and give rise to other persons, who propagate the same republic in the incessant changes of its parts....without losing (its) identity.". *A treatise of Human Nature*, in Fundamental Problems in Philosophy, p. 66.

Ainsi, je compte ébaucher ici quelques rapports entre notre structure cognitive et la structure sociale intervenant dans le processus de constitution de l'identité individuelle. La discussion tournera autour de deux types de description qui sont solidaires et essentiels à ce processus : l'auto-description et la description d'autrui.

II- Percevoir le monde

II.1- La dynamique perceptive et la dimension cognitive :

Dans la littérature scientifique sur la vision⁴, on reconnaît volontiers l'apport, mais pas toujours l'importance, du processus ou traitement cognitif que l'information provenant du monde visuel subi. On parle donc tantôt d'un niveau 'haut' ou 'supérieur'⁵, tantôt d'un système 'central'⁶, tantôt d'un processus 'descendant'⁷, impliqués dans la vision. Dans une formulation classique, on dit que la représentation sémantique repose sur le traitement cognitif des stimuli sensoriels. On peut dire alors que c'est à partir des signaux sensoriels que le système visuel et/ou la conscience perceptive (moyennant le traitement cognitif de l'information) accède à un monde visuel significatif où les choses, les animaux et les personnes ont une identité et un nom⁸. Ce traitement —aussi difficile à modéliser que facile à reconnaître dans la plus simple

⁴La vision est prise ici comme exemple de dynamique perceptive, mais mon propos s'applique aussi, plus ou moins dans la même mesure, aux autres modalités sensorielles

⁵ Cf. par exemple, pp.32-33, in Crick, F., The Astonishing Hypothesis

⁶ Cf. par exemple, pp 262-264, in Sekuler & Blake, Perception

⁷ Cf. pp. 184-188, in Goldman, A., Epistemology and Cognition

⁸Ce niveau correspond au rapport 2 1/2D—>3D de Marr, à la conscience 'épistémique' (access consciousness) de Block, et à la 'digitalisation' de Dretske.

tâche d'identification visuelle et de description— est, par conséquent, constitutif de la dynamique perceptive.

On peut concevoir la dynamique perceptive en tant que lien qui met en rapport la dimension sensorielle (celle des signaux ou des stimuli) avec la dimension cognitive⁹ (celle des phénomènes —les phénomènes étant ce qu'on voit : ce qui se déploie sous nos yeux et/ou ce qui apparaît à notre conscience perceptive). Mais cette dynamique ne saurait avoir un contenu significatif *sémantique* (le ce qu'on voit ne pourrait se nommer ou se décrire) sans les concepts qui lui 'injectent' un sens et qui permettent au percevant de voir les choses *en tant que x*¹⁰. C'est pourquoi la dimension conceptuelle (celle des symboles et de la sémantique, où les phénomènes deviennent intelligibles, prennent un sens, et peuvent être décrits) fait aussi partie intégrante de la dynamique perceptive.

Ces trois dimensions, considérées ici comme parties d'un bloc solidaire, représentent la vision comme étant une fonction qui renseigne le percevant sur l'état du monde à un endroit et à un moment donnés, dans des termes fonctionnellement et/ou symboliquement significatifs. Dans le reste de ce papier, c'est l'aspect symbolique et de haut-niveau de la vision que retiendra mon attention.

Ainsi, c'est la dynamique perceptive, inscrite dans une pratique, ne serait-ce qu'au niveau langagier, qui me permettrait de voir des choses, des événements, des animaux, des personnes, etc., et de les spécifier/individuer sémantiquement au moyen de concepts : voir, par exemple, ma main, Jean venir, un arbre tomber,

⁹ Ici, 'dimension cognitive' désigne la dimension phénoménale : le champ auquel se rapportent les fonctions et les expériences perceptives.

¹⁰ Où 'x' désigne un nom ou une description (définie ou indéfinie).

le président gémir, que la soupe est prête, etc. Notons cependant trois points : 1) que [bien que] le statut ontologique de ces états du monde reste à élucider, 2) cognitivement, ces états sont indissociables de leur formulation linguistique (i.e., [la présentation symbolique de] ce que l'on voit dépend de l'expression descriptive utilisée pour le spécifier¹¹ et, inversement, les descriptions émises dépendent de ce que l'on voit). Cette dépendance réciproque, qui n'est pas forcément symétrique, nous permet de faire le lien entre notre expérience perceptive et nos pratiques langagières. À leur tour, ces pratiques fournissent un moyen de spécifier/individuer sémantiquement le monde perçu. À ce propos, et suivant 'le deuxième' Wittgenstein, on peut dire que

"...les relations sémantiques fondamentales (les connexions entre le langage et la réalité) sont établies et maintenues par les jeux de langage."¹²

Enfin, 3), que les 'renseignements' concernant les états du monde sont toujours ancrés dans un contexte et relatifs, en un sens spécial du terme, à notre point de vue¹³.

II.2- *La dynamique perceptive et les descriptions:*

¹¹ A cet effet, D. Marr dit "...any particular representation makes certain information explicit at the expense of information that is pushed into the background and may be quite hard to recover." Et puis: "...even though one is not restricted to using just one representation system for a given type of information, the choice of which to use is important and cannot be taken lightly. It determines what information is made explicit and hence what is pushed further into the background." (my underlining) p. 21, in *Vision*.

¹² J. Hintikka, <<Semantics: A revolt against Frege>>, p. 60, cité par J. Bouveresse dans *La force de la règle*, p. 58.

¹³ La subjectivité en question n'est pas celle du solipsisme méthodologique, mais, plutôt, celle qui renvoie à deux traits élémentaires de la perception : 1) le percevant occupe un point de vue spatio-temporel et historico-culturel particulier ; 2) le percevant, en tant que point de vue, met en relation divers aspects du monde perceptif de manière à diriger et à planifier ses actions et à trouver un rapport significatif et cohérent entre plusieurs états du monde.

Toute description visuelle repose donc sur un processus perceptif dynamique qui, au niveau cognitif, établit un lien entre (l'information provenant) des états du monde et une expression linguistique, qui, si elle est vraie, leur correspond.

Ainsi, la description d'un objet, d'une scène ou d'une situation visuelle quelconque, en tant que "x", "un x", ou "la x"¹⁴, implique une capacité cognitive qui permet de classer l'information du monde selon la situation du percevant, mais, aussi et surtout, selon la pertinence de cette information relative aux schémas conceptuels et aux *jeux de langage* du percevant. Donc, bien qu'il soit très discutable de savoir si, et dans quelle mesure, les connexions sémantiques *pénètrent*¹⁵ les niveaux bas et moyen de la vision¹⁶, au niveau supérieur symbolique on peut établir à bon droit que la dimension conceptuelle, fixée par nos pratiques langagières, joue un rôle déterminant dans notre manière de percevoir le monde et donc dans notre manière d'agir. Autrement dit, s'il y a une continuité effective et bidirectionnelle entre signal et symbole, alors nos croyances et notre perception doivent guider notre action autant que nos concepts et nos pratiques doivent guider nos croyances et notre perception.

Les notions wittgensteiniennes de 'jeux de langage' et celle, s'y attachant, de 'forme de vie' s'avèrent très éclairantes ici, car elles montrent un possible passage entre la structure sociale et notre structure cognitive. En effet, ces notions montrent que nos

¹⁴ Où *x* désigne un nom commun. Les expressions comportant un article (défini ou indéfini), je les traite *grosso modo* selon la théorie de descriptions de Russell. Cf. Descriptions. Quant aux noms propres je les traite en accord avec la définition de Searle: un nom propre est un 'crochet' sur lequel on pose une description. Cf. Proper names.

¹⁵ Notion d'après Z. Pylyshyn. Cf. pp. 130-145, Computation and Cognition.

¹⁶ Fodor, par exemple, défend un modèle de la perception modulaire, plutôt impénétrable, et largement étanche au niveau symbolique des croyances.

pratiques sociales et nos institutions —elles-mêmes ancrées dans un environnement physique et culturel— déterminent nos schémas conceptuels individuels. Et, comme on l'a vu, ces schémas déterminent le sens avec lequel on perçoit le monde et agit dans celui-ci.

En résumé : les descriptions du monde qu'on formule sont fonction de l'information provenant de celui-ci, mais aussi de la manière dont notre structure cognitive intègre (dans la dynamique perceptive) la dimension conceptuelle, elle-même fixée par nos *jeux de langage*. Cette dimension nous permet d'attribuer un sens à ce qu'on voit ; et c'est en fonction de ce qu'on voit que l'on agit. Perception, croyance et action sont donc intimement liées. La manière dont on décrit le monde reflète la manière dont on perçoit le monde ; la manière dont on perçoit le monde détermine la manière dont on agit.

III- Décrire le monde

Parmi les possibles états perceptibles du monde, je concentrerai mon attention sur ceux attachés aux descriptions émises sur soi-même et sur autrui.

III.1- *Décrire autrui et se référer à autrui :*

Si l'on demande à une personne de décrire quelqu'un qu'elle voit pour la première fois et dont elle ne sait rien, il est probable qu'elle parlera de son âge apparent, de son sexe, de sa couleur, de sa complexion, de sa taille, etc. Il est moins probable qu'elle le décrive dans des termes techniques (p.ex., " c'est un homo sapiens sapiens") ou dans des termes très détaillés (p.ex., "c'est quelqu'un qui a des narines poilues"). Cela met en évidence, comme cela a été

démontré au sujet des objets¹⁷, qu'il y a des catégories plus ou moins 'universelles' pour décrire autrui. Toutefois, ces catégories sont loin d'être les seules pour décrire les gens, mais elles sont de loin le plus utilisées.

Or, cette constatation serait presque triviale si le rapport entre notre perception et les états du monde n'était pas contingent (sur un plan symbolique). Mais, comme nous l'avons vu, la description du monde perceptible n'est pas seulement fonction d'une information externe : elle est aussi fonction de variables conceptuelles, contextuelles, et pratiques, qui spécifient et déterminent le contenu perceptif. Cela revient à dire qu'il n'y a pas de contenu sémantique universel et nécessaire. Pourtant, dès qu'on parle du monde perceptible, on a l'impression qu'on voit tous la même chose¹⁸. Pour expliquer cette 'impression', j'avance que c'est justement le fait de parler du monde perceptible —au moyen des *jeux de langage*— qui structure et conditionne le contenu perceptif de façon à le rendre communicable. Ceci s'oppose à l'idée que la communicabilité des contenus perceptifs est due au partage communautaire d'un noyau informationnel universel et objectif.

Revenons aux descriptions d'autrui. Si la description de quelqu'un qu'on ne connaît pas passe par des catégories 'froides' (du style 'renseignements généraux'), celle de quelqu'un qu'on connaît passe par des catégories 'chaudes' (des catégories où il y a un contenu émotif). Dans ce cas-là, ce sont la connaissance et l'expérience individuelles qui permettent l'accroissement informatif et le 'réchauffement émotif' entre deux descriptions, séparées dans le temps, qui partagent le même référent (p. ex., entre "l'Allemand

¹⁷ Cf. p.ex., pp.79-105, in Boucart, M., La reconnaissance des objets

¹⁸ 'Qu'on voit tous la même chose' est faux et vrai à la fois, car cela dépend de ce qu'on entend par 'voir' —question qui n'est pas possible de traiter ici.

que j'ai rencontré au colloque" [il y a dix ans] et "mon meilleur ami" [maintenant]). Toutefois, sans ou avec connaissance d'autrui, les catégories utilisées habituellement pour le décrire demeurent largement celles provenant des institutions et des moeurs.

C'est-à-dire que, dans la mesure où nos *jeux de langage* déterminent le contenu sémantique de nos descriptions d'autrui, dans cette même mesure on peut dire que notre perception d'autrui est tributaire de nos institutions et de nos pratiques sociales. Le versant conceptuel et symbolique de la dynamique perceptive se trouve donc, *de facto*, 'rempli' par la signification des expressions linguistiques usuelles et disponibles dans une communauté linguistique donnée. Ainsi, par exemple, "le Noir d'à côté" peut devenir, "le Français qui chante sous la douche" pour devenir "le petit ami de la femme de mon ex-mari", etc. On voit comment ces descriptions fixent une identité au référant: description, référence et identité se co-déterminent réciproquement sous l'égide des institutions et des moeurs. Et cette co-détermination, de par le caractère coercitif des institutions et des moeurs, limite le sens et le nombre d'expressions disponibles pour décrire autrement autrui .

III.2- *Indexicaux et noms propres :*

Pour infléchir la prise de l'institution sur nos schémas conceptuels et pour modérer la pression exercée par nos moeurs, on peut développer une stratégie d'usage basée sur les indexicaux et sur les noms propres.

Les indexicaux sont des outils linguistiques qui désignent 'à vide', c'est-à-dire servent pour désigner, sans nommer, quelque chose ou quelqu'un en particulier. 'Aujourd'hui', 'ici', 'eux', en sont des exemples. On peut donc se référer à quelqu'un comme "elle, là-

bas", pour éviter de dire "la grosse à lunettes". Le problème de cette stratégie est qu'elle est totalement contextuelle et demande un signalement ostensif, ce qui n'est pas très pratique.

La stratégie des noms propres offre plus de possibilités communicatives, mais souffre d'un manque d'universalisme qui la rend inintelligible au-delà d'un contexte hyper-spécialisé. Ainsi, qu'un étranger vous parle de Brigitte n'a pas de sens, car 'Brigitte' renvoie à une personne (supposons) qu'on ne connaît pas. Donc, toute la richesse polysémique qu'on puisse obtenir au sujet de Brigitte exige la mise en relation causale, le partage d'une connaissance/expérience, entre les participants d'une conversation. Autrement, l'explicitation de 'Brigitte' devra passer par une description préalable : dans ce cas, on est renvoyé à notre point de départ. Par ailleurs, notons que la complexité de l'identité d'autrui augmente proportionnellement à la connaissance qu'on a de lui.

III.3- *Se décrire soi-même : Identité et Institutions*

Si je me regarde dans la glace et que je me demande "Qui suis-je ?", vais-je dire "Mâle, 45 ans, marié, 3 enfants" ? Vais-je dire "Blanc", "Italien", "catholique", "anarchiste", "boulangier" ? Ou encore vais-je dire "Intelligent", "myope", "frileux", "artiste" ? En fait, même si mon identité 'tombe sous le concept' de ces catégories, je ne me sentirais ni essentiellement ni exhaustivement décrit. Je suppose que notre identité individuelle, pour certains d'entre nous, est si riche et plastique à nos yeux, que l'idée d'une fixation définitive irait à l'encontre de la polysémie de notre personne et du monde. Réciproquement, étant donné le fonctionnement de la dynamique perceptive telle qu'elle a été évoquée plus haut, la

fixation sur une identité individuelle particulière entraînerait une vision appauvrie ou stérile du monde.

C'est maintenant que l'on peut exposer pleinement le rôle des institutions dans le rapport entre la structure sociale et la structure cognitive individuelle. D'une part, les institutions —notamment celle de la langue— servent à structurer une dimension conceptuelle qui est nécessaire pour donner un sens à nos descriptions perceptives. Ces institutions, communes à une culture donnée, homologuent nos rapports communicatifs et rendent intelligibles et partageables les descriptions qu'on émet. D'autre part, les institutions —parce que formelles et normatives— obligent les membres d'une communauté linguistique et culturelle donnée à utiliser une *grammaire* (i.e., les *jeux de langage* et les règles qui les régissent) particulière et à se conformer aux pratiques de l'environnement social. En ce sens, le champ de liberté pour nous décrire et pour décrire autrui s'exprime comme la distance qu'on laisse entre notre perception et l'autorité théorique et pratique à laquelle on obéit pour décrire et communiquer ce que l'on perçoit.

Concernant cette autorité linguistique, et assez paradoxalement, elle est la cristallisation de nos pratiques et non une instance indépendante de celles-ci. En effet, ainsi que le fait remarquer J. Bouveresse à propos de l'autorité de la grammaire :

"C'est en un sens nous qui sommes les maîtres parce que, comme le répète Wittgenstein, nous pouvons jouer ou ne pas jouer le jeu. Mais, d'un autre côté...si un mot que j'utilise doit avoir un sens, il faut que son utilisation implique de ma part un <<engagement>>... Mais s'engager, c'est ici moins s'engager à que s'engager *dans* (une certaine praxis collective)"¹⁹

¹⁹ p. 341, J. Bouveresse, La parole malheureuse

III.4- *Se décrire soi-même : Écologie et changement*

Dans une perspective écologique, l'identité individuelle est directement liée à l'environnement physique et social et aux interrelations dans ceux-ci. Comme Stefan Hormut le dit: "A concept of self arises out of social and efficacious action in an environment that provides stability through implicational links. To change oneself requires changes in the relationship between individual and environment.". Et plus loin: "Active change in relationships with others is difficult as long as the existing ecological system of the self stays in place..." ²⁰

Il ne suffit donc pas de croire que notre identité et la signification de nos descriptions sont données par notre volonté, notre conscience ou notre perception 'interne' de nous-mêmes. En effet, nous entretenons des rapports de dépendance et de détermination tellement étroits vis-à-vis de l'environnement physique, symbolique et social (qui nous accueille et nous transforme depuis notre naissance), que, comme disait Pascal, l'institution devient [souvent à notre insu] comme une deuxième nature humaine. L'idée de liberté et de pouvoir, donc, que nous expérimentons lorsque nous pensons et nous parlons, risque fort d'être un leurre si nous ne connaissons pas d'abord les contraintes inhérentes à notre fonctionnement cognitif et à nos pratiques sociales. C'est pourquoi, jusqu'à preuve du contraire, la force de la règle est maître et chef de nos chantiers perceptifs.

La bonne stratégie pour articuler la prise de l'institution, pour faire jaillir la polysémie du monde, et pour personnaliser les descriptions qu'on fait de soi-même et d'autrui reste à définir. Mais les possibilités qui s'offrent lorsqu'on se regarde dans la glace sont

²⁰ p. 90, The Ecology of the self

déjà une bonne piste, car elles pointent vers l'irréductibilité institutionnelle de la personne — que ce soit soi-même ou autrui — vers la dimension éthique du regard et vers la dimension politique des descriptions.

IV- Conclusion

Pour finir, je voudrais citer quelques passages du livre Ce que parler veut dire, de Pierre Bourdieu, pour montrer quelques prolongations cognitives, conceptuelles, politiques et éthiques, relatives à la question de l'identité individuelle, impliquées par nos descriptions de nous-mêmes et d'autrui. Si nos descriptions décident de notre 'ontologie' et de celle d'autrui, alors nos paroles sont potentiellement porteuses d'identités riches et variées.

"L'institution d'une identité, qui peut être un titre de noblesse ou un stigmate (<<tu n'es qu'un...>>), est l'imposition d'un nom, c'est-à-dire d'une essence sociale. Instituer, assigner une essence, une compétence, c'est imposer un droit d'être qui est un devoir être (ou d'être). C'est *signifier* à quelqu'un ce qu'il est et lui signifier qu'il a à se conduire en conséquence. [...] L'essence sociale est l'ensemble de ces attributs et de ces attributions sociales que produit l'acte d'institution comme acte solennel de catégorisation qui tente à produire ce qu'il désigne. Ainsi, l'acte d'institution est un acte de communication mais d'une espèce particulière : il *signifie* à quelqu'un son identité, mais au sens à la fois où il la lui exprime et la lui impose en l'exprimant à la face de tous (*kategoresthai*, c'est, à l'origine, accuser publiquement) et en lui notifiant ainsi avec autorité ce qu'il est et ce qu'il a à être."²¹

Parler, en fait, n'est jamais innocent : des choses sont dites...et bien plus qu'on ne le soupçonne.

²¹ pp. 125-126, Ce que parler veut dire

V- Bibliographie

- Boucart, M., La reconnaissance des objets, 1996. Presses Universitaires de Grenoble.
- Bourdieu, P., Ce que parler veut dire, 1982. Fayard, Paris.
- Bouveresse, J., La parole malheureuse, 1971. Les éditions de minuit, Paris.
- Bouveresse, J., La force de la règle, 1987. Les éditions de minuit, Paris.
- Crick, F., The Astonishing Hypothesis, 1994. Charles Scribner's Sons, New York.
- Goldman, A., Epistemology and Cognition, 1986. Harvard University Press, Massachusetts.
- Hamilton, E. & Cairns, H., (eds.), Plato, The Collected Dialogues, 12th. ed., 1985. Princeton University Press, New Jersey.
- Hanfling, O., (ed.), Fundamental Problems in Philosophy, 2nd. ed., 1980. Basil Blackwell, Oxford.
- Hormuth, S. The ecology of self, 1990. Cambridge University Press
- Marr, D., Vision, 1982. Freeman, New York.
- McKeon, R., (ed.), The Basic Works of Aristotle, 35th. ed. Random House, New York.
- Pylyshyn, Z., Computation and Cognition, 4th.ed., 1986. MIT Press.
- Russel, B., Descriptions (1919), in *The Philosophy of Language*, Ed. by Martinich, A.P., 1985. Oxford University Press, New York.
- Searle, J., Proper Names (1958), in *The Philosophy of Language*, Ed. by Martinich, A.P., 1985. Oxford University Press, New York.
- Sekuler, R. & Blake, R., Perception, 2nd. ed., 1990. McGraw-Hill

Merci à Olivier Fontaine pour ses commentaires, ses corrections et sa synergie toujours éclairante

Critique de la reconnaissance comme identification

Stéphane HERBIN

Ecole Normale Supérieure de Cachan
Centre de Mathématiques et de Leurs Applications
Groupe DIAM
CNRS, URA 1611
61, Av. Président Wilson
94235 Cachan Cedex
FRANCE
email: herbin@cmla.ens-cachan.fr

Résumé

Cet article se propose de montrer que l'étude et la mathématisation d'une reconnaissance libre et authentique est manquée si elle est assimilée à l'identification conditionnée par un donné empirique d'un modèle d'objet stocké en mémoire.

La science et la pensée rationnelle sont hantées par deux spectres, deux «scandales»: celui de la diversité du monde, et celui de la contingence du changement, d'une aliénation incoercible et incontrôlable. La pensée se confronte à un défaut d'universalité insupportable pour elle: sans unité, pas d'esthétique, sans éternité, pas de repos à espérer. Le réel ne se soumet qu'illusoirement aux idéalités.

L'une des capacités cognitives les plus étonnantes et mystérieuses consisterait à mettre en évidence de l'identique et du même là où un animal, par exemple, ne verrait qu'une pure diversité sans signification. Selon une tradition cartésienne, c'est d'ailleurs par cette faculté que l'unité de l'esprit se manifeste, que l'humain se montre supérieur en développant une intuition catégoriale.

La question critique clé ne concerne-t-elle pas une problématique retournée: comment, en partant de l'unité de l'esprit ou de la conscience, une diversité peut-elle se manifester? Il n'y a rien de surprenant à

ce qu'une unité posée *a priori* se reflète dans le spectacle d'un monde contingent et se contemple Une: il s'agit plus alors d'un argument tautologique ou sophistique que d'une réelle découverte. Ce que cet article se propose de montrer est que l'unité a besoin de divers pour être, que l'identité est intégration de la diversité, que les catégories fondamentales de l'Un et du Multiple ne forment pas une alternative dialectique rigide.

Il n'est cependant pas besoin d'aborder la question de l'identique sous un angle métaphysique pour être en prise avec de réelles difficultés conceptuelles. Cet article se propose d'étudier restrictivement la pertinence du concept d'identification dans le cadre d'un objectif de modélisation mathématique de la reconnaissance.

Il faut bien comprendre quelle peut être la place des mathématiques dans une entreprise de connaissance rationnelle. Leur rôle est réductionniste, non parce qu'elles réduisent une phénoménalité complexe à un jeu de formules symboliques — d'ailleurs, dans ce sens, toute entreprise conceptuelle est réductrice — mais parce qu'elles permettent de révéler les essences, de mettre en évidence ce qui constitue l'architecture des phénomènes. En ce sens, en plus de fournir un moyen de communication irremplaçable car exigeant sur la précision de sa formulation, les mathématiques constituent un outil méthodologique pour *aller à l'essentiel*. Mais pour cela, il importe d'oublier temporairement un siècle de subordination des mathématiques à la logique formelle, provoquée par la recherche des fondements et la constitution de l'informatique, pour leur redonner leurs capacités

descriptives et poétiques. Il n'est pas besoin d'édifier des mathématiques nouvelles, même si la créativité reste bien sûr à l'ordre du jour, mais de renouveler et de retrouver leur rôle *eidétique*.

La question de la reconnaissance renvoie à trois problématiques fondamentales recouvrant une grande partie du questionnement philosophique et métaphysique: la manifestation d'une extériorité objective, une expérience de la connaissance et l'inscription dans une temporalité. L'étude de la reconnaissance peut être vue comme un point de rencontre et de confrontation permettant de mettre à l'épreuve la cohérence extrinsèque des théories. C'est comme révélateur des difficultés conceptuelles liées à une approche représentationnelle dogmatique de la cognition que la reconnaissance sera essentiellement abordée ici. L'enjeu de l'étude de la reconnaissance dépasse de loin les questions techniques de stockage et de rappel de l'information auxquelles les réduisent trop souvent la psychologie et l'intelligence artificielle. C'est l'image même de la cognition qui est ainsi visée.

Nous partons d'abord d'une critique du schéma standard de reconnaissance qui la considère comme appariement adéquat entre une représentation et un donné. Nous mettons en évidence le fait que le passage obligé par la représentation réduit l'activité du sujet à une manipulation normative et conformiste, sans souci ontologique, une syntaxe neutre mais vide. En suivant quelques analyses de Gilles Deleuze nous tenterons de trouver l'origine de l'attrait pour le paradigme représentationnel et par là également la difficulté de s'en dégager.

Ces analyses nous conduiront à présenter dans un cadre phénoménologique le concept d'*altérité circonscrite* pour rendre compte d'une diversité essentiellement intégrée et affirmée de la pensée, condition à la constitution d'un reconnu comme tel. Nous ferons ensuite résonner ce concept avec celui de multiplicité virtuelle tel qu'il est proposé par Deleuze.

La productivité de ces deux concepts sera enfin montrée dans la proposition d'une définition positive du contexte qui ne le réduise pas à un paramètre ou à une variable cachée.

1 Les exigences de la représentation

1.1 L'approche représentationnelle de la reconnaissance

Lorsque je suis mis en présence de quelqu'un et que je suis capable de lui donner un nom, on peut dire que je l'ai reconnu. J'ai utilisé certaines facultés, agencées fonctionnellement, pour dénommer la personne devant moi. Ce quelqu'un que je reconnais, je peux l'avoir déjà vu en personne (un ami), ou non (une célébrité que je rencontre dans un restaurant pour la première fois). Je peux même n'en posséder qu'une description partielle (lors d'un «blind date»). Cette diversité des conditions de reconnaissance conduit alors assez spontanément au concept de *représentation* ou de modèle. Cet objet théorique, est à vocation universelle, et devient le double abstrait de la personne qui se trouve devant moi.

Dans une approche représentationnelle, la reconnaissance est assimilable à un processus d'indexation d'une base de représentations stockées en mémoire. Les problématiques fondamentales liées à l'étude de la reconnaissance — extériorité objective, expérience de la connaissance, inscription temporelle — se réduisent à l'élucidation des caractéristiques du ou des modes de rappel de l'information et de l'organisation du stockage. La reconnaissance consiste alors en l'*identification* conditionnelle d'une représentation.

Les réductions successives de la reconnaissance à une fonction mnémonique puis à un processus de traitement de l'information est une attitude standard en psychologie, neurosciences et intelligence artificielle. Même si la présentation qui en est donnée ici est volontairement caricaturale, il est peu de théories qui s'en écartent radicalement. Pour s'en persuader, il suffit de feuilleter quelques ouvrages récents traitant de la perception, de la mémoire ou de la catégorisation, ou même spécifiquement de la reconnaissance d'objets [2]. Je citerai simplement un paragraphe extrait de l'article PERCEPTION du Dictionnaire de Psychologie qui à mon sens résume l'attitude de sens commun imprégnant la plupart des recherches actuelles sur la reconnaissance, en l'assimilant d'emblée à une identification de représentation:

«Toute identification d'un objet sonore ou visuel, par exemple, nécessite le recours à une représentation en mémoire de cet objet. [...] La remarquable efficacité de

l'identification suppose que nous disposions de représentations mentales préalables des objets et que les mécanismes d'identification consistent à apparier l'information sensorielle actuelle avec ces représentations.»¹

Les sciences cognitives ont établi le concept de représentation comme principe fédérateur de leur objectif commun: une élucidation rationnelle de l'intelligence, sous tous ses aspects. La représentation procure ainsi un noyau de cohérence minimal à même de se renforcer dans l'accord d'autres éléments théoriques. Mais ce recours obligatoire à la représentation conduit à assimiler systématiquement la manifestation d'une extériorité objective à une perception autonome, la connaissance à des processus de catégorisation ou de résolution de problèmes, et la temporalité à la présence d'une mémoire de stockage et de processus séquentiels de traitement de l'information. La représentation comme principe permet une certaine homogénéité des approches, mais n'apporte guère de réponse aux questions philosophiques fondamentales.

Dans la citation précédente, les arguments en faveur de l'usage théorique des représentations, au lieu de les fonder sur un rôle méthodologique pratique limité, somme toute assez raisonnable, s'appuient sur une justification téléologique garantissant leur ontologie. C'est oublier que la représentation est un objet théorique non fondé, au sens que la phénoménologie donne à ce terme, c'est-à-dire sans réalité phénoménale. Ce point est particulièrement sensible lorsque l'on cherche à établir un rôle phénoménal à la mémoire considérée comme stockage de représentations.

1.2 Mémoire de représentations

Pour pouvoir déterminer le nom d'une personne reconnue, il faut bien l'avoir déjà intégré et «stocké», pour lui associer un visage, il faut d'une certaine manière avoir conservé cette relation de référence. Une mémoire de représentations jouerait ainsi un rôle central dans la fonction de reconnaissance.

La réduction de l'activité de reconnaissance, et plus généralement de l'activité cognitive, à un ensemble de processus de stockage, indexation et trai-

tement de représentations génère deux types de problèmes: phénoménologiques et ontologiques.

L'introduction du concept de représentation pourtant est attirante: elle permet de réduire une phénoménalité et une activité complexes en plusieurs composantes susceptibles d'une étude scientifique indépendante. Cependant, cette partition théorique est problématique: phénoménalement, quand je reconnais un ami qui rentre dans mon bureau, c'est lui tout entier qui investit mon esprit, qui le sature, ce n'est pas une représentation ou un modèle que je *retrouve*. Même par une introspection toujours douteuse, biaisée ou illusoire, je suis incapable d'isoler cette représentation impassible postulée qui permettrait d'attribuer un nom à la personne devant moi.

Phénoménologiquement, la mémoire comme stockage de représentations est difficilement intelligible. Même si Husserl utilise le terme de représentation dans ses écrits², ce n'est pas dans le sens que lui donne la théorie moderne de l'information. La représentation n'est pas déjà là, en attente d'être retrouvée: elle est le produit d'une ontogénèse des objets présents à la conscience, d'un «remplissement de l'intention par l'intuition». Le *reconnu*, le pôle intentionnel objectif de la faculté de reconnaissance, est à constituer dynamiquement, activement.

C'est précisément la question de l'intentionnalité de la mémoire qui est la plus problématique. Je ne peux être sûr d'avoir mémorisé un quelque chose que lorsque celui-ci participe à l'intentionnalité d'une autre faculté, comme la reconnaissance. Il n'est pas possible de parler d'un mémorisé en tant que tel, pour lui même, sans le relier à une faculté cognitive intentionnelle, objectivante. Le questionnement philosophique a d'ailleurs toujours posé le problème de la mémoire comme manifestation de la dimension du passé. Les grandes avancées conceptuelles (Kant, Bergson, Husserl, Heidegger) ont consisté non pas à justifier une entité — la mémoire — dans sa participation à la conscience, mais plutôt à montrer que la conscience est d'emblée temporelle, n'est intelligible que comme conservation et accumulation du passé. Ainsi, parler d'une fonction spécifiquement dédiée à la récolte de représentation est-il un non-sens, sauf à la considérer comme la conscience tout entière.

Ceci nous amène à la question ontologique de la reconnaissance comme processus de rappel de l'information. La conception représentationnelle de la

1. in *Grand Dictionnaire de Psychologie*, Larousse, 1994, article PERCEPTION, p. 556

2. cf. le dernier chapitre de la Recherche V.

mémoire lui fait manier des objets dont les conditions d'existence doivent être trouvées ailleurs. La mémoire serait de ce fait, au mieux, *ontologiquement neutre*. C'est là un point essentiel à saisir: l'origine de la puissance de la théorie de l'information est précisément qu'elle ne pose pas de conditions d'existence aux objets. Il n'y a pas de principe d'incompatibilité *a priori* par exemple: rien n'empêche de mélanger des torchons et des serviettes, ou de considérer des «œufs poilus» [14]. Mais le refus de participation de la mémoire représentationnelle à l'être des objets qu'elle stocke suppose le recours à une ontologie *générique*, sans spécificité: tous les objets représentés ont le même statut ontologique.

L'absence de problématique ontologique conduit à une attitude *normative* de la reconnaissance. Ce qui est reconnu doit toujours l'être de la même manière, sans liberté de choix. L'adéquation d'un donné empirique et d'une représentation ne peut être qu'extrinsèque, et soumise au critère de l'erreur. La dénomination est toujours univoque, et constitue d'ailleurs une condition à l'expérimentation psychophysique ou neuropsychologique de la reconnaissance: le sujet, dans une situation donnée, doit reconnaître selon une norme établie par l'expérimentateur. Tout écart ou différence est considéré négativement et permet de mesurer ainsi un degré de normalité. Dans ce cadre, seule une reconnaissance servile qui délègue l'élaboration de ses propres objectifs peut être testée: permet-elle de révéler l'essence du reconnaître en actes?

Le besoin implicite d'homogénéité ontologique, qui justifierait en partie la pertinence de l'approche représentationnelle de la mémoire, est difficile à accorder avec les résultats et modèles de la recherche contemporaine. Les théories ont en effet une tendance assez récente à la fragmentation [13, 1, 3]: on parle de mémoire sémantique, épisodique, procédurale, déclarative, à long ou à court terme, de travail, etc... Les techniques modernes d'imagerie cérébrale et les données neuropsychologiques accentuent encore le dépeçage d'une fonction mnémonique unifiée qu'il est bien difficile d'appeler simplement mémoire, sans se demander de laquelle il s'agit. Plus les connaissances s'accumulent, moins la délimitation définitive d'une *essence* du mémoriser semble accessible. Comment dans cette diversité de modèles et d'opinions parvenir à dégager une unité ontologique de ce qui est mémorisé? Peut-être la fragmentation de la mémoire permettra-t-elle de proposer une alternative au conformisme normatif qui accompagne le

présupposé représentationnel.

Des remarques précédentes on retiendra surtout que la mémoire en tant qu'objet scientifique est problématique, tant qu'elle est approchée comme une entité stockant des représentations. C'est même la pertinence de son étude en tant qu'objet de recherche qui est à questionner. D'une manière générale, il apparaît sensé de soupçonner une incompatibilité entre cognition en tant que processus de traitement de représentations et une véritable problématique ontologique. Bien vécue, cette incompatibilité manifesterait un respect de la critique kantienne. On conteste cependant que l'approche computationnelle ne peut s'empêcher bien souvent de prétendre à une démarche ontologique, à résonances prométhéennes.

1.3 Récognition

La question ontologique dans un contexte représentationnel est problématique. Elle est l'indice d'un mode de pensée dogmatique et «scolaire» qui institue l'erreur comme critère. Cette critique est développée dans les ouvrages de Gilles Deleuze et plus particulièrement dans *Différence et répétition* [4].

Il n'entre pas dans l'objectif de cet article de présenter les conséquences du renouvellement de la pensée que propose Deleuze. Ce qui sera questionné, c'est l'assimilation de la reconnaissance à une identification, à la lumière de certaines de ses analyses.

J'appellerai par la suite modèle de la *récognition* l'assimilation d'une activité libre de reconnaissance détentrice de ses propres objectifs, à une fonction de réponse ontologiquement neutre, et par là servile et normative. En se désaisissant de la question de la constitution des objets de reconnaissance pour se contenter d'un étiquetage nominal, le modèle de la *récognition* élude ce qui devrait constituer sa problématique première: l'ontologie du reconnu comme tel.

Le projet de Deleuze dans *Différence et répétition* est de proposer une manière de penser alternative à celle qui subordonne la compréhension du divers à une partition de concepts, une pensée qui serait «sans image». Il mène ses critiques à l'intérieur d'une analyse du concept de différence en montrant précisément qu'il est évacué dans la pensée représentative car «médiatisé», soumis «à la quadruple racine de l'identité et de l'opposition, de l'analogie et de

la ressemblance.»³ «On confond l'assignation d'un concept propre de la différence avec l'inscription de la différence dans le concept en général.»⁴

L'origine de cette asservissement de la différence est un besoin d'unité, un refus de la singularité, du surgissement de l'altérité. En posant la différence sous l'identité supposée d'un concept, le divers devient harmonieux. Mais du même coup, on s'empêche de penser véritablement la singularité, qui reste de même nature que l'universel. Nous ne pouvons que retrouver un monde conforme à nos aspirations, à nos représentations.

Deleuze définit huit postulats de la pensée dogmatique qui nous empêche de penser pleinement la différence «en elle-même.»⁵ Leur force est de n'être jamais déclarés comme tels: ils apparaissent plus comme heuristique pratique que comme principe. La critique majeure de Deleuze est alors qu'ils «écrasent la pensée sous une image qui est celle du Même et du Semblable dans la représentation mais qui trahit au plus profond ce que signifie penser, aliénant les deux puissances de la différence et de la répétition, du commencement et du recommencement philosophiques.»⁶ Comment dans cette vision conformiste une créativité de la pensée est-elle possible? À côté de la problématique philosophique qui est plus particulièrement celle de Deleuze, la question est d'importance pour les sciences cognitives. Le rôle qu'y prend une représentation allant de soi n'est-il pas fi-

3. p. 45. Toutes les références de page, sauf mention du contraire, concernent *Différence et répétition*.

4. p. 48.

5. «1° postulat du principe, ou de la *Cogitatio natura universalis* (bonne volonté du penseur, et bonne nature de la pensée); 2° postulat de l'idéal, ou du sens commun (le sens commun comme *concordia facultatum*, et le bon sens comme répartition qui garantit cette concorde); 3° postulat du modèle, ou de la reconnaissance (la reconnaissance conviant toutes les facultés à s'exercer sur un objet supposé le même, et la possibilité d'erreur qui en découle dans la répartition, quand une faculté confond un de ses objets avec un autre objet d'une autre); 4° postulat de l'élément, ou de la représentation (quand la différence est subordonnée aux dimensions complémentaires du Même et du semblable, de l'Analogue et de l'Opposé); 5° postulat du négatif, ou de l'erreur (où l'erreur exprime à la fois tout ce qui peut arriver de mauvais dans la pensée, mais comme le produit de mécanismes externes); 6° postulat de la fonction logique, ou de la proposition (la désignation est prise comme le lieu de la vérité, le sens n'étant que le double neutralisé de la proposition, ou son redoublement indéfini); 7° postulat de la modalité, ou des solutions (les problèmes étant matériellement décalqués sur les propositions, ou bien formellement définis par leur possibilité d'être résolus); 8° postulat de la fin ou du résultat, postulat du savoir (la subordination de l'apprendre au savoir, et de la culture à la méthode).» (p. 216-217)

6. p. 217.

nalement le reflet de nos fantasmes et de nos angoisses?

Le postulat qui nous concerne plus particulièrement dans le cadre de notre problématique, la mathématisation d'une reconnaissance libre, est le 3°. Il contient en germes tous les autres, indiquant par là l'enjeu que constitue une analyse fine de la reconnaissance.

Deleuze assimile la faculté de reconnaissance à une reconnaissance. Il ne tente pas de dégager un mode spécifique de la pensée, dans toute sa puissance créatrice, que l'on pourrait appeler reconnaissance. Mais ses analyses permettent justement de mettre en évidence pourquoi cette attitude semble «naturelle» dans une pratique dogmatique de la pensée, pourquoi reconnaissance et reconnaissance libre sont confondues.

La critique de Deleuze porte sur la conformité sous-tendant la reconnaissance comme modèle de la pensée, sous la forme d'un asservissement au *sens commun*. Chez Descartes, ou même chez Kant, il est besoin de poser une unité du sujet pour la retrouver en reflet dans l'accord des facultés sur un même objet⁷. En même temps que cette identité du sujet, ou de la conscience pour se conformer au style cognitif actuel, il est besoin de postuler une identité de l'objet, que l'on pourra appréhender sous des modalités diverses⁸. C'est une banalité, par exemple, de se rendre compte que Je peux lire un livre, le feuilleter, le poser et le reprendre et de penser que ce livre est toujours le même. La puissance de la critique de Deleuze est de mettre en évidence que c'est ce même schéma qui en définitive investit la pensée, et la contraint à la forme du Même⁹.

Le drame de cette pensée dogmatique, paradoxalement, est sa docilité. Son but ultime est la justification *a posteriori* d'un sens commun, raisonnablement partagé. Nous formons bien une collectivité

7. «Tel est le sens du Cogito comme commencement: il exprime l'unité de toutes les facultés dans le sujet, il exprime donc la possibilité pour toutes les facultés de se rapporter à une forme d'objet qui réfléchit l'identité subjective, il donne un concept philosophique au présupposé du sens commun, il est le sens commun devenu philosophique.» (p. 174)

8. «Un objet est reconnu quand une faculté le vise comme identique à celui d'une autre, ou plutôt quand toutes les facultés ensemble rapportent leur donné et se rapportent elles-mêmes à une forme d'identité de l'objet.» (p. 174)

9. «La pensée est supposée naturellement droite, parce qu'elle n'est pas une faculté comme les autres, mais, rapportée à un sujet, l'unité de toutes les autres facultés qui sont seulement ses modes, et qu'elle oriente sur la forme du Même dans le modèle de la reconnaissance.» (p. 175)

d'individus, et nous vivons dans un même monde d'objets auxquels nous donnons des noms et que nous qualifions souvent empiriquement de la même manière. C'est précisément de cette évidence que proviennent les obstacles à une conception de la reconnaissance comme faculté libre.

C'est la philosophie surtout qui est visée dans les propos de Deleuze, en mettant en évidence les filiations conceptuelles qui ont conduit à renforcer la subordination de la différence au concept. Mais les critiques qui sont données peuvent s'étendre aux modes d'analyse scientifique des fonctions cognitives, et à leur modélisation. Si l'ambition des sciences cognitives est une élucidation scientifique de la pensée dans son essence, il importe alors de se méfier du rôle prépondérant que prend la représentation dans les théories.

L'approche informationnelle est le corrélat scientifique du mode de pensée dogmatique. Elle met en pratique explicitement ce conformisme dénoncé par Deleuze. La citation énoncée au dessus constitue un exemple frappant de cette soumission à un sens commun représentationnel, et ce d'autant plus qu'il ne figure pas dans un traité d'ingénierie mais bien dans un ouvrage de psychologie, dont l'ambition est de présenter un aperçu général d'une science de la pensée.

Il y aurait une justification possible de la propagation de la pensée dogmatique à l'entreprise scientifique: ce serait d'admettre une obligation essentielle de la science à se subordonner à la représentation, selon ses quatre éléments: «identité dans le concept, opposition dans la détermination du concept, analogie dans le jugement, ressemblance dans l'objet du concept.»¹⁰ Faire de la science consisterait à suivre ces principes à la lettre, comme méthode. Tout énoncé qui ne serait pas produit conformément à eux ne serait que divagation fantasmatique ou, au mieux, métaphysique.

L'histoire de la créativité scientifique montre bien cependant qu'elle ne se laisse heureusement pas enfermer dans un système de déterminations clos.

2 Intégrer le divers

2.1 Altérité circonscrite

Le paradigme standard de la cognition vu comme processus de traitement de l'information impose à la reconnaissance d'être réduite à un appariement passif entre un donné sensible et un modèle. Nous avons vu que cette subordination à la représentation générerait des difficultés épistémologiques et ontologiques difficilement surmontables.

Pour pouvoir redécouvrir la faculté de reconnaissance dans toute sa spécificité, il apparaît donc nécessaire de s'affranchir de la tutelle d'un mode de pensée sécurisant, mais trop rigide et dogmatique. Le recours à la «méthode» phénoménologique est un chemin vers cette libération.

Une tentative succincte de description de l'essence de la reconnaissance, en suivant les schémas conceptuels proposés par Husserl [7, 8], a été donnée ailleurs [6]. Il était proposé que ce qui constituait un moment noématique du reconnu, non réel, consistait en l'affirmation d'un *parmi d'autres*. Le reconnu comme tel est un complexe stabilisé hylé/noèse/noème spécifique à un mode d'accession à une objectivité que l'on appelle reconnaissance. Le rôle des synthèses noétiques est alors de façonner le pôle intentionnel par des actes que l'on peut qualifier de *distinguants*.

Le reconnu contient ainsi en lui-même une tension d'aliénation virtuelle. Mais l'autre qu'il aurait pu être, la sphère des non distingués, n'est pas quelconque: dans la reconnaissance, il y a affirmation d'un plan, d'une section, d'une région où situer le reconnu. On peut par exemple reconnaître un marteau en tant qu'arme ou en tant qu'outil. Cet «en tant que» indique que l'on place l'indétermination dans une certaine zone de virtualités, *circonscrite*. Le point crucial est de noter que cette tension maîtrisée à être autre chose est un constituant *essentiel* du noème de la reconnaissance, qu'il est un composant de son essence.

On voit ainsi se dessiner un concept s'inscrivant dans une approche phénoménologique indiquant que la reconnaissance ne peut se laisser cerner comme une jonction passive entre une représentation et un contenu. Elle contient en son sein un principe d'indétermination actif qui participe de son essence. Elle n'est pas mise en valeur du même dans le divers, puisque ce dernier doit lui être intégré pour que soit

10. p. 180.

rendue manifeste le reconnu. L'identité est produite après coup, non pas par l'expérience du même mais par la circonscription affirmative d'une altérité au sein de laquelle le reconnu se distingue.

Nous verrons plus loin une présentation plus affinée de cette approche d'une diversité intégrée dans une définition «noético/noématique» du contexte.

2.2 Multiplicité virtuelle

Cette premier pas conceptuel vers la possibilité d'intégration du divers peut résonner également avec certaines des analyses de Gilles Deleuze.

Dans une problématique métaphysique, *Différence et répétition* est une proposition de renouvellement de la relation dialectique traditionnelle entre le Même et l'Autre. Pour mener à bien cette entreprise, nous l'avons vu, il a fallu d'abord libérer la pensée de son asservissement à la représentation. Alors seulement l'Idée, comme véritable support créatif de la pensée, comme puissance subreprésentative, peut devenir objet de réflexion.

Cet article n'a pas pour objectif de présenter en quelques lignes les concepts clés de l'œuvre de Deleuze¹¹, mais de mettre en relief ceux qui semblent pouvoir jouer un rôle clé dans un projet de mathématisation de la reconnaissance. J'insisterai principalement sur le concept de *multiplicité virtuelle* et sur l'idée d'une *implication réciproque* entre différence et répétition dans l'actualisation du virtuel.

Deleuze considère l'Idée comme une multiplicité, non comme une «combinaison de multiple et d'un, mais au contraire [comme] une organisation propre au multiple en tant que tel, qui n'a nullement besoin de l'unité pour former un système.»¹² La multiplicité doit être pensée comme *structure*: un ensemble de places virtuelles, sans signification conceptuelle ni forme, en dépendances réciproques et munies d'un processus d'actualisation spatio-temporel à la fois de ses éléments et des rapports différentiels les déterminant¹³.

Deleuze insiste bien sur la confusion à éviter entre les couples conceptuels possible/réel et virtuel/actuel : «le virtuel possède une pleine réalité, en tant que virtuel, [...] comme si l'objet avait une de

ses parties dans le virtuel, et y plongeait comme dans une dimension objective.»¹⁴ On se rapproche ainsi de l'approche phénoménologique proposée au dessus, la multiplicité virtuelle devenant un constituant à part entière de l'objet comme moment noématique.

Mais une structure, simple multiplicité, est inerte: elle est certes différenciée car elle porte en elle une tension relationnelle entre ses éléments, mais ce sont des éléments fantômes, des spectres sans chair. C'est une différenciation qui leur donnera corps, c'est à dire un processus d'actualisation au pouvoir distinguant.

On peut préciser un peu mieux la nature de cette différenciation. C'est ici qu'intervient la répétition, et qu'il faut à tout prix abandonner l'idée de la reproduction du même, ou de la mise en évidence de l'identique, comme principe de reconnaissance. C'est la répétition qui met en valeur le virtuel en le différenciant, activement¹⁵.

«L'actualisation, la différenciation, est toujours une véritable création. Elle ne se fait pas par limitation d'une possibilité préexistante.»¹⁶ La répétition n'est jamais la même, parce que la structure différentielle qui la supporte est affectée par l'actualisation. Autrement dit, la répétition s'auto-modifie par la médiation de la multiplicité virtuelle, en une réciprocity sans fin. Ainsi, c'est un renversement de la subordination de la différence au monde de l'identité et de la ressemblance qui est réalisé. Le même devient un produit positif de la différenciation, il est une affirmation.

La multiplicité virtuelle et le processus de différenciation introduits par Deleuze et le concept d'altérité circonscrite défini comme moment noématique du reconnu évoqué au dessus participent de la même ambition de ne pas considérer le différent et le divers comme des catégories négatives. *Différence et répétition* ne parle pas cependant, ou du moins explicitement, de cette notion de limitation active contenue dans la circonscription de l'altérité ou de la multiplicité qui me semble essentielle pour la reconnaissance et la cognition en général. La différence puise sa force dans la délimitation affirmative d'un milieu où elle ne se dilue pas dans l'infini des dimensions.

11. On pourra se référer par exemple à [9, 15, 10].

12. p. 236.

13. Pour une présentation du concept de structure on pourra se référer à [5], repris dans [11].

14. p. 269.

15. «La répétition est la puissance de la différence et de la différenciation : soit qu'elle condense les singularités, soit qu'elle précipite ou ralentisse les temps, soit qu'elle varie les espaces.» (p. 284)

16. p. 273.

3 Application: définition du contexte

Cette section se propose de fournir une définition du contexte. On essaiera de montrer comment les concepts présentés plus haut permettent de rendre accessible à la formalisation ce constituant essentiel d'une reconnaissance libre.

Le contexte est un élément déterminant dans les théories des actions cognitives génératrices de sens, et plus particulièrement dans les théories de la signification. La faculté de reconnaissance est bien sûr intimement liée à ces questions. Le contexte est trop souvent une notion fourre-tout, regroupant tout ce qu'une théorie n'est pas capable d'identifier comme invariant ou autonome, ou ce qu'elle décide comme ne faisant pas partie de l'ensemble des objets susceptibles d'être intentionnellement visés. Cependant, le contexte n'est pas une entité nébuleuse qui vient troubler la marche normale des actes cognitifs, il renvoie directement à leur essence.

3.1 Le contexte dans l'*a priori* représentationnel

Afin de révéler ce qu'induit l'*a priori* représentationnel dans la conception du contexte, nous tenterons de montrer formellement comment il est introduit classiquement, même si cela est rarement réalisé explicitement dans les théories dans lesquelles il est censé jouer un rôle déterminant. On le considère souvent selon deux points de vue: comme paramètre ou comme variable cachée.

Dans le premier cas, il participe en tant que variable auxiliaire aux fonctions de significations. Son rôle serait alors double: il autoriserait une économie de signes en permettant à plusieurs signifiés d'être associés au même signifiant, mais aussi il réduirait la complexité du mécanisme générateur de sens en limitant l'ensemble des significations possibles. L'action génératrice de sens peut être comprise comme une fonction d'interprétation I , au sens mathématique, selon le schéma:

$$\text{Signifié} = I[\text{Signifiant}, \text{Contexte}].$$

Le contexte participerait à l'interprétation sémantique des signes comme descripteur des éléments *pertinents* non verbaux ou non lexicaux qui permettent de dégager le sens des expressions "en

situation". Mais que gagne-t-on réellement du point de vue théorique à introduire une variable supplémentaire non directement signifiante appelée Contexte?

D'un point de vue strictement formel, la distinction entre Signifiant et Contexte est arbitraire et devient alors source de conflits théoriques sans fin. Par exemple, dans la définition du signe linguistique, doit-on décider d'inclure l'intonation, les silences, les mimiques, etc... comme partie intégrante du signe ou plutôt comme éléments du contexte. Mais surtout — ce que cache souvent une formulation fonctionnelle mal définie et mal comprise — elle demande que soit posés *a priori* les espaces de Contextes et de Signifiants sur lesquels s'appuie la fonction I d'interprétation. On ne peut se satisfaire d'un espace de "tous les contextes", qui indiquerait que tout, avec le flou conceptuel qui accompagne souvent ce terme, influencerait potentiellement l'interprétation. D'un autre côté, se restreindre à un certain nombre d'éléments *pertinents*, c'est-à-dire susceptible d'influer directement sur l'interprétation, n'est pas toujours satisfaisant car ils ont rarement un pouvoir signifiant homogène et indépendant du signe linguistique qu'ils modulent.

La description explicite d'un contexte signifiant comme variable modulant une fonction d'interprétation est donc problématique car l'espace de variation de ce contexte, élément à part entière d'une définition de la fonction d'interprétation, est difficilement déterminable avec précision. Mais une autre notion de contexte, plus générale, peut également être définie dans le cadre d'une théorie de l'observateur. Ici encore, une description formelle nous aidera à préciser nos critiques.

La plupart des théories des observateurs en interaction avec un environnement se divisent en trois éléments: une description des influences d'un agent sur un monde externe, une description des influences de l'environnement externe sur l'observateur et la définition d'une loi de contrôle régissant les actions de l'agent en fonction de ses observations, la vie de l'agent.

Formellement, un modèle élémentaire d'observateur en relation avec un environnement peut être défini de la manière suivante:

1. L'environnement w se décompose en partie observable o et partie cachée c :
 $w = [o, c]$.

2. L'agent, d'état interne x , est capable d'émettre des actions a .
3. L'évolution de l'environnement externe est régie par une équation dynamique différentielle de la forme: $\dot{w} = F(w, a)$.
4. L'évolution de l'agent suit lui aussi une équation différentielle: $\dot{x} = G(w, x, a)$.
5. L'agent émet des actions selon une loi de contrôle μ : $a = \mu(o, x)$.

Ce schéma formel définit un contexte comme un état caché c , jouant le rôle d'un arrière-fond non observable mais influençant globalement à la fois l'environnement et le comportement de l'agent. Dans ce type de modèle issu de la théorie du contrôle, il est d'usage de chercher à estimer les caractéristiques des relations fonctionnelles F et G à partir d'observations.

La définition du contexte comme variable cachée de l'environnement est formellement mieux justifiée que celle qui le considère comme paramètre modulant une fonction interprétative: le contexte a ici un rôle spécifique et intervient structurellement dans la théorie. Cependant, on peut lui reprocher les mêmes difficultés générées par la nécessité de description *a priori* d'un espace de variations indéterminable.

L'inscription du contexte dans une théorie plus générale de l'observateur est attirante, mais produit en retour un besoin de justification des présupposés philosophiques et métaphysiques si l'on veut être rigoureux. La dichotomie agent/environnement impose l'acceptation d'une vision de sens commun «naturelle» qui fonde l'ontologie de cette division sur un dualisme réaliste naïf. L'histoire de la philosophie montre bien les difficultés posées par ce type de présupposés. Mais plus «scientifiquement», l'existence d'une frontière entre extérieur et intérieur de l'agent, d'une interface, appelle la question de l'emplacement du contexte: est-il dans l'agent ou dans l'environnement? Est-il une représentation interne ou une entité réelle? Ces questions produisent en définitive plus de difficultés que de clarifications sur la place du contexte dans l'activité cognitive, le faisant osciller tantôt comme composant invisible d'un pôle réel, tantôt comme contrainte interne. Il semble nécessaire de revenir sur ces présupposés dualistes pour révéler le rôle *essentiel* du contexte.

3.2 Interprétation phénoménologique du contexte

Dans une approche représentationnelle, le contexte a le statut d'une entité ou d'un état partiel. Il est courant d'entendre des expressions comme «dans ce contexte, ceci est vrai», laissant supposer qu'il peut être dénommé explicitement et qu'il intervient comme condition de vérité. Ce qui est implicitement supposé dans cette vision est sa réalité objective.

Un premier point important est de se convaincre de la nature *théorique* du contexte: il n'est ni une entité, ni un état, ni un objet «fondé», dans un sens phénoménologique, sur une intuition perceptive. C'est là que réside principalement la difficulté de sa définition, car elle ne peut s'appuyer sans médiation sur une réalité empirique.

Ce que je propose, puisque la réalité objective du contexte est difficile à révéler, est de le placer comme composant à part entière d'une description phénoménologique de la reconnaissance, vue comme dénomination. Il s'agit donc de le situer au sein des structures «noético-noématiques» qui permettent de constituer le reconnu comme tel¹⁷.

S'il n'est pas proprement réel, le contexte ne peut donc que participer au noème, comme un de ses moments. Il n'est pas un constituant des synthèses qui façonnent l'objet intentionnel — les noèses — il n'est pas un matériau sensoriel — un donné hylétique — il est un composant essentiel de l'objet intentionnel lui-même. Le noème de la dénomination, le dénommé vécu comme tel, est composé: il est fondé bien sûr sur une dimension perceptive, mais il apporte sa contribution propre. En particulier, il permet de reconnaître un objet «en tant que», indice du rôle essentiel du contexte à la constitution de l'objet.

Le contexte de dénomination sera défini comme *une double spécification noématique de multiplicités noétiques*. Précisons ce que cette définition cherche à proposer.

La première spécification noématique est perceptive et donne les paramètres qui déterminent les synthèses noétiques morphologiques, celles qui constituent l'objet d'une perception pure sans intention signifiante. Un exemple typique d'une telle spécification intentionnelle de la perception est la no-

17. Par manque de place, je suppose que le lecteur possède une idée générale des concepts de hylè, noèse et noème tels qu'ils sont proposés par Husserl, particulièrement dans les *Ideen*[7].

tion d'échelle d'analyse qui permet de révéler des formes en fonction de la finesse des détails fusionnés dans les mécanismes de groupement. La position d'une échelle circonscrit virtuellement l'ensemble des formes constituées, et définit ainsi un plan de variations admissibles, de multiplicités.

L'autre multiplicité, « cognitive », définit le plan de variations du nom de l'objet reconnu: c'est ici que le contexte est chargé de lever les ambiguïtés et l'indétermination de la dénomination. C'est en circonscrivant par l'intention la classe, famille etc... de noms susceptibles d'être donnés à l'objet que la reconnaissance pourra produire une présence pleine de l'objet reconnu. Mais comme dans le cas de la perception pure, la multiplicité cognitive ou dénominative est une spécification des synthèses noétiques qui mettent en forme un matériau brut. Si un rôle phénoménologique peut être donné à une mémoire de stockage — et nous avons vu la difficulté d'une telle interprétation — ce serait justement comme hylè nominale.

La présence d'une double multiplicité noématique n'implique pas cependant un partage des tâches inflexible entre perception et dénomination. L'identité du noème¹⁸ fait que les synthèses noétiques correspondantes doivent « collaborer » pour assurer une adéquation, une stabilité, de l'objectivité. Cette condition eidétique permet en fait de rendre compte de l'influence d'une intention dénominative sur la constitution de la forme: l'objet reconnu peut être influencé par la gamme de noms susceptibles d'être donnés.

Il s'agit de ne pas confondre la position d'un plan de virtualités indéterminées, manifesté par le « en tant que », et la variété de modes d'objectivité qui est révélée par la différence eidétique des noèmes dans l'analyse phénoménologique. Le contexte est un moment *du* noème de la reconnaissance prise ici comme dénomination, de lui seul, et n'a de sens que dans ce mode d'intentionnalité objective.

Une fonction interprétative noématique est indirectement définie par les synthèses noétiques car l'unité du noème est composée: il contient les spécifications de la multiplicité noétique, c'est-à-dire le contexte, mais également le nom donné et l'objet de perception pure. C'est dans la noème que la relation

18. « Le plan noématique serait le lieu des unités, le plan noétique, le lieu des multiplicités "constituantes". La conscience qui unit "fonctionnellement" le multiple, et qui en même temps constitue l'unité, ne présente en fait *jamais* d'identité, alors que l'identité de "l'objet" est donnée dans le corrélat noématique. » [7], p. 343.

entre le nom et l'objet perçu se rejoignent et tissent un lien d'association abstrait qui n'existe pas à l'état de vécu.

Il peut se produire cependant que le noème lui-même ne réunisse pas dans son intention les conditions de plénitude d'une adéquation noétique qui garantisse son unité, empêchant ainsi la constitution de la fonction interprétative: cette instabilité noétique ne peut se résoudre que par un réaménagement de l'intention de reconnaissance en modifiant le contexte selon ses deux composantes perceptives et/ou dénominatives. Ainsi, dans cette interprétation noético/noématique de la dénomination, la reconnaissance ne se résout pas à un appariement passif entre un donné empirique et une représentation mais demande la mise sous tension intentionnelle d'une multiplicité circonscrite de synthèses noétiques. Le contexte apparaît ici comme un élément *essentiel* de la reconnaissance mais à un niveau non réel, noématique.

4 Conclusion

Toute une tradition philosophique occidentale fait reposer l'être sur une impassibilité de l'Un ou du Même. Le multiple et le divers ne seraient que des perturbations désincarnées qui viennent brouiller l'idéal. La cognition n'est jamais autant satisfaite que quand elle isole l'identique, reflet rassurant d'une identité qu'elle pense première, la sienne.

Paradoxalement peut-être, l'objectif de cet article était « pratique » et aurait sans doute voulu le montrer de manière plus évidente. Il visait à critiquer l'approche représentationnelle de sens commun qui considère la reconnaissance comme un appariement d'un donné sensoriel et d'une représentation stockée en mémoire, modèle qui en pratique est source de difficultés artéfactuelles nombreuses. Le poids de la représentation et sa place ubiquitaire dans les théories des cognitions naturelles et artificielles est tel qu'il a fallu essayer de localiser l'origine de sa position dominante selon une approche qui flirtait avec des considérations métaphysiques. L'ambition déclarée de mathématisation de la reconnaissance s'est en fin de compte résumée à la mise en évidence de quelques composants essentiels qui demanderaient à être mieux formalisés.

Le point de départ de cet article était une problématique retournée: comment la cognition est-elle ca-

pable de penser le divers ? Pour élucider cette question, il s'agissait d'opérer une «révolution copernicienne» en poussant l'identité comme principe second, comme principe *devenu*. Le différent devenait alors source de l'identité, le devenir celle de l'être¹⁹.

La cognition libre ne procède pas par identification, mais par appropriation du divers, par affirmation intégrée de la différence. «Elle opère pratiquement une sélection des différences d'après leur capacité de produire, c'est-à-dire de revenir ou de supporter l'épreuve de l'éternel retour.»²⁰ Un programme de recherche fondé sur la représentation et ses avatars ne peut que passer à côté de l'essence de ce qu'est véritablement penser: il lui manquera toujours la mise en valeur d'une créativité authentique.

Pour mener à bien ce renversement de vision, de nouveaux concepts doivent être imaginés. Celui d'altérité circonscrite, comme celui de multiplicité virtuelle, constituait une tentative pour libérer les sciences cognitives du carcan de la représentation. Il s'agissait alors d'en tester la productivité. La définition du contexte dans un cadre phénoménologique visant à en révéler le rôle essentiel constituait un premier pas élémentaire dans cette direction. A poursuivre et à amplifier, bien entendu.

Références

- [1] Baddeley (A.D.). – *Human Memory: Theory and practice*. – London, Lawrence Erlbaum, 1990. (Traduction française : *La Mémoire Humaine: Théorie et pratique*, Presses Universitaires de Grenoble, Grenoble, 1993).
- [2] Boucart (M.). – *La Reconnaissance d'Objets*. – Presses Universitaires de Grenoble, 1996.
- [3] Bruyer (R.). – Les agnosies visuelles. In Seron et Jeannerod [12], chap. 7/partie 2, pp. 179–204.
- [4] Deleuze (G.). – *Différence et Répétition*. – Paris, Presses Universitaires de France, 1968.
- [5] Deleuze (G.). – A quoi reconnaît-on le structuralisme? In : *La philosophie*, éd. par Châtelet (F.). – Paris, Hachette, 1973.
- [6] Herbin (S.). – Etudier la reconnaissance — modèle et perspective philosophique. In : *Actes*

du Deuxième Colloque Jeunes Chercheurs en Sciences Cognitives.

- [7] Husserl (E.). – *Idées Directrices pour une Phénoménologie*. – Paris, Gallimard, 1950. Traduction de P. Ricoeur.
- [8] Husserl (E.). – *Recherches logiques 3, Eléments d'une élucidation phénoménologique de la connaissance (Recherche VI)*. – Paris, Presses Universitaires de France, 1963. Traduction de H. Elie, L. Kelkel et R. Schérer.
- [9] Martin (J.-C.). – *Variations : la Philosophie de Gilles Deleuze*. – Paris, Payot, 1993.
- [10] Mengue (P.). – *Gilles Deleuze ou Le système du multiple*. – Paris, Kime, 1994.
- [11] Petitot-Cocorda (J.). – *Morphogénèse du Sens—I : pour un schématisme de la structure*. – Paris, Presses Universitaires de France, 1985.
- [12] Seron (X.) et Jeannerod (M.) (édité par). – *Neuropsychologie Humaine*. – Liège, Mardaga, 1994.
- [13] Tiberghien (G.). – Psychologie cognitive de la mémoire humaine. In Seron et Jeannerod [12], chap. 9/partie 1, pp. 255–281.
- [14] Treisman (A.). – L'attention, les traits et la perception des objets. In : *Introduction aux sciences cognitives*, éd. par Andler (D.), pp. 153–191. – Paris, Gallimard, 1992.
- [15] Zourabichvili (F.). – *Deleuze : une Philosophie de l'Événement*. – Paris, Presses Universitaires de France, 1994, *Philosophies*.

19. [4], p. 59.

20. p. 60.

L'identité dans un modèle de la dépendance nominale

Pierre Beust, Anne Nicolle

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

Mail : beust@info.unicaen.fr

Tel : 02 31 56 62 17

Fax : 02 31 56 58 14

Résumé: Cet article propose un modèle d'intelligence artificielle pour l'identification de phénomènes de dépendance contextuelle nominale dans des textes ou des dialogues. Dans ce modèle, les choses du monde auxquelles réfère l'énoncé sont représentées par une structure constituée de plusieurs facettes de description. Deux composantes de l'identité d'un objet: une identité individuelle et une identité sortale sont prises en compte dans la représentation des lexèmes. Les dépendances nominales provoquent un enrichissement sémantique des structures évoquées par les mots. Elles précisent de ce fait l'identité des référents au moyen de traits pertinents dans le contexte. Au delà de l'étude des phénomènes de dépendance, c'est un modèle opératoire de la signification qui est visé.

Mots clés: Identité sortale, Identité individuelle, Dépendance Contextuelle, Signification.

Introduction

Cet article présente un modèle pour une automatisé de l'identification de phénomènes de dépendance entre des termes linguistiques dans leur contexte d'énonciation. Quand un sujet produit un énoncé, il utilise des mots et des expressions dans un contexte précis. Les mots et les relations entre les mots participent à la construction de la signification de l'énoncé. Dans la première partie on présentera ces relations de façon générale pour ne retenir que celles qui sont pertinentes au regard d'une représentation de l'identité des objets évoqués. Après la présentation du cadre d'étude de l'identité dans la langue, on examinera, dans la troisième partie, des cas de dépendance nominale directement dus à une prise en compte de l'identité. En dernière partie, on présentera le modèle correspondant à l'identification de la dépendance nominale.

1. Objet de l'étude.

Pour réaliser un système automatique d'identification de relations de dépendance contextuelle, nous cherchons à rendre compte de la nature pragmatique des relations linguistiques faisant apparaître une forme de dépendance. Dans un bon nombre de cas, comme le montre la troisième partie de cet article, la pragmatique va poser la question de l'identité qui amènera un éclairage sur la question de la dépendance.

Les études linguistiques des relations de dépendance sont nombreuses et leur mise en évidence constitue notamment une bonne part de la syntaxe classique. Ainsi, dans les formalismes de grammaire générative, définir une catégorie Phrase à partir de syntagmes nominaux et verbaux, par exemple, revient à établir une relation de dépendance entre ces types de syntagmes. Au niveau syntaxique, on trouve des formes de dépendance liées à un contexte beaucoup plus restreint que celui de la phrase. Par exemple, dans le contexte nominal, un adjectif (non nominalisé) ne peut se passer d'un nom commun qu'il qualifie, de même qu'un nom commun impose la présence d'un déterminant. Dans le contexte verbal, un verbe transitif induit la présence nécessaire d'un objet. Au niveau sémantique, on peut également mettre en évidence des contextes où des mots entretiennent des dépendances. Celles-ci précisent leurs significations. Par exemple, dans le contexte de l'énoncé, la relation entre un verbe et son objet peut induire une propagation de certaines propriétés sémantiques :

(1) *J'ai fini de lire Saussure.*

Dans l'énoncé (1), *Saussure* est interprété comme un livre de Saussure et non comme la personne de Saussure, par exemple. Ceci provient de la relation de dépendance avec le verbe *lire*. À travers ce genre de dépendance, l'interprétant tente de résoudre des ambiguïtés ou des imprécisions sémantiques. Plus largement, dans un contexte extraphrastique, un anaphorique se doit d'être sémantiquement saturé par une

relation qui le lie à un antécédent. C'est le cas de pronoms qui réfèrent à des termes apparus dans un autre énoncé.

définition : On entend par relation de dépendance contextuelle, le fait qu'un morphème ou un syntagme exige la présence d'un ou de plusieurs autres morphèmes ou syntagmes dans son contexte pour déterminer sa signification.

La dépendance contextuelle est tantôt principalement syntaxique tantôt principalement sémantique en fonction du contexte où elle s'établit. Par exemple, un accord dans un groupe nominal est une dépendance syntaxique et constitue une ou plusieurs isosémies (récurrence d'un trait lexical¹, un sème primitif dans la typologie de Rastier) alors qu'une référence anaphorique est une relation beaucoup plus sémantique dont la trace est une ou plusieurs isotopies (récurrence d'un trait sémantique). Ce concept d'isotopie, créé par Greimas ([Greimas 66]), a pour objectif de rendre compte de l'homogénéité du discours par la récurrence de certains traits sémantiques (sèmes). F. Rastier ([Rastier 87]), quant à lui, propose d'étendre ce concept à tous les traits sémantiques et notre modèle suit cette perspective. Une hypothèse que nous cherchons à valider est que cette différence entre isosémie et isotopie ne doit pas être marquée au niveau du mode d'établissement de la relation mais plutôt par la nature des traits qui sont véhiculés. Il nous faut donc distinguer, pour rendre compte de phénomènes d'isotopie aussi bien que de phénomènes d'isosémie, les traits de nature sémantique (/animal/, /animé/, etc.) des traits de nature lexicale (/masculin/, /singulier/, etc.).

La nature des entités entre lesquelles peuvent s'établir des relations de dépendance contextuelle permettent de distinguer plusieurs sortes de relations : les rapports entre syntagmes verbaux dénotant des localisations relatives de certains procès dans le temps constituent des relations de dépendance verbales et relèvent de la sémantique de la temporalité ([Gosselin 96]) et les relations de dépendance entre syntagmes nominaux prennent part à la sémantique des expressions référentielles. C'est dans l'établissement de ces relations de dépendance nominale qu'on trouve des manifestations de l'identité.

Parmi les relations linguistiques de dépendance nominale, la co-référence et l'anaphore sont les plus étudiées, rappelons leur définition.

Co-référence : Il y a co-référence entre deux entités linguistiques A et B quand l'objet référent de A est le même que celui de B.

exemple :

(2) *Je me lave.*

(2') *Il se lave.*

Dans (2), *Je* et *me* sont co-référentiels mais il n'y a pas anaphore. Dans (2'), *Il* et *se* sont co-référentiels et en plus il y a anaphore parce que, à la différence de *me*, le pronom *se* n'est pas sémantiquement saturé d'où son besoin d'entretenir une relation d'anaphore.

Anaphore : Il y a relation d'anaphore entre deux unités A et B quand l'interprétation de B dépend crucialement de l'existence de A, au point qu'on peut dire que l'unité B n'est interprétable que dans la mesure où elle reprend, entièrement ou partiellement, A. ([Milner 82], p. 18)

exemples :

(3) *La veste ne m'a pas coûté cher, elle était en solde.*

(4) *Le mardi était pluvieux, le lendemain était ensoleillé.*

Dans l'exemple (3) il y a anaphore et co-référence entre *La veste* et *elle*. C'est une anaphore pronominale. Dans l'exemple (4) il y a anaphore sans co-référence entre *Le mardi* et *le lendemain*. C'est une anaphore associative. Dans l'exemple (3), la reprise est totale alors que dans l'exemple (4), elle est partielle.

L'objet de notre étude sera la place de l'identité dans la dépendance nominale. En particulier dans les co-références et les anaphores, relations qui sont constitutives de chaînes de référence.

¹ Ce principe d'isosémie se retrouve chez [Coursil 93] dans les notions d'anaphore du nombre, du genre ou encore de la personne verbale.

2. Identité individuelle, identité sortale.

L'identité n'est pas une notion à proprement parler d'ordre linguistique, elle n'est pas marquée comme un temps morphologique, par exemple. Elle relève bien du subtil rapport entre les formes linguistiques proférées et un monde où le changement est incessant. Elle n'est pas dans le matériel linguistique mais bien dans les choses du monde et apparaît dans la façon dont nous les appréhendons. Le rapport entre l'activité langagière et la prise en compte de l'identité place cette étude de la dépendance nominale dans le cadre général de la pragmatique, dont l'objet est l'usage du langage, en opposition à la linguistique qui étudie le système de la langue.

La langue nous permet de parler des choses ou même d'agir sur les choses par l'utilisation de signes car le monde est hors de la langue et seule la représentation qu'on s'en fait est transmissible par elle :

"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." ([Pierce 78], p. 123)

Le rapport que crée un interlocuteur entre les choses et les signes constitue la référence, cette position sur la nature de la référence est celle du courant pragmatique et interactionniste. Ainsi, la référence est la relation qui unit une expression de la langue (dite en général *expression référentielle*) en emploi dans un énoncé et l'objet dans le monde que cette expression désigne ([Moeschler & al. 94]). Elle est ce grâce à quoi le langage permet de dire et d'entendre des choses sur le monde. La référence est une relation très générale qui dans l'effectivité du langage se réalise sous plusieurs aspects. Les formes linguistiques peuvent faire référence à des choses naturelles, à des artefacts, à des "choses" vivantes, à des moments dans le temps, à des idées, à des états mentaux ou même à des formes langagières. Des tentatives de référence sont construites par le sujet qui écoute et, dans un dialogue, la référence est le fruit d'une collaboration entre les différents interlocuteurs. C'est en construisant ensemble la référence qu'ils vont se mettre d'accord sur les "objets"² du monde qu'ils évoquent.

C'est par cette référence que les actants de la communication prennent acte de l'objet visé et pas d'un autre. Le sujet intègre alors une représentation de l'identité de l'objet, une identité qui n'est propre qu'à cet objet. Nous appellerons cette identité que nous appréhendons via la référence, **identité individuelle**.

La recherche de co-référenciation dans une communication langagière demande de la part des actants une coopération. Chacun propose par ses réponses physiques ou langagières une interprétation, soumise à l'évaluation de l'autre et ceci jusqu'à obtenir un consensus. L'interprétation des expressions référentielles, s'appuyant sur la sémantique nominale, est ici au centre du problème pour obtenir une expression qui satisfait les différents acteurs de l'interaction. Une étude sémantique des groupes nominaux dégage les propriétés qui, dans le contexte et pour les actants en cause, ont été pertinentes et suffisantes pour rendre possible cette co-référence. Ces propriétés relèvent de la perception de l'objet par les actants, perception qui sous-tend une catégorisation. La sémantique des expressions référentielles s'appuie sur une identité de l'objet formée de propriétés permettant la catégorisation. Nous appellerons cette identité de l'objet, son **identité sortale**. Elle est fondamentalement différente de l'identité individuelle qui lui préexiste de toute façon et qui est liée aux choses et pas aux signes.

Il apparaît donc que l'identité, telle que la langue permet de la considérer s'articule en deux formes. L'une provient la prise en compte de l'identité individuelle des choses du monde. L'autre provient de la sémantique issue des processus de catégorisation dont la mise en commun dans le dialogue crée l'identité sortale.

On considère que l'identité sortale de l'objet inclut les traits qui représentent les attributs de sa classe ainsi que les valeurs de ces attributs qui lui sont propres. Cette vision de l'identité sortale n'est pas celle que l'on peut retrouver notamment chez Charolles ([Charolles 94]) qui considère une identité sortale comprenant les attributs de la classe et une identité regroupant les valeurs de ces attributs qu'il appelle identité qualitative. Cette distinction se justifie dans le cadre d'une étude linguistique mais ne convient pas à une modélisation en intelligence artificielle car sinon l'identité sortale serait confondue avec la notion de classe de l'objet qui correspond à une description en intention de celui-ci. Nous ne retiendrons donc pas dans notre problématique cette notion d'identité qualitative car il convient de faire la différence entre un type (hors contexte) et un domaine conçu comme une restriction (en contexte) du type par des qualificatifs. La

²"objets" est ici utilisé dans un sens constructiviste comme dans [Mugur-Schächter 89].

sémantique définit le domaine et la référence définit un individu dans le domaine. Notre vision de l'identité sortale est, en quelque sorte, ce pourquoi un objet tombe sous une catégorie et non pas uniquement une description formelle de celle-ci (la classe). L'identité individuelle correspond dans cette perspective à ce pourquoi un objet d'une catégorie n'est pas un autre objet de cette catégorie, c'est à dire les critères qui le différentient des autres objets de sa catégorie ne pouvant être propres qu'à lui. Le fait, par exemple, qu'une chaise soit bleue (critère d'identité qualitative pour Charolles), même si c'est la seule chaise bleue dont on dispose n'est pas un critère à retenir pour caractériser l'identité individuelle de cette chaise car une autre chaise bleue serait possible. Par contre, le fait que cette chaise en question est celle sur laquelle Albert était assis il y a cinq minutes est un critère suffisant pour la différentier de n'importe quelle autre et c'est donc bien un critère d'identité individuelle (c'est un trait de son ontogenèse).

3. L'identité dans les différentes formes de dépendance nominale.

Cette partie décrit comment l'identité, individuelle ou sortale, apparaît dans le système de la langue à travers les relations de dépendance nominale. Après avoir rappelé le rôle du syntagme nominal, on examine l'identité dans les co-références, dans les anaphores associatives, nominales ou pronominales et enfin dans les cas de référents évolutifs.

Les relations de dépendance nominales influent sur la constitution de chaînes de référence. Ces chaînes sont construites à partir de syntagmes nominaux et elles construisent ou maintiennent l'identité des référents. Selon [Kleiber 93], le syntagme nominal qui introduit un référent a deux fonctions :

- il décrit le référent et son contenu descriptif fournit des prédicats sortaux qui permettent de l'identifier,
- il dénomme le référent.

Ces deux fonctions provoquent le dédoublement de la notion d'identité. Dans sa première fonction, le syntagme permet la visée d'un référent comme élément d'une catégorie et ce faisant, il précise son identité sortale. Dans sa seconde fonction, le syntagme dénommant le référent en fait un objet de co-référenciation possible précis. Il ne crée pas l'identité individuelle de l'objet qui lui préexistait mais il la place dans le terrain commun de l'interaction langagière.

Les principales relations de dépendance qui influent sur les syntagmes nominaux sont les co-références et les anaphores. L'une des principales différences entre co-référence et anaphore (dans le sens large) est justement une question d'identité. L'une de ces relations ne peut se passer d'une identité implicite, en l'occurrence la co-référence, l'autre, l'anaphore, ne sous-tend pas une identité et on peut alors justement appeler ces cas des anaphores non co-référentielles (cas de l'anaphore associative, par exemple). La co-référence sans anaphore est une forme particulière où la référence est double mais elle concerne le même objet dont l'identité est figée. C'est donc une relation sémantiquement informative sur l'objet mais qui n'affecte pas particulièrement l'une des deux facettes de son identité. Dans l'exemple suivant les syntagmes nominaux *président de la république* et *ancien maire de Paris* sont co-référentiels et la co-référence est marquée de façon supplémentaire par la copule.

(5) *Le président de la république est l'ancien maire de Paris.*

Si on veut rendre compte d'une trace de l'identité dans la langue, il nous faut la considérer au regard du paradigme discours/récit que l'on trouve en narratologie. Le discours est une forme d'usage de la langue où l'on parle de la nature des choses et pas de leur évolution. Le récit est une forme d'usage de la langue qui marque une différence entre deux situations dans le temps. Ce n'est pas le cas de la co-référence qui est une relation du discours. Le discours est en rapport à un ensemble de connaissances que l'on peut, par exemple, modéliser en terme de logique propositionnelle. Au regard de l'exemple (5) une représentation possible du discours est la suivante :

Président-de-la-république(x) et
Ancien-maire-de-paris(y) et
 $x = y$

Un tel formalisme permet une automatiser des inférences sur les connaissances à partir de faits. Par exemple, on déduit que si *Président-de-la-république(Chirac)* alors *Ancien-maire-de-paris(Chirac)*. La référence est ici supportée par des variables logiques, c'est à dire des entités figées. Le discours impose un accord préalable sur l'identité des choses et après seulement, on discute, on argumente. À l'inverse du discours, l'identité, dans le récit est elle aussi l'objet de la discussion, il n'y a aucun accord préalable sur les choses et c'est pour cette raison qu'il y a nécessité de co-référenciation.

Il apparaît que la distinction discours/récit est un critère primordial dans le choix des relations linguistiques à envisager dans notre étude prenant en compte l'identité. Les relations du discours ne seront pas pertinentes car l'identité y est figée et il convient d'examiner les relations qui structurent le récit car c'est bien dans du changement qu'apparaît ce qui est permanent.

Toutes les formes d'anaphores n'ont pas le même statut en rapport à ce paradigme discours/récit. Les cas d'anaphores associatives, notamment, relèvent plus du discours que du récit.

- (6) *Quand je suis arrivé dans le village, l'église était en travaux.*

Dans l'exemple (6) l'anaphore associative entre *le village* et *l'église* décrit un rapport méréonymique entre les objets référents des syntagmes mais ce n'est qu'une question de prototype et ni l'identité de l'un ni l'identité de l'autre n'est mise en cause par l'anaphore (pas même l'identité sortale, *l'église* n'étant pas un critère ontologique nécessaire à l'objet *village*).

Les anaphores nominales et pronominales n'ont pas ce rapport au prototype et ne sont pas construites, comme l'anaphore associative, autour d'un présupposé atemporel. Ce sont bien des relations du récit et non du discours. L'exemple (7) :

- (7) *Un chasseur s'est blessé ce matin, le jeune homme a été emmené à l'hôpital.*

est une anaphore nominale et on ne peut pas dire qu'il existe une relation sémantique d'ordre prototypique entre *un chasseur* et *un jeune homme*. Le but de la relation n'est plus d'évoquer un objet à travers un autre, il s'agit d'introduire un nouveau substantif pour un même objet. Le nouveau syntagme nominal invoque, comme l'a signalé Kleiber, de nouveaux prédicats sortaux pour identifier l'objet. On a là une évolution, ou plutôt un enrichissement, de l'identité sortale du référent. Les anaphores pronominales :

- (8) *Un chasseur s'est blessé ce matin, il a été emmené à l'hôpital.*

sont aussi des relations du récit. En terme de dépendance sémantique (critère fondamental de la relation d'anaphore) le pronom est un cas extrême de non autonomie et, à l'inverse d'un syntagme nominal, il n'introduit pour son référent aucun nouveau prédicat sortal autre que le nombre ou le genre. Il est, selon la terminologie de [Milner 82], dépourvu de référence virtuelle et appelle une référence actuelle. L'anaphore pronominale ne peut donc être informative sur le référent. Son but, en général, est de replacer le même référent dans plusieurs propositions. L'anaphore pronominale maintient alors l'identité individuelle d'un référent dans le récit. L'anaphore pronominale est même, dans la plupart des cas, la seule façon de maintenir l'identité individuelle.

- (9) *Paul croit qu'il a de l'avenir.*

- (10) *Tous les concurrents espèrent qu'ils vont gagner.*

- (11) *Jean n'a pas épousé une suédoise. Elle est danoise.*

Dans l'exemple (9) notamment remplacer *il* par *Paul* ou encore *le jeune homme* tend à exclure la co-référence et dans les exemples (10) et (11)³ une anaphore autre que pronominale n'est plus du tout possible si l'on veut maintenir cette identité individuelle. Ces cas d'anaphores pronominales sont des exemples classiques mais l'on peut toujours trouver des exceptions comme les exemples (12) et (13).

- (12) *Le premier ministre nous a rendu visite. Elle est charmante.*

- (13) *J'ai acheté une Toyota parce qu'elles sont robustes. ([Kleiber 90])*

Dans l'exemple (12), le pronom est informatif sur le référent et dans l'exemple (13), il ne reprend pas l'identité individuelle d'un référent mais précise les éléments d'un ensemble.

Le devenir de l'identité dans la langue prend donc deux formes. Il s'agit, soit d'enrichir l'identité sortale de l'objet, soit d'assurer le maintien de son identité individuelle.

Nous avons jusqu'ici vu des cas où la continuité référentielle n'est pas problématique et où le maintien de l'identité paraît évident. Dans ces cas de récit, seul le temps change mais il existe d'autres cas où les référents changent également au cours du récit comme dans l'exemple (14).

³ Les exemples (9), (10) et (11) sont repris de [Charolles 91].

(14) *Prenez quatre pommes. Pelez-les, coupez-les et évidez-les. Faites-les cuire pendant une 1/2 heure, broyez-les jusqu'à ce qu'elles soient complètement réduites et, après les avoir laissées refroidir, servez-les avec des petits gâteaux (exemple extrait de [Charolles & al. 93])*

Le problème dans de tels énoncés est d'ordre cognitif. D'un point de vue lexical, le calcul de l'antécédent des pronoms anaphoriques (*les, elles* dans l'exemple (14)) ne pose pas de problème majeur mais, d'un point de vue pragmatique, la question de la nature de l'objet dans la relation de référence est beaucoup plus délicate. Peut-on encore, par exemple, penser que le référent à la fin de l'énoncé est bien le même que celui du début ? On est là face à un problème d'identité. Dans l'exemple (14), il paraît effectivement impossible dans la fin de l'énoncé d'utiliser le syntagme nominal *quatre pommes* à la place du pronom *les*. Dans ces cas de référents évolutifs, l'apparition d'un nouveau substantif est fréquente et précise une nouvelle identité sortale pour le référent. C'est le cas du syntagme *la compote* dans l'exemple (15).

(15) *Prenez quatre pommes. Pelez-les, coupez-les et évidez-les. Faites-les cuire pendant une 1/2 heure, broyez-les jusqu'à ce qu'elles soient complètement réduites et, après les avoir laissées refroidir, servez la compote avec des petits gâteaux.*

On observe alors une différence supplémentaire fondamentale entre l'identité sortale et l'identité individuelle. L'identité sortale est susceptible d'évoluer et même, dans des cas extrêmes, d'être rompue tandis que l'identité individuelle, à partir du moment où elle est précisée perdure quoi qu'il arrive. En fait, l'identité individuelle n'est que très rarement l'objet du problème sauf dans des exemples comme (16).

(16) *Il faut changer l'ampoule.*

Dans cet exemple, il y a d'une certaine façon, ambiguïté sur l'identité individuelle de l'objet à savoir si il faut changer "l'ancienne" ou "la nouvelle ampoule". Il faut ici associer au nom *ampoule* un objet indépendant de l'opération physique de changement et ne retenir que la fonctionnalité lumineuse de l'ampoule, c'est à dire un critère de son identité sortale. Il faut aussi remarquer que dans ce cas, où l'ambiguïté porte sur l'identité individuelle, nous sommes dans une situation d'émergence du discours dans le récit. Dans le récit, l'identité individuelle peut difficilement être remise en cause et si le sujet interprète normalement des situations de référents évolutifs, c'est parce qu'il s'appuie sur l'identité individuelle des objets même dans des cas très extrêmes où l'évolution anéantit le référent.

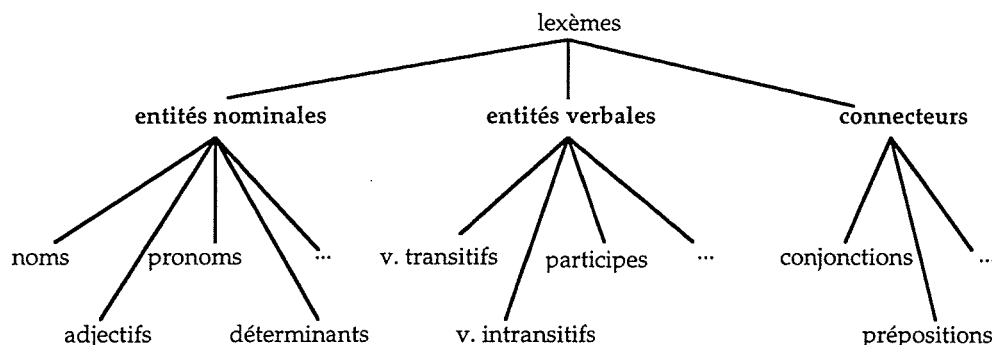
Le devenir dans la langue de l'identité, en général, se résume donc aux modifications subies par l'identité sortale. Dans les cas classiques, on a vu que cela consiste en un enrichissement à but informatif alors que dans des cas extrêmes cet enrichissement entre en contradiction profonde avec les critères sortaux déjà spécifiés. La mutation de l'objet se traduit par des critères sortaux qui s'effacent au bénéfice d'autres que le co-texte propage. D'un point de vue cognitif, ces critères sortaux proviennent d'un examen des différences entre l'objet en question et les autres ([Beust & al. 96]) et prennent part au processus de catégorisation de l'objet. C'est d'ailleurs, une fois qu'un objet est catégorisé qu'on peut évoquer ses critères sortaux (ceux qui le différencient des autres) et c'est donc directement de la catégorisation que provient l'identité sortale de l'objet. Quand l'identité sortale subit une évolution où certains critères sortaux ne peuvent plus être maintenus, les raisons pour lesquelles l'objet tombait sous une catégorie sont en cause. Cet objet prend alors place dans une autre catégorie, nouvellement créée ou déjà existante. Dans des contextes évolutifs, l'évolution de l'identité, donc de l'identité sortale, se traduit par une re-catégorisation de l'objet référent.

Examiner les différentes formes de dépendance nominale avec une problématique centrée sur l'identité permet donc de dégager deux dimensions dans les phénomènes de dépendance. La première est une dimension d'association qui correspond aux situations de discours où l'identité des référents n'est pas mise en cause. C'est le cas des co-références et des anaphores associatives. La seconde est une dimension de devenir de l'identité qui correspond aux situations de récit où soit l'on y précise l'identité sortale, soit l'on y maintient l'identité individuelle. C'est le cas des anaphores nominales et des cas classiques d'anaphores pronominales.

4. Un modèle pour l'identification de la dépendance nominale.

Comme la partie précédente l'a montré, identifier des phénomènes de dépendance nominale revient à se poser un problème sur l'identité sortale et individuelle des référents. Le modèle distingue la représentation des lexèmes, hors contexte, la représentation des mots, comme occurrences des lexèmes dans un énoncé, et la représentation des objets auxquels ils font référence. C'est dans cette représentation des objets que les notions d'identité sortale et d'identité individuelle sont prises en compte.

Dans cette étude des dépendances contextuelles, il est primordial de considérer la catégorie des éléments entre lesquels peuvent s'établir de telles relations. La catégorisation pertinente des lexèmes regroupe certaines des catégories de la syntaxe classique à un niveau superordonné. Cette catégorisation prend en compte la constitution des groupes syntaxiques, lieux des premières dépendances contextuelles. Dans ce cadre, différencier un nom propre d'un nom commun n'est pas pertinent car ils entrent tous deux dans la formation de groupes nominaux, de même pour la distinction entre noms et adjectifs. On propose donc trois classes d'éléments à considérer : les entités à caractère nominal (noms, adjectifs, pronoms, déterminants par exemple), les entités à caractère verbal (verbes transitifs, verbes intransitifs, participes présents par exemple) et les connecteurs (conjonctions de coordination ou prépositions par exemple).



Chaque élément des ces classes inclut des traits sémantiques qu'il est intéressant de catégoriser en fonction de leur caractère informatif en rapport à la classe. En complément des travaux de Rastier on cherche ici à préciser pour chaque lexème les différentes facettes que représentent les traits qu'il porte. Pour représenter ces traits, on considère un point de vue oppositionnel.

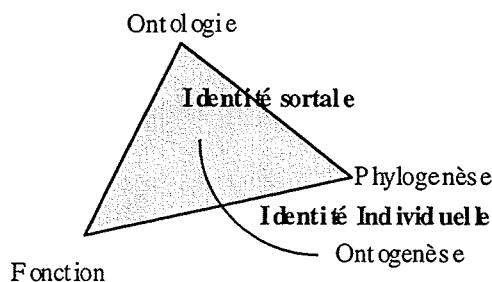
Considérer une approche oppositionnelle pour appréhender des phénomènes linguistiques est une méthode qui a déjà porté ses fruits. Notamment la phonologie de Jakobson en 1963 découle d'une telle approche en définissant les phonèmes d'une langue les uns par rapport aux autres en fonction d'un ensemble fini de traits (sept en ce qui concerne le français contemporain). S'inspirant de cette méthode, la sémantique différentielle ([Pottier 74]), point de départ des travaux de Rastier, propose de décrire des sémèmes (équivalents sémantiques du phonème) par un ensemble de traits permettant de les opposer deux à deux. Ainsi le sémème "chaise" décrit par les traits /avec dossier/ + /sur pieds/ + /pour une personne/ + /pour s'asseoir/ se différencie du sémème "fauteuil" décrit par les traits /avec dossier/ + /sur pieds/ + /pour une personne/ + /pour s'asseoir/ + /avec bras/ par l'unique traits /avec bras/.

Les oppositions dont il est question en sémantique différentielle constituent des oppositions de signes qui donnent lieu à des graphes conceptuels. Nous proposons ici de pousser un cran plus loin l'approche oppositionnelle en considérant que les traits sémantiques peuvent également être décrits en terme d'oppositions et que cela fonde leur pertinence dans le système qu'est la langue. Il ne s'agit pas, ce faisant, de montrer les limites ou les manques de la sémantique différentielle. Simplement, dans une recherche en intelligence artificielle, les buts ne sont plus les mêmes que dans une recherche linguistique. Il ne s'agit pas seulement de pouvoir expliquer précisément les phénomènes de langue mais il faut justifier la pertinence des connaissances que l'on veut donner à une machine. La différence entre les deux disciplines se traduit par la place de l'interprétation dans leurs modèles et n'impose pas les mêmes compétences (ce qui rend les coopérations transdisciplinaires possibles). L'interprétation pour une machine ne peut préexister à l'analyse ce qui contraint l'intelligence artificielle à proposer des modèles effectifs là où la linguistique pourrait se contenter de modèles descriptifs.

Les traits sémantiques que prend en compte notre modèle de la dépendance contextuelle prennent part à la constitution de la signification (et non uniquement à sa description). Ils doivent être des entités à partir desquelles on doit pouvoir construire un type (le sens en étant une instance), c'est à dire d'une certaine façon des contraintes délimitant un domaine.

C'est à travers les facettes des entités nominales que l'on peut trouver un formalisme qui rende compte de l'identité telle qu'on l'a décrite précédemment. Pour décrire les objets du monde, on considère une structure à quatre facettes inspirée de la théorie du système général de Le Moigne ([Le Moigne 84]). Le Moigne propose trois dimensions dans la conception des artefacts : la dimension ontologique (qu'est-ce que c'est ?), la dimension fonctionnelle (qu'est-ce que ça fait ?) et la dimension génétique (d'où ça vient et qu'est-ce que ça devient ?). On distingue deux aspects dans la dimension génétique, celui de l'ontogénèse (d'où vient l'objet et

qu'est-ce qu'il devient) et celui de la phylogenèse (de quelle espèce ça vient ?). Ainsi, les traits de nature ontologiques, fonctionnels et phylogénétiques donnent un cadre à la catégorie de l'objet en cause. Ces trois facettes forment une représentation de l'identité sortale de l'objet. La facette ontogénétique inclut une dimension temporelle dans la représentation des objets et permet de les considérer dans des situations de récit et non plus uniquement dans des situations de discours. L'identité individuelle est, elle, directement introduite par la facette ontogénétique de l'objet. Elle résulte, dans ce modèle, de la "projection" de l'identité sortale sur une ontogénèse. L'identité individuelle est en effet plus qu'une simple instanciation de l'identité sortale. Elle y est rattachée mais elle perdure même si l'identité sortale est rompue car les traits de la facette ontogénétique se conservent quoi qu'il arrive.



Les relations de dépendance contextuelle entre différents mots (occurrences de lexème) vont consister en des propagations de traits de l'un vers l'autre. Autrement dit, la mise en contexte des lexèmes forme des mots qui via les relations vont enrichir leurs contenus sémantiques et préciser, de ce fait, les deux volets de l'identité de leur référent. Chaque mot inclut des traits, qui proviennent de son lexème, que l'on appelle des traits inhérents et les traits que reçoivent les mots par les relations qu'ils entretiennent sont appelés des traits afférents. Ce sont des traits inhérents d'un mot qui vont être afférents chez l'autre et vice versa si la relation est réciproque.

Les relations de dépendance nominale propagent donc les traits d'un mot sur un autre. Les nécessités de cette propagation amènent à approfondir le modèle des traits lui-même. On ne peut donc plus considérer un trait de façon positive qui se suffit à lui-même comme le faisait la sémantique différentielle et considérer de plus qu'un trait ne soit pas affecté lors d'une afférence mais simplement propagé. Nous les considérons ici, dans une vision saussurienne, comme des entités négatives provenant d'une ou plusieurs oppositions. Par exemple considérer le trait /animal/, c'est considérer prioritairement en fonction du contexte un terme de l'opposition /animal/ vs. /objet/ par exemple ou encore de l'opposition /animal/ vs. /humain/ ou même /animal/ vs. /divin/. C'est une de ces trois oppositions qui devra être retenue dans le contexte pour une interprétation non ambiguë du trait /animal/. Cette opposition fonde la pertinence du trait considéré dans le contexte et c'est celle-ci qu'il convient d'y représenter. De même, pour donner un autre exemple, l'interprétation du trait /solide/ résulte d'un choix de l'une des oppositions /solide/ vs. /gazeux/ vs. /liquide/ ou /solide/ vs. /fragile/.

Cette vision oppositionnelle des traits sémantiques se justifie par le potentiel ambigu des sèmes. Ceci s'explique par le fait que la langue naturelle est réflexive et permet d'utiliser des termes de la langue comme traits sémantiques pour étudier la langue. Ceci nous confronte dans les structures opérationnelles du modèle à la polysémie et aux usages métaphoriques puisqu'un trait peut prendre plusieurs significations en fonction de ce à quoi on l'oppose. On est ici face à la dimension deffective de la langue ([Coursil 92]) qui a la capacité de s'autodéfinir à partir de coupures (de différences). Les usages polysémiques et métaphoriques ne constituent pas des formes particulières et rares de la langue ([Lakoff & al. 85]), elles occupent une place centrale dans les processus d'interprétation, et les modèles d'intelligence artificielle appliqués à la sémantique ne peuvent en faire l'économie.

Définir un trait hors contexte revient à mettre en évidence un jeu d'oppositions, la mise en contexte du trait se traduisant par le choix de la bonne opposition. C'est ce choix qui va être le résultat immédiat de l'afférence d'un trait. Dans la suite, on se donne les définitions (récursives) suivantes:

sème : ensemble de traits
 trait : opposition de sèmes

Par exemple le sème /haut/ se définit par les traits qui dénotent un usage possible (souvent métaphorique) d'un lexème qui porte le sème /haut/ :

traits	exemples d'usages (le lexème en italique est celui porte le sème)
/haut/ vs /bas/	il est en <i>haut</i> .
/haut/ vs /tristesse/	Il m'a <i>remonté</i> le moral.
/haut/ vs /maladie/	Il est au <i>sommet</i> de sa forme.
/haut/ vs /soumission/	J'ai pris le <i>dessus</i> sur lui.
/haut/ vs /moins/	<i>Baissez</i> le chauffage.
/haut/ vs /mauvais/	Il a fait un travail de <i>haute</i> qualité.
/haut/ vs /vice/	Je ne m' <i>abaisserai</i> pas à ça.
/haut/ vs /affectif/	Il doit <i>dominer</i> ses émotions.
/haut/ vs /connu/	Le problème est <i>posé</i> .

Le choix de la bonne opposition à considérer intervient lorsqu'un trait est afférent pour un lexème en contexte. Ce choix doit être fonction du type de la relation qui véhicule l'afférence (à savoir si c'est une relation de cohérence du groupe nominal ou anaphorique par exemple) et du contenu inhérent du lexème sur lequel porte la relation. On peut décrire une telle afférence dans l'exemple (17) même si le contexte y est très restreint.

(17) *La nuit fût dure.*

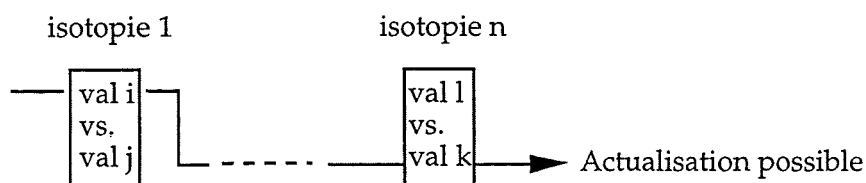
Un trait actualisé dans un lexème se trouve rattaché à une facette du mot et c'est l'influence de ce trait, via la facette, sur la représentation du référent qui peut donner lieu à une négociation sur l'identité de ce référent. Dans l'exemple (17), on peut dégager en terme d'oppositions, pour le nom commun *nuit* et l'adjectif *dure* les contenus sémantiques inhérents suivants :

<i>nuit</i> :	op1 :	nuit vs. jour (facette ontologique)
	op2 :	durée vs. état (facette ontologique)
<i>dure</i> :	op3 :	dur vs. mou (facette ontologique)
	op3bis :	dur vs. facile (facette ontogénétique)

Les mots *nuit* et *dure* sont en relation de dépendance via la facette actancielle du verbe (*fût*). Cette relation entre un nom et un adjectif est unilatérale dans le sens où c'est le contenu de l'adjectif qui vient "enrichir" le contenu du nom et qu'il n'y a pas de réciprocité (une relation anaphorique nom-nom eût été réciproque par exemple). La relation se traduit ici par l'afférence du contenu de *dure* vers *nuit*. Le problème soulevé par cette afférence est de savoir laquelle des oppositions op3 ou op3bis doit constituer l'enrichissement sémantique de *nuit*. La réponse provient du conflit entre les oppositions op3 et op2. L'opposition op3 est caractéristique d'un objet matériel or op2 précise que *nuit* est de nature immatérielle, op2 et op3 sont donc incompatibles dans le même contenu sémantique. L'opposition op3bis ne pose pas ce problème de compatibilité, l'afférence de op3bis est alors recevable. A l'issue de ce calcul on peut dégager le contenu sémantique de *nuit* dans ce contexte :

<i>nuit</i> :	inhérents :	op1 :	nuit vs. jour (facette ontologique)
		op2 :	durée vs. état (facette ontologique)
	afférents :	op3bis :	dur vs. facile (facette ontogénétique)

Il apparaît alors dans l'énoncé une isotopie constituée par la répétition du trait op3bis. Dans un cas où plusieurs isotopies peuvent être extraites, le jeu d'oppositions qui en résulte fournit un cadre où plusieurs sens peuvent être actualisés. Dégager n isotopies permet de constituer un registre (au sens de [Beust & al. 96]) donnant lieu à 2ⁿ actualisations possibles.



Dans l'effectivité de ce modèle, le potentiel de sens de l'énoncé est en relation avec le potentiel de description du registre formé par les isotopies. Plus l'énoncé est porteur d'isotopies, plus on peut actualiser de sens différents, donc plus le potentiel de sens ([Brassac & al 96]) est important et réciproquement.

Pour donner à ce modèle sa dimension opératoire, il faut maintenant pouvoir évaluer automatiquement des conflits entre traits inhérents et traits afférents, c'est à dire mettre en place une algèbre des oppositions. Les premières bases d'une telle algèbre sont peut-être à chercher dans le carré des oppositions de la logique aristotélicienne. Ce modèle du carré des oppositions a pour but une étude formelle de la valeur de vérité d'énoncés de logique propositionnelle par les combinaisons possibles des oppositions entre la partie et le tout et une chose et son contraire. Ce cadre logique ne peut fournir qu'une première ébauche de l'algèbre que nous recherchons car les valeurs qui intéressent notre modèle ne sont plus uniquement des valeurs de vérité.

Conclusion

Si le sens était dans les mots, apprendre une nouvelle langue reviendrait à apprendre un nouveau lexique. Ce n'est pas le cas, car d'une part, les jeux d'oppositions ne sont pas les mêmes et d'autre part, les systèmes de dépendance sont différents. Chercher à préciser le sens d'un mot dans un énoncé amène à résoudre les dépendances qu'il entretient avec les autres mots. Dans notre projet, qui consiste à donner aux machines des facultés d'interprétation, reconnaître et exploiter ce genre de dépendance est nécessaire. Mais, de plus, pour dialoguer en langue naturelle ([Nicolle 93]), une machine doit construire une co-référence avec son interlocuteur. Produire, dans une interaction, des tentatives de référence, suppose que l'identité des choses évoquées ne soit pas partagée préalablement car sinon le problème de la référence ne se poserait pas.

Une approche oppositionnelle est une voie de recherche dans la modélisation du sens qui convient à un traitement interprétatif cohérent des usages métaphoriques. Si les machines que l'on veut placer dans des situations dialogiques n'ont pas cette capacité, l'illusion de dialogue ([Vivier & al. 96]) pour le partenaire humain risque de s'atténuer très vite au cours de l'interaction langagière car la métaphore est constitutive de la langue.

Références :

- [Beust & al. 96] P. Beust, L. Delépine, A. Nicolle, J. Coursil, *Anadia, a Relevance Examination Tool for Representations*, Third European Congress on System Science, Rome, October 1996
- [Brassac & al. 96] C. Brassac, J. Stewart, *Le sens dans les processus interlocutoires, un observé ou un co-construit ?* Cinquièmes Journées de Rochebrune "Du collectif au social", 29 janvier - 3 février 1996
- [Charolles 91] M. Charolles, *La veuve et l'orphelin ou comment les îlots anaphoriques refont surface*, VII^{ème} colloque international de linguistique, Metz, Novembre 1991
- [Charolles & al. 93] M. Charolles, C. Schnedeker, *Coréférence et identité: le problème des référents évolutifs*, Langages, 112, p. 106-126, 1993
- [Charolles 94] M. Charolles, *reprise pronominale et prédications transformatrices: l'interprétation de la référence pronominale à la suite du verbe "changer"*, in Actes du colloque Relations anaphoriques et (i)cohérence, Anvers, 1-2/12/1994.
- [Coursil 92] J. Coursil, *Grammaire analytique du français contemporain - Essai d'intelligence artificielle et de linguistique générale*, Thèse de l'Université de Caen, 1992
- [Coursil 93] J. Coursil, *Rapport en vue de l'obtention du diplôme d'habilitation à diriger des recherches*, Caen 1993
- [Gosselin 96] L. Gosselin, *Sémantique de la temporalité en français, Un modèle calculatoire et cognitif*, Duculot, Louvain-la-neuve, 1996
- [Greimas 66] J. A. Greimas, *Sémantique structurale*, Larousse, Paris, 1966
- [Kleiber 90] G. Kleiber, *Quand "il" n'a pas d'antécédent*, Langages, 97, p. 24-50, 1990
- [Kleiber 93] G. Kleiber, *Anaphore pronominale et référents évolutifs ou comment faire recette avec un pronom*, conférence au colloque "Relations anaphoriques et (in)cohérence", Anvers, 1993
- [Lakoff & al. 85] G. Lakoff, M. Johnson, *Les métaphores dans la vie quotidienne*, Editions de Minuit, 1985
- [Le Moigne 84] J. L. Le Moigne, *La théorie du système général. Théorie de la modélisation*, Presses Universitaires de France, 1984
- [Milner 82] J.C. Milner, *Ordres et raisons de langue*, Editions du Seuil, Paris, 1982
- [Moeschler & al. 94] J. Moeschler, A. Reboul, *Dictionnaire encyclopédique de pragmatique*, Seuil, Paris, 1994
- [Mugur-Schächter 89] M. Mugur-Schächter, *Esquisse d'une méthode générale de conceptualisation relativisée*, dans "Arguments pour une méthode, autour de Edgar Morin", Editions du Seuil Paris 1989
- [Nicolle 93] A. Nicolle, *Towards a Natural Language Dialogue with Machines*, Second European Congress on System Science, Prague, 1993
- [Peirce 78] Charles S. Peirce, *Écrits sur le signe*, rassemblés, traduits et commentés par Gérard Deledalle, Editions du Seuil, 1978
- [Pottier 74] B. Pottier, *Linguistique générale, Théorie et description*, Klincksieck, Paris, 1974
- [Rastier 87] F. Rastier, *Sémantique interprétative*, Presses Universitaires de France, 1987
- [Saussure 86] F. de Saussure, *Cours de linguistique générale*, ed. Mauro-Payot, Paris 1986
- [Vivier & al. 96] J. Vivier, D. Jacquet, *La référenciation dans le dialogue homme-machine et l'utilisation des pronoms personnels*, Third European Congress on System Science, Rome, October 1996

Invariants sensorimoteurs et Perception

C. Lenay (COSTECH-UTC)

L'étude de la perception est profondément renouvelée par la découverte de dispositifs techniques permettant de mettre en place de **nouvelles modalités perceptives**. De tels dispositifs techniques, relativement simples, permettent l'analyse empirique de l'apprentissage perceptif chez l'adulte. Cette approche met en évidence le rôle essentiel de l'action du sujet dans la constitution de ses perceptions. L'objet de cet article est de décrire les perspectives théoriques et expérimentales originales ouvertes par ces techniques.

La perception d'une forme (ou d'un événement) est la reconnaissance d'une même chose dans le divers des sensations. C'est une forme d'**invariance** de la relation du sujet avec son environnement (relativement à d'autres variations internes ou externes). Cette invariance doit être donnée dans un espace et un temps de la perception pour que puisse être pensé l'identité entre deux choses distinctes (données en différents lieux ou à différents moments).

On cherche classiquement à décrire cette invariance comme celle de l'attracteur d'un réseau neuronal interne. Ce qui serait perçu, ce serait une certaine régularité (statique ou dynamique) parmi la diversité des stimuli reçus. Après le réglage adéquat des poids des connexions entre les neurones on pourra retrouver un même attracteur pour diverses occurrences d'une même forme. Ainsi, la perception serait une opération de traitement de l'information reçue par le sujet. L'invariance interne atteinte reflèterait simplement une invariance externe objective.

Une telle perspective doit être critiquée et une alternative proposée.

La critique peut être conduite à partir de diverses considérations philosophiques mais l'on peut aussi partir d'une découverte empirique. Elle est donnée par les dispositifs développés depuis la fin des années 60 aux Etats-Unis par le professeur Paul Bach y Rita, et qui sont improprement appelé "systèmes de substitution sensorielle" (on verra plus loin pourquoi ces termes sont inadéquats). Ces dispositifs transforment les stimuli attachés à une modalité sensorielle (la vision) en des stimuli d'une autre modalité sensorielle (le toucher). Classiquement (Tactile Vision Substitution Systems), une image visuelle captée par une caméra vidéo est convertie en une "image" tactile

composée d'une surface de stimulateurs (en général une matrice de 20 / 20) placée soit dans le dos, soit sur le thorax, soit sur le front (COLLINS 1973).

De tels dispositifs ont été développés pour l'aide aux personnes atteintes de cécité visuelle (congénitale ou acquises), mais ils peuvent aussi être utilisés par des voyants ayant les yeux bandés. Notre travail à Compiègne concerne l'évaluation et l'amélioration de ces systèmes pour handicapés, mais ce qui nous intéresse ici, c'est l'outil expérimental qu'ils offrent pour explorer les mécanismes fondamentaux de la perception.

En effet, les premières utilisations de ces dispositifs ont apporté trois résultats fondamentaux qui servent de point de départ à nos recherches.

i)- Tout d'abord, la présentation de formes à la caméra immobile ne permet qu'une discrimination très limitée des stimuli reçus, et les formes sont perçues à la surface de la peau. Ainsi la simple substitution, d'une entrée tactile à une entrée par le nerf optique, ne donne pas, en tant que telle, accès à une perception spatiale.

ii)- Mais, si l'utilisateur dispose des moyens de manipuler la caméra (mouvement de droite à gauche, de bas en haut, zoom avant et arrière, ouverture), il apprend progressivement comment les variations de ses sensations sont liées à ses actions. Il devient alors capable de reconnaître des formes complexes, même des visages. De plus cette reconnaissance s'accompagne d'une **mise en extériorité** des percepts en des objets distaux avec leurs positions respectives dans l'espace. L'utilisateur **oublie la sensation** de toucher pour **percevoir** des objets et événements à distance devant lui. Ainsi, d'après les témoignages des utilisateurs les irritations proximales que peut provoquer la plaque tactile sont clairement distinguées de la perception proprement dite. Cette localisation subjective des objets dans l'espace se produit rapidement (après 5 à 15 heures d'entraînement). Certaines illusions visuelles classiques sont spontanément reproduites (comme l'effet de cascade).

La nécessité d'une action du sujet sur les capteurs sensoriels pour que se constitue une perception, rend inadéquate l'appellation de "systèmes de substitution sensorielle". En effet, ce que doit amener le dispositif technique c'est autant un accès à de nouvelles données sensorielles que le pouvoir d'agir sur le système récepteur. Nous préférons donc les termes de "dispositifs de couplage sensorimoteurs" ou d'"organes de perception artificiels". Notons cependant, qu'après une période d'apprentissage suffisamment prolongée, la caméra peut rester immobile pour certaines tâches de reconnaissance, comme le suivi d'une balle en mouvement que le sujet doit frapper. On voit que de tels dispositifs expérimentaux donnent un contenu empirique possible à la question de l'intentionnalité, c'est-à-dire de la conscience de quelque chose comme en extériorité (l'apparaître d'un phénomène dans le champ perceptif), puisqu'ils permettent d'en suivre

et reproduire la genèse chez l'adulte. Ils tendent à montrer que la construction d'un espace de représentation se réalise à partir de l'extraction d'invariants sensorimoteurs, ce qui est porteur de profonds changements conceptuels sur la cognition en général.

iii)- Cependant, les premières réussites spectaculaires des systèmes développés pour les handicapés ont été accompagnées d'une déception inattendue: le rejet par les utilisateurs aveugles de dispositifs qu'ils déclaraient décevants et déprimants. En effet, ce que peuvent donner de tels systèmes à un aveugle, ce n'est pas cette vision dont lui parlent tant les voyants. Il n'y a pas de couleur, peu de points et une caméra dont les mouvements sont difficiles et limités, ce qui produit une grande lenteur dans la reconnaissance de la situation. Ce couplage sensori-moteur ressemble par certains aspects à celui de notre vision, mais il ne s'y *substitue* pas. Il réalise plutôt une *addition*, une nouvelle modalité perceptive qui ouvre un nouvel espace de couplage du sujet avec son environnement.

Ce qui manque le plus cruellement au dire des aveugles, ce sont les qualités, les valeurs subjectives des perceptions (par exemple, les sujets découvrent leurs visages ou celui des leurs proches sans y attacher aucune émotion). Mais le contraire aurait été étonnant. La signification n'est pas une simple information à capter. Elle est aussi à construire, soit par interaction interne avec les autres dimensions sensorielles et les sentiments de plaisir et de peine, soit par interaction externe avec d'autres sujets. La question soulevée par les premiers utilisateurs, de la valeur, du sens subjectif accordé ou non à ces perceptions nouvelles, pose donc le problème de la liaison entre représentations et affects, tant sur le plan des émotions individuelles que sur celui des significations collectivement réglées.

Il y a là deux champs de recherche qu'il ne faut pas séparer : l'apprentissage de la perception comme acquisition et composition d'invariants sensorimoteurs de niveaux de plus en plus élevés, et la constitution collective d'un système de valeurs attaché à ces invariants. Pour parler en termes kantien, il s'agit de faire le lien (absent chez Kant lui-même) entre l'esthétique de la sensibilité (Critique de la raison pure) et l'esthétique du jugement (Critique de la faculté de juger).¹

1) Invariants sensorimoteurs individuels

Les dispositifs de substitution sensorielle bouleversent les modes de définition classiques des diverses modalités sensorielles. Si "voir" n'est plus caractérisé par l'utilisation de l'oeil, ni peut-être même par l'activité d'une aire corticale particulière, il faut redéfinir les divers sens par les types de couplages sensorimoteurs qu'ils engagent entre l'organisme et le milieu.

¹ Voir Armen Khatchatourov, Technique et perception, Médiation technique et esthétique des couplages sensori-moteurs, mémoire de DEA, Compiègne, UTC, 1996.

Le rôle essentiel joué par l'action sur les organes sensoriels artificiels dans l'émergence progressive de représentations structurées conduit à une conception de l'apprentissage et de la perception comme un processus d'extraction d'invariants sensorimoteurs. On abandonne ici la conception classique de la perception pour laquelle le système recevrait en entrée une information puis effectuerait un calcul pour identifier des objets, des événements, et produire des représentations avec lesquelles pourrait être conduit un raisonnement qui, le cas échéant, aboutirait à une action adaptée. Ce qui est perçu, reconnu, ne sont pas à proprement parler les invariants de la sensation, mais ce sont les invariants de cercles sensori-moteurs inséparables de l'activité du sujet. La perception d'un objet consiste en l'extraction de ce qui reste constant dans la liaison entre variation d'action (mobilité de l'organe de perception) et variation de sensation. La richesse de la perception dépend donc autant des qualités de l'action, (mobilité, rapidité, zoom, etc.) que des qualités de la sensation (sensibilité, largeur du spectre, nombre de senseurs, etc.). C'est ce que montre ici à l'évidence l'étonnante capacité de reconnaissance des visages à partir de dispositifs artificiels qui ne donnent en entrée sensorielle que quelques 400 points distincts. Le visage reconnu émerge comme un invariant de haut niveau des changements de sensation associés à son exploration active.

L'état perceptif de l'agent est caractérisé par l'état du cercle sensorimoteur dans lequel il est actuellement engagé. Un événement perceptif est une perturbation de ce couplage due à un changement du milieu (ou à l'activité exploratoire de l'organisme) qui conduit à un nouvel invariant sensori-moteur. De telles perturbations sont possibles dans la mesure où l'action ne contrôle pas entièrement la sensation, c'est-à-dire dans la mesure où l'agent est couplé avec un milieu externe toujours inconnu. Mais, c'est l'état du couplage présent qui définit les perturbations possibles : ce qui est accessible à la perception dépend de l'histoire du couplage. A chaque instant, les perceptions possibles sont conditionnées par le type d'action qui a été produite et par la liaison que l'environnement permet entre ces actions et les sensations. La construction des invariants sensori-moteurs n'est déterminée ni par l'organisme, ni par le monde, mais à la fois par l'un et l'autre, dans leur couplage et son histoire. L'objectivité est ni un donné pur, ni une illusion solipsiste, mais une construction effectuée sous des contraintes à la fois réelles et historiques. Elle dépend indissociablement de l'état du milieu (qui a pu être modifié par les actions précédentes) et des actions produites (qui dépendent des perceptions précédentes).

La pleine acceptation d'une telle conception conduit à reconnaître que la perception n'est pas à proprement parler dans le sujet. Elle est dans le couplage lui-même, couplage qui implique autant le sujet que le milieu dans lequel il est incarné.

Dans les perspectives représentationnalistes classiques (computationnalistes ou connexionnistes), on rencontre toujours une aporie des dimensions spatiales et temporelles de l'expérience. D'une part, pour décrire le monde extérieur, on se donne

d'emblée un espace et un temps objectif, indépendant du sujet connaissant. D'autre part, si ce qui est **reçu** par le système cognitif est de l'information (au sens de Shannon), ou des configurations d'activation de neurones récepteurs, il faudrait rendre compte de la constitution indépendante d'un espace et d'un temps interne. Dans le computo-représentationnalisme, la perception correspond en dernier lieu à l'activation d'un symbole mental (état fonctionnellement lié aux autres états du système cognitif). Mais si l'on veut pouvoir représenter que le **même** symbole (la reconnaissance d'une même forme) a été deux fois activé par la **même** chose en deux lieux ou deux moments différents, il faudra différencier ces occurrences en marquant (indexant) différemment ces symboles suivant une règle interne formalisant la simultanéité spatiale ou la succession temporelle. On devra alors se demander de quel point de vue ces symboles différents peuvent cependant être perçus comme les mêmes. Il n'y a pas de spatialité de l'information mais seulement la temporalité externe du calcul rythmé par l'horloge.

Dans le connexionnisme, si l'on admet que la reconnaissance d'une **même** chose en deux lieux ou deux moments différents correspond à l'activation du **même** attracteur de la dynamique interne, on devra se demander comment les distinguer. Faut-il admettre un homoncule, doté d'un espace et d'un temps de la perception pour reconnaître les différentes occurrences du même attracteur ? On peut bien sûr, mobiliser un autre réseau spécifiquement consacré à la localisation des choses perçues. En croisant les résultats des deux réseaux, on pourra donc différencier deux choses semblables qui seraient en deux lieux différents. Mais c'est l'observateur extérieur qui traduit les configurations d'activation du réseau de localisation comme représentant des positions spatiales plutôt que toute autre forme. On ne fait que reproduire le problème initial. Si l'on admet que la reconnaissance de deux lieux différents correspond à l'activation de deux configurations distinctes de ce réseau, on devra se demander dans quel espace elles pourraient être saisies simultanément. On est toujours reconduit au très difficile problème du temps. Aucune solution pour sa construction interne semble pensable. Classiquement, on se donne brutalement le temps externe d'une horloge. Le temps du système cognitif est donc le temps de l'observateur (ou du concepteur) extérieur. On ne voit vraiment pas comment, sans se les donner a priori, les dimensions de la sensibilité pourraient être construites de novo **dans** un sujet artificiel.

Mais si l'on abandonne l'idée que la perception est une simple réception d'information, il devient inutile (et absurde) de vouloir reconstruire un temps et un espace de la perception face au temps et à l'espace d'un monde externe posé indépendamment. Si l'on accepte l'idée que la perception est interne à l'invariant sensorimoteur, c'est-à-dire qu'elle concerne autant l'action que la sensation, alors l'espace de la perception (l'espace dans lequel les perceptions sont données), est en même temps l'espace **dans** lequel le sujet agit (mouvements de l'œil, déplacements du corps). Le sujet **habite** le monde qu'il perçoit. On dispose dès l'origine d'une spatialité

et d'une temporalité, mais leurs formes dépend du mode de couplage, c'est-à-dire essentiellement de la diversité possible des actions.

L'étude de l'enchaînement des différentes étapes de l'apprentissage d'une nouvelle modalité sensorielle montre une progression ordonnée, depuis la reconnaissance de formes simples et l'extraction de contours jusqu'à la compréhension d'une scène complexe parcourue de mouvements. La reconnaissance d'une forme simple (par exemple une ligne droite ou une ligne courbe) qui d'abord exige son exploration active peut, après entraînement, être directement perçue, ce qui permet l'exploration de formes plus élaborées (objet dans l'espace, constance de taille, etc.). Pour maintenir l'idée d'une perception essentiellement attachée à l'action, nous décrivons ce processus comme une **internalisation** de l'action. Les actions (mouvements de l'oeil ou de la caméra) qui dans un premier temps étaient externes, deviendraient progressivement des actions internes de déplacement de l'attention sur l'image reçue. Dès lors les mouvements externes ne seraient que la portion congrue d'une activité du sujet très largement déterminée par son organisation interne.

Cet apprentissage met en évidence une étonnante plasticité du système nerveux central qui doit subir de vastes réorganisations fonctionnelles. L'entrée sensorielle tactile n'a rien à voir avec celle du système visuel, pas plus que le contrôle de la caméra par les mains n'a de relations avec les commandes des muscles oculaires. Pourtant le cerveau se révèle capable d'organiser un monde perceptif dont les formes et événements croisent ceux qui nous sont donnés dans la perception visuelle. De plus, si pour un sujet entraîné, on déplace la plaque de stimulateurs tactiles du thorax vers le dos, et que l'on remplace la caméra qui était tenue dans les mains par une caméra miniaturisée attachée à une monture de lunette, l'adaptation est presque immédiate. En quelques minutes le sujet retrouve une perception distale devant lui. Cette plasticité cérébrale stupéfiante me semble forcer une remise en question de l'approche traditionnelle des recherches sur le cerveau. Classiquement, on cherche à déduire son mode de fonctionnement à partir d'une analyse de sa structure. Ici, au contraire, on voit comment son mode de fonctionnement restructure profondément le cerveau. A la limite, ce n'est plus l'organe qui explique la fonction, c'est la "fonction qui fait l'organe" pour employer l'expression lamarckienne tant décriée du point de vue de la biologie darwinienne. Au minimum, il faut dire que structure et fonction sont réciproquement conditionnées et transformées dans l'histoire du couplage entre l'organisme et son environnement².

² A côté de Lamarck, il faudrait citer Diderot, l'auteur de la Lettre sur les aveugles, qui dans Le Rêve de d'Alembert écrivait: "Les organes produisent les besoins, et réciproquement, les besoins produisent les organes.". Cependant Lamarck est plus précis quand il montre comment "les circonstances créent de nouveaux besoins qui provoquent un effort de l'organisme, c'est-à-dire une mobilisation du fluide nerveux vers l'organe concerné. Le fluide nerveux modifie le cours des fluides ordinaires, ce qui, à la longue, modifie l'organisateur" (J.R. 298). Et pour l'homme Lamarck précise qu'"une race de quadrumanes" est à l'origine de l'espèce humaine qui aura pu "vivre par troupe nombreuses, se sera successivement créé des besoins nouveaux qui auront excité son industrie et perfectionné graduellement ses moyens et ses facultés" (1809, I, p.351)

La modélisation d'un tel apprentissage nécessite d'abord de trouver des outils conceptuels pour penser une telle plasticité et de telles propriétés adaptatives. L'idée qu'un état perceptif correspond d'abord à un invariant sensori-moteur exige que l'on modélise en même temps l'organisme et son environnement (complexe, variable, et modifiable par ses actions).

2) Invariants collectifs

L'absence de valeurs émotionnelles attachées aux perceptions artificiellement acquises peut recevoir différents types d'explications.

- Soit, ce qui fait défaut c'est un apprentissage suffisamment précoce et prolongé pour que soient liées aux perceptions, des **anticipations** de plaisir ou de peine qui les valoriseraient. C'est cette hypothèse qu'explore actuellement Eliana Sampaio à Strasbourg en installant des dispositifs de substitution visuo-tactile sur des nourissons³. Notons que de telles anticipations nécessitent une mémoire individuelle susceptible d'enregistrer les invariants sensorimoteurs atteints pour les lier à leurs effets différés.

- Soit, la signification dont peuvent être porteurs les invariants perceptifs serait liée à l'existence d'une histoire et d'une mémoire collective, mémoire qui devrait pouvoir émerger du jeu des interactions de plusieurs sujets dans un même milieu. En effet, l'idée de couplage sensori-moteur peut s'élargir à celle d'un couplage collectif entre différentes personnes dans un monde partagé, couplage qui donnera une signification aux objets de ce monde. Agissant de telle façon sur tel objet de ma perception je provoque telle réaction d'autrui qui lui même agit sur cet objet. Quand le jeu des interactions se stabilise autour d'objets ou de comportements, chacun peut anticiper les effets de ses actions alors même qu'ils dépendent des perceptions et interprétations des autres. Dans de telles conditions, un sujet peut croire reconnaître ce qui existe pour l'autre comme pour lui. Il y a construction collective d'une objectivité commune et d'une mémoire partagée⁴. Les objets collectivement définis sont à la fois la mémoire de nos actions passées et la condition des actions futures qui pourront les modifier⁵. La

³ Sampaio, E. Gouarir, C. & Mvondo, D. (1995) "Tactile and visual bisection tasks by sighted and blind children". *Developmental Neuropsychology*, 11. 109-127 ; Sampaio, E. & Philip, J. (1995) "Influences of age at onset of blindness on Braille reading performances with left and right hand", *Perceptual and Motor Skills*, 81, 131-141 ; Sampaio, E. (1995) "Les substitutions sensorielles adaptées aux déficits visuels importants". In : A. B. Safran & A. Assimacopoulos (Eds.). *Le Déficit Visuel. Des fondements neurophysiologiques à la pratique de la réadaptation*. Paris. Masson 197-211.

⁴ Si mon action me permet d'anticiper ma perception alors même que celle-ci dépend des actions d'autrui, je crois reconnaître ce qui existe pour l'autre comme pour moi. C'est encore sur l'invariance de ma perception de ce cercle sensori-moteur qui inclut autrui, que je fonde mon critère local (et bien sûr toujours incertain) d'objectivité.

⁵ Pour rendre compte d'une objectivité partagée (Common Knowledge) dans une perspective individualiste sur la représentation, on rencontre inévitablement le problème d'une spécularité nécessaire (et absurde) : chacun doit savoir que l'autre sait qu'il sait que l'autre sait qu'il sait... que telle proposition sur le monde est vraie. Une telle spécularité se trouve ici évitée parce que toujours précédée de la circularité sensori-motrice des interactions des sujets avec un monde concret quoique collectivement construit.

signification devrait ainsi surgir de l'inscription de modifications dans cet espace commun puisque les actions des uns pourraient alors faire sens au regard des autres.

D'ailleurs, il n'est pas nécessaire que les dispositifs de couplage soient strictement identiques. Ce qui est important, c'est que les participants veuillent bien le croire, et que ces dispositifs permettent effectivement l'établissement de couplages sur de mêmes objets dans un environnement partagé. Un daltonien ou un myope n'ont pas le même type de couplage sensori-moteur visuel avec leurs environnements, mais tant qu'ils l'oublient et qu'ils peuvent interagir à travers les mêmes objets, ils ont l'impression de partager un même monde, de voir les mêmes choses. C'est je crois le cas des bébés aveugles auxquels on a donné un système artificiel de couplage visuo-tactile, et qui semblent s'investir immédiatement dans le monde chaleureux des relations avec leur mère.

Ce n'est que dans la situation extrêmement artificielle d'un nouveau couplage sensori-moteur acquis à maturité (soit par un dispositif technique, soit par abatement de la cataracte) que l'on peut envisager de séparer la constitution d'un monde de simples perceptions discriminatives, et la constitution d'un monde de signification. Dans le cas des adultes, il me semble inutile de cacher les différences des techniques de couplage perceptif. Mais si l'emploi du dispositif est partagé, on peut le présenter comme une chose positive, un nouvel espace de rencontre, dans une autre dimension des couplages sensori-moteurs. Pour cela, il me semble que le mieux est de se placer d'emblée dans un mode de couplage qui habituellement n'est accessible à aucun des participants, par exemple dans une perception infra-rouge. Le jeu des interactions ne ressemblera plus à un colin-maillard quelque peu sadique, mais à un jeu d'interaction horizontale où chacun devra chercher à voir les modifications que l'autre apporte au monde partagé. On cherchera alors, de concert, à constituer des significations partagées.

Je cherche actuellement à concilier les deux perspectives présentées ci-dessus à par une théorie du rôle de la mémoire dans la constitution de l'esthétique.

BIBLIOGRAPHIE.-

ALOIMONOS (Y.) 1993, Active Perception, L.E. Hillsdale, NJ.

ANTON J-L, Benali H, Guigon E, Di Paola M, Bittoun J, Jolivet O, Burnod Y (1995)

"Comparison of spatial extent and overlap of activations induced by different manual tasks in human primary sensorimotor cortex". Soumis pour publication.

BACH Y RITA (P.) 1994, The brain beyond the synapse: A review. NeuroReport, 5, 1553-1557.

- BACH Y RITA (P.) 1995, Nonsynaptic Diffusion Neurotransmission and Late Brain Reorganization. New York: Demos Pub.
- BERMUDEZ, MARCEL, EILAN (Eds) 1995, The body and the self, MIT press, Bradford Books.
- BLAKE (A.) & YUILLE (A.) 1992, Active Vision, MIT Press.
- COLLINS (C.C.) and BACH Y RITA (P.) 1973, Transmission of Pictorial Information Through the Skin", *Advan. Biologic. Med. Phys*, 14, 285-315.
- GAZZANIGA (Ed.) 1995, The cognitive neuroscience , MIT press.
- HEUER AND KEELE (Eds.), Handbook of perception and action, Academic Press.
- JEANNEROD (M.) (ed.) 1990, Attention and performance, Motor representation and control, vol.13, Hillsdale, LEA.
- PORT (R.F.) & Van GELDER (T.) (Eds.) 1995, Mind as Motion, MIT Press.
- WHITE (B.W.), SAUNDERS (F.A.), SCADDEN (L.), BACH Y RITA (P.) and COLLINS (C.) 1970, "Seeing with the skin." Perception & Psychophysics, 7, 23-27.

Apprentissage de la commande de robots : le rôle de la boucle sensori-motrice

Dominique Luzeaux*
luzeaux@etca.fr

* Laboratoire Système de Perception

ETCA/CREA/SP, 16bis, av. Prieur de la Côte d'Or, 94114 Arcueil Cedex, France

1 Introduction

La robotique est par essence un domaine traitant de systèmes complexes, exhibant les diverses caractéristiques de complexité pouvant être attendues : la complexité organisationnelle, mais aussi la complexité interactionnelle. En effet, d'une part elle met en jeu des objets complexes du point de vue de leur structure, les robots, alliant et reliant en une même architecture diverses ressources relatives à l'énergie, la motricité, la perception et la mission. D'autre part, du fait même de la conception téléologique des robots, ces derniers n'évoluent pas comme systèmes isolés, mais à l'intérieur d'un environnement avec lequel ils interagissent par leur capacité motrice et par suite de la mission qui définit leur existence. Ceci sans tomber dans un environnementalisme exacerbé qui méconnaîtrait la capacité auto-organisatrice de certains organismes dans leurs interactions avec l'écosystème (nous référons ici à des travaux comme ceux de H. von Foerster ou H. Maturana [16], qui n'attribuent pas à l'environnement le rôle premier dans l'évolution d'un système, mais plutôt un rôle de révélateur ou de « déterminant » d'une structure préexistante).

Cet aspect interactionnel est fondamental, quel que soit le point de vue sur le caractère prééminent des divers acteurs, et il se traduit évidemment par un échange, une manipulation d'information. Dans la suite nous allons nous intéresser à un aspect particulier de cet échange, à savoir sa réalisation par la boucle sensori-motrice. Nous nous focaliserons plus précisément sur la raison fondamentale de ce bouclage, en insistant sur le fait que l'interaction sensori-motrice n'est ni *unidirectionnelle* des capteurs vers les actionneurs, ni *séquentielle*, c'est-à-dire qu'il convient de considérer les ressources liées à la perception et celles liées à l'action, non comme deux éléments ayant des rôles a priori distincts, mais dans leur globalité. Nous verrons notamment comment le deuxième point intervient de manière cruciale dans la commande de robots, puis dans l'apprentissage de la commande de robots. Plus précisément, dans le cadre d'une méthodologie que nous avons développée pour permettre à un robot d'apprendre à remplir des missions prédéfinies (par exemple : survivre, rejoindre un point ou une région, s'asservir autour d'un trajectoire...) dans un environnement inconnu a priori – notons qu'il n'y a aucune incohérence à avoir une mission prédéfinie et un environnement inconnu : il suffit d'imaginer un robot mobile face à un terrain quelconque et une forêt à 300 mètres par exemple, et de lui donner comme mission de rejoindre la lisière de la forêt ; il est alors clair que l'énoncé de la mission n'est pas des plus informatifs quant à la structure de l'environnement pour assurer la progression dudit robot –, nous nous intéresserons au rôle de l'interaction sensori-motrice pour *identifier* un modèle *comportemental* du robot, ainsi qu'un modèle topologique de l'espace d'état du robot (incluant une *représentation* de l'environnement), puis nous montrerons comment à partir de cette première *catégorisation*, il est effectivement possible de résoudre la tâche désirée.

Pour mener à bien cette entreprise, qui vise à montrer le rôle primordial de l'interaction sensori-motrice dans l'organisation d'un système robotisé – révélant ou révélée par l'environne-

ment –, nous insisterons sur les 2 points suivants, servant d'hypothèses de base à notre méthodologie :

- une approche *réursive*¹ (à opposer à une démarche purement adaptative) ;
- une *représentation unique* issue de la théorie des *systèmes dynamiques* intégrant l'évolution du système et un modèle de l'environnement.

Le deuxième point est la clé de voûte de la méthodologie et apporte, comme nous le verrons, des éléments de réponse quant au problème du fondement symbolique (« symbol grounding problem ») et au problème continu/discret, pierre d'achoppement des techniques ascendantes et descendantes. Notre approche, guidée par l'automatique, permet en effet le passage naturel à une discrétisation du problème, cette dernière étant fondée sur la dynamique du système au sein de l'environnement.

2 Boucle sensori-motrice : de l'importance de « motrice »

Avant toutes choses, fixons le cadre de la discussion : nous considérerons dans la suite un robot évoluant dans un environnement, c'est-à-dire, dans des termes empruntés à l'automatique, un système avec des entrées et des sorties, connecté avec un autre système. En fait, nous serons limité dans la nature des environnements : il s'agira de l'espace, tel que nous le vivons quotidiennement, soit par notre expérience directe, soit indirectement par l'expérience d'autrui. Citons à titre d'exemple les environnements structurés : intérieurs de laboratoire, autoroutes, stations de métro, zones urbaines... mais n'oublions pas non plus d'autres environnements rentrant dans la sphère d'intérêt de la robotique : milieu aquatique, milieu terrestre tout terrain, milieu aérien, milieu spatial.

La diversité de ces environnements implique évidemment une diversité quant aux missions que l'on exigera du robot, d'où une nécessité de reconnaître « le même dans le différent, la permanence dans le changement », si l'on désire tirer parti lors d'une nouvelle expérience de réussites passées. Se posent déjà en toute généralité plusieurs problèmes liés au robot, à l'environnement, à l'interaction entre les deux, et à la mission assignée au robot. En particulier, trouver une *représentation commune*, où il sera facile d'exprimer d'une part des contraintes liées à la dynamique du robot, mais aussi à l'environnement, et d'autre part celles relatives à la mission, afin de faire dialoguer efficacement ces différentes composantes. Ceci nous guidera ultérieurement, notamment pour passer de l'identification comportementale du robot et de la représentation topologique de l'espace d'évolution à une solution permettant de remplir la mission prédéfinie.

Ensuite, il s'agit de voir comment s'opère l'interaction entre le robot et l'environnement, et de quelle manière l'un influe sur l'autre et réciproquement : c'est la boucle sensori-motrice qui est au cœur de cette problématique, le robot appréhendant l'environnement au moins par l'intermédiaire de ses capteurs mais aussi de ses actionneurs, comme nous le soulignerons, et agissant sur l'environnement par ses actionneurs. La question que nous abordons est de savoir comment se fait l'appréhension de l'environnement. D'un point de vue psychologique, il est admis que l'« espace perceptif » – antérieur à toute notion d'espace représenté – est un « espace pratique et égocentrique », selon l'expression de Piaget ; pour l'enfant par exemple, l'espace est un milieu structuré lié à son activité. L'élaboration de représentations n'intervient que plus tard, et le passage de l'espace perceptif à l'espace représenté se fait par l'intermédiaire de la motricité. Le rôle capital de ces éléments moteurs a été souligné par de nombreuses expériences dans la perception de la troisième dimension [4]. En ce qui nous concerne, nous insistons sur le fait que la représentation de l'espace, à partir de laquelle le robot devient *situé*, ne s'élabore pas

1. Nous préférons « réursive » à « itérative », car nous n'avons pas a priori de contrôle sur la longueur des boucles ou itérations liées aux expériences.

simplement par l'intermédiaire des capteurs, mais aussi par une utilisation intelligente – au sens étymologique du terme, c'est-à-dire qui comprend, qui saisit globalement – des éléments moteurs, des actionneurs.

Citons à titre d'illustration un exemple du règne animal [22] : l'araignée *Cupiennius Salei*, originaire d'Amérique Centrale, étudiée par Barth et Seyfarth entre autres, s'oriente en intégrant (au sens mathématique du terme) la position et la tension musculaire de ses tibias grâce à ses organes lyriformes. C'est ce mécanisme, basé uniquement sur ses actionneurs de mouvement, qui lui permet de se déplacer entre une proie et son abri, ou d'échapper à une menace et de revenir ensuite vers sa proie.

Ce rôle important que nous faisons jouer aux actionneurs dans l'élaboration de la représentation de l'environnement est également justifié par des considérations théoriques comme celles de l'automatique au sujet de la commandabilité et de l'observabilité (voir par exemple [8] ou [20] pour les définitions de ces notions de base d'automatique) où la considération *simultanée* des entrées et des sorties, donc une vision globale de la boucle sensori-motrice dans la phase d'acquisition de l'information lors de l'interaction entre robot et environnement, se révèle cruciale dans certaines situations et permet notamment de lever certaines difficultés liées au choix pertinent du capteur.

Nous verrons même d'ailleurs à ce propos, dans un paragraphe suivant, comment on peut, à partir de la simple observation de l'enchaînement des actions entreprises sur les actionneurs, reconstruire un certain modèle de l'interaction entre le robot et son environnement. Ce point de vue radicalement opposé aux approches précitées est évidemment un peu excessif, comme nous le soulignerons, mais il permet de montrer le rôle *constructif* que peut aussi jouer la motricité dans l'élaboration d'une représentation de l'environnement, voire d'une catégorisation de ce dernier. Nous nous opposons donc radicalement à des prises de position comme celles de Smithers [21] : « We might talk about [the robot] using information about its motor actions instead, but this would not seem to offer much of an explanation for how it gets to know anything about its environment by this means. (...) To go on to say that it is by this means that the robot picks up the required information from its world for it to form a map seems to be verging on the fanciful. It's certainly unnecessarily convoluted as an explanation of how the robot does what it does. »

3 Identification de dynamiques et catégorisation

Dans cette partie, nous allons nous intéresser à l'apprentissage de la commande de robots. Nous présentons notre méthodologie, détaillée dans [2, 9, 11], en nous focalisant sur trois points :

- le choix d'une représentation commune pour la dynamique du robot et l'interaction sensori-motrice ;
- la catégorisation ;
- l'aspect récursif de la méthodologie.

Le dernier point est important, car il permet potentiellement des apprentissages de plus haut niveau, et découle dans une certaine mesure du premier, car le choix d'une représentation unifiée va permettre de réappliquer, via un codage simple, la même méthodologie sur des données de sémantique plus complexe.

Nous avons illustré la méthodologie à l'aide de plusieurs exemples de difficulté croissante : déplacement d'un robot holonome parmi des obstacles (mission : survie), conduite en simulation de véhicule sur un circuit (mission : boucler des tours de circuit dans un premier temps, puis optimisation du temps de parcours), et commande en vraie grandeur d'un robot mobile, de type voiture, muni d'une caméra, de télémètres à ultrasons et d'odomètres [12, 13] (mission : manœuvre de parking en présence d'obstacles).

3.1 Le schéma de commande dans la boucle sensori-motrice

Il est nécessaire de formaliser l'interaction sensori-motrice si nous voulons être en mesure de l'utiliser dans une méthodologie quelconque. Afin de proposer une représentation de cette interaction, nous nous inspirerons des constatations psychophysiques faites sur la notion de sensation et de ses rapports avec les autres faits psychiques antécédents ou concomitants. En particulier, la *loi de relation* énoncée par Höffding, également appelée loi de contraste, s'interprète en disant que nous sommes surtout sensibles aux différences ou aux rapports, à ce qui change, à ce qui varie, plutôt qu'à des valeurs absolues [4]. En fait, cette loi empirique fait suite à l'énoncé au cours des siècles de diverses lois de la sensation, partant d'une recherche des seuils d'excitabilité, et évoluant, suite aux difficultés de trouver des valeurs pertinentes et répétitives, vers des lois différentielles, où la relation entre l'excitant physique et la sensation n'est plus en intensité, mais en qualité.

Fort de cette constatation, nous proposons donc d'observer prioritairement les variations des diverses mesures issues des capteurs afin de construire toute représentation de l'environnement, et également de commander le robot au niveau de ses actionneurs par une même notion de variation apportée à la valeur courante de l'actionneur.

Nous pouvons alors définir dans un premier temps un schéma de commande comme un couple actionneur/capteur, où les variations de l'actionneur sont pilotées par les variations du capteur. L'aspect dynamique de la chose apparaît quand on introduit la notion d'*espace de phase* utilisée pour interpréter tout système dynamique.

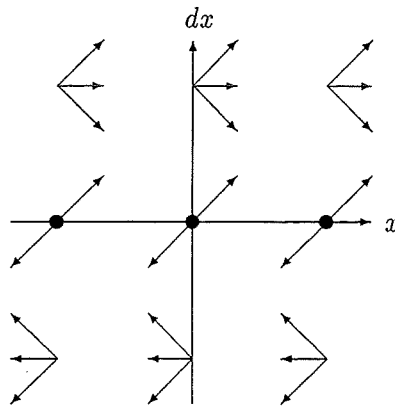


FIG. 1 – Germes de trajectoires dans l'espace de phase.

Si l'on considère par exemple un système où l'entrée u a un comportement de dérivée par rapport à une sortie x (c'est-à-dire $u = f(\frac{dx}{dt})$, f étant une fonction croissante), et que l'on dessine dans le plan $(x, \frac{dx}{dt})$ les germes de trajectoires issus d'un point de ce plan, on remarque que ces derniers sont directement liés à la dérivée seconde $\frac{d^2x}{dt^2}$, et par l'hypothèse entre u et x , $\frac{d^2x}{dt^2}$ est de même signe que $\frac{du}{dt}$, donc les variations de l'entrée déterminent le germe local de la trajectoire. Adoptant maintenant une lecture globale, on conçoit comment on peut, en fonction de la position courante de la sortie dans l'espace de phase, choisir la variation de l'entrée pour « tendre » vers une sortie désirée (certains germes deviennent admissibles ou non selon le domaine de l'espace de phase), et ceci permet de définir des comportements très différents [10, 12, 18], comme l'illustre la figure 2 dans l'exemple simple d'un robot piloté indépendamment en vitesse horizontalement et verticalement.

Évidemment le raisonnement a été mené ici dans le cas d'un modèle du second ordre², mais

2. En effet, au sens de l'automatique, la vraie entrée que l'on gouverne est $\frac{du}{dt} = v$ et on a considéré un système où v est exprimé comme fonction de $\frac{d^2x}{dt^2}$.

il peut être mené quelle que soit la dimension, il suffit d'augmenter ou de diminuer le nombre de dérivées de x qui interviennent. Remarquons que dans le cas du premier ordre, on retrouve le modèle classique (un peu simple du point de vue de la dynamique des systèmes...) de la commande attractive-répulsive.

On peut brièvement se poser la question de la généralité de ce modèle ; pour simplifier la discussion, remplaçons dans la discussion précédente la dépendance d'une fonction croissante par une simple égalité. Un modèle d'ordre n revient alors à modéliser le système entrées-sorties considéré par un ensemble d'équations $\dot{x}_1 = x_2, \dot{x}_2 = x_3, \dots, \dot{x}_n = v$; on parle alors de forme chaînée, et l'on sait qu'une très large classe de systèmes non linéaires peut se ramener, sous certaines conditions de rang, à cette forme [5, 8]. La prudence est cependant de rigueur : si le résultat mathématique justifie la généralité du modèle, il n'en définit pas moins la nécessité d'un changement de variables (la plupart du temps non linéaire, portant à la fois sur les entrées et les sorties), ce qui, en théorie, signifie que le modèle est exact si les *bons* capteurs et les *bons* actionneurs sont disponibles. Dans le cas contraire, ce modèle n'est qu'approximatif (voire invalide dans certains cas), mais précisons cependant – sans y trouver une excuse commode – que de nombreuses prouesses techniques sont réalisées en prenant de telles approximations.

Sans rentrer dans les détails (se référer à [3, 9, 10, 11]) rappelons comment le modèle comportemental précédent peut être identifié par *apprentissage par découverte*. Le système a pour seules données accessibles les valeurs de ses entrées et de ses sorties au cours du temps. N'ayant initialement pas de modèle a priori de sa dynamique, il applique, au niveau de ses actionneurs, des signaux particuliers (échelons, rampes, sinusoides) paramétrés aléatoirement. Sont ainsi stockées des valeurs échantillonnées pour chaque variable (qu'elle corresponde à une entrée ou une sortie) au cours des diverses expérimentations sanctionnées par un échec ou une réussite (ce dernier cas est exceptionnel au début). Chaque suite de valeurs pour chaque variable et chaque expérimentation est alors codée comme une suite de signes $\uparrow, \downarrow, 0$, selon qu'entre deux instants successifs les valeurs ont cru, déchu ou sont restées constantes. Un modèle d'évolution de chaque variable est alors inféré [14], sous forme d'une expression régulière, et les différentes expressions modélisant les diverses variables sont ensuite comparées entre elles. Cette phase permet alors de déduire un modèle qualitatif du système (au sens de la physique qualitative [6, 7, 23]) avec des relations entre variables du type : monotonie croissante ou décroissante, dérivée positive ou négative d'ordre fixé...

Il est à noter que la procédure ci-dessus n'est qu'une solution parmi d'autres pour « apprendre » un modèle dynamique du système, mais elle a l'avantage d'être non supervisée, computationnellement raisonnable (très facilement parallélisable de surcroît), robuste aux erreurs d'observation (démonstration dans [14]) et de s'interfacer naturellement avec la représentation dans l'espace de phase.

Il reste un point non encore déterminé dans le schéma : l'origine de l'espace de phase. C'est cette origine, liée au comportement dynamique choisi, en autorisant selon le domaine tel ou tel germe, qui déterminera complètement le comportement. Un choix naturel est une valeur particulière pour la première composante (obtenue typiquement lors de la représentation topologique de l'environnement, voir ci-dessous) et une valeur nulle pour les autres. C'est le principe de l'asservissement vers une valeur de consigne désirée. Mais on peut tout aussi bien imaginer des contraintes numériques différentes sur les autres composantes, par exemple si l'on veut rejoindre un point donné avec une vitesse imposée.

C'est cette versatilité qui peut d'ailleurs permettre ultérieurement des apprentissages plus complexes : dans un premier temps, on trouve des valeurs de consigne particulières, dépendant éventuellement de la représentation topologique de l'environnement, puis dans le cadre d'une *approche récursive*, on considère ces valeurs de consigne comme les nouvelles entrées, et on cherche à les piloter en fonction des sorties ; ceci autorise une planification non plus statique des comportements, mais dynamique, et une complexité comportementale évidemment accrue. En

fait, cela revient à passer de l'asservissement ponctuel à un asservissement autour de trajectoire, cette dernière étant construite par l'observation originelle de l'interaction sensori-motrice.

Pour conclure, le schéma de commande est un couple de deux paramètres, un certain ordre reliant les dérivées correspondantes d'un des deux paramètres à l'autre, une politique de choix des germes de trajectoire dans l'espace de phase correspondant et une origine de cet espace.

3.2 Représentation topologique et catégorisation

Abordons à présent la notion de représentation topologique de l'environnement, maintenant que l'identification de la dynamique a été menée. Plusieurs solutions sont possibles. Dans [10] et [11], nous nous plaçons dans un cadre de viabilité, c'est-à-dire que le robot se doit de survivre dans un environnement donné : évoluer au sein des obstacles sans les heurter, et rejoindre éventuellement, selon les missions, un point dans l'espace ; c'est l'exemple très classique du robot dans une pièce encombrée qui doit se diriger vers un objet ou une pièce voisine. Nous partons alors de l'idée que l'environnement va se traduire dans l'interaction sensori-motrice lors des échecs (le robot heurte un obstacle), ou si l'on désire un apprentissage en ligne non destructif, lors du dépassement d'une valeur de sécurité fixée a priori par l'expérimentateur sur le capteur adéquat (on voit ici le risque d'interférence entre le processus d'apprentissage de la représentation et l'expérimentateur extérieur au robot, qui pourrait, par sécurité, fournir en fait une information telle que la représentation de l'environnement serait explicite dans les capacités du robot). Si le robot est en situation d'échec, et s'il le sait, il peut par introspection analyser ses entrées et sorties passées pour déterminer ce qui a motivé cet échec.

Le mécanisme que nous proposons consiste à rechercher les paramètres (entrées ou sorties, car rappelons que nous faisons jouer un rôle tout aussi important aux deux) ayant pris une valeur extrême lors de l'échec, puis à déterminer statistiquement au fur et à mesure des expériences des intervalles de confiance au sein desquels le robot estimera être hors de danger, et pourra alors adopter un comportement lié à sa mission ; en dehors de ces intervalles de confiance, un comportement d'évitement pour le paramètre identifié sera adopté, en suivant une démarche similaire à celle décrite dans la partie précédente, si ce n'est que cette fois il ne s'agit pas de tendre vers l'origine de l'espace de phase associé, mais de s'en éloigner.

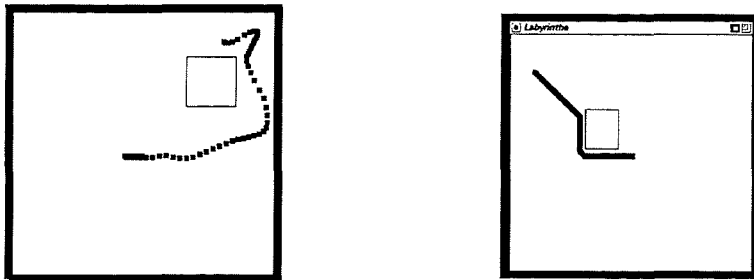


FIG. 2 – *Apprentissage de l'évitement d'obstacles avec divers schémas de commande (sur la figure de gauche, la position initiale est en haut à droite et la vitesse initiale dirigée vers le bas ; sur la figure de droite, la position initiale est en haut à gauche et le robot au repos).*

Une autre méthode, développée dans [18], consiste à partitionner en pavés l'espace d'évolution engendré par tous les paramètres, entrées et sorties, et à chaque élément de la partition pourra alors être associé un comportement élémentaire différent, en particulier en changeant les valeurs de consigne dans les espaces de phase associés. Cette partition a pour objet d'éviter que les domaines explorés au fur et à mesure des évolutions ne contiennent des points d'échec. Cela conduit à cantonner ces points aux frontières des pavés. Lorsque le système franchit la frontière entre

deux pavés, il s'expose donc à rencontrer un obstacle. Par conséquent, il est crucial de signaler à proximité des frontières les zones rendues infranchissables par la présence d'un obstacle. Dans ce but, à chaque fois qu'un échec se produit dans un pavé, on associe à ce point d'échec une zone de sécurité définie comme un pavé dont le centre est le point d'échec, et un coin le point antérieur à l'échec dans l'évolution. Lorsque le système pénètre dans une zone de sécurité, un comportement le repoussera vers le centre du pavé dont il vient, afin que seules soient franchies les frontières le long desquelles aucun obstacle n'a été encore rencontré. Pour résumer, à l'initialisation, la représentation de l'espace d'évolution est constituée d'un seul sous-espace qui s'étend à l'infini dans toutes les dimensions. Le domaine exploré associé à ce sous-espace initial est vide, mais croît au fur et à mesure des évolutions. Lorsqu'apparaît un point d'échec interne au domaine exploré, deux cas se présentent: si le point d'échec est dans une zone de sécurité, on l'ignore; sinon, on segmente chacune des dimensions du sous-espace initial en deux intervalles autour de ce point d'échec, ce qui constitue donc 2^n sous-espaces, si n est le nombre de dimensions de l'espace d'évolution. Dans chaque pavé ainsi obtenu, on recalcule le domaine considéré comme viable en fonction des évolutions qu'il contient (par exemple en prenant les valeurs minimale et maximale dans la dimension considérée). Par ailleurs, si le point d'échec n'est pas ignoré, on recalcule la zone de sécurité associée, et si celle-ci en intersecte une autre, on les remplace toutes les deux par la zone de sécurité englobante. Dans chacun des sous-espaces obtenus, on effectue récursivement le même traitement que dans le sous-espace initial: croissance du domaine exploré local autour des évolutions, jusqu'à ce qu'un point d'échec interne indique la nécessité d'une nouvelle segmentation. Lorsqu'une frontière entre pavés est franchie par une évolution sans rencontrer d'obstacles, cette frontière est supprimée si le résultat de la suppression donne un pavé.

Une fois cette partition apprise, on peut naturellement lui associer un graphe d'accessibilité, et la commande du système se décompose à deux niveaux, en une planification haut-niveau, et une commande bas-niveau à l'intérieur d'un pavé pour rejoindre éventuellement le successeur. Il est intéressant de constater que cette hiérarchisation haut/bas niveau n'est qu'un artifice de langage, car le plan est réévalué à chaque instant, et la commande prévaut le plus souvent sur la planification [18]. En fait, la représentation acquise de l'environnement permet de rajouter un caractère global a priori absent de la représentation sous forme d'espace de phase, tout au moins sous la forme utilisée dans les paragraphes précédents. En effet on peut tout à fait incorporer la représentation topologique de l'environnement dans l'espace de phase en le partitionnant au lieu de ne considérer que l'origine comme référence numérique. Ceci permet d'obtenir des dynamiques complexes comme dans le figure 3 où, selon la région de l'espace de phase, divers schémas de commande sont activés.

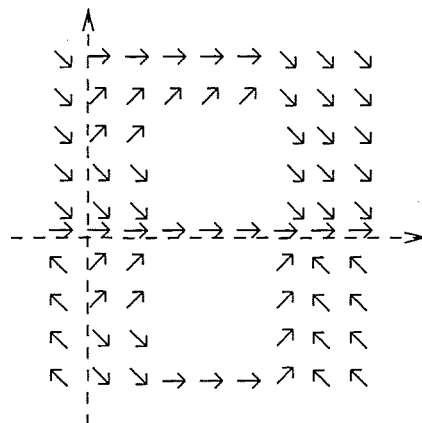


FIG. 3 – Dynamique dans l'espace de phase avec deux points attracteurs, deux obstacles à éviter, et un flot global de gauche à droite.

Le point à retenir de cette discussion est que l'espace de phase permet d'intégrer naturellement la dynamique du système et une représentation de l'environnement, ainsi qu'une expression des contraintes liées à la mission en ajoutant éventuellement de nouveaux points ou régions de référence dans l'espace de phase.

Cet apprentissage d'une représentation de l'environnement, associé à l'identification de la dynamique décrite précédemment, a permis notamment d'apprendre à commander un pendule inversé, à déplacer un robot holonome parmi des obstacles pour rejoindre des points donnés, à réaliser la manœuvre de parking pour un robot mobile en utilisant l'odométrie et la télémétrie à ultrasons, à piloter une petite voiture sur un circuit constitué de lignes de droites, de virages plus ou moins raides et de chicanes.

Dans le cadre de cette dernière application, le robot mobile ayant appris à boucler des tours de circuit, on peut se poser la question de savoir si l'apprentissage sur un circuit particulier peut servir ultérieurement. À ce niveau, il convient de remarquer que la représentation précédente est élaborée par des critères extéroceptifs : la présence d'un obstacle dans l'environnement du système. Une autre représentation fondée sur des critères proprioceptifs cette fois, liés au comportement du système, semble nécessaire [17, 19]. La figure 4 montre d'un côté le circuit, et de l'autre la dérive, c'est-à-dire l'écart en position par rapport au milieu de la route, tel qu'il est obtenu suite à une trajectoire générée par une commande apprise à partir des méthodes exposées précédemment.

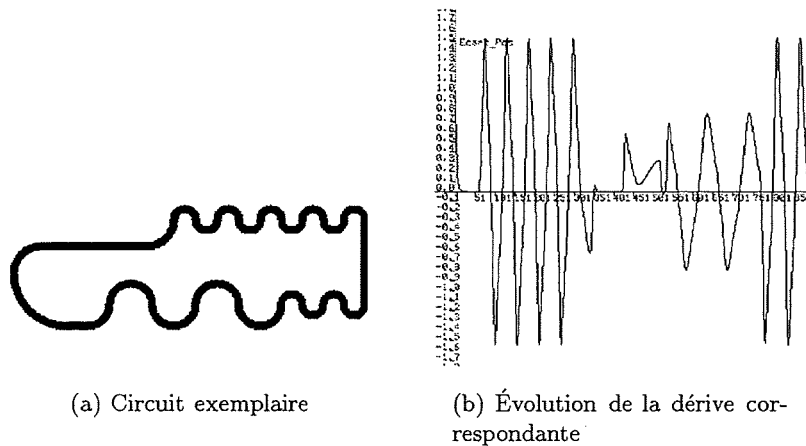


FIG. 4 – La dérive de la voiture permettant de catégoriser l'environnement.

On voit alors qu'une segmentation du comportement qualitatif du paramètre de dérive permet de trouver les trois catégories de ligne droite, virage à droite et virage à gauche, voire, par seuillage, les virages serrés des autres.

Dans [9], nous avons aussi suggéré d'étudier les suites des ordres déclenchés sur les actionneurs lors de la commande du robot pilotant sur un circuit, et une analyse syntaxique permettait de segmenter ces suites de commandes en motifs correspondant qualitativement aux virages et lignes droites. L'appariement précis entre les changements comportementaux et la topologie de l'environnement était cependant beaucoup plus délicat que dans le cas de la figure 4.

À ce point, on s'aperçoit que d'une part le robot est *situé* dans l'environnement, mais également que l'on dispose d'un moyen de faire *émerger explicitement* des comportements liés à la topologie de l'environnement. De plus un tel comportement est construit à partir des comportements élémentaires, qui sont définis par un schéma de commande et des valeurs de consigne, et peut alors dans l'absolu servir de base à un apprentissage de plus haut niveau symbolique.

4 Conclusion

Nous avons insisté dans les paragraphes précédents sur l'importance d'une approche issue de la théorie de la commande pour parvenir à une vision cohérente de la robotique autonome. En identifiant un modèle dynamique du robot et en adoptant une représentation continue grâce à l'interprétation de cette dynamique au sein de l'espace de phase, nous avons mis en évidence la possibilité d'une catégorisation de l'environnement. Notons au passage que la discrétisation du problème ne se fait pas indépendamment sur les actionneurs et les capteurs, comme c'est le cas dans les formalisations classiques par agents, où les comportements sont typiquement définis comme une association entre une représentation construite à partir des capteurs et une partition des actions possibles ; ici l'aspect discret intervient par l'intermédiaire du concept d'espace de phase (qui est d'ailleurs construit de manière étroite avec la représentation topologique), mais ce dernier autorise toujours une multitude de trajectoires, voire de famille de trajectoires.

L'espace de phase n'est pas simplement un moyen de faire émerger certains comportements élémentaires. En effet, si au lieu de construire une représentation de l'environnement et d'utiliser au sein de chaque élément de cette représentation un espace de phase local – déterminant ainsi en quelque sorte un comportement local selon l'environnement –, on partitionne directement l'espace de phase, et on choisit les germes de trajectoire non plus simplement avec le souci d'un asservissement vers l'origine de l'espace, on devient alors en mesure d'obtenir des comportements bien plus complexes.

Cette manière de faire coopérer la boucle sensori-motrice de bas niveau avec des représentations topologiques est très fréquente en automatique, notamment dans la synthèse de lois de commande pour systèmes non linéaires. De telles directions semblent prometteuses (la notion de décomposition d'un comportement global en comportements locaux par l'intermédiaire d'un oracle fondé sur une partition de l'espace d'état a été étudiée dans [1] sur la base de [15]) et permettraient de promouvoir l'idée que l'interaction sensori-motrice n'est pas uniquement le premier niveau nécessaire à l'apparition de mécanismes cognitifs apparemment plus complexes, mais qu'elle peut aussi jouer un rôle dans leur élaboration, voire en faire partie intégrante.

Références

- [1] J.-F. Antonioti. *Formalisation de la notion de comportement élémentaire dans la famille des trajectoires d'un système dynamique*. DEA, Paris VII, 1995.
- [2] B. Burg, D. Luzeaux, and B. Zavidovique. Apprentissage de règles de comportement destinées au contrôle d'un système. In *Journées Françaises de l'Apprentissage n° 4*, St-Malo, France, 1989.
- [3] B. Burg, D. Luzeaux, and B. Zavidovique. Candide: a learning system for process control. In *International Conference on Industrial and Engineering Applications of Artificial Intelligence and Expert Systems*, volume 1, pages 270–7, Tullahoma, TE, USA, 1989. ACM.
- [4] A. Cuvillier. *Cours de philosophie*. Armand Colin, 1954.
- [5] A. Isidori. *Nonlinear control systems*. Springer Verlag, 2nd edition, 1989.
- [6] B. Kuipers. Qualitative simulation. *Artificial Intelligence*, 29, 1986.
- [7] B. Kuipers. *Qualitative Reasoning: Modeling and Simulation with incomplete knowledge*. The MIT Press, 1994.
- [8] W. S. Levine, editor. *The Control Handbook*. CRC Press, 1996.

- [9] D. Luzeaux. *Pour l'apprentissage, une nouvelle approche du contrôle: théorie et pratique*. Doctorat de l'Université Paris XI, Orsay, 1991.
- [10] D. Luzeaux. Let's learn to control a system! In *IEEE International Conference on Systems Man Cybernetics*, Le Touquet, France, 1993.
- [11] D. Luzeaux. Process control and machine learning: rule-based incremental control. *IEEE Transactions on Automatic Control*, 39(6), 1994.
- [12] D. Luzeaux. Rule-based control and learning : application to mobile robotics. In *7th International Conference on Artificial Intelligence and expert systems applications*, San Francisco, CA, USA, 1995.
- [13] D. Luzeaux and S. Meunier. Rule-based incremental control for the parking maneuver of a mobile robot. In *International Conference on Recent Advances in Mechatronics*, Istanbul, Turkey, 1995.
- [14] D. Luzeaux and B. Zavidovique. Robust grammatical inference. In *AFCET, Paris*, 1989.
- [15] E. Martin, D. Luzeaux, and B. Zavidovique. Learning and control from a recursive viewpoint. In *IEEE International Symposium on Intelligent Control*, Glasgow, Scotland, 1992.
- [16] E. Morin and M. Piattelli-Palmarini. *L'unité de l'homme, tome 2: le cerveau humain*. Éditions du Seuil, 1974.
- [17] O. Sigaud. *Apprentissage : de la commande au comportement*. Doctorat de l'Université Paris XI, 1996.
- [18] O. Sigaud and D. Luzeaux. Control and learning: from local actions to global behaviors. In *3rd European Control Conference*, Roma, Italia, 1995.
- [19] O. Sigaud and D. Luzeaux. Learning hierarchical controllers for situated agents. In *Proceedings of the 14th International Congress on Cybernetics*, Bruxelles, Belgium, 1995.
- [20] E. D. Sontag. *Mathematical control theory: deterministic finite dimensional systems*. Springer Verlag, Texts in Applied Mathematics, volume 6, 1990.
- [21] L. Steels and R. Brooks, editors. *The artificial life route to artificial intelligence*. Lawrence Erlbaum Associates, 1995.
- [22] A. Teyssèdre. *L'orientation des animaux: méthodes et mécanismes*. Nathan, collection Science et Nature, 1996.
- [23] D.S. Weld and J.H. Kleer, editors. *Readings in Qualitative Reasoning about Physical Systems*. Morgan Kaufmann, 1990.

Problèmes liés à l'identité dans la construction de Systèmes à Base de Connaissances

Jean-Marc Fouet

Laboratoire d'Ingénierie des Systèmes d'Information

Université Claude Bernard et INSA de Lyon

43 Bd du 11 Novembre 1918

69622 Villeurbanne Cedex

tel : 04 72 44 83 69

fax : 04 72 43 15 36

email : fouet@lisisun.univ-lyon1.fr

URL : <<http://www710.univ-lyon1.fr/~fouet>>

Résumé

Le processus d'identification, s'il est indispensable, a toutefois des inconvénients. Dans la mesure où il permet à deux personnes de construire des ensembles intensionnels similaires à partir d'ensembles extensionnels différents, il est l'un des piliers de la communication ; mais en revanche, le fait que ces deux personnes croient que leurs deux ensembles sont identiques est à la source des difficultés de communication.

Le présent article a pour objectif de montrer que, lorsqu'on remplace (pour quelque raison que ce soit) l'une de ces deux personnes par un Système à Base de Connaissances en cours d'élaboration, le problème peut devenir crucial, mais aussi susciter des éléments intéressants de solution.

Après avoir brièvement rappelé en quoi un SBC se distingue d'un logiciel algorithmique, on analyse l'impact de l'identité sur les trois aspects des SBC : recueil, stockage, et exploitation des connaissances. La conclusion reviendra sur la confusion ambiguë que l'article entretient entre "identité" et "identification".

Abstract

The process of identifying, although necessary, has its drawbacks. Inasmuch as it allows two persons to build similar intensional sets from different extensional sets, it constitutes one of the foundations of communication ; but on the other hand, the fact that people tend to believe their sets to be identical explains most communicating difficulties.

This paper aims at showing that when one of these people is -for any reason- replaced with a Knowledge Based System under construction, the problem may become critical, but may also suggest interesting elements of solution.

After having briefly summarised how a KBS differs from an algorithm, we analyse how identity impacts against the three aspects of KBSs : acquisition, storage and working of knowledge. To conclude, we come back on the ambiguous confusion this article maintains between "identity" and "identification".

1. Opposition entre Algorithme et SBC

Considérons un algorithme a . Lorsqu'on lui soumet un jeu de données d , représentant un problème p , a fournit une solution s (exacte, approchée, ou fausse) de p . L'ensemble D de tous les jeux de données admissibles par a définit l'ensemble P de tous les problèmes que a peut traiter.

Par exemple, si a est un algorithme de tri, d sera constitué d'un entier positif n , inférieur à une limite N fixée par l'auteur de l'algorithme, et d'un vecteur de n nombres, obéissant à des règles définies par l'auteur (par exemple, flottants simple précision). D sera l'ensemble de tous les couples $(n, \text{vecteur de longueur } n)$; P sera "le problème du tri d'un vecteur de flottants de longueur inférieure à N ". De ce point de vue, tous les p sont donc identifiés en P .

Si l'on implémente a sur un ordinateur et qu'on lui soumet successivement :

- un vecteur déjà dans le bon ordre ;
- un vecteur dans un ordre aléatoire ;

- un vecteur exactement dans l'ordre inverse de celui que l'on veut,
les performances de a pourront étonner l'observateur naïf ; il se pourrait en particulier que le temps mis par a soit maximal dans le troisième cas, voire même supérieur au temps nécessaire à l'observateur pour résoudre ce problème "à la main".

Dans cette hypothèse, la supériorité de l'observateur sur l'algorithme vient de ce que, pour lui, le problème posé n'appartient pas à la classe P .

La situation au 40^{ème} coup de la partie d'échecs jouée en 1914 par X contre Y n'a absolument rien à voir avec la situation du 10^{ème} coup de Z contre T hier ; en regroupant toutes ces situations sous le nom de "problème de jouer un coup aux échecs", c'est à dire en les identifiant, on aboutit à l'impasse d'un algorithme de complexité exponentielle. A contrario, un bon joueur ne réfléchira que quelques secondes dans chaque cas, mais suivant un chemin très différent dans chaque cas.

Les Systèmes à Base de Connaissances ont précisément pour objet de résoudre des problèmes dont le regroupement en une classe P n'est pas légitime, et conduit à un algorithme inefficace¹.

La question est donc maintenant de recueillir, de stocker et d'exploiter les connaissances permettant de résoudre des problèmes plus vite qu'avec un algorithme.

2. Recueil de connaissances

Considérons un conducteur parisien expérimenté. Demandons lui comment il fait pour traverser la Place de l'Etoile à 18h. Il est incapable de nous l'expliquer, parce que *il n'y a pas* de "Place de l'Etoile à 18h" : il y a chaque soir une situation différente, pour ne pas dire une situation différente à chaque seconde. Donc accompagnons le un soir, et demandons lui de nous dire ce qu'il fait et pourquoi... Il est probable qu'au bout de quelques secondes nous préférons qu'il ne cherche plus à verbaliser. Tournons-nous vers le moniteur d'auto-école : il est, lui, capable de nous dire tout ce qu'il faut faire, mais si nous le plaçons à 18h Place de l'Etoile, force est de reconnaître qu'il n'est pas très bon, ou qu'il ne fait pas ce qu'il a dit. Comment, dans ces conditions, donner à un robot les connaissances nécessaires pour qu'il puisse piloter une automobile à 18h Place de l'Etoile? La réponse actuelle de la plupart des chercheurs en Intelligence Artificielle est : on ne peut pas les lui donner, il faut qu'il les découvre [Pitrat 90].

Et pourtant, le problème est simple : nous avons n véhicules... Non! La Jaguar grise, là, il a peur pour sa peinture, il va freiner ; le livreur de pizza, il ne m'a probablement pas vu ; GTI rouge, méfiance ; la 104 qui était à droite est maintenant à gauche, c'est un slalomeur...

En d'autres termes, il convient de distinguer plusieurs niveaux d'abstraction : identifier tous les véhicules serait une erreur ; par contre, identifier les GTI rouges comme potentiellement dangereuses est en général justifié et salutaire.

Ce qui fait la force de l'expert d'un domaine, c'est précisément qu'il n'identifie pas au même niveau que nous ; en contrepartie, c'est ce qui va rendre délicate l'extraction de ses connaissances ; car, comme nous, il n'est conscient ni des processus qu'il met en oeuvre pour identifier, ni du résultat de ses identifications. Tout au plus est-il capable de rejeter les identifications que nous lui proposons pour montrer que nous croyons avoir compris. ("Non, ce n'est pas si simple". "Attendez, vous mélangez tout"...).

Mais les experts ont également leurs faiblesses, en ce qu'ils identifient des situations dont l'expérience montre qu'elles sont fondamentalement différentes. Les roulements des pompes des premières fusées à propulsion liquide ont été calculés en appliquant des lois universellement admises, mais en n'ayant pas conscience que ces lois étaient valables en présence de lubrification, et étaient totalement fausses dans le vide. Le premier tir

¹Et non pas, absolument pas, de modéliser les processus cognitifs.

d'Ariane 5 a été un lamentable échec, parce que quelqu'un avait oublié que la poussée d'Ariane 5 n'avait aucune commune mesure avec celle d'Ariane 4. L'identification résulte de l'abstraction (*abstracto* = j'enlève) des différences ; quand celles-ci sont significatives, l'identification aboutit à une catastrophe. Aussi l'un des rôles de l'Ingénieur de la Connaissance consiste-t-il à faire dire le non dit, à faire préciser le contexte des connaissances, leur domaine de validité [Fouet& 95].

D'autre part, et notamment lorsqu'ils s'expriment par écrit, les détenteurs de la connaissance ont une fâcheuse² tendance à faire des effets de style, et notamment à employer des synonymes et des périphrases pour éviter les répétitions. La détection automatique des synonymes s'appuie sur la recherche de similarités de structures entre les Objets (voir §3 ci-dessous) construits à partir du texte [Nouira& 96]. Dans l'état actuel de nos travaux, elle exige une confirmation humaine. L'automatisation du traitement des périphrases n'est pas actuellement à notre portée.

A contrario, le même mot peut désigner des concepts différents chez des personnes différentes, voire chez le même locuteur ; la détection automatique de tels homonymes s'appuie sur l'accumulation anormale d'attributs au sein d'un même objet.

3. Stockage

3.1. Rappel sur les formalismes actuellement utilisés

Le stockage en machine des connaissances fait largement appel, de nos jours, au couplage de la Logique et des Objets.

Les Objets se caractérisent par leur structuration en hiérarchies. On trouve notamment souvent une hiérarchie d'héritage et une hiérarchie d'instanciation, parfois confondues.

La hiérarchie d'héritage permet de ne pas stocker au niveau de chaque objet des propriétés communes à d'autres objets : il suffit de savoir de quel objet un objet hérite pour (récurivement) retrouver les propriétés voulues. Noter que si l'héritage est multiple (un objet héritant de plusieurs autres), des problèmes de cohérence peuvent se poser.

La hiérarchie d'instanciation permet de mémoriser la filiation entre objets (quel objet a créé quels objets). Elle peut être utile pour décrire, dans la partie logique, un ensemble d'objets en intension (au lieu d'énumérer les objets, on parle de "tous les objets instances de tel objet").

3.2. Utilisation "anti-identitaire" de ce modèle

Une version faible de ce modèle consiste à n'avoir, comme feuilles de ces arbres hiérarchiques, que des objets "concrets", les niveaux supérieurs étant dénommés "classes". Une version forte consiste à ne pas limiter a priori la profondeur de l'arbre. Ainsi, tel "objet concret" pourra lui-même donner naissance à des instances.

Par exemple, l'objet représentant mon véhicule pourra avoir comme instances *mon-véhicule-propre* et *mon-véhicule-sale*. Au lieu d'avoir à remettre à jour une grande quantité de faits chaque fois que je lave mon véhicule, il suffira de désactiver le *sale* et d'activer le *propre*. On résout ainsi élégamment des difficultés lourdes attachées à la Logique Non Monotone, en introduisant la notion de "point de vue", notion qui s'oppose à celle d'identité.

De même, une clef de 13 n'aura pas du tout les mêmes attributs selon qu'elle est décrite par celui qui l'a conçue, celui qui l'a fabriquée, celui qui l'a vendue, un garagiste, le responsable-qualité qui veut savoir pourquoi elle a cassé, ou le policier qui la trouve tachée de sang à côté d'un cadavre.

² pour nous qui cherchons à passer du *mot* de la langue au *terme* d'un langage, cf [Raccah 96]

4. Exploitation

4.1. Inférences

Classiquement, l'exploitation des connaissances consiste à faire appliquer aux Objets les règles énoncées en Logique, par le truchement d'un moteur d'inférence. Le problème essentiel de ce moteur est de trouver à chaque instant quelle est la règle à appliquer. Ici encore, on paie le tribut à l'identification, en considérant toutes les règles sur le même plan. Dans la Machine Gosseyn [Fouet 93], on renverse la proposition, en dotant chaque règle de son mécanisme d'inférence, voire même en la transformant en réflexe associé à certains objets, ou en en déduisant une modification de la hiérarchie [Medjahed 96]. Ainsi, pour prendre un exemple trivial, la règle "tout homme est mortel" peut-elle se traduire par le fait que la classe des humains hérite de celle des mortels. Le coût d'application de cette règle est alors ramené à celui du mécanisme d'héritage.

Autre exemple : la règle "si un objet x est lié à un objet y et si la position de x est z, alors la position de y est z" peut être traduite par un réflexe "si-besoin" attaché à la position de y ; on réalise ainsi une "évaluation paresseuse" qui permet des gains de temps de calcul quasi-infinis (puisque l'on ne calcule que si on a besoin du résultat, au lieu de tenir constamment à jour).

En d'autres termes la notion de "règle", si elle facilite le recueil des connaissances, doit se voir substituée des notions plus fines pour une exploitation efficace ; on en revient au fait que deux objets (ici les règles) identiques pour le profane sont différents pour celui qui s'en sert. Et une même règle sera transformée différemment suivant l'utilisation qui en est régulièrement faite.

4.2. Interfaces

Un autre méfait de l'identification conduit les auteurs de logiciels à construire une interface pour "l'utilisateur type". Il est préférable au contraire de considérer que tous les utilisateurs sont différents, et varient dans le temps. Les Interfaces Auto-Adaptatives [Berthomé 95] observent les utilisateurs et se modifient en conséquence (présentation de l'écran, forme des questions, réponses par défaut, procédures personnalisées). Dans un sens, cela revient à distinguer l'identité de chaque utilisateur, auquel est associé un fichier résumant son histoire et ses aptitudes, sorte de "carte d'identité". Mais, au fur et à mesure que l'utilisateur interagit avec l'application, que progresse l'appropriation mutuelle, ce fichier est modifié (et l'interface se modifie en conséquence) ; l'"identité" de l'utilisateur, loin d'être un invariant, est donc pour nous une fonction du temps³.

Conclusion

Pour transmettre des connaissances à mes étudiants, je joue sur deux registres : le cours, intensionnel, et les Travaux Pratiques, extensionnels. Dans le premier cas je fais appel à leurs capacités de déduction conscientes, dans le second cas à leurs capacités d'induction inconscientes. Dans le premier cas je donne une identité à un concept⁴ avant de le leur imposer par le truchement du langage, dans le second cas je les amène à faire abstraction de ce qui diffère entre plusieurs exemples, cette identification résultant -à leur insu- en un concept.

Pour transmettre des connaissances à une machine, je peux lui fournir un algorithme intensionnel, mais le résultat est souvent catastrophique. Or cette machine n'a, à ce jour, aucune capacité d'induction, ce qui rend la deuxième voie impraticable. Il nous faut donc nous livrer à un patient travail sur la personne dont on veut capturer le savoir, l'amenant à nous révéler les connaissances que recouvre chaque concept, donc à en casser l'identité.

³A tel point que, s'il revient après une longue absence, on ne sera pas surpris que sa connaissance de l'application aît régressé.

⁴Pléonasme?

- Berthomé 95 Berthomé-Montoy, A., Une approche descriptive de l'auto-adaptativité des interfaces homme-machine, Thèse, Université Claude Bernard Lyon 1, Octobre 1995.
- Fouet 93 Fouet, J-M., Manuel d'utilisation de la Machine Gosseyn. Rapport interne, LISI, 1993.
- Fouet& 95 Fouet, J-M., Noura, R., **Acquiring and Structuring Knowledge Using Metaknowledge**, IJCAI Workshop on Machine Learning and Engineering, Montréal, Août 1995.
- Medjahed 96 Medjahed, F., Compilation de bases de connaissances dans un environnement objets/acteurs, thèse de Magister, Université Houari Boumedienne, Alger, Avril 1996.
- Noura& 96 Noura, R., Fouet, J-M., **A Knowledge Based Tool for Checking Large Knowledge Bases**, 9th International Conference on Industrial & Engineering Applications of AI, Fukuoka (Japon), Juin 1996
- Pitrat 90 Pitrat, J., Meta-connaissances. Futur de l'intelligence artificielle, Hermès, 1990.
- Raccach 96 Raccach, P.Y., **De la sémantique théorique à la terminologie**, Actes du 3ème Congrès Européen de Systémique, Rome, Octobre 1996.

Catégorisation adaptative de données sensori-motrices pour un système d'apprentissage par renforcement

Rémi Munos*

CEMAGREF, LISC, Parc de Tourvoie,
BP 121, 92185 Antony Cedex, FRANCE.

Tel : 01 40 96 61 21, Poste 64 14. Fax : 01 40 96 60 80

E-mail : Remi.Munos@cemagref.fr

Résumé :

Cet article propose d'utiliser le formalisme de l'Apprentissage par Renforcement pour mettre en évidence certains processus cognitifs mis en jeu lors de la prise de décisions d'un système en interaction avec son environnement. Nous pensons que la capacité d'apprendre par l'expérience nécessite une coopération entre les deux processus dynamiques d'apprentissage et de catégorisation. Nous définissons la fonction valeur, fonction construite par les techniques d'apprentissage par renforcement, qui anticipe les renforcements à venir et permet de déterminer le comportement adéquat. Enfin nous interprétons la fonction valeur comme une perception émotionnelle d'états corporels et pensons que la capacité de ressentir des émotions est utilisée dans les processus de prise de décisions. Aussi, nous proposons l'idée que le sentiment d'identité corporel pourrait naître des capacités d'anticipation et d'action sur l'environnement donc de la bonne adéquation de la valeur émotionnelle perçue en fonction de la situation vécue et des dynamiques d'interaction système-environnement.

1 Introduction

Dans cet article, nous souhaitons mettre en évidence certains processus cognitifs permettant à un système agissant au sein d'un environnement d'acquiescer, à

*DASSAULT-AVIATION, DGT-DTN-EL-Et.Avancées, 78 quai Marcel Dassault, 92214 Saint-Cloud

partir de l'expérience de ses interactions sensori-motrices, un ensemble de compétences ou d'aptitudes qui lui permettent d'agir de manière adéquate. Nous illustrons les mécanismes nécessaires à ces capacités d'apprentissage en proposant des modèles algorithmiques que l'on souhaite être les plus simples possibles et dont les simulations sur ordinateur rendent compte de tels comportements.

Un courant de l'Intelligence Artificielle s'intéresse à modéliser ces processus cognitifs par des algorithmes dits "d'apprentissage par renforcement". Il s'agit de considérer un *système artificiel*, ou *agent*, "plongé" dans un environnement, dont le comportement est défini par des séquences d'actions (ou décisions ou contrôles). L'information dont dispose le système pour améliorer ses performances est le *renforcement*, une rétribution occasionnelle et a posteriori de son comportement, qui représente une émotion ou une perception affective. C'est un scalaire positif lorsqu'il simule un encouragement (ou un bienfait, par exemple la découverte de nourriture) ou négatif lorsqu'il s'agit d'une sanction (ou un comportement néfaste, par exemple le fait de tomber lorsque l'on apprend à marcher).

Cette rétribution permet au système d'évaluer les prises de décisions passées, lui permettant de rectifier son comportement en renforçant les séquences d'actions qui mènent à des renforcements avantageux.

Nous pensons que de telles capacités d'apprentissage peuvent avoir lieu si le système contient les deux processus dynamiques suivants :

- **Une dynamique d'apprentissage** qui permet d'anticiper les renforcements à venir en construisant itérativement une fonction d'évaluation, qui associe à chaque donnée sensori-motrice (c'est à dire à chaque couple perception-action) une estimation du renforcement à venir.
- **Une dynamique de catégorisation** qui permet d'établir des catégories sensori-motrices homogènes selon le critère de l'estimation des renforcements à obtenir.

A partir de données neurophysiologiques concernant l'observation de modifications comportementales chez des personnes ayant subi certaines lésions cérébrales, nous tenterons de donner une interprétation biologique de ces deux dynamiques.

Nous avancerons l'idée que le sentiment d'identité, tout au moins corporel, nécessite un mécanisme qui permet de percevoir une émotion relative à l'"état corporel" du système, à partir de l'appréciation des dynamiques d'interactions avec son environnement, afin que l'émotion ressentie soit cohérente avec la situation vécue. Nous pensons qu'une telle capacité de ressentir des émotions est utilisée dans les processus de prise de décisions et permet au système d'engendrer des stratégies d'action sur son environnement.

La *partie 2* fournit une introduction aux techniques d'apprentissage par renforcement et décrit le formalisme employé.

La *partie 3* traite des deux dynamiques d'apprentissage et de catégorisation et fait référence à deux réalisations informatiques mettant en application ces mécanismes.

La *partie 4* fournit une interprétation biologique de la fonction d'évaluation et des deux dynamiques précédentes.

2 Introduction aux méthodes d'apprentissage par renforcement

Les modèles d'apprentissage par renforcement sont issus des travaux en psychologie expérimentale (voir par exemple, [RW72]) et en neurophysiologie. Ils sont essentiellement utilisés en *théorie du contrôle* comme méthode adaptative pour déterminer le contrôle optimisant le gain (voir [Bar90] par exemple). Ils fournissent également des algorithmes d'adaptation (par exemple l'algorithme du "Bucket Brigade" utilisé dans les systèmes de règles, voir [Gol89]) pour des systèmes artificiels dits *animats* (modèles informatiques et robotiques qui s'inspirent de comportements animaux; voir les conférences "Simulation of Adaptive Behavior", par exemple [CHMW94]).

L'apprentissage par renforcement s'intéresse à l'acquisition de compétences d'un système en interaction avec un environnement. Il s'agit davantage de l'acquisition d'un "savoir-faire" ou d'une habileté à réagir en situation (par exemple les coordinations sensori-motrices nécessaires pour tenir en équilibre sur une bicyclette ou pour jouer au tennis), qu'à l'apprentissage de connaissances symboliques (au sens représentationnel ou linguistique).

Lors de son développement, l'individu acquiert, conditionné par les contraintes et possibilités de son être physique en interaction dans un environnement, des compétences motrices qui lui permettent d'agir de manière adéquate en fonction de ses propres capacités et limites.

L'apprentissage par renforcement, tente modestement de modéliser le processus de prise de décisions qui permet de générer des comportements pertinents afin d'optimiser la satisfaction générale, définie assez librement en termes de renforcement, d'un système. En effet, la récompense peut provenir directement de l'environnement (guidage par encouragement ou sanction, méthodes de conditionnement, de dressage) mais elle peut aussi modéliser une motivation interne, propre au système (par exemple selon un principe de plaisir), ce qui lui octroie une certaine autonomie, dans les limites des contraintes externes. En général, les modélisations informatiques prennent en compte un renforcement (interne ou externe) qui ne dépend que de l'état présent.

La forme d'apprentissage par renforcement n'exclut pas l'utilisation d'autres formes d'apprentissage (symbolique, supervisé, par analogie, etc.) ni la notion de connaissance a priori, lorsqu'elles sont disponibles. Voici le formalisme mathématique généralement employé.

2.1 Formalisme utilisé

2.1.1 Positionnement du problème

A l'instant t , le système reçoit une certaine perception (de l'environnement et de lui-même) qui définit son *état* $x(t) \in X$ (où X est l'espace d'état). En fonction des compétences actuelles Q (notation détaillée plus tard), le système choisit ou sélectionne une action (ou un contrôle ou un comportement) $a(t) \in A$ ou A est l'ensemble des actions possibles.

Suite à cette action, en fonction de la *dynamique d'interaction agent-environnement*, le système perçoit un nouvel état $x(t+1)$ et peut éventuellement recevoir un renforcement $r(x(t), a(t))$. Selon les processus que l'on cherche à contrôler, on utilise un modèle du temps discret ou continu :

- **Modèle discret.** Si la dynamique est stochastique, l'état résultant $x(t+1)$ est défini selon une distribution de probabilité qui dépend de $x(t)$ et $a(t)$. Si la dynamique est déterministe, l'état résultant s'exprime comme une fonction successeur : $x(t+1) = \text{succ}(x(t), a(t))$.
- **Modèle continu.** Dans le cas déterministe, la dynamique s'exprime comme une équation différentielle dépendant de l'action a :

$$\frac{dx}{dt} = f(x(t), a(t))$$

Dans la suite, nous nous limiterons à étudier les dynamiques déterministes sachant que les conclusions auxquelles nous aboutirons seront généralisables au cas stochastique.

Ainsi, pour tout état initial x , le choix d'une séquence d'actions $a(t)$ génère une séquence d'états $x(t)$ appelée *trajectoire*.

Pour définir la valeur, ou le gain d'une trajectoire, on peut comptabiliser la somme des récompenses obtenues au cours de la trajectoire pondérée par un taux d'actualisation $\gamma < 1$ (notion courante en économie). On définit ainsi le *gain* :

$$J(x; a(.)) = \sum_0^{\infty} \gamma^t r(x(t), a(t)) \text{ ou } J(x; a(.)) = \int_0^{\infty} \gamma^t r(x(t), a(t)) dt$$

selon le formalisme employé. Le facteur γ permet d'établir une préférence pour une récompense immédiate plutôt que pour une récompense de même valeur mais différée dans le temps. On peut alors formuler l'objectif de l'apprentissage par renforcement ainsi :

Trouver pour chaque état x , la séquence d'actions optimale $a^*(t)$ qui permet de maximiser le gain $J(x, a(t))$.

2.1.2 Méthode de la Programmation Dynamique

Le principe de la Programmation Dynamique (voir [Ber87]) consiste à remplacer le problème d'optimisation global précédent (dont l'espace de recherche est l'espace énorme de tous les comportements possibles) par une condition locale (simple) à satisfaire en tous points de l'espace d'état.

Pour cela, on construit une fonction d'anticipation, appelée la *fonction valeur* V , qui représente l'espérance maximale de gain, c'est à dire le gain obtenu lorsque le système effectue le comportement optimal à chaque instant. Ainsi, la fonction valeur est définie par :

$$V(x) = \sup_{a(.)} J(x; a(.))$$

La condition locale est appelée équation de la Programmation Dynamique (formalisme discret) ou équation de Hamilton-Jacobi-Bellman (formalisme continu) et porte sur la fonction valeur :

- **Equation de la Programmation Dynamique :**

$$V(x) = \sup_{a \in A} [r(x, a) + \gamma \cdot V(\text{succ}(x, a))]$$

- **Equation de Hamilton-Jacobi-Bellman :**

$$V(x) \ln \gamma + \sup_{a \in A} [r(x, a) + DV(x) \cdot f(x, a)] = 0$$

où $DV(x)$ représente le gradient de V en x .

Ainsi, nous avons une équation de "propagation" de la fonction valeur le long de la dynamique d'interaction agent-environnement, et en certains états nous connaissons les renforcements correspondants. Ainsi, en partant des renforcements et en les propageant de proche en proche dans tous l'espace d'état, le système peut construire le paysage de l'espérance de gain à venir.

L'approximation de la fonction valeur permet d'anticiper et de prévoir les renforcements à venir. Nous l'interpréterons (voir la *partie 4*) comme le fait de "ressentir émotionnellement" la valeur d'une situation ; par exemple si elle est négative, elle peut correspondre à une peur à l'approche d'un danger.

Cependant, pour un état donné, la seule information de la fonction valeur ne suffit pas pour déterminer quelle est la meilleure action à choisir. Pour cela on utilise la *fonction qualité* $Q(x, a)$ définie pour chaque couple (état, action) et qui permet de déterminer l'action optimal en chaque état. Cette fonction attribue à chaque donnée sensori-motrice (état x , action a) une estimation du meilleur gain lorsque que le système perçoit l'état x et choisit l'action a . Considérons ici le cas discret, l'équation de la programmation dynamique s'écrit :

$$\begin{aligned} Q(x, a) &= r(x, a) + \gamma \cdot V(\text{succ}(x, a)) \\ \text{avec } V(x) &= \sup_{a \in A} Q(x, a) \end{aligned}$$

Ainsi, la fonction qualité donne une évaluation des couples de données perception-action et permet de définir l'action optimale $a^*(x)$ en chaque état x :

$$a^*(x) = \arg \sup_{a \in A} Q(x, a)$$

Cependant, afin de permettre une exploration de son environnement (notamment en début d'apprentissage), le système peut avoir intérêt à choisir des actions sous-optimales, le choix pouvant se faire, par exemple, selon une distribution de Boltzmann :

$$\text{L'action } a \text{ est choisi selon la probabilité } p(x, a) = \frac{e^{\frac{1}{T} Q(x, a)}}{\sum_{a' \in A} e^{\frac{1}{T} Q(x, a')}}$$

qui permet de générer un bon compromis (en faisant varier la valeur de la "température" T) entre l'exploration de l'environnement (T élevée) et l'exploitation en fonction des connaissances actuelles (T proche de 0). Voir [Meu96] pour une étude sur le dilemme exploration/exploitation dans les systèmes d'apprentissage par renforcement.

Le fait que le choix des actions $a^*(x)$ dépend de l'évaluation $Q(x, a)$ des situations-actions sera rapproché (à la *partie 4*) à la corrélation, observée chez des personnes ayant subi des lésions cérébrales, entre une déficience de leur capacité de prendre des bonnes décisions et une absence de perception émotionnelle correcte des situations.

Nous allons maintenant décrire un exemple d'algorithme d'apprentissage par renforcement, c'est à dire une méthode qui permet au système en interaction avec l'environnement de construire la fonction d'évaluation.

2.2 Un exemple d'algorithme : le Q-learning

Un *algorithme d'apprentissage* est une loi de modification itérative des compétences du système ; en particulier l'algorithme du Q-learning (créé par [Wat89]) fournit la règle suivante portant sur les qualités :

Lorsque le système se trouve dans l'état x , choisit l'action a et se retrouve dans l'état résultant y en percevant éventuellement une récompense $r(x, a)$ alors la qualité $Q(x, a)$ est modifiée selon :

$$\Delta Q(x, a) = \alpha [r(x, a) + \gamma V(y) - Q(x, a)]$$

(avec α un coefficient d'apprentissage décroissant au cours du temps). Cet algorithme est intéressant car on connaît les hypothèses pour lesquelles l'apprentissage converge vers la solution optimale.

Ainsi, un algorithme d'apprentissage par renforcement construit *itérativement*, à partir des *successions des perceptions-actions vécues*, une fonction d'anticipation qui attribue à chaque donnée sensori-motrice une estimation des rétributions à venir.

3 Une double dynamique d'adaptation

3.1 Nécessité d'une catégorisation

Dans le cas général, l'espace des perceptions-actions est immense, ce qui entraîne un problème de représentation sur un ordinateur de la fonction qualité ainsi qu'une divergence du nombre d'itérations à effectuer pour que les valeurs convergent. Il est alors tentant de chercher les régularités de la fonction qualité et d'utiliser ces propriétés pour représenter la fonction avec un nombre limité de paramètres.

Ainsi, on peut essayer de regrouper des perceptions-actions différentes ayant des qualités similaires, donc créer des sous-ensembles de l'espace des perceptions-actions appelés *catégories sensori-motrices*. L'apprentissage consiste alors à déterminer la qualité de chaque catégorie sensori-motrice.

Ainsi, une possibilité serait de faire coopérer les deux processus dynamiques suivants (voir la figure 1 pour une illustration) :

- Une *dynamique d'apprentissage* qui évalue, via le renforcement, les qualités des catégories sensori-motrices existantes.
- Une *dynamique de catégorisation* qui partitionne l'espace des données sensori-motrices en catégories contenant des états dont les gains espérés sont homogènes.

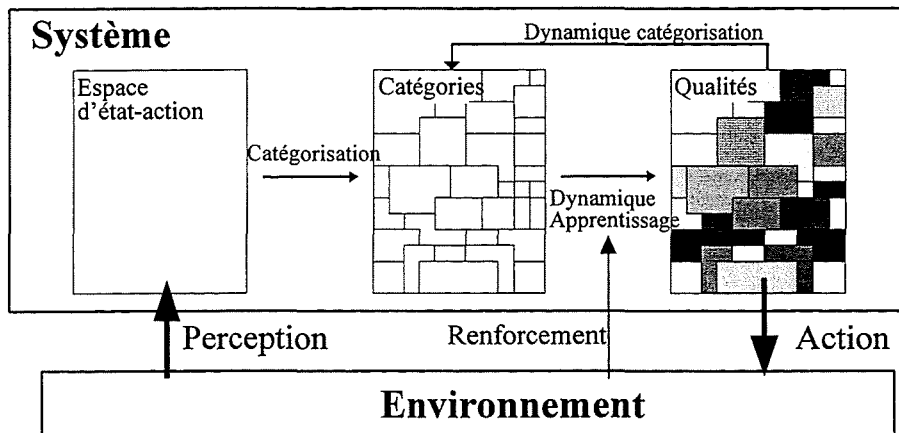


Figure 1: *Le processus de catégorisation des données sensori-motrices via le renforcement.* L'ensemble des perceptions et actions possibles définit l'espace d'état-action. Cet espace est partitionné en régions appelées catégories. La dynamique d'apprentissage permet d'évaluer les qualités (en gris) des catégories. La dynamique de catégorisation modifie le partitionnement pour avoir des catégories homogènes.

Remarque : la modélisation du processus de catégorisation décrite ici est outrageusement simplifiée. En effet, en général, la simple perception de l'état du système met en jeu des processus complexes (détection de traits, discrimination, segmentation, etc.) qui est déjà un mécanisme de catégorisation. Nous n'étudierons ici, pour des raisons de simplicité, qu'un processus de catégorisation qui ne fait que regrouper des états (i.e. les catégories sont des parties de l'espace d'état-action).

3.2 Modèles informatiques

Nous avons étudié les deux réalisations suivantes :

- Le *Q-learning partitionné* (défini dans [MP94]) utilise un formalisme discret et définit une fonction qualité constante à l'intérieur des catégories. Celles-ci sont représentées sous forme de règles et utilisent le formalisme des "Classifiers Systems". La dynamique d'apprentissage consiste à modifier les qualités des règles sélectionnées en fonction du renforcement obtenu. Le processus de catégorisation permet de remplacer une règle générale (règle mère) par une série de règles plus spécifiques (règles filles).
- Le *Finite-Element Reinforcement Learning* avec triangulation adaptative (voir [Mun96] et [Mun97]) utilise un formalisme continu et approche la fonction valeur par des méthodes aux Eléments Finis. Pour une triangulation donnée de l'espace d'état, la fonction qualité est approchée par une fonction linéaire à l'intérieur des triangles (qui définissent les catégories) et la dynamique d'apprentissage consiste à modifier les valeurs aux sommets de la triangulation. Le processus de catégorisation consiste à ajouter des points à l'ensemble des sommets et ainsi permet de construire une nouvelle triangulation plus fine.

Pour ces deux réalisations, la dynamique de catégorisation est identique. Il s'agit d'un processus de spécialisation (descendant) des connaissances, c'est à dire que le système est initialisé avec un petit nombre de "grandes" catégories (ce qui représente une connaissance potentielle grossière et de faible pouvoir discriminant, i.e. connaissance du novice) et le processus de catégorisation consiste à subdiviser certaines catégories (selon un critère estimant l'hétérogénéité de l'espérance de gain dans chaque catégorie) en plusieurs catégories plus fines. Pour déterminer quelles sont les catégories qu'il est nécessaire de subdiviser, on utilise un critère local représentant l'erreur d'approximation de l'espérance de gain dans une catégorie : l'écart-type des modifications de la qualité de cette catégorie. Ainsi, si une même catégorie mène à des renforcements différents, cela signifie qu'elle contient des états dont l'estimation de gain est variable, sa qualité sera modifiée constamment et elle sera subdivisée.

Ainsi, la connaissance acquise après la dynamique d'apprentissage devient plus précise et pertinente. Ce processus modélise l'acquisition de connaissances expertes.

Cependant ces réalisations souffrent d'une absence de processus ascendants pour lesquels des catégories spécialisées pourraient se regrouper en une catégorie globale afin d'établir un mécanisme de généralisation.

4 Interprétation biologique

4.1 Faits neurophysiologiques

A. Damasio (voir [Dam95]) rapporte le cas de certains patients qui, suite à des lésions cérébrales (causées par des maladies neurologiques ou des accidents ayant détruit des régions corticales), ont subitement été handicapés dans l'aptitude à prendre des décisions. Pourtant, ces personnes semblaient tout à fait normaux par ailleurs ; leurs capacités linguistiques, de mémorisation, d'attention, de raisonnement logique étaient intactes. Seul un symptôme accompagnait cette déficience à agir de manière adaptée : une incapacité à ressentir la moindre émotion.

Ainsi, des lésions cérébrales à différents niveaux :

- régions ventro-médianes du cortex préfrontal,
- aires corticales somatosensorielles dans l'hémisphère droit,
- système limbique (notamment l'amygdale)

entraînent une déficience dans la prise des bonnes décisions (au niveau social, personnel ou de la motricité) ainsi qu'un affaiblissement de la perception des émotions.

Ainsi, on peut considérer cette incapacité à agir de manière pertinente dans leur environnement comme une impossibilité d'attribuer des poids différents aux diverses solutions possibles de sorte que le paysage émotionnel où s'effectue les prises de décisions est plat ou pire, ne correspond pas à la situation vécue.

4.2 Interprétation

Considérons l'"état corporel" : ensemble des perceptions internes provenant de l'état du corps, des viscères, des muscles, de la peau, des organes sensoriels, etc.

- Nous identifions l'état du système x à cet état corporel.
- Nous identifions le renforcement $r(x, a)$ à une *émotion primaire* ressentie de façon réflexe et préprogrammée en réponse à la détection de certains traits de l'état corporel, par exemple, de certains sons (rugissement d'un lion), de certains types de mouvements (reptiles), de certains contacts physiques avec des corps extérieurs (douleur, plaisir), etc.

Selon Damasio, la perception d'une émotion correspond à l'information sensorielle provenant d'une partie de l'état corporel. Ainsi,

- nous identifions la fonction valeur $V(x)$ (resp. la fonction qualité $Q(x, a)$) à la perception émotionnelle de l'état corporel x (resp. de l'état x et de l'action a).

Il s'agit d'une *émotion secondaire* apprise par le système et qui résulte d'une propagation des émotions primaires (les renforcements) selon sa dynamique d'interaction avec l'environnement. Ainsi une peur sera apprise à l'approche d'un danger. Un appétit à l'approche d'un repas.

En poursuivant l'analogie, les deux dynamiques peuvent être interprétées de la manière suivante :

- La dynamique d'apprentissage permet d'attribuer à chaque catégorie corporelle la valeur émotionnelle adaptée à l'état actuel, c'est à dire celle qui satisfait une bonne adéquation du système en interaction avec l'environnement. Ainsi, un algorithme d'apprentissage par renforcement (par exemple le Q-learning) permet de construire itérativement cette perception émotionnelle.
- La dynamique de catégorisation permet de regrouper les états corporels en fonction du seul critère de similarité de leur valeur émotionnelle. Ainsi, des états seront perçus comme étant différents uniquement s'ils ont des influences différentes sur le renforcement à venir.

Il nous semble que le sentiment d'identité corporel peut apparaître lorsqu'un système développe des capacités d'anticipation et d'action sur l'environnement lui permettant d'assurer une certaine autonomie. D'après les faits neurophysiologiques précédemment cités, il apparaît que la capacité de ressentir des émotions est utilisée dans les processus de prise de décisions à différents niveaux. Ainsi, nous proposons l'idée selon laquelle l'apparition de l'identité corporelle d'un système nécessite l'existence d'un mécanisme de perception émotionnelle ainsi qu'un processus d'acquisition par l'expérience de ces émotions (par les deux dynamiques d'apprentissage et de catégorisation) afin d'assurer une bonne adéquation de la valeur émotionnelle perçue par rapport à la situation vécue.

5 Conclusion

Nous avons succinctement proposé une modélisation possible des mécanismes cognitifs mis en jeu dans l'apprentissage par l'expérience en distinguant deux processus complémentaires. La dynamique d'apprentissage permet de construire pour chaque catégorie sensori-motrice une évaluation émotionnel afin de permettre des prises de décisions et comportements satisfaisants en fonction du renforcement à venir. La dynamique de catégorisation regroupe des données sensori-motrices en fonction du seul critère de similarité de leur espérance de renforcement.

Ce critère de regroupement est le plus faible si l'on considère le seul objectif de l'optimisation des rétributions à venir.

En effet, un critère plus fort consisterait à regrouper des perceptions-actions menant à des situations similaires, mais ceci nécessiterait de construire au cours de l'apprentissage, un modèle de l'environnement (voir par exemple [Sut90]), ce qui serait utile, mais coûteux d'un point de vue cognitif (puisque cela utiliserait un processus de création de cartes mentales ainsi qu'un processus de planification).

Sous le formalisme étudié ici, la faculté de créer un modèle de l'environnement n'est pas utile car la modélisation de la dynamique agent-environnement ($r(x, a)$, $succ(x, a)$ et $f(x, a)$) ne dépend que de la perception et de l'action à l'instant présent (ce sont des modèles markoviens) ce qui est une simplification abusive pour la plupart des cas réels. Dans les cas où il y a une dépendance du passé ou lorsque la perception est imparfaite (voir les travaux sur les processus partiellement observables de [LCK95] et sur les variables cachées de [McC95]), il semble nécessaire de construire (au moins partiellement) un modèle de l'environnement.

Ainsi, il apparaît que les dynamiques d'apprentissage et de catégorisation représentent les capacités cognitives les plus simples qui rendent compte de comportements d'apprentissage par renforcement et qu'il faut leur associer, selon la complexité de l'environnement dans lequel le système évolue, divers "modules" cognitifs.

Nous avons proposé une interprétation de la fonction valeur comme étant une perception émotionnelle d'un état corporel et soutenu que les processus de prise de décisions dépendaient de l'adéquation de l'émotion perçue avec la situation vécue, relativement à la dynamique d'interaction système-environnement, c'est à dire du bon fonctionnement des dynamiques d'apprentissage et de catégorisation.

References

- [Bar90] Andrew G. Barto. *Neural networks for control*. W.T. Miller, R.S. Sutton, P.J. Werbos editors. MIT press, Cambridge, Massachusetts, 1990.
- [Ber87] Dimitri P. Bertsekas. *Dynamic Programming : Deterministic and Stochastic Models*. Prentice Hall, 1987.
- [CHMW94] D. Cliff, P. Husbands, J.A. Meyer, and S.W. Wilson. *From Animals to Animats 3: Proceedings of the Third International Conference on Simulation of Adaptive Behavior*. The MIT Press/Bradford Books, 1994.
- [Dam95] A. Damasio. *L'erreur de Descartes, La raison des émotions*. Odile Jacob, 1995.
- [Gol89] D.E. Goldberg. *Genetic Algorithms in search, optimization and machine learning*. Addison Wesley, 1989.

- [LCK95] M. Littman, A. Cassandra, and L.P. Kaelbling. Learning policies for partially observable environments: Scaling up. *International Conference on Machine Learning*, 1995.
- [McC95] R.A. McCallum. Instance-based utile distinctions for reinforcement learning with hidden state. *International Conference on Machine Learning*, 1995.
- [Meu96] N. Meuleau. *Le dilemme Exploration/Exploitation dans les systèmes d'apprentissage par renforcement*. PhD thesis, Université de Caen, 1996.
- [MP94] R. Munos and J. Patinel. Reinforcement learning with dynamic covering of the state-action space : Partitioning q-learning. *Simulation of Adaptive Behavior*, 1994.
- [Mun96] R. Munos. A convergent reinforcement learning algorithm in the continuous case : the finite-element reinforcement learning. *International Conference on Machine Learning*, 1996.
- [Mun97] R. Munos. Finite-element methods with local triangulation refinement for continuous reinforcement learning problems. *European Conference on Machine Learning*, 1997.
- [RW72] R.A. Rescorla and A.R. Wagner. A theory of pavlovian conditioning: variations in the effectiveness of reinforcement and nonreinforcement. *Classical conditioning II: Current research and theory*, 1972.
- [Sut90] R.S. Sutton. Integrated architectures for learning, planning and reacting based on approximating dynamic programming. *International Conference on Machine Learning*, 1990.
- [Wat89] Christopher J.C.H. Watkins. *Learning from delayed reward*. PhD thesis, Cambridge University, 1989.

Architecture Multi-Agents de Robots Mobiles Autonomes

VINCENT DECUGIS[†] & JACQUES FERBER[‡]

[†] Laboratoire Système de Perception, ETCA,
16 bis Avenue Prieur de la Côte d'Or, 94114 ARCUEIL Cedex France

[‡] Laboratoire d'Informatique et de Robotique et de Microélectronique de Montpellier,
161, rue Ada, 34392 Montpellier Cedex 5

27 octobre 1996

1 Introduction

Une branche relativement récente de la robotique remet en question un certain nombre de principes qui étaient à la base de l'approche classique. Le point de départ de ces architectures repose sur la constatation du neurobiologiste Braitenberg, selon laquelle des réflexes extrêmement simples implantés sur un robot suffisent à donner à celui-ci un comportement utile et très efficace [1]. On doit citer dans ce contexte le travail de Brooks qui a su le premier poser les bases de cette approche nouvelle dans une série d'articles fondamentaux ([3], [2], [4], [5]). Cette approche utilise une métaphore biologique pour justifier sa démarche, soit en comparant directement le comportement des robots à des comportements animaux, en s'appuyant sur des modèles éthologiques, neurobiologiques ou psychologiques, soit en utilisant des techniques informatiques issues de modèles biologiques comme les réseaux de neurones, l'évolution artificielle ou la programmation génétique.

Néanmoins, l'approche biologique ne fournit pas aussi rapidement qu'espéré des méthodes de programmation simples et efficaces pour les robots mobiles autonomes. En fait, s'il est relativement facile de concevoir des robots parfaitement autonomes, présentant une certaine viabilité dans un environnement donné, il est souvent impossible de faire accomplir au robot une tâche précise définie a priori. Il faut que de l'accomplissement de cette tâche dépende la survie du robot, que celui-ci est en quelque sorte «intéressé» à ce qu'elle s'accomplisse. De plus, le comportement du robot étant difficilement prévisible, la mise au point ne peut en général se faire que de manière empirique.

L'objectif général de notre travail est de trouver une méthode de programmation de robots autonomes qui possède les propriétés suivantes:

- Assurer la survie du robot dans un environnement dont la topologie est inconnue, mais dont la nature est connue;

- Mener à bien un objectif utile au concepteur;
- Posséder un comportement relativement stable et prévisible;
- Permettre une programmation simple et modulaire. En particulier, limiter au maximum le nombre de paramètres à régler empiriquement pour obtenir un effet désiré.

Afin de préciser notre démarche pour aborder cet objectif, nous allons examiner plusieurs questions auxquelles les deux approches doivent répondre.

La première concerne la manière de représenter les informations que le robot extrait de son environnement. La robotique classique utilise des procédés de traitement d'image, de fusion de données et de reconnaissances des formes très sophistiqués pour doter le robot d'un modèle géométrique ou symbolique de l'environnement aussi précis que possible. Toutes ces informations servent à prévoir l'évolution de l'environnement en fonction des actions que le robot doit entreprendre pour atteindre l'objectif qui lui a été fixé. Les problèmes engendrés par cette méthode tiennent au fait que les représentations employées sont arbitraires. L'utilisation d'une carte tridimensionnelle ou d'un ensemble de symboles abstraits comme description de l'environnement est non seulement très difficile à obtenir, mais aussi très délicat à utiliser. Le résultat est souvent une grande lenteur du robot, car les besoins en temps de calcul introduit par la représentation sont énormes.

La réponse de l'approche biologique à ce problème consiste à éliminer au maximum le recours à ces représentations a priori. Ceci est possible en rendant le robot totalement réactif. Un robot réactif réagit instantanément aux sollicitations de l'environnement, sans chercher à prévoir les conséquences à long terme de leurs actions. Le robot décide alors de l'action à entreprendre à partir des données brutes de ses capteurs à chaque instant. La réactivité se révèle très efficace dans les situations dangereuses, car elle fournit

au robot une réponse instantanée. Cependant elle ne fournit pas de réponse satisfaisante quand l'environnement est non-markovien¹, ce qui est toujours le cas dans des situations réalistes. Prenons l'exemple d'un robot purement réactif cherchant à mettre un objet dans un tiroir. On pourra supposer qu'il dispose de primitives réactives lui donnant la possibilité de chercher cet objet, de le transporter, d'ouvrir le tiroir, de le placer dans le tiroir et de refermer le tiroir. Chacune de ces actions élémentaires est déclenchée par un stimulus; par exemple ne pas avoir d'objet dans la main activera le comportement de recherche d'objet. Si tout se passe bien, le robot cherchera l'objet, l'emmènera au tiroir, l'ouvrira, posera l'objet, refermera le tiroir. Une fois le tiroir refermé, le robot ne percevra plus l'objet, et n'aura aucun moyen de déceler sa présence dans le tiroir. Il se retrouvera donc exactement dans les conditions initiales de l'expérience, et par conséquent il se remettra à rechercher l'objet déjà dans le tiroir.

La réactivité pure a donc un prix. Néanmoins, la solution au problème particulier soulevé par notre exemple semble relativement simple. Il suffit en effet de permettre au robot de mémoriser le fait qu'il a déjà mis l'objet dans le tiroir. Nous introduisons par là une forme de représentation de l'environnement qui indiquera au robot si le tiroir contient l'objet ou non. Contrairement aux représentations de la robotique classique cette façon de représenter l'environnement a posteriori est adaptée aux problèmes réels rencontrés par le robot. Une architecture de contrôle incluant la possibilité de construire ce genre de représentation ne tombe pas dans le piège du fondement symbolique, car elle se fonde sur un caractère pragmatique: elle servira à résoudre un problème précis que le robot rencontre dans son interaction avec l'environnement. Par contre, les «concepts» que le robot est ainsi capable d'appréhender sont plus limités. Ainsi, on ne pourra pas apprendre à un tel robot ce qu'est une chaise. Il faudrait pour cela qu'il ait la capacité de s'asseoir, la «chaise» serait alors l'endroit où il pourrait utiliser cette capacité. Mais sans celle-ci, la chaise ne sera pour lui tout au plus qu'un obstacle d'une forme particulière.

La notion d'action est donc au centre du problème de la représentation. On peut la considérer selon deux perspectives différentes. La première définit l'action du robot par sa fonction. Le robot possédant un objectif final, chacune des actions remplit une étape bien précise dans l'accomplissement de ce but. C'est le genre d'action que nous avons employé dans l'exemple du tiroir. Cette conception est bien adaptée à l'utilisation d'une représentation symbolique. L'action est alors décrite sous forme d'une liste de prédicats rendus vrais ou faux par son exécution. Les méthodes

de planification symbolique utilisent cette conception très fréquemment. Ce point de vue ne s'intéresse pas à la manière dont l'action est exécutée. Il pourra être qualifié de «descendant», puisqu'on part du problème global sans tenir compte explicitement des moyens initialement à notre disposition pour le résoudre. Si une telle forme d'action se prête très bien aux simulations informatiques, elle s'avère souvent insuffisante pour réaliser un vrai robot.

A l'inverse, on peut se focaliser sur la manière de réaliser un but partiel, et considérer que l'ensemble des moyens pour y parvenir constitue une action. Un asservissement sur une trajectoire particulière, ou un couplage sensori-moteur direct² constituent des exemples d'une telle forme d'action. Un véhicule de Braitenberg sera ainsi un robot implémentant une seule action. Cette conception se rapproche de la notion de comportement. Ce point de vue est alors «ascendant», car on part des structures de réalisation. La fonction résultante n'en est plus qu'une conséquence. Le problème apparaît quand on veut combiner plusieurs actions différentes. En particulier, il est très difficile de déterminer à quel moment une action doit être interrompue. Si le but partiel est rempli, la question ne se pose pas. Mais les choses se déroulent rarement aussi simplement, et le robot a peu de moyens de s'apercevoir d'un échec. Il est donc hasardeux de planifier une séquence de plusieurs actions de ce type. C'est pourquoi cette conception est plutôt employée dans les robots uniquement réactifs où il n'y a pas de planification des actions.

La réalisation d'un robot réel répondant aux spécifications que nous nous sommes fixées implique l'intégration de ces deux idées dans une architecture cohérente. Le point de vue ascendant étant notre point de départ nous tenterons de résoudre le problème délicat de la planification. Les notions de représentation que nous avons proposées devraient permettre de dépasser les limitations des robots réactifs.

Enfin, les deux approches ont adopté des démarches fondamentalement différentes pour donner un forme d'«intelligence» à leurs robots. Comme la robotique classique utilise les techniques de l'intelligence artificielle, elle a toujours découpé le problème de manière fonctionnelle, chacune des parties de l'architecture devant résoudre un sous-problème clairement identifié. A l'inverse, l'approche biologique a privilégié le découpage structurel, se concentrant sur les structures de base de la cognition [16]. Ceci conduit à des architectures décentralisées: chaque fonction est répartie sur plusieurs structures, et chaque structure participe à la réalisation de plusieurs fonctions. De nombreuses architectures décentralisées ont été proposées à partir de ce principe ([6], [8], [14], [15], ...). L'étude des systèmes multi-agents [7] fournit un formalisme intéressant à toutes ces approches en mon-

1. Environnement dans lequel la perception instantanée du robot ne suffit pas à prédire statistiquement l'état suivant de l'environnement.

2. Couplage entre les valeurs des capteurs et les consignes des effecteurs

trant comment des propriétés globales peuvent naître de l'interaction complexe de plusieurs éléments relativement indépendants entre eux. Dans la pratique, les systèmes multi-agents permettent de mieux comprendre les relations entre la structure d'une architecture et les fonctions qui en découlent. Ils présentent en tout cas la possibilité de mieux saisir les ressemblances entre les différentes architectures distribuées, en s'abstrayant de leurs différences formelles.

La représentation de l'environnement, la sélection d'action, l'organisation structurelle de l'architecture de contrôle et l'adaptativité du robot sont des problèmes fondamentaux et liés. Nous proposons de les aborder par une approche sensori-motrice et ascendante, en proposant une architecture dans laquelle les relations entre la perception et l'action se construisent en même temps que la représentation que le robot se fait du monde. Notre but est de fournir ainsi au robot une base sur laquelle le développement de capacités cognitives simples pourraient être envisagée.

2 Architecture multi-agents

2.1 Principe

L'architecture que nous proposons s'inspire largement de celle proposée par Maes (cf. [10], [11], [12]). Ce choix a été essentiellement dicté par la méthode de sélection d'action sous-jacente qui permet d'obtenir un compromis paramétrable entre la réactivité et la planification. Le problème principal a été de transposer cette architecture prévue pour un environnement et des actions symboliques à un cas plus réaliste. Or celle-ci a déjà été utilisée avec succès pour modéliser un ensemble de réflexes spinaux de la grenouille [9]. On s'est aperçu expérimentalement que la grenouille possédait un ensemble de réflexes lui permettant d'éliminer un contact gênant dans le dos. Lorsqu'elle ressent une sensation de ce genre, la grenouille envoie une de ses pattes postérieures vers la gène afin d'éliminer sa cause. Giszter a proposé un modèle de contrôle de l'ensemble de ces réflexes basé sur l'architecture de Maes. Chaque réflexe est codé sous la forme d'une carte transformant les sensations proprioceptives dues à la position de la patte en une commande motrice des muscles de celle-ci. Il est donc parfaitement possible d'utiliser l'architecture de Maes dans un cas continu. Nous avons donc essayé de résoudre les problèmes de segmentation de l'action, des perceptions et du temps qui sont associés à ce passage à une forme d'action ascendante.

En passant de la notion d'action symbolique à la notion d'action réflexe, c'est-à-dire à celle de couplage sensori-moteur direct, les deux premiers problèmes ont été résolus. Alors que la prise de décision dans l'architecture de Maes se fait à partir d'informations symboliques sur l'environnement, notre architecture de sélection d'action a pour rôle de choisir parmi un

répertoire de réflexes disponibles celui qui correspond le mieux à la situation. Les perceptions sont choisies pour refléter une information pertinente à l'application d'un réflexe particulier. Elles sont construites à partir des données brutes délivrées par les capteurs du robot. Par exemple, pour déclencher un réflexe d'évitement d'obstacle, on utilisera une perception déterminant si un obstacle est perçu ou non par le robot.

2.2 Description de l'architecture

L'architecture est organisée sous la forme d'un réseau dont les noeuds sont :

- des *perceptions* combinant les valeurs des capteurs pour déterminer si une proposition logique sur celles-ci est vraie. Les perceptions possèdent un niveau d'activation α_P . Si \mathcal{S} est l'espace de toutes les valeurs possibles des capteurs, et P la perception considérée, nous noterons $v_P : \mathcal{S} \rightarrow \{0, 1\}$ la fonction permettant de calculer la valeur de vérité de P .
- des *comportements réflexes* qui possède un niveau d'activation α_R , et une fonction de couplage sensori-motrice $b_R : \mathcal{S} \rightarrow \mathcal{E}$, ou \mathcal{E} est l'espace des sorties motrices du robot.
- des *buts* qui possède une force α_B et un ensemble de perceptions et de réflexes cibles.

Ces noeuds sont reliés par des liens orientés, des perceptions vers les réflexes ou des réflexes vers les perceptions. Le lien d'une perception vers un réflexe peut être obligatoire, dans le sens où ce réflexe ne pourra pas être sélectionné si la perception origine n'est pas vérifiée. A chaque pas de temps, les liens permettent la propagation des niveaux d'activations à travers le réseau, dans le sens direct du lien suivant la formule :

$$\alpha_{ij} = \alpha_i f_{ij} \theta_{direct}$$

où i est l'origine du lien, j son extrémité, f_{ij} sa force et θ_{direct} un paramètre de l'architecture. Le lien permet également une rétro-propagation :

$$\alpha'_{ij} = \alpha_j f_{ij} \theta_{inverse}$$

En parallèle à ces phénomènes de propagation de l'activation, une introduction d'activation est opérée à chaque instant dans le réseau :

- par les buts à leurs réflexes et perceptions cibles, de valeur α_B
- dans les perceptions quand elles sont vérifiées, de valeur 1.

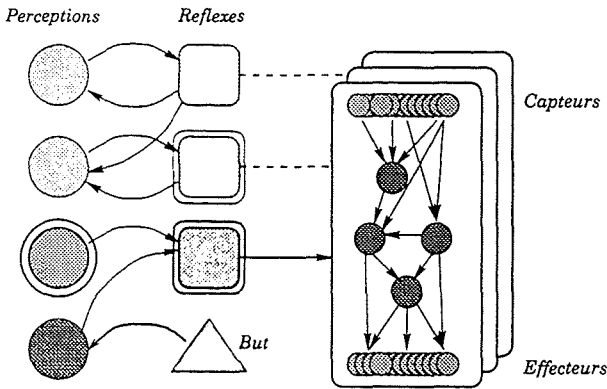


FIG. 1 - Réseau de sélection d'action. Le réseau est présenté à gauche avec ses perceptions et ses réflexes. A droite, le détail d'un des réflexes est montré: il s'agit d'un couplage sensori-moteur implémenté sous la forme d'un petit réseau de neurones (ce n'est qu'un exemple). Les quatre types de liens du réseau sont représentés: lien direct perception vers réflexe, et son lien réciproque accolé en pointillé; lien direct réflexe vers perception, et son lien réciproque accolé en pointillé. Le niveau de gris indique le niveau d'activation: plus il est foncé, plus le niveau est élevé. La ligne entourant la perception indique qu'elle est vérifiée. Celles entourant les réflexes indiquent qu'ils sont sélectionnables.

Les équations d'évolution des niveaux d'activation dans les perceptions et les réflexes sont donc :

$$\begin{cases} \alpha_{P_j}^{t+1} = \alpha_{P_j}^t + \theta_{direct} \sum_{C_i \in \mathcal{C}} f_{C_i P_j} \alpha_{C_i}^t \\ \quad + \theta_{inverse} \sum_{C_k \in \mathcal{C}} f_{P_j C_k} \alpha_{C_k}^t \\ \quad + \sum_{B \in \mathcal{B}_{P_j}} \alpha_B + v_{P_j}(s^t) \\ \alpha_{C_j}^{t+1} = \alpha_{C_j}^t + \theta_{direct} \sum_{P_i \in \mathcal{P}} f_{P_i C_j} \alpha_{P_i}^t \\ \quad + \theta_{inverse} \sum_{P_k \in \mathcal{P}} f_{C_j P_k} \alpha_{P_k}^t \\ \quad + \sum_{B \in \mathcal{B}_{C_j}} \alpha_B \end{cases}$$

avec \mathcal{C} l'ensemble des réflexes, \mathcal{P} l'ensemble des perceptions, \mathcal{B}_{P_j} et \mathcal{B}_{C_j} l'ensemble des buts se projetant sur P_j et C_j , s^t la valeur des capteurs à l'instant t , et en considérant que si un lien n'existe pas, cela équivaut à lui donner une force nulle.

Finalement, on normalise l'activation pour les perceptions et les réflexes afin d'éviter une divergence. Le réseau se stabilise au bout de quelques pas quand les conditions temporelles sont stables. On considère ensuite l'ensemble des réflexes sélectionnables, c'est-à-dire ayant toutes leurs perceptions obligatoires vérifiées. On choisit alors parmi eux le réflexe qui va être exécuté, soit de manière déterministe, en prenant celui dont le niveau d'activation est le plus fort, soit de manière stochastique, en utilisant une distribution de probabilité proportionnelles aux niveaux d'activation des réflexes sélectionnables.

2.3 Interprétation

Les réflexes et les perceptions que l'on donne au robot sont choisis par le concepteur. Il fixe également la connectivité du réseau, ainsi que la force des liens. Enfin, il fixe les buts à atteindre et oriente donc le comportement du robot pour remplir la mission qui lui est attribuée. Pour arriver à fixer tous ces paramètres, il faut mieux comprendre comment s'effectue la sélection de l'action.

On peut interpréter le mécanisme que nous venons de décrire en considérant que les perceptions cherchent à devenir vraies, et les réflexes à être choisis pour diriger l'action du robot. Il s'opère donc entre une compétition au sein de chaque groupe. La situation instantanée du robot et ses objectifs biaisent cette compétition. Les perceptions reçoivent de l'activation des réflexes, et réciproquement. On a donc également une coopération entre les deux groupes grâce aux liens. Ceux-ci fonctionnant à double sens, et pouvant avoir deux orientations possibles, ils se répartissent en quatre types différents dont les rôles peuvent être distingués :

- Lien direct d'une perception P vers un réflexe R . Le niveau d'activation de P est transmis par ce lien à R . Ce lien représente en quelque sorte la pertinence d'utiliser R dans la circonstance représentée par P . Si ce lien est obligatoire R ne pourra être choisi que si P est vérifiée ;

- Lien réciproque de R vers P . R aide par ce lien P à devenir active par rétro-propagation, et donc à être lui-même plus facilement activable, en particulier si le lien direct est obligatoire;
- Lien direct d'un réflexe R' vers une perception P' . Ce lien existe quand le déclenchement de R' peut provoquer au bout d'un certain temps la vérification de P' ;
- Lien réciproque de P' vers R' . P' aide ainsi R' à devenir plus actif, et donc plus facilement sélectionné. Si R' est choisi par le robot, P' a plus de chance de devenir vraie, et ce lien lui permet donc de faciliter sa vérification.

Le mécanisme tient compte à la fois de l'environnement et des objectifs du robot. Si l'influence de l'environnement n'intervenait pas, seuls les liens réciproques seraient nécessaires à la réalisation des objectifs du robot, en rétro-propageant l'activation introduite par les buts. On a donc une forme de planification qui s'opère dans le réseau. Mais, grâce à l'activation introduite par les perceptions, le robot vérifie les conséquences de ces actes, ce qui lui permet de corriger son comportement en cas de problème. Il n'y a donc pas de plan à proprement parlé puisque le choix est remis en question à chaque instant. Le mécanisme conserve donc une certaine réactivité. On peut moduler l'influence relative de la planification et de la réactivité en réglant le rapport $\theta_{direct}/\theta_{inverse}$. En augmentant θ_{direct} , on favorise la réactivité, en augmentant $\theta_{inverse}$, on favorise la planification. Le réflexe ayant la plus grande activation est alors le meilleur compromis entre l'accomplissement des objectifs et l'exploitation des circonstances actuelles.

Si on est obligé d'utiliser une sélection d'action stochastique³, on ne peut pas effectuer de choix de comportement à chaque instant. En effet, contrairement à ce qui se passe dans une simulation symbolique des actions, l'effet des réflexes n'est pas instantané. Il faut leur laisser un certain temps d'exécution pour qu'il puisse obtenir les résultats souhaités. Comment savoir quand effectuer une nouvelle compétition? On ne peut pas attendre que toutes ses conséquences soient vérifiées, car celles-ci varient en fonction des circonstances, et leur vérification n'est qu'une probabilité. De plus, si un échec survient, on peut attendre un temps infini avec un tel système. La première solution que nous avons adoptée a été d'effectuer une compétition à chaque fois que la valeur de vérité d'une perception change ou que le robot se trouve immobile⁴.

3. Par exemple quand on veut faire de l'apprentissage.

4. Et donc que les perceptions ne peuvent plus évoluer du fait de l'action du robot.

2.4 Interprétation en terme d'agents

L'interprétation en terme d'agents de notre architecture est assez naturelle. Les perceptions et les réflexes de notre réseau cherchent à s'exprimer soit en rendant leur valeur de vérité vraie pour les premiers, soit en prenant le contrôle de l'action du robot pour les seconds. On peut donc les considérer comme des agents cherchant à satisfaire des objectifs individuels. Dans le cadre de cette interprétation nous qualifierons les perceptions d'agents perceptifs, et les réflexes d'agents comportementaux. Le réseau définit le mode d'interaction de ce système multi-agents. Les agents perceptifs et comportementaux se livrent à une coopération par le biais de la circulation d'activation. Les agents comportementaux se trouvent en compétition à chaque changement d'agent perceptif. La prise de décision est distribuée et émerge de la dynamique des interactions entre agents. Bien que cette interprétation ne soit pas indispensable pour la compréhension des phénomènes, elle la facilite et permet de faire le lien avec les réflexions sur l'émergence de propriété dans les systèmes distribués.

2.5 Apprentissage

L'apprentissage permet soit d'affiner les liens existants au départ soit de créer de nouveaux liens. L'apprentissage se fait en ligne, c'est-à-dire qu'il n'y pas de distinction entre une phase d'apprentissage et une phase d'exploitation, comme cela est le cas pour les méthodes de rétro-propagation de gradient dans les réseaux de neurones à couches; autrement dit un lien peut être exploité en même temps qu'il apprend.

Plus précisément, les liens des réflexes vers les perceptions sont appris par estimation statistique. On évalue statistiquement la probabilité d'obtenir la vérification ou l'infirmité d'une perception P après l'exécution d'un réflexe R . Si un lien de R à P existe déjà, sa force prendra la valeur de l'estimation de la probabilité:

$$f_{RP} = \frac{n_{succès}}{n_{essais}}$$

où $n_{succès}$ est le nombre de fois où R a entraîné la vérification de P , et n_{essais} le nombre de fois où R a modifié P . S'il n'existe pas, il est créé avec une force 1. La loi des grands nombres assurera la convergence de la force du lien. Si cette probabilité est conditionnée par des paramètres cachés au robot, l'estimation s'adaptera progressivement au cours des variations de ces paramètres, toujours d'après la loi des grands nombres.

D'autre part, les liens des perceptions vers les réflexes peuvent être appris par renforcement. Toute la difficulté vient de l'obtention du signal de renforcement⁵, qui dépend du but. En général, on peut

5. Signal permettant de mesurer la progression en direction d'un but

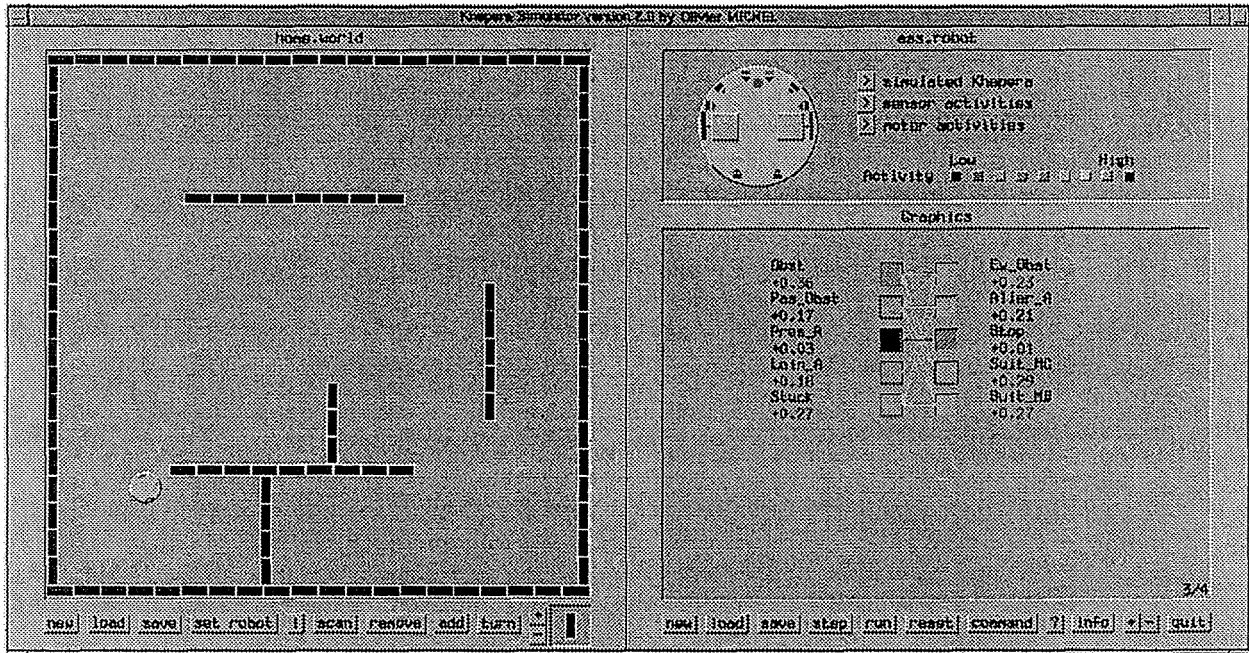


FIG. 2 - Vue du simulateur de robot Khepera qui nous a servi de base pour nos expériences.

utiliser un renforcement négatif au cours du mouvement si le robot vient à nuire à son intégrité physique, par exemple en heurtant un obstacle. En revanche, les renforcements positifs ne sont faciles à obtenir qu'à l'accomplissement du but. Il faut alors disposer d'un algorithme de rétro-propagation du renforcement, mais cela est toujours hasardeux. Pour avoir un renforcement instantané, il faut disposer d'un signal de renforcement disponible directement à partir de l'environnement, ce qui est une hypothèse très forte. L'apprentissage par renforcement n'a pas encore été utilisé dans l'architecture.

2.6 Support Expérimental

Nous avons mené deux séries d'expériences pour tester et mettre au point les idées précédentes. Ces expériences ont été effectuées en simulation sur un logiciel reproduisant un robot de type Khépéra, écrit par O. Michel.

Les comportements de base que nous avons implantés sont tous organisés autour d'une même architecture neuronale, comportant deux moto-neurones dont les sorties sont reliées directement aux moteurs, et dont les entrées sont les capteurs de luminosité et de proximité du robot, ou éventuellement des capteurs virtuels que nous avons rajouté au simulateur. La fonction de transfert de ces moto-neurones est une double sigmoïde :

$$f(x) = \frac{a}{1 + e^{T-bx}} - \frac{a}{1 + e^{-T+bx}}$$

où a est une constante de calibration pour les entrées des moteurs, b une constante servant à la fois à calibrer les entrées et à jouer sur la raideur de la pente, et enfin T est une constante servant éventuellement à induire un seuil de réaction dans le neurone. En effet, comme on peut le voir sur la figure 3, la réponse du neurone pour une excitation entre $-T/2$ et $T/2$ est négligeable.

Les poids des différentes liaisons des capteurs vers les moto-neurones sont notés ainsi :

- luminosité: w_{l1}^i et w_{l2}^i pour la connexion du i -ème capteur respectivement au neurone 1 et 2.
- proximité: w_{p1}^i et w_{p2}^i pour la connexion du i -ème capteur respectivement au neurone 1 et 2.
- constante: o_1 et o_2

Nous avons réussi à partir de ce formalisme à produire un certain nombre de réflexes: *évitement d'obstacle, suivi de mur à gauche ou à droite, suivi de couloir, aller vers une balise, tourner sur place à gauche ou à droite, aller tout droit*. Le tableau 1 récapitule les poids de connexions. Seuls ceux du neurone 1 sont montrés, car ils sont en général symétriques. Ces valeurs sont données à titre indicatif. Il est sans doute possible d'obtenir des comportements similaires, c'est-à-dire que nous qualifierions de la même manière, en modifiant substantiellement ces valeurs. Il est à noter que ces noms n'indiquent pas que les actions associées soient descendantes. Ils ont été donnés après avoir observé les comportements résultants.

Comportement	w_{p1}^0	w_{p1}^1	w_{p1}^2	w_{p1}^3	w_{p1}^4	w_{p1}^5	w_{p1}^6	w_{p1}^7	o_1	o_2
Évitement d'obstacle	0	0	0	1	1	1	0	0	0.5	0.5
Suivre un mur gauche	0	0	0	0	1	1	0	0	0.5	0.6
Suivre un mur droite	0	0	0	0	1	1	0	0	0.6	0.5
Suivre un couloir	1	0.5	-1	-1	-0.5	1	1	1	0	0
Tourner sur place g.	0	0	0	0	0	0	0	0	-1	1
Tourner sur place d.	0	0	0	0	0	0	0	0	1	-1
Aller tout Droit	0	0	0	0	0	0	0	0	1	1

TAB. 1 - Coefficients des réflexes du Khépéra.

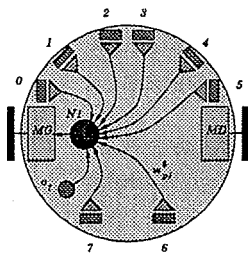


FIG. 3 - Notation des différents éléments de l'architecture.

2.7 Expérimentation

La première expérience avait pour but de vérifier l'efficacité de la sélection d'action. Le réseau très simple que nous avons employé comportait un réflexe d'évitement d'obstacle et un réflexe permettant au robot de se diriger vers une balise. Il disposait de perception lui permettant de déterminer la présence ou l'absence d'obstacle, et de déterminer s'il était près ou loin de la balise. L'objectif du robot était d'être près de la balise, c'est-à-dire de voir la perception *près-de-la-balise* vraie. Nous avons constaté un bon comportement du robot dans un grand nombre de cas. Le robot évite correctement les obstacles, ce que nous savions puisque le comportement avait été testé seul. Il se dirige correctement vers la balise. L'ennui vient bien évidemment du fait que n'ayant aucune représentation globale de l'environnement, le robot ne choisit pas un chemin globalement optimal. Il peut même rester coincé dans un cul-de-sac. Ceci survient quand la balise est orientée de manière défavorable. La solution la plus économique pour résoudre ce problème est de remplacer le réflexe d'évitement d'obstacle par un réflexe de suivi de mur. Le robot remonte alors le long des parois du cul-de-sac, au lieu de tourner en rond au fond de celui-ci. De même, il arrive que le robot reste coincé dans un coin. Ceci provient du fait que le comportement d'obstacle possède un point d'équilibre. On voit donc que les problèmes

rencontrés sur ce premier exemple très simple sont imputables au fait qu'un des comportements les plus élémentaires possède des défauts, ou qu'il manque un comportement élémentaire adéquat pour résoudre un problème. Pour pallier à ces défauts, nous avons développé une gamme de comportements élémentaires plus étendue que nous avons mis au point chacun indépendamment. Nous avons essayé ces nouveaux comportements dans de nombreux exemples d'environnement, ce qui nous permet d'obtenir un robot circulant aussi bien dans un labyrinthe étroit que dans un espace dégagé.

Nous avons ensuite examiné les résultats de l'apprentissage des liens des réflexes vers les perceptions chargés de prédire les conséquences des réflexes. Ces résultats ne nous sont pas apparus très satisfaisants. En effet, le comportement de direction vers la balise est un comportement à long terme censé aboutir à la perception *près-de-la-balise*. Malheureusement, le robot rencontrant beaucoup d'obstacles sur sa route, il apprend que la conséquence du réflexe *aller-vers-la-balise* est de rencontrer un obstacle. Il faudrait pour obtenir un meilleur comportement utiliser une perception déterminant si on se rapproche de la balise. Ceci fournirait un résultat immédiat qui permettrait d'apprendre correctement les conséquences à court terme du réflexe. Le fait d'arriver près de la balise doit se trouver à un niveau temporel plus élevé.

3 Architecture à plusieurs niveaux

3.1 Motivation et Principe

Dans le type de réseau que nous avons exposé, le problème de la segmentation temporelle n'est que partiellement résolu. En effet la disparité entre les temps nécessaires aux réflexes pour s'exécuter correctement est telle que les résultats de l'apprentissage deviennent incohérents : il favorise les réflexes donnant des résultats à court terme et rend improbable la sélection de comportements plus lents, mais indispensables. Le résultat final est que le robot apprend très peu de choses.

Tyrrell met en lumière le problème précédent ([18], [19]) ainsi que d'autres inhérents à l'architecture de Maes. La solution envisagée passe par l'étude d'autres mécanismes de sélection d'action proposés notamment par des éthologues. Ces mécanismes présentent tous la particularité d'être organisés sous forme d'une hiérarchie. Ainsi, le modèle de Baerends décrivant le comportement de nidification des oiseaux ou celui plus général de Tinbergen [17] proposent une hiérarchie fonctionnelle de comportements. Un comportement plus général utilise plusieurs sous-comportements plus élémentaires de manière ordonnée afin de réaliser la satisfaction des contraintes motivationnelles. C'est aussi une hiérarchie de comportements que propose Minsky [13] pour rendre compte du comportement de jeu des enfants, et plus généralement de l'organisation des comportements « intelligents ». Il existe deux points communs à tous ces modèles hiérarchiques : plus on se place haut dans la hiérarchie, plus les phénomènes d'évolution sont lents (temps caractéristiques d'évolution grands). D'autre part plus on est à un niveau élevé de la hiérarchie, plus les comportements qui le composent sont abstraits. Prenons un exemple éthologique : le comportement d'alimentation pourra être considéré de plus haut niveau que les comportements plus élémentaires qui le constituent : rechercher de la nourriture, la transporter dans sa tanière pour la stocker, la manger, etc. Chacun de ces comportements plus élémentaire pourra être lui-même décomposé à des niveaux encore inférieurs.

Nous avons donc modifié l'architecture précédente pour y introduire une hiérarchie de comportement. L'architecture est désormais répartie sur plusieurs niveaux de compétition. Chacun des niveaux est organisé comme le réseau de sélection d'action que nous avons déjà présenté. Chaque niveau est donc composé d'agents perceptifs et comportementaux, ainsi que d'un ou plusieurs buts. Les agents comportementaux de chaque niveau peuvent maintenant être de deux types : des réflexes comme précédemment, ou des réseaux de niveau inférieur. Quand un agent implémentant un réseau est sélectionné, il permet à ce réseau de mener lui-même une compétition. L'ensemble des réseaux est donc par le biais de ces agents comportementaux organisé sous la forme d'un arbre dont les feuilles sont les réflexes. En terme de système multi-agents, ceci revient à considérer que chaque système peut être lui-même considéré comme un agent de niveau supérieur. En quelque sorte, les propriétés fonctionnelles émergentes du système de niveau inférieur sont matérialisées dans une structure de niveau supérieur. Les propagations d'activation se font dans chaque réseau de la même manière que pour l'architecture à un seul niveau. Les agents perceptifs des niveaux supérieurs changent moins fréquemment de valeur de vérité, ce qui provoque des compétitions moins fréquentes. On a donc bien une dynamique de plus en plus lente quand on considère les niveaux su-

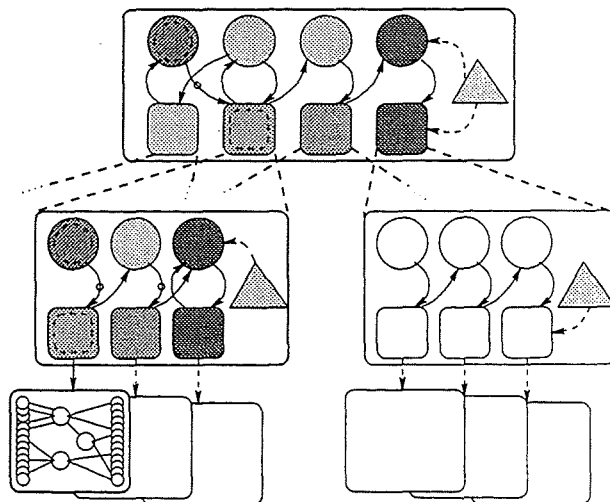


FIG. 4 - La nouvelle architecture est organisée sous la forme d'une hiérarchie de réseaux de sélection d'action. Les agents comportementaux sont soit des réflexes, soit des « sous-réseaux » qui sont reliés à l'agent les représentant par des pointillés. Les sous-réseaux choisis à chaque niveau forment une chaîne partant du niveau le plus élevé et aboutissant au réflexe en cours d'exécution.

périeurs.

3.2 Résultats expérimentaux

Notre deuxième expérience visait à tester l'architecture hiérarchique. Cette architecture devait résoudre le problème précédent, ainsi que le problème de la combinaison de deux buts à long terme. Dans cette expérience le robot a deux objectifs permanents à remplir : arriver à proximité de deux balises émettant des signaux de natures différentes. Le résultat de ce genre d'expérience avec une architecture à un seul niveau est très mauvais. Le réseau remet en question son objectif à chaque compétition de sorte que le robot oscille entre les deux objectifs sans parvenir à en remplir un complètement (cf. figure 5). Ce phénomène d'oscillation avait déjà été rapporté sur l'architecture de Maes [19]. L'effet de la hiérarchie est de répartir correctement ces remises en question de but. L'architecture utilisée s'organise en deux niveaux. Le plus lent regroupe les perceptions *près-de-la-balise-A*, *loin-de-la-balise-A*, *près-de-la-balise-B*, *loin-de-la-balise-B*, et deux comportements implémentant deux sous-réseaux. Chacun de ces sous-réseaux permet de diriger le robot vers une des deux balises. Il comprend les réflexes *évitement d'obstacle* et *aller-vers-A*, et les perceptions *près-de-la-balise*, *loin-de-la-balise*, *absence-d'obstacle* et *présence-d'obstacle*. Le robot ef-

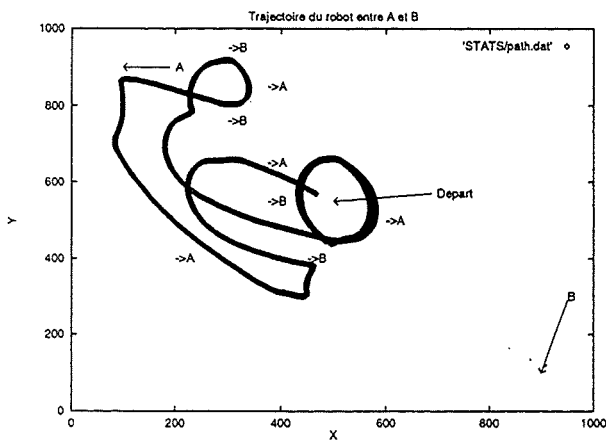


FIG. 5 - Trajectoire dans l'expérience 2 avec une architecture à un seul niveau. Le robot change d'objectif quasiment à chaque fois qu'il rencontre un obstacle. Les buts instantanés sont indiqués à côté de chaque portion de courbe. L'environnement utilisé est celui de la figure 3.

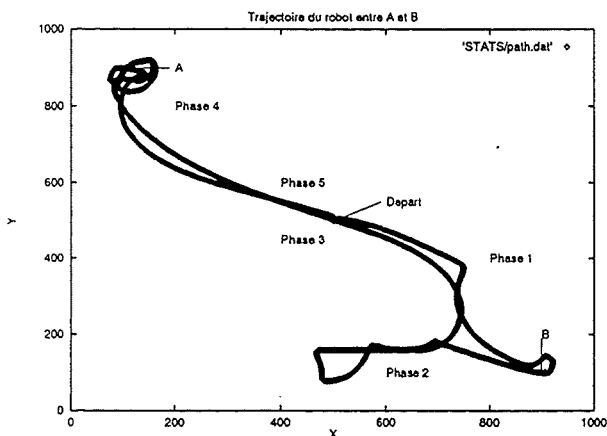


FIG. 6 - Trajectoire de l'expérience 2 avec l'architecture hiérarchique. Le robot arrive à concilier les deux buts. Le fait d'entrer dans un cul de sac (Phase 2) ne remet pas en cause le but aller en A.

fectue alors des allers et retours entre A et B sans remettre en question son comportement lorsqu'il rencontre un obstacle. La seule chose qui peut le faire changer de comportement de niveau supérieur est un blocage durable du robot.

4 Discussion

Nos résultats prouvent que nous avons bien dans le cadre de notre simulation une méthodologie de programmation du robot remplissant plusieurs des critères que nous nous étions fixés. La méthode est modulaire, puisqu'elle permet de mettre au point séparément plusieurs réflexes ou plusieurs réseaux, et de les combiner ensuite dans une sélection d'action. Elle permet au robot de remplir un objectif fixé a priori dans un environnement de topologie inconnue. Le comportement du robot est invariant dans une large gamme d'environnement. Une dégradation substantielle des performances constatées empiriquement signifie une différence de nature de l'environnement et la nécessité de fournir au robot des comportements adaptés à ce changement. L'architecture permet de combiner les objectifs fixés avec des contraintes de viabilité du robot. La méthode de conception est ascendante, dans la mesure où les représentations utilisées par le robot sont largement émergentes de son interaction avec l'environnement et résulte de l'apprentissage des résultats de ses comportements.

L'architecture est compatible avec l'utilisation de fonctions préexistantes sur un robot. En effet, des programmes fournissant au robot des capacités motrices ou perceptives particulières peuvent être réutilisés dans notre architecture en les encapsulant dans un agent perceptif ou comportemental. Il faut néanmoins tenir compte du temps d'exécution de ce programme pour l'introduire au niveau hiérarchique adéquat.

L'exemple de simulation que nous avons choisi manque certainement de réalisme et de complexité pour conclure de manière définitive à l'efficacité de la méthode dans un cas réel. En effet, le monde réel est en général beaucoup plus difficile à appréhender. Les capteurs sont souvent très imparfaits et les résultats des actions beaucoup plus aléatoires que dans une simulation. Une implantation de l'architecture est donc prévue sur une plate-forme réelle.

Notre architecture permet de concilier deux aspects fondamentaux des approches classiques et biologiques, la planification et la réactivité. Le formalisme employé est compatible avec les deux approches, puisqu'un comportement peut être aussi bien vu comme un couplage sensori-moteur que comme une règle dans un graphe de décision. De plus, le caractère récursif de la hiérarchie laisse entrevoir la possibilité d'augmenter la complexité des comportements de manière importante.

La possibilité de considérer un système multi-agent comme un agent de niveau supérieur ouvre la voie à une complexification des tâches que l'on peut espérer accomplir grâce à notre architecture. Notre deuxième expérience montre que l'accumulation de perceptions et de comportements sans organisation hiérarchique n'améliore pas forcément les performances de l'architecture. Il faut donc disposer d'une méthode de hiérarchisation des comportements. Pour l'instant, celle-ci est réalisée par le concepteur qui évalue empiriquement les performances du robot. Si nous disposions d'une méthode automatique de hiérarchisation, un véritable développement serait possible en permettant la création de niveau de plus en plus élevé de comportements. La représentation apprise par le robot ne serait plus restreinte à un seul niveau temporel, mais constituerait véritablement un modèle prédictif de l'environnement lui permettant peut-être de construire des comportements cognitifs.

Une autre limitation du système à l'heure actuelle se trouve dans la catégorisation a priori des perceptions que le concepteur fournit au robot. Le dépassement de ce problème semble possible grâce à des techniques neuronales d'apprentissage non-supervisé.

Pour prendre un exemple concret, la complexification de l'architecture pourrait ainsi permettre au robot d'apprendre à reconnaître des objets caractéristiques de son environnement. En effet, on pourrait dire qu'il reconnaît ces objets s'il associe à leur présence un comportement particulier utile et stable. Par exemple, cet objet pourrait lui servir à se repérer spatialement. Toute la difficulté vient du fait qu'il faut pouvoir extraire des informations sensorielles brutes les régularités indiquant cette présence. Ceci doit être fait par le robot de manière totalement autonome, à l'aide d'essais et d'erreurs, de la même manière qu'un enfant apprend à connaître le monde qui l'entoure. On passerait ainsi d'un robot réagissant instinctivement et immédiatement à son environnement à un robot capable de discerner des régularités sur lesquelles il peut construire une stratégie d'action. Les objets ainsi extraient ne seraient pas alors des représentations symboliques a priori, mais contiendraient une signification puisque le robot aurait le moyen de les relier à une expérience sensible.

Bien que ces travaux soient trop peu avancés pour donner une réalité à cette dernière réflexion, il est clair que notre problématique se rapproche de la construction d'une forme d'identité simple pour le robot. Afin de lui permettre une adaptation adéquate aux évolutions de son environnement, celui-ci doit pouvoir retrouver à travers toutes les informations qu'il reçoit les invariances et les liens qui les relient à son action. La construction de cette représentation est un développement puisqu'il doit agir en même temps qu'il apprend.

5 Conclusion

Les approches classiques et biologiques ont conduits jusqu'à présent à des méthodes très différentes de conception des architectures robotiques. Des différences terminologiques et de principes ont handicapé leur rapprochement. Cependant, une analyse de leurs défauts et de leurs avantages sur le plan théorique permet d'aboutir à un formalisme intermédiaire évitant les principaux écueils. Ce formalisme, sans juxtaposer naïvement les deux modèles dans une organisation commune, propose une méthode récursive et incrémentale devant permettre de combler le fossé entre les descriptions fonctionnelles du système artificiel que l'on souhaite réaliser et les éléments structurels comportementaux disponibles actuellement en robotique. Des arguments expérimentaux ont pu être avancés pour montrer en quoi cette approche est prometteuse. Enfin, les thèmes abordés dans cette étude peuvent être rapprochés des réflexions menées par les sciences cognitives sur la structuration du comportement et de l'identité.

Références

- [1] V. Braitenberg. *Vehicles: Experiments in synthetic psychology*. MIT Press, Cambridge, Massachusetts, USA, 1984.
- [2] R.A. Brooks. *Achieving artificial intelligence through building robots*. A.I. Memo 899, MIT, 5 1986.
- [3] R.A. Brooks. A robust layered control system for a mobile robot. *IEEE Journal of Robotics and Automation*, RA-2(1):14-23, 3 1986.
- [4] R.A. Brooks. *Intelligence without reason*, 1990.
- [5] R.A. Brooks and L.A. Stein. *Building brain for bodies*. A.I. Memo 1439, MIT, 1993.
- [6] H.J. Chiel, R.D. Beer, and R.D. Quinn. Robustness of a distributed neural network controller for locomotion in a hexapod robot. *IEEE Transactions on Robotics and Automation*, 8(3):293-303, 6 1992.
- [7] J. Ferber. *Les systèmes multi-agents*. InterEditions, 9 1995.
- [8] C. Ferrell. *Robust agent control of an autonomous robot with many sensors and actuators*. Master's thesis, MIT, Massachusetts Institute of Technology Artificial Intelligence Laboratory, 1993.
- [9] S. Giszter. Reinforcement tuning of action synthesis in a virtual frog. In D. Cliff, P. Husbands, J.A. Meyer, and S.W. Wilson, editors, *From animal to animat 3*, pages 293-300, 1994.

RÉFÉRENCES

- [10] P. Maes. How to do the right thing. *Connection Science Journal*, 1(3), February 1990.
- [11] P. Maes. Situated agents can have goals. *Robotics and Autonomous Systems*, 6:49–70, 1990.
- [12] P. Maes. Learning behavior networks from experience. In F. Varela and P. Bourguine, editors, *Towards a Practice of Autonomous Systems*, pages 48–57, 1991.
- [13] M. Minsky. *La société de l'esprit*. InterEditions, 1989.
- [14] D.W. Payton. An architecture for reflexive autonomous vehicle control. In *Proceedings of IEEE Conference on Robotics and Automation*, pages 1838–1845, 1986.
- [15] J.K. Rosenblatt and D.W. Payton. A fine-grained alternative to the subsumption architecture for mobile robot control, 1988.
- [16] Rumelhart and McClelland. *Parallel distributed processing*. MIT Press, 1986.
- [17] N. Tinbergen. The hierarchical organization of mechanisms underlying instinctive behaviour. *Experimental Biology*, (4):305–312, 1950.
- [18] T. Tyrrell. *Computational Mechanisms for Action Selection*. PhD thesis, University of Edinburgh, 1993.
- [19] T. Tyrrell. An evaluation of maes's bottom-up mechanism for action selection. *Adaptive Behavior*, 2(4):307–348, 1994.

IDENTITY AS FUNCTION OR STRUCTURE: AN ILLUSTRATION ON ROBOT DEVELOPMENT

Eric Dedieu, Olivier Lebeltel, Pierre Bessière

LEIBNIZ laboratory — IMAG, Grenoble, France
email: eric.dedieu@imag.fr

Introduction

Robot design is always oriented toward specific and predefinite purposes. Infrared sensors are dedicated to measuring distances to obstacles, and their sensitivity to colours is considered an inconvenience. Wheels are built to let robots move in a well-defined way, and slipping is considered a parasitic move. Light-following reflex behaviours are intended to let robots follow lights, not to let them avoid obstacles — and despite the amount of work in robot navigation, few people have noticed that following lights often leads to avoiding obstacles (because of their shades - see below).

So, in robot design, all devices or programs are given a definite function (or task) which can be considered their “identity”, in the sense that it is a specific invariant characteristic that should guide their use and development.

However this functional notion of identity is an engineering notion. Nature does not seem to work in that way. For example, wings were not “designed” to let animals fly. They probably were at first heat-regulating devices, the structure of which happened unexpectedly to give some gliding capacities. Because it gave a selective advantage, this “side-effect” evolved and became the main function of the structure — eventually yielding the wings we know (Gould 1991b).

Many phenomena, from biological structures to key arrangements on keyboards (Gould 1991b), are better explained a-posteriori, through such an evolutionary perspective, than by the pure functional view that they were a-priori

designed for their current use.¹ “Contingency” is the notion that the origins of important facts are often unexpected events — “accidents”, as it were —, which yet happened to yield important consequences in the particular context of their occurrence.

Though being an essential aspect of animal evolution, contingency has never been considered in building robots. A robot wing would be once and for all labeled as a device intended to let the robot fly or glide. Such an “engineering” frame of mind could never lead to achieve the wing solution opportunistically through contingent evolution — which was Nature’s way, though.

The goal of this paper is to put into question the usual functional basis for identifying robotic resources (i.e., naming, using, designing, developing them). We propose a structural basis instead, allowing functions to co-evolve with structures, rather than control their evolution. We illustrate this with a robotic experiment on the Khepera robot, to show how such concerns may be approached in robotics.

Embodiment and opportunistic behavior

The fact that a robot is in sensori-motor interaction with its environment is usually considered a *problem*, rising many hard and tedious constraints like calibration, mechanical or electronic adjusting, motor

¹ See the books from S.J. Gould (1988, 1991a, 1991b) for arguments against the “adaptationist” view that every detail of evolution was directed toward optimal adaptation.

powering, sensor noise, etc. These constraints are regarded as taking time and, all things considered, *parasitic* compared to “truly” interesting cognitive issues like high-level navigation, planning or task selection. Thence, in order to “get to the point” more easily, many robotic systems are developed as computer simulations, somewhat assuming that a physical implementation would raise purely technical problems. Doing this, however, just hides the specificity of robotics compared to e.g. artificial intelligence, which is, indeed, the embodied nature of robots. The consequences of embodiment are not just parasites — they are the very ones to be understood and, even, *exploited*. Far from being just a technical engineering problem, embodiment can be considered a scientific subject to study and experiment with.

This idea, however, requires an unusual perspective on robotics. This field has been usually approached with an engineer's point of view: given a task to perform, the engineer tries to build a robot or computer program achieving it. Robotic systems, thus, are designed as problem-solving or task-achieving ones. But robotics can also be considered within *cognitive science*. The purpose of cognitive science is to model animal and human cognitive abilities, i.e., understand and reproduce the mechanisms of “intelligent” behaviour. Though task achievement is an aspect of intelligent behaviour, there are other aspects that can be approached through building robot. In particular, our own research subject is *opportunistic behavior*.

“Opportunistic behavior”, here, does not mean watching for an expected opportunity to do something that you were somehow intending to do. Rather, the idea is, in response to an unanticipated situation, to really invent or improvise a new task, or to alter an existing one. This is in contradiction with the engineer's “functional” approach, where design is completely guided by and toward

predefinite high-level tasks. Tasks are no longer predefinite, but may change through a development process. It is a major conceptual change in the usual design of robots, proposing that the very tasks of the robot might change opportunistically, and *no longer be a constant guide for design*.

Yet, animal evolution suggests to robotics that this idea is worth trying. The wing example above showed a given physical device having its function modified for contingent reasons (from heat regulation to gliding). Generally, the functional evolution of biological structures seems an important process for animal diversity and adaptation (Gould 1988, 1993). The first lesson of animal evolution that we wish to consider in artificial systems is that structures are not developed toward predefinite functions, but that there is a co-evolution of structures and functions.

Let's illustrate by an example how this abstract idea can lead to concrete research in a robotics framework.

A true story

Some years ago, we observed an unusual robot behavior, which was the origin of our current concerns.

The KitBorg is a small mobile robot², commanded by its left and right wheel speeds, and using three photocells, one frontal, one 90° right and one 90° left. It was behaving according to the most activated cell: for the frontal cell it drove ahead, for the left cell it turned left, for the right cell it turned right. This was just a test reflex to make the robot follow a lamp. But when putting out the lamp, the robot did something totally unexpected — which the witnesses recalled spontaneously several years later. The robot turned back, heading toward a large window *and avoiding obstacles in its path*, then turned toward a half-closed door. It managed the

² Designed and developed by Aleph Technologies, Grenoble, France.

narrow passage to the office room behind, and finally kept “dancing” under the light bulb, as if showing that it had reached its goal. This was what everyone usually try to do with mobile robots!

After thought, this behavior is no longer surprising. Obstacles were avoided because of their shades, then the robot was caught by the light cone coming through the door, which guided it to the office. The final dance was just indecision in a local maximum of light intensity.

But, though every witness knew how the KitBorg was programmed, this interpretation in navigation terms forced itself upon each of them.

To bring out the lesson of this story, let's focus on the unexpected property that a light-seeking behaviour frequently drives a robot to avoid obstacles. Here, obstacle-avoidance was not intended, but the particular context of the experiment let it *emerge* from another behaviour. This was a good surprise, but thanks to luck only; and there was no systematic way to exploit this fact.

However, the robot was equipped with sonars — useless this far. Imagine that somehow it could automatically observe, while moving, an interesting relation between the sonar signals and its rotation speed:

- o when the right sonars are more excited than the left ones, I turn left,
- o when the left sonars are more excited than the right ones, I turn right,
- o when there is no significant difference, the movement is unconstrained.

This is a description of an empirically observed relation. By changing this description to a specification for a reflex (i.e., how to act given an immediate sensory situation), this relation could be exploited to produce a new behavior, namely obstacle avoidance.³

³ Note that such a transformation is not immediate, all indeterminations have to be solved first. For example, the movement in free space is unconstrained, but a definite value must eventually

In a sense, this dependency can be seen as *representing an emergent functionality, because it now allows to exploit it, to take advantage of it*, by making a new task available (avoiding obstacles).

This example illustrates an incremental methodology for robot opportunistic development, which is one of our major research areas.

Roughly, the idea is to learn “internal representations” of a robot's behaviour as *relations between sensori-motor variables that can be empirically observed in given contexts*. These are called “sensori-motor representations”, because they are expressed in purely sensori-motor terms. Technical and experimental developments are outside the scope of this paper; let's just say that we use probabilistic structures on which we can compute probabilistic inference.⁴

The fundamental point is that the interpretation (function) of any such inference is not fixed but context-dependant, and can be opportunistically used to program different behaviours in different contexts. If the context changes due to unexpected and contingent causes, the robot's working can be opportunistically “tinkered” to keep on being relevant. This kind of development process, which is usually regarded as a “debugging” process, may now be approached as an open opportunistic (if not creative) methodology.

Let us just emphasize some more points about this approach that we are developing:

- o Opportunistic development is not an automatic process. For the moment, it

be chosen to command the robot. We have chosen a probabilistic framework to tackle these issues.

⁴ For French readers, full developments and experimental descriptions are available in the author's PhD thesis (Dedieu 1995), which can be found on the website:

<http://leibniz.imag.fr/LAPLACE/Mediatheque/Bibliot heque/Abstracts/Dedieu95a.html>.

is still a methodology. However, we are trying to make this methodology as *systematic* as we can, and to make as much as we can automatic (for example, the learning processes to experimentally identify the sensori-motor representations).

- o The concern underlying this approach is the origin and genesis of sensori-motor representations more than their performances.
- o The inspiration from animal evolution has to be adapted to the robotics framework. We focus on software development, because it is more amenable to experimentation than animal-like hardware evolution. However, doing this, our purpose remains to understand the specific nature of a sensori-motor (rather than symbolic) system in a natural (non-controlled) environment.
- o The inability to recognize and manage *unexpected* (i.e., non a-priori modeled) situations or events is currently one of the strongest limits of artificial vs. natural systems. Contingent re-interpretation of opportunistically developed sensori-motor representations seems a first step to tackle this hard problem.
- o This can be related to the "symbol grounding problem" (Harnad 1990), which is the difficulty to relate functional terms to sensori-motor terms and experience. The solution sketched here is to let the robot develop its own representations (i.e., based only on sensori-motor terms). Only after that, we may try to interpret them depending of how they *actually* may be used.

Let's give another example, where the "parasitic" effect of unexpected events is more obvious than in the story above. A mobile robot was doing path-planning and path-following in our laboratory, using a geometric map of it. When passing a door, however, a small bulging copper door-step caused it to slip and collide the wall.

Such an unexpected event can typically be dealt with by removing the door-step (thus controlling the environment to make the model valid), or by adding more rubber to the wheels to prevent slipping (thus altering the hardware to make the model valid). A possible suggestion, however, would be to use this empirical observation to help recognise a door location (if you slip, you're likely to be at a door-step). Of course, this was not an intended capacity for the robot, and it hinders the intended navigation abilities. Yet, it allows a *continued incremental software development of the robot, without controlling the environment nor having to re-design the robot.*

Conclusion

Robots, unlike computer simulations or symbolic systems, are in sensori-motor interaction with their environments. Sensori-motor interactions are rich, complex, and hard to model. Their nature is badly conveyed by functional approaches, because these approaches are guided by predefined tasks or functions, which lead to regard unexpected events and incompleteness of models as parasites to get rid of. We would like, by contrast, to study and explore the contribution (rather than interference) of embodiment in developing cognitive abilities. The notion of contingency in animal evolution suggests that it is worth considering a co-evolution of biological structures *and their functions together*. This amounts to (1) redefining the identity of robotics within cognitive science as a framework to let such issues emerge and get studied, and to (2) redefining a structural notion of identity for robot development or evolution, in addition to the usual functional one. This paper presented a tentative approach to these issues in a robotics framework.

References

- Dedieu 1995 E. Dedieu, *La représentation contingente : vers une réconciliation des*

approches fonctionnelles et structurelles de la robotique autonome. Thèse de l'Institut National Polytechnique de Grenoble. Disponible à: <http://leibniz.imag.fr/LAPLACE/Mediatheque/Bibliotheque/Abstracts/Dedieu95a.html>.

Gould, S.J. 1988 *Le sourire du flamant rose*. (org. The Smile of the Flamingo). Seuil.

Gould, S.J. 1991a *Wonderful Life*. Penguin.

Gould, S.J. 1991b *Bully for Brontosaurus*. Penguin.

Harnad, S. 1990 The Symbol Grounding Problem. *Physica D*42.

Identification d'un opérateur pour le contrôle de processus dynamiques

Gilles Coppin^{1,2}, Eric Faure¹, Pascale Kuntz¹ and Dominique Snyers¹

¹ LIASC, Télécom Bretagne, B.P. 832, F-29285 Brest Cédex, France

² Thomson-CSF RCM CEB, 10 Av. de la 1^{ière} D.F.L., 29283 Brest Cedex
email: Gilles.Coppin@enst-bretagne.fr

1 Introduction

La coopération homme-machine met en jeu un ou plusieurs couples (opérateur, machine) devant remplir un certain nombre de tâches qui leur sont allouées. Et, devant la complexification croissante des outils logiciels nécessaires à la réalisation de ces tâches est apparue le besoin de mieux adapter le fonctionnement de la machine à l'opérateur. Afin de pouvoir proposer des aides adaptées à différents niveaux de connaissances et de performances ou de suivre son évolution dans le temps pour lui proposer des commandes spécifiques à ses besoins réels, il devient nécessaire de pouvoir caractériser avec précision un opérateur en interaction avec une machine.

Comme l'indique Visetti à propos du cadre spécifique des systèmes à base de connaissances ([12]), nous considérons ici que l'ajustement doit être effectué au sein même de la machine, et sur la base de l'observation des actions – conduites – comportements de l'opérateur. Ainsi, dès lors que la machine a identifié l'opérateur, et modifié son état ou ses fonctionnalités en conséquence, on arrive par rétroaction à des modifications du comportement de l'opérateur lui-même, par adaptation au nouvel état de la machine. Ce processus d'adaptation mutuelle est donc à l'origine d'une dynamique d'évolution permanente interne au couple (opérateur, machine) (cf. figure 1).

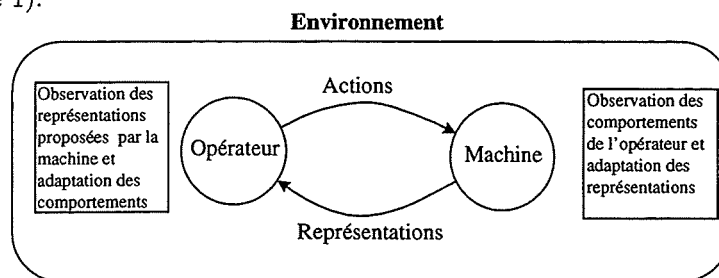


FIG. 1 – Description générale de la coopération homme-machine

Cette évolution conjointe permet en fait de considérer l'entité (opérateur, machine) comme un véritable système dynamique, dont les paramètres caractéristiques se situent au niveau des représentations (fournies par la machine) et des actions (effectuées par l'opérateur). Ainsi, «l'image cybernétique de la boucle de régulation qui symbolise entre les représentations et actions et donne l'idée que [la machine] cherche à maintenir un état stable en compensant les perturbations est dépassée [...] pour traduire la construction d'un équilibre dynamique en perpétuelle évolution entre l'opérateur et son environnement» ([13]).

Application au contrôle de processus

Bien que l'analyse d'un phénomène de coopération homme-machine doive à terme se placer au niveau global (opérateur, machine), il est nécessaire au préalable de munir la machine des moyens d'assurer le retour d'information participant à la dynamique

du système. Il faut par conséquent implanter dans la machine des mécanismes d'observation du comportement de l'opérateur qui puissent aboutir à des modifications adéquates des représentations. On doit donc avant toute chose définir ce que l'on entend par «action» ou «comportement de l'opérateur».

Nous nous plaçons pour cela dans le cadre du contrôle de processus dynamique, où les «actions» de l'opérateur peuvent être définies par des séquences d'événements élémentaires. Les événements élémentaires peuvent être des actions sensori-motrices élémentaires (dans le cas du pilotage ou du contrôle de systèmes réactifs par exemple), ou des choix élémentaires (au sein d'une stratégie de résolution de problème par exemple). Ce type de description est proche de la notion de « pattern de conduite » utilisée couramment dans les études d'ergonomie et de psychologie du travail ([13]).

Les travaux présentés ici concernent l'analyse de ces séquences d'événements élémentaires, et constituent la première étape de l'analyse globale du processus de coopération homme-machine (cf. figure 2).

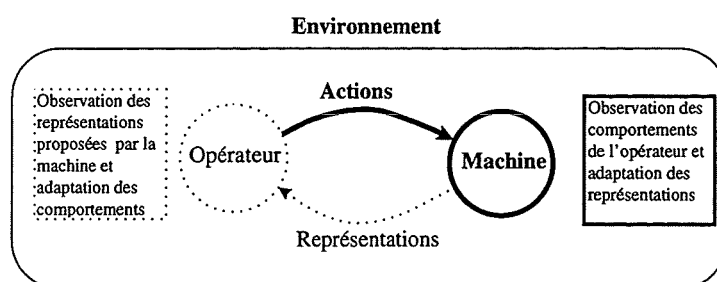


FIG. 2 – Limitation de l'étude dans le cadre de la démarche générale

Dans une première partie, nous présenterons les lignes directrices de notre méthode d'analyse de comportements. Puis nous présenterons le système expérimental que nous avons choisi, ainsi que les premiers résultats obtenus. Enfin, une dernière partie exposera les perspectives de l'étude en rapport avec la démarche générale de coopération homme-machine décrite ci-dessus.

2 Analyse des comportements des opérateurs

Dans cette communication, l'identification des opérateurs en phase d'interaction avec la machine consiste d'une part, à détecter des régularités chez un même opérateur soumis séquentiellement à différentes phases d'interaction et d'autre part, à catégoriser différents opérateurs d'un même système. La méthode développée ici repose sur deux hypothèses fortes :

- l'opérateur établit et consolide des procédures personnalisées qui, même si elles se placent au niveau sensori-moteur, sont significatives des schémas d'actions et de la sémantique du contrôle du processus ([10]);
- l'observation des conduites de l'opérateur au niveau sensori-moteur et de leur évolution au cours de temps, permettent la caractérisation de ses schémas d'actions.

Ainsi, nous restreignons l'analyse de la dynamique d'interaction opérateur-machine à l'étude de la trace du comportement sensori-moteur de l'opérateur en action et n'utilisons pas de données de verbalisation issues d'entretiens in situ comme il est classique de le faire dans la tradition issue des systèmes experts.

2.1 Espace de représentation de la trace comportementale d'un opérateur

Soit E l'espace des états du système dynamique (opérateur, machine). Une phase d'interaction entre l'opérateur et la machine peut être modélisée par une trajectoire dans E . En pratique, la détermination de E est un problème très délicat. Elle dépend évidemment des outils d'observations de l'analyste. De plus, la prise en compte d'une précision maximale – au vu de ces outils – peut engendrer une complexité inextricable pour les algorithmes de traitement de données dont on dispose actuellement. C'est pourquoi la méthode proposée ici consiste à partitionner l'espace d'états en un nombre relativement restreint de sous-espaces codés par un symbole. La dynamique de l'opérateur est alors représentée par une séquence de symboles associés aux sous-espaces atteints. C'est ce que l'on appelle la dynamique symbolique du système ([14]). Le choix du partitionnement est également critique. Dans cette communication il est donné a priori par un spécialiste du système étudié. Les travaux menés en mécanique computationnelle ([9], [7], [15]) suggèrent toutefois une méthode plus systématique qui examinerait des niveaux de granularité successifs pour le partitionnement de l'espace E .

2.2 Catégorisation / identification de l'opérateur à partir des séquences symboliques

Suivant l'hypothèse cognitive selon laquelle tout opérateur s'appuie sur des schémas d'actions englobant le contexte [10], il semble naturel de vouloir identifier les opérateurs par des schémas d'actions types. Afin d'évaluer quantitativement la pertinence de cette hypothèse, nous définissons une mesure de proximité entre des séquences symboliques décrivant différentes phases d'actions d'un opérateur puis, nous recherchons des sous-séquences communes à cet ensemble de séquences. La représentation de la mesure de proximité dans un espace mathématique ad hoc (ici un X-arbre, [8]) nous permet ensuite d'identifier des classes d'opérateurs (e.g. novice, expert).

2.3 Modélisation prédictive des comportements de l'opérateur

Après cette première étape de catégorisation, l'analyse peut se poursuivre par l'extraction de la "signature" de certains comportements types de l'opérateur en vue d'une modélisation prédictive devant permettre à la machine une meilleure adaption de sa représentation à l'opérateur. Pour ce faire, nous envisageons l'utilisation de modèles de Markov cachés sur les séquences de symboles d'une manière similaire à celle utilisée en reconnaissance automatique de parole ([5]). Les résultats de ces recherches en cours seront présentées ultérieurement ([6]).

3 Application de la méthode à un jeu video

Le système dynamique choisi pour expérimenter la méthode décrite ci-dessus est un jeu vidéo populaire, XGalaga. Le principe du jeu est simple : un joueur peut, à l'aide d'une souris, déplacer latéralement un vaisseau qui se trouve au bas de l'écran et doit détruire des vaisseaux ennemis qui arrivent du haut de l'écran en se dirigeant vers lui selon des trajectoires pseudo-aléatoires (figure 3). Le but du jeu consiste pour le joueur à rester "en vie" le plus longtemps possible et à détruire le plus de vaisseaux ennemis possibles.

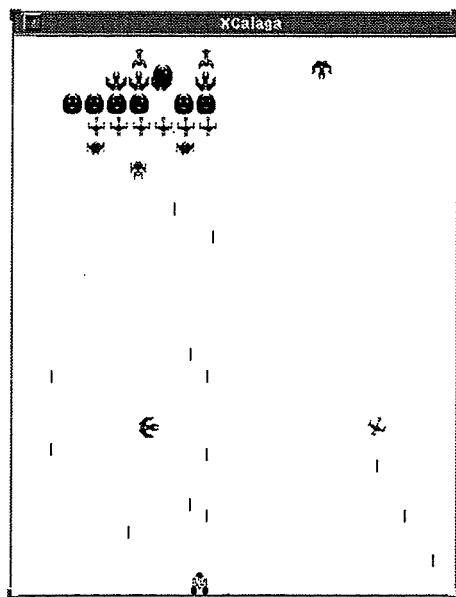


FIG. 3 – Ecran de jeu Xgalaga

Le code du programme de ce jeu étant accessible, nous avons pu en déterminer tous les paramètres de façon à permettre des comparaisons de parties dans des conditions expérimentales strictement identiques.

3.1 Espace de représentation

Le comportement spatio-temporel du système “Xgalaga-joueur” correspond à une trajectoire dans l’espace d’état E associé. Une représentation symbolique de cette dynamique s’obtient, par exemple, en décrivant par un vecteur binaire, à chaque instant t , la présence/absence de missiles défini en chaque zone de l’écran et les commandes effectuées par le joueur sur la souris (mouvements gauche/droite et tir/non tir). Cependant, une étude partielle préalable montre que l’information n’est pas traitée à ce niveau de précision par les joueurs ([4]). Nous nous contentons de considérer ici un partitionnement de notre espace qui a été défini en dialoguant avec des experts du système. L’écran de jeu est découpé verticalement en cinq zones en fonction de la position du joueur à chaque instant : le joueur étant très attentif aux vaisseaux qui sont dans son voisinage immédiat, la zone qui est au dessus de lui est plus fine que celles qui sont plus éloignées (figure 4). Pour chacune des zones, l’information retenue est la présence ou l’absence de missile. Le comportement du joueur est décrit par ses mouvements relatifs par rapport à sa position à l’instant précédent : se déplacer à gauche, à droite ou rester statique. Le joueur est supposé tirer de façon continue. L’état du système en chaque instant t est donc un symbole : $x_t \in \{0, 1\}^5 \times \{gauche, droite, statique\}$. La longueur de la séquence x_1, \dots, x_N décrivant la dynamique dépend de la durée de vie du joueur : N est l’instant auquel le joueur est atteint par un missile ennemi.

3.2 Mesure de dissemblance et comparaison des comportements

Afin de comparer les comportements de différents joueurs et/ou de mettre en évidence des caractéristiques de stabilité (versus instabilité) comportementale chez un

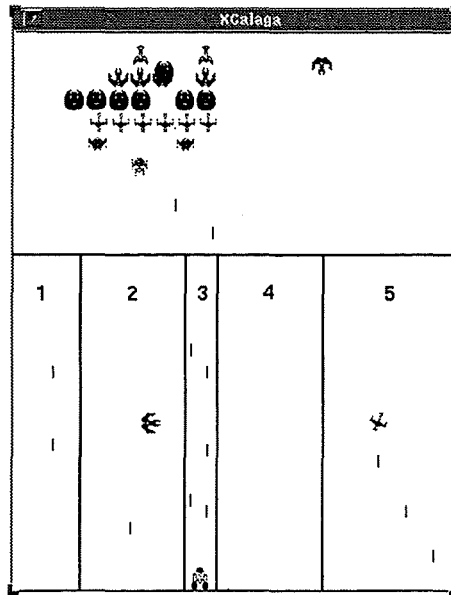


FIG. 4 – Découpage de l'écran de jeu en zones pertinentes. Seules les zones de la partie inférieure de l'écran sont prises en compte. Cette situation correspond au symbole $x_n = (1, 1, 1, 0, 1, \text{droite})$

même individu, nous utilisons un indice de dissemblance d qui est calculé entre des paires de parties décrites par des séquences de symboles de $\{0, 1\}^5 \times \{\text{gauche}, \text{droite}, \text{statique}\}$. La distance d'édition est un indice classique lors de la comparaison de séquences ([11]). Il se calcule par programmation dynamique en se basant sur les transformations les "plus économiques" qu'il faut réaliser pour passer d'une séquence à l'autre. Les transformations admissibles sont la suppression d'un symbole, l'insertion d'un symbole virtuel "blanc" et la substitution d'un symbole par un autre. Chaque transformation est associée à un coût et, la distance d'édition donne la succession de transformations de coût total minimum qui permet de passer d'une séquence à l'autre. Pour le système étudié ici, les coûts sont les suivants :

- les coûts d'insertion et de suppression sont égaux à 1;
- deux positions contigües du joueur sont considérées comme équivalentes; deux positions éloignées de deux unités – sur une échelle de quarante – ont un coût de 1; sinon le coût est infini;
- les missiles éloignés du joueur ayant peu d'influence sur la stratégie de ce dernier, le coût associé au changement de valeur binaire (présence/absence de missiles) d'une zone est une fonction de l'éloignement de la position du joueur à cette zone.

L'algorithme qui calcule la distance d'édition permet également de chercher une sous-séquence commune maximale entre deux séquences, i.e. la plus grande suite de symboles que l'on peut obtenir à partir de chaque séquence en supprimant certains d'entre eux. L'extraction de sous-séquences communes permet de mettre en évidence des schémas d'actions chez un même individu ou des comportements communs caractéristiques d'un groupe de joueurs (figure 5). En abscisse sont représentés uniquement les mouvements du joueur (sixième coordonnée de x_t) et en ordonnée, le temps t . La figure 5.a correspond à un joueur expert, la figure 5.b à un joueur

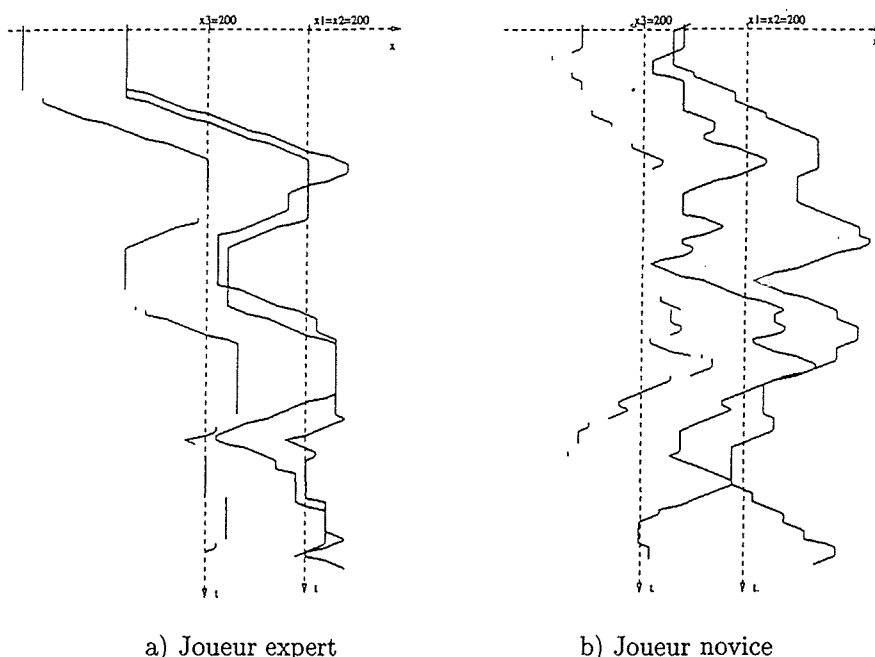


FIG. 5 – Sous-séquences communes (courbes décalées à gauche) entre deux séquences associées à des joueurs novices et des joueurs experts (courbes de droites).

novice. Pour chacune, sont tracées, décalées à gauche pour plus de lisibilité, les sous-séquences communes à deux séquences du même joueur présentées à droite. On peut remarquer que le joueur expert engendre plus de sous-séquences communes que le joueur novice.

3.3 Une représentation arborée de l'indice de dissemblance

Afin d'identifier chaque joueur, nous calculons l'indice de dissemblance sur un nombre fini de parties (10 dans notre expérience) et nous recourons à une représentation arborée pour mettre en évidence les informations pertinentes de ce tableau. Les arbres additifs (ou X-arbres [8]), support d'une telle analyse, représentent des entités que l'on cherche à comparer comme les feuilles d'un arbre, de façon à ce que la distance entre deux feuilles corresponde à la similarité existant entre les entités correspondantes. Cette représentation peut de plus avoir un rôle explicatif. Elle fait en effet apparaître, lors de la transformation du tableau de dissimilarités en une structure d'arbre, des sommets internes à l'arbre (encore appelés *sommets latents*) que l'on peut interpréter comme :

- soit des objets «intermédiaires» non recensés mais expliquant le passage d'une feuille à l'autre (il s'agit alors de «chaînon manquant» notamment dans le cas de construction d'arbres additifs relatifs à l'évolution d'espèces ou de filiation de manuscrits ([1]))
- soit des objets prototypiques représentant la «catégorie» des objets feuilles qui lui sont rattachés.

Ce dernier point de vue a été notamment utilisé dans différentes études de psychologie cognitive portant sur la mémoire sémantique. Barthélemy, Abdi et Luong ont ainsi montré ([2], [8]) que les arbres additifs étaient adaptés à la représentation des concepts en mémoire épisodique sur la base des expériences de Bartlett ([3]), portant sur l'évaluation de distances entre textes mémorisés.

Notre approche consiste à étendre ces méthodes, qui ont fait leurs preuves pour la structuration de concepts en mémoire, à celle des schémas d'action (cf. [10]) des opérateurs au sein d'un système. Nous pensons que, de la même façon qu'une représentation arborée appliquée aux concepts peut mettre en exergue des sommets latents interprétables en termes de catégories, une représentation arborée des comportements d'un opérateur en charge du contrôle d'un processus dynamique pourra

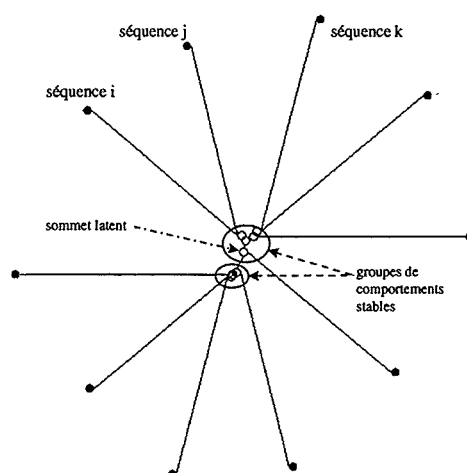


FIG. 6 – Représentation arborée des comportements d'un expert

mettre en évidence des catégories de comportements stables et prototypiques. Pour cela, nous souhaitons pouvoir représenter visuellement les relations entre parties sous-jacentes dans la matrice de dissimilarité établie lors de la mesure de dissimilance. On plonge la dissimilarité d sur un arbre de sorte que les sommets de l'arbre représentent les parties et la distance D du plus court chemin sur l'arbre entre ces sommets soit "la plus proche possible" de d . Pour effectuer cette approximation, nous avons utilisé l'heuristique des plus proches prédécesseurs ([8]).

Les premiers résultats montrent que l'arbre représentant les comportements d'un joueur expert met en évidence un petit nombre de comportements prototypiques, et fortement regroupés. La figure 6 permet de représenter l'ensemble des séquences enregistrées comme les feuilles de l'arbre, et de définir deux groupes principaux de sommets latents. Ces groupes peuvent être interprétés comme les deux grands types de comportement que l'expert met en œuvre face aux contextes présentés lors des expériences.

A l'inverse, les comportements du joueur novice aboutissent à une représentation arborée possédant un nombre de sommets latents beaucoup plus important et distants les uns des autres : les comportements de référence semblent en quelque sorte "répartis", et trop nombreux pour être significatifs et prototypiques (en vue d'un "ancrage" comportemental). La figure 7 montre en effet des groupements des sommets latents de l'arbre beaucoup moins compacts et moins interprétables que dans le cas de l'expert.

4 Perspectives

Nous envisageons de développer nos travaux selon plusieurs axes. Un premier axe consiste à poursuivre l'étude du système expérimental Xgalaga :

- en raffinant l'espace de représentation. Les résultats présentés dans cette communication semblent démontrer que la méthode permet de distinguer les opérateurs experts des opérateurs novices, mais ne permettent pas encore d'isoler clairement des stratégies expertes intelligibles. Un travail plus en profondeur sur l'espace de représentation devrait aboutir à des résultats plus probants sur ce point.
- en développant une modélisation prédictive des comportements de l'opérateur. Une approche statistique à l'aide de chaînes de Markov cachées est actuelle-

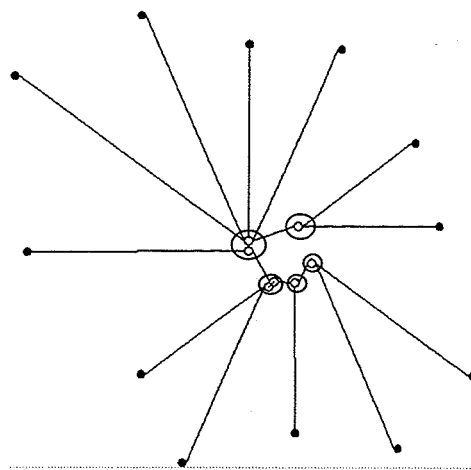


FIG. 7 – Représentation arborée des comportements d'un novice

ment à l'étude et sera présentée prochainement ([6]). Elle devrait permettre une représentation des schémas d'actions opérationnels indispensable à la mise en œuvre du retour d'information de la machine vers l'opérateur telle qu'annoncée dans l'introduction (cf. figure 1).

- en mettant en place le retour d'information vers l'opérateur nécessaire à une véritable coopération homme-machine. Ce retour d'information peut prendre la forme d'une aide en ligne (la machine peut signaler l'apparition d'écarts par rapport aux stratégies expertes usuelles). Mais il peut également s'agir d'adaptation des stratégies du jeu (c'est-à-dire de la machine) à celles du joueur expert, afin de rendre plus attrayante la poursuite du jeu. Ce développement permettrait d'aboutir à une véritable analyse de la coopération homme-machine telle que définie en introduction, ce qui reste l'objectif premier de l'étude.

Un second axe consiste à appliquer la méthode à d'autres types de processus dynamiques, comme notamment le parcours d'opérateurs au sein d'un système d'information, ou dans le cadre de contrôle de processus industriels.

Remerciements :

Nous remercions vivement Xavier Simon et Jean Bertrand pour les développements informatiques réalisés dans le cadre de ce projet.

Références

- [1] H. Abdi. *Additive-tree representations*, volume 84 of *Lectures Notes in Biomathematics*. Springer Verlag, 1990.
- [2] H. Abdi, J.P. Barthélemy, et X. Luong. Tree representation of associative structures in semantic and episodic memory research. In E. Degreef et J. Van Buggenhaut, editors, *Trends in Mathematical Psychology*, pages 3–31. North Holland, Amsterdam, 1984.
- [3] F.C. Bartlett. *Remembering*. Cambridge University Press, Cambridge, 1932.

- [4] J. Bertrand et X. Simon. Etude du comportement d'un joueur. Rapport de stage, Ecole Nationale Supérieure des Télécommunications de Bretagne, Brest, 1996.
- [5] H. Bourlard, Y. Kamp, et C. Wellekens. Speaker dépendent connected speech recognition via dynamic programming and statistical methods. In M. Schroeder, editor, *Speech and Speaker Recognition*, pages 115–148. Karger, Bâle, 1985.
- [6] G. Coppin, E. Faure, and P. Kuntz. Sequences of actions in dynamical environment and typology of behaviours. In *Decision Support Systems*, Bruges, 1997. Association of European Operational Research Societies, à paraître.
- [7] J. P. Crutchfield et K. Young. Computation at the onset of chaos. In W. Zurek, editeur, *Entropy, Complexity, and the Physics of Information*, volume VIII of *SFI Studies in the Sciences of Complexity*, pages 223–269. Addison-Wesley, Reading, Massachusetts, 1990.
- [8] J.P. Barthélemy et A. Guénoche. *Les Arbres et les représentations des proximités*. Masson, Paris, 1988.
- [9] H. Packard, P. Crutchfield, J.D. Farmer, et R.S. Shaw. Geometry from a time series. *Phys. Rev. Letters*, 45(742), 1980.
- [10] J.F. Richard. *Les activités mentales. Comprendre, raisonner, trouver des solutions*. Armand Colin, Paris, 1990.
- [11] D. Sankoff. Matching sequences under deletion-insertion constraints. In *Proceedings of the National Academy of Sciences of the USA*, volume 69, pages 4–6, 1972.
- [12] Y.M. Visetti. Des systèmes experts aux systèmes à base de connaissances : à la recherche d'un nouveau schéma régulateur. In *Expertise et sciences cognitives*, volume 12, *Intellectica*, pages 221–279. 1991.
- [13] A. Weill-Fassina, P. Rabardel, et D. Dubois. *Représentations pour l'action*. Octares, Paris, 1993.
- [14] S. Wiggins. *Introduction to Applied Nonlinear Dynamical Systems and Chaos*. Springer Verlag, Berlin, 1990.
- [15] K. Young. *The grammar and statistical mechanics of complex physical systems*. PhD thesis, Univ. of California, Santa Cruz, 1991.

Construction et interaction de modalités perceptives Définition d'un cadre théorique et propositions pour un atelier expérimental

Sabine Ploux
laboratoire ELSAP URA 1234
Université de Caen
14 032 Caen cedex
e-mail ploux@info.unicaen.fr

1 Présentation

Nous cherchons, dans cet article, à poser le cadre d'une modélisation du rapport de détermination qui doit exister entre un système et le type de cognition qu'il est susceptible de déployer. Plus spécifiquement, nous nous fixerons comme objet, le résultat cognitif qui peut se dégager d'une interaction constructive entre des schèmes sensori-moteurs décrits en fonction de l'« anatomie » et de la « physiologie » du système choisi. Dans cette perspective, l'environnement et sa représentation ne préexistent pas à l'élaboration cognitive. Celui-ci est initialement donné par l'activité qu'il imprime sur les zones sensibles, et son émergence est ici ramenée (paramétrée) par le système et son évolution. D'autres travaux font de l'activité sensori-motrice un principe de construction cognitive. Nous ne pouvons être exhaustif et nous avons choisi de mentionner au cours de cet exposé certains points relatifs à l'approche sélectionniste défendue entre autres par G. Edelman [REEK, SPORNS, EDELMAN 1990].

Pour présenter nos choix en matière de construction cognitive, nous nous placerons suivant plusieurs points de vue.

D'abord seront énoncées les motivations suscitées par notre propre parcours scientifique. Elles nous ont poussée à trouver une alternative aux méthodes qui, utilisant une représentation logique ou mathématique, symbolique et discrète, deviennent inefficaces pour résoudre certains problèmes posés en I.A.

Puis vient une étude comportementale de la cognition suivie de la caractérisation du système capable de supporter un apprentissage constructif. Pour ce qui est du comportement, nous expliquerons comment le point de vue piagetien¹, constructif et explicatif, a conforté et influencé notre recherche.

Il faut pour décrire le système et son évolution, un cadre qui soit robuste aux spécificités modales et qui puisse s'adapter à la variation de la complexité

¹Nous avons trouvé récemment chez E. Thelen [THEL 92] (et je remercie ici Y. M. Visetti de m'en avoir informée) un modèle qui associe à une démarche piagetienne (pour ce qui est de l'intelligence sensori-motrice), une description dynamique

initiale du système. Nous avons choisi de représenter l'espace cognitif par un système dynamique dont le contrôle sera initialement déterminé par l'espace comportemental. Ceci constitue notre perspective mathématique.

Toutes ces approches reflètent, parce que son unité est encore en perspective, la nature proprement interdisciplinaire de l'objet d'étude. Elles permettront, nous l'espérons, de mieux suivre l'élaboration de l'atelier expérimental que nous développons actuellement.

2 L'objet et son espace de contrôle

En prenant pour exemples des thèmes sur lesquels nous avons travaillé, nous voulons ici rappeler que la représentation qui, à un objet associe un symbole unique sans dimension intrinsèque, sans variation possible, n'a pas permis (sinon dans des formes extrêmement figées) la reproduction par les machines actuelles de fonctions cognitives comme le sont la démonstration mathématique, la reconnaissance de formes graphiques, la représentation sémantique de la synonymie.

- En démonstration automatique, la seule donnée des axiomes et des règles de déduction est insuffisante à la production d'un bon nombre de démonstrations, même simples. Les systèmes se noient dans une complexité d'objets et de propositions dénués de toute pertinence pour l'objectif à atteindre.

Contrairement au fonctionnement de ces systèmes, la découverte d'une démonstration est souvent accompagnée d'une représentation graphique (effective ou intériorisée) comme le sont les diagrammes de Venn, le tracé de fonctions, etc... Ces représentations, pour certains aspects, synthétisent mieux la situation.

Nous avons mis en œuvre un programme [PLOU 92] qui, utilisant un équivalent de ces représentations, permet de guider une démonstration en évaluant la pertinence des objets et des propositions construits.

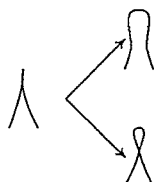
En somme, nous défendons l'idée que des représentations perceptives et comportementales pour lesquelles le support est continu, permettent de reconnaître des rapports constitutifs des objets entre eux de façon plus immédiate que la seule utilisation des formes logiques. Dans le système réalisé, l'équivalent des espaces, ici géométriques, participent au contrôle de l'objet dans l'évaluation des formes logiques qui le représentent.

- Prenons comme deuxième exemple la reconnaissance des formes et plus spécifiquement la reconnaissance de l'écriture manuscrite sur laquelle nous avons travaillé.

Certains systèmes [WAGN 74, WANG 91] fonctionnent en utilisant la cooccurrence d'indices comme les croisements, les boucles, les points de rebroussement et des règles de production que l'on peut en tirer (telle configuration d'indices correspond à telle lettre). Mais une même lettre peut être représentée par différentes configurations (par exemple, un *u* peut contenir 0, 1 ou 2 boucles).



Ces descriptions de l'objet, ici une lettre, n'expliquent pas quelles sont les déformations possibles (celles qui conservent à l'objet son identité) et quelles sont celles qui ne le sont pas. Elles ne permettent pas de comprendre l'invariance de la désignation de l'objet à travers ses différentes réalisations et comment l'on passe d'une configuration à une autre : par exemple, comment un point de rebroussement peut devenir une boucle ou un arc.



Nous avons proposé un modèle mathématique différentiel pour représenter la topologie du geste qui produit une écriture. Il permet de rendre compte de la morphogenèse des formes graphiques à partir de déformations de points singuliers comme l'est un point de rebroussement. Dans notre modèle, ces déformations sont contrôlées par la variation des représentations des vitesses gestuelles. P. Viviani et N. Stucchi [VIVI 92] ont proposé des lois de reconnaissance à partir d'observations. Ces lois mettent en évidence des transferts intermodaux entre des données visuelles (le rayon de courbure d'un tracé) et des données gestuelles (la vitesse du geste qui a produit ce tracé). Ces travaux, conformes à notre point de vue, nous poussent, en généralisant le travail déjà effectué, à proposer un modèle du contrôle de l'objet et de sa reconnaissance, par une construction interactive entre différents espaces modaux.

- Notre dernier exemple portera sur la sémantique linguistique. Nous avons pu constater à l'occasion d'un projet visant à organiser l'espace synonymique d'un mot [PLOU 96] (mais bien d'autres l'avaient déjà écrit [FUCHS, VICTO 94]) que des méthodes qui font d'un terme linguistique donné, une unité insécable, un *point mathématique*, échouent à la construction d'un espace qui articulerait entre eux les sens synonymes. Cette structure devrait, à l'image d'un article de dictionnaire, regrouper des termes proches par le sens et séparer ceux qui ne le sont pas. Cet échec tient à la variabilité sémantique que tout terme porte en lui : on ne peut modéliser le déploiement du sens d'un mot sans se donner les outils théoriques qui le permettent.

Ces trois exemples nous ont conduite à penser la catégorisation, la reconnaissance et les autres fonctions cognitives comme des processus qui se fonderaient, non pas seulement sur un calcul symbolique, mais aussi sur l'utilisation d'un contrôle cognitif, issu de toutes les formes pratiques du comportement (et dont la langue fait partie). C'est cet espace de contrôle, que nous poursuivons. Pour commencer, nous nous attacherons à décrire ses aspects sensori-moteurs.

3 Un cadre descriptif pour les premières étapes cognitives

Nous avons trouvé chez Piaget une description constructive qui relie l'expérience du sujet à l'évolution de ses fonctions cognitives. D'autres points de vue psychologiques [STRER, SPELK 1989] mettent en évidence des transferts inter-modaux qui seraient premiers à une expérience des interactions modales. Ces recherches font l'hypothèse d'un traitement central et amodal de la perception², mais n'offrent pas à l'heure actuelle et autant que nous en ayons connaissance de modèle véritablement explicatif.

Nous avons ici fait le choix scientifique et donc réducteur (pour ce qui est de la psychologie du vivant) de n'étudier que la part interactive de la construction cognitive.

Nous proposerons, d'abord une description des constituants anatomiques, ensuite des principes pour l'évolution cognitive des rapports entre ces constituants. Cet exposé sera justifié par des exemples tirés des deux premiers chapitres de « *la naissance de l'intelligence chez l'enfant* ». (Nous ne voulons bien sûr ni réécrire, ni résumer, ni paraphraser Piaget ; nous voulons mettre en évidence ce qui, dans son analyse, peut éclairer une démarche artificielle.)

La notion de réflexe sera notre point de départ, nous poserons ensuite les lois de sa transformation ; ces mêmes lois serviront ensuite à construire l'interaction entre des réflexes déjà adaptés.

3.1 Le réflexe et son évolution

Voici posée une définition pour un système artificiel du réflexe : le réflexe est un comportement cyclique d'un système donné. Il est composé :

- d'une surface sensitive appelée *rétilne* elle même centrée sur une *fovea*.
- d'une unité motrice dont l'intensité d'activation varie suivant un paramètre linéaire
- d'une horloge interne.

Sa fonction est de réguler le système.

Au moment initial les caractéristiques du comportement réflexe sont :

- l'activation de l'unité motrice par excitation de la zone foveale.
- l'activation à vide de l'unité motrice par simple appel de l'horloge.

Les réflexes innés sont des comportements complexes du point de vue de la coordination musculaire impliquée, mais ils sont pris dans l'œuvre de Piaget comme des globalités qui « préparent[...]l'individu à s'adapter au milieu externe et à acquérir les conduites ultérieures, caractérisées précisément par l'utilisation progressive de l'expérience ». L'étude vise alors à comprendre leur évolution, leur interaction, leur capacité à s'organiser en de nouvelles globalités sensori-motrices, de nouvelles Gestalt.

Voici reprises quelques caractéristiques des réflexes de l'enfant :

²ce qui pourrait être s'expliquer par une étude embryologique de l'histoire commune des zones perceptives

Les réflexes peuvent fonctionner à vide (exemple : la succion à vide) ; ils sont déclenchés par une stimulation de la zone réflexe (une main qui frôle les lèvres déclenche la succion, une pression légère sur la paume de la main déclenche la fermeture de la main).

La zone réflexe sensitive est au début très localisée, le comportement moteur est isolé.

3.2 Évolution du comportement réflexe par interaction avec l'environnement

Nous proposons maintenant un schéma pour l'évolution comportementale d'un système artificiel.

L'évolution du réflexe est déterminée par la recherche de sa réalisation. Cette recherche se réalise par un appel progressif :

- aux unités motrices voisines de l'unité réflexe initiale en suivant l'anatomie du système ;
- aux unités sensibles voisines de la fovea.

de telle sorte qu'à chaque excitation de ces zones sensibles étendues sera associé un comportement qui rapprochera le système de la réalisation fonctionnelle attachée au réflexe.

En somme le réflexe va élargir ses zones sensibles et motrices graduellement et le rayon de cette expansion est donné par une topologie somatopique.

À l'exercice réflexe tel qu'il apparaît à la naissance succède l'élaboration d'une recherche réflexe. Il suffit par exemple, à l'enfant d'avoir heurté des lèvres le mamelon pour tâtonner la bouche ouverte jusqu'à réussite. Les mouvements de la tête sont ainsi impliqués dans la réalisation du réflexe. La zone sensitive s'élargit. Un contact avec la joue entraîne une recherche. Au début cette recherche est aléatoire. L'enfant tourne la tête du mauvais ou du bon côté. Puis elle devient efficace et l'enfant se tourne du bon côté. Il y a donc, par utilisation de mouvements extérieurs au réflexe initial, sélection d'un bon comportement moteur et corrélation entre une zone sensitive voisine et des unités motrices. Le réflexe élargit et construit une dynamique sensori-motrice qui à une excitation sensitive fait correspondre un déplacement moteur.

Pour le réflexe de succion sont successivement impliqués les mouvements latéraux de la tête puis le lever de la tête quand le seul déplacement latéral n'a pas suffi.

3.3 Interaction entre réflexes

On supposera ici que le système a développé indépendamment plusieurs réflexes. Voici énoncés leurs modes d'interaction.

L'interaction est rendue possible par un recouvrement synchrone de deux réflexes : l'excitation d'une zone sensitive d'un réflexe R1 par une zone sensitive d'un réflexe R2.

Cette synchronisation réflexe constitue une nouvelle unité, un nouveau schéma qui pareillement va étendre ses zones sensibles et motrices.

Les lois d'extension sont celles précédemment données ; elles suivent la topologie anatomique.

Les horloges biologiques déterminent les priorités entre les différentes régulations.

Piaget décrit ce type d'interaction dans le second chapitre de « la naissance de l'intelligence chez l'enfant ». Chez le nourrisson, elles aboutissent à des comportements qui peuvent se décrire ainsi : saisir pour porter à la bouche, saisir pour regarder, regarder pour saisir...

Il est difficile (tant est déjà complexe et divers l'ensemble des dynamiques qui se sont précédemment constituées), d'utiliser la description des événements qui se déroulent à ce stade pour suivre pas à pas, en utilisant les principes énoncés, la construction cognitive. La vision, par exemple, fait intervenir de nombreux réflexes pupillaire, palpébral, d'accommodation... une sensibilité à l'éclairement, à la variation lumineuse.

Nous tenterons donc d'annoter quelques étapes constitutives de la construction de nouveaux schèmes psychologiques. Ces observations serviront de pierre de touche pour prévoir et tester l'évolution d'un système artificiel simplifié.

Au début de façon imprévisible et ensuite entretenues se produisent des stimulations mutuelles des zones sensibles liées à des réflexes initialement indépendants : mettre les doigts dans la bouche, saisir le nez, les joues, les mains...

[Ces contacts seront ici interprétés comme des synchronisations perceptives entre des zones sensibles liées, d'une part à la succion (pour le visage) et à la préhension (pour la main saisie), d'autre part à la préhension pour la paume de la main qui saisit.]

Cette synchronisation perceptive va permettre la constitution d'un comportement cyclique autonome :

- * sucer son poing (puis/ou) son pouce. [Le mouvement du bras, voisin pour l'anatomie corporelle des mouvements de la tête eux mêmes déjà impliqués dans la succion, est à son tour utilisé pour la réalisation du nouveau schème.]*
- * joindre les deux mains. [Ici le mouvement des épaules est impliqué dans le nouveau schème construit à partir du réflexe d'agrippement]*
- * Puis il arrivera que l'enfant tenant un objet, exécute le schème de succion du poing ce qui l'amènera à sucer un objet tenu [nouveau schème (sucer un objet tenu) qui est un produit de la synchronisation de la succion et de la préhension. Ce nouveau schème va devenir autonome et la zone sensible de l'agrippement (la paume) va être impliquée dans sa réalisation par association au déplacement du bras.]*
- * Un autre nouveau schème va alors se produire : regarder un objet saisi. Voyons un événement qui permet cette création et quelques schèmes qui seront impliqués dans la recherche de sa réalisation. Le regard se porte sur l'objet qui dans sa trajectoire est amené à la bouche ou qui en sort. [Nouvelle synchronisation qui cette fois implique la vision]. Ce nouveau schème va se réaliser par appel à des formes variées du comportement. Les unes par exemple impliquent dans un premier temps le schème précédemment constitué (sucer un objet saisi). D'autres utiliseront le regard : la main et l'objet côte à côte dans le champ visuel, seront regardés alternativement réalisant ainsi les schèmes du regard d'un objet ou du regard de la main. Les informations proprioceptives du regard vont servir de zones sensibles étendues pour effectuer le nouveau schème par association aux petits mouvements du bras qui va attraper l'objet.*

L'enfant apprendra ensuite, par extension topologique, à saisir un objet de plus en plus éloigné de la main.

** En dernier lieu les deux types de recherches vont se combiner pour la réalisation du schème quand la main n'appartient pas au champ visuel.*

Nous le répétons, ceci n'est qu'une ébauche, qu'un argument pour motiver le développement d'une recherche approfondie des principes que nous allons maintenant énoncer et qui nous ont paru (bien que la complexité des interactions du vivant avec son environnement ne permette pas d'énumérer manuellement et de façon exhaustive l'application de ces principes) conformes à l'idée piagetienne de la construction cognitive. Par ailleurs, nous attendons des simulations sur machines, une aide, un outil pour tester la validité de l'évolution des systèmes dynamiques (que nous ne pouvons résoudre *à la main*) obéissant aux règles choisies.

Nous reprendrons donc pour les généraliser les idées que nous avons commencé à dégager pour les réflexes

3.4 Principe d'évolution du système et de ses fonctions

On suppose que le système est doté d'un ensemble de régulations fonctionnelles qui cherchent à se réaliser périodiquement. La réalisation de ces fonctions a une facette énergétique. Elle permet au système de « se nourrir ». Si le déclenchement du schème fonctionnel ne suffit pas à alimenter le système, il y a mise en œuvre d'une recherche. Cette recherche est d'une part un appel progressif aux autres schèmes d'autre part une utilisation elle aussi progressive des autres zones sensibles.

Nous proposons une complexification qui suit la topologie corporelle. Le réflexe va élargir ses zones sensibles et motrices graduellement et le rayon de cette expansion est donné par une topologie somatopique.

Les schèmes voisins, les surfaces sensibles voisines, investissent le schème initial et le réorganisent.

Nous allons maintenant proposer une anatomie et une physiologie du système susceptible de permettre la réalisation de ce principe.

4 Caractérisation du système

Nous proposerons d'abord une anatomie support des réflexes initiaux et de leur fonctionnement ensuite les principes d'apprentissage qui permettront la construction des connexions nouvelles entre des zones distinctes.

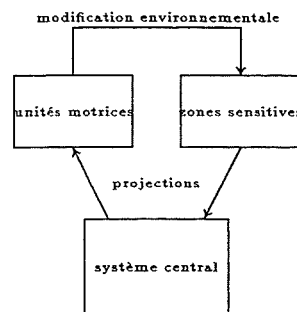
4.1 L'anatomie

- Une anatomie du système comportera :
 - les zones sensibles
ce sont des zones planaires, on les supposera dans une simulation artificielle constituées d'un pavage d'unités répondant à une intensité qui sera ici représentée par une valeur mathématique.

- les unités motrices
ce sont des unités répondant à un paramètre linéaire et dont l'activité provoquent une modification de l'activation des zones sensibles.
- un système central

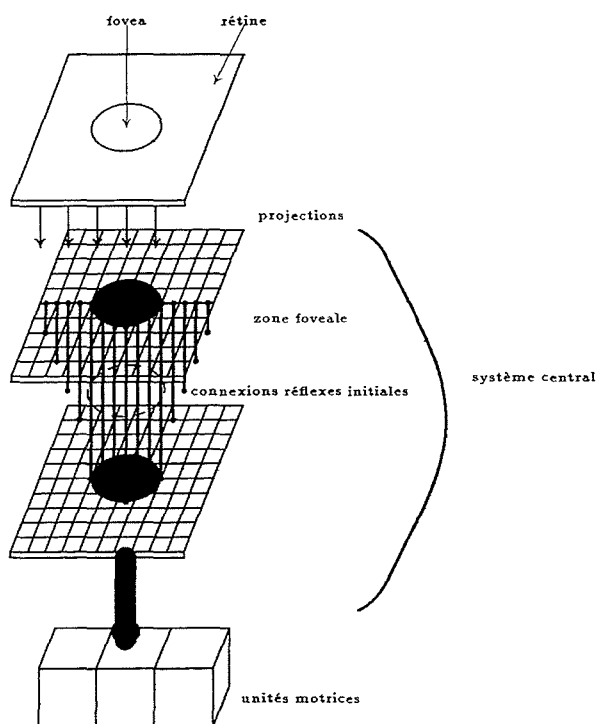
le système central est constitué :

- des couches « neuronales ». Chaque unité de ces couches est munie d'une fonction de transfert qui à une combinaison des intensités des liaisons entrantes associera une valeur qui sera transmise aux liaisons sortantes.
- des liaisons définies par leur direction, le point initial et le rayon du champ récepteur autour de ce point, le point final et le rayon du champ effecteur autour de ce point.



- À l'état initial sont constituées :

- les projections centrales
Toutes les couches sensibles se projettent sur des couches centrales et la topologie de cette projection définit les voisinages des zones sensibles entre elles.
À toutes les unités motrices sont associées au niveau central des couches qui se projettent sur ces unités motrices. La topologie de cette projection définit le voisinage des zones motrices.
- les connexions réflexes
Les réflexes sont matérialisés par des liaisons constituées entre les zones centrales associées aux foveas et les zones centrales qui se projettent sur les unités motrices du réflexe.
- les horloges qui codent une fonction périodique d'appel au déclenchement d'un schème.



4.2 Les lois d'apprentissage

Pour répondre aux principes topologiques énoncés dans la description comportementale on pose :

- règle topologique

L'expansion des liaisons entre deux couches suit la topologie définie par les cartes sensibles et motrices. La réalisation d'un réflexe attendu par l'horloge apporte au voisinage des jonctions, qui ont permis cette réalisation, une énergie (valeur mathématique) qui est traduite par une progression des liaisons voisines aux liaisons efficaces. La direction de cette progression est parallèle à la direction des liaisons efficaces. Les liaisons qui touchent une nouvelle unité créent ainsi une connexion.

- règle de synchronie

Une jonction efficace est une jonction qui a été activée pendant le cycle précédant de l'horloge interne.

L'horloge interne définit la synchronie de coactivation.

La règle topologique est une alternative aux principes de sélection proposés par exemple par G. Edelman. Le point de vue sélectionniste impose une surproduction initiale de connexions entre « neurones », connexions qui seront ensuite triées en fonction d'une interaction avec l'environnement et des satisfactions que cette interaction a pu apporter au système. L'extension topologique évite cette complexité initiale. Nous avons trouvé dans [QUART, SEJNO 1996] une critique du point de vue sélectionniste et un

analogue des règles ici énoncées. Ces auteurs proposent un modèle neuro-physiologique de la croissance progressive des connexions neuronales. Cette croissance régulée par l'interaction que l'entité entretient avec l'environnement permettrait ainsi la croissance progressive des capacités cognitives du cortex. La règle de coactivation que nous proposons est cependant différente de celle qu'ils formulent puisqu'ici l'horloge est une unité de satisfaction globale du système et non plus une coactivation neuronale fondée sur un temps physique externe au système.

Ceci permet de construire des liaisons entre des zones sensibles et des zones motrices en fonction d'une satisfaction réalisée après l'activation du circuit neuronal.

5 Modèle mathématique, simulation et perspectives

Pour décrire les réflexes et leur évolution nous choisirons un modèle dynamique. La relation entre l'excitation d'une surface sensitive et le comportement correspondant sera représentée par un champ vectoriel sur cette surface.

Les zones foveales sont les attracteurs de ces dynamiques.

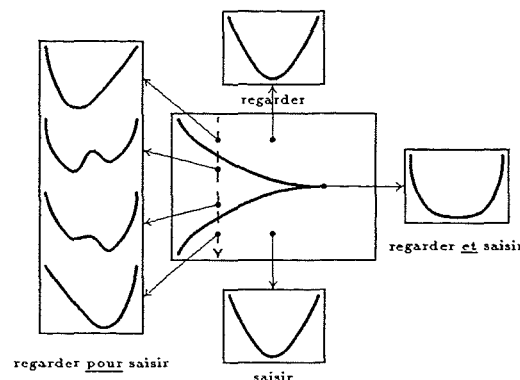
Dans une première étape plusieurs dynamiques sont apprises indépendamment ; ceci correspond à la phase d'adaptation des réflexes.

Ensuite la synchronisation des excitations des zones sensibles conduit à l'élaboration de schèmes qui sont un produit des dynamiques concernées.

Notre objectif serait alors d'utiliser la théorie des singularités pour mieux comprendre ces produits de dynamiques et leurs formes possibles à partir des dynamiques initiales.

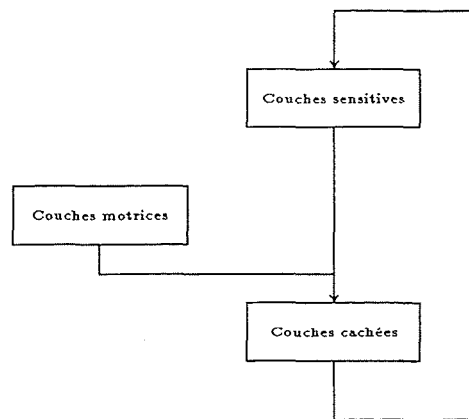
Prenons un exemple simple pour illustrer cette perspective. Au départ on a deux dynamiques : l'une représente le regard, l'autre l'agrippement. Le nouveau schème est une synchronisation des deux dynamiques : regarder et saisir. L'apprentissage du contrôle de ce nouveau schème conduit à des comportements du type : regarder pour saisir ou saisir pour regarder... on pourrait représenter ici cette structure par une fronce.

Nous voudrions comprendre si le produit, la synchronisation de deux dynamiques associées à deux schèmes n'est pas le centre organisateur d'une structure plus complexe qui s'exprime en langue par de nouvelles unités composées et qui permet ainsi une description véritablement morphogénétique de la structuration schématique et cognitive.



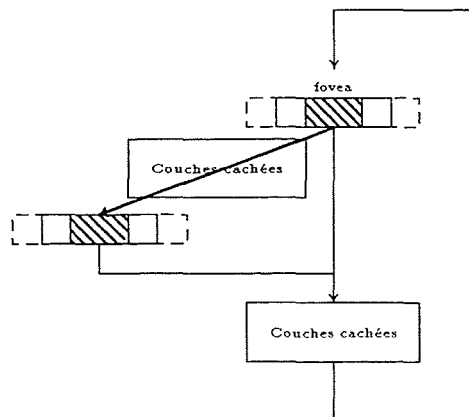
B. Victorri et Y. M. Visetti [VICT 96] [FONT 95] ont proposé et dirigé la réalisation d'une architecture de réseaux connexionnistes récurrents pour simuler des systèmes dynamiques. Ce cadre permet de tester, parmi les propositions énoncées, ce qui relève de l'extension et de l'interaction des dynamiques réflexes. C'est sur ces premières vérifications que nous concentrons actuellement nos efforts³. Voici résumée en quelques étapes la structure évolutive d'un tel réseau.

- Paramétrage du système par le choix des fonctions de transfert entre le système et son "environnement".



En jouant sur ces fonctions (homothéties, translations, rotations), on pourra étudier la variation induite sur les comportements sensori-moteurs.

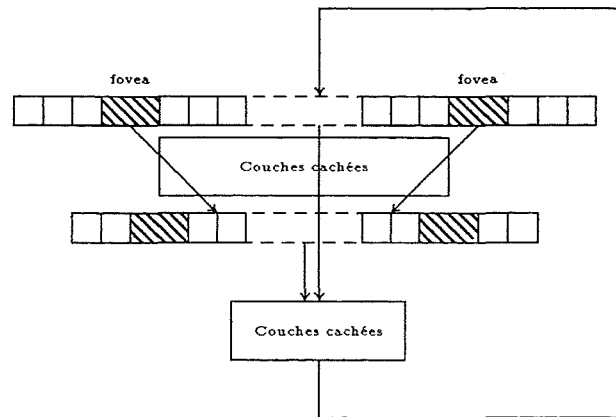
- Apprentissage de la réalisation de réflexes indépendants.



En procédant par paliers, pour réaliser l'extension topologique, le système apprendra à converger vers les attracteurs matérialisés par les foveas.

³Nous ne disposons pas d'un dispositif robotique qui, par l'intermédiaire de capteurs, ferait de l'environnement une donnée. Dans un premier temps la transformation de l'activation des zones sensibles après réalisation d'un comportement moteur sera programmée. Cette réduction provisoire du projet induit des remaniements qui devront être évalués.

- Couplage des réseaux par extension.



En se fixant comme nouvel attracteur la cooccurrence des deux attracteurs appris indépendamment, le système apprendra une dynamique produit dont, comme nous l'avons dit, on étudiera les caractéristiques par appel à un modèle morphogénétique.

Références

- [FONT 95] F. Fontaine. *Modélisation de l'interaction local/global en sémantique à l'aide de réseaux connexionnistes récurrents* Rapport de stage de D. E. A. sous la direction de B. Victorri et Y. M. Visetti. Université de Caen.
- [FUCHS, VICTO 94] C. Fuchs et B. Victorri (eds). *Continuity in linguistic semantics*. *Linguisticæ investigationes: supplementa*. John Benjamins Publishing Company. 1994.
- [PIAGE 77] J. Piaget. *La naissance de l'intelligence chez l'enfant*. Delachaux et Niestlé. 1977.
- [PLOU 92] S. Ploux. *Les modèles analogiques et le contrôle dans les systèmes à base de règles*. Thèse Lyon I. 1992.
- [PLOU 96] S. Ploux. *Une étude pour le traitement informatique de la synonymie*. Rapport interne. 1996.
- [QUART, SEJNO 1996] S. R. Quartz, T. J. Sejnowski. *The neural basis of cognitive development: a constructivist manifesto*. *Behavioral and brain sciences*. 1996.
- [REEK, SPORNS, EDELMAN 1990] G.N. Reeke, O. Sporns, G.M. Edelman. *Synthetic Neuronal Modeling: The "Darwin" Series of Recognition Automata*. *Proceedings of the IEE.*, Vol 78, N° 9, Septembre 1990.
- [STRER, SPELK 1989] A. Streri, E. Spelke. *Effects of motion and figural goodness on haptic object perception in infancy*. *Child Development*. 1989, 60, 1111-1125.

- [THEL 92] E. Thelen *Time-Scale Dynamics and the Development of an Embodied Cognition* in *Mind as Motion Explorations in the Dynamics Cognition*. eds R. P. Port, T. van Gelder. MIT Press.
- [THOM 77] R. Thom. *Stabilité structurelle et morphogenèse* 1977, InterEditions, Paris.
- [THOM 1980] R. Thom. *Modèles mathématiques de la morphogenèse* 1980, Christian Bourgeois, Paris.
- [VICT 96] B. Victorri. *Polysémie et construction dynamique du sens* Hermès 1996.
- [VIVI 92] P. Viviani and N. Stucchi. *Biological Movements Look Uniform : Evidence of Motor-perceptual Interactions* *Journal of experimental psychology : Human perception and performance*. 1992, Vol. 18. pp 603-623.
- [WAGN 74] R.A. Wagner. *Order- n Correction for Regular Languages*. *Communication of the ACM*, 17, 1974.
- [WANG 91] P. Wang and A. Gupta. *An Improved Structural Approach for Automated Recognition of Handprinted Characters*. *World Scientific Series in Computer Science*. Vol 30. 1991.

