

5. Quatrième étape. L'enaction : une alternative à la représentation

D'autres motifs d'insatisfaction

Il est tentant de voir les STC actuelles comme deux camps en lutte : un paradigme prédominant bénéficiant des faveurs académiques et des investissements technologiques à grande échelle, et un groupe plus petit, plus jeune, de chercheurs et d'investisseurs commerciaux audacieux, avançant tous les deux dans une direction opposée.

Mais, pour deux bonnes raisons, cette description des STC est trop superficielle. D'abord, parce que la plupart des chercheurs appartenant au courant secondaire croient possible la synthèse entre le cognitivisme et le connexionnisme, les assimilant respectivement aux approches complémentaires descendante et ascendante. Le statut conceptuel d'une telle synthèse n'est pas clair, comme nous l'avons vu au dernier chapitre. Une alliance que permettrait

une forme de compromis entre un cognitivisme moins orthodoxe, suffisamment large d'esprit pour inclure le parallélisme distribué de bas niveau qu'offre l'approche auto-organisationnelle, demeure possible.

Une seconde et plus importante raison pour rejeter une telle description des STC est qu'il manque *encore* certaines dimensions essentielles de la cognition à chacune des deux orientations (et donc à leur synthèse éventuelle). Il faut au contraire encourager une approche entièrement différente en STC, qui vient d'une insatisfaction plus profonde que la quête du parallélisme distribué, et plus proche des véritables fondements des systèmes représentationnels. Il faut espérer que cette tendance naissante, qui profite aujourd'hui de quelques ressources, ne subira pas le même sort que jadis les idées sur l'auto-organisation, mises au rancart pendant trente ans.

La redécouverte du sens commun

L'insatisfaction principale de ce que nous appelons ici l'approche de l'enaction est simplement l'absence complète de sens commun dans la définition de la cognition jusqu'à ce jour. Pour le cognitivisme comme pour le connexionnisme actuel, le critère d'évaluation de la cognition est toujours la représen-

tation adéquate d'un monde extérieur prédéterminé. On parle soit d'éléments d'information correspondant à des propriétés du monde (comme les formes et les couleurs), soit de résolutions de problèmes bien définis qui impliquent un monde aussi bien arrêté.

Cependant, notre activité cognitive quotidienne révèle que cette image est par trop incomplète. La plus importante faculté de toute cognition vivante est précisément, dans une large mesure, de *poser* les questions pertinentes qui surgissent à chaque moment de notre vie. Elles ne sont pas prédéfinies mais *enactées*, on les *fait-émerger** sur un arrière-plan, et les critères de pertinence sont dictés par notre sens commun, d'une manière toujours contextuelle. Ces deux termes, enaction et faire-émerger, ne sont pas transparents dans le présent contexte. Le but de ce chapitre est de les expliciter en les examinant à la fois conceptuellement et par le biais d'exemples spécifiques.

Avant de nous lancer dans cette entreprise, il n'est

* Le premier de ces deux néologismes reproduit le terme anglais « enaction », en tentant de préserver la proximité entre « action » et « acteur » ; le second nous vient de la tradition de la phénoménologie qui traduit *hervorbringen* par « faire-émerger », ce terme s'opposant d'emblée à l'émergence par le verbe « faire », qui implique un geste du sujet et ne dépend pas uniquement d'une qualité intrinsèque de ce qui émerge (voir par exemple la traduction du texte de M. Heidegger, « Aletheia », dans *Conférences et Essais*, Paris, Gallimard, 1981) (N.d.T.).

pas inutile d'insister sur le fait que ceci est une critique de l'usage de la notion de représentation au cœur des STC, puisque seul un monde prédéfini peut être représenté. Si le monde dans lequel nous vivons se réalise naturellement plutôt que d'être prédéfini, la notion de représentation ne peut plus dorénavant jouer un rôle aussi central. Il ne faut pas sous-estimer le poids des conceptions auxquelles nous allons ici nous attaquer. Notre tradition occidentale tout entière a privilégié (avec, bien sûr, des variantes) cette idée que la connaissance est un miroir de la nature⁵⁴. Ce n'est que dans les plus récents travaux de certains penseurs continentaux (plus particulièrement M. Heidegger, M. Merleau-Ponty et M. Foucault) que la critique explicite de la représentation a commencé. Ces penseurs se préoccupent du phénomène de l'*interprétation* tout entier, dans son sens circulaire de lien entre action et savoir, entre celui qui sait et ce qui est su. Nous nous référons à cette circularité totale de l'action/interprétation par le terme de faire-émerger⁵⁵. De plus,

54. Pour une discussion lumineuse sur ce sujet, d'un point de vue anglo-saxon, voir R. Rorty, *Philosophy and the Mirror of Nature*, Princeton University Press, 1981 ; trad. française, *L'Homme spéculaire*, Seuil, Paris, 1990.

55. Le travail le plus conséquent sur cette question est celui de H.G. Gadamer, *Vérité et Méthode*, Paris, Éditions du Seuil, 1976. Pour une introduction claire à l'herméneutique, voir Palmer, *Hermeneutics*, Northwestern University Press, 1979. L'articulation de ce chapitre doit beaucoup à l'influence de F. Flores : voir

puisque cette perspective analytique se préoccupe spécialement de faire prédominer le concept de l'action sur celui de la représentation, il convient d'appeler cette nouvelle approche des STC *l'enaction*⁵⁶. Dans le monde anglo-saxon, traditionnellement plus proche de l'empirisme logique, ces thèmes ont le plus souvent été ignorés.

Au cours des dernières années, cependant, certains chercheurs en STC ont fait des propositions concrètes, tirant cette critique des sphères philosophiques et la reportant dans le laboratoire et sur la réévaluation de l'IA. Il s'agit d'une divergence plus radicale que jamais des STC, et qui va plus loin que les thèmes discutés pendant les premières années. En même temps, elle intègre naturellement plusieurs des outils développés dans le cadre connexionniste, comme nous allons maintenant le voir.

Le problème de la résolution de problèmes

On a toujours présupposé en STC que le monde peut être divisé en *domaines* : des régions d'éléments

T. Winograd et F. Flores, *Understanding Computers and Cognition*, *op. cit.*

56. L'appellation est loin d'être établie. Je la suggère ici pour les besoins de la discussion, jusqu'à ce qu'on en propose une autre qui soit plus adéquate.

discrets et des tâches que le système cognitif s'impose lui-même, agissant au sein d'un certain « espace » de problèmes : la vision, le langage, le mouvement, etc. Il est relativement facile de définir le domaine des échecs : c'est l'ensemble des états possibles dans l'« espace » du jeu d'échecs. Il y a des pièces et des positions sur l'échiquier. Il y a des règles gouvernant les déplacements et l'alternance des joueurs. Les limites sont clairement définies. Comme on le voit dans la figure 9, le « monde des échecs » est presque cristallin. Il n'est donc pas étonnant que le jeu d'échecs par ordinateur avance à grands pas, comme on le sait. Par opposition, cette approche s'est avérée vaine quand on l'a reportée sur, disons, le domaine des robots mobiles. Bien sûr, il est possible là encore d'identifier des éléments discrets (comme un châssis de métal, des roues et des fenêtres dans une chaîne de montage de voitures). Mais il est aussi clair que, tandis que le monde des échecs finit bien quelque part et sans équivoque, celui du mouvement parmi des objets est sans fin. Notre sens commun est constamment requis pour configurer notre monde d'objets. Par exemple, doit-on inclure les piétons dans notre monde de la conduite automobile ? Il apparaît tout de suite que la réponse à cette question ne peut qu'être tirée d'une masse confuse de considérations semblant irrémédiablement contextuelles : où est-on ? quelle heure est-il ? de quelle type de route est-il question ? ainsi de suite. Contrairement au domaine des échecs, celui

la conduite automobile ressemble plus à une courbe fractale infiniment détaillée qu'à un cristal nettement taillé (figure 9)⁵⁷. Le sens d'un mot en langue naturelle est peut-être le meilleur exemple de tous les éléments qui peuplent notre monde naturel : il faut bien connaître une langue pour appréhender les sens multiples d'un mot, et le même mot contribue par ailleurs à définir le sens de tous les autres mots. *Aucun* aspect de notre monde naturel et vivant ne peut être classifié à partir de délimitations nettes : on ne peut en faire un domaine dont on tracerait la carte.

En fait, dans les années soixante-dix, après deux décennies de très lent progrès, il apparut à plusieurs chercheurs en IA que même la plus simple action cognitive demande une quantité de connaissances apparemment infinie, que nous prenons pour acquise (qui est en fait tellement évidente qu'elle en est invisible), mais qui doit être servie à l'ordinateur à la petite cuillère. Comme nous l'avons déjà dit, le puissant héros est devenu l'enfant innocent qui apprend à se déplacer et à parler. Au début des années soixante, la recherche était inspirée par l'espoir cognitiviste d'un dispositif de résolution de problèmes *général*, un moteur logique qui pourrait être appliqué à la résolution de tout problème quel qu'il

57. Voir P. Biere, « The Professor's Challenge », *AI Magazine*, hiver 1985, p. 60-70.

soit. Lentement et humblement, ce premier rêve a dû être ramené à des domaines de connaissance strictement *locaux* et des problèmes à résoudre bien

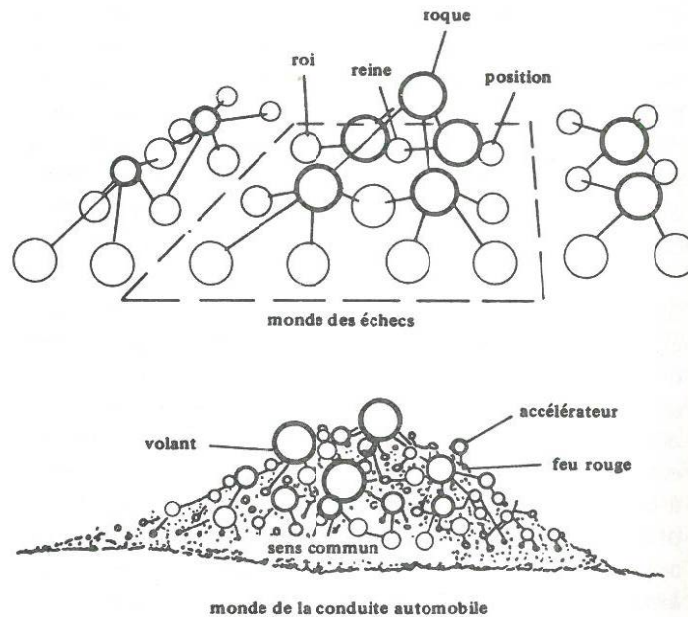


Figure 9 : Pour l'espace « échecs », il semble possible de dessiner un réseau de relations dont les nœuds représentent chaque élément pertinent. Dans le cas de l'espace « conduite automobile », une semblable tentative montre bien que, au-delà de quelques éléments isolés, le réseau évolue rapidement vers une masse non-circonsrite de sens commun. (D'après P. Biere, *op. cit.*)

définis, permettant au programmeur d'injecter dans la machine autant de son savoir général qu'il était nécessaire — par exemple : un système expert dédié uniquement à la réservation de billets d'avion. De même, la stratégie connexionniste demande qu'on restreigne l'ensemble des attracteurs possibles à partir de présuppositions sur les propriétés connues du monde, qui sont intégrées comme règles supplémentaires⁵⁸. Dans les *deux* cas, l'ambiguïté inabordable du sens commun est reléguée en marge de l'investigation, dans l'espoir d'une élucidation ultérieure.

Ces préoccupations ont une contrepartie philosophique bien définie. On doit aux phénoménologistes une discussion détaillée de la question : « Pourquoi la connaissance tient-elle du fait que notre monde soit inséparable de notre corps, de notre langage et de notre histoire sociale⁵⁹ ? » Il s'agit d'une interprétation continue qui ne peut être adéquatement encapsulée dans un ensemble de règles et de présuppositions, puisqu'elle dépend de l'action et de l'histoire ; c'est un monde de significations qu'on s'ap-

58. Pour un énoncé concis et formel, voir T. Poggio, V. Torre et C. Koch, *Nature*, 317, 1985, p. 314-319. L'inspiration originale vient de D. Marr, *Vision*, *op. cit.*

59. Les principales références que nous avons en tête sont M. Heidegger, *Basic Writings*, San Francisco, Harper and Row, 1977 ; M. Merleau-Ponty, *Phénoménologie de la perception*, Paris, Gallimard, 1976 ; Michel Foucault, *Surveiller et Punir, naissance de la prison*, Paris, Gallimard, 1975 ; H. Dreyfus, *Why Computers Can't Think*, New York, Macmillan/The Free Press, 1984.

propre par imitation et qui devient partie intégrante de notre monde préexistant. De plus, nous ne pouvons nous exclure du monde pour comparer son contenu avec ses représentations : nous sommes toujours immergés dans ce monde. En posant des règles pour exprimer l'activité mentale et des symboles pour exprimer les représentations, on s'isole justement du pivot sur lequel repose la cognition dans sa dimension vraiment vivante. Cela n'est possible que dans un contexte limité où presque tout est statique (les philosophes parlent de la condition *ceteris paribus*). Le contexte et le sens commun ne sont pas des artefacts résiduels pouvant être progressivement éliminés grâce à des règles plus sophistiquées. Ils sont en fait l'essence même de la cognition *créatrice*.

Si cette critique est juste, il s'ensuit que notre compréhension du fonctionnement normal de la cognition (et non de la cognition opérant exclusivement au sein d'environnements extrêmement limités) ne progressera pas à moins que nous partions d'autres préceptes que celui de la représentation du monde extérieur.

Exeunt les représentations

Le vrai défi posé aux STC par cette approche est sa mise en cause du préjugé le plus enfoui de notre

tradition scientifique, à savoir que le monde tel que nous le percevons est indépendant de celui qui le perçoit. Si nous devons au contraire conclure que la cognition ne peut être adéquatement comprise sans le sens commun, qui n'est rien d'autre que notre histoire physique et sociale, il nous faut en déduire que celui qui sait et ce qui est su, le sujet et l'objet, sont la spécification réciproque et simultanée l'un de l'autre. En termes philosophiques : le savoir est *ontologique*.

Dans notre discussion sur le cognitivisme, nous avons distingué deux notions de représentation. L'une est la notion généralement admise de représentation en tant qu'interprétation : la cognition correspond toujours à l'interprétation — ou à la représentation — de l'état du monde. L'autre, beaucoup plus forte, pose que cette activité cognitive ne s'explique qu'au prix de l'hypothèse selon laquelle un système agit à partir de représentations internes. Comme les implications de ces deux idées peuvent sembler identiques, nous allons en affiner la définition.

Nous pouvons commencer par le sens « faible » et non controversé du concept de représentation. Ce sens est purement *sémantique* : il réfère à tout ce qui peut être compris comme étant à propos de quelque chose. C'est là la définition de la représentation en tant qu'« interprétation », puisque rien n'est à propos de quelque chose sans être d'abord interprété

comme étant dans tel ou tel état. Une carte, par exemple, est exécutée à propos d'une certaine région géographique; elle représente certains aspects du territoire, et suggère une interprétation de ce territoire. De même, les mots sur une page représentent des phrases qui, à leur tour, représentent — ou sont à propos de — autre chose encore. Cette définition de la notion de représentation peut encore être précisée. Si, par exemple, nous nous intéressons aux langages dans un cadre plus formel, nous pouvons dire que les énoncés d'un langage représentent leurs conditions de satisfaction. Ainsi, l'énoncé : « La neige est blanche » — pris littéralement — est satisfait si la neige est blanche; l'énoncé : « Ramasse tes chaussures » est satisfait si les chaussures sont emportées par la personne à laquelle on s'adresse.

Cette acception de la notion de représentation est « faible » parce qu'elle ne véhicule aucune implication épistémologique ou ontologique. Il est ainsi parfaitement acceptable de parler d'une carte représentant un territoire sans se préoccuper de savoir, par exemple, comment les cartes acquièrent leur signification. Il est aussi parfaitement acceptable de penser à un énoncé comme à la représentation d'un ensemble de conditions sans se préoccuper davantage de savoir si telle est bien la nature du langage, ou s'il existe des faits dans le monde qui soient indépendants du langage et qui puissent être représentés par les énoncés d'un langage. En d'autres mots, ce sens faible de la notion de représentation est pragma-

tique; nous y recourons constamment sans problème.

La simple évidence d'une telle idée se transforme toutefois rapidement en une acception de la notion de représentation beaucoup plus forte, qui véhicule effectivement de lourdes implications ontologiques et épistémologiques. Cette acception « forte » survient quand par généralisation de l'acception faible on formule une théorie complète des mécanismes perceptuels, linguistiques ou cognitifs en général. Les implications ontologiques et épistémologiques sont principalement les deux suivantes. Nous présupposons que le monde est prédéfini, c'est-à-dire que ses propriétés sont établies préalablement à toute activité cognitive. Alors, pour expliquer la relation entre cette activité cognitive et un monde prédéfini nous posons l'hypothèse de l'existence des représentations mentales au sein du système cognitif (peu importe pour l'instant qu'il s'agisse d'images, de symboles, ou de schémas sub-symboliques d'activité distribuée sur un réseau). Nous disposons alors d'une théorie complète qui dit 1. le monde est prédéfini, 2. notre cognition concerne ce monde — même partiellement — et 3. notre cognition de ce monde prédéfini s'accomplit à partir de la représentation de ses propriétés, puis d'une action fondée sur cette représentation.

Pour les écoles représentationnistes, une entité cognitive est pour l'essentiel parachutée dans un monde préexistant. Cette entité n'y survivra qu'à la

condition de posséder une carte et d'apprendre à agir en fonction de celle-ci. En termes scientifiques, cela veut dire que la carte est un système de représentation qui est soit inné (à propos, disons, de l'espace, du temps, des formes et des odeurs), soit acquis par apprentissage (les schémas moteurs et émotionnels) et mis à jour (l'acquisition d'une langue spécifique ou d'une tâche) au fil de l'ontogénèse.

Plusieurs chercheurs en STC répliqueront que nous n'avons présenté qu'une caricature. Ne présumons-nous pas que la représentation soit statique, ce qui occulterait la richesse de détails de la structure interne du système cognitif, et représenterait injustement la représentation comme un simple miroir? N'est-il pas bien connu, par exemple, que la perception visuelle est le produit d'une fonction des schémas d'énergie physique, qui stimulent la rétine, sur les représentations de l'environnement visuel qui sont ensuite utilisées pour l'inférence et éventuellement pour produire un jugement perceptuel? La perception est un processus actif dans la production d'hypothèses, et non le simple miroir d'un environnement donné.

Cette objection est assez juste mais elle est à côté de la question. Nous ne cherchons pas à caricaturer un programme de recherche sophistiqué mais simplement à expliciter de la manière la plus claire possible certains présupposés épistémologiques non formulés. Ainsi, bien que chacun — sauf peut-être quel-

ques behavioristes d'arrière-garde — souscrive à l'idée que la représentation est un processus actif, ce processus est néanmoins envisagé comme un processus de « recouvrement » ou de « reconstitution » extrinsèque, indépendant des propriétés de l'environnement. Dans le domaine de la vision par exemple, on parle de « recouvrir la forme à partir de l'ombre », ou « la couleur à partir de la luminosité ». Ces propriétés sont ici considérées comme extrinsèques à l'environnement qui fournit l'information requise pour en recouvrir les propriétés visuelles d'« ordre supérieur », comme la forme et la couleur. L'idée fondamentale d'un monde aux propriétés prédéfinies demeure⁶⁰.

Le faire-émerger : des exemples

Considérons le cas de la vision : lequel vient en premier, le monde ou l'image? La réponse que donne la recherche en vision (qu'elle soit cognitiviste ou connexionniste) est indiquée sans équivoque par les noms des tâches étudiées : le « recouvrement de

60. Pour cette conception de la vision, voir David Marr, *Vision...*, op. cit. (New York, W. H. Freeman, 1982), spécialement l'introduction. Pour une explication philosophique de ces idées, voir Fred I. Dretske, *Knowledge and the Flow of Information*, Cambridge, Mass., Bradford Books/MIT Press, 1981.

la forme à partir de l'ombre », ou de « la profondeur à partir du mouvement », ou de « la couleur à partir de l'illumination ». C'est ce que nous pouvons appeler l'extrême de la *poule* :

La position de la *poule* : le monde extérieur comporte des règles fixes ; il précède l'image qu'il projette sur le système cognitif dont la tâche est de le saisir — le monde — de manière appropriée (que cela soit au moyen de symboles ou d'états globaux).

Remarquons maintenant à quel point cela semble raisonnable, et comment il serait difficile d'imaginer les choses autrement. On pourrait croire que la seule autre approche possible serait la position de l'*œuf* :

La position de l'*œuf* : le système cognitif crée son propre monde, et toute son apparente solidité repose sur les lois internes de l'organisme.

L'approche de l'enaction propose une voie moyenne⁶¹, ouvrant la voie au-delà de ces deux extrêmes en réalisant que (comme le savent tous les agriculteurs) l'*œuf* et la *poule* se *définissent l'un l'autre*, ils sont corrélatifs. C'est le processus continu de la vie qui a modelé notre monde par ces aller

61. Ceci est discuté dans mon « Living Ways of Sense Making : A Middle Way Approach to Neuroscience », dans P. Livingstone (ed.), *Order and Disorder*, Stanford, Anma Libris, 1984.

et retour entre ce que nous appelons, depuis notre perspective perceptuelle, les contraintes extérieures et l'activité générée intérieurement. Les origines de ce processus sont perdues à jamais et notre monde est, à toutes fins pratiques, stable (... sauf quand il s'écroule). Mais cette apparente stabilité ne doit pas faire oublier la recherche des mécanismes qui l'ont faite. C'est cet accent sur la codétermination (au-delà de la poule et de l'*œuf*) qui démarque le point de vue de l'enaction de toute forme de constructivisme⁶² ou de néo-kantisme biologique⁶³. Il est important d'en prendre note, car la philosophie plus ou moins réaliste qui sévit en sciences cognitives aura tendance à présumer que toute remise en question de la représentation vient *ipso facto* de son antipode, que hante le spectre du solipsisme.

Exprimons les mêmes idées à l'aide d'un exemple. Prenons le simple système auto-organisé introduit à la figure 6, et *relâçons* la contrainte qu'il doit être conçu *pour* une tâche particulière. Disons que le système existe, simplement, au sein d'un milieu contenant un assortiment d'éléments aléatoires. Retraçons ensuite l'historique du couplage de ce système particulier avec ce milieu aléatoire. Comme le

62. Voir par exemple P. Watzlawick (ed.), *L'Invention de la réalité*, Paris, Éditions du Seuil, 1988 ; New York, Norton, 1985.

63. Le plus clairement illustré par l'école de Vienne de Konrad Lorenz, comme on le voit par exemple dans *L'Envers du miroir*, Paris, Flammarion, 1975.

montre la figure 10⁶⁴, parmi l'infinité des séquences possibles, le système change d'état global (se déplace d'un attracteur à un autre) seulement lorsqu'il rencontre deux perturbations successives. En d'autres termes, le système a donné à un sous-ensemble d'événements (ceux qui sont constitués de deux perturbations successives) une signification particulière pour lui, puisque seul ce type d'événement altérera sa configuration de manière régulière. Cet exemple très simple illustre la naissance d'un « monde » pour le système, qu'il fait-émerger à partir de circonstances aléatoires, au fil de l'historique du couplage. Il n'y a, bien sûr, aucune « représentation » de ce sous-ensemble de séquences sélectionné parmi toutes les séquences aléatoires possibles du système, et ce n'était pas non plus sa tâche de le reconnaître. C'est simplement l'existence du système lui-même qui les a fait émerger à partir d'une masse indéfinie de possibles.

Pour prendre un exemple plus proche de nous, considérons le monde des couleurs que nous percevons tous les jours. Elles envahissent tellement notre vie qu'il est tentant de dire que les couleurs, comme nous les voyons, *sont* vraiment les couleurs du monde. Il est normalement admis que la couleur est

X 64. Figure tirée de F. Varela, *Autonomie et Connaissance*, op. cit., chapitre 11. Pour une discussion détaillée, voir F. Varela, « Structural Coupling in a Cellular Automaton », in E. Secarz, F. Celada, et M. Mitchinson (eds), *Semiotics and Cellular Communication*, Berlin, Springer-Verlag, 1988.

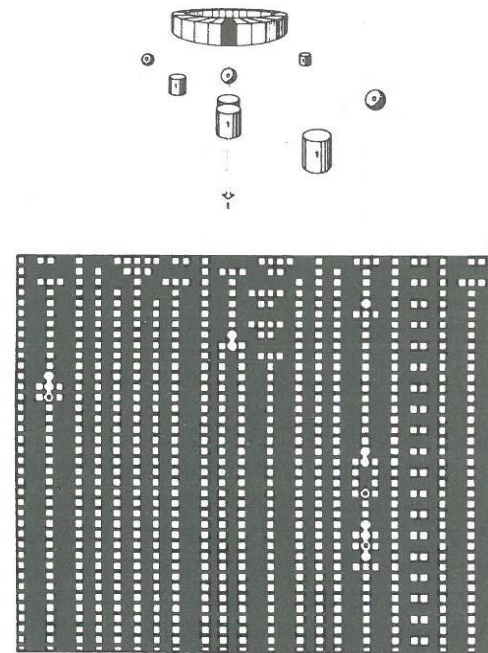


Figure 10 : Un anneau — une chaîne circulaire —, construit selon le principe illustré à la figure 5, est maintenant enveloppé d'un monde de perturbations aléatoires — représentées dans la partie supérieure du diagramme par des 0 et des 1. L'anneau entre en contact avec son monde avec chaque nouvelle perturbation, qui change l'état d'un neurone donné. Ainsi, au fil du temps, l'historique du couplage entre l'anneau et son monde s'ajoute à la dynamique interne propre à l'anneau lui-même (cf. figure 5). La partie inférieure de la figure donne l'exemple d'un tel historique de couplage. Ici, l'occurrence d'une perturbation est indiquée par des cercles. On remarque qu'une seule perturbation ou trois perturbations successives sur un neurone donné ne provoquent pas de changement global de la configuration de l'espace, tandis qu'une séquence de deux perturbations modifie cette configuration. Par un tel changement global, le système a sélectionné une configuration spécifique d'événements (à savoir deux perturbations successives) dont l'émergence est donc inséparable de l'historique du couplage du système avec ce monde aléatoire.

un attribut de la longueur d'onde de la lumière réfléchie par les objets, que nous captons et analysons en conséquence. En fait, comme il a maintenant été démontré, la couleur d'un objet telle que nous la percevons est dans une large mesure indépendante de la longueur d'onde⁶⁵. A la place, c'est un processus complexe (et seulement partiellement expliqué) de comparaison coopérative entre plusieurs ensembles de neurones du cerveau⁶⁶ qui spécifie la couleur d'un objet selon l'état cérébral global correspondant à la fois à une image rétinienne et à une certaine expectative de ce que l'objet est censé être.

Ainsi, par exemple, quand une feuille de papier gris est placée sur un fond rouge, le papier (physiquement) gris prend une teinte verdâtre, bien qu'aucune longueur d'onde verte n'ait bien sûr été ajoutée à la région. Ce phénomène est appelé l'« induction de la couleur » et est généralement interprété comme une illusion. Pourtant, ce type de processus est le cœur même de la perception de la couleur puisqu'elle apparaît lorsqu'un processus similaire active trois catégories cellulaires distinctes : leur activité relative spécifie la couleur en jeu. Les théoriciens modernes parlent alors de la couleur comme d'une représentation, non pas de la longueur d'onde mais

des propriétés réfléchissantes d'une surface, puisque cette propriété, indépendante de l'illumination, caractérise plutôt l'objet. Bien que cela soit clairement plus satisfaisant que l'explication de la longueur d'onde, la question n'est que repoussée d'un cran : comment le système définit-il préalablement ce qui doit être reconnu comme une surface ? Têtu, après qu'on l'ait mis à la porte, le sens commun rentre par la fenêtre.

Ce qu'on peut dire, c'est que notre monde chromatique est *viable* : puisque notre lignée biologique s'est perpétuée, c'est qu'il fonctionne. Ces opérations neuronales coopératives qui sous-tendent notre perception de la couleur sont le résultat d'une longue évolution biologique de notre branche de primates. Mais d'autres espèces ont élaboré des mondes chromatiques différents, en effectuant à partir de leurs organes sensibles des opérations neuronales coopératives différentes. Par exemple, plusieurs oiseaux sont apparemment tetrachromats (utilisant quatre couleurs primaires) alors que nous sommes trichromats (trois couleurs primaires nous suffisent)⁶⁷. Cela veut dire que, dans le domaine de la couleur, les oiseaux et nous ne sommes pas simplement plus ou

65. Voir par exemple E. Land, *Proc. Natl. Acad. Sci. (USA)*, 80, 1983, p. 5163-5169.

66. P. Gouras et E. Zenner, *Progr. Sensory Physiol.* 1, 1981, p. 139-179.

67. Pour un panorama comparatif de la vision de la couleur, voir G. Jacobs, *Comparative Color Vision*, Academic Press, 1983. Pour une discussion de la physiologie de la couleur chez les oiseaux, voir F. Varela *et al.*, *Arch. Biol. Med. Exp.*, 16, 1983, p. 291-303.

moins « fidèles » à un domaine supposé « identique » : nous habitons deux mondes perceptuels qui ne sont pas superposables puisque leurs dimensions sont différemment articulées.

Formulé autrement : l'historique bien différent du couplage structurel des oiseaux et des primates a fait émerger un monde de pertinence pour chacun d'eux qui est inséparable de leur vécu. La seule condition requise est que chaque itinéraire soit viable, c'est-à-dire qu'il soit constitué d'une séquence non interrompue de changements structuraux. Les mécanismes neuronaux sur lesquels repose la perception de la couleur ne correspondent pas à la solution à un problème (comme de trouver les propriétés chromatiques prévalentes des objets), mais à l'apparition simultanée de la perception de la couleur chez l'homme ou l'oiseau et de ce qu'on peut appeler les attributs chromatiques du monde habité.

L'olfaction est une autre dimension perceptuelle où ces idées sont à l'œuvre, non pas sur le plan de l'éventail comparatif que nous offre la phylogénèse, mais dans un champ d'exploration ouvert par les techniques électrophysiologiques nouvelles. Au cours de plusieurs années de travail, Freeman⁶⁸ a réussi à insérer un réseau d'électrodes dans le bulbe olfactif d'un lapin de sorte qu'une petite portion de l'activité globale puisse être mesurée tandis que

l'animal se comporte librement. Il a été trouvé qu'aucun schéma clair n'apparaît dans le bulbe à moins que l'animal entier et éveillé soit exposé aux odeurs. De plus, ces schémas émergents semblent se détacher d'un arrière-plan d'activité incohérente pour former un attracteur cohérent. Comme dans le cas de la couleur, l'odeur se révèle, non pas comme une fonction passive des propriétés extérieures, mais comme l'articulation créatrice du sens à partir de l'historique⁶⁹.

Il apparaît donc à cette lumière que le fonctionnement du cerveau est axé sur la constante enaction de mondes différents, fondée sur l'historique de lignées viables : un organe qui construit des mondes plutôt que les réfléchir.

Ébauche de la doctrine

L'idée fondamentale est donc que les facultés cognitives sont inextricablement liées à l'historique de ce qui est vécu, de la même manière qu'un sentier au préalable inexistant apparaît en marchant. L'image de la cognition qui s'ensuit n'est pas la résolution de problèmes au moyen de représentations,

68. W. Freeman, *Mass Action in the Nervous System*, Academic Press, 1975.

69. W. Freeman et C. Skarda, *Brain Res. Reviews*, 10, 1985, p. 145-175. Il est significatif qu'une section de cet article soit intitulée : « A Retraction on 'Representation' » [« Un désaveu de la 'représentation' »] (p. 169).

mais plutôt le faire-émerger créateur d'un monde, avec la seule condition d'être opérationnel : elle doit assurer la pérennité du système en jeu⁷⁰.

Question 1 : Qu'est-ce que la cognition ?

Réponse : L'action productive : l'historique du couplage structurel qui enacte (fait-émerger) un monde.

Question 2 : Comment cela fonctionne-t-il ?

Réponse : Par l'entremise d'un réseau d'éléments inter-connectés, capable de subir des changements structuraux au cours d'un historique non interrompu.

Question 3 : Comment savoir qu'un système cognitif fonctionne de manière appropriée ?

Réponse : Quand il s'adjoint à un monde de signification préexistant, en continu développement (comme c'est le cas des petits de toutes les espèces), ou qu'il en forme un nouveau (comme cela arrive dans l'histoire de l'évolution).

On remarque que deux notions nouvelles pour les STC apparaissent dans ces réponses. D'abord, puisque la représentation ne joue plus un rôle clef,

70. Pour une introduction complète à ce point de vue, voir H. Maturana et F. Varela, *The Tree of Knowledge : The Biological Roots of Human Understanding*, Boston, New Science Library, 1986.

l'intelligence ne se définit plus comme la faculté de résoudre un problème mais comme celle de *pénétrer* un monde partagé. Ensuite, c'est *l'évolution* qui remplace maintenant l'idée d'une structure fonctionnelle que justifie la réalisation d'une tâche. Bref, de même que le connexionnisme est issu du cognitivisme et d'une plus grande proximité avec le cerveau, l'approche de l'enaction fait un pas de plus dans la même direction pour englober aussi la temporalité de la vie, qu'il s'agisse d'une espèce (évolution), d'un individu (ontogenèse) ou d'une structure sociale (culture).

Travailler sans représentations

La recherche d'alternatives à la représentation pour étudier les phénomènes cognitifs (et il faut reconnaître que cela constitue une étiquette plutôt vague, un peu comme le connexionnisme) attire un relativement petit groupe de chercheurs venus de divers domaines. Il est intéressant de constater que ce groupe a grandi constamment au cours des dernières années. De plus, plusieurs des outils de la perspective connexionniste traditionnelle peuvent être reformulés dans ce cadre, de sorte que les lignes de partage sont aussi nettes ici qu'elles l'étaient entre les approches symboliques et connexionnistes.

Il est clair que la stratégie de l'enaction n'est

viaable pour l'IA qu'à la condition d'oublier les contraintes opérationnelles de la résolution de problèmes spécifiques prédéfinis par l'ingénieur. C'est le cas, par exemple, de ce que l'on appelle « systèmes de classification », dont l'objectif est l'articulation significative d'un monde indéfini auquel le système est confronté⁷¹. D'une manière plus générale, les simulations prolongées qui mettent en jeu l'histoire des couplages et les stratégies d'évolution, et qui ont pour but la découverte des principaux éléments de l'émergence du comportement intelligent, sont encore pour plus tard. Mais puisque ce cadre de recherche ne comporte pas de retombées immédiates, il intéresse peu le secteur compétitif de l'IA et peu de travail a été fait dans cette direction.

Une autre aire d'influence importante pour l'approche de l'enaction est celle de l'ordinateur et du langage. En fait, dans cette perspective, l'acte de communiquer ne se traduit pas par un transfert d'in-

71. Voir J. H. Holland, « Escaping Brittleness », *Machine Learning — Part 2*, R. Michalski, J. Carbonnel, et T. Mitchell (eds), Morgan Kaufmann, Los Altos, Californie, 1984. Pour plus de discussions sur certaines de ces idées (pas nécessairement dans la perspective proposée ici), voir J.D. Farmer et N. Packard (eds), *Evolution, Games and Learning: Models for Adaptation in Machines and Nature*, Physica D, 1986. Pour une idée de cette approche dans le cadre de l'étude du système immunologique, voir F. Varela, V. Sanchez-Leighton et A. Coutinho (1988), « Adaptive Strategies Gleaned from Immune Networks », *Evolutionary and Epigenetic Complexity*, Edinburgh University Press, A. Waddington Memorial Symposium, pp. 112-113.

formation depuis l'expéditeur vers le destinataire, mais plutôt par le modelage mutuel d'un monde commun au moyen d'une action conjuguée : c'est notre réalisation sociale, par l'acte de langage, qui prête vie à notre monde. Il y a des actions linguistiques que nous effectuons constamment : des affirmations, des promesses, des requêtes, et des déclarations. En fait, un tel réseau continu de gestes conversationnels, comportant leurs conditions de satisfaction, constitue non pas un outil de communication, mais la véritable trame sur laquelle se dessine notre identité. On peut ainsi exprimer l'essentiel de la vie quotidienne d'un bureau d'un point de vue qui s'apparente beaucoup plus à sa dynamique réelle que les classiques organigrammes de gestion⁷². Cette approche a donné naissance à une toute nouvelle espèce d'outils IA que l'on appelle les *coordonateurs*⁷³, dont le principe de base est de remplacer le courrier électronique habituel par un logiciel spécialisé dans l'analyse du flux conversationnel d'une entreprise. Les coordonnateurs sont un exemple de réorientation de l'IA, où l'on cesse d'exiger de l'ordinateur ce qui est peut-être impossible — à savoir la tâche non circonscrite de l'analyse linguistique —, et où on l'utilise plutôt comme véhicule structuré pour

72. Pour une discussion de cette idée, voir L. Smircich et C. Stubbart, *Acad. Manag. Rev.*, 10, 1985, p. 724-736.

73. T. Winograd et F. Flores, *Understanding Computers and Cognition*, *op. cit.*

la reconnaissance explicite de notre continuelle enaction linguistique.

Rapprocher l'émergence et l'enaction

Le rapprochement entre l'émergence et l'enaction dépend de l'idée qu'on se fait de la fonction d'un système distribué. En insistant sur le fait qu'un processus historique fait émerger des régularités sans contrainte de finalité arrêtée, on conserve la notion biologique d'un monde non circonscrit. En mettant en valeur, au contraire, l'acquisition par un réseau d'une faculté bien spécifique dans un domaine défini, on retrouve le principe de la représentation, ainsi qu'une appréhension plus habituelle des modèles connexionnistes. Toutefois, la première interprétation implique aussi une perspective complètement nouvelle sur la nature de la cognition, comme nous l'avons vu plus haut.

Ainsi, la route à suivre dépend largement de l'intérêt qu'on aura à se rapprocher de la réalité biologique, et de s'éloigner des considérations pragmatiques de l'ingénieur. Bien sûr, la définition d'un domaine circonscrit dans lequel un système connexionniste pourrait fonctionner reste possible, mais cela occulterait les questions plus profondes sur les origines, qui sont fondamentales dans la perspective de l'enaction.

Considérons par exemple la théorie de l'harmonie

de Smolensky. Son point de vue sur la computation sub-symbolique pour modéliser l'intuition semble tout à fait en phase avec la perspective de l'enaction, ce qui en fait le cas idéal pour notre comparaison. Cependant, même la théorie de l'harmonie est évaluée en référence à un aspect de la réalité de l'environnement qui est pris pour acquis : des propriétés exogènes correspondant à des propriétés prédéfinies du monde, et une activité endogène qui atteint au fil de l'expérience un état de signification abstraite, une « codification optimale de la régularité de l'environnement ». L'objectif est de trouver une activité endogène qui corresponde à la « caractérisation optimale » du milieu⁷⁴. La perspective de l'enaction demanderait ici que ce genre de système cognitif soit transposé dans une situation où l'endogène et l'exogène se définissent mutuellement au fil d'un historique prolongé, ne nécessitant qu'un couplage viable, et ignorant toute forme d'adéquation optimale⁷⁵.

Certes, du point de vue d'une IA aux contingences

74. P. Smolensky, *op. cit.*, p. 260.

75. Il vaut la peine de remarquer que des arguments similaires peuvent s'appliquer à la pensée évolutionniste. Pour les parallèles entre le représentationnisme cognitif et l'adaptationnisme évolutionniste, voir F. Varela, in P. Livingstone (ed.), *Order and Disorder*, *op. cit.* Pour une excellente discussion sur les mêmes questions telles qu'elles s'appliquent à l'évolution et au développement, voir S. Oyama, *The Ontogeny of Information*, *op. cit.*

pragmatiques, disposant d'un court délai pour produire un système qui sera opérationnel dans un domaine donné, cette approche semble sans intérêt. Mon argument est que les propriétés cognitives émergent des systèmes vivants indépendamment de telles préoccupations d'optimisation. Elles découlent de l'historique de compensations viables qui crée des régularités, mais il est loin d'être évident qu'elles puissent être associées à un référent unique.

Il existe donc une tension entre deux mondes parallèles de la recherche, où le choix pour ou contre l'approche de l'enaction est influencé à la fois par les complications qu'implique un remaniement conceptuel et par le monde technologique, dont la camisole de force de la mise en opération immédiate limite l'exploration. Il me semble probable que cette tension sera résolue à mesure que se creusera davantage le fossé entre les éléments technologiques et scientifiques des STC⁷⁶.

76. Voir aussi les remarques de Roger Schank dans *AI Magazine*, automne 1985, p. 122-135.

6. Conclusions

Nous sommes partis du noyau dur des STC pour nous déplacer vers ce qui peut être considéré comme leur périphérie, c'est-à-dire la prise en compte du contexte immédiat et les effets de l'historique biologique et culturel sur la cognition et sur l'action. Bien sûr, ceux qui tiennent à la représentation comme idée pivot considèrent que ces préoccupations ne sont que temporairement exclues de la région mieux définie de la résolution de problèmes qui leur semble plus accessibles ; d'autres vont jusqu'à prétendre que des enjeux aussi « embrouillés » et « philosophiques » n'ont pas de place dans les sciences cognitives sérieuses.

Certaines des distinctions qui sont à l'origine de ces tensions peuvent être énoncées de la façon suivante :