

John STEWART

Un système cognitif sans neurones : les capacités d'adaptation, d'apprentissage et de mémoire du système immunitaire

I. Introduction

Les sciences cognitives sont encore jeunes, et il n'existe pas de consensus établi concernant les disciplines qui en font partie. On considère souvent comme allant de soi qu'il n'y aurait pas de cognition sans neurones (ou des automates neuro-mimétiques dans le cas de l'intelligence artificielle), et donc que les neurosciences font automatiquement partie des sciences cognitives. Autrement dit, on considère que des neurones constituent une base nécessaire et suffisante pour des phénomènes cognitifs, et que la biologie contribue aux sciences cognitives essentiellement par le biais des neurosciences. Dans cet article, je souhaite remettre en question ces présupposés en fournissant un contre-exemple possible.

L'immunologie est une discipline biologique caractérisée par le fait que le discours spontané des immunologistes attribue au système immunitaire des capacités "cognitives" : la reconnaissance, l'apprentissage, la mémoire et la distinction soi/non-soi (Klein, 1982). Afin d'examiner sérieusement la possibilité que le système immunitaire puisse être légitimement considéré comme "cognitif", je proposerai une définition théorique de ce que l'on peut entendre par le terme "cognition". Cette définition fournit des critères qui permettront ensuite d'examiner à quelles conditions les neurosciences peuvent être considérées comme appartenant pleinement aux sciences cognitives.

II. Le système immunitaire

Tous les animaux vertébrés possèdent un système immunitaire composé essentiellement de *lymphocytes*, des cellules différenciées qui circulent librement dans le sang et dans la lymphe. Les lymphocytes fabriquent des protéines très particulières, les *immunoglobulines*, qui existent sous deux formes : soit intégrées dans la membrane cellulaire, soit sécrétées dans le sang. Les immunoglobulines membranaires fonctionnent comme des *récepteurs* : quand ceux-ci sont "occupés" par d'autres molécules, la lymphocyte est activée, soit pour produire et sécréter ses propres immunoglobulines en grande quantité, soit pour se multiplier. L'ensemble des cellules produites par la multiplication d'une même lymphocyte ancestrale forme un *clone*. Au sein de chaque molécule d'immunoglobuline, on distingue deux régions : l'une, dite région "d'ossature", est la même chez toutes les immunoglobulines et leur confère les mêmes structures et formes globales ; l'autre, dite région "variable", est différente d'une immunoglobuline à une autre¹ et forme le "site de combinaison" responsable pour les interactions spécifiques de l'immunoglobuline avec d'autres molécules. La diversité des régions variables est produite par un processus remarquable de recombinaison somatique de l'ADN des lymphocytes ancestrales. Cette diversité est si grande qu'il n'existe pas deux clones ayant des sites de combinaison identiques ; en même temps, dans une population de 10^8 lymphocytes², il existe au moins un clone capable d'interagir spécifiquement, par le biais de ses immunoglobulines membranaires, avec n'importe quelle molécule chimique d'une taille suffisante. Les immunologistes décrivent cette situation en disant que le "répertoire" du système immunitaire est "complet" (Coutinho et al., 1984).

¹ La région variable est identique chez toutes les immunoglobulines appartenant à un même clone.

² Le nombre de lymphocytes chez une souris.

III. Les fonctions du système immunitaire

Le système immunitaire possède au moins deux modes de fonctionnement, dont chacun constitue un domaine phénoménologique caractéristique.

III.1. L'immunologie classique : les réponses immunitaires

Depuis ses origines au début du vingtième siècle, l'immunologie classique s'est focalisée principalement sur l'activation du système immunitaire par des "antigènes" extérieurs, notamment des bactéries et des virus. Quand un antigène est introduit dans le système, les clones de lymphocytes qui interagissent³ avec lui sont stimulés, de sorte qu'ils se multiplient⁴ et secrètent leurs immunoglobulines spécifiques en grande quantité. Ainsi, en l'absence d'interaction des lymphocytes entre elles sur le mode d'un réseau (voir ci-dessous), l'introduction d'un antigène provoque une forte "réponse immunitaire" qui déclenche la destruction de l'antigène. Ce mode de fonctionnement du système immunitaire est connu sous le nom de la *théorie de la sélection clonale* (Burnet, 1959).

III.2. Le système immunitaire comme réseau

Depuis quelques années, notre compréhension scientifique de l'organisation du système immunitaire s'est enrichie par le développement d'une théorie supplémentaire. Le point de départ de cette nouvelle théorie fut donné par Jerne (1974), qui faisait remarquer si le répertoire du système immunitaire est effectivement complet, alors les immunoglobulines spécifiques produites par un clone (qui sont, après tout, des molécules chimiques) vont nécessairement interagir avec les immunoglobulines d'un autre clone. Cette remarque a conduit à l'élaboration d'une théorie du système immunitaire comme un *réseau* capable de s'activer lui-même de

³ Les immunologistes parlent de "reconnaissance" de l'antigène ; j'examinerai à l'occasion de ma discussion si l'emploi de cette métaphore cognitive peut être justifié.

⁴ D'où le terme "sélection clonale", en analogie avec la théorie néo-darwinienne de l'évolution.

manière autonome, indépendamment de toute interaction du système avec des antigènes extérieurs. Afin de se faire une idée du comportement possible d'un tel réseau, nous avons construit un modèle mathématique du système immunitaire (Varela et al. 1988, Stewart & Varela 1989). Même dans sa forme la plus simple, présentée ci-dessous, ce modèle constitue un système dynamique hautement non-linéaire⁵. Des solutions analytiques étant donc hors de portée, nous avons étudié ce système au moyen de simulations sur ordinateur.

Les éléments de ce modèle sont des clones de lymphocytes, numérotés de 1 à n . Pendant toute la vie d'un animal vertébré, de nouveaux clones sont produits en grand nombre dans la moelle osseuse, et présentés comme candidats à un recrutement parmi la population de lymphocytes activées. Afin de modéliser les interactions entre les clones, nous avons adopté le concept d'un "espace de formes" (Stewart & Varela 1991) selon lequel la "forme" du site de combinaison de chaque immunoglobuline peut être représentée par p paramètres stéréo-chimiques. Les valeurs de ces paramètres peuvent être prises comme coordonnées dans un espace abstrait à p dimensions, de sorte que chaque forme moléculaire correspond à un point dans cet espace. Ainsi, des molécules ayant des formes semblables (et par conséquent des profils d'interaction avec d'autres molécules également semblables) seront représentées par des points rapprochés. Par ailleurs, deux molécules ayant des formes complémentaires (et par conséquent une forte interaction due à leur affinité chimique) sont représentées par des points proches dans l'espace des formes, mais appartenant à des "plans" différents que l'on désigne conventionnellement par "noir" et "blanc". Ainsi, l'affinité m_{ij} entre deux molécules i et j est donnée par :

$$m_{ij} = \exp(-d_{ij}^2) \quad \text{Equation [1]}$$

où d_{ij} est la distance dans l'espace des formes entre un point noir (ou blanc) i et un point blanc (ou noir) j .

⁵ De surcroît, la liste des variables n'est pas fixe mais évolue elle-même au cours du temps.

Dans la forme très simplifiée du modèle présentée ici, il n'y a pas de prolifération : les clones sont soit présents (à concentration 1) soit absents (à concentration 0). L'effet combiné de tous les autres clones j sur un clone donné i est défini par le champ h_i , où

$$h_i = \sum m_{ij} \quad \text{Equation [2]}$$

(la sommation est effectuée sur $j = 1 \dots n$).

Sur ces bases, le modèle postule que les critères pour le maintien d'un clone existant et pour le recrutement d'un nouveau clone sont les mêmes : si le champ h_i est compris dans une "fenêtre" entre un seuil inférieur et un seuil supérieur, le clone est maintenu/recruté dans la population actuelle ; si le champ est extérieur à la fenêtre, le clone est éliminé/rejeté comme candidat à recrutement. Dans la mesure où le concept même d'un espace de formes est valide, le nombre de dimensions est probablement relativement élevé. Néanmoins, les simulations que l'on présentera ici se déroulent dans un espace de formes à deux dimensions, car cela permet une visualisation graphique des résultats qui contribue grandement à leur interprétation.

Chaque simulation commence avec un seul clone "noir" au centre de l'espace de formes. Des clones "noirs" et blancs" sont générés à des points aléatoires dans l'espace de formes et proposés comme candidats au recrutement. Après chaque recrutement, tous les clones présents dans la population sont ré-évalués, et ceux dont le champ est extérieur à la fenêtre sont éliminés. Le résultat de ce processus est illustré dans la figure 1 ; schématiquement, il comporte deux phases.

Dans un premier temps, la population augmente jusqu'à ce que la totalité de l'espace soit occupée, mais de manière plutôt désorganisée (figures 1a et 1b). Pendant cette phase, on observe assez souvent un phénomène "d'effondrement local", pour la raison suivante : (i) la présence d'un premier clone (par exemple, "noir") crée un champ permettant le recrutement dans son voisinage d'un deuxième clone (blanc) ; (ii) ce recrutement peut avoir l'effet d'augmenter le champ pour le premier clone au-dessus du seuil supérieur, de sorte que ce clone est éliminé ; (iii) cette élimination

abolit le champ pour le deuxième clone, de sorte que celui-ci est également éliminé. Autrement dit, le processus exhibe une forme de "criticalité auto-organisée" caractérisée par de nombreuses "avalanches" de plus ou moins grande ampleur.

Figure 1. Quatre stades successifs dans la constitution auto-organisée d'un réseau immunitaire, modélisé par simulation informatique dans un espace de formes à

2 dimensions. Après une première période chaotique (1a et 1b), on voit l'apparition (1c) et la stabilisation (1d) de chaînes parallèles de clones noirs et blancs. D'après Stewart & Varela (1991).

Cette phase initiale continue jusqu'à ce que cette réorganisation répétée conduit à l'apparition de certaines configurations qui sont quasi-stables (figures 1c et 1d). La nature de ces configurations apparaît clairement si des "effets de bord" sont éliminés par fusions des bords opposés de l'espace, de sorte que l'espace de formes possède la topologie d'un tore (figures 2).

Ces configurations quasi-stables sont formées de chaînes linéaires de clones noirs et blancs ; chaque chaîne noire est accompagnée d'une chaîne blanche parallèle. La région "derrière" une chaîne noire (ou blanche) est bornée par une deuxième chaîne noire qui fait face à sa propre chaîne blanche, et ainsi de suite. La quasi-stabilité de cette configuration s'explique de la manière suivante. Dans la région entre deux chaînes de couleurs opposées, le champ est au-dessus du seuil supérieur, et donc trop élevé pour que des clones de l'une ou l'autre couleur puissent y être recrutés. Dans la région "derrière" une chaîne, le champ est trop bas pour recruter des clones de la même couleur. Autrement dit, la configuration d'une paire de chaînes parallèles de couleurs opposées est stable parce que chaque chaîne crée un champ (à une distance correspondant à celle entre les deux chaînes) dans la fenêtre qui permet le recrutement/maintien des clones de l'autre couleur. Finalement, il reste une forme de recrutement qui pourrait éventuellement conduire à la dislocation de cette configuration globale : celle de clones blancs dans la région "dos à dos" bordée par des chaînes noires (ou de clones noirs dans une région "blanche"). Sporadiquement, un tel recrutement peut se produire (par exemple, le clone noir isolé dans la figure 2b). Toutefois, le clone isolé crée immédiatement un champ qui conduit au recrutement avoisinant de clones de l'autre couleur ; à cause de l'asymétrie créée par les chaînes qui entourent le clone isolé, le champ pour celui-ci dépassera systématiquement le seuil supérieur,

ce qui conduit à son élimination par un processus semblable à celui des "avalanches" évoquées ci-dessus⁶.

⁶ Des connaisseurs du jeu de Go reconnaîtront des analogies plaisantes non seulement avec les configurations en chaînes noires et blanches, mais aussi avec les raisons stratégiques sous-jacentes à la formation de ce genre de configuration.

Figure 2. Deux exemples supplémentaires des configurations émergentes dans un réseau immunitaire simulé, dans le cas où l'espace de formes possède la topologie d'un tore. D'après Stewart & Varela (1991).

En conclusion, on peut dire que certaines configurations caractérisées émergent précisément à cause de leur relative stabilité. Ce processus s'apparente donc à une forme de sélection naturelle, mais qui diffère de la "sélection clonale" de la théorie classique sur

deux points. Premièrement, les unités de sélection ne sont pas des clones individuels, mais une entité d'un niveau d'organisation supérieur, à savoir des configurations composées d'un ensemble de clones. Deuxièmement, l'agent de la sélection n'est pas une entité fixe, pré-donnée, extérieure au système ; les contraintes sont endogènes, générées par le fonctionnement du système lui-même.

Figure 3. L'effet du couplage d'un réseau immunitaire, simulé comme dans les figures 1 et 2, avec des antigènes fixes (carrés noirs et blancs). 3a et 3b : deux exemples qui montrent l'incorporation des antigènes dans des chaînes de même

couleur. 3c : la configuration de 3b reste quasi-stable même après enlèvement des antigènes, ce qui correspond à une "mémoire" du système. 3d : l'introduction d'antigènes discordants dans la situation précédente (3c) conduit à un réajustement des chaînes. D'après Stewart & Varela (1991).

Sur la base de cette étude des propriétés auto-organisatrices du réseau, nous sommes maintenant en mesure d'aborder la question de l'interaction du système avec des antigènes. Dans notre modèle, des antigènes sont des formes moléculaires dont la présence est indépendante du champ qui s'exerce sur elles ; ils sont représentés par des carrés afin de les distinguer des clones de lymphocytes représentés par des ronds. Puisque des antigènes peuvent *a priori* posséder des formes moléculaires quelconques, il y aura des antigènes "noirs" et "blancs". Les résultats de simulations où l'ontogénèse du réseau se déroule en présence de plusieurs antigènes sont illustrés dans la figure 3.

Ces résultats montrent que le réseau conserve ses propriétés auto-organisatrices, avec l'émergence de configurations semblables à celles déjà étudiées. L'élément nouveau est que les antigènes noirs et blancs semblent être systématiquement incorporés dans des chaînes de la même couleur (figures 3a et 3b). Autrement dit, le couplage du système avec les antigènes crée une contrainte sur l'emplacement des chaînes. La figure 3c montre que cet emplacement contraint est maintenu même si les antigènes en question sont enlevés⁷. La figure 3d montre qu'il s'agit non pas d'une simple inertie, mais d'un phénomène dynamique : l'introduction d'antigènes discordants (un antigène noir dans une chaîne blanche, ou d'un antigène blanc dans une chaîne noire) conduit à un réajustement des chaînes de sorte qu'elles incorporent à nouveau des antigènes de la même couleur.

IV. Discussion

IV.1. Une définition de la "cognition"

⁷ J'examinerai dans la discussion si on peut considérer qu'il s'agit là d'un exemple de "mémoire".

Nous sommes maintenant en mesure d'aborder la question centrale de cet article : peut-on légitimement considérer que le système immunitaire est une entité "cognitive" ? Toute réponse à cette question requiert évidemment une définition convenable de ce que l'on entend par le terme "cognition". Dans cet article, je propose d'examiner la fécondité de la définition suivante : une entité est "cognitive" si elle est capable à la fois de perceptions et d'actions ; de plus, les actions doivent être guidées par les perceptions de telle sorte qu'une contrainte significative est satisfaite. Ce concept est proche de celui de "systèmes adaptatifs" proposé par Meyer et Guillot (1990) ; il s'agit évidemment d'une définition minimale, le "degré zéro" de la cognition. Cette définition n'est pas pour autant triviale, comme on peut le constater en approfondissant la notion d'une contrainte "significative".

Par "significative", j'entends une contrainte *proscriptive*, par opposition avec une contrainte *prescriptive* qui spécifie, en détail, exactement ce que le système doit faire dans chaque situation. Dans l'Intelligence Artificielle traditionnelle (la "Good Old-Fashioned Artificial Intelligence" de Boden, 1987), les systèmes sont typiquement soumis à des contraintes prescriptives. Afin de couvrir toutes les situations possibles, la spécification de telles contraintes est souvent longue et fastidieuse, ce qui peut faire perdre de vue un point bien plus important : puisque ces systèmes fonctionnent par la simple intériorisation de ces contraintes, en fin de compte ils sont strictement déterminés par l'extérieur. On peut même considérer qu'ils sont en fait assez stupides, car ils font exactement ce que l'on leur dit de faire, ni plus ni moins. Pour résumer, ces systèmes sont *hétéronomes* (Varela 1989).

Depuis quelques années, ce genre de cognitivisme s'est considérablement enrichi par l'avènement du néo-connexionisme. En particulier, la technique d'ajustement des poids de connexion dans un réseau multi-couche par rétropropagation des erreurs contourne un goulot d'étranglement, car il est dorénavant possible d'associer des entrées sensorielles avec des sorties symboliques sur la base d'une série d'exemples plutôt que par une spécification prescriptive exhaustive. Toutefois, dans la mesure où une série appropriée d'exemples doit être fournie de l'extérieur du système lui-même, de

tels réseaux néo-connexionnistes restent essentiellement hétéronomes.

Une contrainte proscriptive, comme le terme l'indique, se prête souvent à une formulation négative. Ainsi, par exemple, on peut exiger d'un robot qu'il ne doive pas heurter des obstacles, qu'il ne doive pas tomber par-dessus le bord de la table sur laquelle il est posé, sans pour autant simplement rester immobile. Malgré cette apparence négative, la capacité à satisfaire une contrainte proscriptive est en réalité un exploit positif, précisément parce que la stratégie à suivre n'est pas dictée de l'extérieur. Des systèmes qui satisfont une contrainte proscriptive sont donc *autonomes*. Cet aspect créatif des systèmes autonomes est particulièrement évident dans l'exemple prototypique d'une contrainte proscriptive : la capacité des organismes vivants de rester dans leur domaine de viabilité (Aubin 1991). La capacité de satisfaire une contrainte "négative" (ne pas se départir de leur viabilité) constitue de toute évidence un exploit remarquable : rester en vie ou, dans les termes de Maturana et Varela (1980) maintenir un état d'autopoïèse. Par contraste avec des contraintes prescriptives, qui possèdent en général une solution optimale unique, une contrainte proscriptive peut être satisfaite par une pluralité de stratégies. Les milliards d'espèces et de stratégies de vie des organismes vivants sur la planète Terre offrent une illustration saisissante de cette pluralité.

Il est clair que cette proposition correspond à une définition élémentaire et minimale de la cognition, qui est satisfaite par les organismes vivants les plus primitifs (mais non pas par des virus isolés). Par exemple, les bactéries peuvent agiter leurs cils vibratiles soit de manière cohérente et ondulatoire, ce qui provoque un déplacement en ligne droite, soit de manière désordonnée, ce qui provoque une réorientation aléatoire : elles sont donc capables d'actions différenciées. Par ailleurs, les bactéries sont également capables de percevoir différenciellement si la concentration en sucre de leur milieu est en train d'augmenter ou de diminuer. Enfin, les bactéries sont aussi capables de guider ces actions par ces perceptions de telle sorte qu'ils se rapprochent systématiquement d'une source de nutrition sucrée, ce qui contribue manifestement à leur viabilité. Selon notre définition minimale, les bactéries sont donc des agents cognitifs.

Toutefois, afin de dissiper tout malentendu, il est important de souligner que je ne veux absolument pas suggérer que la cognition ne serait jamais autre chose, qu'elle se réduirait nécessairement à cette seule définition élémentaire. Au contraire, au cours de l'évolution ultérieure, toute une série de domaines cognitifs émergent successivement, et les sciences cognitives telles que je les envisage se doivent d'étudier chacune de ces émergences. Citons quelques-uns de ces événements marquants : l'apparition d'organismes multicellulaires dotés d'une ontogénèse ; la communication sémiotique des animaux sociaux ; l'apparition du langage au cours de l'homínisation ; l'invention de l'écriture et l'entrée dans l'histoire ; la conscience réflexive et l'invention de la philosophie ; la modernité et la naissance de la science moderne. Il ne serait nullement absurde de prendre l'une ou l'autre de ces étapes comme marquant la frontière de la "cognition" proprement dite. Cependant, je soutiens que toutes ces formes ultérieures de la cognition sont enracinées dans la forme élémentaire de la définition proposée. De plus, l'idée radicale que l'on peut résumer par la formule "Vie=Cognition" (Stewart, 1992, 1993a) entraîne quelques conséquences fortes et inattendues ; en particulier, comme nous allons voir, cela signifie que ni des cerveaux, ni des ordinateurs, ne sont en eux-mêmes "cognitifs".

Je propose donc d'accepter cette définition de la cognition comme "actions guidées par des perceptions sous contrainte proscriptive", et d'examiner sa fécondité en cherchant à l'appliquer au cas du système immunitaire. Il est à noter que cette définition n'est pas simplement fabriquée sur mesure pour les besoins de cet article. Premièrement, pour anticiper sur ma conclusion, l'état actuel de l'immunologie ne permet pas d'affirmer que le système immunitaire parvient effectivement à se qualifier comme "cognitif" selon les termes de cette définition. Par ailleurs, cette définition est conforme à l'approche générale du champ nouveau et dynamique que l'on appelle "Vie Artificielle" (Langton 1989, ECAL 91, ECAL 93). Enfin, cette approche comporte des implications positives et précises pour l'étude de la cognition supérieure, notamment les phénomènes de la communication, des représentations, et du langage. Je reviendrai sur ce point dans ma conclusion.

IV.2. Le système immunitaire est-il "cognitif" ?

Si l'on définit la cognition comme "actions guidées par des perceptions sous contrainte proscriptive", il faut donc répondre à trois questions : premièrement, le système immunitaire perçoit-il son environnement ? deuxièmement, est-il capable d'actions différenciées ? et troisièmement, les actions sont-elles guidées par les perceptions de telle sorte qu'une contrainte de viabilité est satisfaite ?

Premièrement donc : le système immunitaire perçoit-il son environnement ? "Percevoir", dans ce contexte, signifie que des interactions entre le système immunitaire et son environnement conduisent, de manière systématique et fiable, à des modifications caractéristiques de l'état interne du système. Nous ne nous plaçons donc pas dans l'optique d'une transmission "d'information" de l'environnement vers le système (Stewart et al., 1988) ; on considère plutôt que si des interactions entre système et environnement déclenchent des transitions d'un état interne du système à un autre, ce sont les propriétés auto-organisatrices du système lui-même qui déterminent ce que sont ces états internes (Maturana & Varela 1980).

Afin de répondre correctement à la question posée, il est donc nécessaire de prendre en considération le fait que le système immunitaire exhibe deux modes de fonctionnement : le mode "classique" de la sélection clonale, et le mode du "réseau" modélisé dans les simulations décrites ci-dessus. Dans le premier mode, le système "perçoit" un antigène en produisant une réponse immunitaire, c'est-à-dire en produisant une concentration élevée d'immunoglobulines ayant une forte affinité pour l'antigène. Dans le deuxième mode, le couplage avec un antigène conduit à un réajustement (dans l'espace de formes) des chaînes de clones "noirs" et "blancs" de sorte que les antigènes sont incorporées dans des chaînes de la même couleur. Dans les deux cas, on peut conclure que dans les termes employés ici, le système immunitaire "perçoit" effectivement son environnement.

Deuxièmement, le système immunitaire est-il capable "d'actions" guidées par ses perceptions ? La réponse est encore "oui", à condition que l'on accepte de décrire le fonctionnement du système

immunitaire dans des termes qui sont inhabituels pour l'immunologie classique. Par des mécanismes qu'il n'est pas nécessaire de détailler ici⁸, une réponse immunitaire déclenche la destruction de l'antigène qui l'a suscité. Dans le cas du deuxième mode, celui du "réseau", une réponse positive dépend d'une hypothèse spéculative : il est possible que l'incorporation d'antigènes dans des "chaînes" d'immunoglobulines crée des connexions fonctionnelles entre les antigènes en question (Varela et al., 1992). Brièvement, l'idée est celle-ci : si deux antigènes, situés à des endroits éloignés dans l'espace des formes, sont reliés par de telles chaînes, alors une variation dans la concentration de l'un des antigènes produira une variation dans le champ ressenti par l'autre. Si les chaînes sont déjà en place — la mise en place étant un processus ontogénétique se déroulant sur une échelle de temps de semaines ou de mois — alors ce genre "d'action à distance dans l'espace des formes" peut se produire en quelques secondes ou minutes par des processus moléculaires d'association/dissociation des complexes formés d'immuno-globulines "noires" et "blanches". Dans le cas où les antigènes ainsi reliés sont des molécules ayant des rôles de "messagers chimiques" (Sinding, 1993) — des hormones, des neurotransmetteurs, leurs récepteurs, etc. — de telles connexions pourraient jouer une fonction d'intégration physiologique, et fournir un mécanisme permettant par exemple d'expliquer des phénomènes psychosomatiques.

Troisièmement, est-ce que la manière particulière par laquelle les perceptions interviennent pour influencer les actions permet au système de se conformer à une contrainte de viabilité ? Cette question cruciale nous amène au cœur de la discussion. Le mode "classique" de fonctionnement du système immunitaire est approprié dans le cas d'antigènes extérieurs comme les bactéries et virus : la destruction de ceux-ci, déclenchée par la réponse immunitaire, défend l'organisme contre des maladies infectieuses. Le mode de fonctionnement en "réseau" est approprié dans le cas où les antigènes sont des molécules faisant partie de l'organisme lui-même, ce que l'on appelle communément les "antigènes du soi". Par contre,

⁸ Brièvement, il s'agit des processus d'inflammation, d'ingestion par des macrophages, et de déclenchement de la cascade du complément (Klein 1982).

une réponse à des bactéries et des virus sur le mode du réseau ne permettrait pas à l'organisme de se défendre contre les maladies infectieuses ; et une réponse à des antigènes du soi sur le mode de la réponse immunitaire provoquerait de graves maladies auto-immunes. On peut en conclure que le système immunitaire dans son ensemble est effectivement "cognitif" si mais seulement s'il entre systématiquement dans le mode de fonctionnement approprié.

Au niveau purement empirique de l'observation du résultat final, il y a lieu de penser que le système immunitaire parvient effectivement à satisfaire cette contrainte. Cependant, cette conclusion ne sera réellement acquise, et l'immunologie n'appartiendra pleinement aux sciences cognitives, que dans la mesure où les *mécanismes* qui conduisent à ce résultat peuvent être identifiés. Cette question est habituellement désignée par le terme "discrimination soi/non-soi" (Bernard et al. 1990) ; elle est d'une importance capitale pour notre discussion, et nous l'examinerons de plus près.

Pour commencer, il est à noter que les deux modes de fonctionnement du système immunitaire sont mutuellement exclusifs : si le système répond à un antigène donné sur le mode du réseau, cela l'empêche de développer une réponse immunitaire. Pour comprendre pourquoi il en est ainsi, il est utile de considérer le sort d'une lymphocyte fraîchement produite par la moelle osseuse. Si cette lymphocyte est recrutée par le réseau, elle sera activée et deviendra partie du "Système Immunitaire Central" ou SIC (Varela & Coutinho 1991). Si elle n'est pas recrutée, elle survivra sous la forme d'une petite cellule inactive avant de mourir au bout de deux ou trois jours. Le taux de production de nouvelles lymphocytes par la moelle osseuse est extrêmement élevé (il est tel que le système entier peut être reconstitué en moins d'une semaine), de sorte que ces cellules inactives et moribondes constituent 80% de la population totale à un moment donné ; elles forment le "Système Immunitaire Périphérique" ou SIP. Il est à noter que même si les lymphocytes du SIP sont le plus souvent moribondes, chacune d'elles peut être activée dans le cas où elle rencontre un antigène ayant une forte affinité avec ses immunoglobulines membranaires ; dans ce cas, n'étant pas soumise à la régulation qui découle d'une insertion dans le réseau du SIC, elle produira une importante réponse immunitaire.

Il est important de bien saisir le rapport entre le SIC et le SIP. Le SIC est défini positivement : il est composé de lymphocytes intégrées dans un réseau autonome d'activations mutuelles. Le SIP, quant à lui, est défini négativement : il est simplement composé d'un résidu de lymphocytes qui ne sont pas incorporées dans le réseau. Puisque le répertoire de toutes les lymphocytes produites par la moelle osseuse est "complet", il s'ensuit que le répertoire du SIP est "complet *moins* le répertoire du SIC".

Nous sommes maintenant en mesure de spécifier plus précisément dans quelles conditions la répartition des rôles entre le mode de fonctionnement en réseau (instancié par le SIC) et le mode de fonctionnement classique (instancié par le SIP) sera "approprié" au sens défini ci-dessus. Essentiellement, le répertoire du SIC devrait inclure tous les antigènes du soi, sous peine de déclencher une maladie auto-immune ; mais ce répertoire ne devrait pas s'étendre trop au-delà, car sinon le répertoire du SIP sera diminué d'autant et ne couvrirait plus la totalité des micro-organismes pathogènes que le SIP doit normalement reconnaître et vouer à la destruction. L'inclusion des antigènes du soi dans le répertoire du SIC ne pose pas un très grand problème ; le SIC se développe le premier, au cours de l'embryogénèse péri-natale, à un moment où seuls des antigènes du soi sont présents. La restriction du répertoire du SIC est beaucoup plus problématique, et constitue un sujet majeur de nos recherches actuelles. Le problème est que dans le modèle très simple présenté ci-dessus, le "réseau" s'étend pour couvrir la totalité de l'espace des formes, ce qui réduit le répertoire du SIP à néant. Le problème, cependant, est bien posé ; il devrait être possible, par une extension appropriée du modèle⁹, de le résoudre.

IV.3. Identité du soi et apprentissage

En résumé, il semble qu'il est à la fois plausible et fécond d'envisager l'hypothèse selon laquelle le système immunitaire est

⁹ En termes techniques, une voie prometteuse consiste à inclure des interactions entre les lymphocytes B productrices d'immunoglobulines (seules prises en considération jusqu'ici) avec des lymphocytes T qui reconnaissent des fragments digérés des protéines "présentés" sur des surfaces cellulaires par les molécules HLA.

effectivement "cognitif" au sens défini ici. Cela signifie que l'emploi de métaphores cognitives en immunologie pourrait correspondre à une réalité scientifique rigoureuse. Nous avons vu qu'en immunologie, une question centrale est celle de la constitution positive d'une identité du soi ; la capacité de réaliser une distinction soi/non-soi en découle comme conséquence secondaire. Il est également à noter que puisque (dans une population naturelle qui est génétiquement diversifiée) chaque individu est génétiquement unique, étant en particulier différent de ses deux parents, la capacité à distinguer "soi" et "non-soi" ne peut être encodée en tant que telle dans l'ADN ; cette capacité cognitive résulte d'un processus "d'apprentissage" épigénétique. En faisant fonctionner la métaphore cognitive à rebours, notre analyse du système immunitaire suggère que l'apprentissage lui-même est un phénomène épigénétique, fondamentalement inséparable de l'organisation de l'ontogénèse du système dans son ensemble. Cette perspective s'accorde bien avec la suggestion que la "mémoire" ne correspond pas à des "informations" stockées ; elle est plutôt un reflet du fait que l'état actuel d'un système cognitif découle de l'histoire de sa constitution ontogénétique (Rosenfield 1989).

V. Conclusions

V.1. Implications générales

Nous avons cherché à examiner, de manière systématique et théoriquement fondée, s'il peut être scientifiquement acceptable de considérer que le système immunitaire est une entité cognitive. Je souhaite à présent reformuler les conclusions de cette étude du cas spécifique du système immunitaire en termes potentiellement applicables aux systèmes cognitifs en général.

Premièrement, cette étude suggère que l'on ne peut répondre à la question concernant l'éventuel statut cognitif d'un système en analysant les seuls composants du système. Des molécules et des cellules différenciées (comme par exemple des lymphocytes, mais aussi des neurones) peuvent éventuellement faire partie d'un système cognitif, mais *en tant que telles* elles n'ont absolument rien de

cognitif. Il ne suffit pas non plus d'analyser les interactions locales entre ces composants. Ainsi, par exemple, les interactions stéréochimiquement spécifiques des immunoglobulines, entre elles et avec d'autres molécules, sont fondamentales pour le fonctionnement dynamique du système ; mais encore une fois, *en tant que telles*, elles ne sont pas cognitives. Il est extrêmement important de bien comprendre qu'à *ce niveau d'organisation* il n'y a absolument rien qui distingue un "antigène du soi" d'un "antigène externe" : les forces physico-chimiques qui interviennent sont rigoureusement de même nature.

Si des interactions moléculaires, en tant que telles, ne sauraient être cognitives, l'activation d'une cellule par l'interaction de ses récepteurs moléculaires situés dans la membrane avec d'autres molécules n'est pas non plus cognitive, quelle que soit la spécificité de cette interaction. Il ne devient légitime de parler de telles activations comme les instances d'une "reconnaissance" que si elles sont situées dans le contexte d'un système global qui est effectivement cognitif. Même le système immunitaire tout entier n'est pas, en lui-même, cognitif : comme nous l'avons vu, les questions décisives portent sur les interactions entre le système et son environnement. Par conséquent, pour qu'il soit possible de parler de cognition, il faut se situer au niveau d'organisation des perceptions et des actions d'un système *dans sa relation à son environnement* : la question clé est de savoir si le couplage entre action et perception est tel que le comportement émergent du système *dans son milieu* satisfait une contrainte significative.

S'il est possible de répondre positivement à cette dernière question — et j'ai suggéré qu'il se peut que ce soit le cas pour le système immunitaire — alors (selon la définition employée ici) il est légitime de parler de cognition. Dès lors, il se pose une question du plus grand intérêt : quel est "l'objet" de cette cognition ? Dans le cas du système immunitaire, la réponse est celle-ci : l'incorporation des antigènes appartenant au corps de l'organisme dans les configurations dynamiques de clones de lymphocytes qui sont constitutives d'une identité du soi immunologique ; et la distinction entre ces antigènes et des antigènes externes qui déclenchent des réponses immunitaires destructrices. Il convient de souligner un trait tout à fait remarquable de cet "objet" : il ne peut être spécifié que

par référence à des catégories constituées par le fonctionnement du système immunitaire lui-même. Ainsi, le système immunitaire est capable de perception et d'apprentissage ; il est certainement adaptatif, dans le sens qu'il contribue de manière importante à la survie de l'organisme dont il fait partie ; mais il ne perçoit, il n'apprend, et il ne s'adapte à rien de pré-existant ou de définissable dans le domaine d'une réalité purement objective et indépendante. Pour résumer par une phrase : l'objet et le sujet de la cognition immunitaire sont inséparables.

V.2. Le cas de la cognition neuronale

Peut-on étendre ces conclusions à la neurophysiologie ? Examinons de plus près quelles en sont les implications.

Pour commencer, cette perspective signifie que des neurotransmetteurs et des canaux ioniques dans des membranes neuronales ne sont pas, en eux-mêmes, cognitifs. Des neurones et des synapses ne sont pas non plus cognitifs ; ni même, considéré isolément en tant que tel, le cerveau¹⁰. Afin d'être cognitif, le système nerveux doit être considéré dans le contexte fonctionnel d'un animal *dans son milieu naturel* réalisant des actions guidées par ses perceptions. A cet égard, il est certainement significatif que des neurones font leur première apparition au cours de l'évolution dans le rôle de relier des entrées sensorielles à des actions locomotrices. On remarquera ici que des actions de ce type donnent lieu, via les déplacements de l'animal dans son milieu, à des modifications corrélatives de ses perceptions sensorielles. Dans une étude classique, Held et Hein (1958) ont montré que des chatons ne peuvent apprendre à voir que si leurs perceptions sont corrélatives de leurs propres actions. La signification d'un système nerveux central doté d'une architecture connexionniste est peut-être qu'il permet à l'animal d'identifier ce genre de corrélations entre ses actions et ses perceptions, et par conséquent de développer des catégories conceptuelles fondées sur son engagement actif dans le monde. Nous verrons par la suite (section V.3b) que cette

¹⁰ Comme Maturana (communication personnelle) aime à le rappeler, "l'esprit n'est pas dans le cerveau".

perspective ouvre la voie à une nouvelle hypothèse concernant la nature des "représentations". Dans cette optique, le fait d'interpréter le fonctionnement du système nerveux en termes de "neurones grand-mère", ou de "détecteurs" de traits existant objectivement en tant que tels dans le monde extérieur, est une erreur malencontreuse.

Cette brève évocation suffit peut-être pour suggérer que la perspective proposée ici, bien qu'apparemment négative à première vue en ce qui concerne une grande partie des recherches contemporaines en neurophysiologie, peut servir en fin de compte à promouvoir l'intégration des neurosciences au cœur des sciences cognitives en focalisant l'attention sur ce qui est *cognitivement significatif* dans les mécanismes moléculaires et cellulaires sous-jacents.

V.3. *Les implications pour la cognition supérieure*

J'ai indiqué, à la fin de la section IV.1, ci-dessus, que l'approche proposée ici comporte des implications positives et précises pour l'étude de la cognition supérieure. Je souhaite à présent examiner ce qui peut justifier cette affirmation.

Dans les théories computationnelles de l'esprit habituelles, les représentations symboliques sont définies comme des éléments primitifs, dotés *a priori* d'une structure syntaxique analogue à celle des mathématiques formelles (Lakoff 1987). Par contre, la définition élémentaire de la cognition proposée ici ne fait appel ni à des symboles, ni à des représentations. La voie est ainsi ouverte pour l'étude de l'émergence des différents aspects de la cognition supérieure comme autant d'événements naturels survenant au cours de l'évolution biologique et de l'histoire humaine. J'évoquerai brièvement trois phénomènes émergents d'une importance particulière : la communication sémiotique ; les représentations ; et le langage. Je terminerai par quelques commentaires sur les applications technologiques qui peuvent découler de cette approche.

a) *La communication.* Etant donnée une population d'agents cognitifs interagissant entre eux, le schéma élémentaire de perceptions-actions contraintes peut s'élargir pour inclure, dans le répertoire des actions, *l'émission de signaux*, et dans celui des perceptions, *l'interprétation de ces signaux*. Dans certaines

conditions très restreintes, portant sur les perceptions particulières qui déclenchent l'émission d'un signal, et les modifications particulières des actions qu'entraîne la perception d'un signal, il peut en résulter une *coordination des actions* des agents qui contribue de manière déterminante à leur viabilité individuelle et collective. Un bon exemple est fourni par "l'intelligence d'essaim" exhibée par les insectes sociaux (Deneubourg et al. 1991). Cette approche de la compréhension de la communication comme phénomène émergent se prête bien à la simulation par ordinateur (Booth & Stewart 1993, Victorri & Cazoulat 1993). Il est à noter que les signaux en question (par exemple les phéromones des fourmis) sont tout à fait arbitraires (Saussure 1985) ; leur signification sémiotique ne peut être attribuée à une quelconque "information" encodée dans leur structure (Stewart et al. 1988), mais plutôt à leur rôle comme médiateurs d'une coordination d'actions (Maturana et Varela 1980). Il est également à noter que, contrairement à ce qu'avancent les théories computationnelles de l'esprit, l'organisation sociale est fondamentale tant comme condition de possibilité que comme conséquence de la communication (Havelange 1994).

b) Les représentations. Les théories computationnelles de l'esprit sont dans l'obligation de recourir à un postulat métaphysique extrêmement fort, à savoir qu'il existerait une réalité purement objective pouvant servir de référent aux connaissances d'un sujet. Dans cette optique, les représentations symboliques acquièrent leurs contenus sémantiques en vertu de leurs relations de correspondance avec différents aspects de cette réalité objective. Par contre, dans l'approche présentée ici, les "objets" des représentations ne sont pas séparables du sujet cognitif. Etant donné un agent cognitif élémentaire, engagé dans des perceptions-actions de telle sorte qu'il parvienne à maintenir sa viabilité, les actions de l'agent conduisent à des *conséquences* déterminées pour ses propres perceptions ultérieures. Dans cette autre optique, les "représentations" sont des représentations non pas d'une réalité externe "en soi", mais de ces conséquences qui, remarquons-le, sont très commodément formulées dans des termes directement accessibles à l'agent. Un agent doté de ce genre de capacité représentationnelle sera capable d'anticipations, et même d'élaborer une séquence d'actions destinée à réaliser un état perceptuel spécifié à l'avance comme but à

atteindre. Ainsi, les "représentations" correspondent à une sorte de "dédoublé interne" de la cognition élémentaire. En pratique, la construction de telles représentations requiert probablement un cerveau dans le cas des organismes vivants ou bien, dans le cas des robots artificiels, un réseau connexionniste qui génère de manière autonome une série indéfinie d'états perceptuels anticipés et actualisés pouvant servir comme base d'exemples pour un processus d'apprentissage des poids synaptiques (Mel 1990). Il s'ensuit que l'émergence des représentations, et d'actions intentionnelles corrélatives, est certainement un événement tout à fait précis et identifiable, qui est probablement survenu au cours de l'évolution des vertébrés quelque part entre les lamproies (les vertébrés les plus primitifs, qui n'ont pratiquement pas de tête) et les grands singes qui possèdent manifestement des intentions (Lestel 1993). Dennett (1987) a proposé de réduire l'intentionnalité au statut d'une simple projection épistémique : selon lui, qualifier une entité "d'intentionnelle" signifie seulement qu'il nous est commode de décrire son comportement en employant des termes intentionnels. Par son propre aveu, ceci impliquerait que des thermostats, par exemple, sont "intentionnels" ; ce qui revient à virtuellement éliminer l'intentionnalité comme concept scientifique. Selon l'optique envisagée ici, ceci relève d'un pessimisme indu ; ni des thermostats, ni des bactéries, ne possèdent des "représentations" (dans le sens précis employé ici), et par conséquent il existe des raisons scientifiques sérieuses pour dire que leurs actions ne sauraient être intentionnelles. A l'inverse, dans les cas où il peut être démontré que des représentations sont effectivement employées pour élaborer des séquences d'actions destinées à réaliser un but spécifié à l'avance, l'intentionnalité peut être réhabilitée comme un concept significatif.

c) *Le langage*. Quelle est la différence entre la communication sémiotique des animaux, telle que nous l'avons évoquée dans (a) ci-dessus, et la communication proprement linguistique si caractéristique des êtres humains ? Je suggère qu'une différence déterminante réside dans la très grande *variabilité* des significations associées aux signaux employés dans la communication linguistique. Dans la communication animale telle que nous l'avons définie dans la section V.3a ci-dessus, les conditions d'émission des signaux, et l'interprétation des signaux, sont toutes les deux non seulement

étroitement contraintes, elle sont extrêmement *stéréotypées*. Pour prendre un exemple autre que les fourmis, certains singes émettent trois types de cris d'alarme : l'un pour des aigles, un autre pour des serpents, et un troisième pour des félins. Les actions déclenchées sont également stéréotypées : respectivement se blottir par terre, sauter dans les arbres, et se cacher en scrutant anxieusement. Les actions ainsi coordonnées contribuent manifestement à la viabilité des animaux en leur permettant d'éviter au mieux chacun des trois types de prédateurs ; il s'agit donc d'un exemple de communication selon la définition proposée. La nature stéréotypée de cette communication n'est probablement ni innée ni acquise, mais résulte de l'organisation particulière du processus de l'ontogénèse qui garantit la régularité de la forme adulte pour des caractères tant anatomiques que comportementaux (Oyama 1985, Stewart 1993b). Dans la communication linguistique humaine, par contre, la signification des mots est tellement dépendante du contexte qu'à la limite on n'utilise guère deux fois le même mot pour dire exactement la même chose. Cette variabilité est à mettre en relation avec la nature profondément métaphorique des langues (Lakoff 1987). Etant donnée cette énorme variabilité, il est probable que nous ne parvenons à nous comprendre que parce que, dans toute conversation normale, nous sommes constamment en train de communiquer à *propos de* notre communication (Reddy 1979). Des phrases comme "Est-ce que tu veux dire que...?", "Non, je ne te suis pas", "Oui, je vois", etc, sont tout à fait essentielles à la communication ; sans parler des signaux non-verbaux (mais, selon cette optique, profondément linguistiques) comme des sourcils froncés, des clignotements des yeux, des hochements de tête, etc. Autrement dit, la communication linguistique se distingue de la communication animale par le fait qu'elle inclut normalement une méta-communication (Maturana et Varela 1980).

Ces caractéristiques de la communication linguistique humaine, qui sont intégrées très naturellement dans l'approche émergente esquissée ici, servent à consolider un point fortement souligné par Raccah (1993), à savoir que les langues naturelles (l'anglais, le français, le chinois etc.) sont des entités radicalement différentes des langages formels utilisés dans la logique mathématique et la programmation informatique. Ce point est capital, car les théories

computationnelles de l'esprit présupposent que les langues naturelles peuvent être assimilées à des langages formels (Montague 1970), avec comme corollaire que la syntaxe prime dans l'étude du langage (Chomsky 1957 ; pour un point de vue opposé, voir Lakoff 1987). En fait, ces théories computationnelles présupposent, de plus, que toute cognition s'exprime en termes d'un "langage de la pensée" (Fodor 1975) qui se prête à des computations symboliques. Il est salutaire de se rappeler qu'en réalité les langages formels sont une invention humaine extrêmement récente : ils n'ont abouti à leur forme canonique qu'au début du vingtième siècle, dans le contexte très particulier du formalisme mathématique et logique (Lakoff 1987). Leur redéploiement comme fondement d'un paradigme particulier en sciences cognitives est dû principalement à Turing et sa notion de "computation universelle". Il devrait être clair que, du point de vue exprimé dans cet article, le fait de prendre une théorie computationnelle de l'esprit comme fondement des sciences cognitives dans leur ensemble revient à mettre la charrue avant les boeufs.

d) Perspectives technologiques. Les théories computationnelles de l'esprit possèdent des applications technologiques immédiates et bien connues, comme les systèmes experts et l'Intelligence Artificielle. L'approche alternative proposée ici possède ses propres applications technologiques caractéristiques, notamment dans le champ des robots mobiles autonomes, qui sont moins bien connues et qui méritent d'être mentionnées. Un pionnier renommé dans ce champ a déclaré très explicitement que lors de la conception de tels robots, il est avantageux d'abandonner totalement les représentations symboliques du monde extérieur (Brooks 1987). En superposant des couches successives de couplages entre perceptions et actions, Brooks a pu fabriquer des robots capables d'apprendre à marcher sur leurs jambes dans un terrain irrégulier et inconnu. Les couplages entre perceptions et actions dans les robots de Brooks doivent être spécifiés par l'ingénieur humain ; des travaux récents de Cliff et al. (1992) suggèrent qu'il est possible de surmonter cette limitation en faisant évoluer une population de robots simulés en utilisant les techniques des algorithmes génétiques. Un autre développement récent est l'expérimentation avec des populations de robots mobiles qui "communiquent" selon la définition proposée dans la section

V.3a ci-dessus (Deneubourg, communication personnelle). Finalement, un projet de recherche à long terme vaut d'être mentionné en raison de ses implications théoriques. Dans la mesure où des robots autonomes du type mentionné ici "interprètent" des signaux en des termes qui sont directement reliés à leur propre "cognition", il n'est pas inconcevable qu'ils puissent un jour "comprendre" une langue appropriée d'une manière qualitativement différente des systèmes contemporains de l'IA, qui ne "comprennent" pas plus les langues que l'opérateur machinal dans la "chambre chinoise" de Searle (1987).

V.4. Conclusions générales

Ce survol rapide de quelques-unes des implications de la définition élémentaire de la cognition proposée ici pour l'étude de la cognition supérieure est partiel et extrêmement schématique. Néanmoins, il suffit peut-être pour montrer que mon but en proposant cette définition n'est nullement de réduire l'ensemble des phénomènes cognitifs au niveau de comportement exhibé par des bactéries. En fait, mon but est autrement plus ambitieux : il s'agit de renouveler la totalité des sciences cognitives en proposant une alternative radicale et complète aux théories computationnelles actuelles.

En conclusion, je souhaite réitérer dans une forme plus générale le résultat le plus important de l'analyse détaillée du système immunitaire qui constitue le sujet principal de cet article. En fait, cette conclusion générale découle inéluctablement de la définition élémentaire de la cognition comme perceptions-actions bouclées sous contrainte proscriptive. La perception, l'apprentissage et l'adaptation sont tout sauf arbitraires : ils sont fortement contraints par un "principe de réalité" qui, dans le cas des organismes vivants, concerne la viabilité du sujet cognitif (Aubin 1991, Varela 1989). Une condition *générale* pour la réalisation effective d'un phénomène cognitif est que, dans chaque cas, il instancie une stratégie *particulière, spécifique*, pour satisfaire une contrainte proscriptive. A l'intérieur du cadre fourni par une stratégie donnée, la contrainte prend la forme de limitations sévères sur les valeurs admises de tous les paramètres qui influencent l'émergence du comportement global

requis du système ; il se peut donc que ces paramètres donnent l'impression d'être soumis à une optimisation conduisant à une solution unique. Toutefois, il faut souligner qu'une telle optimisation est toujours *relative* à la stratégie particulière en question ; et qu'en général, il existe toujours une *pluralité* de stratégies potentielles. Suivant la stratégie trouvée et adoptée par un sujet cognitif, les objets de sa cognition seront différents. Il s'ensuit que les objets de la cognition (avec ou sans neurones) ne sont ni pré-existants, ni définissables en termes d'une réalité totalement indépendante et purement objective. Ces objets sont amenés à exister par les perceptions et les actions couplées du système cognitif lui-même.

John STEWART
Unité d'Immunobiologie
CNRS, Institut Pasteur, Paris

Remerciements

Les travaux sur la modélisation du système immunitaire décrits dans cet article ont été réalisés en collaboration avec Francisco Varela et Antonio Coutinho, et n'auraient pas été possibles sans leur soutien, amitié, et stimulation intellectuelle.

Bibliographie.

- Aubin J.P. (1991). *Viability Theory*. Birkhauser.
- Bernard J., Bessis M. & Debru C. (Eds) (1990). *Soi et non-soi*. Paris, Editions du Seuil.
- Boden M.A. (1987). *Artificial Intelligence and Natural Man*. 2nd Edition, Londres, MIT Press.
- Booth M. & Stewart J. (1993). Un modèle de l'émergence de la communication. Actes des Premières Journées Francophones "Intelligence Artificielle Distribuée et Systèmes Multi-Agents", Toulouse 7-8 avril 1993, pp. 9-18.
- Brooks R.A. (1987). Intelligence without representation. MIT Artificial Intelligence Report.
- Burnet F.M. (1959). *The Clonal Selection Theory of Acquired Immunity*. Cambridge, Cambridge University Press.

- Cliff D., Harvey I. & Husbands P. (1992). Incremental evolution of neural network architectures for adaptive behaviour. University of Sussex School of Cognitive and Computing Sciences Technical Report CSRP256.
- Chomsky N. (1957). *Syntactic Structures*. La Haye, Mouton.
- Coutinho A. Forni L., Holmberg D., Ivars F. & Vaz N. (1984). From an antigen-centered, clonal perspective of immune responses to an organism-centered, network perspective of autonomous activity in a self-referential immune system. *Immunol. Revs.* **79**, 151-169.
- Deneubourg J.L., Theraulaz G. & Beckers R. (1992). Swarm-made architectures. In : P. Bourguine & F. Varela (Eds), *Towards a Practice of Autonomous Systems*. Proceedings of the First European Conference on Artificial Life. MIT Press/Bradford Books.
- Dennett D. (1987). *The Intentional Stance*. Cambridge, MIT Press.
- ECAL 91. *Towards a Practice of Autonomous Systems*. P. Bourguine & F. Varela (Eds), MIT Press/Bradford Books.
- ECAL 93. *Self organisation and life : from simple rules to global complexity*. May 24-26, 1993, Brussels, Belgium.
- Fodor J. (1975). *The Language of Thought*. Crowell, New York.
- Havelange V. (1994). Sciences cognitives et tradition sociologique. *Revue Internationale de Systémique*, sous presse.
- Held R. & Hein A. (1958). Adaptation of disarranged hand-eye coordination contingent upon re-afferent stimulation. *Perceptual-Motor Skills* **8**, 87-90.
- Jerne N.K. (1974). Towards a network theory of the immune system. *Ann. Immunol. (Paris)* **125C**, 373-389.
- Klein J. (1982). *Immunology: the science of self-nonsel self discrimination*. New York, Wiley.
- Lakoff G. (1987). *Women, Fire and Dangerous Things*. Chicago, Chicago University Press.
- Langton C. Ed. (1989). *Artificial Life*. Addison Wesley.
- Lestel D. (1993). Une multimodalité problématique : communications symboliques des primates non humains, tromperie tactique et socialité postulée. 4^e Ecole d'Eté de l'ARC, "Communication et multimodalité dans les systèmes naturels et artificiels", Bonas, 4-17 July 1993.
- Maturana H.R. & Varela F.J. (1980). *Autopoiesis and Cognition : the realization of the living*. Dordrecht, Reidel.
- Mel B. (1990). *Connexionist robot motion planning : a neurally-inspired approach to visually-guided reachng*. New York, Academic Press.
- Meyer J.A. & Guillot A. (1990). Simulation of adaptive behavior in animats: review and projects. In Meyer & Wilson (Eds), *From animals to animats*. Proceedings of the First International Conference on the Simulation of Adaptive Behavior. MIT Press/Bradford Books.

- Montague R. (1970). Universal Grammar. *Theoria* **36**, 373.
- Oyama S. (1985). *The Ontogeny of Information : Developmental systems and evolution*. Cambridge, Cambridge University Press.
- Racah P.-Y. (1993). Significado e Inferencia Argumentativa. In: Bustos E., Garcia-Bermejo J.C., Perez Sedeno E., Rivadulla & Zofio J.L. Eds., *Panorama de logica, metodologia y filosofia de la ciencia*. Ciglo XXI, Madrid.
- Reddy M.J. (1979). The conduit metaphor - a case of frame conflict in our language about language. In: Ortony A. Ed., *Metaphor and Thought*. Cambridge, Cambridge University Press, pp. 284-324.
- Rosenfield I. (1989). *L'invention de la mémoire : le cerveau, nouvelles donnes*. Paris, Eshel.
- Searle J. (1987). Esprits, cerveaux et programmes. pp. 354-373 dans : Hofstadter D. et Dennett D. (eds.), *Vues de l'Esprit*. Paris, InterEditions.
- Sinding C. (1993). Les métaphores en biologie : analogies ou outils de pensée ? *Intellectica* **16**, 85-99.
- Stewart J., Andreevsky E. & Rosenthal V. (1988). Du culte de l'information en biologie et en sciences du langage. *Revue Internationale de Systémique* **2(1)**, 15-28.
- Stewart J. & Varela F.J. (1989). Exploring the meaning of connectivity in the immune network. *Immunol. Revs.* **110**, 37-61.
- Stewart J. & Coutinho A. (1991), A hundred years of immunology : paradigms, paradoxes and perspectives. In : P.A.Cazenave & P.Talwar Eds., *Immunology : Pasteur's Inheritance*. New York Wiley.
- Stewart J. & Varela F.J. (1991) Morphogenesis in Shape-space. *J. Theoret. Biol.* **153**, 477-498.
- Stewart J. (1992). Life=Cognition: the epistemological and ontological significance of Artificial Life. In: P. Bourguin & F. Varela (Eds), *Towards a Practice of Autonomous Systems*. Proceedings of the First European Conference on Artificial Life. MIT Press/Bradford Books.
- Stewart J. Ed. (1993a). *Biologie et Cognition*. *Intellectica* **16**.
- Stewart J. (1993b). Au-dela de l'inné et de l'acquis. *Intellectica* **16**, 151-174.
- Varela F. J. (1989). *Autonomie et Connaissance: Essai sur le Vivant*. Paris, Editions du Seuil.
- Varela F., Coutinho A., Dupire B. & Vaz N.M. (1988). Cognitive networks: immune, neural and otherwise. In: *Theoretical Immunology*. A.S. Perelson Ed., New York, Addison-Wesley, Vol II pp. 359-374.
- Varela F.J. & Coutinho A. (1991). Second generation immune networks. *Immunology Today* **12**, 159-166.
- Varela F., Coutinho A. & Stewart J. (1992). What is the immune network for? In: W. Stein & F. Varela (Eds), *Thinking about biology: an invitation to current theoretical biology*. Addison Wesley (SFI Series on Complexity), New Jersey.

Victorri B. & Cazoulat R. (1993). Auto-organisation et émergence des symboles.
Actes des Journées "Formation des symboles dans les modèles de la cognition",
Grenoble 10-11 juin 1993.