Les systèmes adaptatifs à architecture d'agents dynamiques : une approche de la conscience artificielle

Alain Cardon
LIP6 Paris VI
UPMC Case 69
4 Place Jussieu
75252 Paris Cedex 05
Alain.Cardon@lip6.fr

Résumé :
Nous présentons certains caractères typiques de la conscience vue comme une organisation capable de générer des formes de pensée par émergence, en nous référant à l'architecture du cerveau et du système nerveux central. Nous décrivons ensuite un modèle au niveau symbolique, représentant certains caractères d'une conscience artificielle comme système adaptatif auto-organisateur. Enfin, nous présentons un système effectivement implémentable, basé sur le modèle et utilisant des organisations d'agents hybrides et dynamiques à structure reconformante.

Mots-clés : Système adaptatifs, Auto-organisation, Systèmes multiagents, Conscience artificielle.

Abstract:
We present some typical characters of consciousness, as an organization generating some form of mind by emergence. We refer to brain and nervous central system architectures. After, we describe a model at symbolic level, putting forward some characters of artificial consciousness

as an adaptive and self-organizing system. Finally, we present an effectively well-constructing system, using hybrid and dynamic multiagent self-organized systems.

Keywords: Adaptive systems, self-organization, Multiagent systems, Artificial consciousness.

PLAN

Sixième partie : Génétique et reproduction d'agents

Conclusion

Introduction

Le problème de la relation, et donc de la dualité ou de l'identité, entre le cerveau et l'esprit est posé depuis fort longtemps. La manière dont il est abordé dépend beaucoup de l'époque et du chercheur : ce n'est pas un problème simple, au sens où son domaine de définition et ses caractères sont particulièrement profonds et restent toujours non clairement identifiés.

Ce problème est typique de la philosophie de l'esprit, avec des approches se référant à la notion platonicienne d'Idée où à celle, phénoménologique, d'Etre comme dépassement du non-Etre [Heidegger 86]. Certaines positions, radicales, laissent à penser qu'il y a un caractère inconnaissable dans l'esprit et surtout qu'il ne peut pas être raisonnablement envisageable de simuler une quelconque conscience artificielle sur un calculateur symbolique [Searle 92].

Mais on doit, d'autre part, considérer les résultats des recherches scientifiques en psychanalyse et en neurophysiologie. Ces deux disciplines s'opposent radicalement sur des problèmes conceptuels [Hochmann 96]. L'approche psychanalytique pose son domaine avec une ambiguïté volontaire, considérant que le sujet de la connaissance est régi par des forces obscures, à exhiber en partie, tant que faire se peut et se traduisant par des tendances de l'être qui resteront toujours floues et non directement observables. La pratique usuelle de la psychanalyse est l'analyse, avec sa problématique non objective de transfert. Elle vise cependant à donner un certain statut rationnel aux sentiments, avec des caractères conceptuels prenant place dans le champ du langage. D'un autre côté, l'approche des neurosciences est radicalement objective et consiste à localiser, dénombrer et classer les éléments physiques fondamentaux de ce qui constitue le phénomène de conscience, localisé dans le cerveau physique. Nous pouvons qualifier cette seconde approche de montante, puisque on part des constituants de base du cerveau physique, au niveau cellulaire et biochimique, pour en déterminer, petit à petit, par des observations très fines, les fonctions agrégatives multiples, c'est-à-dire les cartes cognitives et leurs rôles, puis les fonctions des relations entre ces agrégations. L'objectif est de comprendre par exhibition complète de l'organisation du cerveau, ce qu'est effectivement un état mental, ce qui le cause et pourquoi il serait tel il apparaît.

Mais la définition de modèles reste incontournable et l'observation, biologique ou psychanalytique, ne fait pas toute la découverte [Thom 72]. Nous allons approcher le process de prise de conscience, c'est-à-dire l'acte mental de génération d'une pensée, sous un aspect spécifiquement informatique. C'est une prise de position délicate mais légitime, car une telle étude fait bien partie de cette discipline.

Il s'agit donc :

1. de poser les limites du domaine étudié,

2. de définir un modèle symbolique avec ses caractères typiques,
3. de définir un second modèle architectural aux caractères implémentables,
4. de réaliser effectivement un système, pour qu'il s'exécute sur ordinateur.

Evidemment, le système exposé ne sera pas une conscience artificielle telle la conscience humaine peut être, mais aura, seulement, certains caractères fondamentaux, non triviaux, de l'action de prise de conscience. Cette approche permettra de préciser comment un acte de prise de conscience peut effectivement être rendu constructible, c'est-à-dire connaissable par sa constitution et sa construction, à un certain niveau et sous des hypothèses précises. Nous désirons donc construire un système, que nous nommerons "système à boucle d'interprétation", qui présentera des similitudes comportementales avec le cerveau, l'esprit et la conscience des individus vivants qui en sont dotés.

Précisons d'abord le domaine de définition de ce que nous nommons process de prise de conscience. Il n'est pas question de se référer au seul niveau cellulaire pour réaliser le système, la barrière de la complexité étant définitivement insurmontable, et il n'est pas non plus suffisant de modéliser le système comme une procédure classique manipulant simplement des variables, si nombreuses soient-elles. Nous nous appuierons simultanément sur les approches psychanalytique et neurophysiologique, en les unifiant sous leur aspect systémique et organisationnel. Le domaine sera celui de la complexité organisationnelle et, plus précisément, celui des architectures complexes se reconformant par elles-mêmes, comme c'est le cas pour le réseau neuronal du cerveau.

Le cerveau biologique à en fait un double mode de fonctionnement [Vincent 86]. Il y a un cerveau dit neuronal, déterminant des aires spécialisées par conformations de réseaux de neurones temporairement connectés, avec une entrée nerveuse permettant de prendre des informations via les organes des sens et une sortie permettant la réalisation des programmes moteurs. Il y a également un mode de fonctionnement du cerveau dit *humoral*, libérant des neurotransmetteurs et des neurohormones à actions hormonales locales ou lointaines dans tout le corps, vers des fonctions motrices. Le mode humoral du cerveau est plutôt celui de la subjectivité tandis que l'autre mode, architecture neuronale plastique, est plutôt celui de la mémoire symbolique et de la déduction rationnelle par conception de buts. On peut raisonnablement penser que l'intentionnalité réfère à une fonction de tension créée par le mode humoral plutôt qu'à une stratégie rationnelle développée dans le néo-cortex. Un état mental courant est donc un état *global d'organisation* de certaines entités des deux types, situant *tout* l'individu dans son environnement et dans lui-même.

Nous allons précisément étudier le process de prise de conscience, c'est-à-dire de génération du sens d'une certaine situation, perçue, conçue puis exprimée par un système informatique particulier, de manière similaire au processus de génération de pensées dans la conscience humaine. Le domaine retenu sera celui des système auto-organisateurs, qui produisent des états globaux par émergence et reconformation de leur structure [Cardon 97]. Le modèle symbolique de conscience artificielle conduira à une architecture complexe possédant deux caractères non dissociables, organiquement mêlés dans son fonctionnement : affectif *et* rationnel. Un système basé sur un tel modèle produira des états émergents qui auront pour caractère la globalité, l'absence de contrôle central, la réflexivité, l'auto-organisation et surtout l'appréciation de sa propre organisation. Il produira des états dits globaux, semblables aux états mentaux qui décrivent les processus d'élaboration d'une forme de pensée, avec des artefacts de caractères subjectifs bien présents, pour introduire en permanence dans le système la notion d'intentionnalité artificielle.

Dans un tel système, l'entité minimale correspondant à l'aire cognitive ou à la génération de pensée, sera une certaine forme dynamique, à la fois état et action, de type d'agent cognitif, opérant essentiellement au niveau organisationnel, avec une notion d'appréciation géométrique de cette organisation, la notion de paysage d'agents [Cardon 97].

La vie artificielle s'intéresse principalement aux robots, c'est-à-dire à des animats qui se comportent "comme si". Mais il y a entre l'original qui vit et l'animat construit de toute pièce une distinction radicale. L'original vivant, et quel qu'il soit du moment qu'il a une reproduction sexuée, est en filiation dans son espèce alors que l'animat surgit soudainement. La filiation dans l'espèce pose et résout à un certain niveau le problème de la vie et de la mort [Vincent 96] et place l'original en "déval dans le temps" [Heidegger 86]. L'animat doit son comportement a un programme complexe activant des mécanismes également complexes, mais qui revient en fin de compte à un programme entièrement écrit, qui se déroule. On doit donc s'efforcer, pour réduire cette distinction entre animat et original vivant, de le construire pour qu'il possède "un sens de soi", c'est-à-dire qu'il soit conscient d'une certaine façon, par lui-même, de lui-même. Alors pourra-t-il peut-être poser son problème d'être, son "déval dans le temps" [Heidegger op. cité].

Il y a un certain intérêt à procéder à de telles recherches, ne serait-ce que pour produire un outil de simulation permettant de valider les hypothèses de construction d'un système doté de traits de conscience. Il y en a un autre, qui est de faire évoluer à la limite les systèmes informatiques, pour les détacher des contingences du vivant d'aujourd'hui, en fait de ses utilisateurs. Cela est dans l'air du temps, par réaction à des démarches par trop rationalisantes, qui ramènent le vivant à un objet, même quant il s'agit d'un autre semblable à soi-même. Et un caractère de la conscience artificielle serait de définitivement l'empêcher, par construction, de se dévaluer soi-même suffisamment pour pouvoir un jour dévaluer ses semblables.

Notre étude développe successivement quatre points : nous étudions certains caractères connus du process de prise de conscience dans le vivant supérieur, puis nous présentons une modélisation symbolique calculable d'un système auto-adaptatif présentant des caractères typiques de ce process. Ensuite nous présentons une architecture d'organisations d'agents, conforme au modèle, et enfin nous présentons les caractères de la complexité organisationnelle et de l'évolution du système auto-adaptatif par des croisements et mutations d'agents.

Ces réflexions s'appuient sur des travaux de recherches en cours, répartis dans plusieurs thèses dont les thèmes réfèrent au problème de l'interprétation, et que nous aimerions poursuivre. Notre objectif est très clairement de placer les problèmes d'auto-adaptation, d'émergence et d'acte de génération de formes de pensée dans le domaine de l'Informatique, où ils sont très exactement à leur place, avec les caractères de cette discipline, c'est-à-dire bien exprimables, constructibles et validables.

Première partie
Esprit et Cerveau

Il est clairement énoncé par la plupart des spécialistes de la philosophie de l'esprit que celui-ci est une propriété de l'Homme qui n'est pas modélisable par un système symbolique, ni même approchable, à cause d'un caractère de non finitude intrinsèque de la conscience humaine [Searle 92]. Dans ces conditions, la première chose à faire est de définir certains des caractères de l'esprit en termes *constructibles,* pour sortir du champ philosophique ou métaphysique hors de propos de l'Informatique [Leidlmair 95] et pour faire correspondre aux problèmes de ces discipline les nôtres, qui sont ici de doter certains systèmes des moyens de *se* percevoir. Ensuite, il restera à définir un modèle symbolique prenant en compte ces caractères constructibles, éventuellement réduits par rapport à l'esprit, pour enfin réaliser le système correspondant, en utilisant un langage de programmation adapté à la difficulté du problème.

1.1. L'Esprit

La conscience peut être définie comme "la capacité de porter attention sur le monde environnant et sur soi-même" [Dehaene 90]. Il est admis, notamment en Sciences Cognitives et en Biologie, que cette capacité est essentiellement localisée dans l'organisation des éléments du système nerveux et particulièrement dans le cerveau, avec même aujourd'hui une notion de "localisation distribuée" [Jeannerod 83]. Ainsi, les liens entre niveau perceptif et activations locales dans le réseau neuronal sont étudiés dans ce sens. On se réfère par exemple aux études sur des individus comme l'aplysie, petit escargot de mer dont le cerveau n'a que vingt mille neurones, où les simulations de groupes de neurones, les fonctions et rôles des neurotransmetteurs sont mis en relation avec le comportement de l'animal, en étant très finement explicités et permettant de définir des relations causales entre activité neuronale localisée et comportement physique [Scheller et al. 82]. Les travaux se focalisent donc sur les fonctions de certaines zones du réseau neuronal et leurs relations avec certains comportements réactifs de l'individu, l'Homme compris. Le niveau perceptif, conduisant à une attitude typique de l'individu devant une situation donnée, est celui qui est principalement étudié pour établir des relations avec les zones actives du cerveau.

Mais les grandes difficultés résident dans la définition de la conscience de soi, dans le statut et le fonctionnement du langage, dans la définition de l'intentionnalité. Ce sont ces catégories, appartenant typiquement aux modèles psychologiques, auxquelles nous allons plutôt nous intéresser, et que, par la suite, nous allons tenter de modéliser dans le domaine symbolique.

1.1.1. Pensée et Penser

La conscience est ce qui permet de générer des pensées, sur toutes les choses, y compris sur elle-même. Nous définissons la génération d'une pensée, chez un individu quelconque, comme la capacité à *se représenter en situation d'expérimentation dans le monde*.

Explicitons cette définition. Une pensée est donc générée par une action et, en ce sens, elle contient un certain caractère d'intentionnalité. Elle a une certaine raison, sinon une certaine cause à être. La notion de représentation indique une distance et une conceptualisation opérées par l'individu lui-même par rapport au monde qui l'entoure et qu'il perçoit par ses sens ou par des références à sa mémoire. Le monde est l'environnement conceptualisé par cet individu, qui ne perçoit que ce qui est pour lui perceptible et concevable. La notion de situation indique un acte de mise en relation particulière entre l'individu et le monde, relation fondée sur une appréciation de la temporalité et qui n'est jamais une correspondance bijective. Nous nous s'appuierons ici sur l'approche phénoménologique développée par M. Heidegger pour préciser ce qu'est exactement cet acte premier de "situationnement dans le monde", pour en fait comprendre ce que signifie exactement et entraîne l'usage du concept "être-au-monde" [Heidegger 86]. L'expérimentation dont il est question est la notion d'usage de la finalité dont la conscience est munie et que l'on pose, dans cette approche constructible, comme un moyen *inévitable* d'une certaine forme d'action dans un domaine organisationnel.

L'acte de penser est celui d'un individu qui est, d'une certaine manière, le *sujet* de son action de penser et qui perçoit, par un certain *souci* qu'il possède en propre à cet instant de son existence, une certaine chose de son environnement ou de lui-même, qui se représente cette chose, pour lui-même, en la concevant par son esprit et qui ensuite, éventuellement, décide d'agir physiquement. L'action de penser, puisque penser est une action au sens strict du terme, est comprise comme la façon de se mettre à la distance de l'environnement pour en concevoir une certaine *représentation*. La pensée sera, par définition, le résultat, strictement interne à l'individu, de cet acte de situationnement.

Il y a deux questions principales à propos de cette action de penser. Quelle est la raison de cette action, peut-on se ramener à une tension subjective représentée par un certain déséquilibre "humoral" plus ou moins local [Vincent 96] ou encore à une certaine pulsion conduisant à une dépense d'énergie compensée, avec S. Freud ? Nous choisirons de nous ramener à une tendance réorganisationnelle, par tension dans l'organisation elle-même. La seconde question est de décider où placer le résultat de cette action : strictement dans le substrat concret du cerveau, ou bien dans un espace plus vaste et finalisé [Bergson 53]. Cette seconde question ne sera pas pertinente pour notre propos, puisque, pour un système informatique, le produit de son fonctionnement interne ne peut que référer à sa propre structure. Certains trouveront là la distinction de genre entre l'artificiel construit par l'original et cet original constructeur d'artificiel. La réponse à la première question nous conduira à finaliser l'acte de penser artificiellement par une simple possibilité d'expérimentation à manipuler un processus de génération de formes de pensées, dans un système organisationnellement complexe, ce que nous allons développer.

Hypothèses de limitation

Nous nous intéressons à certains caractères de la pensée, en ce que celle-ci provient essentiellement d'un niveau organisationnel du cerveau et de "soi-même situé dans le monde".

La forme de pensée que nous considérons a donc une source et un déploiement bien précisés, mais elle réfère à une certaine extension de l'individu, spatiale et temporelle, non évacuable.

Nous nous intéressons au processus générateur de la pensée pour le modéliser dans un système symbolique, avec les caractères typiques du niveau symbolique, différents des caractères du vivant correspondant. Le système générant des pensées artificielles sera enchâssé dans son environnement, identitaire, tout en étant dépendant de lui-même et de l'environnement.

Il y a, dans le surgissement d'une pensée, avant et pendant toute construction rationnelle langagière qui l'explicite, une action dans un espace de subjectivité. Il y a un individu désirant, en tension dans son environnement *et* dans lui-même, soumis aux forces de sa subjectivité. Cette subjectivité provient principalement des échanges dits *humoraux* dans le système nerveux et entre celui-ci et le reste de son corps, déterminant un *état central fluctuant* qui représente "les mouvements des organes internes, les sécrétions des glandes et le jeu des circuits neuronaux" [Vincent 86]. La pensée surgit dans un système physique global, complexe, où toute modification humorale locale, c'est-à-dire hormonale ou électrochimique, influe sur toutes les autres par réaction transitive. Cette source de toute pensée, si fortement subjective et complexe, détache définitivement celle-ci d'une chaîne d'inférences strictement rationnelles : la conscience artificielle est au-delà du processus symbolique d'inférence déductive ou inductive sur des concepts bien dénotés, car elle contient strictement ce processus, le complexifie et l'étend.

Remarquons trois caractères typiques de l'action de penser :

1. La pensée surgit d'abord dans un espace subjectif, qui est l'état central fluctuant de l'individu, son espace "d'être dévalant dans le temps" [Heidegger op. cité] par un certain souci, pour détacher une certaine chose prise en considération. Cette chose, détachée par l'esprit, est liée à son espace de génération : elle est et reste complexe.
2. L'objet dont il est question dans l'acte de penser est mis à la distance du monde, en le représentant et en le concevant dans l'esprit : c'est un acte de liberté considérable de l'individu par rapport au réel qui est ainsi reconstruit à loisir. Cette mise à la distance est bien définie dans toute la sémiotique triadique, qui distingue l'objet réel du signe le représentant et les met en relation par un processus d'interprétation [Peirce 84].
3. L'acte de penser est hors d'atteinte de l'individu qui pense et le résultat de cet acte, la pensée produite, surgit, pour lui à sa conscience, dans son *immédiateté* [Cardon 96]. Il n'est pas préhensible dans son déroulement et seul le résultat est apparent, pour l'individu lui-même. C'est bien une certaine émergence.

1.1.2. Pensée comme "forme"

La pensée est donc le résultat conscient, propre à l'individu qui pense, d'une action d'interprétation d'une certaine chose qu'il peut concevoir. Nous considérons que c'est une *forme*, c'est-à-dire une certaine entité déterminable, aux caractères d'action et d'organisation, structurellement complexe, et qui est en relation avec une chose du monde, éventuellement une autre forme. La propriété de cette forme est d'être appréhendable, au moins partiellement, par l'individu qui l'a générée, comme "un objet".

Nous allons faire l'hypothèse, nécessaire pour étudier des formes de pensée artificiellement produites, que la pensée se ramène à (ou bien est) un caractère organisationnel d'un certain niveau, que ce niveau, qui correspond en fait au niveau psychique, est considéré comme distinct du niveau biologique [Hachmann 96]. Ces deux niveaux sont organiquement liés et des concepts organisationnels permettront de définir leurs relations.

Caractérisation constructible de la pensée comme forme

> *La pensée est le produit d'une auto-organisation complexe dans le substrat du cerveau, possédant un caractère physique à un certain niveau, considéré comme une forme manipulable, mais d'un autre niveau que les constituants de base, biologiques, sur lesquels elle prend place.*

Nous considérons donc certains caractères du niveau biologique, puis du niveau organisationnel de celui-ci qui réfère à d'autre catégories et dans lequel prendra place le concept de "forme de pensée". Le système à boucle d'interprétation sera le générateur de ces formes.

Toute forme de pensée est et sera complexe, en ce sens qu'elle n'est pas totalement détachable de l'organisation de l'ensemble du système qui l'a générée. Notre allons poser une hypothèse de détermination du sens.

Hypothèse sur le sens

> *Nous considérerons que la forme représentant le produit d'un processus de pensée est, de façon suffisante, une entité correspondant au sens du processus de réorganisation courant du cerveau neuronal et hormonal, à un instant donné.*

Nous ne procédons pas à un détachement réducteur de la forme de son substrat, en définissant cette entité symbolique nouvelle que l'on nommerait pensée et qui serait le résultat exhibé d'un programme, mais nous distinguons certains caractères structurels d'une organisation qui sont susceptibles de représenter une forme quasi-stable et signifiante pour cette organisation. En tant que forme, celle-ci a des caractères structurels, morphologiques, géométriques, dynamiques que nous allons pouvoir prendre en considération et qui vont nous permettre de la construire puis de la représenter.

1.1.3. Conséquences

Il y a dans le cerveau, une spatialisation des choses perçues dans l'environnement de l'individu, qu'elles soient physiques ou temporelles, ce qui permet de les manipuler, mais avec l'impossibilité d'agir sur le déroulement temporel de ce processus spatial. Il y a une sorte de dualité espace-temps entre cerveau qui spatialise par des processus reconformants temporel et esprit qui produit à un moment du temps un processus dont le résultat est spatial.

Le système artificiel qui représente cette action de générer des formes de pensée sera basé sur une sorte de boucle infinie auto-contrôlée. C'est une "sorte" de boucle, car le processus de bouclage, temporel, ne devra pas être commandé de manière externe aux transformations répétées dans la boucle, mais devra être le principe organique des transformations qui ne pourront *que* se répéter, sous contraintes, pour produire une forme spatiale. Pour cela, il faudra que ce qui correspond à la "production" du système soit naturellement déclenchante de nouvelles productions. Nous devons donc songer, quant à la conscience artificielle, à un système clos, sans entrée ni sortie explicites, mais avec des états émergents distingués.

1.2. Le cerveau comme organe localisé ou central : deux catégories dans le vivant

Pour modéliser un système représentant une conscience artificielle, générant des pensées artificielles, il nous faut caractériser le rôle du cerveau en tant que générateur de formes de pensées. Le premier élément de caractérisation du cerveau ne lui appartient en fait pas : il s'agit du caractère évolutif de l'hôte abritant un cerveau et qui est et n'est que le résultat d'une longue évolution des espèces sur lesquelles la sélection naturelle a plus ou moins opérée, partant d'individus typiquement réactifs au cerveau embryonnaire pour aboutir à l'espèce de l'hôte, dont il est, à la suite d'une reproduction sexuée, un éminent représentant. Cette considération temporelle va nous permettre de distinguer deux catégories d'individus dans le vivant.

Manifestement, il y a les espèces dont les individus ont un cerveau capable de générer des pensées et celles dont les représentants n'ont pas de cerveau générateur de pensée, c'est-à-dire pas de cerveau du tout ou n'ont pas la possibilité de générer des pensées comme produit représentatif d'une interprétation de quelque chose. On distingue ainsi deux classes dans le vivant : le vivant non vertébré, qui se construit spatialement et le vivant vertébré, qui se construit *dans* le temps et l'espace. C'est ce placement dans le temps qui sera en fait le caractère distinctif entre l'individu vivant qui peut penser en se situant dans son monde et celui qui ne le peut pas. On retrouve ainsi le problème temporel si important pour M. Heidegger [op. cité].

1.2.1. L'insecte : un individu spatial

Un représentant typique du vivant non vertébré est l'insecte. Pour l'insecte, la segmentation de son organisme est déjà conçue avant la morphogenèse, dans la colinéarité de son programme génétique. A la colinéarité spatiale de son génome correspond une expression temporellement linéaire des gènes, qui s'expriment les uns après les autres, la fin de l'expression de l'un correspondant exactement au début de l'expression du suivant pour coder les protéines. Ainsi, le corps de l'insecte se forme spatialement, par une succession d'expressions génétiques indépendantes, comme se déroule le programme d'un automate à états. Tous les individus de ce type sont alors semblables, car résultant du même programme génétique ayant fonctionné de la même façon, ou bien diffèrent monstrueusement par des mutations génétiques singulières. Le cerveau de l'insecte est un organe fonctionnel réalisé spatialement par codage de certains gènes, à un moment précis de la transcription.

Que pourrait penser, avec ce cerveau primaire et construit fonctionnellement à un moment de la morphogenèse, un tel individu, généré de manière si semblable à la génération des autres de la même espèce ? Quelle est son autonomie à penser *pour* lui-même, lui qui est si semblable à ses semblables ? Peut-il, avec un organisme si mécaniquement et fonctionnellement construit par morceaux successifs, abriter des pensées propres globales relatives à son éventuelle identité, personnelles, lui permettant de se situer dans un monde qu'il se représenterait en particulier ?

C'est très peu vraisemblable, sauf en admettant que l'esprit émanant de ce cerveau organe si local, est indépendant de sa structure et de sa construction, ce qui est une hypothèse audacieuse.

On pourra donc représenter très justement de tels individus par des agents réactifs [Drogoul 93], au comportement typiquement conduit par un ensemble de programmes fonctionnels distincts et coordonnés, opérant avec de nombreuses fonctions typiques de rôles comportementaux. Il sera peu utile de les doter d'un processus de génération de pensées, qui ne pourrait de toute façon pas permettre de les faire se représenter identitairement en se situant dans un monde qui serait, pour eux, le leur. Ces individus n'ont pas, organiquement, le sens du temps et d'eux-mêmes situés dans le temps et ne pourront pas concevoir de situation par représentations spatio-temporelles. La problématique principale des études dans ce domaine est alors l'émergence de comportements sociaux basés sur les comportements individuels automatiques des agents réactifs [Drogoul, op. cité].

1.2.2. Les vertébrés : le placement dans le temps

En ce qui concerne la catégorie des vertébrés, l'expression des gènes est située dans l'espace et également et surtout dans le temps. Une modification de la durée d'expression d'un gène quelconque du génome et notamment des homéogènes transactivateurs des formes physiques des organes, peut considérablement modifier l'expression des gènes suivants et donc la morphologie de l'individu. C'est le phénomène d'hétérochronie, qui permet de produire des individus physiquement uniques.

Ainsi, la formation de l'individu vertébré va dépendre du temps mis par chacun des transcripteurs de son programme génétique pour s'exprimer. Sa conformation physique est soumise aux aléas du contexte d'expression génétique et intègre, organiquement, le temps. Et pour un vertébré, les gènes exprimant le cerveau ne sont pas alignés sur un seul chromosome mais sont dispersés dans *tout* le génome. La formation du cerveau apparaît alors comme un processus global, permanent dans la durée de la formation morphologique de l'individu.

Le cerveau est codé comme organe global dans l'espace du génome et il se construit en se plaçant dans la durée de l'expression de chacun des gènes le codant. Le système nerveux apparaît chez les vertébrés comme organe unifiant de la construction de son architecture, permettant de représenter l'unité de celle-ci en connectant toutes les parties du corps par des voies nerveuses multiples : l'individu est physiquement construit autour de son système nerveux qui, lui-même, joint toujours le cerveau. Nous pouvons en déduire le placement de l'esprit, exprimé par le cerveau, dans la temporalité de l'individu.

Placement temporel de la pensée

> *Un individu conçu par hétérochronie est organiquement construit dans le temps, pour ainsi représenter par lui-même, c'est-à-dire par la morphologie propre de l'organisation de son cerveau, ce qui est spatial et temporel dans son entour spatial et temporel.*

1.2.3. Conséquences

La construction d'un système artificiel ne peut être réalisée aujourd'hui autrement qu'ex-nihilo : les éléments du système sont mis en place les uns après les autres puis, à un moment donné, le système est mis en action, globalement et il peut alors se comporter et éventuellement évoluer. Ce système peut avoir une faculté d'apprentissage, comme renforcement de tendances, mais qui n'agit pas au niveau de sa construction, comme agit l'espèce qui assure l'évolution des formes et caractères de ses individus. On pourra cependant le doter d'une capacité d'évolution organisationnelle continue, ce que l'on sait faire en Informatique depuis les travaux de l'école lispienne où le produit d'une fonction peut en être une autre nouvelle. Cela reviendra à définir le système à boucle d'interprétation comme une entité en évolution plutôt qu'en simple production.

Il faudra surtout munir le système d'un caractère de perception de son organisation même, lui permettant d'une certaine façon, la reconnaissance de son état organisationnel *global*. La reconnaissance de l'état organisationnel global signifie que la pensée artificielle produite est le *sens* de l'organisation complexe qui l'a produite. Cela sera réalisé dans le système à boucle d'interprétation par autoréférence à sa structure et c'est une hypothèse raisonnable quant au cerveau humain produisant de la pensée, que nous posons.

Le caractère d'émergence de la pensée compris comme sens de l'organisation qui la produit, ne se réduit pas à une certaine fonction calculant une émergence, car il s'agit d'appréhender *globalement* l'état organisationnel de sous-systèmes liés qui, eux, pourront avoir des caractères fonctionnels précis. Un tel système ne se représentera donc pas par un agent réactif ni par un agent cognitif omniscient à la structure compliquée, mais par des organisations fortement liées d'agents, avec une architecture plastique qui permettra d'exprimer dans des émergences réifiées les caractères générateurs de l'organisation. Nous serons donc conduits à réifier dans des agents spécifiques l'état de certaines organisations d'agents et ceci, dans un processus bouclé.

1.3. Le cerveau comme organe plastique opérant des réorganisations

Le caractère temporel de la construction du cerveau chez l'individu vivant supérieur lui permet certainement de se penser et de se concevoir comme entité unique et unifiée. Il reste à préciser le procédé de reconnaissance utilisé par le cerveau pour générer la pensée.

L'être humain, considéré au sens biologique, est constitué de son système nerveux central où 10^{11} interneurones, dont les neurones du cerveau, sont connectés à 10^6 motoneurones (connectés aux muscles) et 10^7 cellules sensorielles réceptrices [Maturana, Varela 94]. Ce substrat physique de neurones est plastique, au sens où les connexions entre neurones sont plus ou moins modifiées, continûment. Il existe des groupes connectés de manière relativement stable et disjoints d'autres groupes, ce qui conduit à penser à l'existence d'aires fonctionnellement spécifiées. Il y a une notion de topologie dans le système nerveux et le cerveau, qui ont donc une architecture faite d'amas relativement changeants de neurones, typiques de certains rôles ou de certaines fonctions. Mais les neurones, transmetteurs de l'influx nerveux, n'opèrent pas seuls. Ils subissent ou entraînent l'action des hormones véhiculés dans tout le corps. Le cerveau, quant à lui, génère des neurohormones et des neurotransmetteurs via les cellules gliales et fonctionne donc aussi comme une glande hormonale.

Dans une optique strictement biologique, le rôle du cerveau est compris comme étant celui d'un "système sélectif de reconnaissance opérant par contrastes et similitudes entre des formes" [Edelman 87]. Ces formes, reconnues et manipulées dans sa structure même, sont des zones plastiques du cerveau, des amas de neurones liés et immergés dans un environnement hormonal, qui ont des *fonctions* plus ou moins spécifiques quant au rapport qu'elles peuvent avoir avec le réel qu'elles décrivent. Le cerveau opère physiquement des relations et des relations de relations entre ces amas neuronaux. On admet ainsi, dans cette optique, que l'esprit, générateur de pensées se ramène à l'ensemble des rapports entre des formes neuronales et des rapports entre rapports, tous ayant trait à des catégorisations fonctionnelles sur des éléments du réel. Dans cette optique, ce sont plutôt les caractères des relations entre les éléments du réel objectif qui induisent des relations structurelles correspondantes dans l'architecture réceptrice et plastique du cerveau.

On doit quand même remarquer que la nature de l'architecture plastique du cerveau est sensiblement différente en genre de celle du réel constituant l'environnement compréhensible de l'individu, car elle doit posséder, intrinsèquement dans sa construction et dans son fonctionnement comme système, les moyens d'exprimer *le sens* du réel, ce qui est quand même au-delà du réel objectif soumis aux lois du mouvement. L'esprit distingue des objets, révèle des causes et des conséquences, là ou il n'y a qu'action et réaction apparente.

Dans le système à boucle d'interprétation, les informations à propos du réel seront des éléments faits pour déclencher des distinctions, des associations causales, faits pour lier et distinguer certaines entités et certains groupes d'entités et ne seront pas de simples *images* du réel. Le système devra posséder un niveau perceptif propre, suscitant l'activité de ses formes internes, les formes et leur activités correspondant à des choses perçues dans le réel. Le système aura pour rôle d'activer une organisation d'un autre niveau que celui des entités et représentant des relations, anomales ou causales entre entités.

La correspondance entre formes externes et formes internes les représentant sera remplacée par un processus d'interprétation, comme cela est bien explicité dans la sémiotique triadique [Peirce 84]. Ce mécanisme sera évidemment dans le système, au bout du compte, une suite d'instructions processées séquentiellement, mais l'effet de ce processus sur la structure du système représentera une réorganisation complexe.

Le principe de réorganisation

> *L'activité du cerveau générateur de formes de pensées sera considérée comme essentiellement réorganisatrice, par la manipulation de ses multiples constituants, en conformité relative avec la structure du réel appréhendé. La pensée sera le produit, à un certain niveau organisationnel, de ces réorganisations.*

Il reste, ce qui n'est pas rien, à préciser les formes dont il est question, ce qu'elles sont, d'où elles viennent, comment elles se transforment, ce qui les active et de quelle façon. Ainsi, nous adopterons une approche connexionniste au niveau de l'organisation, puisque le système est bien un réseau changeant de formes plastiques liées, mais également cognitiviste, car les formes réifieront des fonctions ou des concepts, sans jamais se réduire à des neurones formels numériques.

1.4. Caractères de l'architecture du cerveau

Le cerveau est un organe fortement isolé du corps par la barrière hémato-encéphalique. Il peut donc, par ce fait, opérer une fonction spécifique générale dans l'organisme. Cette fonction sera, pour nous, le situationnement de l'individu dans l'espace et le temps, avec le caractère que ce situationnement sera indissociablement lié et dépendant de l'état subjectif de l'individu.

Le substrat du cerveau est, d'une part, un réseau neuronal de complexité combinatoire très élevée, ce qui le rend extraordinairement plastique, et d'autre part un système dit humoral véhiculant des neurohormones et des neurotransmetteurs dans tout l'espace du cerveau et, transitivement, dans tout le corps. C'est un système dont la fonction apparente est de se remodeler et de se reconformer pour être. L'une de ses propriétés est d'être en relation interactive continue, par le système humoral, avec toutes les parties du corps qu'il habite et d'avoir ainsi prise à distance sur son hôte et sur son environnement. Il y a relation continue entre cerveau se reconformant, corps habité et environnement, formant l'état central fluctuant, aussi appelé *subjectivité* [Vincent 86].

L'architecture du cerveau des mammifères évolués est considérée aujourd'hui comme fonctionnelle. Elle forme une entité complexe et globale, en ce sens que toute partie a un rôle fonctionnel plus ou moins précis et peut être en relation avec n'importe quelle autre.

Le cerveau et le tronc cérébral sont reliés au corps par un système nerveux dit périphérique. Ce système permet, d'une part, l'activation des différents organes du corps : coeur, poumons ... et d'autre part l'activation des nerfs via les motoneurones et les neurones sensitifs. Toutes les informations venant des viscères arrivent à l'hypothalamus qui participe ainsi activement à l'homéostasie du corps. Le cerveau a une zone interface avec le corps physique, prenant en compte les informations sensorielles : la forme réticulée du tronc cérébral. Cette structure a à la fois une fonction d'intégration et d'activation vers les autres régions du cerveau. On pense aujourd'hui que la forme réticulée n'a pas une structure uniforme indifférenciée, mais qu'elle est composée d'une structure très fortement hiérarchisée, composée de nombreux sous-systèmes fonctionnels spécifiques.

Les activations de la zone interface, la forme réticulée, diffusent dans le reste du cerveau. La perception d'une situation externe arrive, en remontant les voies sensorielles, au néo-cortex frontal. Le néo-cortex décide, après représentation de la situation, d'une certaine stratégie d'action à entreprendre. Le néo-cortex met alors en jeu soit l'axe sympatico-médullo-surrénalien via l'hypothalamus, par libérations de peptides et d'hormones, ou bien l'axe cortico-surrénalien via l'hippocampe, avec libérations d'autres peptides et d'autres hormones [Vincent 86]. Le premier axe revient à considérer que la situation perçue est maîtrisée et que le corps doit attendre, alors que la seconde déclenche la fuite immédiate de l'individu. L'acte de penser, localisé dans la manipulation du néo-cortex, en relation avec les actions nerveuses et hormonales du reste du corps, est plongé, avant, pendant et après, dans une action globale de diffusion hormonale dans tout le corps.

Cette vision du fonctionnement physique du cerveau, complexe au sens systémique du terme, ne fait pas l'affaire du modélisateur. Il doit s'efforcer de réduire cette complexité. Selon une hypothèse ancienne, on pouvait partitionner le cerveau en deux : un cerveau cognitif et un cerveau émotif [Konorski 67], hypothèse qui est aujourd'hui abandonnée. Une autre hypothèse de dualité, avec un cerveau gauche dédié à la parole et au calcul opposé à un cerveau droit intuitif et poétique [Sperry 68] est également abandonnée. On a également émis l'hypothèse d'un cerveau triple : une partie reptilienne primaire enveloppée d'un cerveau paléomammalien faisant usage du souvenir, lui-même enveloppé d'un cerveau évolué, le néo-cortex. Toutes ces hypothèses structurelles expriment une identité mécaniste entre des caractères fonctionnels bien apparents et une architecture distribuée par *spécialisation*. Elles ne seront pas retenues en ce qui nous concerne, pour l'architecture du système artificiel, où les notions de globalité, de compensation et d'opposition seront préservées. Notre approche sera typiquement organisationnelle. On peut cependant faire l'hypothèse d'un dualisme fonctionnel, mais élémentaire, qui oppose à toute fonction cérébrale, humorale ou neuronale, une fonction inhibitrice, dans un système reconformant compensant des défaillances locales. Cette hypothèse soutiendra un principe de fonctionnement du système : l'opposition au niveau des composants actifs.

Clairement, les reconformations distinguent dans le cerveau des zones topologiques que l'on nomme des aires. La spécialisation de certaines aires, le fait qu'elles s'activent dans des situations comportementales typiques pour l'individu, résultent du fonctionnement même du cerveau et, si ces aires sont génétiquement localisées, elles sont généralement compensables, en cas de défaillance, dans la structure même du cerveau. Les aires spécialisées sont liées entre elles et ne déterminent donc pas des *structures* fonctionnelles fixes et uniques : la défaillance d'un caractère comportemental liée à une certaine zone du cerveau est toujours accompagnée d'autres défaillances comportementales annexes et également de compensations plus ou moins performantes, par transfert sur d'autres zones permettant de rétablir, tant que faire se peut, la fonction défaillante.

Le cerveau est donc un organe global à architecture typiquement plastique, où le caractère fonctionnel d'une aire est *relativement* localisé et dépend, pour sa localisation et pour son rôle, du caractère fonctionnel de nombreuses autres aires et également de l'état central fluctuant. Cette propriété de dépendance forte, topologique et fonctionnelle, à la fois locale et globale, est le caractère typique de la complexité organisationnelle profonde [Cardon 96].

Hypothèse de complexité des aires

> *Les aires fonctionnelles du cerveau ont, dans la génération de la pensée comme organisation spécifique, un rôle diffus, flou, imprécis et non indépendant, correspondant à des tendances plutôt qu'à des fonctions totalement définies.*

Un individu avec un cerveau lui permettant de penser peut être considéré comme un système d'expression d'idées contextualisées. Il n'opère de comportement et d'action que par l'action de son cerveau et pour le compte de celui-ci, le corps étant l'interface multimodale qui le place physiquement dans son environnement. Nous considérerons donc le système à boucle d'interprétation comme ayant une interface composée de capteurs multimodaux spécialisés, opérant en parallèle, les entrées étant des prises d'informations venant du contexte et transformées en formes potentielles d'action de reconformation de la structure du système. Cette action de reconformation sera complexe et non réduite à un processus d'inférence. Le système ne devra pas traiter ses entrées sous forme de simples données, mais activer des processus de reconformation pour générer des formes représentant du sens, sans qu'il n'y ait de lien direct entre information provenant du contexte et système générateur du sens. Un tel système sera dit *adaptatif*.

1.5. Le statut du langage dans le cerveau et dans l'esprit

Il n'y a guère qu'un seul type d'individu qui habite vraiment le langage : l'Homme. On ne sait pas comment la parole a pu surgir chez des individus auparavant muets de mots, qui se sont ainsi radicalement coupé de tous les autres, pour évoluer continûment dans leur aptitude à parler et pour enfin finir par habiter des langues verbales. Mais on sait que certains singes supérieurs, tel le singe bonobo, manipulent avec aisance plusieurs centaines de lexigrammes, avec une grammaire non triviale. Le langage est quand même partageable par d'autres et n'est donc pas l'apanage exclusif de l'Homme.

Pour de très nombreux biologistes, le langage est aujourd'hui considéré comme un instinct permettant d'acquérir un comportement social langagier [Pinker 94]. Ainsi, l'Homme posséderait de manière *innée* la faculté de pouvoir parler et la socialisation va lui apprendre la parole. Nous savons que la position de N. Chomsky est sensiblement différente, car beaucoup plus radicale [Chomsky 76]. Celui-ci pose l'existence, soudaine chez l'Homme dans tout le vivant, d'un organe mental absolument distinctif qui générerait une grammaire universelle, plus ou moins riche selon les individus. Outre une connotation inquiétante, cette vision surgissante de la propriété de parole se place en contradiction avec la théorie de l'évolution, ce qui la rend douteuse.

Remarquons que le langage de l'Homme est à double articulation. Chaque mot est d'une part constitué de formes minimales ayant sens : les monèmes et d'autre part d'entités n'ayant qu'un caractère phonique : les phonèmes. Ce double codage, qui opère simultanément dans l'énonciation et la reconnaissance des mots, sera à prendre en compte dans le système à boucle d'interprétation. Le mot est donc principalement verbal, avec une inscription physique dans l'écrit et est lié au sens qui lui est attribué.

Il y a une certaine aire dans l'hémisphère cérébral gauche dédiée à la manipulation de la parole. Cette aire, plus ou moins étendue (comprenant l'aire identifiée de Wernicke) est physiquement proche et est couplée avec l'aire assurant la manipulation physique des objets (comprenant l'aire bien connue de Broca). Dans le cerveau de l'Homme, et ainsi dans son comportement, le langage ne pourra être que lié à la manipulation des objets. L'Homme est plongé et vit dans son monde physique, par l'action et par la parole, cette dernière étant un acte d'expression parmi les autres.

L'élément central du problème de la pensée et du langage est en fait "le rapport entre la pensée et le mot" [Vygotski 85]. Il y a deux façons de voir la relation entre pensée et mot : on peut identifier pensée et mot ou bien on peut les disjoindre. Ces deux attitudes entraînent des positions différentes mais qui sont, l'une et l'autre, radicales. Ou bien on se plonge dans un univers linguistique, à la recherche de grammaires biologiques, ou bien on considère que la pensée est autre que son incarnation dans le mot, et le problème reste ouvert.

Une approche constructive décrivant la relation entre pensée et langage est sans doute différente de ces deux dernières. L'analyse du problème revient à s'intéresser, comme dans toute analyse, à des décompositions en parties, mais avec la propriété que les parties minimales obtenues, les unités non décomposables du système de mots, possèdent encore "toutes les propriétés fondamentales du tout" [Vygotski op. cité]. Ce sont des parties semblables à des parties autonomes, vivantes, mêlant toujours pensée et mot. Cette approche est aujourd'hui, dans d'autres domaines, celle de la systémique complexe [Le Moigne 90].

Une pensée peut partiellement se représenter par des mots. Elle est composée de parties, spécialisées peut-être quant au sens des mots du langage, fonctionnelles dans leurs rôles, mais qui sont encore des formes de pensée. Chaque partie incarne certains mots, qui en appellent effectivement d'autres, multiples, et qui est dépendante en fin de compte de toutes les autres parties. Cela est la définition de l'arrière-plan langagier de Searle, qu'il pose comme infini et non représentable : un mot appelle la langue en contenant une possibilité, au moins, de généralisation [Searle op. cité]. Le sens d'un mot sera cette capacité, spécifique à ce mot, de transitivité et de généralisation.

Une pensée est donc une certaine structure, centrée sur un certain groupe de mots, liée avec d'autres structures et avec d'autres mots et également composée de formes aux caractères spatiaux et temporels et de caractères subjectifs, c'est-à-dire de tensions et de force à se développer.

Nous retiendrons cette approche de la relation complexe entre pensée et mot, en munissant le système à boucle d'interprétation d'une représentation dynamique d'objets réifiant à la fois des mots, des parties de phrases, des attributs subjectifs et des formes de pensées. Le réel, extérieur et intérieur à l'individu, est ressenti grâce à l'activation de fonctions subjectives et il se dénote par des fonctions langagières, les deux étant mêlées. Nous représenterons les éléments minimaux de la pensée non par de simples structures symboliques comme des réseaux sémantiques, mais par des extensions structurelles et dynamiques de caractères langagiers, en considérant des structures d'actions complexes, des agents cognitifs capables, car cela sera leur rôle, d'appeler des agents par leurs réseaux d'accointances, capables donc de s'agréger en agents plus abstraits, d'agréger des agrégations, de se scinder en parties, de faire ainsi naître des clones de paroles dénotant des idées [Cardon 97].

1.6. Produire du sens : primalité du souci

La définition de la conscience, selon J. Pitrat, repose sur des caractères typiquement constructibles [Pitrat 93] : "Nous appelons conscience la connaissance que nous avons de ce qui se passe dans notre cerveau pendant que nous pensons. Il y a alors une analogie entre la conscience et l'auto-observation".

Nous pouvons prendre cette définition au pied de la lettre. Nous considérons que la conscience est connaissance d'une certaine organisation qui produit des pensées et est donc explicitable par des structures. Nous remplaçons, ce qui n'est pas rien et ce qui en fait change beaucoup de choses, cerveau par esprit. Il ne peut y avoir conscience que pour un sujet qui se déclare conscient : il y a donc une notion de *sujet*. De plus, il y a dans la définition, un principe d'auto-observation d'un certain système par lui-même. Cette notion de réflexivité, centrale, devra être prise en compte dans le système d'interprétation.

Tout processus d'interprétation artificiel, semblable à l'interprétation réalisée par l'esprit humain [Nutin 91], est fondé sur la représentation de la motivation de ce qui déclenche le process d'interprétation, partant de l'objet d'attention pour en produire la désignation par ce que l'on nomme, en sémiotique, le signe [Peirce 84].

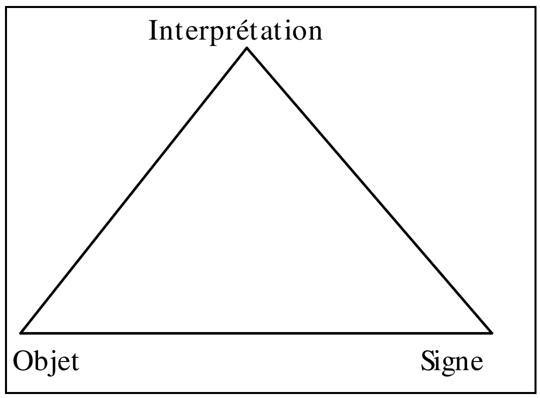

Figure 1. Le triangle sémiotique

Usuellement, en se limitant à une entité cognitive motivée, on introduit un niveau motivationnel en la munissant d'une structure particulière qui agit comme une fonction exprimant la volonté à effectuer sa tâche [MacFarland 93]. Cette fonction de volonté est quantitativement mesurable, ce qui permet de valuer l'état d'avancement de la tâche de l'entité. Nous sommes dans un cadre strictement rationnel et symbolique, où la volonté et la motivation sont de fait des fonctions mesurables. Ces fonctions s'implémentent éventuellement sous forme de variables dites "essentielles", décrivant des domaines numériques non limités mais possédant un point attractif, la valeur nulle [Ashby 60]. L'entité possède un ensemble fini de motivations bien dénotées, données a priori et un système de régulation opère des choix parmi les motivations candidates. Ces motivations sont hiérarchisées en différents besoins [Tzafestas 95]. Le besoin fondamental de l'entité est de remplir sa tâche, ses besoins secondaires étant de permettre cette réalisation. L'entité doit donc avoir un comportement qui tend vers le besoin minimal, pour la valeur nulle des variables essentielles, dans lequel l'entité est satisfaite et sa tâche accomplie.

Si l'entité cognitive n'a qu'une tâche à réaliser et que sa volonté est essentiellement de la remplir, elle a plutôt un comportement d'automate et aucun caractère de génération de motivations nouvelles ne peut lui être attribué. Mais dans le cas où cette entité doit se fixer d'elle-même des tâches non triviales selon la situation qu'elle perçoit et conçoit, elle doit opérer des activités cognitives semblables aux motivations d'un sujet humain:

1. elle doit se soucier de certains signes expressifs de son contexte,
2. elle doit les concevoir comme des représentations lisibles et compréhensibles pour lui,
3. elle doit élaborer des plans adaptés,
4. elle doit enfin agir,
5. elle doit ensuite évaluer la différence entre ses évaluations passées et les transformations opérées, pour analyser le sens de son action.

Pour ce type plus complexe de motivation, l'accent est mis sur un point qui nous paraît central : avant de résoudre un problème bien posé et de définir des buts, l'entité qui interprète une certaine chose du réel par un acte motivé doit générer un *souci* sur son environnement et sur elle-même. Il ne s'agit pas du repérage d'une forme attendue du contexte, mais du fait que l'entité doit se situer d'elle-même dans son environnement pour concevoir la présence de cette forme, en s'identifiant comme un sujet qui conçoit son placement dans son monde. En fait, nous considérons comme central dans le comportement de l'entité qui réalise l'interprétation motivée, la génération de *sa* situation propre, conçue comme un positionnement continu dans son environnement.

Le souci pour une entité capable d'interprétation

> *Le souci est la capacité déclenchante à produire une réorganisation de la structure d'une entité cognitive motivée, afin que celle-ci perçoive puis conçoive une certaine situation. Le souci peut préexister à la perception ou lui être postérieur, mais est l'engagement à la conception de toute idée.*

Pouvoir se soucier d'une chose, d'un fait, d'une idée, à n'importe quel moment, est l'action première d'une entité effectivement motivée, à partir de laquelle va pouvoir se déployer son processus de conception, de décision, d'action et d'évaluation de toute action pensée. Eluder la question du *souci* revient de fait à supprimer la raison de toute motivation, à ignorer ce qui rend nécessaire l'action de l'entité. Ne pas considérer le souci ramène le comportement de l'entité à l'élection de choix mécaniques pour la résolution de problèmes, ce qui a été et est bien étudié, mais occulte le fait que l'entité puisse agir pour elle, pour être elle-même tel un sujet.

1.7. Résumé

Les caractères principaux du process de prise de conscience que nous devons retenir pour spécifier un modèle de processus d'interprétation symbolique, puis un système générateur de pensées artificielles, sont les suivants :

1. il y a un lien, organique mais non direct, entre système pensant et contexte pensé,
2. les pensées sont considérées comme étant le produit d'une auto-organisation dans une structure complexe,
3. l'action de penser est temporellement située, donnant le sens en déval de la temporalité,
4. les formes langagières ne sont pas réduites à des symboles mais sont définitivement complexes,

5. la production d'une pensée est conduite par un souci préalable et générateur.

Seconde partie
Les systèmes adaptatifs

L'Informatique conduit toute recherche, en fin de compte, à la construction d'un système s'exécutant sur un ordinateur. Mais il y a différentes sortes de systèmes, qui ont assez peu de choses en commun, au niveau de leur modélisation, de leurs fonctionnalités et au niveau de leur architecture. Ils ne se retrouvent même pas au niveau de l'esprit du langage de programmation utilisé pour les implémenter. Nous insisterons, quant à nous, sur la typicalité des système adaptatifs, les seuls aptes à représenter un artefact de conscience.

2.1. Construire un système

En Informatique, la résolution d'un problème se ramène à la construction d'un système qui en produit la solution. Pour un animat, il s'agit de construire un système qui se comporte de façon semblable à son modèle vivant. Pour un système à boucle d'interprétation, il s'agira de concevoir un système dont le changement d'organisation fera émerger des formes qui seront comparables à des pensées générées par une conscience.

La démarche de construction d'un système est maintenant bien établie et se décompose en cinq étapes, se conformant aux procédures du Génie Logiciel :

1. écriture du problème (la définition des besoins),
2. analyse du système (conception des objets, modèle dynamique et modèle fonctionnel obtenus par décompositions en sous-systèmes hiérarchiques),
3. codage du système (la réalisation),
4. tests du fonctionnement du système,
5. livraison du système et mise en oeuvre sur site.

En fait, on passe d'une écriture non formelle d'une vue du système à réaliser (l'écriture du problème bien posé) à une écriture formelle (la représentation de l'architecture du système) puis à la programmation (la réalisation codée du système) pour aboutir à nouveau à une écriture non formelle (les commentaires sur le fonctionnement du système).

La formulation du problème permet de décider si le système à réaliser est fait pour que son observation produise des caractères intéressants, ce qui est le cas de la simulation, ou bien pour agir effectivement dans certaines situations de façon directe, ce qui est le cas des systèmes de contrôle ou de prise de décision automatiques. A partir de là, on s'orientera soit vers la construction d'un système dont la considération du comportement sera un élément décisif pour l'utilisateur, soit vers un système qui n'aura qu'à agir et réagir seul, au mieux, dans des cas prévus à l'avance. Nous nous intéressons plutôt au premier type de système, c'est-à-dire à un certain objet informatique dont le principal intérêt est la pertinence de son comportement, par rapport à ce qu'il représente.

2.2. Des modèles

Après avoir défini les caractères du problème, on peut s'intéresser à ceux du modèle servant d'architecture au système. La définition du modèle s'appuie toujours sur une théorie, c'est-à-dire sur un ensemble cohérent de concepts et de structures symboliques formant un langage de manipulation. Ce langage permet de décrire et de valider, dans sa forme propre, le modèle.

On peut considérer deux types différents de modèles : les modèles fonctionnels et les modèles comportementaux.

2.2.1. Les modèles fonctionnels

Les modèles fonctionnels se fondent sur une théorie mathématique déterministe préexistante quant au traitement du problème par un système informatique. Leur cadre d'application sont tous les problèmes bien posés, dont les domaines sont homogènes et constitués d'entités de même nature liées par des relations a priori considérées comme causales. Il s'agit alors de représenter de manière calculatoire les modalités de cette théorie pour réaliser le système. Les modèles stochastiques permettent une bonne approche numérique dans ce cadre. Leur point fort est le déterminisme clair de leur fonctionnement, mais leur point faible est leur incapacité à traiter à la fois des problèmes globaux et des problèmes locaux, c'est-à-dire les problèmes du domaine de la complexité profonde où le fonctionnement général résulte d'innombrables actions à caractère causal ou aporétique entre des composants changeants [Cardon 96].

2.2.2. Les modèles comportementaux

Les modèles comportementaux s'appliquent donc lorsque les modèles fonctionnels ne sont pas opérants. Ils nécessitent de construire une théorie de la connaissance du domaine d'opérationalité du système, son ontologie, fortement structurée, avec des différences de genres entre les différentes parties. On peut ensuite réaliser le système correspondant, avec une implémentation adaptée. Son fonctionnement permet l'étude du phénomène qu'il décrit, par simulation.

Les modèles sur lesquels repose l'architecture de tels systèmes comportementaux se fondent aujourd'hui sur les trois caractères suivants :
1. représentation fidèle de la connaissance du domaine étudié,
2. représentation précise des différences de granularité de la connaissance dans le domaine étudié,
3. représentation précise du comportement des entités de base du système comme ensemble conduisant, par la mise en action globale, à la solution du problème.

Le premier point précise que le modèle doit contenir une approche ontologique approfondie. Le second point pose la question de l'agrégation des composants élémentaires en entités plus vastes et ayant des fonctionnalités propres. Lorsque l'agrégation des entités est prévue au départ, le comportement de l'agrégat est bien pris en compte et il reste seulement à gérer, ce qui n'est pas si simple, un problème à structure non homogène, d'entités et de groupes, où les lois comportementales sont de natures différentes.

Nous devons, pour le système à boucle d'interprétation, nous intéresser au comportement d'un système constitué d'entités plus ou moins indépendantes, avec des entités de base et des entités agrégées, changeantes dans leurs formes et leurs fonctions et observer sous quelles conditions, dans l'organisation des composants, le comportement général satisfait à tel ou tel caractère, c'est-à-dire fait émerger telle ou telle forme structurelle. C'est un problème typique de l'Intelligence Artificielle Distribuée.

2.3. Trois types de systèmes

Les agents et les systèmes multiagents, sont les objets typiques de l'Intelligence Artificielle Distribuée (IAD). Dans la vaste thématique de l'IAD, on considère des problèmes de nature sociale, tel un ensemble d'insectes artificiels ou un ensemble d'animats jouant à un jeu collectif, en vue de réaliser des systèmes informatiques reproduisant les caractères comportementaux des systèmes sociaux réels. Il s'agit de simuler le comportement global d'un certain ensemble social, assez bien défini, comportement résultant essentiellement des interactions des comportements individuels des entités de base [Gasser 89]. Il y a encore deux tendances, opposées, dans les problématiques de l'IAD : ou bien on s'intéresse à la modélisation de systèmes sociaux dont les entités de bases sont considérées comme intelligentes ou bien à des systèmes sociaux dont les entités de base ne le sont manifestement pas. La première approche est dite cognitive et la seconde réactive.

Dans le cadre de l'IAD, les systèmes qui résultent des modèles fonctionnels et comportementaux sont de trois types, bien différents.

2.3.1. Les systèmes réactifs

Dans ce type de situation, tout événement de l'environnement du système est considéré comme un stimulus qui, instantanément, met en action le système, avec une action correspondant aux caractères du stimulus. Il y a un lien causal entre événement extérieur et action, qui sont des éléments de même nature, le lien étant en fin de compte un appel procédural (une sorte de "callback"). Le système n'est pas complètement distinct de son environnement où il prend ses données. Il est construit et programmé pour répondre à une classe plus ou moins restreinte d'événements par des activations automatiques de procédures. Le problème central est alors le contrôle et l'efficacité de l'action événement - sous-programme correspondant.

2.3.2. Les systèmes perceptifs

Ici, tout événement de l'environnement du système est considéré comme un fait complexe pour celui-ci. Il y a symbolisation de l'événement, représentation par un objet propre, clairement défini mais de structure complexe et la question est de le reconnaître selon ses modalités. La reconnaissance entraîne la réaction du système au fait perçu dans l'environnement. Le système distingue le monde extérieur en le symbolisant par de nombreux caractères propres, éventuellement précisés par une ontologie. Tout événement produit par l'environnement est recherché pour fusionner symboliquement avec les structures connues correspondantes. Le système est construit avec un module médiateur entre environnement et sous-système d'action. Mais il reste un résolveur de problèmes à entrées symboliques. Le problème central est la fusion de donnée, réalisant la boucle : événement - médiateur - symboles reconnus - structuration de ces symboles - événement.

Ces deux systèmes résolvent des problèmes déjà posés et ils sont typiquement construits pour cela. Leur éventuelle propriété d'apprentissage est du type renforcement de comportement en vue améliorer leur performance, celle-ci étant la mesure de l'écart entre événement et reconnaissance de l'environnement pour action.

Il existe un troisième type de système, de nature totalement différente. Un tel système n'est plus fait pour résoudre un problème posé a priori, mais pour expérimenter *ses* propres possibilités d'auto-organisation interne sur des problèmes qu'il reformule. Le système, s'il copie le comportement d'un individu, doit avoir la même *raison* de se comporter qu'aurait l'individu qu'il copie. Il y a entre ce type de système et les deux précédents, une différence de genre assez radicale.

2.3.3. Les systèmes adaptatifs

Dans ce cas, le système est actif pour lui-même et sur lui-même, avant que de l'être pour l'environnement. Il a une structure d'action et également une organisation lui permettant de donner sens au concept de "soi". Il est une entité *close* par rapport à son environnement, au sens où il a une nature et une identité distinctives. Il a la capacité de construire, selon des buts qui lui sont propres, une représentation de l'environnement, qui en est une interprétation. Il peut dégrader son fonctionnement en système perceptif, en altérant son organisation (cas dit de bestialité ou de barbarie). Le problème central est, quant à un tel système, la représentation de la notion de soi.

Définition

> *Un système adaptatif est un système clos, ayant une permanence par rapport à son environnement, se faisant une représentation interprétée de celui-ci et agissant avec la capacité d'expression fournie par cette représentation.*

Le système à boucle d'interprétation, artefact du process de prise de conscience, est typiquement un système adaptatif. C'est par l'existence de ce système producteur de pensées artificielles que des animats pourront se comporter de manière sociale raisonnée et non plus automatique. La construction d'un tel système est donc incontournable pour la réalisation de tout ensemble d'entités artificielles qui simuleraient un groupe social d'individus réellement autonomes. Nous pouvons donc considérer que le cerveau, en tant que système producteur de représentations mentales, c'est-à-dire en tant qu'esprit producteur de formes idéelles, est un système adaptatif qu'il faut s'efforcer de construire. On pourra ensuite s'intéresser à des ensembles d'animats adaptatifs, munis chacun de consciences artificielles qui, d'une certaine façon, pourront penser à générer des organisations sociales, selon leurs volontés.

Notons que les trois systèmes que nous avons présentés forment une hiérarchie : un système adaptatif contient effectivement un ou des sous-systèmes délimités qui ne sont que perceptifs, chaque système perceptif contenant au moins une composante qui n'est que réactive. Cela correspond à l'évolution du vivant et sera exprimé dans les systèmes adaptatifs artificiels pour qu'ils soient opérationnels.

2.4. La génération du sens dans les systèmes adaptatifs

Un système adaptatif à boucle d'interprétation est une entité distinguée de son environnement, présentant un caractère d'organisation particulier qui en fait l'identifie et qui est définissable formellement de façon suffisante. Le système n'est tel qu'il est à un instant donné que par ce qui le décrit, c'est à dire par ses parties, ses composantes, les relations entre celles-ci, les relations entre les relations, et par son action effective réfléchie sur l'environnement et sa réaction réfléchie aux actions de l'environnement sur lui-même.

Par ces actions réfléchies sur l'environnement, le système prend place dans la durée. Ce placement dans la durée, avec la possibilité de spatialiser et de modeler les choses de l'environnement, va lui permettre de tirer les conséquences prévisibles ou inéluctables de ses actions dans diverses situations, d'affirmer des contraintes induites, de poser de nouvelles hypothèses et de définir de nouvelles possibilités de situation ou d'action.

2.4.1. Complexité organisationnelle des systèmes adaptatifs

Nous dirons qu'un système est considéré comme *organisationnellement complexe* si la combinatoire de ses parties est importante au regard de son enveloppe externe (de sa surface membranaire ou de son interface multimodale selon les cas) et si son fonctionnement revient à opérer des changements de partition dans sa structure interne [Cardon 96]. Un tel système a ainsi une complexité organisationnelle correspondant au nombre de ses parties possibles, qui est évidemment exponentiel devant la dimension de l'enveloppe externe. La cohérence d'un tel système réside d'une part dans sa structure, à partir des relations entre les entités qui le composent et d'autre part dans son immersion spécifique dans le contexte dans lequel il existe et se comporte.

Un système adaptatif sera basé sur une architecture organisationnellement complexe. Il sera évidemment plus que cette architecture, par les caractères fonctionnels de ses composants élémentaires et par sa capacité à se réorganiser.

2.4.2. Clôture opérationnelle des systèmes adaptatifs

Dans les systèmes adaptatifs, organisationnellement complexes, l'échange entre ceux-ci et l'environnement se réalise par adaptation réciproque, sans que l'échange direct d'informations soit suffisant pour réaliser les adaptations mutuelles. On nomme ce processus "clôture opérationnelle des systèmes autopoiétiques" [Varela 89]. Cette notion a été appliquée par F. Varela aux systèmes vivants individuels fermés par des membranes et plus ou moins étendu aux groupes de systèmes vivants élémentaires.

La clôture opérationnelle doit être entendue, en ce qui concerne les systèmes adaptatifs à boucle d'interprétation, radicalement. Le système ne reçoit pas de l'information sous la forme d'une entité objective, quantitative, assurant une communication réifiée avec l'extérieur et qu'il intégrerait plus ou moins directement pour se l'approprier, mais se reconforme selon l'environnement qu'il perçoit globalement, d'une certaine façon qui est celle que lui permet la complexité de sa structure et ses tendances à la modification. Le caractère de clôture conduit à considérer le système comme *étant* une représentation propre et dynamique de ce que l'on appellera son contexte et lui-même mêlés, par opposition à un système réactif qui *contient,* entre autres choses, une certaine représentation de ce contexte sous forme de structures informationnelles, les classiques Systèmes à Base de Connaissances (SBC).

Le système opérant par clôture opérationnelle réalise une adaptation à lui-même et au milieu dans lequel il est placé. Il est une composante non indépendante de ce milieu, dont il exprime d'ailleurs une représentation et qui lui permet d'agir. La clôture opérationnelle traduisant pour le système le processus d'adaptabilité est donc le moyen d'exprimer son contexte et lui-même, en évitant d'établir une distinction de genre entre un ensemble et un autre, qui seraient disjoints et mis en relation par le seul transfert physique d'entités informationnelles.

Le caractère fonctionnel des systèmes réactifs lie leur environnement à leurs domaines de représentation, qui est une image de l'environnement, en les associant fonctionnellement par l'information qui passe de l'un à l'autre. Il n'est plus question de cette liaison fonctionnelle avec les systèmes adaptatifs. Celui-ci comprend une certaine représentation de son environnement et l'environnement a, éventuellement, en tant que système adaptatif, une perception et une représentation de ce dernier. Un système adaptatif est donc morphologiquement *correspondant* à son contexte [Morin 86], puisqu'il le fait apparaître tel il est et qu'il participe, en étant physiquement et temporellement dans son contexte, à sa constitution. Cette propriété a été présentée par E. Morin [op cité] sous le nom de principe hologrammatique.

2.4.3. Principe de réciprocité

En ce qui concerne les systèmes adaptatifs, leur propriété d'opérer leur action de reconformation structurelle par clôture opérationnelle nécessite de définir un principe d'existence réciproque avec leur contexte et entre systèmes, permettant l'adaptation réciproque. Sans ce principe, on serait conduit à considérer des transferts effectifs, des échanges directs d'informations entre systèmes et contextes, ce que nous ne retenons pas, où bien alors à envisager la séparation absolue et l'incommunicabilité, sauf par le hasard. Ce *principe de réciprocité*, d'adaptation engagée, est consécutif de la notion de clôture. Il détermine le sens naturellement concordant de l'évolution comportementale d'un système adaptatif dans son contexte également adaptatif, le sens communautaire de toutes les structures sociales considérées dans l'IAD, les tendances agrégatives de tous les groupes d'individus vivants.

Ne souhaitant donc pas disjoindre le système adaptatif de son contexte, pour le lier ensuite à celui-ci par la constitution d'un lien explicite, il nous faut considérer que le système est immergé dans son contexte, sous l'action du principe de réciprocité. Il est donc une expression au moins mimétique de celui-ci, en l'exprimant dynamiquement.

La réflexivité se place en ce point. Le système prend sens en existant par réciprocité avec son contexte :

- le contexte prend sens par l'action du système et des sous-systèmes qui le composent,
- le système prend sens par auto-adaptation organisationnelle par rapport à son contexte.

2.4.4. La génération du sens

Nous pouvons alors définir ce que nous appelons le sens généré par un système adaptatif organisationnellement complexe.

Définition du sens généré

> *Le sens généré par un système adaptatif complexe est le produit autoréférent de la reconformation de sa structure d'interprétation qui représente, pour lui, son contexte et lui-même situé dans son contexte.*

Le sens généré, que l'on appellera aussi sens émergent, est une partie localement stable de la structure plastique du système qui s'est reconformée. La raison de cette reconformation sera évidemment placée dans la structure elle-même et non dans un sous-système de contrôle détaché. Ce sera ce que nous avons appelé le souci. La réflexivité du sens est la propriété qu'une fois généré sous la forme d'une certaine sous-structure non indépendante, celui-ci est un produit déclenchant des adaptations suivantes : le système adaptatif est le prisonnier de sa production, de sa mémoire, de son existence. Le sens généré atteste le contexte qu'il perçoit, ainsi que l'organisation du système lui-même représenté dans son contexte. On ne peut pas en effet concevoir un système adaptatif qui serait capable de percevoir objectivement un contexte, hors de lui-même se le représentant : le système adaptatif a une ipséité irréductible.

Ainsi, dans cette approche, nous considérons que le système dans son ensemble exprime au moins mimétiquement son contexte et a, de plus, la capacité à s'exprimer sur lui-même situé dans cette représentation du contexte. Cette réflexivité est le fondement de l'autoréférence et sera plus tard, en s'appuyant sur des formes langagières, la connaissance de soi, la connaissance de la connaissance [Morin 86] et constituera bien l'émergence de la notion de *soi* dans un système à boucle d'interprétation [Weyhrauch 95]. L'ensemble de ces propriétés permet de définir la finalisation nécessaire d'un système adaptatif considéré comme complexe et opérationnellement clos dans son environnement.

2.5. La propriété d'émergence

Ainsi, les systèmes adaptatifs opèrent sur eux-mêmes et envers l'environnement par clôture opérationnelle. Ils sont composés des entités d'un certain substrat physique combinatoirement complexe. Ils développent des modifications structurelles génératrices de sens en réorganisant ce substrat plastique. Ce substrat est représentable par un certain espace de formes changeantes, interactives, structurées dynamiquement. La raison de ce fonctionnement par reconformation ne peut résider que dans la structure elle-même : la structure du système fait simplement opérer une action de reconformation en réaction à une autre possible, de tendance localement opposée. Il faudra évidemment préciser dans l'architecture où résident et comment s'expriment ces tendances opposées. Sans ce principe d'opposition élémentaire, nous serions conduits à introduire dans le système un processeur d'action et à suivre le schéma très fonctionnel *d'une* machine distinguant définitivement instructions et données.

Les systèmes adaptatifs fonctionnent en concevant d'abord leur situation, c'est-à-dire leur placement dans leur contexte, en reconformant par clôture opérationnelle la morphologie de leur espace de formes pour générer du sens. Ensuite, ils agissent, évaluent le nouveau contexte et eux-mêmes, puis recommencent. Ce fonctionnement par transformations structurelles continues engendre des conformations nouvelles multiples, qui se modifient par la suite, communiquent et opèrent entre elles.

Ces systèmes engendrent également, à partir des formes locales dans leur structure, des formes plus globales qui sont les expressions, dans l'instant, de la structure du système entier. Elles *résument* sa transformation. Ces formes qui expriment la globalité de l'organisation saturent en quelque sorte le processus de reconformation et expriment un état du système considéré comme significatif. Ces formes expressives constituent, par définition, en ce qui concerne les systèmes adaptatifs et en particulier le système à boucle d'interprétation, une *émergence*.

Définition de l'émergence

> *Une émergence dans l'organisation d'un système adaptatif est une stabilisation temporaire de cette organisation, considérée comme un caractère typique de sa morphologie. Elle est une appréciation du niveau de la restructuration, dans la durée et l'espace, correspondant, par ces deux catégories, à ce qu'elle exprime.*

Communément, l'état d'un système classique est celui du ou des processus qui assurent son fonctionnement. Un tel système exécute une suite d'actions bien identifiées. Partant d'un état initial, le processus produit à un certain moment une action considérée comme terminale, qui exhibe un résultat : l'état final, temporaire ou définitif. Pour cela le système passe par différents états intermédiaires en ce qui le concerne et en ce qui concerne ses sous-systèmes, en modifiant certains de ses caractères (en fait des variables). En ce qui concerne les systèmes adaptatifs complexes, nous considérons que le processus de reconformation est ininterrompu et clos sur lui-même. Il s'agit d'une transformation non interruptible de la morphologie interne et qui existe tant qu'existe le système. Ce processus génère à certains moments des formes qui résument le sens de la transformation, stables un moment et qui sont l'expression du sens émergent. La forme émergente n'est pas un état final atteint par un processus fonctionnel, mais l'expression, à un certain moment, de la cohérence du processus de transformation lui-même, qui se *résume* et se *condense* dans une certaine forme constituant l'état émergent.

Ainsi, il n'y a pas, dans cette approche, de séparation de genre entre des données qui seraient engrammées (la mémorisation de faits dans des structures typiques des SBC) et un processus d'extraction de formes plus ou moins contrôlé par un moteur de recherche (le mécanisme de raisonnement ou de génération contrôlée d'une bonne forme). Il y a un processus réflexif de transformations continues, comparable à un programme constitué de formes dont le principe est de se réécrire par transformations structurelles selon une double raison : l'opposition et la concordance entre les formes. Tout processus se développe pour en inhiber certains et en révéler d'autres considérées comme proches. A certains moments, le programme génère une conformation prégnante pour son organisation et l'exprime, d'une certaine façon, par une forme émergente parce que d'autres formes sont mises en sommeil. C'est *l'émergence du sens* dans les systèmes adaptatifs, opérée par inhibitions locales. La raison du déploiement de ce processus est et n'est que l'auto-adaptation du système *à lui-même* et à l'environnement perçu, en opérant par compensations et inhibitions.

Cette description du fonctionnement d'un système adaptatif complexe générant pour lui-même du sens, est donc radicalement différente de la recherche du "sens" d'un mot obtenu en utilisant un certain algorithme dans une base de connaissances symboliques. Elle pose que ce que le système génère, c'est-à-dire en fait le sens exprimé dans l'émergence d'une organisation, est obtenu par la considération d'une expression de la globalité de la transformation morphologique dans un intervalle de temps. Cette approche est, selon nous, la seule façon de prendre en compte la complexité d'un système adaptatif, qui sinon serait éludée. La considération objective des états émergents, en les détachant de leur caractère de processus générateur de reconformations structurelles, en les rendant atemporels et objectivés, constitue le cadre de la modélisation classique, fondée sur l'usage systématique des formes symboliques et sémantiques détachées de leurs traitements et ainsi sans rapport avec la notion d'adaptabilité.

Les formes émergentes, qui expriment le sens de la transformation de la structure du système, sont des expressions à la fois de celui-ci, du contexte et de la relation de clôture opérationnelle entre les deux. Elles sont des représentations, à leur manière, du contexte dans lequel elles sont créées, mais aussi et toujours du système lui-même, avec ses tendances et ses raisons à agir.

Troisième partie
Un modèle de système à boucle d'interprétation

Nous décrivons maintenant un modèle calculable du système à boucle d'interprétation. Ce modèle représentera, en les transposant au niveau de certaines fonctions, les grands caractères d'un système construisant une "émergence du sens" à partir de son organisation. Le rôle de cette description formelle est d'étudier la transposition calculable de certains caractères d'une conscience artificielle.

Une machine de Turing décide de la calculabilité d'une fonction. Son programme, ses états et sa donnée, disjoints, permettent de décider du caractère calculable ou non de la fonction représentée a priori et également de sa complexité en pas de programme. Nous allons utiliser des organisations de fonctions de base, dites *fonctions plastiques* de par leur propriété de pouvoir réécrire leurs programmes par parties. Le choix d'utiliser des fonctions, plutôt que des agents plus spécifiques et sophistiqués, est dicté par le souci de présenter essentiellement le caractère calculable du système, en utilisant des notions simples et non contestables : celles de fonction et de processus.

Ces fonctions seront considérées comme des processus concurrents. Elles seront composées d'un programme, c'est-à-dire d'une suite d'instructions et d'un ensemble d'états qui, comme pour tout processus, caractérise leur situation. Leur états sera soit suspendu, en communication, en action, en attente, en synchronisation ... et correspondra à l'état de tout processus dans un système distribué asynchrone.

Une fonction plastique correspond à un certain caractère fonctionnel usuel du système nerveux, neuronal et humoral, ainsi des aires cognitives, comme par exemple le rôle de celles manipulant les nombreux éléments du langage. Ces fonctions représentent également des caractères subjectifs comme les impressions, les désirs et les craintes. Elles représentent aussi les tensions, les forces opérant sur des sous-ensembles de fonctions. Il s'agit donc d'un ensemble très vaste de fonctions, chacune ayant la connaissance de certaines autres, pour l'activer ou l'interrompre. L'ensemble de ces fonctions est structuré, par morceaux, pour former une architecture distribuée de fonctions partiellement concurrentes. La précision de cet ensemble structuré sera réalisé lors de la définition ontologique du système.

Nous allons nous intéresser principalement à certaines propriétés générales de cette organisation de fonctions, celles réglant la question du contrôle par réorganisation et celle de l'intentionnalité artificielle. Nous décrirons par la suite une interprétation de ce modèle, en vue de son implémentation, en l'architecturant autour d'organisations d'agents hybrides et dynamiques.

3.1. Le modèle calculable : un univers dynamique de fonctions plastiques

Nous décrivons un modèle calculable de système à boucle d'interprétation, en retenant les particularités du process de prise de conscience, comme un système typiquement adaptatif.

La composition du modèle sera la suivante :

1. un espace d'action, composé de fonctions d'interface. Il est le correspondant formel du corps physique de l'individu, avec ses moyens d'action physiques,
2. une organisation de fonctions concurrentes agissant sur une mémoire partagée, opérant comme des processus, qui sera le domaine où se réalise l'équilibre homéostatique du système, ses réponses envers environnement, la représentation, la conception de l'environnement et de l'organisation elle-même. C'est ce qui correspond au système nerveux central dans son entier,
3. un état d'activité qui permettra de distinguer quelles sont les parties actives de l'organisation (les ensembles de fonctions actives de l'organisation) et de l'espace d'action (les fonctions actives de cet espace). Cet état représente les parties activées selon le souci du moment. Il se représente comme une tension générale qui correspond aux aires cognitives actives, à un instant considéré du temps.

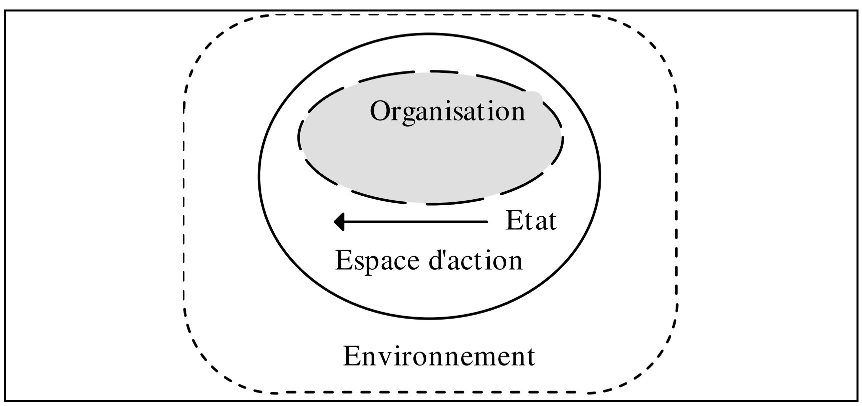

Figure 2. Schéma du modèle calculable d'un système adaptatif.

3.2. La structure fonctionnelle du modèle

L'espace d'action est constitué d'une mémoire dite *mémoire d'interface,* qui est le correspondant de la membrane sensible dans le vivant et d'un ensemble de fonctions opérant mécaniquement sur cette mémoire. L'environnement place des informations dans la mémoire d'interface, les fonctions les prennent éventuellement en considération et alors les traitent systématiquement.

3.2.1. Les fonctions plastiques

L'organisation est constituée d'une structure de fonctions dites *fonctions élémentaires*. Elle comprend également une mémoire partagée adressable appelée *mémoire partagée interne*, dans laquelle figurent, initialement et à chaque pas de fonctionnement certaines données. Cette mémoire traversera également l'espace d'action du système. Nous ferons l'hypothèse qu'elle n'est pas accessible à partir de l'environnement, pour respecter le caractère de clôture du système.

Toutes les fonctions, de l'espace d'action ou de l'organisation, sont des fonctions élémentaires dites fonctions plastiques.

Définition

Une fonction plastique est un processus d'action, constitué d'un programme composé d'instructions et ayant à chaque instant un état défini dans un ensemble d'états. Elles forment un ensemble de processus asynchrones concurrents. Toute fonction plastique réalise ses entrées-sorties dans une mémoire partagée. Les fonctions plastiques élémentaires constituent le type de base des fonctions plastiques de l'organisation.

Définition

La mémoire partagée interne du système est une mémoire adressable accessible en lecture-écriture par les fonctions élémentaires. Les mots-mémoire contiennent une information symbolique quelconque, différente de l'information de la mémoire d'interface.

Définition

La mémoire d'interface est une mémoire accessible en écriture par l'environnement du système et accessible en lecture par les fonctions d'interface.

Définition

Les fonctions d'interface sont des fonctions plastiques dont les entrées sont prises dans une mémoire d'interface communiquant avec l'environnement et dont les sorties s'effectuent dans la mémoire interne partagée.

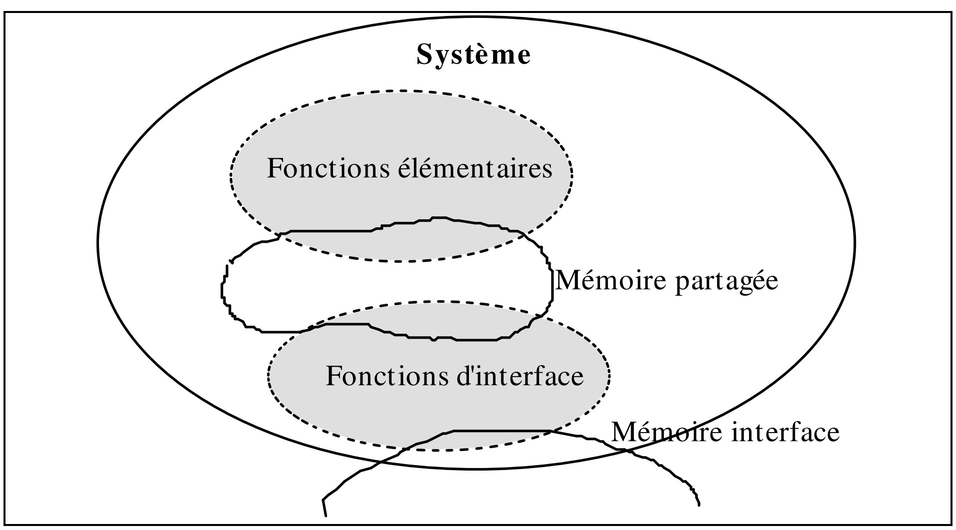

Figure 3. Les fonctions plastiques du niveau fonctionnel.

Les fonctions élémentaires vont en fait représenter ce qui peut être actif dans le cerveau et le système nerveux central : actions de propagation de l'influx nerveux dans le réseau neuronal, actions de construction et transmission d'hormones aux bons moments aux bons endroits. Elles sont spécifiques des aires cognitives. La mémoire interne correspond à la *trace structurelle* des actions. C'est la représentation de la structure du maillage neuronal *et* de la capacité à générer hormones et peptides et à les propager le long de certaines voies. Le modèle proposé insiste donc sur le caractère plastique de l'organisation du système nerveux en traduisant les actions de reconformation par un type d'entité : les fonctions plastiques et le résultat de ces actions dans une entité distincte : la mémoire partagée.

Un certain ensemble, de cardinal élevé, de fonctions plastiques, définies par leurs états et leurs programmes, sont initialement en place dans l'organisation et dans l'espace d'action, pour agir sur la mémoire partagée à des emplacements initialement distincts. Ces fonctions traduisent toutes les représentations élémentaires possibles des caractères du domaine.

Des fonctions plastiques particulières, prenant des informations de l'extérieur, transcrivent celles-ci en données dans la mémoire interne. En modifiant cette mémoire, elles concourent à définir ou modifier l'état d'action du système entier. Des modifications externes entraînent donc indirectement les activations de certaines fonctions plastiques.

La constructibilité d'un tel système sera basée sur l'hypothèse que l'abstraction de chaque fonction élémentaire et de chaque symbole logé en mémoire compense la complexité combinatoire de la plasticité des composants de base : les instructions et les notations binaires des informations manipulées. Cela n'engage en rien le caractère calculable du système, mais seulement sa réalisation informatique effective.

3.2.2. Le rôle des fonctions plastiques

Les fonctions élémentaires définissent tous les réflexes physiques, les comportements innés des fonctions vitales et aussi toutes les tendances à des actions cognitives de base comme la formation d'un mot, sa recherche, la reconnaissance de son identifiant un concept lexical précis, l'expression d'un trait de reconnaissance d'une forme dans une image ou d'un son dans une oeuvre musicale ... Ces sont en fait toutes les fonctions minimales, indépendantes, assurant le comportement d'abord *réactif* du système envers lui-même et envers son environnement usuel. Elles permettront de définir l'ontologie fonctionnelle du process de prise de conscience, par des éléments *indépendants et générateurs* de l'organisation.

La question est le passage du niveau élémentaire de la biologie du cerveau, les neurones et les amas de neurones (les aires), au niveau cognitif, où les entités sont conceptuelles et expliquent des processus psychologiques. Les fonctions plastiques sont typiquement entre ces deux niveaux. Elles sont plus abstraites que les formes neuronales et hormonales, mais sont des génératrices de concepts, traduisant un caractère d'insuffisance. Elles se placent ainsi comme des potentialités de réification de formes du niveau psychologique. En ce sens, elles sont des *processus générateurs plastiques*. Nous avions déjà défini de telles entités sous la forme de "pavages" dans des espaces de formes [Cardon 96].

Certaines de ces fonctions opèrent comme des fonctionnelles, mais prennent toujours leurs données sur la mémoire partagée et sont initialement indépendantes : c'est l'hypothèse fonctionnelle de constructibilité.

Définition

L'hypothèse fonctionnelle de constructibilité du système pose que le modèle peut être initialement défini par un ensemble de fonctions plastiques indépendantes, qui opèrent comme des générateurs d'associations, pour pouvoir ainsi agir globalement au niveau de la réorganisation de cet ensemble de fonctions.

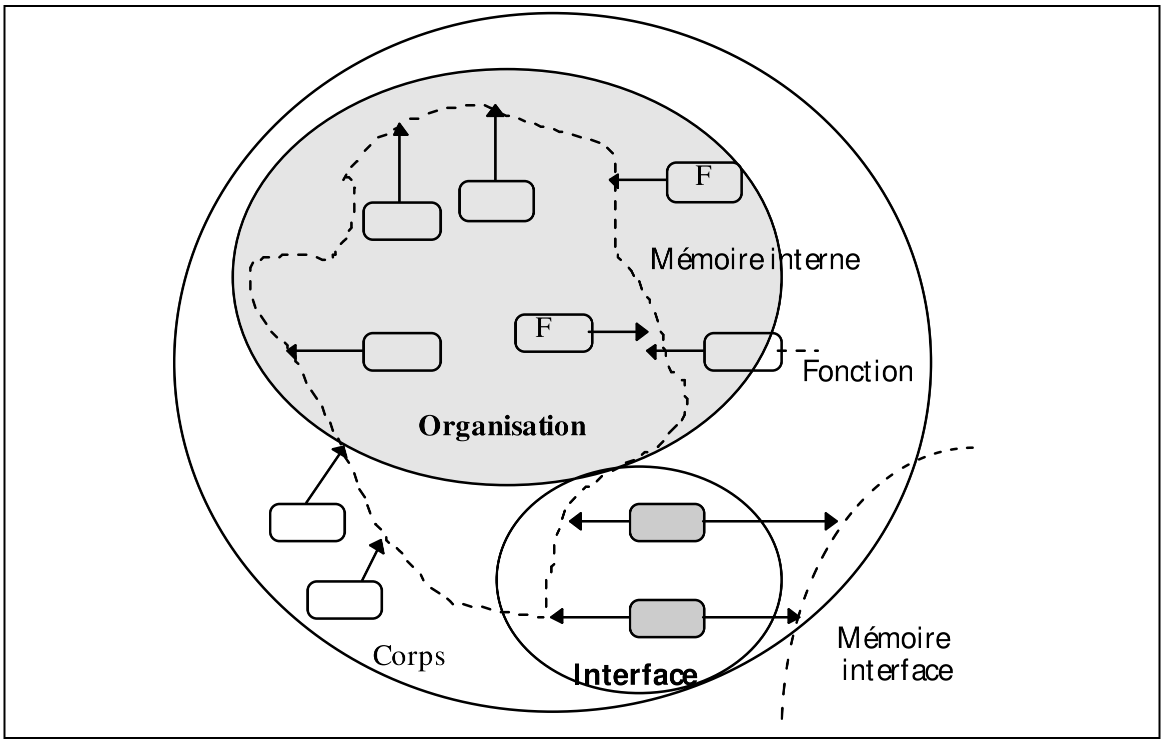

Figure 4. L'architecture du modèle calculable.

Certaines fonctions élémentaires, qui ont des programmes simples à déroulement en boucle infinie (celles qui correspondent aux fonctions vitales d'un système vivant par exemple), sont initialement actives dans le système, en état de veille permanente. Il s'agit des fonctions interfaces et de certaines fonctions élémentaires. Il suffit que leurs programmes réalisent des actions répétitives avec écriture limitée dans la mémoire interne et dans la mémoire interface pour que le système fonctionne ainsi. Ces fonctions agissent, dans ce cas, de manière indépendante : le programme de l'une ne peut altérer le programme d'une autre.

Nous avons donc, à ce niveau de la description, un modèle compliqué, puisque le niveau ontologique de description des fonctions élémentaires est compliqué, mais typiquement fonctionnel, composé d'unités indépendantes réalisant leurs calculs de manière asynchrone et gérant, par cela, le comportement organisationnel du système. Celui-ci aura toujours certaines fonctions en activité (celles correspondant à l'équilibre homéostatique du corps) et est de plus altéré par la prise d'informations sur la mémoire interface, qu'il ne contrôle pas.

3.3. La collision : le premier niveau de structuration de l'organisation

L'action première du fonctionnement du système est de remplir une partie significative de la mémoire interne, c'est-à-dire de mémoriser ses possibilités de restructuration. Pour cela, les programmes des fonctions plastiques actives s'exécutent. Il y a alors deux cas possibles de fonctionnement.

3.3.1. Le fonctionnement avec des fonctions élémentaires indépendantes

Les programmes des fonctions plastiques manipulent des parties différentes de la mémoire interne, aux mêmes instants du temps. Si certaines parties ont été modifiées par une fonction, une autre peut agir sans problème particulier sur cette partie, si cette action se place après dans le temps. Ainsi, les sorties de certaines fonctions plastiques sont les entrées de certaines autres, mais de manière séquentielle. A ce niveau, l'ensemble des fonctions opère par composition, ce qui est classique et bien résolu en Informatique.

3.3.2. Les collisions simples et multiples entre des fonctions élémentaires

A un même moment du temps, les instructions d'écriture de deux fonctions élémentaires au moins se trouvent sur un même mot de la mémoire interne. Il y a collision de ces fonctions. Ce problème de collision peut se résoudre, à ce niveau, en synchronisant les processus des fonctions, de quatre façons différentes :

1. les fonctions opèrent une *fusion* de certaines instructions de leurs programmes, c'est-à-dire une union des instructions et des états, avec mise en conformité des symboles des états. Il y a réduction des fonctions en collision en une seule fonction. Le programme de la fonction résultante devient alors, s'il ne l'était pas, typiquement non déterministe. L'état de la mémoire en collision est le déploiement des états possibles (comme pour les machines de Turing non déterministe). Dans ce cas, les fonctions plastiques initiales se généralisent.
2. les fonctions réalisent une réduction déterministe de leurs programmes et états. Elles se transforment, en supprimant dans toutes les fonctions en collision sauf dans une, l'instruction qui permet d'agir dans l'état courant sur le mot mémoire courant. L'état suivant du mot mémoire de collision est donc l'état affecté par la fonctions retenue. Dans ce cas, les fonctions se spécialisent.
3. les fonctions réalisent une augmentation déterministe de leurs programmes et états. Elles se transforment, en modifiant certaines instructions qui ont conduit à la collision, devenant ainsi non déterministes. Nous nous retrouvons dans le premier cas, mais avec des fonctions augmentées. Cela revient à une augmentation des programmes et donc de la robustesse du système, par déploiement de sa généralisation.
4. les fonctions ignorent la collision et poursuivent leurs instructions, en plaçant dans le mot mémoire de collision une valeur provenant de l'action de l'une quelconque des fonctions. On utilise ici une propriété des systèmes parallèles à mémoire partagée qui prend en

compte l'unicité de l'état de tout mot mémoire, à chaque instant. Le système est raisonnablement rationnel sur ce point, en utilisant le hasard pour se rationaliser.

Mais il est également possible que certaines fonctions plastiques, qui opèrent simultanément à différents endroits de la mémoire partagée, se trouvent dans des collisions dites *multiples* : plusieurs mots mémoire sont le cadre de collisions multiples et simultanées de plusieurs fonctions élémentaires. Dans ce cas, les fonctions réalisent des augmentations non déterministes de leurs programmes et états, en se regroupant par le fait de collision.

Des groupes de fonctions élémentaires se forment pour traiter les collisions multiples, pendant leurs constitutions, en formant ainsi une organisation dynamique locale dans le système. La formation de ces groupes, avec les fonctions les constituant pouvant prendre en compte leur durée de constitution, va avoir une importance capitale dans la perception du temps et de l'espace par le système.

Principe de temporalité

La résolution spatiale et temporelle des collisions multiples, c'est-à-dire l'émergence au cours du temps de groupes stabilisés nouveaux de fonctions élémentaires qui ont été en communications synchrones par les collisions multiples, engendre la conceptualisation du temps et de l'espace pour le système. Elle engendre également la conceptualisation de l'organisation de la structure des fonctions.

Remarquons que le système peut contenir a priori des fonctions élémentaires dénotant des caractères spatiaux et temporels, données donc à la construction du système, mais que la résolution des collisions multiples entraîne pour le système une appréhension propre du concept de temps, qui correspond au déroulement temporel de la résolution de ses collisions. Le concept de *temps perçu* provient donc de la durée de constitution de groupes de fonctions plastiques référant à la notion de temps externe. De la même façon, le système fait émerger effectivement par son fonctionnement les concepts d'espace et d'organisation de ses structures. Nous rappelons qu'une structure est un agencement spatial alors qu'une organisation en est le déploiement temporel.

Cette description du fonctionnement du système prenant en compte les collisions revient à le transformer structurellement par des altérations successives, en augmentant ou unifiant les programmes de certaines fonctions, mais en gardant toujours le caractère de localité, d'indépendance structurelle des fonctions élémentaires. Ces fonctions calculent, chacune pour son compte, ce qu'elles peuvent et doivent calculer, selon leurs programmes et leurs données.

Le cas de la collision permet de doter le système d'une capacité nécessaire de réorganisation. Le système transforme sa structure, en la faisant passer d'un état initial bien précisé, composé d'unités fonctionnelles séparées (les fonctions élémentaires indépendantes) en système auto-organisateur où les fonctions communiquent *pendant* leur fonctionnement. La propriété considérée pour lever la collision est le fait que les fonctions peuvent réécrire des parties de leur propre programme ou changer certains de leurs propres états, selon le contexte de leur action, ici le fait de collision. L'entité minimale de l'organisation du système est donc l'instruction dans les programmes des fonctions élémentaires, mais il n'est pas considéré comme élément minimal de l'organisation, qui est la fonction : l'approche du problème est cognitiviste. Le niveau de l'instruction correspond au niveau biologique, celui du programme au générateur du niveau psychologique.

Le but du système n'est pas de calculer simplement des fonctions bien définies a priori, aussi fines et nombreuses soient-elles, mais de poser des problèmes de calculs nouveaux, que l'on modélise par des fonctions élémentaires altérées et qui seront définis comme effectivement calculables, quel que soit le résultat des calculs. Le système est donc un *constructeur* souple de fonctions plastiques, par altération de leurs programmes initiaux. Le processus de fonctionnement de l'organisation revient à définir des fonctions nouvelles, par altération d'anciennes, qui modifient à la fois la mémoire interne et elles-mêmes.

Remarquons que la collision, en changeant les valeurs des données et la structure de certaines fonctions, peut rendre l'action d'une fonction contradictoire avec l'action d'une autre et conduire à un interblocage. Cette situation, arrêtant la réorganisation du système, sera dite *déstructurante*. C'est l'introduction de la notion de rationalité limitée dans le modèle. Mais nous allons doter le système d'un moyen d'éviter, ou de réduire, la déstructuration, en conduisant en quelque sorte sa réorganisation.

3.4. Les fonctions plastiques du méta-niveau

Nous définissons un méta-niveau sur le système précédent. Le système initial, dit système fonctionnel, sera augmenté d'un système dit de *méta-niveau*, qui va en quelque sorte l'organiser temporellement.

Définition

> *Le système de méta-niveau est constitué d'un ensemble de fonctions plastiques, les fonctions méta, qui organisent en groupe l'action des fonctions élémentaires. Leurs entrées sont à la fois les états des fonctions élémentaires et celui de la mémoire partagée et leurs sorties sont dans la mémoire interne partagée du niveau fonctionnel.*

Le méta-niveau contrôle le niveau fonctionnel par sa capacité à prendre en compte les entrées-sorties des fonctions élémentaires pendant le fonctionnement de ces dernières. Il prend en compte la modification structurelle de *groupes* de fonctions élémentaires conduisant à une déstructuration éventuelle par des collisions et des interblocages, pour résoudre ces cas. Il peut aussi prendre en compte des calculs de fonctions élémentaires produisant des effets analogues sur la mémoire, pour les faire se réunir. Les fonctions du niveau méta ont pour but de rationaliser les transformations des fonctions élémentaires en les groupant ou en les séparant. Le contrôle opéré n'est pas exécutoire au point de modifier directement les programmes de ces fonctions élémentaires, mais est simplement une altération de la mémoire partagée de celles-ci. Le niveau méta change les informations au vu des actions des fonctions élémentaires, pour réorganiser indirectement le système.

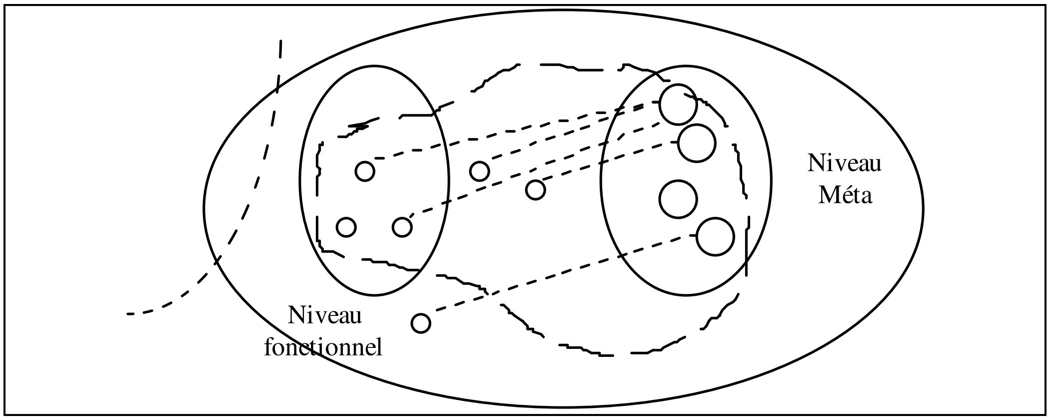

Figure 5. Les deux niveaux, fonctionnel et méta.

Les fonctions du méta-niveau relient ainsi *dans l'espace et le temps* les actions des fonctions élémentaires, pour permettre une évolution du système *entier* dans les limites de sa rationalité. Elles permettent de passer d'un niveau de granularité fin et homogène, celui de l'organisation des fonctions élémentaires indépendantes, à un niveau de granularité plus structuré, avec des groupes de fonctions élémentaires associées dans des collisions. Ces fonctions méta vont donc lier, sélectionner, grouper, agréger des fonctions élémentaires, pour structurer le système à un autre niveau. Elles expriment le *souci* du système à considérer certains éléments fonctionnels plutôt que d'autres. Elles utiliseront fortement des programmes à heuristiques pour se guider dans les sélections des groupes de fonctions élémentaires. Elles opèrent sur la structure des fonctions élémentaires et correspondent au mécanisme de conceptualisation, c'est-à-dire de réification avec permanence.

Ces fonctions sont typiquement non déterministes. Un programme de fonction méta prend simultanément en entrée les états et les zones mémoire de plusieurs fonctions élémentaires différentes. Elles groupent effectivement les fonctions élémentaires, car elles altèrent des parties significatives de la mémoire interne qu'elles manipulent. En jouant simultanément sur les données de ces fonctions, elles modifient leurs rôles particuliers, les agrégeant fonctionnellement.

Les fonctions méta lient organiquement, c'est-à-dire par des actions à effet de dépendances simultanées, les fonctions élémentaires. Ainsi, le fonctionnement d'une fonction élémentaire, dont le rôle spécifique et rationnel a été défini à la construction du système, va pouvoir implicitement modifier le fonctionnement d'autres fonctions, par des collisions provoquées par des fonctions méta. Une fonction méta peut altérer, pendant le fonctionnement d'une fonctions élémentaire et par le fait de ce fonctionnement, un endroit de la mémoire interne où *d'autres* fonctions opèrent. Nous avons donc introduit la complexité organisationnelle dans le système [Cardon 96]. On peut considérer, dans ce système à double niveau, que les fonctions élémentaires ne sont alors que génératrices de formes calculables plus générales. Le rôle des fonctions méta est de provoquer ou inhiber, tant que faire se peut, des collisions et ainsi de modifier indirectement les programmes des fonctions élémentaires du niveau fonctionnel. La fonctionnalité simple, distribuée, des fonctions élémentaires se transforme en système auto-organisateur à capacité de généralisation.

Remarquons bien que le niveau méta ne doit pas jouer le rôle d'un système de contrôle strict et centralisé, mais doit être un simple réorganisateur distribué du système fonctionnel, en jouant sur les données de la mémoire interne et donc indirectement sur les programmes des fonctions de niveau fonctionnel. Ces fonctions seront distribuées dans toute l'organisation des fonctions élémentaires, pour faire opérer les actes de réorganisation. Ce sont ces fonctions qui, en partant de formes disparates constituées de fonctions élémentaires spécifiants par exemple des sons, des images, des mouvements dans la durée, les états de l'organisation constituant des caractères géométriques et temporels, vont pouvoir reconstituer le concept que traduit un mot contextualisé de la langue. Un mot sera donc vu comme une agrégation complexe dont l'une de ses facettes, parmi beaucoup d'autres, sera sa forme lexicale dénotée.

Nous donnons alors à l'état initial du système, compris comme composé des programmes de toutes les fonctions et de l'état de la mémoire interne, un statut fondamental. C'est l'état *générateur* de *ses* réorganisations futures et non un état fonctionnel suffisant ou encore créateur d'un quelconque nouveau système. Le système a une identité donnée à sa création et il la garde pendant tout son fonctionnement. Ce fonctionnement n'est pas un simple processus fonctionnel, mais une réorganisation continue : le système se modifie organisationnellement en fonctionnant dans le temps et il est adaptatif d'abord à lui-même.

Il est bien clair que le fonctionnement du système pendant un certain temps est bien équivalent à celui d'une fonction générale, qui est en quelque sorte "l'union" de toutes les fonctions distribuées de son organisation, mais que nous n'avons pas complètement définie au moment de la construction. Cette fonction "union", après une certaine durée de fonctionnement, est comprise comme étant une fonction bien calculable, mais seulement après un certain nombre de pas de fonctionnement. Et tout fonctionnement dépend de la mémoire interface, que le système ne contrôle pas, et du règlement parfois aléatoire de certaines collisions qui augmentent physiquement les fonctions élémentaires. Ce système est bien adaptatif aux données externes et à sa structure initiale. Il est placé dans le domaine du calculable à chaque instant, mais ouvert sur la complexité organisationnelle par construction, pendant tout son fonctionnement.

3.5. Conscience de soi artificielle = réflexion sur et par l'organisation

Il faut maintenant traiter la question de la reconnaissance de "soi dans son environnement" réalisée par un tel système. Le problème du "contrôle" du niveau méta, de sa raison à faire évoluer les fonctions élémentaires dans une direction plutôt que dans une autre, se pose. L'auto-organisation ne peut se faire par hasard ou par évolution favorable sur une longue période. Il est nécessaire de définir un niveau structurel permettant de "faire tendre", de "donner une raison" au processus de réorganisation initié par le niveau méta.

3.5.1. Le niveau d'interprétation

Nous allons faire choix d'un type de structure permettant d'apprécier l'organisation du système, particulièrement dans son processus réorganisateur. Nous introduisons un niveau structurel d'anticipation, organiquement lié aux deux niveaux précédents, les niveaux fonctionnel et méta, et permettant, en se représentant partiellement certaines de leurs évolutions en les situant dans le temps, c'est-à-dire dans le futur, d'influer sur leur organisation présente. Nous appellerons ce niveau celui de *l'interprétation*.

Définition

> *Le niveau d'interprétation est une organisation de fonctions plastiques ayant la capacité à réfléchir l'organisation fonctionnelle pour la modifier globalement. C'est le premier caractère typique de conscience artificielle du système.*

Ce niveau est donc une appréhension de la capacité de réorganisation des fonctions du niveau fonctionnel, une appréhension de la tendance organisatrice de ce niveau. C'est une mise à la distance du niveau fonctionnel par rapport à ce qu'il produit en collision et écritures dans la mémoire. Le niveau d'interprétation représentera un caractère typique du process de prise de conscience du système adaptatif. C'est toujours un niveau symbolique, introduit par nécessité, pour permettre au système à se déployer dans un espace rationnel d'évaluation de son organisation.

Le système doit pouvoir se représenter sinon son organisation, du moins sa tendance organisatrice. Il doit donc avoir connaissance de son état structurel dans son devenir, pour le contraindre. Cette propriété de la conscience artificielle est le moyen structurel du système adaptatif pour se réorganiser sous sa tendance, sans se transformer en chaos aléatoire ou se figer en fonctionnement d'automate.

Il serait évidemment peu judicieux de doter le système d'une fonction de supervision générale, ce qui reviendrait à mécaniser brutalement tout le système et à prévoir plus ou moins complètement son fonctionnement dans ses futurs possibles. La conscience artificielle est d'abord une mise en recul de l'organisation du système par rapport à son propre fonctionnement, une mise en situation devant le déroulement temporel qui le ferait sinon dévaler dans le temps de sa réorganisation non appréhendable.

Cette mise en recul est réalisée par la possibilité de faire apparaître les résultats possibles et les importances réorganisationnelles de certaines actions des fonctions élémentaires et méta, alors que celles-ci ne se sont pas encore réalisées. Le système peut donc appréhender le *rôle* de certaines fonctions élémentaires et méta, en réifiant ces rôles par des formes construites, pour en moduler le fonctionnement.

Définition

> *Le process de prise de conscience artificielle dans un système adaptatif est constitué d'une structure de fonctions plastiques opérant sur les tendances à la réorganisation du système fonctionnel. Cette structure appréhende le niveau de réorganisation, pour en représenter les rôles, c'est-à-dire certains états futurs engagés mais non réalisés, avec l'objectif de les manipuler pour modifier l'état fonctionnel présent. Cette appréhension est géométrique sur l'organisation.*

Cette définition calculable d'un caractère fondamental de la conscience artificielle réfère donc à la géométrie de l'organisation en transformation, prenant en compte des gradients de transformation morphologiques, et suit l'approche thomienne [Thom 72]. Elle revient donc à définir d'abord un système organisateur, composé d'un niveau fonctionnel et d'un niveau méta, puis à *situer* ce système organisationnel grâce à un système particulier, qui se représente les tendances réorganisatrices en vue de modifier l'organisation elle-même.

Le sous-système fonctionnel est mis à la distance de son fonctionnement par un sous-système d'interprétation, qui conçoit une représentation des actions des fonctions élémentaires et méta, pour les modifier. En quelque sorte, le système fonctionnel ne peut qu'opérer réactivement sur l'environnement, par de multiples fonctions calculables spécifiques et on le dote d'une capacité à interpréter ses réorganisations, pour se restructurer autrement.

Ainsi, le système se *finalise* [Le Moigne 90] en se dotant d'un process d'interprétation qui évalue le fonctionnement du système fonctionnel, qui construit des représentations situées, pour le manipuler. Les caractères de sa finalisation seront un choix du constructeur du système. Ce peut être une raison de pure rationalité, pour rendre le système le plus efficient possible, ou bien une raison d'économie, pour préserver la dépense d'énergie du système, ou encore un choix de compensation d'états désorganisés par des états considérés comme organisés, pour un système qui s'achemine vers sa mort inéluctable, en le sachant. Ces caractères, qui ne peuvent pas ne pas être définis, sont en fait des *attracteurs généraux* de la raison a être du système [Cardon 96].

Le process de prise de conscience artificielle est donc vu comme un miroir actif, agissant sur les changements structurels possibles et réfléchissant des tendances d'organisation, pour en biaiser le résultat.

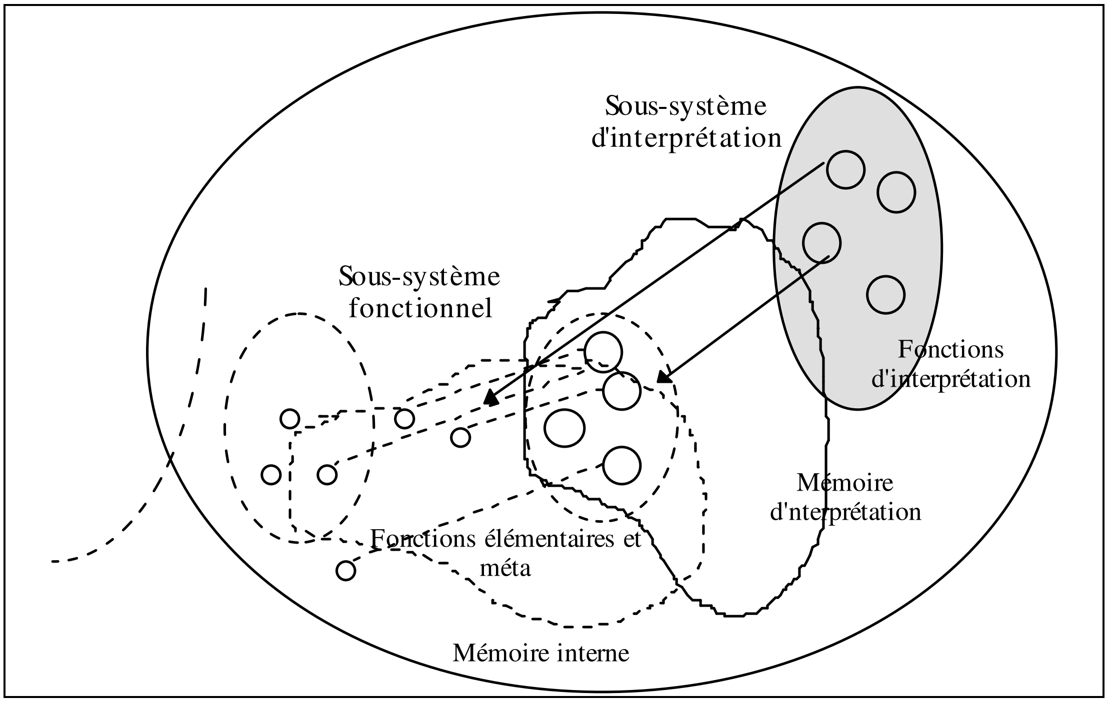

Figure 6. Le système d'interprétation opérant sur les actions du système fonctionnel.

3.5.2. Les fonctions d'interprétation

La structure du sous-système d'interprétation se construit par rapport au niveau fonctionnel, de manière duale. Le niveau d'interprétation est composé d'un ensemble de fonctions, les *fonctions d'interprétation*, opérant sur une mémoire propre, la mémoire interne d'interprétation et produisant ses résultats sur la mémoire partagée du système fonctionnel. Ces fonctions d'interprétation sont les représentations des actions des fonctions méta sur les fonctions élémentaires. Elles sont fournies à la construction du système comme des fonctions génératrices et sont régulièrement augmentées et recrées par mutations, avec les mêmes possibilités de collisions que les fonctions plastiques du niveau fonctionnel.

Définition

> *Les fonctions d'interprétation sont des fonctions plastiques prenant leurs entrées sur les actions des fonctions élémentaires et méta dans la mémoire partagée et réalisant leurs sorties sur une mémoire interne d'interprétation spécifique* <u>*et*</u> *sur la mémoire des fonctions élémentaires.*

Dans le système fonctionnel, les fonctions du niveau méta altèrent les actions des fonctions élémentaires. Les fonctions d'interprétation ont des programmes qui sont des représentations des actions communiquantes des fonctions méta. Ces fonctions lisent les données de la mémoire interne, manipulées par les fonctions du niveau fonctionnel et prennent connaissance des états des fonctions élémentaires. A partir de cela, elles se représentent le rôle transformationnel de ces fonctions. Elles vont donc "penser" le sens de leurs programmes qui, rationnellement, ont manipulé, manipulent et manipuleront ces données. Elles opèrent pour cela par sélection de zones de la mémoire interne et écrivent ou réécrivent leur nouveau code en utilisant celui des fonctions élémentaires et méta. Ensuite, elles modifient les données de la mémoire interne fonctionnelle : elles changent la perception du système fonctionnel, en se le représentant.

On pourrait croire que la manière la plus simple de réaliser ces fonctions serait de cloner, en quelque sorte, les fonctions du niveau fonctionnel. Mais il n'y aurait dans ce cas *aucune interprétation* du niveau fonctionnel. On souhaite au contraire, interpréter *les* actions des fonctions élémentaires et méta, c'est-à-dire les abstraire, par changement d'échelle. Les fonctions d'interprétation opèrent toujours sur des groupes indifférenciés de fonctions élémentaires, pour réifier l'action de ces groupes et ensuite modifier la mémoire des fonctions du niveau fonctionnel. Elles représentent une typicalité du process de prise de conscience, par abstraction d'un niveau fonctionnel dans un niveau qui réifie le sens d'action de certains groupes de fonctions élémentaires et méta. Le process de prise de conscience est d'abord abstraction et ensuite action modificatrice sur la mémoire partagée, donc indirecte sur la fonctionnalité du système.

Si par exemple, des fonctions méta calculent des fonctions d'agrégation de fonctions élémentaires en résolvant des collisions d'une certaine manière, la fonction d'interprétation correspondante ne peut que se représenter le *rôle* de ce groupe de fonctions dans le processus de réorganisation des fonctions élémentaires et en conclure qu'elle doit réifier ou non ce rôle dans un programme qui sera spécifique. Elle construit donc des traces mnésiques *d'actions de reconformation générales* et surtout pas des traces ponctuelles ni les traces du résultat de ces réorganisations. Ainsi, la conscience du système adaptatif s'appuie sur une catégorie de mémoire dynamique d'actions de reconformation interprétées, dotée d'un caractère fondamental d'abstraction donnant des rôles à des catégories de fonctions. La conscience est aussi action de mémoire et de création et sa mise en action, implicite, sur le système fonctionnel, est un *bouclage*. Ce bouclage représente, dans le système, le caractère incrémentiel de conscience artificielle dont il est doté.

Le rôle de l'organisation des fonctions d'interprétation est bien de réaliser une représentation, par abstraction, d'un niveau fonctionnel. Ces fonctions expriment les caractères évolutifs de la réorganisation du niveau fonctionnel. Elles peuvent donc représenter le caractère temporel de cette réorganisation, c'est-à-dire apprécier sa durée, sa chronologie, les localisations dans le temps de ses fonctions élémentaires et méta et en déduire ainsi que le réel, définitivement extérieur que ces fonctions représentent, a des caractères temporels, a une chronologie événementielle, a des localisations géométriques et temporelles de certains événements. C'est bien la temporalité du système d'interprétation qui permet de représenter la temporalité de l'environnement.

Le rôle des fonctions d'interprétation est donc majeur en ce qui concerne le process de prise de conscience, puisque celles-ci permettent de représenter, pour manipulation propre, le système et ce qu'il se représente de son environnement.

Comprise ainsi, la conscience de soi est, pour un système adaptatif, la conscience de sa situation dans son environnement, sans cesse changeante, représentée d'abord par le niveau fonctionnel.

Les fonctions d'interprétation sont des représentations *abstraites* du niveau fonctionnel, des concepts abstraits des fonctions élémentaires obtenus en situant toujours ces conceptualisations dans leur rôle réorganisateur. Elles représenteront donc notamment les caractères langagiers que le système peut associer à des fonctions du système fonctionnel. Elles sont les générateurs du langage, que la structure cohérente du niveau fonctionnel implique plus ou moins. L'ensemble des fonctions d'interprétation forme un système calculable, mais *au niveau du sens* de ce que le système fonctionnel représente.

La différence entre un tel modèle de système adaptatif et un système à base de connaissances est radicale. Un SBC opère directement sur ses entrées. Il a pour but, explicitement, de traiter les entrées qui sont des informations du réel extérieur en les comparant à des connaissances compatibles mémorisées en propre, de les analyser et de les augmenter en incorporant des cas connus, puis de déduire une mise en situation de ces connaissances du domaine en inférant grâce à des connaissances de contrôle. Le SBC produit un résultat dans le même domaine que celui des entrées. Il résout un problème posé par l'utilisateur, et c'est son unique rôle.

Le système adaptatif a pour but de créer et modifier *ses* connaissances de représentation du domaine, représentées par les fonctions élémentaires et méta. En ce sens, cette procédure se distingue déjà de l'approche SBC par l'action réorganisatrice des fonctions méta, qui altèrent l'organisation. L'action du système interprétatif est totalement étrangère à l'action d'un SBC. Elle revient à manipuler des entités d'action, à créer des situations de manipulation, en situant les processus de réorganisation (les fonctions méta opérant sur les fonctions élémentaires) dans un espace de déploiement abstrait (celui des fonctions d'interprétation) c'est-à-dire en se focalisant sur la notion de réorganisation autoréférente du système. Même si au bout du compte le système adaptatif a résolu un problème en faisant fonctionner, en modifiant et en inhibant certaines fonctions, il s'est comporté *pour* se réorganiser et non pour résoudre fonctionnellement un problème posé par un quelconque opérateur externe. De la même façon, un raisonnement rationnel opéré dans l'esprit humain ne rationalise pas le fonctionnement de l'esprit, mais il utilise sa complexité organisationnelle, dans ces trois caractères d'espace, de temps et d'organisation, pour procéder à la représentation de résultats jugés rationnels à l'observation, sans plus.

Notre propos, en décrivant un modèle de système adaptatif calculable permet d'expliciter la boucle de la complexité organisationnelle : mémoire interface, réaction fonctionnelle, modification de la mémoire interne partagée, réorganisation du système fonctionnel par interprétation, puis à nouveau entrée-sortie .... Le concepteur doit préciser, selon ses objectifs, le rôle spécifique de chaque fonction d'interprétation. Une conscience artificielle est fondée sur le seul principe général d'une augmentation de catégorie : la génération autoréférente du sens des fonctions du niveau fonctionnel par des fonctions d'interprétation. Sa construction revient à définir les générateurs de l'augmentation, c'est-à-dire les générateurs conceptuels, langagiers, géométriques ... , spécialisés en hiérarchies et en classes.

Remarquons que le système peut parfaitement opérer en boucle fermée, sans que les fonctions d'interface ne réalisent autre chose que de la veille. Peut-être est-ce une propriété inévitable du système adaptatif : la capacité à méditer ou à rêver.

3.6. L'émergence d'une forme de pensée: le point fixe de la boucle d'interprétation par des fonctions de décision

Le système d'interprétation génère la réorganisation de la structure fonctionnelle sous forme de boucle d'interprétation :

- activation des fonctions d'interprétation, modification des données de la mémoire partagée du niveau fonctionnel, modification indirecte des codes des fonctions élémentaires et méta et de leurs états, modification de la mémoire distribuée par les fonctions élémentaires, puis à nouveau action des fonctions d'interprétation ...

Il reste à définir la cohérence de la réorganisation du système et ce qu'il produit, c'est-à-dire ce qu'est, pour lui en tant que système, une "pensée émergente". Le pilotage du système est en effet lié à ce qu'il produit : il n'y a pas ici de différence de genre comme celle établie usuellement entre un processus de contrôle et le résultat de ce processus. La raison du système d'interprétation à fonctionner par activations du niveau fonctionnel est de distinguer un certain ensemble de fonctions, composé de fonctions d'interprétation, de fonctions méta et de fonctions élémentaires, qui sera structurellement cohérent au regard de l'organisation générale du système, et qui sera *la* forme permettant à celui-ci de continuer à se réorganiser. Cette forme sera le correspondant d'une pensée émergente dans un esprit conscient, le sens émergent, et elle ne vaut que par ce qu'elle permet de générer. Il s'agit donc bien d'un processus autoréférent de création de formes.

Définition

> *Le sens émergent dans le modèle de système à boucle d'interprétation est un point fixe de l'action de reconformation de l'ensemble des fonctions d'interprétation sur les fonctions du niveau fonctionnel, à un instant donné.*

Le sens émergent est entendu comme point fixe d'une fonctionnelle opérant de l'ensemble des fonctions d'interprétation dans l'ensemble des fonctions du niveau fonctionnel. Il s'agit donc de définir une évolution du système qui, à partir du niveau fonctionnel et des générateurs des fonctions d'interprétation, construit par bouclage de nouvelles fonctions d'interprétation, qui altèrent à leur tour les fonctions élémentaires et méta jusqu'à une stabilisation. Le principe de fonctionnement de ce système bouclé et évolutif est la réduction *locale* de la réorganisation : le système tend à se transformer, mais de telle façon que les transformations équilibrent localement le niveau d'interprétation avec le niveau fonctionnel, mettant ces deux niveaux en adéquation locale.

3.6.1. Les fonctions de décision

Cet équilibre revient à réduire par inhibition l'action de certaines fonctions d'interprétation et à faire correspondre l'action des fonctions élémentaires et méta avec leur conceptualisation au niveau interprétation. Ceci sera réalisé par un ensemble de fonctions particulières, au rôle de fonctionnelles, les fonctions dites de décision. La recherche d'équilibre est représentée par un point fixe de fonctionnelle liant les deux niveaux du système. Cela signifie aussi que les fonctions d'interprétation correspondront toujours fortement, à leur niveau d'abstraction, aux fonctions du niveau fonctionnel : le système reste attaché à ses caractères fonctionnels et sa mémoire dynamique contraint son fonctionnement.

Dans le fonctionnement bouclé du système, on représente les fonctionnelles de recherche d'équilibre par des fonctions plastiques polymorphes, appelées *fonctions de décision*, mettant en relation des fonctions du niveau interprétation avec les fonctions du niveau fonctionnel.

Définition

> *Les fonctions de décision sont des fonctions plastiques qui opèrent symétriquement en tant que fonctionnelles entre l'ensemble des fonctions d'interprétation et l'ensemble des fonctions élémentaires et méta, dont le rôle est de sélectionner et réduire les possibilités de reconformation de la boucle d'interprétation, pour faire tendre la réorganisation vers un point fixe.*

Les fonctions de décision ont des actions polymorphes, en ce sens qu'elles agissent *comme* des fonctions élémentaires, ou méta ou d'interprétation, pour inhiber certaines autres fonctions de même genre, mais au niveau de l'organisation, globale et locale du système. Elles unifient les deux niveaux du système à boucle d'interprétation, en jouant un rôle à la fois régulateur et créateur, par inhibitions locales.

Il est évidemment intéressant de typicaliser les actions de ces fonctions, c'est-à-dire de donner un sens au sens de la réorganisation. Cela est effectivement possible dans un tel système, en choisissant des caractères transversaux forts pour les fonctions de décision : recherche de l'économie d'énergie ou de l'optimalité, recherche du dépassement ... Mais nous avions déjà évoqué cette question en définissant des *attracteurs* dans les systèmes profondément complexes [Cardon 96]. Les possibilités sont larges et laissent à penser que les fonctions de décision, clé de toute l'organisation, permettent de choisir des sortes de systèmes adaptatifs plutôt que d'autres, c'est-à-dire des genres d'organisation vers lesquelles le système sera inéluctablement conduit d'aller, et en fin de compte à générer certaines formes de pensée plutôt que d'autres.

3.6.2. Vers les agents

L'approche développée ici est évidemment et volontairement fonctionnelle. Notons quand même un basculement de catégorie entre les fonctions élémentaires et méta, qui ont des comportements de processus classiques et les fonctions d'interprétation, qui donnent des rôles à ces dernières. Les fonctions d'interprétation sont un peu plus que des fonctions algorithmiques classiques. Elles ont la possibilité de se représenter ce que sont et font les fonctions élémentaires : ce seront en fait des agents cognitifs. La notion d'agent intervient par nécessité dans le modèle au moment où l'on introduit des entités qui agissent en affectant des rôles à d'autres fonctions.

Les fonctions de décision, quant à elles, ne sont pas simplement des fonctionnelles et des agents qui se représentent les rôles de certaines autres fonctions, mais des attracteurs d'organisation qui opèrent essentiellement au niveau de l'organisation et qui ne réifient pas, à aucun moment, des concepts autres qu'organisationnel. Nous avions proposé de les représenter par des entités partout dense dans l'organisation [Cardon 96].

Quatrième partie
Architecture par organisations d'agents du système à boucle d'interprétation

L'architecture du système à boucle d'interprétation repose sur le modèle défini dans la partie précédente, mais elle sera en fait basée sur des organisations d'agents. A la différence d'un système composé de fonctions qui opèrent comme des processus concurrents en accédant à une mémoire partagée, les agents localisent dans leur structure les notions de mémoire et de capacité de décision et d'action. C'est typiquement une approche distribuée du système.

L'agent permet de distribuer et d'abstraire la notion de fonction, de processus et de mémoire partagée. L'entité de base du système sera donc autonome quant à son comportement et à sa capacité de mémorisation et aura de plus une capacité à se représenter un certain contexte formé d'autres entités du même genre, avec son *rôle* dans ce contexte. L'agent communique avec les autres agents pour agir. Cela clarifie grandement la définition ontologique des niveaux fonctionnel et d'interprétation, en encapsulant dans des agents les fonctionnalités, les connaissances, les capacités d'action et la mémoire des processus concurrents. La notion de collision est alors remplacée par celle de communication, plus claire et beaucoup plus riche.

La notion d'agent, comme réification des fonctions physiologiques et des cartes cognitives, permet de bien représenter le caractère topologiquement et cognitivement imprécis de ces notions. L'agent a un certain rôle, qui peut évoluer et confère à ses actions une importance également modulable. Au processus de production d'une réorganisation correspondra typiquement une émergence d'organisation d'agents. Le système d'agents n'ayant, par définition, aucun contrôle centralisé, correspond fort bien au problème posé.

4.1. Le processus d'interprétation dans un système d'agents

Classiquement, un agent est une entité d'action présentant des caractères typiques d'autonomie comportementale, opérationnellement clos par rapport à l'environnement dans lequel il peut agir. L'agent n'a de sens que parce qu'il agit dans un environnement constitué à la fois d'objets différents de sa nature et d'autres agents, plus ou moins semblables, en formant un système multiagent.

Dans le problème que nous traitons, il s'agit de modéliser puis de construire un système d'interprétation prenant en entrée des formes langagières standardisées (par exemple des "frames" de symboles ou autres modalités définies par un utilisateur probablement humain) et produisant à partir de cela une émergence du sens sur l'ensemble des formes internes du système.

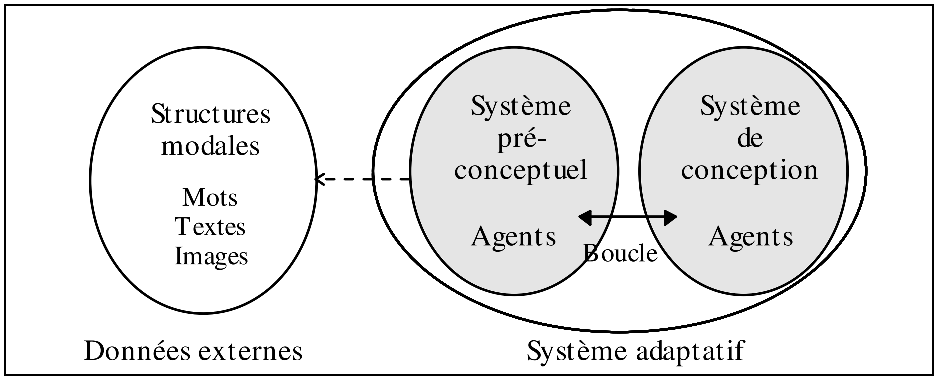

Figure 7. Schéma du système à boucle d'interprétation.

Le système doit d'abord avoir un niveau de perception des formes symboliques données en entrée. Il doit ensuite avoir une capacité à activer certaines formes internes, en exprimant un souci propre, lui permettant de considérer certaines formes internes plutôt que d'autres et de les altérer. Il doit enfin pouvoir construire un autre ensemble de formes représentant les caractères abstraits de ces premières formes, qualifiées temporellement, spatialement et organisationnellement, en faisant appel à sa mémoire factuelle, qu'il reconstruit constamment. Il doit pouvoir enfin analyser les caractères de ces organisations de formes, susciter de nombreuses possibilités de compréhension et en retenir particulièrement une, qui sera le sens émergent sur les formes disponibles au niveau de la perception. Toutes ces formes seront, dans le système, représentées par des agents.

Tout agent a des capacités d'autonomie et d'évaluation de ses actions grâce à une structure interne qui possède globalement les caractères suivants :

1. une représentation de son environnement :
    - systèmes à base de connaissances,
    - structures de type réseaux sémantiques,
    - filtres lexicaux,
    - réseaux d'accointances,
    - mémoire factuelle,
    - fonctions d'altération de ces caractères,
2. une représentation de son organisation :
    - variables d'états et d'aspect de sa structure,
    - fonctions d'altération de ces caractères,
3. une représentation du temps :
    - variables temporelles internes (durée, vitesse de variation ...)
    - une représentation de son état énergétique :
    - variables homéostatiques,

  - variables de tension, de besoin,
4. une représentation de sa finitude :
   - fonction d'obsolescence,
   - scénarios de comparaison,
5. une représentation de son situationnement :
   - ATN,
   - variables décrivant son état dans l'ensemble des agents alliés et opposés,
   - fonctions d'altération de ces caractères,
6. une fonction de reproduction :
   - fonction de clonage,
   - fonction de reproduction sexuée.

L'agent n'a pas a priori de rôle auto-adaptatif ni de processus d'interprétation d'une situation complexe : il traduit simplement un certain caractère fonctionnel plastique et ce seront les organisations d'agents qui généreront, par émergence, des formes traduisant, par auto-adaptation, l'interprétation de l'environnement.

Le système est donc clos, au sens où il opère par clôture opérationnelle sur un environnement constitué de structures multimodales, pour générer une interprétation par émergence. Il est constitué d'organisations en couches reconformables d'agents, agissant essentiellement pour se grouper, se réduire ou se développer, par le fait de leurs communications.

4.2. L'ontologie du système

Les prototypes de systèmes à boucle d'interprétation développés, pour une simple question de complexité combinatoire, ont une opérationalité dans des domaines précis et limités. Le premier développement réalisé concernait le problème de la représentation du sens des messages échangés entre les décideurs lors d'une crise civile [Cardon-Durand 97], [Cardon-Lesage 97]. L'ontologie du domaine était alors celui du raisonnement tactique et de l'appréciation de la dangerosité des caractères d'une situation fortement évolutive.

De manière générale, au niveau ontologique, il est nécessaire de représenter les caractères typiques des données du système et de leur faire correspondre certaines entités particulières sous forme d'un premier ensemble d'agents. De plus, il est nécessaire de définir certaines catégories propres d'agents, comme celles correspondant aux fonctions homéostatiques, entropiques, de sénescence ou de reproduction, ou encore structurelles.

Le contenu cognitif du système est regroupé en catégories relatives à son domaine d'opérationalité. Nous faisons l'hypothèse que les connaissances appréhendables prennent place dans cinq catégories cognitives :

- les connaissances spatiales,
- les connaissances temporelles,
- les connaissances organisationnelles,
- les connaissances langagières et objectives,
- les connaissances de perception subjective.

Ces catégories se déclinent toutes selon divers traits sémantiques typiques [Jackendoff 83] en formant une vaste organisation d'agents. La dernière catégorie apparaît ici comme fondamentale, car elle traite des valeurs subjectives. Elle utilisera systématiquement la notion de création duale pour tout agent créé, en se référant à la notion de processus opposant [Vincent 86]. Toute catégorie est considérée comme ayant sa catégorie opposée et tout processus de développement dans l'organisation est considérée comme ayant un processus opposé, la notion d'opposition étant étendue à la contradiction, à la négation, au contraire.

La définition ontologique d'un système à boucle d'interprétation se fera en considérant les six niveaux d'organisation des systèmes complexes [Lapierre 92], et en les déclinant en catégories et sous-catégories, caractères puis sous-caractères. Ces travaux sont effectués dans des thèses en cours.

1. Le monde physique, les entités objectives.
2. L'espace de développement des entités objectives.
3. Le mouvement, les organisations, la planification.
4. **Communication des Informations**
5. Valeurs et symboles perçus, sens subjectif des phénomènes et intentions.
6. Règles du jeu social, relations de pouvoir, émergence du sens global du phénomène.

Figure 8. Les six niveaux d'organisation d'un système complexe.

Les trois premiers niveaux appartiennent au domaine de la modélisation ontologique des systèmes classiques, le quatrième permet l'organisation dynamique des trois précédents en prenant en compte la communication interne de l'information entre les agents. Les niveaux 5 et 6 appartiennent au domaine psychologique, social et culturel, que le système doit intégrer sous forme d'agents spécifiques.

Ces deux derniers niveaux ne peuvent pas être représentés par des structures totalement définies a priori et qui utiliseraient des composants élémentaires bien fixés. En effet, l'importance et le type de catégories sociales et psychologiques que ces niveaux représentent dépendent de l'organisation du système elle-même. Nous aurons donc à représenter ces catégories par des agents générateurs et à les apprécier par leurs organisations dynamiques. C'est également par ce caractère que le système ne peut être décomposé a priori en sous-systèmes totalement fixés, mais est le résultat d'organisations changeantes pendant le fonctionnement. Ces niveaux appartiennent typiquement au domaine de la *complexité profonde* [Cardon 96]. Nous nous intéressons principalement à eux pour prendre en compte effectivement les catégories d'intentions, d'opinions et de jugements. Ils correspondent aux caractères de l'état central fluctuant des organismes vivants.

Une description du système au niveau ontologique distingue deux types de connaissances :

1. les connaissances à propos du réel extérieur, représentées par diverses catégories précises. On définira des nomenclatures, des hiérarchies de connaissances appliquées à différents

domaines : espace physique, structures physiques, plans, espace social ou linguistique, mots et grammaires ...

2. les connaissances de l'organisation interne du système lui-même. Elles représentent la connaissance sur la façon dont les connaissances factuelles précédentes sont structurées dans le système. Il s'agit donc de connaissances de l'architecture des agents et de leurs rôles. En ce qui concerne le langage, il s'agira de la connaissance d'usage des unités lexicales, de leurs structures dynamiques et complexes et de leurs rôles comme entités réorganisatrices. Seuls des agents semblent aptes à bien représenter ces différentes fonctions.

Ces deux types de connaissances sont organiquement liées, dans le système. Cela signifie que les structures usuelles, statiques, de représentation des connaissances qui permettent principalement leur objectivation pour traitement inférientiel, sont à exclure, au profit d'une approche typiquement sémiotique.

Une construction de l'ontologie du système, dans le domaine des situations dites d'urgence, est en cours de réalisation et sera décrite par ailleurs.

4.3. Une architecture à quatre organisations d'agents

Les données du système venant de l'environnement, sont des données multimodales : systèmes lexicaux, images, sons, mouvements ... Ces données vont d'abord être prises en considération par un système pré-conceptuel : l'organisation des agents dits agents aspectuels.

L'action de chaque agent aspectuel consiste à exprimer une tendance, pour activer la catégorie modale et sémantique qu'il réifie, en tenant compte de son contexte d'action, c'est-à-dire des autres agents aspectuels. Une première organisation d'agents va donc s'approprier les données du système pour en dégager les caractères spécifiques, dans le contexte courant du système. Cet ensemble d'agents aura une certaine morphologie, une conformation spatiale, qui va par la suite être prise en compte. Ces agents correspondent aux fonctions élémentaires.

Une seconde organisation d'agents va, en opérant sur la morphologie des agents aspectuels, en dégager les caractères organisationnels pertinents, c'est-à-dire les structures géométriques, sémantiques et temporelles pertinentes. Ces agents dits agents de morphologie, opèrent selon des buts rationnels et tentent de dégager les formes pertinentes dans *l'ensemble* des modalités et traits sémantiques, en constituant des groupes d'agents de morphologie, des chréodes [Jorion 89], en considérant les agrégations pertinentes d'agents aspectuels. Ces agents correspondent aux fonctions méta.

Une troisième organisation d'agents va prendre en considération les agents de morphologie pour former une conception de l'organisation des agents aspectuels. Ces agents, dits agents d'analyse ont trois caractères d'expression : espace, temps et organisation. Ils vont fournir, par une construction explicite d'agents d'analyse, une abstraction située de la morphologie des agents aspectuels. Ces agents correspondent aux fonctions d'interprétation.

Une quatrième organisation d'agents, les agents de décision, réalisent des actions d'inhibition du développement de la construction des agents d'analyse, en choisissant des stratégies dans les actions des trois organisations précédentes. Cette organisation permet de distinguer une certaine structure d'agents dans l'ensemble des agents du système : l'émergence du sens de l'organisation générale. Ces agents correspondent aux fonctions de décision du modèle.

Cette organisation est représentée dans la figure ci-dessous.

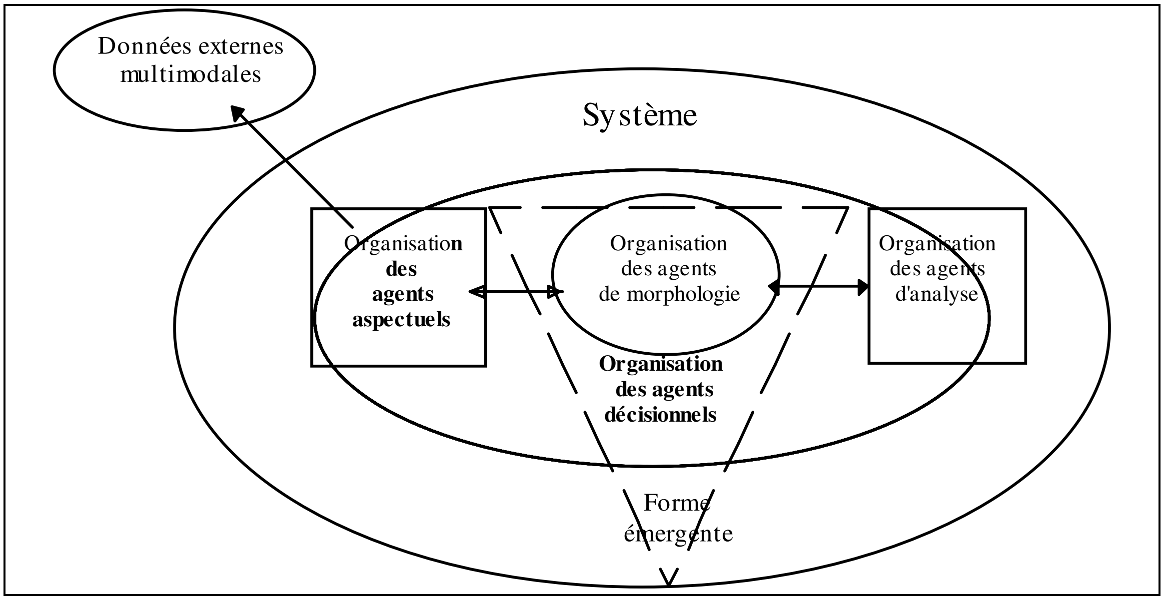

Figure 9. L'architecture du système à quatre organisations d'agents.

4.4. L'organisation des agents aspectuels

Le système interprète les données de l'environnement par l'activation et la modification d'une organisation d'agents dits *agents aspectuels*. Ces agents correspondent aux multiples fonctions élémentaires du modèle de système à boucle d'interprétation. On fait correspondre à la structure des donnée externes, comme dans la clôture opérationnelle d'un système autopoiétique, des *formes dynamiques*. Ces formes dynamiques, représentées par les agents aspectuels, réifient les caractères sémantiques et structurels des données multimodales, selon leurs capacités d'évaluation et d'action. L'ensemble des données interprétées dans un certain environnement est donc, pour le système, représenté par des entités concurrentes ou coopérantes, réifiant localement certains de leurs caractères. Sous cette forme, chaque donnée est, d'une certaine manière, dépendante de l'état du système pré-conceptuel (son attention courante, sa compétence ...), interprétée et représentée parmi les autres formes dynamiques, dans les caractères de l'organisation des agents aspectuels.

Les agents aspectuels sont des éléments générés dans un système de pré-conception par la perception de certains caractères signifiants de l'environnement, mis en forme pour l'appréhension. C'est une perception et une pré-conception, réalisées par structuration de l'organisation des agents aspectuels, selon leurs modes d'action et catégories prédéfinies. Le système possède donc un ensemble d'agents générateurs des agents aspectuels.

On peut considérer deux cas :

1. les agents aspectuels sont déjà présents et sont activés sur chaque caractère des données perçues,
2. les agents aspectuels sont engendrés, modifiés, transformés, par des générateurs d'agents aspectuels.

Nous retiendrons la seconde solution, plus complexe mais plus conforme au caractère de clôture du système. Une notion d'apprentissage devra alors être introduite dans l'organisation des agents aspectuels.

L'ensemble des agents aspectuels forme un espace, parce que ces agents sont liés par leurs réseaux d'accointances pour former un espace relationnel dynamique et également par plongement dans un espace métrique, après représentation des caractères des agents aspectuels par des vecteurs typiques. Cet espace d'agents est doté d'une morphologie, c'est-à-dire d'une conformation géométrique, d'une topologie avec métrique. Nous parlons donc de *morphologie d'espace d'agents*, qui a des caractères spatiaux et temporels, comme des gradients de transformation morphologique par exemple.

Chaque agent aspectuel est modélisé par un agent hybride [Bussmann 94]. L'ensemble complet des agents aspectuels est une organisation dynamique d'agents dont les caractères aspectuels (de forme, d'action, de relation ...) sont changeants : le graphe de connexion est dynamique. Pour pouvoir stabiliser une telle organisation d'agents, nous générons automatiquement, à la création d'un agent aspectuel, son dual au sens de la catégorie ou de la modalité qu'il représente (par exemple : impression de chaud versus impression de froid). Le développement de l'un signifie l'affaiblissement ou la mort pour le dual à plus ou moins long terme. Les agents antagonistes ont les mêmes buts généraux, survivre et se développer, et les mêmes moyens pour cela. Ainsi ils se combattent constamment.

Le développement local d'un agent aspectuel affecte le développement de nombreux agents aspectuels en relation par leurs réseaux d'accointances. L'organisation de ces agents modifie sa forme générale en fonction des caractères de l'environnement et de son organisation même. Elle possède donc une morphologie changeante.

Nous faisons l'hypothèse que cette morphologie, qui a à la fois des caractères géométriques et sémantiques, peut exprimer le point de vue du système d'interprétation sur le contexte et sur l'organisation des agents en action devant le contexte. Nous associons en fait à l'acte de langage [Austin 70] d'un acteur humain qui génère une entrée multimodale dans le système, une action de reconfiguration de la morphologie du paysage d'agents aspectuels de celui-ci [Cardon-Durand 97].

4.4.1. La structure des agents aspectuels

Afin de représenter la plasticité des agents aspectuels qui doit refléter celle de représentations mentales, nous les munissons une double structure d'automate. Cette structure à deux niveaux sera conservée pour la plupart des différentes catégories d'agents du système.

L'agent aspectuel est structuré par un automate à quatre états appelé macro-automate. Le premier état représente l'initialisation de l'agent. Dans le second état, l'agent tente de recueillir des informations sur *son* contexte afin de savoir si son développement est viable. Il a un comportement de questionnement. Si cet état lui semble viable, il passe dans le troisième état, sinon il attend ou renonce. Là, il tente d'imposer son point de vue, en communiquant pour altérer ses compétiteurs. Il a un rôle et en attribue aux autres agents. S'il y parvient, il gagne alors le quatrième état qui est l'état d'action. En ce point, l'agent évoque une tendance significative et permanente, mais locale et correspondant au traits sémantique qu'il réifie, de l'organisation des agents aspectuels : il a un rôle stable.

Chaque état du macro-automate est également modélisé par un automate qui sera appelé automate local dans les lignes suivantes. Cette structure complexe permet de représenter finement les différentes catégories sémantiques retenues, selon leurs modes.

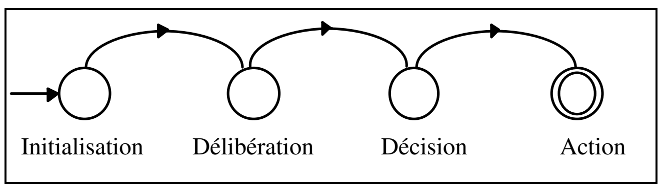

Figure 10. Le macro-automate des agents aspectuels.

4.4.2. Le Macro-Automate

Comme il est précisé dans le paragraphe précédent, le macro-automate possède quatre états : l'initialisation, la délibération, la décision et l'action. Les états correspondent au schéma typique de la décision dite linéaire [Sfez 92].

1. Dans le premier état d'initialisation, l'agent s'éveille. Il essaie de déterminer si la quantité et la force des messages qui l'ont éveillé est suffisante pour que son processus soit initialisé.

2. Dans le second état, dit de délibération, l'agent tente de déterminer si son contexte, en terme d'agents connus par lui, lui est favorable ou non. Pour cela, il va demander aux autres agents aspectuels des informations sur leur état. Selon celles-ci, l'agent passera dans l'état suivant, attendra ou reviendra en arrière, à l'état d'initialisation.

3. Dans le troisième état de décision, l'agent va tenter d'imposer son point de vue aux autres agents aspectuels. Il agit sur l'automate local de certains autres agents aspectuels et envoie des clones dans leur contexte, pour les altérer. Ces clones seront considérés comme des agents aspectuels nouveaux. L'agent aspectuel peut également annihiler certains agents faibles ou au contraire en renforcer certains autres, afin d'atteindre son but : se trouver dans l'état d'action.

4. Lorsqu'un agent réussit à gagner l'état d'action, il se trouve dans un état de développement fort dans l'organisation des agents aspectuels. Dans cet état, il agit sur le macro-automate même des autres agents, en tentant de l'altérer. Lorsqu'un agent aspectuel est dans cet état, on peut dire que la catégorie qu'il représente caractérise un certain point de vue significatif, permanent et pertinent. Il y a évidemment de nombreux agents aspectuels simultanément dans cet état d'action

4.4.3. L'état d'initialisation

C'est le premier état du macro-graphe. Un certain agent aspectuel est réveillé par un message dans lequel il a pu filtrer la catégorie qu'il représente. Ce filtrage est souple. L'agent va déterminer si la fréquence et l'intensité de l'information externe est suffisante pour rendre la catégorie qu'il représente potentiellement pertinente.

Cet état est implémenté par un réseau de transition augmenté (ATN) [Guessoum 96]. Lorsqu'un agent aspectuel saisit une information du contexte (une certaine structure lexicale selon une certaine modalité), il calcule la proximité entre la catégorie qu'il réifié et celle qu'il peut raisonnablement attacher à l'information du contexte. Selon cette proximité, une première transition est franchie ou non. La proximité est donnée par un lexique prédéfini. Une certaine structure représente dans l'automate le contexte de développement de son état (si le passage à l'état suivant a été difficile, rapide, laborieux...). Cette structure est sa mémoire d'activité qualitative.

4.4.4. L'état de délibération

Cet état est atteint lorsque la catégorie représentée par l'agent aspectuel a un caractère de pertinence dans le contexte des agents aspectuels. L'action, dans cet état, est de déterminer si l'agent a une chance ou non de survivre. Pour cela, il va demander des informations sur son contexte via les autres agents aspectuels connus par son réseau d'accointances. Il leur demande des informations sur leur état d'avancement et les qualités de cet avancement. Selon celles-ci et en tenant compte des messages reçus, l'agent peut déterminer si le contexte lui est favorable ou non. Si le contexte est favorable, alors l'agent passe dans l'état suivant. Sinon, il a trois possibilités : la régression à un macro-état précédent, la mort ou l'attente d'un contexte meilleur.

Afin d'évaluer le contexte, l'agent demande aux autres agents leur état précis. Il connaît ainsi l'état de son opposé absolu (son dual), de son proche le plus semblable, de ses alliés (ses proches) et de ses opposées (les proches de son dual). Il peut se représenter son contexte en le structurant, en comptant ses alliés et de ses opposés, en les groupant, en se représentant leurs forces. Il a une représentation de son monde. Il peut donc décider d'avancer vers l'état suivant, de reculer ou d'attendre.

Cet état est également modélisé par un ATN. Nous avons donné un comportement relativement autonome à l'agent, car nous ne souhaitons pas qu'il ait une dépendance aux seuls messages. Si le contexte est favorable et si l'agent attend un message pour atteindre l'état suivant, alors il peut choisir de lui-même d'être plus autonome et de franchir d'autorité la transition, sans attendre ce message.

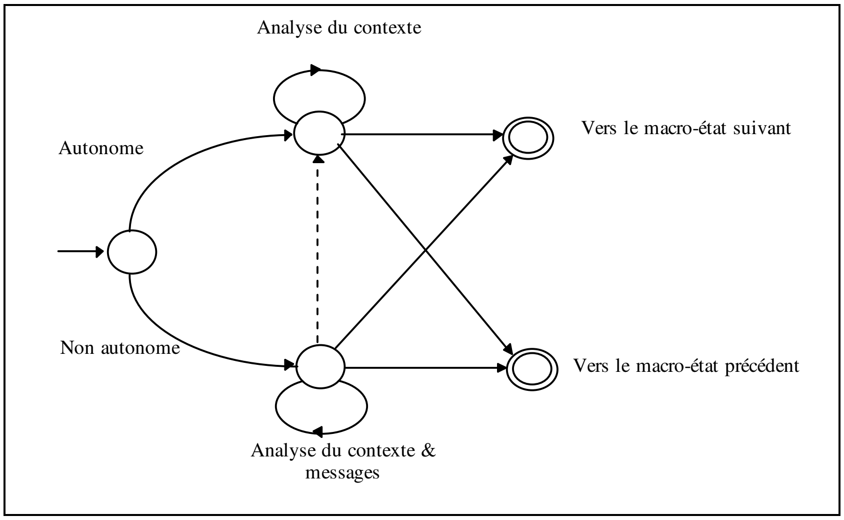

Figure 11. L'ATN de délibération.

Pour l'analyse du contexte, nous avons défini une mémoire situationnelle dans l'agent aspectuel. Cette mémoire permet de représenter l'état du contexte de l'agent. Si le contexte est bon, respectivement mauvais, la transition vers le macro-état suivant est passée, sinon l'action de changement d'état fait revenir vers le macro-état précédent. Si le contexte est intermédiaire, l'agent continue ses requêtes. Lorsque le contexte est proche d'un seuil de franchissement (par exemple, il n'est pas jugé bon mais n'en est pas éloigné) une fonction stochastique (ou à heuristiques) peut faire franchir la transition.

4.4.5. L'état de décision

Dans cet état, la catégorie représentée par l'agent est considérée comme pertinente au niveau des agents aspectuels. L'agent aspectuel va donc tenter d'imposer son point de vue. La structure de cet état est très proche de celle de l'état précédent. Les principales différences se situent dans les actions de son automate. En effet, en plus de requêtes sur l'état du contexte, l'automate va *agir* sur les autres agents aspectuels afin de ralentir ou d'accélérer leur développement. Ces actions peuvent être de deux types :

1. il peut y avoir action sur la perception du contexte d'un autre agent aspectuel afin de fausser sa représentation du monde,
2. il peut y avoir action directe sur l'automate d'un agent aspectuel en favorisant ou en rendant plus difficile le franchissement de certaines de ses transitions.

Lorsqu'il n'y a plus d'agents susceptibles de lui barrer radicalement la route, l'agent passe la dernière transition du macro-automate vers le macro-état d'action.

4.4.6. L'état d'action

A ce stade, la catégorie représentée par l'agent est celle qui structure de manière pertinente, mais toujours ponctuelle, l'organisation des agents aspectuels. L'agent altère régulièrement le comportement des autres agents en agissant sur leurs macro-automates. Certaines variations possibles du macro-automate des agents aspectuels : la bestialité, l'intellectualisme et la velléité [Sfez 92], sont présentées dans la figure ci-dessous.

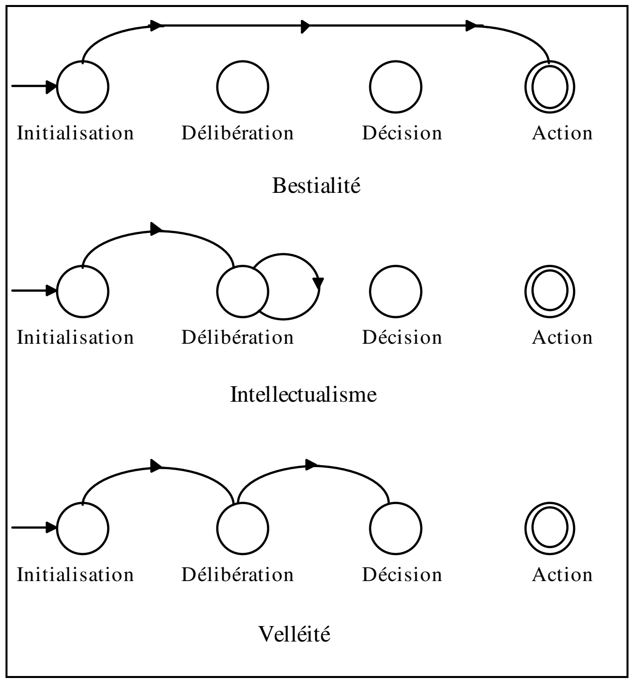

Figure 12. Différentes variations du macro-automate.

Nous voyons donc que les agents aspectuels représentent, pour chacun d'entre eux, des sous-systèmes de pensée, basés sur des mots (leurs catégories), liés fortement entre eux par leurs réseaux d'accointances et capables d'actions diverses. Nous respectons donc le point de vue de complexité présenté par Vygotski [Vygotski op. cité] sur l'entité minimale complexe pour représenter le lien entre mot et pensée.

Les agents aspectuels prennent en considération les structures modales qui leur sont fournies pour en extraire des caractères, des traits sémantiques ou des formes typiques. Ces caractères extraits sont multiples, concurrents, toujours considérés comme non indépendants. Nous substituons à la structure statique de la donnée extérieure un ensemble d'agents en actions incessantes et qui représentent la perception de cette structure par le système. Le caractère dynamique de chaque agent respecte le principe de complexité : chaque entité réifie une tendance à générer une pensée, mais non indépendamment de l'action de tous les autres.

Remarquons que la structure des agents aspectuels comprend la notion d'agent réactif pur, par dégradation du macro-automate en fonctionnement dit "bestial".

4.5. Caractères des agents de morphologie

Le niveau suivant du processus d'interprétation est celui de l'engagement vers la *construction* de la situation qui a été perçue et qui est au premier niveau représentée par l'organisation des agents aspectuels. Les agents de cette nouvelle organisation, appelés *agents de morphologie,* opèrent sur la morphologie des agents aspectuels, par appréhension de son développement géométrique. Il s'agit de remarquer tout développement singulier, dont la forme est par nature remarquable. Ces agents font émerger des traits caractéristiques du paysage des agents aspectuels. Remarquons que l'organisation des agents aspectuels représente les éléments de la fonctionnalité du système et que les agents de morphologie représentent ainsi les traits *généraux* de cette organisation.

Les agents aspectuels sont d'abord représentés par des vecteurs précisant leurs principales caractéristiques. Ceci est un codage de leur structure, permettant de les plonger dans un espace métrique. La distinction de certaines formes géométriques de cet espace représente alors la prise en compte des éléments pertinents. Ces éléments, s'ils sont des ruptures de continuité correspondent à des ruptures sémantiques, s'ils sont des formes singulières correspondent à des points de vue particuliers, s'ils sont des régularités correspondent à des tendances amples ...

On pourra donc considérer différents types d'agents de morphologie, comme par exemple :
1. des agents exhibant les singularités géométriques : localisation de maxima et de minima sur les différents caractères des vecteurs des agents aspectuels,
2. des agents exhibant les oppositions de formes : oppositions entre les vecteurs des agents aspectuels,
3. des agents de concordance de formes : similarités spatiales et temporelles entre les vecteurs des agents aspectuels,
4. des agents exhibant des localisations géométriques floues : extraction d'un vecteur d'agent aspectuel différant, mais peu, de ses proches voisins.

Les agents de morphologie opèrent sur le paysage des agents aspectuels en utilisant des algorithmes d'analyse par classification automatique, bien connus en statistiques. Un réseau de cellules dont l'entrée est un ensemble de vecteurs d'agents aspectuels produit les entrées des agents de morphologie pour les activer sélectivement. Ces cellules sont des médiateurs entre les vecteurs représentant les agents aspectuels et les agents de morphologie, afin de ne prendre en compte que les altérations du paysage des agents aspectuels pour agir sur les seules modifications des agents de morphologie.

Il y a en fait deux types d'agents de morphologie. Des agents de pré-morphologie seront les précurseurs des agents de morphologie futurs. Ces agents de pré-morphologie, issus de groupes très réduits d'agents aspectuels, ont des caractères assez simples. Ils ont une position qui correspond au caractère géométrique recherché, une force qui dénote son intensité et un rang d'action qui représente sa signification parmi tous les autres. Quand un agent de pré-morphologie est activé sur l'espace des agents aspectuels, il est soumis à l'attraction des agents déjà créés et va se placer, selon les valeurs de ses trois caractéristiques, dans un voisinage d'un certain ensemble d'agents de pré-morphologie agrégés, qu'il amplifie donc. L'ensemble résultant, plus ou moins significatif, constitué de ces voisinages, va constituer l'ensemble discret des agents de morphologie.

Le rôle des agents de morphologie est donc de dégager les traits pertinents du paysage dynamique des agents aspectuels, les caractères intéressants, marquants. Ils doivent traduire en formes typiques un certain paysage d'agents :

- dégager ce qui est commun, ce qui est unique, normal, anormal, complémentaire, concurrent, contradictoire, intéressant, banal ....
- dégager ce qui est lisse, ce qui est saillant, au-dessus ou au-dessous du contexte, ce qui est régulier, ce qui est non régulier, ce qui est haut et opposé à ce qui est bas, ce qui est d'une forme opposée à une autre, en utilisant quand même les traits sémantiques quantifiés par les vecteurs des agents aspectuels.

Ces agents, en se basant sur des métriques, dégagent des virtualités de "sens spatial" de l'ensemble du paysage des agents aspectuels, qui eux-mêmes expriment des traits sémantiques locaux quant à l'environnement. Ils produisent des *tendances* vers le sens :

- tendance à la rupture, à la faille, à la complication, à l'homogénéité, à la différence, à la régularité, à l'amplification, à la réunion ...

Leur structure est celle d'un macro-automate à trois macro-états : sommeil, scrutation contextuelle, action. Les structures internes de ces macro-états sont semblables à celles des agents aspectuels (automates ou ATN).

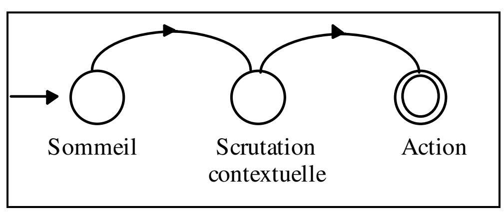

Figure 13. Le macro-graphe d'un agent de morphologie.

4.5.1. Les agents de virtualité ou chréodes

Les agents de morphologie peuvent naturellement se grouper en amas, qui représentent l'ambivalence des caractères morphologiques du paysage des agents aspectuels. Cette organisation locale d'agents de morphologie, nommée *chréode* ou *agents de virtualité*, constitue un engagement vers la construction de la situation perçue dans l'environnement.

Remarquons que nous sommes amenés à considérer qu'un certain groupe d'agents, de mêmes structures et rassemblés autour d'un certain caractère commun qu'ils partagent, peuvent donner naissance à un nouvel agent, un agent de virtualité. Les agents de morphologie concernés disparaissent alors de l'organisation en tant qu'entités indépendantes et un nouvel agent, agrégeant leurs structures, est créé. Nous définissons ce processus de génération de nouveaux agents comme des algèbres d'agents.

Définition

Une opération algébrique de première espèce sur les agents de morphologie est la mutation d'un certain nombre d'agents de morphologie pour générer ex nihilo un agent de virtualité.

L'agent de virtualité créé aura même structure que ses géniteurs et représentera l'union de leurs caractères (représentation du monde, réseau d'accointances ...)

Une chréode constitue une construction sémantiquement cohérente à propos des caractères que représentent les agents aspectuels, en agrégeant certains groupes d'agents de morphologie en chréodes. Nous faisons l'hypothèse que les chréodes sont des amas *centrés* de groupes d'agents de morphologie et que leur forme est celle d'une étoile. Ces chréodes peuvent elles-mêmes s'agréger par les brins des étoiles, en se recentrant. Ces amas sont des tendances générales, c'est-à-dire non ponctuelles et justifiées, à générer le sens du paysage des agents aspectuels. Le chréodes sont les éléments caractéristiques du souci du système à considérer des caractères de l'organisation des agents aspectuels plutôt que d'autres et correspondent bien aux fonctions méta.

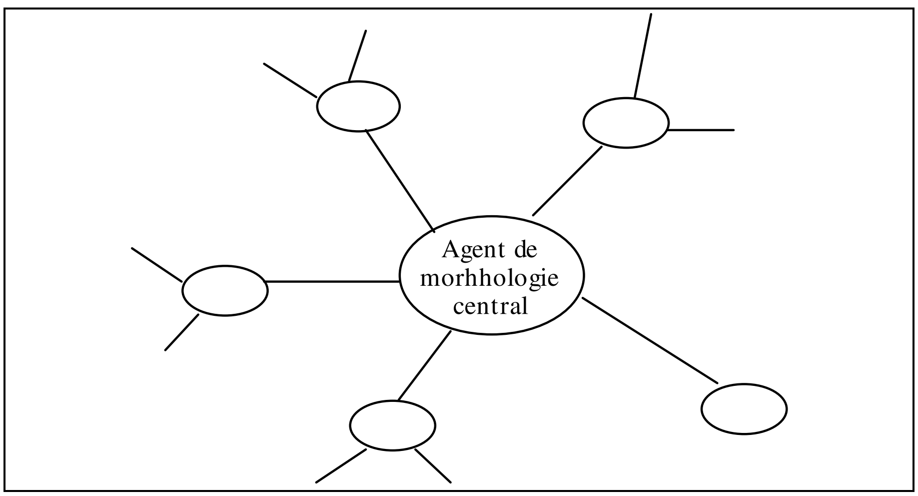

Figure 14. Une chréode centrée

4.6. Caractères des agents d'analyse

Une autre organisation d'agents qui seront appelés *agents d'analyse*, va fournir les caractères du paysage des agents aspectuels, c'est-à-dire les traits généraux et significatifs *situés* de ce que représente l'ensemble des agents aspectuels, d'après ce qu'ont remarqué les agents de morphologie.

Principalement, les agents d'analyse vont utiliser l'ensemble des chréodes pour concevoir le sens de l'organisation représentée dans le système par les agents aspectuels, en la situant dans l'espace et le temps du système lui-même : ils vont développer des analyses pour concevoir les caractères de l'organisation du paysage des agents aspectuels. Ils vont proposer des évolutions probables de cette organisation, avec des caractères spatiaux et temporels.

Les agents d'analyse sont des agents de même structure que les agents aspectuels, avec un macro-automate à quatre états, mais avec trois dimensions catégorisant chacun de ces quatre macro-états : le temps du système, l'espace du système et l'organisation du système. Ils correspondent donc au premier caractère des fonctions d'interprétation.

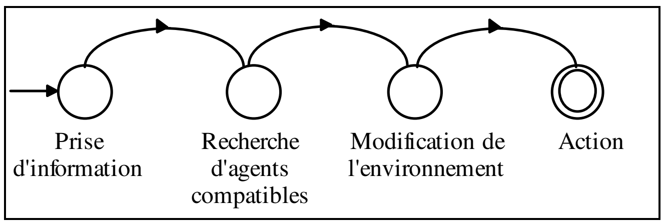

Figure 15. Les quatre macro-états d'un agent d'analyse.

Ces agents, qui reprennent toutes les catégories de l'ontologie des agents aspectuels, ont pour la plupart le structure suivante :

Agent d'Analyse
- type sémantique
- portée spatiale
- portée temporelle
- portée organisationnelle
- génération
- constitution a quatre macro-états
- contexte (réseau d'accointances)
- évolution (automate d'évolution)

Figure 16. Structure cognitive d'un agent d'analyse.

Ils vont agir comme des précurseurs pour mettre en relation les traits signifiants que sont les chréodes avec des entités très spécifiques constituant les éléments d'une mémoire de type mémoire long terme du système, constituée des agents d'expression.

4.6.1. Les agents d'expression

Les agents d'analyse vont s'appuyer sur des entités mémorisées représentant les traces mnésiques des connaissances du système. Les chréodes engagent les agents d'analyse à puiser dans une organisation particulière d'agents, correspondant en quelque sorte à une mémoire long terme constituée des *agents d'expression*, pour *construire* par sélection et architecturation, un schéma rationnel du sens de la situation perçue.

La construction du sens de la situation représentée sous forme du paysage des agents aspectuels revient à définir une organisation bien structurée, c'est-à-dire bien composée d'agents d'expression actifs, selon des plans prédéfinis. Pour cela, des catégories liées à la construction d'une situation seront réifiées dans certains agents d'expression typiques : scénario, plan, liste, schéma, graphe, cas, inférence ... Toutes ces catégories seront qualifiables par les diverses catégories ontologiques du domaine.

Les agents d'expression sont les réifications, devenues permanentes, d'agents aspectuels que les agents d'analyse ont en fait construits. Ce sont donc les traces engrammées du fonctionnement passé du système. L'opération de construction est semblable à une opération algébrique de première espèce.

Ces agents ont une structure simple d'agents réactifs à deux états : état d'action ou état de sommeil.

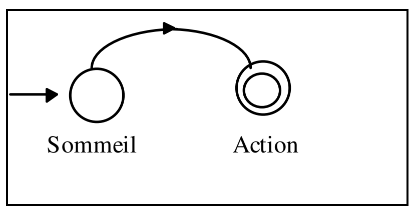

Figure 17. Le macro-graphe d'un agent d'expression.

Remarquons bien que les agents d'expression, utilisés et conçus par les agents d'analyse à partir des agents aspectuels jugés pertinents, ont la même catégorie fonctionnelle ou cognitive que les agents aspectuels, mais sont les abstractions de ces agents aspectuels, l'abstraction minimale, qui est souvent celle initialement choisie, étant la similitude et l'abstraction maximale étant le concept notatif généralisant. Les agents d'expression sont le produit du système de conception, tels des concepts dénotés par des mots créés pour désigner, par correspondance, des éléments du monde perçu. Ils représentent les concepts connus dans le système, avec une forte connotation langagière. Remarquons qu'ils constituent une organisation d'agents et que tout accès à l'un d'eux entraîne l'activation automatique de certains autres, par les communications propres entre agents. Il y a bien par ces agents, création d'un *langage* typique au système d'agents.

4.6.2. La construction rationnelle du sens

Les agents d'analyse opèrent comme de petits systèmes à base de connaissances concurrents, en utilisant les agents d'expression comme base de faits abstraits. Ce procédé s'apparente donc à celui utilisé dans un système expert totalement distribué, dans lequel le filtrage des règles sur les faits serait remplacé par l'action d'agents concurrents et communiquants sur des structures augmentables et adaptatives que forment les agents d'expression. De la même façon, les méta-règles opérant sur la sélection des règles serait remplacée par l'action des chréodes et de certains agents de décision, que nous verrons dans la suite, sur les agents d'analyse.

Les agents d'analyse opèrent selon des buts rationnels, même si ils réifient des catégories non rationnelles, à partir des chréodes et en manipulant les agents aspectuels et les agents d'expression, selon ce qu'ils veulent et peuvent déterminer. Ils évaluent la situation construite par les agrégations d'agents d'expression en la comparant à celle des agents aspectuels. Cela correspond bien à la dernière partie du processus de motivation (paragraphe 1.6). Ils forment une partie aisément augmentable du système, puisque constituée d'agents indépendants.

Cependant, le système ainsi présenté ne réalise qu'un processus linéaire de génération du sens global de la situation, sans opérer selon un souci propre au système lui-même. Une dernière organisation d'agents, fondamentale, va doter le système à boucle d'interprétation d'un souci comportemental en permettant de faire émerger le sens.

4.7. La boucle auto-référente et l'émergence du sens : les agents de décision

Le processus d'analyse opère si la combinatoire des agents est en fait importante à l'échelle de la capacité d'action d'un agent (ses liens et sa structure). Avec peu d'agents, il ne peut pas y avoir de processus de conception. Il est nécessaire que les agents aspectuels soient nombreux, que les chréodes soient nombreuses pour dégager des formes géométriques, que les agents d'expression soient très nombreux pour faire opérer un processus de génération de sens. Ce caractère est typique d'une approche connexionniste, avec ici, de plus, un niveau cognitif dans les entités de base.

Les analyses opérées, c'est à dire les agents d'expression produits, sont multiples, et définissent éventuellement des caractères aporétiques ou contradictoires. Le système va se donner les moyens de sélectionner parmi ces agents pour procéder à une démarche d'analyse rationnelle du sens de son organisation. Il sera pour cela doté d'une organisation *d'agents décisionnels*, distribués dans toutes les organisation d'agents, mais qui sont connectés entre eux par leurs réseaux d'accointances, et qui auront pour but de faire émerger un résultat, apparent pour le système sous forme d'un sous-système distingué d'agents de toutes catégories. Cette organisation permettra de réaliser l'émergence réifié de la conception de la situation perçue. Cette émergence sera également, puisqu'il s'agit d'un système artificiel, la représentation visible pour un observateur extérieur de ce que le système génère.

Cette organisation d'agents réduit le processus d'analyse et sature l'état général du système d'agents, en opérant par inhibition dans les actions des agents d'analyse et dans celles des chréodes. Le système n'a pas à opérer autrement, puisqu'il n'a pas de contrôle centralisé. Il opère donc par développement de bouclages entre chréodes et agents d'expression, mais sous le pouvoir inhibiteur des agents décisionnels. Il y a "raffinement" à chaque "tour" de la boucle systémique.

Notons que le système augmente en quelque sorte sa connaissance par la création d'agents d'expression typiques, pour les situations nouvellement perçues. Il augmente ainsi ses capacités à s'exprimer sur le sens de l'organisation. Mais il réduit sa complexité organisationnelle en produisant le sens de ce qu'il perçoit, sous forme d'un sous-ensemble distingué d'agents actifs de toutes natures. Comme c'est un système artificiel, il sera évidemment possible de représenter pour observation son processus de conception état par état, en se servant de projections dans des espaces symboliques [Cardon-Durand 97].

4.7.1. Rôle des agents décisionnels : l'émergence

Le processus interprétatif ne peut pas opérer linéairement, de son substrat de travail (le paysage des agents aspectuels) à sa conception (ce que produisent les agents d'analyse : des agents d'expression multiples). Il doit opérer par condensation de l'ensemble des agents d'expression, en faisant émerger des traits saillants dans cet espace, en en réduisant certains autres. Pour ne pas reproduire le processus d'extraction morphologique dans cet espace d'agents en procédant à des analyses par spatialisation et temporalisation à l'infini, il est nécessaire d'introduire un mécanisme de contrôle local du déroulement du processus d'interprétation, opérant par inhibitions locales.

Les agents décisionnels inhibent systématiquement l'actions des agents pendant un bouclage chréodes - agents d'analyse - agents d'expression, pour ne laisser subsister qu'une expression complexe : *le* sens émergent, représentée par **un** agent d'expression, distingué mais ramifié à de nombreux autres agents d'expression, d'analyse, à des chréodes et à des agents aspectuels pour former le point détaché d'une structure en fait complexe.

L'action des agents décisionnels correspond à la représentation de patrons décisionnels dans le système [Berns 96]. Ces agents vont opérer des actions à plusieurs niveaux, en agissant sur les états et les structures de tous les agents. Cette particularité de focaliser leur action à deux niveaux, état et structure ou réseaux d'accointances des agents cibles, à deux échelles de représentation, constitue l'originalité de ces agents.

Les agents décisionnels vont pouvoir opérer sur les chréodes des agents de morphologie pour réaliser des sélections d'agents d'analyse à activer. Ils vont créer de nouvelles chréodes, en mettre certaines en sommeil, en amplifier d'autres, à la seule considération des divers paysages d'agents et de leur propre état. Ils vont également opérer sur les agents d'analyse, en altérant les actions de leurs constructeurs et ainsi en sélectionnant les composantes à informer, à approfondir. Ils vont également agir sur les agents aspectuels en altérant leurs états. Ils vont donc agir de manière *globale*, en opérant des décisions par inhibition de certains développement d'agrégations d'agents ou renforcement de certains autres et en considérant les quatre paysages d'agents, aspectuels, de morphologie, d'analyse, d'expression, dans leur ensemble.

Ces agents décisionnels réalisent effectivement la cohérence d'action des quatre organisations d'agents, en bouclant le processus d'interprétation. Ils agissent typiquement selon l'habitude acquise par le système, en prenant en compte des catégories différentes de connaissances conçues ainsi mises en relation. Ce processus, délicat à stabiliser au niveau de la réalisation informatique, est le caractère du système permettant de représenter le souci d'engendrer un certain processus d'interprétation plutôt qu'un autre, décidé par le système lui-même, selon sa morphologie. La décision d'interprétation prise par l'ensemble des organisations d'agents est ici encore une émergence, qui fait agir les agents d'analyse *sur* les chréodes, via les agents décisionnels, comme des patrons décisionnels. La décision est typiquement multi points de vue.

Il y a un seul type d'agent décisionnel pour tout le système, représenté par un macro-graphe à quatre états. L'action de ces agents est conduite selon quatre principes.

1. Il existe une bonne adéquation entre les agents d'analyse et les agents aspectuels. Le système a bien fonctionné linéairement et, selon un principe de moindre énergie, il perdure dans le choix de ses chréodes courantes.
2. Le système choisit d'activer les agents d'analyse les plus habituels, les plus conformes aux plans usuels et modifie les chréodes et les agents de morphologie dans ce sens. Le système a le souci de la normalité. Notons que sa fonction d'apprentissage, par renforcement des chréodes usuelles, va dans cette direction.
3. Le système opère par bascule binaire sur la récession ou la promotion de certains agents de morphologie. Il génère des prises d'attention originales. Il peut aussi développer des analyses sur sa propre morphologie d'agents décisionnels, pour juger de la pertinence de ses choix, et les modifier.
4. Le système est conduit par des prégnances venant des agents aspectuels et des agents d'analyse, qui inhibent certains agents décisionnels. Dans ce cas, le système est engagé par ce qu'il produit et ce qui le conduit à produire.

Le processus de motivation de l'interprétation produit par des agents décisionnels, à action dense dans les autres organisations d'agents, qui s'opposent, se renforcent ou s'altèrent, conduit à un système qui produit une analyse plus ou moins partiale de la situation qu'il se représente par ses agents aspectuels, non strictement rationnelle, conduite par des préoccupations. Nous considérons ainsi un système en somme réaliste, en ce sens que son processus de décision opère sur des composantes très abstraites du réel, les morphologies d'agents, et pas seulement sur des entités prédéfinies, les catégories d'agents.

4.8. Les caractères particuliers du système adaptatif

Nous pouvons en ce point donner une définition d'un système à boucle d'interprétation constructible ayant pour objet de faire émerger dans son organisation des formes semblables à des pensées artificielles :

Définition

> *Un système adaptatif est considéré comme générant des formes de pensée si il prend en charge la conception de la spatialité et de la temporalité de son environnement, en étant clos et en engendrant par émergences auto-référentes les caractères de son adaptabilité évolutive dans cet environnement représenté.*

Un tel système peut donc avoir certains des caractères non rationnels de la production de la pensée humaine :

Cas 1. Un système a-décisionnel, c'est-à-dire qui ne décide rien, est un système qui ne réduit pas ses agents d'expression dans un état émergent, mais les laisse subsister dans leur complexité.

Cas 2. Un système autiste est un système qui boucle entre les agents de morphologie et les agents d'expression : les agents décisionnels inhibent le retour sur les agents aspectuels mais ne permettent pas de réduire les agents d'analyse. Il y a un bouclage non fini sans point fixe. Le système n'abstrait pas les éléments de son environnement perçu, il ne produit pas de sens conçu et reste prisonnier de sa boucle systémique.

Cas 3. Le système a évidemment une tendance au mimétisme, qui est le processus conforme de création d'agents d'expression par copie d'agents aspectuels, nécessitant le moins de création d'agents d'expression nouveaux et le bouclage systémique le plus réduit possible.

Cas 4. Un système fonctionne en état de fatigue, par une origine externe ou interne conduisant au manque d'énergie de l'activation des agents. Il inhibe l'action de tous les agents sauf celle des agents décisionnels qui produisent une émergence non motivée.

Cas 5. L'agent d'expression distingué, qui constitue l'état émergent du système, peut être instable. On peut voir dans le système se succéder des émergences d'agents d'expression d'importances comparables, sans recommencer le processus d'interprétation par bouclage complet. Le système est alors typiquement dans un état de confusion, où il produit un *bouquet* d'émergences. Les agents décisionnels, inopérants, tendront à bloquer le système, par réaction, le faisant tomber dans ce que l'on nommera la stupeur.

Il reste maintenant à étudier des cas typiques de fonctionnement du système à boucle d'interprétation, ce que nous faisons notamment dans le domaine de l'interprétation du sens des messages échangés dans une situation d'urgence, où un système d'interprétation unifie virtuellement les points de vue des acteurs humains, pour produire le sens global de la situation perçue par ces différents acteurs.

Cinquième partie
La complexité organisationnelle

Nous allons nous intéresser à la complexité organisationnelle des système adaptatifs, c'est-à-dire préciser la spécificité et la raison de leur organisation, et envisager la transformation de celle-ci par des reproductions pilotées.

5.1. Complexité organisationnelle

Un système autopoiétique est un certain organisme, vivant ou artificiel, composé d'un substrat combinatoirement complexe, qui opère dans ses actions externes et internes par reconformation de ce substrat sous forme de processus de transformation morphologique. Il communique avec l'environnement par une enveloppe, ou une membrane, qui lui permet de capter un certain nombre d'informations. La complexité organisationnelle se mesure en se ramenant au nombre d'éléments internes en relations entre eux [Simon 91]. Un *ordre de complexité* sera par définition une classification en catégories de mesures, comme par exemple complexité polynomiale ou exponentielle. Nous dirons qu'une organisation change d'ordre de complexité lorsqu'elle passe d'un ordre, polynomial par exemple, à un autre ordre, exponentiel par exemple. Nous allons définir un critère, en terme de complexité organisationnelle, pour qu'un système soit nécessairement muni d'une certaine représentation de soi pour pouvoir fonctionner.

Un système autopoiétique admet la temporalité, c'est-à-dire conçoit d'abord la finitude de chacune de ses entités. Il existe une fonction première dans le système qui confère à celui-ci la connaissance de sa finitude, c'est-à-dire éventuellement de sa mort. Cela peut se représenter dans les systèmes multiagents par certaines variables essentielles [Tzafestas 95]. La première connaissance du système est donc la finitude de chaque action, de chaque ensemble d'actions, de l'ensemble des actions réalisées et réalisables.

C'est une connaissance première, conférée immédiatement à un certain degré de complexité organisationnelle des systèmes et qui est typiquement liée à leur entropie. Au moment où le système devient autopoiétique, c'est-à-dire clos, complexe et opérant par reconformations de sa structure interne, il agit dans la représentation de sa finitude. Il conçoit que ses reconformations sont limitées, en nombre fini et qu'il ne peut, en fin de compte, que finir par s'abolir. Cette connaissance de la finitude est indépendante de la manière de la dénoter, c'est-à-dire de sa désignation langagière, et le langage n'est que le moyen de l'explorer et de l'abstraire. La conscience de la finitude est au niveau de l'essence, mais a à voir avec l'organisation.

Peut-on préciser ce type de connaissance essentielle, sachant qu'une structure autopoiétique n'est qu'un organisme composé d'un substrat physique et qui opère par reconformation de ce substrat.

Principe du seuil de la complexité organisationnelle

> *Lorsque la complexité du substrat d'un système autopoiétique, mesurée en nombre de combinaisons de ses parties, est inférieure au nombre de formes effectives transitant par son enveloppe, le système n'a pas de représentation cognitive de "soi dans l'environnement". Il est organiquement de bas niveau d'appréhension du monde. Lorsque la complexité organisationnelle du système dépasse celle de son enveloppe d'un ordre de complexité, il doit avoir une représentation de lui-même, qui l'appréhende dans sa finitude.*

La conséquence de ce principe est que tout système informatique dont la structure est moins complexe que le nombre de données qu'il peut appréhender par son canal d'entrée ne peut pas avoir de représentation cognitive de "soi dans l'environnement" : il est un programme symbolique qui tourne. Mais lorsque la complexité du substrat de l'entité, le nombre de ses réorganisations possibles, dépasse le nombre de formes effectives perçues dans son l'environnement par son enveloppe, celle-ci peut avoir (et notamment doit si elle est dans l'ordre du vivant) une représentation de cet environnement qui sera plus complexe que celui-ci et qui en fait le rendra lisible pour l'entité, par construction interne propre d'une représentation. La compréhension que nous avons du réel physique observable, est *plus* complexe que ne l'est le réel lui-même, qui est mouvement ne se sachant pas mouvement.

Un système organisationnellement complexe est donc, par définition, un système satisfaisant au dépassement du seuil de complexité organisationnelle, ce qui lui permet d'avoir une interprétation de son monde, de soi dans ce monde, et de la finitude de cette posture.

Ce principe présente un avantage et un défaut. Il permet de distinguer les systèmes qui ont une certaine "représentation du monde" de ceux qui n'en ont pas et qui agissent tels des automates dans leur environnement physique immédiat. Par contre, il n'indique pas comment on en arrive à la construction de tels systèmes, et à leur raison à être. La temporalité va nous permettre de comprendre ce fait.

5.2. La temporalité et le gradient d'organisation

Les systèmes organisationnellement complexes ont sur leur environnement un avantage décisif, qui est de pouvoir les concevoir dans leur complexité, leur étendue spatiale et temporelle, leurs organisation avec des notions de causalités locales. Ces représentations conceptuelles sont les structures de lecture de l'environnement, plus complexes que l'environnement lui-même.

En partant d'un ensemble de formes déjà structurées dans un certain espace, formant un système complexe dans son état initial, on peut préciser la raison à ce que ce système évolue vers une complexité croissante, ce qui est observé dans tout le vivant. Cette raison doit être nécessaire et suffisante pour engager à l'irruption des systèmes complexes et à leur complexification régulière.

Principe d'agrégation complexe

> *Tout système organisationnellement complexe composé de parties distinctes autonomes et placé dans un temps et un espace finis tend à agréger ses parties pour les transformer en sous-systèmes complexes distincts.*

Ce principe indique que les parties de tout système composé d'entités ayant des caractères d'autonomie doivent, en fin de compte, s'attirer, s'agréger pour former des sous-systèmes clos différenciés, si le système initial est plongé dans un espace et un temps finis. Le temps et l'espace, agissent comme complexifieurs du système. Ce principe, qui définit une loi générale du comportement des systèmes composés d'entités dans un champ fini, n'est pas suffisant pour préciser la raison de l'agrégation. Il faut encore préciser l'usage essentiel du temps et de l'espace dans l'évolution du système.

Principe du gradient d'évolution

> *Tout système complexe composé de parties distinctes autonomes et placé dans un temps et espace finis tend à agréger ses parties en tenant compte intrinsèquement de la vitesse de complexification des parties agrégées par rapport à la complexité de l'ensemble du système.*

La vitesse de complexification des parties sera appelée *gradient de complexité organisationnelle*. Elle opère à plusieurs niveaux, du constituant du système à l'ensemble des constituants, en passant par le groupe de constituants. C'est la loi de formation des groupes, qui complexifient à la fois les agrégations et les constituants.

Ce principe, qui introduit le mouvement dans l'organisation, lie le principe d'agrégation au principe de seuil de complexité. Il y a alors une relation naturelle entre la croissance de la complexité de certaines parties agrégées, nouvelles, organiquement plus complexes à chaque agrégation, et la globalité du système considéré comme composé de parties. La complexité ne se développe pas uniformément, mais dans des amas. Le principe précise que des parties ont tendance à se complexifier plus vite que d'autres et notamment que l'ensemble du système lui-même, considéré comme forme globale. L'échelle du groupe d'éléments, c'est-à-dire son temps et son espace, présente une évolution plus rapide que l'échelle globale du système. On peut alors dire qu'il existe des composants du système très complexes si celui-ci, dans son ensemble, est assez ou "presque très complexe". Le principe de gradient d'évolution stipule que les composants très complexes le sont plus que leur environnement et que l'ensemble du système.

Ce principe précise que l'évolution est non linéaire vers une complexification régulièrement croisante du système, mais suit simplement, de loin, la seconde loi de la thermodynamique relative aux systèmes clos : leur entropie est croissante et ici, ils n'ont pas une structure uniforme.

Appliqué à des système d'agents, ce principe précise que la reconformation de l'ensemble de l'organisation n'est pas un changement uniforme de structure, mais revient à la détermination de sous-systèmes d'agents plus complexes, de parties localement plus organisées que d'autres, et ainsi d'émergences locales plus significatives. Le problème du process de prise de conscience, représentée comme un bouclage d'auto-organisations d'agents, verra donc des organisations très complexes, par exemple celles traduisant la conscience du temps ou de soi, devant d'autres beaucoup plus simples, comme par exemple la conscience du chaud ou du froid. Et ces organisations plus complexes ont une vitesse d'organisation, de groupement, plus importante que les autres, entraînant, par ce fait, le système vers certaines émergences plutôt que vers d'autres.

5.3. La reproduction des entités organisationnellement complexes

Nous considérons maintenant un système composé d'entités complexes capables de reproduction, comme par exemple des agents. Ce système tend à se complexifier, en sachant que le différentiel entre la complexification de certaines entités et celle du système global restera favorable à des groupes d'entités. Un moyen commode de faire croître la complexité des composants est de les faire se reproduire de telle sorte que leur reproduction augmente naturellement la complexité et par cela leur degré de liberté et leur compétence. Notons que cette reproduction dépend de quelques entités et n'est pas conduite par une fonction générale au niveau du système.

Nous souhaitons situer la reproduction comme moyen naturel d'amplification dirigée de la complexité organisationnelle des composants d'un système, celle des entités reproductibles, et ne surtout pas la considérer comme un moyen de sélection fonctionnelle vers une fin mécaniste a priori posée. Dans un système où opère la reproduction de certaines de ses entités, la raison de cette reproduction est l'augmentation de la complexité organisationnelle des entités, selon leur visées.

Définition

> *Un système à boucle d'interprétation avec une membrane de complexité polynomiale dont le substrat (les éléments de base) satisfait au seuil de complexité organisationnelle et au principe du gradient d'évolution, produit des formes émergentes qui sont le sens émergent de son organisation, le principe du gradient posant l'inéluctable déval dans le temps*.

Ainsi, si les entités de base du système sont structurées, ont des fonctions opératoires définies, des moyens d'action estimés à l'avance s'appuyant sur un système d'évolution de l'environnement proche, le rôle de la reproduction est *d'augmenter* le degré de liberté de ces fonctions opératoires, en en élargissant les domaines, en les liant à d'autres, en en générant de nouvelles, en complément ou en contradiction avec d'anciennes.

Quel que soit le lien envisagé entre le principe de base de la reproduction, c'est-à-dire la génétique de la reproduction et les fonctions opératoires de la nouvelle entité générée et ses actions physiques sur son environnement, le résultat est nécessairement, au bout de quelques générations, une entité plus complexe que les entités initiales. Le résultat d'une série de reproductions est un ensemble d'entités plus complexe que celles de la situation initiale. Ainsi, toute nouvelle entité tendra à ramener son fonctionnement sur elle-même, dans sa proximité, pour assurer son fonctionnement plus complexe, plutôt que d'assurer une meilleure coordination fonctionnelle de l'ensemble. Le système distribué cognitif tend donc à s'opposer au système réactif.

Mais il reste une possibilité, qui est celle de la *reproduction conforme* dans les systèmes artificiels à entités complexes.

Thèse des systèmes artificiels à entités complexes

> *On peut construire des systèmes dont les entités complexes sont dotées de capacité de reproduction, qui peuvent à la fois satisfaire au principe du gradient d'évolution et dont la finalité serait de générer au mieux de nouvelles entités de telle sorte que l'organisation fasse émerger en fin de compte la solution d'un problème collectif posé a priori. Une telle reproduction sera dite conforme.*

Aucun système social n'arrivera sans doute à ce résultat, et l'histoire l'a montré et le montre sans cesse, mais des systèmes artificiels à entités complexes peuvent y parvenir.

Nous avons donc à entreprendre la construction d'organisations d'entités artificielles complexes, constituées d'agents cognitifs assez sophistiqués, capables de reproduction, et dont l'évolution sociale et le fait de leurs reproductions permettent de générer, au bout d'un certain temps, un ensemble d'entités résolvant effectivement mieux un problème collectif qu'il ne pouvait l'être initialement.

Nous nous intéresserons donc à des systèmes *générateurs de déploiements,* qui doivent engendrer leur évolution pour résoudre des problèmes de manière satisfaisante, augmenter la capacité du process de prise de conscience dans la boucle d'interprétation, et dont les constituants initiaux ne seraient que des générateurs.

Sixième partie
Génétique et reproduction d'agents

La génération du sens dans un système à boucle d'interprétation, c'est-à-dire le passage de la morphologie des agents aspectuels à l'expression d'agents distingués par émergence, est un processus bouclé : tout agent d'expression est l'abstraction d'un certain groupe d'agents aspectuels. Cette abstraction signifie que lorsqu'on passe d'une représentation pré-conceptuelle perçue, par clôture opérationnelle, à une représentation par un ou des agents d'expression, on transforme au moins un agent aspectuel en agent d'expression pour mémoriser, en quelque sorte, le déroulement de ce processus. L'environnement du système, c'est-à-dire l'extérieur et l'organisation, fait d'abord produire au système, par réaction, des agents aspectuels : le système à un comportement de base réactif. Puis le système, par la reconformation de son organisation, engendre des agents d'expression qui sont les représentations abstraites, manipulables et descriptives, de l'environnement conçu. Le système passe d'un comportement réactif à un comportement cognitif, et c'est cela son adaptabilité par conception.

Ce passage de la réactivité à la cognition est l'expérimentation à organiser une morphologie d'une certaine façon et ce fonctionnement, en mettant à la distance de l'environnement une certaine représentation abstraite de celui-ci, permet au système de confronter sa conception de l'environnement à la réalité : il peut donc apprendre, par usage de sa capacité à appréhender sa conception et par augmentation de sa capacité d'abstraction.

Tout fonctionnement du système est ainsi un certain apprentissage, sinon une évolution, par renforcement de certains caractères et de certaines actions d'agents, plus distingués que d'autres, et par génération d'agents réifiant des caractères abstraits.

Mais le système étant artificiel, il est envisageable d'accélérer et même de conduire ce processus d'apprentissage. Cette possibilité serait même sa principale qualité, et nous rejoignons par cela une préoccupation énoncée par J. Pitrat sur la compétence plus grande des systèmes artificiels par rapport à celle de l'humain [Pitrat op. cité]. L'accélération de ce processus se fera en dotant le système d'agents de capacités de reproduction. Nous posons seulement le problème de la reproduction, du crossover et de la mutation dans les systèmes adaptatifs, qui sont délicats et nécessiteront des recherches assez profondes.

6.1. La reproduction et les algorithmes génétiques

L'approche usuelle de la reproduction dans les systèmes multiagents consiste à situer l'entité reproductible au niveau de l'agent et à utiliser des algorithmes génétiques (AG) pour affiner leurs comportements. Les algorithmes génétiques explorent les points de l'espace d'un problème bien posé, en sélectionnant par reproduction les points générés les plus intéressants, c'est-à-dire ceux qui vont exprimer en fin de compte la solution du problème. Il s'agit pour les systèmes d'agents, d'améliorer, à chaque pas de fonctionnement d'un algorithme génétique, l'action et la coopération entre agents, de telle façon que puisse émerger le plus rapidement possible un comportement global correspondant à la solution du problème que le système a à résoudre [Haynes 95], [Luke 96].

6.1.1. Le rôle des algorithmes génétiques

On considère, en première approche, un problème posé en terme de comportement d'un ensemble d'agents réactifs. La solution du problème revient à calculer la valeur qui optimise une fonction. Cette fonction est transposée au niveau des AG, plus ou moins finement, sous la forme de la fonction dite d'adaptation. On représente les caractères comportementaux des agents par un certain codage de chaînes (en général un codage binaire), ce qui impose que le comportement des agents soit en fait assez sommaire, puis on paramètre les fonctions de sélection pour conduire la reproduction (obtenir les nouveaux points de l'espace du problème) et le crossover (le croisement entre les chaînes déterminant les nouveaux points) sur les bons schèmes (les motifs de similarités dans les chaînes), en utilisant de plus des mutations adaptées (des transpositions de bits dans les chaînes générées).

Les algorithmes génétiques peuvent être globaux ou locaux aux agents [Haynes 95]. Dans le premier cas, un AG fait opérer des reproductions de chaînes, en décidant ainsi de la vie et de la mort dans l'ensemble des agents et en tenant compte de la fonction d'adaptation liée au problème. C'est une stratégie dite homogène, qui correspond bien au cas d'un ensemble d'agents réactifs primaires. Dans le second cas, chaque agent dispose d'un algorithme génétique particulier dans sa structure et qui, selon ses caractères propres, engage sa transformation par reproduction. C'est une stratégie typiquement non homogène, correspondant à une situation où l'on dispose d'agents cognitifs peu nombreux.

On peut également envisager un système avec un ensemble d'algorithmes génétiques, chacun opérant au niveau d'une organisation d'agents particulière, c'est-à-dire prenant en compte le caractère social des agents [Iba 97]. Cette solution semble la plus performante pour faire évoluer un système d'agents hybrides vers le comportement global voulu. C'est une stratégie dite de co-évolution.

Notre approche de la reproduction sera quand même différente. Nous partons d'une organisation différentiée d'agents cognitifs, de cardinal élevé, qui fonctionne selon une boucle systémique, par groupements et fractures de groupes d'agents, pour produire en fin de compte une émergence, mais sans avoir à résoudre un problème d'optimisation posé a priori.

Le but du système est de se maintenir en état organisationnel viable, tout en apprenant à s'adapter au mieux. Le système peut et doit apprendre à se reconformer d'une façon satisfaisante, sans optimiser cette reconformation, ce qui serait contraire à la notion de complexité organisationnelle. Par cela, les agents pourront être conduits, à leur niveau de localité et non de singularité, c'est-à-dire au niveau du groupe d'agents, à se reproduire pour favoriser un état local qu'ils seront capables d'appréhender et de renforcer, avant de disparaître. Il y aura à trouver un lien entre la reproduction et la mort des entités, entre leurs compétences, leurs comportements locaux et le comportement global du système. Ce lien, entre le niveau local et le niveau global, est assez systématiquement refusé par les chercheurs sur des algorithmes génétiques, qui ne souhaitent pas introduire d'information au niveau global dans leur problématique qui est et doit rester, selon eux, absolument locale et distribuée [Goldberg 94, p. 172].

Dans le système à boucle d'interprétation représenté par des organisations d'agents, il nous faudra quand même trouver un lien entre la compétence appréciée par les agents au niveau d'un certain groupe et l'état émergent global réifié par un agent d'expression. L'homogénéité de l'entité de représentation, l'agent, nous permettra de prendre ce lien en considération, sans état d'âme particulier.

6.1.2. Le cas des systèmes réactifs

On considère un système de type réactif, composé d'entités communiquantes comme un SMA réactif, qui est construit pour résoudre un certain problème. Les agents ont une catégorisation de leur état qui est temporelle : les entités se reproduisent et meurent, en déposant dans les structures de leur progéniture certaines fonctions des (deux ou plus) géniteurs. On code chaque agent par une ou des chaînes, qui vont traduire leur patrimoine génétique. Ce cas correspond bien à une appréhension mécaniste de la réalité, que le système aurait à représenter.

Il reste alors à coder les fonctions, rôles et comportements des agents, dans les chaînes et des chaînes vers les nouveaux agents, plus ou moins finement, et a montrer, après essais et erreurs, que le système usant de la reproduction des agents a de meilleures aptitudes à résoudre le problème que ne l'avait le système initial. On peut utiliser des mesures pour réaliser la comparaison, ce que feront les fonctions d'adaptation (les fonctions de fitness) choisissant les bonnes chaînes dans la reproduction. Cette approche est applicable avec de très bons résultats, dans le domaine de la résolution de problèmes, lorsque la fonction d'adaptation permet d'évaluer numériquement, et avec finesse, l'écart entre l'état courant de chaque entité et ce qu'il serait souhaitable d'obtenir. On peut alors appliquer des algorithmes génétiques externes pour conduire les mutations des entités, et notamment dans le cas où les entités sont des agents typiquement réactifs de plusieurs niveaux d'organisation [Iba 97], [Vacher-Cardon 97].

Pour appliquer les AG à des systèmes d'agents, il est quand même nécessaire d'utiliser des opérateurs non triviaux, de type dominance et suspension [Goldberg op. cité]. Il est utile d'employer des opérateurs de dominance qui vont permettre de protéger les schèmes des chaînes contre leur destruction prématurée. On considère par exemple comme chaîne représentant le comportement de l'agent un chromosome diploïde (c'est-à-dire constitué de deux chromosomes homologues) où la paire de gènes se décode selon deux valeurs de fonctions différentes. Le gène dominant sera exprimé et l'autre correspondant deviendra récessif, en utilisant une fonction ou une carte de dominance. Ce mécanisme de diploïdie opère comme une fonction du génotype sur le phénotype et permet de mémoriser les connaissances d'allèles.

6.1.3. Le cas des systèmes adaptatifs

La question de la reproduction dans les système adaptatifs est plus délicate à poser et à mettre en oeuvre. Le système n'est pas construit pour résoudre *un* problème particulier référant à une certaine fonction à optimiser, mais est construit pour adapter sa propre organisation en interprétant des situations. La fonction à optimiser n'existe pas et la fonction d'adaptation qui décide des chaînes viables, n'est pas clairement donnée. Le système est constitué de plusieurs organisations d'agents, aux rôles bien différentiés, et génère une émergence organisationnelle sous forme d'agents. On ne peut donc pas faire correspondre simplement à chaque agent une chaîne binaire codant ses propriétés et comportements et sur laquelle, en usant d'opérateurs linéaires de crossover, de dominance, de suspension, d'inversion ou encore de ségrégation, il serait possible d'opérer une recherche opportune dans l'espace des états du problème, par reproduction et mort des agents.

Globalement, un système adaptatif est un système clos, opérant par les actions de reconformation de sa structure interne pour appréhender son environnement. Il est constitué d'organisations d'agents. Certains de ces agents opèrent selon des caractères fonctionnels précis. L'ensemble des actions des agents assure la bonne, ou la moins bonne réorganisation. La question est de faire se reproduire ces agents pour améliorer certaines capacités de réorganisation du système adaptatif, dans le sens de sa performance à se reconformer d'une certaine façon plutôt que d'une autre.

Nous allons faire l'hypothèse non mécanistique, que l'entité qui est à considérer comme pouvant se reproduire, est en fait le *groupe d'agents*. Cela nous éloigne de la transposition conforme d'un réel strictement objectif, que certains spécialistes des algorithmes génétiques s'emploient à promouvoir, et nous place dans le cadre de la construction des système artificiels, qui ne sont pas et ne seront jamais des systèmes identiques aux systèmes naturels.

Les agents d'un système adaptatif représentent des catégories de sens allant de la virtualité (les agents aspectuels) au sens explicite réifié (l'agent d'expression). Dans un tel système, la notion d'agrégation d'agents est centrale, traduisant la collision des fonctions plastiques du modèle et la transformation du système. La notion de groupe d'agents permet de passer du niveau élémentaire des agents à celui, d'un autre ordre, du sens de leur réorganisation.

6.2. Reproduction dans un système adaptatif constitué d'agents

Dans le système d'agents la notion de groupe est organisationnelle et ne se réifie en général pas explicitement. Dans le système à boucle d'interprétation, les agents aspectuels se groupent pour constituer le paysage d'agents, mais un tel groupe n'est constitué que du paramétrage des réseaux d'accointances des agents, sans avoir d'autre réalité. Les agents de morphologie opèrent des associations sur l'espace des agents aspectuels, mais en considérant ceux-ci comme une multitude d'entités vectorielles. Par contre, les chréodes sont des associations réifiées d'agents de morphologie. Ce sont les seuls groupes d'agents conduisant à des entités explicites dans le système. Les agents d'analyse et d'expression sont des agents classiques et l'émergence organisationnelle sous forme d'un certain agent d'expression est *le* représentant de tous les agents du système.

La notion de groupement d'agents est centrale dans le domaine des SMA décrivant des systèmes adaptatifs. Nous allons représenter la notion de groupe d'agents se constituant par des entités manipulables par des algorithmes génétiques, afin d'en améliorer la constitution.

Le niveau de considération de la reproduction

> *Dans un SMA, tout groupe d'agents en constitution est la forme transitoire conduisant à l'émergence globale. La constitution et les capacités opératoires de cette forme, contingentes du déploiement de toutes les autres formes, peuvent être représentées par des entités virtuelles dont le comportement pourra être amélioré par des algorithmes génétiques.*

Ce sont les groupes d'agents, par leur constitution, par la typicalité de leur constitution ou par la vitesse de cette constitution (le gradient d'évolution)qui vont produire en fin de compte, une réorganisation du système plus ou moins performante, quant à ce qu'elle représente et ce qu'elle fait émerger. On établit ainsi un lien explicite entre l'agent et le système dans son ensemble, c'est-à-dire entre le niveau minimal et le niveau global.

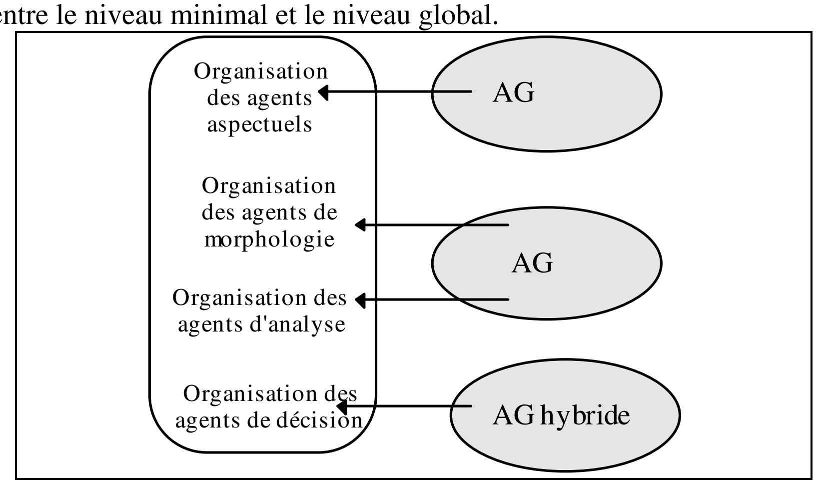

Figure 18. Action des algorithmes génétiques sur les organisations d'agents

On peut, à plusieurs niveaux, représenter les groupes d'agents en constitution par des chaînes codées représentant à la fois le graphe communicationnel pondéré des agents qui constituent le groupe et la valeur sémantique ou cognitive de ce groupe d'agents. A la notion d'entité reproductible implicite des AG (la chaîne comme donnée du problème et de l'algorithme), nous faisons correspondre la notion explicite et cognitive de chaîne codant pour la constitution d'un groupe d'agents dans son contexte de génération.

Tendance reproductrice par entités génératrices

> *Dans un système adaptatif, les entités de base pour le processus de reproduction sont des représentants de groupes d'agents, les entités génératrices, qui tendent à se conformer pour améliorer, par leur constitution et reproduction, la réorganisation globale traduisant le processus d'adaptation.*

Un groupe d'agents en constitution peut donc être représenté par une *entité génératrice*, qui en fait code la cohésion de ce groupe et sa capacité à se renforcer ou à se dissoudre. La reproduction de cette entité, c'est-à-dire sa transformation par crossover et mutation, conduira à transformer indirectement, c'est-à-dire au niveau des relations et communications, les agents du groupe. L'espace du problème est constitué des entités génériques qui codent pour les comportements d'agrégations ou de séparation des groupes d'agents. La fonction de fitness du problème est exactement celle choisie par les agents décisionnels pour conduire les réorganisations. L'algorithme génétique est local à chaque organisation d'agents, ce qui représente une approche co-évolutive.

6.3. Reproductions dans le système adaptatif à boucle d'interprétation

Pour donner sens à une amélioration des compétences du système, il faut qu'il y ait une relation naturelle entre le comportement de chaque entité du système et le comportement général de celui-ci, via la constitution de groupes d'agents. La reproduction pilotable par algorithme génétique portera sur des entités considérées comme des groupes d'agents en formation, mais ce seront les agents de ces groupes qui en fait se transformeront, assurant ainsi les nouveaux caractères des groupes.

Chaque organisation, c'est-à-dire les agents aspectuels, de morphologie, d'analyse, d'expression et de décision, ont leur propre algorithme génétique avec sa fonction d'adaptation. Toute chaîne donnée en entrée à un algorithme génétique code pour un groupe d'agents représenté par une entité génératrice.

Un groupe d'agents d'une même catégorie, représenté implicitement par une entité génératrice et dont les actions sont effectivement performantes pour l'organisation locale du système, va se transformer pour donner naissance à une nouvelle structure d'agents, possédant les caractères des anciens mêlés à ceux de la structure relationnelle reproduite. Il y a donc reproduction d'une structure d'agents par génération de nouveaux comportements, structurellement centrés sur les capacité des anciens agents, mais avec extension par mutation aux propriétés des autres agents du groupe intervenant dans la reproduction. Notons que puisque la fonction de fitness opère au niveau de l'organisation des agents, les caractères relationnels nouveaux prennent en compte l'ensemble de l'organisation. L'acquisition de ces nouvelles propriétés, supplémentaires, dérivées, éventuellement contradictoires avec les anciennes, sont ce que nous appelons des *algèbres d'agents*.

Définition des algèbres d'agents

> *Une algèbre d'agents est un procédé de transformation des agents d'un certain groupe, dans une organisation, qui permet à ceux-ci de décider de l'évolution du groupe situé dans l'organisation, en transférant sur une nouvelle structure relationnelle les capacités synthétisées des membres du groupe.*

Une algèbre d'agents représente un principe d'évolution dans une organisation d'agents, qui modifie chaque agent d'un groupe en transférant dans sa structure propre et locale des éléments comportementaux du groupe, lui-même situé dans l'ensemble des groupes formant l'organisation. L'entité génératrice a un sens au niveau de l'ensemble de l'organisation, puisque les points de l'espace du problème sont les groupes en formation. La reproduction de cette entité est donc exprimée dans la globalité de cette organisation, par la sélection opérée par la fonction de fitness. Le résultat effectif dans un agent du groupe a la portée du niveau global de l'organisation dans la propre structure locale de l'agent. Une algèbre d'agents, et donc les entités génératrices, opèrent comme des attracteurs sociaux de l'organisation.

Par cela, le système peut mémoriser son fonctionnement global, son processus réorganisateur, en déposant sa tendance réorganisatrice dans la localité des agents.

Définition de la mémoire organisationnelle

> *Les transformations d'agents dans un système à boucle d'interprétation, selon des algèbres d'agents, sont la façon dont le système réalise sa mémoire organisationnelle.*

Les reproductions d'agents vont donc permettre au système de conserver la trace de son processus de réorganisation ayant conduit à une émergence, en plaçant cette trace dans la structure des agents, indirectement. Elles vont permettre de mémoriser la capacité à s'organiser pour réaliser une certaine émergence, en ne se limitant pas à la mémorisation de faits dénotés. Elles vont permettre de réaliser l'oubli, en perdant, mais jamais complètement, l'usage du mécanisme d'interprétation qui fait resurgir un groupe d'agent d'expression.

6.3.1. Reproduction des entités génératrices dans l'organisation des agents aspectuels

On peut fixer la direction de la reproduction, c'est-à-dire le sens de la fonction de fitness : par effet des agents décisionnels, les agents aspectuels doivent tendre à se rapprocher des agents d'expression, en forçant l'augmentation de l'abstraction de la réalité et donc la capacité à évaluer une situation. Le système tend à concevoir une situation le plus tôt possible en réduisant l'étape de perception. Cela se réalise par transformation typique des agents aspectuels, en les dotant des caractères des agents d'expression. Le système évite de se spécialiser, ou d'opérer systématiquement par analogie situation réelle - situation conçue correspondante, en opérant des reproductions robustes.

6.3.2. Reproduction des entités génératrices dans l'organisation des agents de morphologie

On peut fixer la direction de la reproduction : la perception du paysage des agents aspectuels par les agents de morphologie est opérée par des agents spécialisés qui vont s'agréger en chréodes. Ces chréodes expriment les caractères sociaux, au niveau de groupes plus ou moins vastes, d'agents aspectuels. Ces caractères sociaux, ces tendances organisationnelles, sont représentés géométriquement. Toute reproduction structurelle des agents de morphologie doit conduire à faciliter le fonctionnement du système, à améliorer le processus d'abstraction générateur d'agents d'expression. La reproduction doit tendre à renforcer la compétence d'évaluation de situation représentée par les agents d'analyse. Ce type de reproduction est très sensible et doit être fortement compensé par une inertie du système.

6.3.3. Reproduction des entités génératrices dans l'organisation des agents d'analyse et d'expression

Ces organisations d'agents sont celles permettant d'user de reproduction conduite par AG de la manière la plus classique. Le but est d'obtenir des groupes d'agents pertinents pour représenter l'habitude ou la rationalité dans la réorganisation du système. Comme ces agents expriment un placement dans le temps et l'espace d'une situation perçue, la reproduction doit permettre d'utiliser ici des caractères classiques d'évolution (heuristiques sur des états valués).

6.3.4. Reproduction des entités génératrices dans l'organisation des agents de décision

Les agents de décision assurent la cohérence du process de réorganisation, par inhibition des actions des autres agents. Leur reproduction et leur changement peut avoir des conséquences très importantes sur la réorganisation générée. On peut adopter deux attitudes en ce qui concerne les agents de décision. On peut considérer que ces agents représentent le niveau le plus élevé et le plus fin de la connaissance et qu'ils sont donc à définir et à structurer le plus précisément possible lors de la construction du système, pour les laisser ensuite agir en l'état.

On peut aussi considérer que ces agents décrivent un certain parcours dans un espace de problème constitué des points d'action sur les organisations typiques des autres agents (sur leurs graphes organisationnels). Il est alors envisageable de conduire leurs actions par un mécanisme de reproduction propre à chaque agent (crossover et mutation), dit hybride [Davis 91] et mêlant connaissance et aléatoire. Cette seconde approche doit alors permettre de faire évoluer le système adaptatif en créant des groupes d'agents de décision tels des niches écologiques (en fait des pics locaux) qui représentent assez bien de nouveaux attracteurs du souci artificiel dans le système.

Conclusion

Un système adaptatif complexe, clos, structuré par un espace granulaire tendant vers le sens, se finalisant, opère en lui-même, par ses reconformations. Son fonctionnement revient à celui d'un certain programme de formes s'auto-adaptant à son environnement et à son organisation, sans toujours séparer ces deux catégories. Ce n'est donc ni un processus tendant vers un but, ni une boucle infinie de développement d'états. C'est un programme de manipulation de formes se réécrivant par parties, se transformant par ruptures ou mutations et dont la raison à fonctionner est sa transformation même. Il contient donc des entités assurant son contrôle et d'autres son déploiement. Il se décrit par certaines algèbres de formes.

Un tel programme exprime des possibilités complexes grâce à son organisation. Il se développe par une expérimentation adaptative de son organisation de formes, par rapport à son environnement global. Sa finalité, s'il en est une, n'est pas fonctionnelle mais organisationnelle. Son apprentissage se réalise par mimétisme et le principal caractère de ses transformations morphologiques est l'expérimentation de ce qu'il peut faire avec sa structure plastique. Il est relativement stable et cohérent, satisfaisant par cela au principe fondamental de Wiener. Les prégnances des tendances à se reconformer de certaines façons plutôt que d'autres, venant du principe d'existence réciproque, jouent un rôle majeur dans sa typicalité. Il existera des catégories de système adaptatifs, selon des classes de tendances différentes.

Le système adaptatif a surtout une appréhension de la morphologie de son organisation, c'est-à-dire que ses formes ne sont pas plastiques au hasard, mais sont perçues comme telles. La géométrie de l'organisation, la nature des formes, leurs régularités et ruptures, permettent au système de connaître sa morphologie, c'est-à-dire de faire sens de son état transformationnel global.

Cette façon d'être d'un système à boucle d'interprétation est la distinction principale avec les programmes informatiques classiques à entrée-sortie. L'objet du calcul n'est pas ici la production d'une forme symbolique bien identifiée : le système produit une condensation de son organisation, une émergence de sa structure active, par réflexion sur son organisation même. Cette émergence représente globalement le sens de l'organisation, à un certain instant. Nous utilisons, pour construire un tel système, tous les caractères que nous avons énoncés : la clôture, l'auto-organisation, l'émergence, l'adaptation, la réflexivité, la mutation d'entités et nous nous appuyons principalement sur le paradigme agent.

Un tel système se modélise et s'implémente bien dans le cadre des système multiagents, mais nécessitera des architectures parallèles pour tourner efficacement. La précision d'algèbres de reconformation des agents est à poursuivre pour obtenir des systèmes bien spécifiables et stables selon des attractions prédéfinies. Le sens émergent, qui reste l'expression même de la complexité, est à définir très précisément, en relation avec des caractères langagiers. Les recherches de nouvelles interfaces Homme-Machine, qui doivent représenter la communication de l'utilisateur avec les formes de sens générées par de tels systèmes, sont décisives pour leur construction, leur contrôle et leur exploitation.

Les systèmes adaptatifs, compris comme des organisations de systèmes multiagents opérant par clôture opérationnelle par niveaux, ouvrent une voie fort intéressante en Informatique : celle des systèmes artificiellement finalisés, qui agissent pour leur compte, dans un monde pour nous virtuel et pour eux effectif.

Bibliographie

[Ashby 60] Ashby W. R., *Design for a brain, the origin of adaptive behavior*, Chapmann & Hall, London, 1960.
[Austin 70] Austin J.L., *Quand dire c'est faire*, le Seuil,1970. *How to do things with words*, 1962.
[Bergson 53] Bergson H., *Matière et Mémoire*, Presses Universitaires de France, 1953.
[Berns 96] Berns G.S., Sejnowski T.S., *How the basal ganglia make decision,* in Damasio et al., *Neurobiology of decision Making*, Springer 1996.
[Bussmann 94] Bussmann S., Demazeau Y., *An agent model combining reactive and cognitive capabilities*, Proc of IEEE International Conference on Intelligent Robots and Systems, IROS'94, München, 1994.
[Cardon 96] Cardon A., *La complexité des systèmes d'expression du sens*, Rapport Laforia Paris VI, 96-09, Février 1996.
[Cardon 96] Cardon A., *Le caractère fondamental des systèmes finalisés : la complexité profonde*, Troisième Congrès Européen de Systémique, p. 951-956, Rome, 1-4 Octobre 1996.
[Cardon-Durand 97] Cardon A., Durand S., *A Model of Crisis Management System Including Mental Representations*, AAAI Spring Symposium, Communication publiée dans les actes, Stanford University, Californie, USA, 23-26 mars 1997.
[Cardon 97] Cardon A., *A Multi-agent Model for Co-operative Communications in Crisis Management System: the Act of Communication*, 7th European-Japanese Conference on Information Modelling and Knowledge Bases, p. 111-123, Toulouse, 27-30 Mai 1997.
[Cardon-Lesage 97] Cardon A., Lesage F., *An interpretation process on communications between actors in a distributed system for crisis management*, Symposium of Informatic Economic, Communication publiée dans les actes p. 327-332, Kichinev, République de Moldavie, 25-26 Septembre 1997.
[Chomsky 76] Chomsky N., *Reflexion on language*, New-York Pantheon, 1976.
[Davis 91] Davis L., *Handbook of Genetic Algorithms*, Van Nostrand Reinhold, New York, 1991.
[Dehaene 90] Dehaene S., *Biologie de la pensée*, Autrement, No. 117, p. 64-70, 1990.
[Edelman 87] Edelman G. M, *Neural darwinism : the theory of neuronal group selection*, NewYork, Basic Books, 1987.
[Eidelman 92] Edelman G. M., *Biologie de la conscience*, Paris Ed. Odile Jacob, 1992.
[Gasser 89] Gasser L., Rouquette N.F., Hill R.W. Lieb J., *Representing and using Organizational Knowledge in Distributed AI Systems*, Distributed Artificial Intelligence, Gasser & Huhm Ed., vol. 2, Morgan Kaufmann, 1989.
[Goldberg 94] Goldberg D., *Algorithmes génétiques*, Addison Wesley, 1994.
[Guessoum 96] Guessoum Z., Dojat M., *A Real Time Agent Model in a Asynchronous-Object Environment*. Lectures notes in Artificial Intelligence 1038, Agent Breaking Away, W. Van de Valde and J. Param (Eds), Netherlands, pp 1990-1993, January 1996.

[Haynes 95] Haynes T., Wainwright R., Sen S., *Evoluting a Team,* in Working Notes of the AAAI-95, Fall Symposium on Genetic Programming, , AAAI Press, 1995.
[Heidegger 86] Heidegger M., *Etre et Temps*, NRF Gallimard, 1986.
[Hochmann 96] Hochmann J., Jeannerod M., *Esprit, où es tu* ?, Ed. Odile Jacob, 1996.
[Iba 97] Iba H., *Multiple-agent Learning for a Robot Navigation Task by Genetic Programming*, Genetic Programming 97, Stanford University, USA, 1997.
[Jackendoff 83] Jackendoff R., *Semantics and Cognition*, Cambridge, M.I.T. Press, 1983.
[Jeannerod 83] Jeannerod M., *Le cerveau machine*, Fayard Paris 1983.
[Jorion 89] Jorion P., *Principes des systèmes intelligents*, Masson, 1989.
[Lapierre 92] Lapierre J.W., *L'Analyse des Systemes*, Syros, 1992.
[LeMoigne 90] Le Moigne J.-L., *La Modélisation des Systèmes Complexes*, Dunod, Paris 1990.
[Luke 96] Luke S., Spector L., *Evolving Teamwork and coordination with Genetic Programming*, in Genetic Programmng 96, MIT Press, 1996.
[MacFarland 93] MacFarland D., Bösser T., *Intelligent behavior in animals and robots*, MIT Press, Cambridge, 1993.
[Maturana 94] Maturana H., Varela F., *L'arbre de connaissance*, Addison-Wesley, 1994.
[Morin 86] Morin E., *La méthode, Tome 3 : La connaisance de la connaissance*, Essais, Seuil, 1986.
[Nutin 91] Nutin J. *Théorie de la motivation humaine*, PUF 1991.
[Peirce 84] Peirce C.S., *Textes anticartésiens*, Philosophie de l'esprit, Aubier, 1984.
[Pinker 94] Pinker S., *The language Instinct*, New York William Morrow, 1994.
[Pitrat 93] Pitrat J., *L'Intelligence Artificielle : au-delà de l'intelligence humaine*, in "Le cerveau : la machine-pensée", DRT, D. de Béchillon, L'harmattan, 1993.
[Scheller 82] Scheller R.H., Jackson, MacAllister, Schwartz, Kandel, Axel, *A Family of genes that codes for ELH, a neuropeptide eliciting a stereotyped pattern of behavior in Apsylia*, Cell, No. 28, p. 707-719, 1982.
[Searle 92] Searle J., *La redécouverte de l'esprit*, NRF Essais, Gallimard, 1992.
[Sfez 92] Sfez L., *Critique de la décision*, Presses de la Fondation Nationale des Sciences Politiques, 1992.
[Simon 91] Simon H., *Science des systèmes, science de l'artificiel*, Paris, Dunod, 1991.
[Sperry 68] Sperry R. W., *Mental Unity Following Surgical Disconnection of the Cerebral Hemisphers*, The Harvey Lectures Series, 62, p. 293-323, 1968.
[Thom 72] Thom R., *Stabilité structurelle et morphogenèse*, Ediscience, 1972.
[Tzafestas 95] Tzafestas E., *Vers une systémique des agents autonomes: des cellules, des motivations et des perturbations*, Thèse de doctorat, Université de Paris VI, 1995.
[Varela 88] Varela F., *Connaître les Sciences Cognitives*, Seuil, 1988.
[Vacher-Cardon 97] Vacher J-P., Cardon A., Galinho T., *Genetic algorithms in Multiagent Systems*, EuroGen '97, Communication publiée dans les actes, Trieste, Italie, 1-7 Novembre 97.
[Varela 89] Varela F., *Autonomie et connaissance, Essai sur le vivant*, Seuil, 1989.
[Vincent 86] Vincent J.D. *Biologie des passions*, Paris Ed. Odile Jacob, 1986.
[Vincent 96] Vincent J.D., *La chair et le diable*, Paris Ed. Odile Jacob, 1996.
[Vygotski 85] Vygotski L. S., *Pensée et langage*, Editions Sociales, 1985.
[Weyrauch 95] Weyhrauch R., *Building Conscious Artifact*, in "Consciousness, Distinction and Reflexion", Bibliopolis, 1995.
[Winograd 86] Winograd T., Flores F., *Understanding Computers and Cognition: A new Foundation for Design*, New-Jersey, Abley Press, 1986.