Cognition située

(morceaux choisis et librement traduits/commentés du livre de Clancey (Clancey 1997))

Clancey, W. J. (1997). Situated Cognition: On Human Knowledge and Computer Representations, Cambridge University Press.

1 Introduction

La théorie de la cognition située, telle que présentée ici, prétend que chaque pensée ou action humaine est adaptée à son environnement, c'est-à-dire située, parce que ce que les personnes perçoivent, comment ils conçoivent leur activité, et ce qu'ils font physiquement, tout cela se met en place ensemble. De ce point de vue, penser est une compétence physique comme faire de la bicyclette.

En faisant de la bicyclette, chaque coup de pédale et la posture sont contrôlées non pas par manipulation d'équations apprises à l'école, mais par une re-coordination de postures précédentes, manières de voir et séquences de déplacements. De manière similaire, dans le raisonnement, lorsque nous donnons un nom aux choses, assemblons des phrases en paragraphes, et interprétons ce qu'une phrase veut dire, chaque étape est contrôlée non pas par l'application de règles de grammaires ou de schémas précédemment appris, mais par re-coordonnant en les adaptant de précédentes manières de voir, parler et se déplacer. Tout action humaine est au moins partiellement improvisée par le couplage direct de la perception, de la « conception » et du déplacement – un mécanisme de coordination non médié par des descriptions d'associations, de lois ou de procédures. Ce mécanisme complète les processus inférentiels de délibération et de planification qui forment la colonne vertébrale des théories de la cognition fondée sur la manipulation de descriptions. Le couplage direct de la perception, de la conception et du mouvement implique une sorte « d'auto-organisation avec une mémoire » que nous ne savons pas encore émuler avec un ordinateur et d'ailleurs avec aucune autre machine.

Selon l'idée simple que la connaissance consiste en modèles descriptifs – descriptions de comment le monde apparaît (tels que des symptômes associés à une maladie infectieuse donnée-) et en descriptions de comment se comporter dans certaines situations (comme comment un médecin sélection un antibiotique pour traiter telle ou telle maladie) – les sciences cognitives firent de rapides progrès dans les années 70 et 80. Cette approche, souvent appelée modélisation cognitive symbolique, étayée par des études empiriques, a contribué à notre compréhension de l'expertise humaine. Les études cognitives ont révélé à quel point le raisonnement d'un novice est différent du raisonnement d'une expert, à quel point les personnes apprennent de leurs erreurs, et comment un enseignant sélectionne des exemples pour corriger un mauvaise compréhension d'un élève. Plus généralement, la modélisation descriptive a révélé comment les personnes mettent en relation les mots et leur signification quand on les lit, comment , en résolution de problème les buts sont mis en relation, de manière opportuniste et stratégique, avec des plans et des ressources limitées, comment les décision sont prises face à des données ambiguës, incertaines, etc. De nombreux logiciels exploitent des modèles descriptifs pour automatiser les opération de routine en science, gestion et ingénierie, y compris le contrôle-commande d'usines de fabrication, l'évaluation de tableaux, la découverte de motifs dans les bases de données, etc. Bien que de manière moins

omniprésente que prévu dans les années 70, les assistants « intelligents », instructeurs, simulateurs… ont progressivement orienté l'ingénierie logicielle vers le paradigme « fondé sur les connaissances ». Bien évidemment donc, les approches descriptives sont valables et utiles, et toute révision de la théorie cognitive doit se bâtir sur elles.
Grossièrement, la cognition située est une perspective philosophique et une méthodologie d'ingénierie qui reconnaît la valeur des modèles descriptifs de la connaissance comme des abstractions, mais tentent de construire des « robots » d'une manière différente. En contraste avec l'approche symbolique (qui sera nommée « approche descriptive » dans la suite du texte), la théorie de la cognition située prétend que quand on modélise la connaissance humaine avec un ensemble de descriptions, telles qu'une collection de faits et de règles dans système expert, on décrit abstraitement comment le logiciel doit régir à certaines situations, mais on ne tient pas compte de toute la flexibilité venant du fait que perception, action et mémoire sont associées chez l'homme (dans le cerveau dit l'auteur…). Pour reprendre ce que disait Alfred Korzybski, la carte n'est pas le territoire. La conceptualisation humaine possède des propriétés qui sont en relation avec la coordination physique qui rend lla connaissance humaine différente des procédures écrites et des réseaux de mots dans un logiciel. Selon cette théorie qui trouve ses racines dans la psychologie fonctionnaliste de William James et Frederic C Bartlett et dans la philosophie pragmatique de Charles S. Peirce et John Dewey, le mécanisme de mémoire qui coordonne la perception et l'action humaines sont tout à fait différente d'une mémoire stockant des descriptions comme le suppose les modèles descriptifs de connaissance. Les descriptions sont en fait centrales dans le comportement humain, mais leur rôle n'est pas directement en rapport avec le contrôle de ce que nous faisons (même en tant qu'instructions). Plutôt, dans notre discours comme dans nos écrits, les descriptions nous autorise à étendre notre activité cognitive à notre environnement, maintenu actif, et organise des conception alternative dans notre processus mental, et ainsi se transforme en routines « non-pensées » et réactives.
De manière un peu en avance sur la suite du livre, il y a ici une question centrale, particulièrement excitant et source d'inspiration pour les constructeurs de robot : si la connaissance humaine ne consiste pas à stocker des descriptions, qu'est-ce qui fait la relation entre ce que nous disons et ce que nous faisons ? Parler doit être vue non comme montant ce qui serait déjà « à l'intérieur », mais comme une manière de changer ce qui « est » à l'intérieur. Parler n'est pas reprendre ce qui a été déjà rangé inconsciemment à l'intérieur du cerveau, mais est en soi une activité de représentation. Les noms que nous donnons aux choses et ce qu'ils signifient, nos théories, et nos conceptions se développent au travers de notre comportement lorsque nous intéragissons et percevons à nouveau ce que nous et les autres ont déjà précédemment dit et fait. Cette interaction causale est différente d'un processus linéaire « décrire ce que je perçois » ou « regarder ce que je conçois ». A la place de ce processus, les processus de regarder, percevoir, comprendre et décrire se déroulent simultanément en se modelant mutuellement. C'est ce qui est appelé la « perspective transactionnelle ».
Il y a donc une sorte d'entrelacement entre les manières de se développer des processus neuronaux et des comportements. Les processus neuronaux sont plus flexibles et adaptatifs dans la manière dont ils mettent en relation la conception et la perception qu'un ensemble de descriptions l'autorise, mais l'acte de décrire est néanmoins crucial pour réorienter le comportement humain. La première étape en dénouant cette relation récursive est de distinguer mémoire et stock de descriptions. Comprenant la nature de la mémoire humaine comme un mécanisme de re-coordination, on peut alors distinguer les rôles relatifs de la catégorisation neuronale et les manipulations de représentation dans l'environnement (telles que dessiner ou écrire). L'acte même de décrire la connaissance humaine change ce que nous connaissons (et c'est tant mieux). De plus reconnu comme un stock de pensées formalisées,

les description rendent obscures notre expérience effective (qui ne peut pas toute entière se mettre en mots).
Pour faire court, la congintion située est l'étude de comment la connaissance humaine se développecomme un moyen de coordonner l'activité depuis l'intérieur de l'activité elle-même. Cela signifie que le feedback – survenant de manière interne et avec l'environnement selon les moments- a une importance fondamentale. La connaissance possède donc une dynamique aussi bien dans sa formation que dans son contenu. Ce changement de perspective d'une connaissance comme un artefact stocké à la connaissance comme des capacités construites dans l'action inspire une nouvelle génération de cybernéticiens dans les champs des robots situés, de la psychologie écologique et de la neuroscience computationnelle. Les études empiriques considèrent ensemble la compréhension d'interactions survenant aux différents niveaux à l'intérieur comme à l'extérieur du cerveau.
La connaissance humaine est, naturellement, plus complexe que l'exemple de la pratique de la bicyclette le suggére : contrôler les capacités sensorimotrices, créer et interpréter les descriptions (tels que les explications et conseils des parents sur l'équilibre), et participer à une matrice sociale (telle que devenir un « grand » en rejoignant la bande en bas de l'immeuble), tout cela est mis en relation et dynamiquement recomposé par une coordination conceptuelle. Ces organisateurs de comportement interviennent en parallèle, tout le temps, s'influençant les uns les autres. Trois formes de feedback sont ainsi mises en évidence :

- Les action à court terme change le flux des données sensorielles (les forces et les vues sont changées par le déplacement à venir de la bicyclette).
- La perception et la conception sont dynamiquement couplées (le bord du trottoir apparaît comme une limite ou comme une occasion de faire un saut selon la manière dont vous concevez comme dangereux le trafic qui est derrière vous).
- Les buts et les significations sont reconçues selon la manière dont les transformations faites à l'environnement sont reperçues avec le temps (de nombreuses traces de vélo sur une colline peuvent indiquer l'endroit comme « super pour faire du vélo » pour les enfants et apparaître comme une indication qu'il va falloir entretenir ce endroit pour le conseil municipal).

En réexaminant la nature de feedback, la recherche sur la cognition située explore l'idée que la connaissance conceptuelle, en tant que capacité à coordonner et séquencer le comportement, est intrinsèquement vue comme une partie de et au travers des performances physiques. La formation de catéforisations perceptuelles et leur couplage aux concepts procurent le matériau pour rasionner (inférence), ce qui alors change ce que nous cherchons et ce que nous sommes capables de trouver.
Dans ce livre,, j'expose ces idées en réexaminant la nature de la modélisation cognitive descriptive, interprétant la preuve biologique et les arguments philosophiques, et en critiquant la conception des nouveaux robots. En élaborant les notions de feedback et de couplage causale, je montre comment la cognition située facilite la résolution de vielles controverses sur la nature du sens et éclaire bien les récents débats sur les bases du symbole, la perception direct et l'apprentissage situé.

Comparaison de la connaissance humaine et les représentations en machine
Pour conclure cette introduction, je fais un petit rappel pour les lecteurs qui ne seraient pas familiers avec l'histoire de la modélisation cognitive descriptive et introduis ma propre approche.
Les communications scientifiques et les livres sont finalement des déclarations personnelles, situant le développement des pensées de l'auteur le long d'un chemin qui est maintenant vu comme naïf vers ce qui est vu comme une redirection pleine d'espoir. De 1974 à 1987, je

faisais partie d'une communauté de chercheurs en Intelligence Artificielle qui développèrent des logiciels qui étaient capables de diagnostiquer des maladies, permettre d'enseigner en s'appuyant sur des études de cas, de modéliser les méthode de résolution de problème des étudiants, … Seon la rugrique « systèmes à base de connaissance », nous pensions non seulemet que la connaissance pouvait être représentée par des règles, mais également que ce stock de règles serait fonctionnellement équivalent à ce qu'un médecin expert peut faire. Nous savions que le médecin savait plus de choses, mais nous faisons l'hypothèse que cette connaissance représentait plus de règles.
L'hypothèse que la connaissance humaine consiste exclusivement de mots organisés en réseaux de règles et de schémas de descriptions (frames) guida la création de centaines de logiciels, décrits dans des douzaines de livres… Bien entendu, ces chercheurs rélaisaient que les processus de coordination physique et de perception impliqués dans les capacités motrices ne pouvaient pas facilement être reprodutis par des schémas et des règles. Mais de tels aspects de la cognition étaient considérés comme périphériques ou des problèmes d'implantation. Selon ce point de vue, l'intelligence est mentale, et le contenue de la pensée consiste en réseaux de mots, coordonné par une architecture pour apparier, rechercher et pour l'application de règles. Ces représentations, décrivant le monde et comment se comporter, servent de connaissance pour la machine, tout comme ils sont la base du raisonnement et du jugement humain. En utilisant cette approche symbolique pour constuire une intelligence artificielle, les modèles descriptifs non seulement représentent la connaissance humaine, ils correspondent à la manière d'une carte aux structures stockées dans la mémoire humaine. Selon ce point de vue, un modèe descriptif est une explication d'un comportement humain parce que le modèle EST la connaissance de la personne – stockée chez lui, elle contrôle directement ce que la personne voit et fait.
La distinction entre les représentations (connaissance) et l'implantation (biologie ou silicone), appelée l'hypothèse fonctionnaliste (Edelman,92), prétend que bien que les ingénieurs puissent apprendre beaucoup des processus biologiques pertinents pour comprendre la nature de la connaissance, ils seront pourraient au bout du compte fabriquer une machine exhibant des capacités humaines qui ne serait ni biologique ni organique. Cette stratégie rencontre un soutien considérable, mais malheureusement, la tendance été d'ignorer les différences entre la connaissance humaine et les logiciels et au contraire de « vendre » les logiciels existant comme « intelligents ». En insistant sur les ressemblances plutôt que sur les différences entre les personnes et les ordinateurs, plutôt que sur les différences, les chercheurs en IA on adopté une posture assez ironique qui consista à donner comme central dans la recherche la formalisation d'une analyse –fins/moyens comme une méthode de résolution de problème. Progresser en résolution de problème peut être vu comme décrire la différence entre l'état courant et le but courant et donc de faire une modification qui tente de combler la différence. Étant donné le focus mis sur l'inférence symboliques, les études cognitives se sont focalisées sur les aspects de l'intelligence qui ont un rapport avec les modèles descriptifs tels que les mathématiques, la science, l'ingénierie, la médecine, les domaines de l'expertise humaine… Se focaliser sur l'expertise humaine permet de soutenir l'idée que la connaissance se confond avec des modèles stockés, et en conséquence la distinction entre capacités physiques et connaissance est fondé sur une hypothèse, qui est instillée à l'école d'ailleurs, que la « connaissance réelle » consiste en faits scientifiques et théories. De ce point de vue, l'intelligence est concernée uniquement par des croyances bien construites et par des hypothèses raisonnées.
Mais comprendre la nature de la cognition nécessite de considérer plus que la résolution de problèmes complexes et l'apprentissage à partir de l'expertise humaine et leurs étudiants. D'autres sous-domaines de la psychologie cherche à comprendre des aspects plus généraux de la cognition, tels que la relation entre les primates et les humains, les disfonctionnements

neuronaux, et l'évolution du langage. Chacun de ces exemples nécessite des considérations plus importantes du fonctionnement du cerveau, et à son tour procure des éclairages utiles pour les constructeurs de robots. À cet égard, l'approche fins/moyens que je promeus est la continuation du but original de la cybernétique : comparer les mécanismes naturels et artificiels.
En confrontant les logiciels à ce que l'on a appris de d'études plus générales de la cognition, les chercheurs en science cognitive peuvent chercher à comprendre les différences entre la connaissance humaine et les meilleurs modèles cognitifs. Bien que des questions sur la relation entre langage, pensée et apprentissage sont très anciennes, les modèles computationnels offrent une occasion de tester des théories d'une nouvelle façon – en construisant un mécanisme en dehors des descriptions du monde et comment s'y comporter et voir si ça marche bien (voir Howard Gardner).
Gardner conclut que c'est à partir de ces comparaisons que les scientifiques cognitivistes devraient élargir substantiellement leur point de vue sur les processus mentaux. Ce livre est dans le même esprit, montrant notre talent à partir de ce que les logiciels IA font pour se demander comment de tels modèles de la cognition relèvent de la connaissance et de l'activité humaines. Je formule des stratégies pour combler le gap et proposer des avis pour utiliser de manière appropriée en utilisant la technologie du moment (à jour).
……

Le reste de l'introduction continue à argumenter sur l'importance de l'approche sans apporter d'autres démonstrations importantes

2 Représentations et mémoire

Aaron's drawing

Aaron est un robot conçu pour produire des dessins originaux. Harold Cohen est l'artiste et le programmeur qui développe Aaron depuis les années 70. En réalité, l'objectif de l'artiste est de « trouver la configuration minimum pour que des tracés soient compris comme des images ». En particulier, serait-il possible de faire percevoir des choses en 3D si le système n'avait pas a priori de représentations de ce qu'est le 3D ? En pratique, Aaron manipule des représentations 2D (il en a donc des modèles), mais leur disposition (la façon de les dessiner successivement) est pilotée pour que les objets à dessiner soient plus ou moins en arrière plan…
Cet exemple simple permet de montrer la différence entre un mécanisme interne simple et une observation extérieure qui prête à ce mécanisme une complexité qu'il n'a pas (en effet, les images dessinées automatiquement semblent l'œuvre d'un logiciel d'une grande complexité).

Les modèles internes utilisés par Aaron concernent des proportions entre taille des membres et du corps pour une personne, formes de base d'une plante, formes de base d'un arbre, etc…

Il est clair que l'artiste reste Harold Cohen. Pourquoi ? Parce que Aaron n'a aucune autonomie, n'a pas conscience de lui-même, ne sait pas « voir » ce qu'il fait et n'a aucune capacité d'apprentissage.

Pour synthétiser :

Tableau 1 Trois perspectives sur la cognition située

¨Perspective (point de vue)	Interprétation
(Analyse **fonctionnelle** de la forme) Sociale	Organisée par l'action et la perception interpersonnelle ; conceptuellement autour des relations sociales (normes, roles, motivations, chorégraphies, dispositifs participatifs)
(Analyse **structurelle** des mécanismes Interactive Prêt à fonctionner	 Relations dynamiquement couplées état-sens-action ; co-organisation réactive Connexion physiquement couplée, non objectivée (sans but explicite)
(Analyse **comportementale** de contenu) Inscrite dans le monde physique (grounded)	Situé dans quelque activité physique quotidienne, un cadre interactif spatio-temporel.

Si nous reprenons ces trois perspectives :

Perspective « fonctionnelle » : En psychologie, une fonction est liée à un but cognitif appelé tâche. Ainsi la fonction d'artiste entraîne en général un comportement lié à cette fonction (activité bien sûr, mais aussi comportement vestimentaire, occupations, loisirs, tout ceci étant socialement situé ; c'est un artiste !). Sa connaissance et son identité sociale sont étroitement liées.

Perspective « structurelle » : cette perspective s'intéresse à la manière dont la perception, la conception et l'action sont physiquement coordonnées. Par exemple, un visiteur d'une exposition va « voir » dans un dessin des personnages et des plantes alors qu'avec un effort d'abstraction, on ne pourrait voir que des lignes tracées (si par exemple, il est difficile d'identifier des formes connues).C'est « l'expérience » qui couple la perception et l'interprétation conceptuelle (conception). Les structures « toutes faites, prêtes à fonctionner » n'ont pas besoin d'expérience pour accomplir leur action, elles sont faites pour (le mécanisme de dessin par exemple, n'est pas une structure coordonnée au moment de l'action).

Perspective « comportementale » : ce troisième niveau met en relation la cognition et le cadre spatio-temporel. A la différence de l'aspect fonctionnel, qui prend en compte grossièrement la signification de l'action, ou l'aspect structurel qui considère les mécanismes internes, l'aspect comportemental considère le feedback local et la sensibilité au temps et au lieu de l'action. Par exemple, dans Aaron, le logiciel trace ses traits en fonction de ce qui a déjà été tracé à un endroit particulier (il va ainsi faire croire qu'un personnage est derrière un autre avec la simple règle de ne pas faire apparaître les poins qui seraient sur une zone déjà dessinée (ce qui revient à dire que les motifs dessinés les premiers sont au premier plan). De ce point de vue le comportement est réflexif et continuellement ajusté.

Pour construire un robot (un logiciel embarqué) capable d'apprendre et coordonner son comportement « comme un humain », nous devons tenter de comprendre mieux les trois déclinaisons du fait d'être en situation :

- Comment perception et déplacement sont reliés (aspect structurel)
- Comment ce processus de coordination physique est relié aux activités de conceptualisation, dont le contenu est intrinsèquement social (aspect fonctionnel).

- Comment les processus subconscients de perception et de conception sont reliés aux processus intrinsèquement conscients de représentation en parole, texte, dessin, etc.. (aspect comportemental).

Deux idées peuvent synthétiser les choses :
- Chez une personne, la re-coordination implique habituellement la conceptualisation (l'interprétation),
- La compréhension conceptuelle d'un lieu, d'une activité, d'un rôle et d'une valeur est développée et constituée socialement.

Il est important de noter que « l'identité » d'une personne est perpétuellement affectée par son activité et ce qu'elle perçoit de l'environnement.

La carte de Mycin

Introduction (simple) à la représentation des connaissances

Mycin est un logiciel développé dans les années 70.

De la même façon que l'on pouvait dire que Aaron était un artiste, on est tenté de dire que Mycin est un médecin.

Ce système est bien connu et s'exprime sous forme d'un ensemble de règles de production, de faits et d'un mécanisme général d'inférences. Les règles et les faits sont considérés comme des représentations de connaissance. Ces représentations sont exprimées dans un « langage » propre à Mycin. Ce langage est lui-même construit sur le langage LISP (et on pourrait continuer jusqu'à la représentation binaire qui sera « machinalement » traitée par l'ordinateur.

Une base de connaissance est une sorte de carte

Si on considère le plan d'un campus, on peut le voir comme une structure devant aider les gens à trouver leur chemin pour trouver des bâtiments et des lieux pour stationner. Si on cherchait du pétrole à cet endroit, il faudrait une autre carte…
Les systèmes à base de connaissance sont donc des modèles particuliers pour faire quelque chose de précis. La notion d'heuristique est ici principale puisqu'elle permet de « couper » l'espace des situations possibles en tenant compte de situations pour lesquelles on sait comment chercher une solution efficacement.

Modèles généraux versus spécifiques à une situation

Quand un logiciel du type Mycin tourne, il y a deux types de descriptions dans la mémoire du programme : les descriptions générales sur une maladie par exemple (règles) et des descriptions spécifiques à un patient par exemple (faits).
Tout ce qui est enregistré à propos d'une situation spécifique (le résultat des dialogues qui sont tracés) est appelé modèle spécifique à une situation.
Les règles générales s'appellent le modèle général.

Chercher dans un modèle descriptif

Pour trouver son chemin sur une carte, il faut pouvoir se repérer, reconnaître des points clés en général définis dans une légende.
Dans une base de connaissance, il faut aussi pouvoir disposer de clés d'accès, qui seront les termes à utiliser dans une requête. Pour connaître les clés possibles, il faut également une sorte de liste des clés possibles (ce qui est procuré par les interfaces des systèmes à base de connaissance). A partir d'une requête exprimée par les (mots) clés, on pourra déclencher les règles qui la satisfont qui à leur tour détermineront d'autres clés constituant des requêtes, jusqu'à aboutir à une association entre la requête initiale et un ensemble de recommandations liées aux heuristiques insérées dans la base de connaissance. En intelligence artificielle, on dit que l'on explore un espace de recherche pour la solution (la réponse à la requête). Les heuristiques orientent la recherche.
Les capacités « d'interprétation » d'un programme sont limitées aux relations que l'ingénieur concepteur a décrites entre les différents symboles utilisés pour représenter les éléments « cartographiées ». A partir d'un modèle général et d'éléments singuliers (spécifiques à une situation) pris comme « entrée » de la requête, le système cherche à indiquer quelle sera la situation singulière conforme au modèle général.

La carte n'est pas le territoire

La question est de savoir quelle peut être la relation qui existe entre les représentations internes (ou cartographiées de manière extérieure) d'une machine avec la connaissance d'un utilisateur ? Au début de l'ingénierie des connaissances, on considérait que les règles « externes » étaient issues de l'expert et qu'en conséquence si on codait correctement ces règles, la représentation interne était équivalente à la connaissance de l'expert (sur ce point particulier). En fait, ceci n'est pas vrai. La plupart du temps, la phase « d'acquisition des connaissances » est l'occasion de faire une synthèse de la connaissance des experts qui va bien plus loin que leur expertise : par exemple, les relations causales, temporelles sont mises en évidence alors qu'elles étaient probablement fortement implicites auparavant.
Il est maintenant admis que la représentation de la connaissance n'est évidemment pas la connaissance elle-même.

Expliquer les règles

Si on remplace les symboles « signifiants » pour l'homme (comme « hémorragie ») par un code non signifiant (comme « G189 »), ceci ne changera rien au déroulement du programme, mais il faudra faire le mapping vers un symbole signifiant pour l'utilisateur. Il n'est donc pas possible de vérifier en lisant le programme que « G120 implique G189 » est VRAI.
Un système ne sait donc pas une règle, mais il la mémorise et saura produire le calcul inférentiel si cette règle est sélectionnée. Le système ne comprend pas la règle (il ne peut pas l'expliquer). Ce serait pourtant utile, si on ne veut pas appliquer cette règle dans une situation où elle ne serait pas valide. Ceci peut être réglé par des « méta-règles » qui sont des règles sur les règles. Ces méta-règles envisagent différentes classes de situation (reconnues par des éléments d'une requête = clés) et sélectionnent les paquets de règles qui sont reconnus comme s'appliquant dans ces situations. Bien entendu, la complexité s'accroît et il est particulièrement difficile d'être exhaustif dans l'étude des classes de situation. Cela dit, il y a quand même là un grand pas en avant par rapport aux logiciels codant des situations « en dur ».

Tout irait à peu près bien si l'utilisateur exprimait ses requêtes en étant totalement d'accord (et conscient) des relations exploitées par le SBC pour faire ses calculs inférentiels. Mais l'utilisateur formule les symboles de sa requête en imaginant des relations qui peuvent être différentes de celles qui sont inscrites dans le SBC. Dans ces conditions, il peut y avoir difficulté d'utilisation, voire impossibilité si l'utilisateur ne « se représente pas » convenablement les relations inscrites dans le SBC.

Interprétation sémantique et syntaxique

Nous voyons bien que nous utilisons le mot « interpréter » différemment quand nous parlons d'un logiciel qui fait des inférences et d'un homme qui raisonne. Dans le premier cas, il s'agit d'une interprétation syntaxique (la reconnaissance des symboles déclenche le calcul des relations vers d'autres symboles qui…) sans qu'il puisse y avoir de problème avec ce qu'ils représentent dans le monde réel. Quand un utilisateur lit une règle (utilisant des termes signifiants), il commente la règle, explique pourquoi elle est vraie, comment on a pu l'établir, etc. (même dans le cas d'utilisateurs « exécutant », l'expérience leur fait découvrir des explications –parfois fausses- qu'ils donneront si on leur demande le bien fondé d'une règle).

Une conséquence pratique est qu'il faudrait (au moins) que les sources (documents donc !) expliquant l'origine d'une règle, puisse être gardées en même temps que la règle.

Pour régler en partie ce problème (de différence entre la représentation interne de la connaissance et l'idée que se fait un utilisateur de cette représentation), on peut voir les symboles comme des « formes » dénotant un concept : le concept peut être représenté différemment selon le moment, selon le lieu etc… Le comportement pourra ainsi être défini « en contexte ».
Cela laisse supposer toutefois que le niveau symbolique est stable et que le raisonnement serait donc de la manipulation de symbole. Dans ces conditions, il s'agit toujours de calculs syntaxiques à différents niveaux de granularité.

Clancey pense qu'il ne s'agit pas seulement d'une différence de représentation interne, mais que les personnes fondent leur raisonnement sur autre chose. Quoi ? est la question qui lui semble la plus importante à traiter.

Se souvenir des controverses

Arguments pour une mémoire dynamique plutôt qu'une structure de stockage

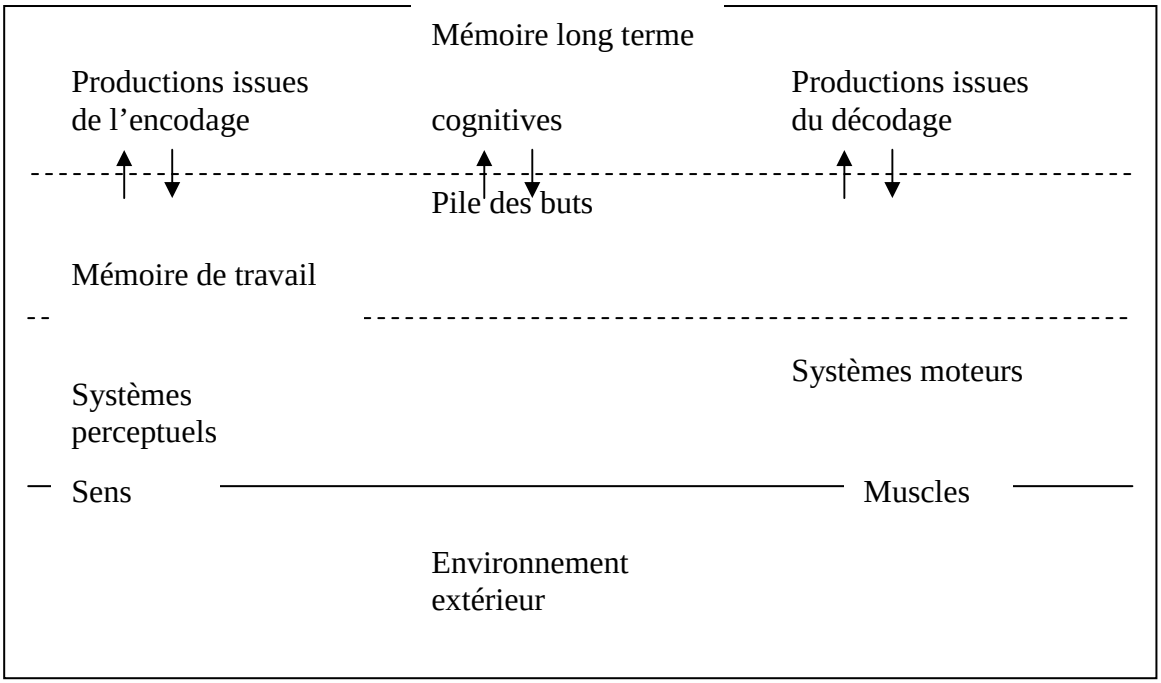
Dès 1932, on savait que la mémoire ne stockait pas des schémas indexés que la remémoration permettait de retrouver et donc de rendre à nouveau conscients. L'idée de construction dynamique de ce qui pouvait être formulé comme un schéma était défendue par Bartlett (inventeur de la notion de schéma).
Il est alors admis que la mémoire est « associative » et que les traces stockées sont activées pour reconstruire une image de ce qui a été mémorisé.

Le point de vue de ceux qui parlent de représentations

Pour tenir compte de la complexité du raisonnement symbolique, les tenant de la représentation des connaissances proposent une organisation en sources de connaissances coopérant autour d'un « réseau sémantique ». On y trouve les connaissances de domaines, les méta-connaissances, les stratégies de modélisation du diagnostic, les stratégies de résolution de problème, les stratégies d'explication, les stragégies de tutorat, l'histoire de l'utilisateur, un modèle différentiel de l'utilisateur, etc…). Tout cela serait localisé quelque part dans le cerveau sous une forme synthétisée par des réseaux neuronaux naturellement.

Le point de vue psychologique d'un système physique de symboles

Le modèle « système cognitif total » :



Qu'est-ce qui ne va pas avec le connexionnisme simple ?

Le connexionnisme est le nom d'un ensemble de méthodes computationnelles fondées sur la représentation de l'information dans un grand réseau de processus parallèles, vaguement inspiré (et informellement appelé) « réseaux neuronaux ». Grossièrement, un logiciel connexionniste simple construit une mémoire de paires d'entrées/sorties, de telle façon que ce qui est retenu n'est pas des unités indépendantes stockées, mais plutôt un enregistrement cumulatif des relations entre éléments de l'entrée et de la sortie correspondante. Le connexionnisme simple est ainsi un pas en direction de la vue contextualiste de la mémoire comme une gestalt ou une intégration de stimuli au travers du temps.
Le problème est que les entrées et les sorties sont considérées comme prédéfinies alors que les contextualistes prétendent que les associations apprises ne sont pas des couples entrées/sorties fournis au préalable mais des contrastes ou des distinctions signifiantes pour le sujet.
L'apprentissage d'un réseau neuronal revient à construire un « motif » dans le réseau capable de discriminer des entrées pour produire des sorties. En ce sens, il revient à stocker une

structure de classification et donc en quelque sorte à refaire un « matching » selon une structure particulière.
Un réseau neuronal simple est donc par nature associationniste et non pas contextualiste.

La question est de rendre les dispositifs sensori-moteurs et mnésiques dynamiques et indissociablement liés dans cette dynamique. En d'autres termes :

- Pourquoi un organisme est-il attentif à certains stimuli et pas d'autres ?
- Comment créons nous de nouvelles façons de voir le présent, de nouvelles sortes de généralisations ?
- Comment les formes d'activité prennent-elle sens dans un contexte particulier ?

(Questions posées par Rosenfield en 1988)

De telles questions rejoignent celles de l'acquis et de l'inné.

Cartes sensorimotrices versus encodages

Von Foerster et Bateson : Descriptions de l'information

L'idée que l'information est une substance physique est fondamentale dans la psychologie du traitement de l'information. L'information est quelque chose qui peut être stockée, codée, comparée et affichée. En psychologie, l'information est constituée des noms des choses ou des événements, de valeurs numériques ou d'abstractions qualitatives de paramètres et même d'explications causales d'observations. En fait, l'idée d'information est souvent regroupée avec celle de données, représentation, modèle ou connaissance. Les seules distinctions résident dans ce qui est entrée ou sortie dans une situation.

Von Foerster insiste essentiellement sur le fait que l'information ne saurait être « stockée » puisqu'elle n'a de sens que pour l'homme (il s'agit de psychologie).[1]
La critique essentielle vient du fait que l'on stocke une information dont il reste très mystérieux de savoir comment elle a été fabriquée !
Si la mémoire n'était que le stockage, il faudrait donc un « démon » pour retrouver à tout moment ce qui est utile pour gérer une situation courante. Il s'agit de l'objection bien connue dite de « l'homunculus » (petit homme) qui serait nécessaire quelque part chez l'homme pour interpréter ce qui est stocké (et chez ce petit homme, un autre ?...).

Bateson renforce cette idée en proposant de distinguer l'information et la description de l'information (ce qui était déjà largement dit dans le début de cet ouvrage à propos de la connaissance…).

En conclusion, il semble démontrer que l'information est quelque chose qui n'a pas grand-chose à voir (en terme de substance) avec la connaissance. L'information est intrinsèquement biologique et donc ne concerne pas que l'homme (à la différence de la connaissance qui ne serait dicible que par l'homme ? [remarque personnelle]).

[1] Je me demande pourquoi la théorie de l'information n'est pas ici citée ? Il semble que l'attaque aille surtout vers les psychologues qui appellent information ce que nous aurions tendance à appeler « connaissances » en IA ?

La carte sematosensorielle du singe

La carte cérébrale de la main d'un singe a été réalisée. Les différentes zones proches sur la main correspondent à différentes zones proches dans le cerveau. Il n'y a pas de description symbolique de la main du type (NEXT_TO THUMB FIRST_FINGER). Le système nerveux physiquement relie la main et la partie cérébrale associée. Il apparaît alors que c'est le corps du singe qui fait que cerveau et main sont associés et peuvent être observés ainsi.

Processus systémiques et dynamiques de la représentation

Le processus analytique, linguistique de la description ne peut pas être identifié au processus physique survenant dans le système nerveux.
Maturana insiste sur l'aspect linguistique de la modélisation :
« L'opération basique qu'un observateur réalise (bien que cette observation n'est pas exclusive aux observateurs) est l'opération de distinction ; c'est-à-dire, la mise en évidence d'une unité en exécutant une opération qui définit ses limites et la sépare de l'environnement » [Maturana, 1975]
« Nous parlons habituellement et proposons des explication pour des phénomènes perceptifs comme si nous, en tant qu'observateurs, et les animaux que nous observons existaient dans un monde d'objets, at comme si le phénomène de perception consistait à saisir les caractéristiques des objets du monde, parce qu'il existe un moyen qui autorise ou spécifie cette saisie » [Maturana, 1983]
Nos modèles et diagrammes descriptifs portent à voir la perception comme un processus de mise en correspondance de caractéristiques les unes avec les autres plutôt qu'à expliquer comment le processus de perception lui-même crée ces caractéristiques.

Les systèmes capables de se réorganiser en permanence (y compris dans leurs capacités de perception) sont appelés « autopoiëtiques » par Maturana : le réseau de composants et de relations qui constitue le système est continuellement régénéré par la mise en œuvre du système lui-même.

Compte-tenu des aspects dynamiques des processus mentaux de haut niveau de catégorisation et de raisonnement, il semble raisonnable d'imaginer que la représentation conceptuelle et procédurale possède au moins autant de plasticité que celle observée dans les cartes sensorimotrices.

Une distinction fondamentale est faite : « les interactions avec l'environnement ne sont pas médiées par des encodages qui iraient de l'extérieur à l'intérieur. L'environnement est directement perçu ».
Comprendre les idées de fermeture d'information et de perception directe est essentiel pour comprendre le type de mécanisme de coordination mémoire que Merzenich a mis en évidence.
Pour expliquer l'idée de perception directe, Winograd et Flores (1986) paraphrasent l'un des exemples de Maturana :
« Quand la lumière frappe la rétine, elle altère la structure du système nerveux en déclenchant des changements chimiques dans les neurones. Cette structure modifiée conduit à des formes d'activité différentes de celles qui auraient été générées sans ce changement, mais ce serait une simplification qui nous induit en erreur que de voir ce changement comme une perception de la lumière. Si nous injectons un produit irritant sur le nerf, il déclenchera un changement

dans les formes d’activité, mais que nous hésiterions à considérer comme la perception du produit irritant ».

Cette façon de voir les choses est totalement convaincante chez les hommes, mais ne s’applique pas aux ordinateurs ! Pour les ordinateurs, données et informations ont un sens que les théories décrivent !

Pour construire un artefact informatique, il n’est donc pas question d’évacuer la notion d’information, mais il est particulièrement utile de comprendre ce fonctionnement émergent dans la pensée humaine pour s’en inspirer dans la réalisation de « robots ».