

MODELISER ET CONCEVOIR UNE MACHINE PENSANTE

APPROCHE CONSTRUCTIBLE DE LA CONSCIENCE ARTIFICIELLE

Alain CARDON, 2004, Coll. Automates intelligents, Vuibert, Paris

Note de lecture de Gérard Sabah

Mél : sabah@limsi.fr

Le livre d'Alain Cardon relève d'une réflexion approfondie sur la question de l'intelligence et de la conscience. La construction méthodique qu'il propose — en 550 pages — vise à montrer le caractère automatisable de ces facultés, en principe réservées à l'Homme.

Parfois ardu, parfois manquant de justification (il s'agit d'une construction montante, dont l'aspect suffisant est généralement convaincant, sans que la nécessité des propositions avancées soit toujours évidente...), ce livre est tout à fait fascinant par son ampleur, courageux par les thèmes qu'il aborde et constructif par ses propositions concrètes et solides.

Les principales hypothèses qui fondent l'ensemble sont : *une pensée artificielle est calculable ; elle nécessite l'interaction physique d'un corps matériel avec son environnement.*

Pour justifier sa première affirmation, l'auteur met en avant divers arguments opposés aux trois raisons qui, à son avis, peuvent faire croire à l'impossibilité de mécaniser la pensée :

- Les croyances religieuses et spirituelles,
- La physique (par exemple si le continu de la matière et du temps était nécessaire),
- Même si la pensée est constituée de processus discrets finis non continus, sa complexité pourrait être inatteignable pas les ordinateurs actuels.

La seconde hypothèse est argumentée en se référant à Damasio : un organisme ne peut penser que s'il est d'abord capable d'éprouver sans cesse des émotions relatives à l'état et à la posture de son corps physique (Damasio, 1999).

En résumant le modèle de départ de façon excessivement caricaturale, on peut dire que les composants de base sont des agents actifs, proactifs (capables, *de leur fait propre*, de mouvements et de communications), symboliques, évolutifs et communicants ; ils sont regroupés en ensembles et sous-ensembles (nommés *agents aspectuels*), eux-mêmes communicants, au fonctionnement automatique, inévitable.

Un autre ensemble d'éléments représente l'état et le fonctionnement des agents aspectuels (les *agents morphologiques*). Ces deux systèmes sont co-actifs et

s'influencent réciproquement. La stabilisation de l'ensemble correspond à un « état de pensée ».

L'objet de ce livre est de proposer une architecture informatique qui permettrait de développer un tel système. Dans ce dessein, est affirmée la nécessité d'un robot élaboré et d'un système de calcul massivement parallèle. Le livre se présente en quatre parties : des généralités sur les modèles et la complexité ; l'explicitation de relation entre pensée, émotion et conscience, et les réalisations informatiques pertinentes ; les modélisations des notions d'émotion artificielle et de « proto-soi » ; les modélisations de la pensée artificielle et du « soi ». J'ai tenté la gageure d'en faire un résumé en quelques pages, émaillé de quelques commentaires personnels **(en gras dans le texte)**.

1. Première partie (définition de quelques concepts essentiels)

Un phénomène correspond à l'observation d'une forme et de sa transformation dans l'espace spatio-temporel tandis que les effets correspondent à des caractères qui permettent de distinguer un phénomène d'un autre. Un modèle d'un phénomène est un système formel (un ensemble d'états et de règles de changements d'état) qui le décrit **[la définition n'est pas si claire qu'elle le semble, ce dernier terme n'étant pas précisé...]**. On distingue les modèles prédictifs (permettant de représenter l'état du système à n'importe quel instant), les modèles discrets, les modèles équationnels...

Les modèles formels rencontrent de grandes difficultés pour représenter les phénomènes complexes où existent d'innombrables interactions. Un autre type de modélisation est alors nécessaire : *le traitement effectif* (il s'agit d'un modèle d'exécution, de fonctionnement du phénomène, fondée sur une analogie de comportement : on ne décrit pas entièrement le phénomène, mais on exprime sa dynamique de fonctionnement). La complexité se représente ainsi par du calcul effectif et non par la syntaxe du modèle.

Différents types de *systèmes* (un ensemble de différents modules qui interagissent), sont considérés : les systèmes fermés (complètement décrits par leurs fonctionnalités) où l'ordre des modules est prédéfini, les systèmes instables (dont l'état du système ne peut être prévu à partir des seules conditions initiales), et les systèmes ouverts (la fonction de leurs composants n'est que partiellement déterminée à leur création, l'ordre de déclenchement résulte essentiellement des interactions coordinatrices entre les composants et non pas de leurs fonctionnalités).

[Ici, l'affirmation : « un système ouvert conduit à la recherche d'états stables pour n'être ni aléatoire, ni chaotique (et cela est une caractéristique fondamentale du vivant) » demanderait à être mieux justifiée (surtout la parenthèse) ...]

Les *systèmes complexes virtuels* sont des modèles de ce type, structurellement cohérents avec le phénomène et qui décrivent effectivement son évolution. Ils se caractérisent par :

- un changement continu des relations entre leurs composants,

- l'impossibilité d'observer les éléments constitutants sans influencer sur leur comportement.

Dans un système, plusieurs types d'entités se distinguent par leurs rôles fonctionnels : les *entités de base* (au comportement prédéfini), les *entités informationnelles* (envoyées et reçues par les entités de base pour les informer des actions en cours) et les *entités régulatrices* (qui observent les comportements des entités informationnelles).

Les entités informationnelles permettent un couplage fort (mais variable dans le temps) entre les entités de base et les entités régulatrices. Ces dernières permettent de modifier les comportements locaux du système en tenant compte de son fonctionnement passé.

L'environnement est représenté par l'ensemble des stimuli qu'il produit sur le système.

La *modélisation de phénomènes complexes* est ainsi fondamentalement différente des modélisations équationnelles :

- Définition des éléments générateurs des mouvements de base du phénomène.
- Définition des règles de communication et de synchronisation entre ces éléments.
- Mise en mouvement des éléments de base (évolution temporelle du phénomène).
- Observation de ces mouvements.

Au niveau de l'implémentation également, l'approche se distingue du génie logiciel classique (analyser, concevoir puis programmer un système opérationnel et mis en fonctionnement une première fois) : ici, l'auteur propose un système opérationnel dès le début, dont le composant de base doit être capable de se spécialiser continûment par différenciations, et qui s'adjoint des agents adaptatifs en fonctionnant. Il nomme ce mode de programmation : *agentification incrémentielle par fonctionnement*.

Un tel système est constitué d'ensembles de processus récrivant des programmes de machines de Turing (il n'est pas réductible à un seul programme, car il est incapable de déterminer lui-même son état suivant). Pour que ce type de système se stabilise après une perturbation, il est nécessaire qu'il soit composé de très nombreux processus.

[Tous ces éléments sont passionnants, de mon point de vue probablement vrais, mais ne sont pas toujours réellement justifiés...]

2. Deuxième partie

Centrée sur la construction d'une pensée « artificielle », cette partie caractérise la pensée humaine par l'auto-observation et par le fait qu'elle permet d'énoncer (et de croire à) des choses non prouvables ; imagination et irrationalité en sont des bases essentielles !

La différence entre le cerveau et la pensée est ramenée à la différence entre une structure très plastique et sa mise en mouvement très dynamique. La construction d'une pensée artificielle est alors fondée sur une hypothèse de compensation (les

agents informatiques sont beaucoup moins nombreux mais nettement plus complexes que les neurones) et sur une hypothèse géométrique (l'activité des agents *représente* le flux de l'influx nerveux).

Une analyse succincte du système nerveux débouche sur une première approche d'une émotion (non consciente, considérée comme une réponse comportementale automatique visant à s'adapter à la situation), d'un sentiment (conscient, déclenché par une émotion), de la connaissance d'un sentiment (une première approche de la conscience) et du calcul d'un plan d'action (raisonnement). La *conscience noyau* est définie comme la perception de soi-même en état d'appréhension du monde [analogie forte avec la conscience de premier niveau d'Edelman (Edelman, 1992)]. La *conscience étendue* est alors la mise en situation de l'individu dans la temporalité.

La suite de cette partie concerne essentiellement des questions d'informatique et de morphologie. L'auteur insiste sur le rôle *actif* des agents et sur les possibilités de reconfiguration qu'offrent les systèmes multi-agents. Il définit les *agents aspectuels* : des entités sociales, réactives et proactives, percevant leur environnement et communiquant par un langage non ambigu. Il propose un système où un ensemble de telles entités a en plus la possibilité *d'engendrer* de nouveaux agents, et où chaque sous-groupe évalue sa cohérence et son adéquation par rapport à l'ensemble, de façon dynamique et continue.

[là aussi, j'ai été légèrement gêné par le manque de définition précise de l'« environnement » et de la forme que peut prendre sa représentation ; la méthode utilisée pour engendrer de nouveaux éléments n'est pas précisée non plus. Enfin, j'aurais aimé voir discuté l'évaluation dynamique de la cohérence de façon plus approfondie : je crois cette proposition inutilement complexe et peut-être même dangereuse ; elle est probablement nécessaire partiellement mais pas sur l'ensemble du système...].

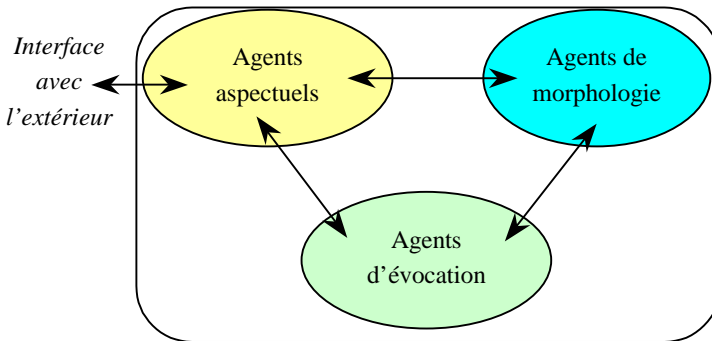


Figure 1 — Les différentes catégories d'agents d'une organisation adaptative :

- *agents aspectuels* : des entités sociales, réactives et proactives, ayant une perception de leur environnement et communiquant par un langage non ambigu ;
- *agents morphologiques* : représentent les activités des agents (des *cartes d'activité*) ;
- *agents d'évocation* : fournissent une vue cognitive des informations sémantiques.

Après avoir souligné l'instabilité d'un système multi-agent massif (fonctionnement non-déterministe, sensible aux conditions initiales), l'auteur cherche à réduire cette instabilité en rendant le système *auto-adaptatif*. Pour cela, il introduit la notion de *carte d'activité* qui permet de représenter l'évolution temporelle des caractères significatifs de l'ensemble des agents. Dans l'espace morphologique défini par ces caractères significatifs, des agents particuliers (les *agents morphologiques*) représentent les activités des agents, grâce à un couplage fort entre agents aspectuels et agents morphologiques. Des regroupements de ces derniers révèlent alors un caractère émergent de l'ensemble ; le fonctionnement de l'ensemble se ramène à une succession d'émergences.

Les *agents d'évocation* examinent ces éléments émergents et en fournissent une interprétation cognitive, ce qui permet un certain auto-apprentissage.

Ainsi, on peut dire que l'activation d'un groupe d'agents aspectuels *représente* un phénomène réel si le groupe est systématiquement activé (grâce à une *raison* interne) quand le système est confronté au phénomène.

3. Troisième partie

Cette partie est centrée sur les émotions artificielles et le « proto-soi » et vise à en proposer une modélisation en rupture nette avec les modélisations classiques.

Pour un tel système, il est nécessaire de disposer des éléments suivants :

- une représentation intentionnelle,
- une représentation de la situation de l'organisme dans son environnement,
- une formulation explicite d'une idée,
- une prise de conscience et une mémorisation de cette formulation.

Caractérisé par un fonctionnement non stable de relations entre sa structure et un environnement, interprété pour permettre des références, un tel système ne dispose pas de représentations figées sous forme classique : toute la pensée est (et n'est que) processus. Il s'agit de systèmes adaptatifs qui contrôlent leur activité selon les sollicitations de l'environnement et qui nécessitent donc une auto-observation (une « conscience de soi »).

[Si je suis d'accord avec ces propositions, il me semble toutefois que tout cela ne représente qu'une catégorie de pensées, il en est d'autres qui ne me semblent pas entrer dans ce cadre ! (rêves, envies...). Par ailleurs, cette conception s'oppose radicalement à celle de Minsky qui propose que quantité d'idées élémentaires soient produites plus ou moins en vrac et que seules celles qui sont cohérentes émergent consciemment. Une discussion sur ces points serait utile ici].

Un corps ne pouvant être un assemblage de parties construites indépendamment (car alors, il faudrait expliciter *toutes* les relations entre *tous* les éléments du système, en particulier au niveau neuronal — ce qui semble totalement impossible), il apparaît nécessaire que chaque partie d'un corps soit construite simultanément aux autres. La structure fondamentale proposée est alors celle du « multi-robot » : ensemble de robots autonomes, spatialement dispersé, contrôlé par un système unique de génération d'émotions et de pensées artificielles.

[Sur ce point, il me semble qu'on trouverait quelques éléments pertinents chez Edelman (qui est par ailleurs peu cité dans ce livre), en particulier à propos de la représentation des connaissances par des éléments minimaux, proposée ici et qui rejoint sa conception dynamique de la mémoire.

L'idée du multi-robot est intéressante, à mon avis probablement viable, mais elle n'est pas vraiment justifiée. Une telle structure est-elle la seule possible ? Par ailleurs, elle pose le difficile problème de l'unité de la perception et de la cohérence de l'ensemble, problème qui n'est pas discuté ici.]

Une émotion est alors définie comme une réaction automatique, non consciente, face à une situation inattendue **[pourquoi forcément inattendue ? certaines situations attendues produisent aussi des émotions]**. Elle demande que soient synchronisées l'activité calculatoire avec l'activité perceptive et l'action physique. D'où le principe de co-activation et de co-construction : *le corps et les systèmes de génération d'émotions et de pensées sont liés ; ils ne peuvent exister indépendamment et doivent être conçus en co-construction stricte (il doit y avoir une relation organique entre une partie calculable et une partie matérielle).*

Une émotion est donc relative à une représentation *directe* de son corps et il faut une autre structure pour la *ressentir*. Une *carte corporelle* est alors l'ensemble des représentations liées à toute prise d'information sensible du corps. Ces éléments sont *symboliques*, ce qui les distingue fondamentalement des réseaux de neurones. La notion d'intégrité d'un organisme résulte ainsi de la multiplicité et de la cohérence des perceptions.

La mise en relation de ces diverses cartes corporelles spécialisées (co-activation ou désactivation, selon que les émotions correspondantes sont liées ou exclusives) est assurée par des éléments de synchronisation qui s'activent en cas d'intensité particulière d'un ensemble morphologique, ou bien de la reconnaissance d'un type particulier de morphologie, ou encore de la constatation d'une faiblesse anormale d'une morphologie particulière. Ces éléments — qui prennent des décisions selon leurs connaissances et l'état du contexte — sont représentés par des agents cognitifs qui analysent les activités des cartes et qui contrôlent et régulent l'activité émotionnelle selon le contexte. Ce mécanisme produit des modifications de cartes spécifiques, qui ne sont pas induites par les résultats de leurs capteurs ; d'où un comportement qui n'est pas forcément rationnel ! Ainsi, les cartes qui arrivent le plus vite à un état remarquable imposent-elles leur comportement. **[Je note ici une analogie significative avec les mécanismes d'interprétation sémantique proposés par un de mes doctorants (Alejandro Bassi, 1995) où là aussi, les processus qui arrivent le plus vite à un résultat sont ceux qui font appel aux connaissances les plus accessibles et ont donc de plus grandes chances d'être les plus pertinents.]**

Grâce à un bouclage dynamique sans fin, une carte corporelle générale représente l'état global des activités de ces cartes corporelles, et leurs évolutions. Cette carte n'est pas une structure symbolique ou numérique donnant une image de l'état des capteurs et des processus internes, c'est une représentation *située* de l'activation des agents.

Une *carte émotionnelle* est définie comme une carte corporelle spécialisée où le processus miroir et le couplage conduisent à la modification continue de l'organisation aspectuelle.

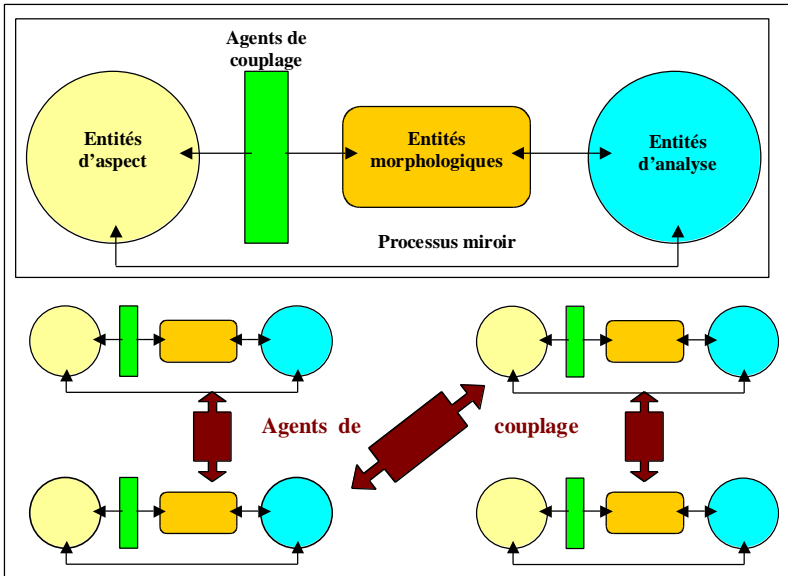


Figure 2 — Une carte émotionnelle spécifique et des couplages entre plusieurs cartes émotionnelles

On distingue les cartes spécifiques de situation des cartes spécifiques d'état interne. De même que pour les cartes corporelles, la synchronisation et la mise en cohérence de l'ensemble des cartes émotionnelles spécifiques fournit une carte émotionnelle générale d'où émerge l'état émotionnel de l'entité. Le chapitre 12 illustre ces points de façon claire par des cartes émotionnelles particulières (comme : quiétude, gêne, malaise, peur, plaisir...).

L'auteur introduit ensuite des *tendances fondamentales* [à mettre en parallèle avec les valeurs d'Edelman], qui sont nécessaires pour l'appropriation de l'espace et du temps. Elles se traduisent par la génération de formes remarquables (issues de cartes très générales considérées comme bornes supérieures des cartes émotionnelles) dans les ensembles morphologiques des cartes émotionnelles [cela est à comparer également avec les hiérarchies de contextes de Baars (1988)]. L'équilibre entre les deux types de cartes traduit le contrôle que l'entité a sur ses pulsions et relève du système de génération de pensées.

4. Quatrième partie

Le modèle du système générateur de pensées présente une architecture similaire à celui du générateur d'émotions, mais plus complexe, à cause 1) de la nécessité de représenter l'évolution même des réorganisations des cartes émotionnelles et 2) de l'exigence que le système dans son ensemble (pensées + émotions) soit conçu comme un tout. La mise en correspondance de ces deux sous-systèmes ne relève pas d'un superviseur mais d'un mécanisme fortement distribué permettant une auto-organisation adaptative.

Dans le système générateur de pensées, une structure de signification est un ensemble de *grains de signification* (entités aspectuelles) associé à un ensemble morphologique et à une mémoire, où sont enregistrées les associations entre éléments émotionnels et éléments cognitifs correspondant aux états de stabilité. Une pensée artificielle est alors une forme dynamique d'organisation des structures de signification engendrant un état global continu. L'oubli correspond à une surcharge de certaines entités de mémoire (diminution d'accessibilité ou perte d'accès). Les éléments de synchronisation agissent sur un ensemble de structures de significations et peuvent les co-activer de manière minimale, les synchroniser plus ou moins fortement ou les fusionner. L'état courant de la pensée de l'organisme est représenté par un point fixe du processus miroir bouclé entre les entités agrégées et la représentation de l'activation de ces entités. **[On trouve ici une analogie forte avec le modèle du carnet d'esquisses, que j'avais proposé il y a quelques années (Sabah, 1996)].**

Mais, en outre, *pour penser, il faut être capable de se savoir penser*. Cela représente un saut qualitatif considérable qui introduit la notion d'intention (alors que les émotions sont automatiquement engendrées et ne sont pas perçues, les pensées sont suscitées par des intentions et correspondent à un état de saturation et à une brève fermeture sur le monde extérieur).

La notion de *scène latente* est définie comme l'élément sous-jacent qui donne une unité à l'ensemble des pensées issues des perceptions successives ; cette scène n'est pas perçue, mais conduit à la perception. Les altérations de cette scène sont issues d'actions intentionnelles : une pensée est un état de stabilité (*scène signifiante*) qui se réfère à quelque chose que le système *voulait* se représenter et qu'il *ressent* en l'exprimant dans ses entités aspectuelles.

Pour caler le système générateur de pensées sur le réel, le système doit être capable d'interrompre ses processus miroirs, sans faire effondrer le système ; d'où la nécessité de la redondance et de l'inertie.

Le système « se sait penser » à cause de la tension entre la stabilité qui produit une pensée et le besoin de déstabilisation nécessaire à l'appréhension du monde et à l'intention de penser à nouveau ; d'où un choix, que le système doit effectuer intentionnellement.

L'aspect créateur de la pensée implique que l'état déclencheur n'est pas de même nature que les états engendrés et est forcément non perceptible (sinon on aurait un déterminisme absolu). Cela est traduit par les *ensembles d'anticipation* : des entités particulières qui, agissant sur les entités morphologiques des structures de signification, engagent *a priori* à la représentation d'une certaine scène, en modifiant les processus miroirs. Ces ensembles s'activent à partir de caractères typiques des ensembles morphologiques et agissent de manière continue et impérative (*il s'agit là de l'un des éléments essentiels qui définissent le « moi »*). L'autocontrôle du système résulte de l'uniformité des éléments de base du système (agents proactifs communicants) ayant un fonctionnement parallèle et concurrent, ainsi que de la possibilité pour quelques agents d'infléchir le fonctionnement de l'ensemble.

Tout ce processus est régi par un ensemble de prégnances fondamentales ou relevant de divers systèmes de valeurs. **[Ce point est à mettre en parallèle à la fois avec les « valeurs », centrales dans le modèle d'Edelman et les hiérarchies de « buts » de Baars...]**

Les *sensations* (la traduction au niveau conscient des émotions, qui sont non conscientes) résultent de la co-activation des systèmes d'émotions et de pensées, ce qui provoque un mouvement réciproque entre eux. La perception de ce mouvement (dont l'intensité est proportionnelle à l'intensité de l'émotion correspondante) est alors considérée comme une *sensation artificielle*.

Le « Soï » du système est alors la sensation qu'il a de s'auto-contrôler par un bouclage systémique continu, en activant des éléments d'anticipation en réaction à l'activation d'éléments morphologiques qui s'activent, eux, à cause de l'émergence d'organisations aspectuelles de structures de signification. La machine pensante étant alors à la fois le metteur en scène, l'acteur et l'observateur de ces auto-contrôles peut ainsi se « savoir être ».

Le langage est également abordé et soulève un autre problème fondamental : quelles sont les relations entre les idées et les mots de la langue ? **[À part une note d'optimisme précisant qu'il n'est peut-être pas insoluble, je trouve que ce problème est traité d'une façon trop périphérique, car, à mon sens, cette partie ne reflète pas le rôle essentiel que joue le langage dans le système de pensée des humains ! Quelques raccourcis saisissants auraient sans doute bénéficié ici d'une articulation approfondie avec les thèses d'Edelman...].**

Pour rendre compte de la faculté de langage, l'auteur propose d'introduire des agents aspectuels de type « mot » qui se réfère, en fait, à un symbole (phonétique ou écrit) désignant un mot. Un ensemble d'agents aspectuels de ce type constitue une *structure de signification*, nécessairement liée à des structures non langagières, et dans l'étape initiale de conceptualisation, à des cartes émotionnelles. L'auteur insiste sur l'importance des agents d'anticipation dans les processus de compréhension.

Les structures langagières sont considérées comme une action au second degré de la machine pensante sur le réel (elles sont une *abstraction* des structures de signification ordinaires). Alors que les scènes signifiantes ordinaires conduisent en général à des mouvements et se réfèrent à des phénomènes présents, les scènes langagières deviennent abstraites et produisent de nouveaux objets, eux-mêmes abstraits.

À propos du fonctionnement de l'ensemble du système, l'auteur insiste sur le fait qu'il **ne s'agit pas** d'une machine de Turing, même extrêmement compliquée ! Pour lui, le système pourrait se ramener à un *ensemble de réseaux* de machines de Turing, interconnectées, communicantes et contrôlées par des éléments en prise directe avec le monde (de ce dernier point, résulte le fait fondamental qu'il est impossible de prévoir complètement le fonctionnement effectif du système, qui, de plus, s'altère lui-même). Ainsi, la machine pensante est-elle considérée comme un *système calculable, localement de Turing*.

En conclusion la caractéristique essentielle d'une telle machine est d'engendrer continuellement une scène latente et des scènes signifiantes, et de transformer sans cesse et de façon inéluctable ses productions. Elle mémorise partiellement ces productions, apprenant et se modifiant constamment.

[Bien entendu, cette machine n'est pas construite et le modèle demande encore une validation par l'expérience, mais il est extrêmement élaboré et, personnellement, me semble convaincant.]

Je ne peux que recommander très fortement la lecture de ce livre qui donne un point de vue éclairant sur une conception matérialiste de la conscience, de la pensée et donc de l'intelligence, ainsi que sur les possibilités d'automatisation de ces facultés. Les seuls regrets que j'ai eus en le lisant sont le manque de références à Edelman et Baars, l'absence d'un index ainsi que d'exemples plus concrets qui seraient souvent bien utiles pour illustrer les différents concepts présentés].

Références bibliographiques

- Baars B. (1988) *A cognitive theory of consciousness*. Cambridge University Press, Cambridge
- Bassi Acuña A. (1995) *Un modèle dynamique de la compréhension de texte intégrant l'acquisition des connaissances*. Thèse d'université, Paris XI, Orsay, France.
- Damasio A. (1999) *Le sentiment même de soi*. Odile Jacob, Paris
- Edelman G. (1992) *Biologie de la conscience*. Odile Jacob, Paris
- Sabah G. (1996) Le « carnet d'esquisses » : une mémoire interprétative dynamique. Actes *RFIA'1996*, Rennes, France.

L'auteur de la revue critique



Né en 1948, diplômé de l'École Polytechnique en 1971, **Gérard Sabah** obtient son doctorat de troisième cycle en 1973 (reconnaissance des formes et apprentissage automatique), puis son doctorat d'état ès sciences en 1978 sur le thème du traitement automatique des langues. Spécialiste donc d'informatique linguistique depuis plus de vingt-cinq ans, il est l'auteur d'environ deux cents articles et de plusieurs ouvrages ; il a dirigé plus de soixante thèses sur le sujet.

Il fut président de l'ARC (Association pour la recherche cognitive) de 1987 à 1991, puis membre du bureau de l'AFIA (Association française d'intelligence artificielle) en 1995 et rédacteur en chef de son bulletin (de novembre 1997 à septembre 2001).

Actuellement directeur de recherche au CNRS, il est chargé de mission au département STIC du CNRS (domaine « interactions humaines et cognition ») et membre du groupe de travail « interactions homme-machine » de l'Académie des Technologies. Depuis mars 2003, il est nommé par la directrice générale du CNRS directeur du programme interdisciplinaire de recherche TCAN (Traitement des Connaissances, Apprentissage et Nouvelles technologies de l'information et de la communication).