

La Cognitive

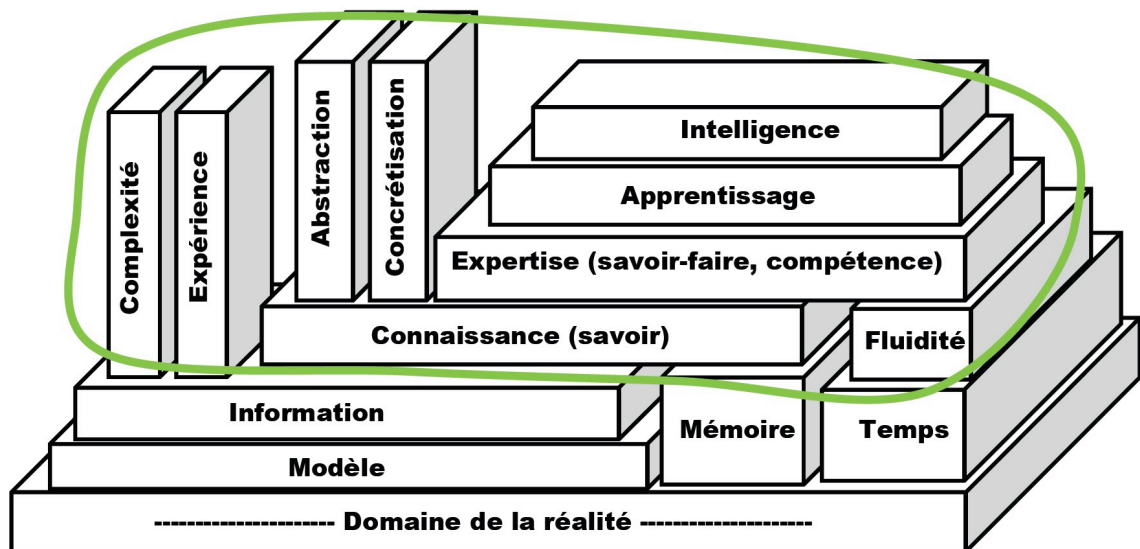
Remarques éventuelles à l'auteur bienvenues
(<mailto:Jean-Daniel@Dessimoz.org>)

Sites de référence pour mises à jour futures:
<http://cognitics.org>;
<http://cognitive.populus.ch>

Jean-Daniel Dessimoz

La Cognitique

Définitions et métrique pour les sciences
cognitives
et la cognition automatisée



Editions Roboptics

Première édition août 2008

No réservé : ISBN 978-2-9700629-0-5

Copyright 2008 - Editions Roboptics

CH-1400 Cheseaux-Noréaz, Suisse

Tous droits réservés

Reproduction, même partielle, sous quelque forme
ou sur quelque support que ce soit,

Interdite sans l'accord écrit de l'éditeur.

Imprimé en Suisse.

Préface

Le monde de la cognition est passionnant. Alors que les activités cognitives étaient largement réservées à l'humain par le passé, les progrès actuels en matière de technologies et de théories permettent de transférer beaucoup de ces activités humaines à des systèmes artificiels ; nombreux, bon marché, très performants.

Dans la démarche cognitique, deux mouvements complémentaires apparaissent, opposés et pourtant aussi fascinants l'un que l'autre . Dans un sens, l'humain sert de référence, définit les tâches à réaliser, les buts à atteindre, parfois pour le remplacer, en général pour l'assister ; et dans l'autre sens, la métrique développée dans la présente théorie « MCS », et la rigueur d'analyse exercée sur les systèmes artificiels pour les construire, les dépanner et les améliorer s'avèrent réciproquement utiles pour mieux connaître les propriétés cognitives humaines.

Cet ouvrage se destine à un large public et privilégie la discussion des fondamentaux plutôt que les spécificités d'aspects élémentaires ou applicatifs réservés à d'autres cercles, tels que conférences spécialisées, ou journaux scientifiques.

L'objectif principal est que le lecteur ordinaire (« l'honnête homme » de la Renaissance) puisse comprendre et exploiter dans sa vie de tous les jours les apports généraux de la présente théorie et des commentaires formulés. Un deuxième objectif concerne un public plus restreint, peu ou prou lié à la cognitique, et consiste à jalonner de nouvelles pistes pour améliorer la conception et le développement des robots, des systèmes informatiques et des systèmes automatisés en général.

J.-D. Dessimoz

1. Introduction

L'humanité a évolué en inventant une variété toujours croissante d'outils et de méthodes. Cela lui a permis de croître en nombre, de vivre plus longtemps, et d'explorer une portion toujours plus grande du monde.

Schématiquement, et pendant longtemps, deux sortes de progrès se sont faits de façon largement indépendante, touchant dans un cas le monde immatériel des idées, et dans l'autre cas le monde des objets physiques. Seuls les humains semblaient avoir la faculté d'établir un lien entre ces deux mondes, notamment par le langage oral, l'écriture et le dessin, voire la sculpture et l'architecture.

Durant le vingtième siècle, la révolution des communications à longue distance s'est accompagnée de la formalisation d'une première articulation entre monde physique et monde immatériel : l'information. Il était essentiel d'établir une correspondance précise entre l'idée à communiquer et l'objet physique servant de message, et cela a été fait.

Aujourd'hui une nouvelle étape s'ouvre devant nous, où des systèmes artificiels peuvent non seulement passer, souvent avec une grande facilité, du monde physique au monde des idées, et réciproquement ; mais en plus il est même possible pour de tels systèmes de traiter par eux-mêmes des idées ; de tirer des conclusions, d'induire des précédents, d'abstraire ou au contraire de concrétiser. Il s'agit ici du domaine de la cognitique, c'est

à dire de la prise en charge automatisée des activités cognitives traditionnellement typiquement, voire exclusivement, réservées à l'humain.

Pour progresser en cognitique, il est opportun d'y définir de façon rigoureuse les concepts essentiels, ainsi que d'y consacrer un système de mesure, une métrique appropriée. C'est ce qu'apporte la présente théorie, intitulée « MCS », initiales des termes « Model for Cognitive Sciences ». MCS se présente ici en plusieurs sections ; la compréhension devrait être facilitée par un survol initial des grandes étapes du développement de cette théorie ; puis les caractéristiques principales du modèle s'esquissent, et enfin, de façon progressive et structurée, chacun des concepts principaux est abordé précisément.

Mais avant de déboucher sur les vastes plaines de la cognitique, il est nécessaire de passer deux cols classiques, curieusement beaucoup plus difficiles à franchir qu'il me le semblait a priori. De nombreuses discussions avec des interlocuteurs variés m'ont convaincu que de bien passer ces cols préalables s'avère une condition nécessaire pour ensuite apprécier avec confort et sans réserve le paysage nouveau proposé. Je m'apprête donc à accompagner le lecteur qui le souhaite sur les terrains sensés connus, voire archiconnus, de l'information et du modèle.

2. Terrains connus ?

L'expérience montre que la plupart des gens ne connaissent pas bien les concepts « d'information » et de « modèle », pourtant établis et même classiques, depuis des décennies voire des millénaires. Il est donc utile ici de revisiter ces notions.

Ce sera ensuite en s'appuyant sur ces notions classiques, que les nouveaux concepts de la cognitique pourront se définir.

Notons que certaines notions mentionnées ci-dessous doivent se comprendre de façon intuitive au premier passage, et pourront éventuellement être revisitées par la suite, lorsque les définitions formelles proposées plus bas auront été vues. En général, il ne faut pas s'attendre à de franches incompatibilités, mais plutôt à une utile précision nouvellement gagnée pour bien des cas demeurés flous ou ambigus dans les approches traditionnelles.

2.1 L'information

Définitions:

L'information, c'est ce qui permet au système cognitif qui la reçoit (CS), de se créer et de mettre à jour la représentation qu'il se fait d'un certain domaine cognitif, son modèle ad hoc. De façon intuitive on pourrait dire que l'information forme l'opinion.

L'information est convoyée par des messages. Et la quantité d'information contenue dans un message a été

définie comme correspondant à une probabilité. L'essentiel est lié à l'attente qu'a le récepteur du message qui arrive.

Lorsque le message est parfaitement prévisible, la quantité d'information est nulle. Alors que si le message est très inattendu, c'est une grande quantité d'information que l'on perçoit comme arrivant.

La fonction fondamentale définissant la quantité d'information contenue dans un message, Q , remonte aux années 50 du vingtième siècle et est due à Claude Shannon.

Un message totalement prévisible n'apporte pas d'information. Au contraire, un message très inattendu surprend. Une faible probabilité correspond en fait à beaucoup d'information. La fonction mathématique qui rend compte du phénomène, c'est d'abord l'inverse de la probabilité d'occurrence du message¹ :

$$Q = f(1/p)$$

En plus de cela, il semble judicieux que si plusieurs messages arrivent, leurs quantités d'information respectives s'additionnent. Or si l'on en reste aux probabilités, l'opérateur adéquat est le produit : par exemple si deux signaux indépendants ont une probabilité individuelle d'occurrence de un sur trois (un tiers), c'est un sur neuf qui exprimera les chances pour que ces deux messages se produisent simultanément.

¹ La définition fondamentale est formulée, comme ici, pour le cas de messages discrets (énumérables). Mais en fait ceci n'est pas vraiment une limitation car des passerelles existent pour l'étendre à d'autres cas, tels notamment les signaux continus .

Pour pouvoir continuer d'utiliser la simple addition, il faut au préalable passer par les logarithmes² :

$$Q = \log_2 \left(\frac{1}{p} \right) \text{ [bit]} \quad (1)$$

L'unité dans le cas de l'équation 1, c'est le « Bit », contraction de « Binary digiT » rappelant que le logarithme utilisé dans cette équation a pour base le nombre 2 ^{cf. note³}.

On remarque au passage que la formule donne bien une quantité d'information nulle (zéro bit) pour le cas du message totalement prévisible (probabilité égale à un).

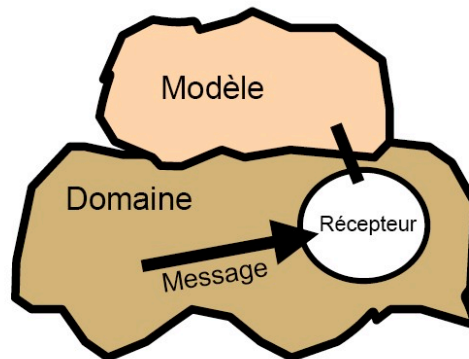


Fig.1 - L'information. L'information, convoyée par des messages, permet au récepteur de former son opinion, son modèle interne, sa représentation schématique (d'un certain domaine) du monde réel.

² Pour approcher de façon intuitive la quantité d'information contenue dans un message, il convient de compter combien de zéros significatifs se trouvent dans le nombre exprimant la probabilité d'occurrence du message. Voir aussi l'annexe C.

³ En théorie, les logarithmes en base 10 et en base « e » ont aussi été utilisés, conduisant dès lors aux unités « dit » et « nit ». Mais en pratique ces variantes ne se répandent pas.

Commentaires

La définition classique de l'information a fait ses preuves, et il est nullement question ici de la modifier. Néanmoins, voire au contraire, il convient de bien comprendre deux de ses propriétés essentielles: ses caractères périssable et subjectif.

Théorème 1 – L'information est instantanément périssable

Preuve: L'équation 1 définit la quantité d'information convoyée par un message sur la base de sa probabilité d'occurrence, telle qu'estimée *avant réception* par le récepteur. Or le message a précisément pour fonction, et donc typiquement pour effet, de modifier cette probabilité. A réception, ce qui n'était auparavant qu'une probabilité pour le récepteur se transforme désormais en certitude. La probabilité *a posteriori* du message reçu vaut « 1 ». L'équation 1 donne zéro bit dans ces nouvelles circonstances. Le message est désormais connu et ne contient plus d'information.

Discussion:

Considérons, à titre d'exemple, un domaine cognitif correspondant à un seul lancer, aléatoire, d'une pièce à « pile ou face ». Avant réception, deux messages sont possibles (pile, et face). Ils sont habituellement attendus avec, pour chacun, une probabilité de $\frac{1}{2}$. Après réception du message par contre les probabilités se modifient, la probabilité de l'un des messages (par exemple "pile") devenant respectivement 1, et pour l'autre 0. Pour ce lancer unique, il est inutile de répéter le message. L'équation 1 donne effectivement une quantité d'information valant 1 bit pour le message initial, puis 0 bit pour toutes ses répétitions éventuelles.

Il est important de bien comprendre cette particularité de l'information, car cela tranche par rapport à l'expérience que l'on a des autres unités de mesure, dans le monde physique : répéter une estimation de poids, de longueur, ou de temps donne en principe toujours le même résultat, en kilos, mètres ou secondes. Répéter un même message dans les mêmes circonstances n'apporte par contre plus aucune information additionnelle ; 0 bit pour les messages répétés.

Voici quelques exemples informels où le caractère variable au cours du temps de l'information joue un rôle particulièrement manifeste:

- En pratique, on ne lit en principe pas deux fois le même journal.
- Il est généralement malvenu de révéler la fin du film à ses amis alors qu'ils s'apprêtent justement à aller le voir.
- Il est difficile de planifier collectivement une surprise pour un tiers.
- Les opérations en bourse sont interdites aux insiders.

Théorème 2 – L'information est essentiellement subjective

Preuve: La quantité d'information d'un message est donnée par l'équation 1. On y observe qu'elle se base sur la probabilité d'occurrence telle que perçue par le récepteur. L'information a donc un caractère essentiellement subjectif.

Discussion :

Le caractère objectif des messages reçus est nullement garanti par l'équation de base. Intuitivement cependant

les gens ont tendance à croire à leur caractère objectif, notamment pour les deux raisons qui suivent :

- Dans les domaines techniques simples, tels que ceux où la théorie de l'information s'est d'abord développée, les modèles sont standardisés, définis de façon rigide, en conformité pour les émetteurs et les récepteurs, dans le cadre de systèmes cohérents.
- En général les membres d'un groupe ont vécu des expériences en grande partie similaire et tendent donc à développer une certaine uniformité de leurs modèles respectifs.

Et pourtant un même message peut avoir autant de probabilités différentes qu'il y a de récepteurs différents.

Par exemple dans le domaine du lancer de pièce, considérons deux récepteurs très différents. L'un associe typiquement, comme plus haut, a priori, la probabilité de $\frac{1}{2}$ à l'état pile. Par contre le deuxième est un plaisantin qui a fourni la pièce, pièce spéciale frappée des deux côtés de l'état « pile ». Pour ce dernier, la probabilité d'un message « pile » est déjà a priori de 1, 100%. Dès lors le message (« pile ») apporte 1 bit d'information au premier observateur, et 0 bit au second.

Il est utile de bien voir le rôle crucial du récepteur. Même si souvent en pratique l'information a un caractère très objectif, par exemple lorsqu'elle concerne les unités de mesure et les objets très courants, il existe tout un spectre de situations, allant à l'autre extrême jusqu'aux objets d'art dit moderne, voire aux tests de Rorschach.

2.2 La notion de modèle

Définitions:

Un modèle, c'est généralement une représentation simplifiée de la réalité ; typiquement élaborée pour atteindre un certain but. Parfois, la correspondance avec la réalité n'est pas recherchée et il s'agit alors, par extension, de la représentation d'autres mondes, virtuels.

La correspondance entre modèle et réalité sous-jacente, définit la notion de sens.

Dans la mesure où un modèle permet d'atteindre un certain but, il peut se qualifier de bon pour ce but.

Théorème 3 – L'information requiert la notion de modèle

Preuve: La définition même de l'information requiert les notions de message(s) et de probabilité(s) associées, estimées de façon quantifiée dans une représentation propre au récepteur (cf. équation 1). Cet ensemble d'éléments (messages, probabilités, représentation propre) constituent de facto un modèle (cf. figure 1).

Discussion : La tentation est constante pour l'homme de vouloir établir un pont direct entre monde cognitif et réalité. Mais c'est pratiquement impossible. Des philosophes aussi célèbres que Socrate, Kant ou Hegel sont particulièrement représentatifs d'efforts faits pour formaliser cette problématique et proposer des solutions. Socrate est contraint de constater que la portée de notre perception se limite aux reflets sur la caverne ; Kant postule l'existence de catégories déjà établies pour l'esprit humain en préalable à toute perception; et pour Hegel l'importance des représentations est telle que celles-ci font l'essentiel de notre monde, allant jusqu'à

éclipser la réalité, dont on peut à l'extrême douter de l'existence.

Dans notre approche également, on souhaiterait pouvoir appliquer la métrique de l'information directement à la réalité, mais cela est impossible. Nous reviendrons sur ce point par la suite, dans la discussion du théorème 4. Contentons-nous pour l'instant de vivre avec le théorème 3.

Théorème 4 – Sous réserve d'un but atteint de façon similaire, le modèle préféré, c'est le plus faux

Preuve: La qualité essentielle attendue d'un modèle, c'est qu'il permette d'atteindre un certain but. Dans ce sens, il est bon. Or si le but peut être atteint de façon similaire à l'aide d'un modèle plus simple, c'est à dire avec un modèle que l'on peut décrire avec moins d'information, le modèle est généralement considéré comme préférable. Pour se simplifier, le modèle doit ignorer certains aspects, devenir donc moins complet par rapport à la réalité. Et si un modèle est plus incomplet, il est plus faux. Ainsi, sous réserve d'un but atteint de façon similaire, le modèle préféré c'est le plus faux.

Discussion :

Il est classique de dire qu'il faut faire simple (cf. notamment la « loi de parcimonie », ou « la loi d'Occam ») et cela c'est sans doute une qualité pour un modèle. Mais dans cette formulation, on occulte combien on s'abstrait de la réalité. Einstein avec son aphorisme « il faut faire simple, mais pas plus » lève un peu le voile sur le risque à trop s'abstraire de la réalité. La difficulté croît si plusieurs objectifs sont considérés : un modèle judicieusement simple pour un but, s'avère

trop simple pour un autre but. Malheureusement dans tous les cas, certains éléments de la réalité s'éclipsent, et donc comme George Box le formule : « Tous les modèles sont faux ; certains sont utiles ». Le présent théorème, et de façon plus générale, la démarche visant une évaluation quantitative des grandeurs cognitives poussent encore cette affirmation plus loin . En substance oui, on peut affirmer qu'il peut y avoir de l'utile ou du bon à faire simple, mais il faut aussi voir qu'en termes de correspondance à la réalité, modéliser reste toujours extraordinairement lacunaire.

On dit parfois qu'une qualité spécifique de l'expert, c'est qu'il sait focaliser son attention de façon très sélective sur les dimensions critiques du domaine, sachant ainsi ignorer tous les autres aspects, qui rendent la situation confuse pour le débutant.

Une erreur commune c'est de penser qu'un modèle aurait des qualités de vérité et de capacité de représentation à la fois compacte et exhaustive de la réalité; qu'il pourrait retenir sans pertes la "quintessence" de la réalité qu'il représente.

Si l'on entreprend une estimation quantitative de la chose, on est forcé de constater que c'est impossible.

Quelles que soient les contraintes et les restrictions que l'on mette à un domaine de la réalité, il resterait une quantité d'information infinie nécessaire pour décrire de façon exhaustive ce domaine. Prenons l'exemple de la fameuse pipe de Magritte, (cf. figure 2). Même s'il ne s'agit de décrire qu'un certain millimètre cube du foyer de la pipe réelle, la quantité d'information nécessaire pour ce faire explose: quelles sont les fibres du bois? y a-t-il des agents conservateurs?, d'où provient le bois dont elle est faite, la pipe est-elle assurée? les employés

qui ont fabriqué le produit ont-ils été traités de façon éthique? etc.

En pratique, on recourt à des modèles: suivant l'objectif courant, c'est tel ou tel aspect très particulier de la réalité que l'on retiendra, de façon aussi exclusive que possible. Ainsi la pipe sera bien décrite par un numéro de commande pour le service comptable; par l'information de visibilité ou non, voire un code normalisé de couleur, pour un rendu pictural; par l'information de présence ou non d'un faux-goût pour le fumeur de pipe; etc.



Fig.2 – Modèle ou réalité. Il y a toujours une grande différence entre un objet réel (par exemple une pipe permettant de fumer réellement du tabac), et sa représentation picturale (ici le fameux tableau de Magritte, [4]) ; entre la réalité et n'importe lequel des modèles qu'on définit pour la décrire.

Pour l'intérêt pratique, il faut insister une fois encore sur la nécessité d'être toujours très clair par rapport au domaine auquel on s'intéresse, au modèle qu'on adopte, au but que l'on poursuit. Les jésuites se limitaient modestement au « hic et nunc », ici et maintenant ; les méthodes récentes « d'extreme management » en génie logiciel réclament de façon similaire qu'on se limite au plus serré des spécifications, sans aucun « bonus » vers

des recherches de solutions plus universelles, considérées comme illusoires par principe.

Par exemple s'il s'agit du poids d'une personne, est-ce habillé ? le matin au réveil ? sur Terre ou sur Mars ? Et le message donnant le numéro de téléphone, donne-t-il directement le numéro de téléphone (par ex. +123 123 123) ou le donne-t-il de façon indirecte (par ex. « c'est le numéro de John »).

2.3 Mémoire

Définition:

Une mémoire, c'est un support dont la caractéristique essentielle est le maintien de l'information au cours du temps.

Discussion:

La mémoire mérite peut-être un commentaire particulier. Elle est schématiquement représentée sur figure 3 – Pyramide cognitive. En tant que support physique à long terme, tel un menhir, la mémoire n'a pas de grand intérêt du point de vue cognitif. Simplement, on attend, de ce point de vue, une grande stabilité du support au cours du temps. Par définition, on s'attend à pouvoir relire par la suite exactement ce que l'on a écrit dans un premier temps; dès lors la quantité d'information générée est nulle.

Par ailleurs si l'on observe une mémoire microélectronique, on remarque le rôle extrêmement important des circuits d'adressage, ainsi que des circuits d'écriture et de lecture. De façon générale, dans la mesure où l'on intégrerait les processus d'adressage, d'écriture et de lecture à la notion de mémoire on déboucherait alors sur un ou des systèmes cognitifs

relativement complexes. Par exemple, il s'agirait non plus du menhir tout seul, mais aussi « d'Obélix », son sculpteur. Pour une bibliothèque, il s'agirait non seulement d'une collection de livres sur les étagères, mais aussi du bibliothécaire capable d'aller judicieusement classer l'information, puis la retrouver.

Dans la théorie MCS, la fonction de rémanence caractérise de façon essentielle la mémoire. Cette fonction n'est pas jugée devoir faire ici l'objet de grands développements. Par ailleurs les processus d'adressage, d'écriture et de lecture qui, traditionnellement, sont parfois associés à la mémoire, peuvent se traiter séparément, comme n'importe quel autre processus cognitif.

3. Grandes étapes de développement de la théorie MCS.

Pour que le lecteur comprenne plus facilement la théorie MCS, il vaut sans doute la peine de refaire, dans les grandes lignes, de façon d'abord quelque peu informelle, le cheminement chronologique du développement de la théorie MCS. Nous retiendrons trois grandes étapes : connaissance, expertise, et « autres développements ».

3.1 Vers le concept de connaissance

La notion d'information étant bien établie, scientifiquement et techniquement, la notion nouvelle à explorer était celle de connaissance.

Dans le langage courant, le mot connaissance, d'ailleurs généralement pris à la forme plurielle, s'utilise de façon relativement peu différenciée, recouvrant indifféremment des notions multiples, notamment l'« information », le « savoir » ou encore les « compétences »⁴. Or il vaut la peine de bien distinguer ces notions, ce que la présente théorie MCS contribue à faire. Les nuances sont importantes, et l'on verra par la suite sur le plan métrique que les unités proposées sont fort différentes également.

Si le terme de connaissance ne correspondait qu'à celui d'« information », on pourrait s'en dispenser. Et effectivement, dans la théorie MCS on n'utilise pas le

⁴ On s'en tient ici au domaine de la cognition. Ne parlons donc pas ici du sens du mot « connaissances » au sens de « personnes » que l'on connaît (acquaintances, en anglais).

mot connaissance dans cette acception. Mais il est néanmoins déjà intéressant de constater un voisinage manifestement très proche entre la notion de connaissance et celle d'« information », à laquelle on souhaite pouvoir s'ancrer.

La (les) connaissance(s) c'est le savoir ? Qu'est-ce que la connaissance dans ce sens ? Le déclic fondamental a été de constater que l'information s'avérait à nouveau la bonne voie, cette fois-ci non pas comme moyen de définition identitaire (l'information *n'est pas* la connaissance, dans la théorie MCS) mais comme moyen de *cerner* ce concept différent : l'observateur va estimer l'existence de connaissances (au sens de savoir), dans un système, sur la base de l'information qu'il en obtient. C'est notamment ainsi qu'à l'école l'évaluation des élèves se fait, typiquement, de façon traditionnelle ! L'idée forte est ici qu'un système doté de connaissance, qu'un système cognitif donc, a la faculté de *produire*, à la demande, *l'information* souhaitée. Intuitivement, ces notions soulignent la différence entre un musicien jouant en play-back (« information » pré-stockée) ou au contraire en live (« connaissance », capacité à générer l'information à la demande). La fonction cognitive est encore plus mise en évidence lors d'improvisation ou de jazz !

Pour constater une fois de plus, sous une forme un peu différente, comment le concept d'information permet de cerner celui de connaissance, considérons à quoi servent les connaissances ? A fournir l'information pertinente, lorsqu'on le souhaite.

Par exemple si on connaît une ville, on peut par exemple la décrire ; indiquer si une photo donnée la représente ; ou donner au chauffeur de taxi les indications pour s'y

déplacer. Dans tous ces cas, de façon typique, les connaissances permettent de générer l'information pertinente, à la demande. Brièvement dit, les connaissances permettent de « faire juste ».

3.2 Vers le concept d'expertise

L'étape précédente, comprenant non seulement la définition rigoureuse de la notion de connaissance, mais aussi la métrique associée (ces points sont développés plus bas) ont apporté tout un lot de résultats intéressants. Mais d'autres éléments ont à leur tour provoqué la surprise !

Notamment, il est parfois apparu une grande différence entre les quantités estimées par la théorie, et ce que l'intuition indiquait : telle tâche semblait exiger une énorme quantité de connaissances d'après les équations MCS et pourtant cela semblait facile à faire, intuitivement⁵.

Dans ces cas-là, il est apparu particulièrement pertinent de prendre aussi en compte la quantité de temps nécessaire au système cognitif pour élaborer l'information délivrée ; ou de façon similaire son inverse, la fluidité. Cette nouvelle grandeur de synthèse caractérise la capacité d'un système cognitif à faire à la fois « juste et vite » ; c'est l'expertise (synonymes : savoir-faire, compétences, etc.). Cette propriété est sans doute la plus importante pour un système cognitif et ses mérites ne s'arrêtent pas à éclairer les cas surprenants évoqués plus haut.

Il peut a priori sembler étonnant que le temps ait quelque chose à voir avec la cognition ; qu'on y soit très

⁵ Etant moi-même enseignant, je me suis parfois permis d'appeler cela ironiquement le paradoxe du professeur !

sensible. Et pourtant si l'on observe un match d'échecs, ou un examen scolaire, il est évident que la durée des plages de temps y jouent un rôle critique.

3.3 Autres développements

La définition de la notion de connaissance, avec son unité associée, le « lin » qui sera définie plus bas, puis celle du concept d'expertise, sont sans doute les éléments les plus révolutionnaires de la théorie MCS. Et sur cette lancée, héritant par ailleurs des notions classiques d'information et de modèle, de multiples entités du monde de la cognition se définissent avec facilité et limpidité : abstraction, concrétisation, expérience, apprentissage, intelligence, simplicité, etc. Une véritable ontologie pour les sciences cognitives peut se mettre en place. Dans des publications dont le contenu n'est pas repris ici pour des raisons de limite thématique, des extensions vont plus loin encore, débordant sur les domaines de l'économie, de l'automatique, de l'éthique et d'autres domaines encore, dans doute avec pertinence et contributions également à plus de clarté.

La propriété encore qui mérite ici une mention spéciale est celle de complexité. Dans les références classiques, la théorie qui me paraît la plus proche de la théorie MCS, c'est celle de Chaitin-Kolmogorov (CK). Il y a toutefois une différence importante concernant la façon de définir la complexité dans les deux contextes :

- Chez CK, la « longueur d'un programme minimal » caractérise la « complexité de la chaîne qu'il produit », et ces grandeurs ne peuvent pas se calculer.
- La présente théorie permet facilement de caractériser la « complexité d'un programme », et, de façon

similaire et indépendante, la « complexité de la chaîne qu'il produit » ; par ailleurs, elle définit la « réductibilité » de la chaîne comme étant le rapport de cette seconde complexité sur la première. De plus, la complexité se définit ici en prolongement direct et naturel du concept d'information, au point d'en hériter la même unité de mesure

Après ce survol des axes de développement de la théorie MCS, il est maintenant temps de considérer concrètement les caractéristiques principales du modèle proposé.

4. Caractéristiques générales du modèle proposé

Notons trois caractéristiques générales du modèle proposé : ses caractères comportemental, indépendant de l'échelle, et conceptuellement indépendant de la nature physique du support d'implémentation.



Fig. 4 – Modèle comportemental. Un système cognitif est principalement caractérisé, dans la théorie MCS, par l'information présente en entrée, I_{in} , et, respectivement, en sortie, I_{out} . Essentiellement, il est capable de générer l'information pertinente, I_{out} .

4.1 Modèle comportemental

Dans la théorie MCS, le modèle de référence est comportemental. Les caractéristiques cognitives nouvellement définies se basent essentiellement sur la prise en compte des messages et des flux d'informations entrant (I_{in}), et respectivement sortant (I_{out}), d'une boîte a priori « noire⁶ » : un système cognitif.

L'ensemble des messages possibles en entrée, avec leurs messages de sortie correspondants, constitue un domaine cognitif, D.

La caractéristique essentielle d'un système cognitif, c'est de pouvoir générer l'information pertinente en sortie (I_{out}), c'est-à-dire l'information correspondant au domaine D considéré.

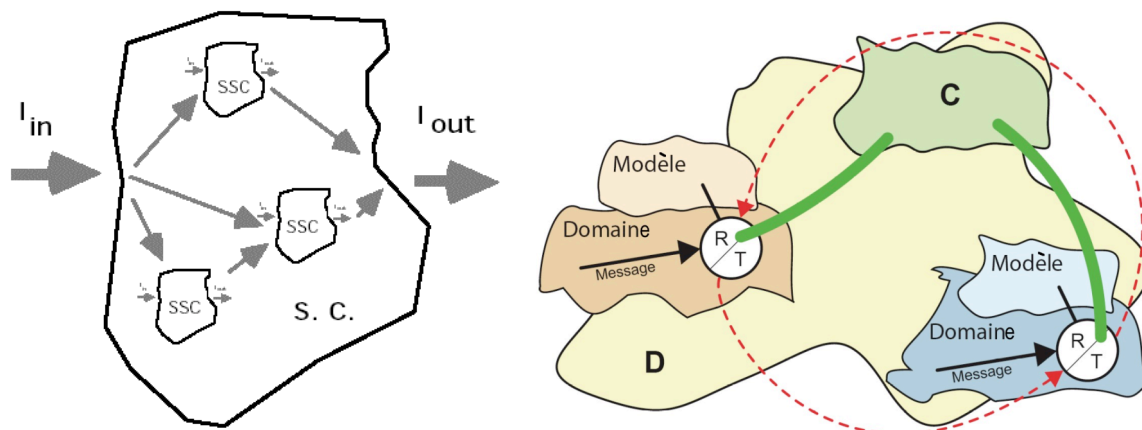


Fig. 5 – Granularité libre. On retrouve le schéma de la figure 4, à gauche dans une structure interne détaillée, et à droite pour un groupe intégrant plusieurs individus et ayant un modèle culturel commun.

Notons qu'il peut en principe arriver que la quantité d'information en entrée d'un système cognitif soit nulle.

⁶ Une « boîte noire », c'est un système dont les structures internes ne sont pas explicitées.

Par contre un système cognitif ne délivrant aucune information en sortie n'a aucun intérêt.

4.2 Modèle applicable à tout niveau de détail (granularité)

Le système cognitif considéré dans MCS est récursif : Il est bien entendu possible d'affiner l'analyse et de la continuer à l'intérieur de la boîte noire définie plus haut, pour en analyser la structure et d'éventuels sous-systèmes. Et réciproquement il est possible de prendre du recul pour éventuellement traiter à un niveau holistique, toujours de façon comportementale, c'est-à-dire sous l'angle des entrées-sorties globales, des systèmes plus grands, tels des groupes comportant un ensemble de systèmes cognitifs individuels.

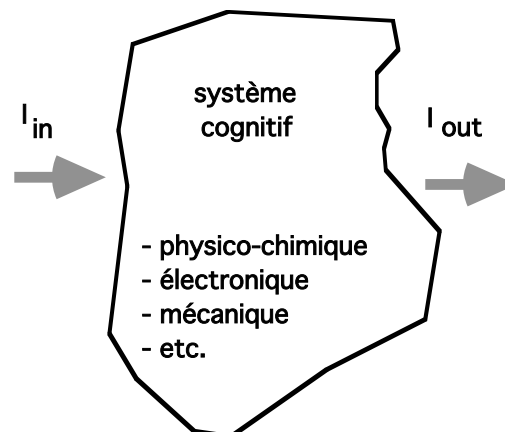


Fig. 6 – Nature immatérielle des systèmes cognitifs.

Comme dans le cas de l'information, un système cognitif, dans la théorie MCS, se définit indépendamment de la nature du support physique nécessaire à l'implémentation.

4.3 Modèle en principe indépendant de la nature physique du support d'implémentation

La portée du modèle de référence ne se limite pas pour MCS à une implémentation physico-chimique, comme

pour le cas de l'homme, ou microélectronique, tel le cas des circuits logiques ou les ordinateurs. Le modèle de référence proposé pour la cognition, comme d'ailleurs c'est le cas classiquement pour l'information, n'est en principe pas lié à une implémentation de nature physique particulière.

4.4 Exemples de systèmes cognitifs

La figure 7 présente plusieurs cas représentatifs de systèmes cognitifs. On y voit des systèmes à échelles très différentes : microscopique, puis macroscopique, pour le cas de l'homme ; et des implémentations de natures physiques différentes : physico-chimique ou microélectronique.

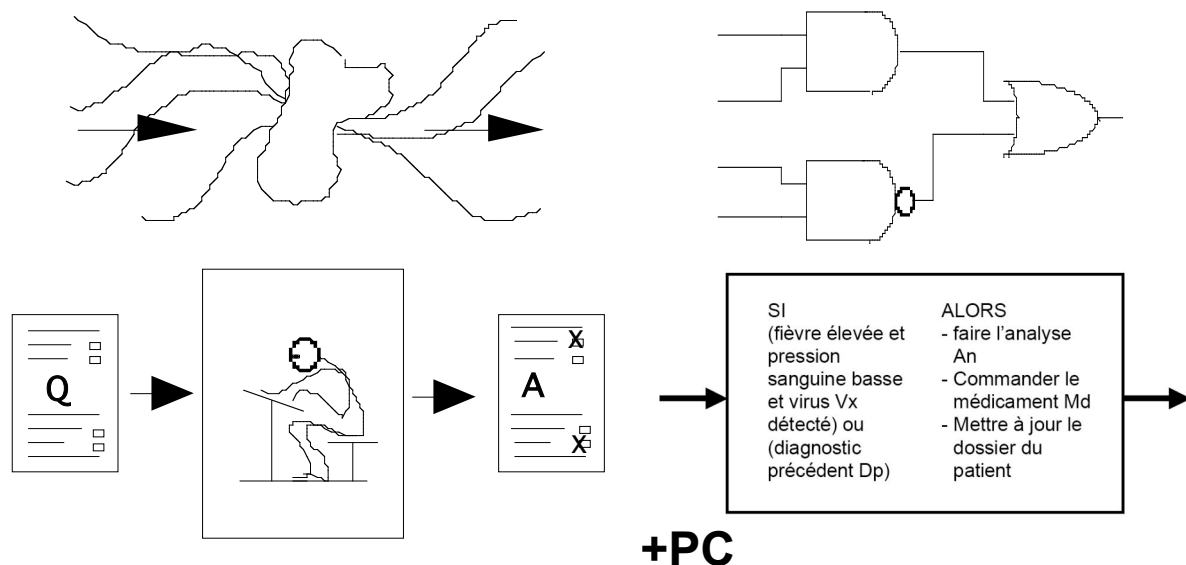


Fig. 7 – Exemples de systèmes cognitifs.

Représentation schématique d'un neurone (en haut à gauche), d'un circuit logique combinatoire (en haut à droite), d'une personne remplissant un QCM (en bas à gauche), d'un ordinateur exécutant une règle de production (en bas à droite). Tous ces cas sont modélisables de façon similaire, et leurs grandeurs

cognitives respectives peuvent s'estimer: complexité, quantités de connaissances, d'expertise, indice d'abstraction, etc.

Il est à noter qu'aujourd'hui les connaissances sont très souvent transmises sous forme d'information, par exemple comme méthodes, données ou code exécutable, et qu'elles ne prennent effet que lorsqu'elles sont implémentées par un agent cognitif, personne humaine ou ordinateur notamment. Ainsi, pour être complet dans le quatrième exemple donné, mettant en évidence une instruction de programme, une règle de production, il ne faut pas oublier de mentionner l'ordinateur et ainsi l'ensemble des ressources nécessaires à véritablement réaliser le système cognitif évoqué.

4.5 But de la théorie MCS

Enfin, nous l'avons vu⁷, un modèle ne peut avoir de qualité (caractère « bon ») que pour un but donné. Il convient donc de préciser ici le but poursuivi par la présente théorie, MCS : Il s'agit de définir de façon précise les grandeurs cognitives, de façon à pouvoir les évaluer quantitativement. Ceci doit sans doute s'avérer utile pour le contexte des sciences cognitives classiques, pour toutes les grandeurs et tous les processus intellectuels donc (par ex. connaissance, expertise, ou apprentissage), mais plus encore cela devient tout à fait nécessaire dans l'optique d'automatiser ces processus cognitifs ; pour la cognitique donc⁸.

⁷ cf. notamment le §2.2 qui présente la notion de modèle

⁸ Une première étape étant maintenant franchie, il apparaît qu'au-delà de la simple cognition, l'étude du/des but(s)

5. Concepts principaux de la théorie MCS

Les notions classiques de modèle et d'information étant familières, nous sommes bien équipés pour aborder la théorie présentée ici pour le domaine de la cognition, la théorie « MCS » (Model for Cognitive Sciences). Parmi les dimensions de base, seul le temps sera encore à prendre en compte par la suite.

Il faut (malheureusement?) accepter d'entrée que les caractères apparemment limitatifs des notions de modèle et d'information se reportent également sur tous les concepts cognitifs qui se basent sur eux : non-stationnarité (caractère périssable), subjectivité, incomplétude par rapport à la réalité.

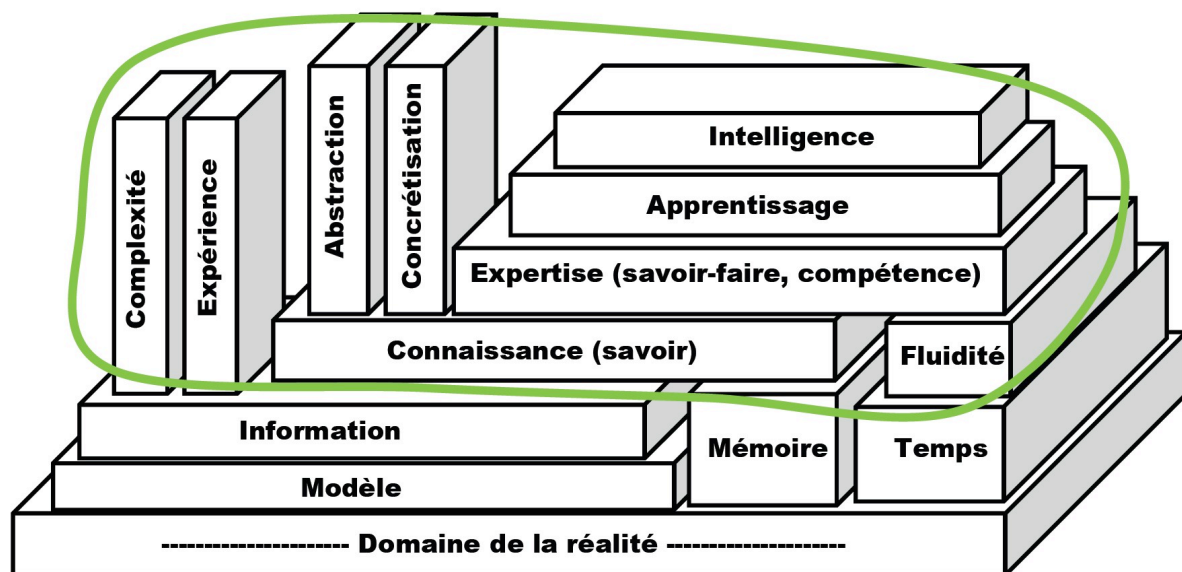


Fig. 3 – *Pyramide cognitive*. Les concepts cognitifs importants définis dans la théorie MCS sont

pertinents pour la cognition ne peut guère s'esquiver, et c'est le domaine de l'éthique qui logiquement devrait suivre. De façon similaire, la *mise en œuvre* des résultats de la cognition nécessite d'autres dimensions : énergie, éléments structurels et mécaniques, émotions, etc.

regroupés dans le périmètre vert de la figure, et se basent sur des grandeurs classiques, notamment l'information et le temps.

Dans la première section, le terme de cognitique est d'abord défini. Ensuite, successivement, une vingtaine de concepts centraux en cognition sont tour à tour introduits. Pour chacun d'eux, le titre est suivi d'un nombre indiquant l'ordre logique selon lequel les définitions s'imbriquent au mieux.

La figure 3 donne une impression d'ensemble des concepts développés.

5.1 La cognitique (0)

Définition :

La cognitique est le domaine des sciences et des techniques relatives à la cognition automatisée.

Discussion :

La cognitique apporte une réponse appropriée au défi posé par la complexité, qui, elle, ne cesse de croître⁹.

En cognitique, les opérateurs puissants comprennent notamment ceux qui permettent un accès rapide à l'information, tels que les classificateurs de listes notamment par ordres alphabétique, numérique ou chronologique. Un paradigme très puissant pour l'apprentissage automatique est celui des mémoires-caches. Dans ce dernier cas, un deuxième accès à une

⁹ On constate que de plus en plus d'information s'accumule, l'exploration scientifique par l'homme se poursuit, les réseaux se développent. Sous cet angle donc, nos représentations s'enrichissent. Et pourtant la réalité ne se révèle en fait encore que de façon infinitésimale.

donnée mémorisée peut être plus rapide que l'accès initial, ce qui se traduit, selon la théorie MCS développée plus bas, par une augmentation d'expertise, c'est-à-dire, par définition, par une capacité d'apprentissage. Un autre exemple est celui de la gestion de signets et de liens informatiques préférés.

Les avancées en cognitique sont rendues plus faciles par la mise à disposition, grâce à la théorie MCS, de définitions formelles et de métriques pour les propriétés cognitives, telles que complexité, connaissance, abstraction ou intelligence. Dans le contexte de la cognitique, il apparaît que les systèmes qui démontrent un haut niveau d'expertise sont en principe ce qui se fait de mieux. Les systèmes intelligents, c'est-à-dire les systèmes capables d'apprendre, sont en général comparativement moins intéressants. Le domaine de la cognitique recoupe en grande mesure celui de l'intelligence artificielle, tels qu'on l'entend généralement. Mais il est plus général, et héberge beaucoup d'autres concepts, notamment ceux de connaissance, d'expertise, d'apprentissage, d'abstraction ou de concrétisation par exemple.

5.2 Modèle (1)

Définition :

Comme introduit de façon générale, au §2.2, un modèle, c'est généralement une représentation simplifiée de la réalité élaborée pour atteindre un certain but.

Dans la théorie MCS, un domaine D de la réalité est en principe représenté comme D_m , un ensemble de N associations, A_i de type « message d'entrée - message de sortie »:

$$A_i : (I_{in}, I_{out})_i \quad [\text{bit}] \quad (2)$$

$$D_m : \{(I_{in}, I_{out})_1, (I_{in}, I_{out})_2, \dots, (I_{in}, I_{out})_n\} \quad (3)$$

Discussion :

La notion de domaine est introduite ici pour évoquer le fait qu'à un moment donné l'attention ne se porte en principe pas sur la réalité dans son tout, mais seulement sur une part très limitée de celle-ci, un certain domaine.

Il y a donc déjà une première contrainte à toujours garder à l'esprit : de quel domaine parle-t-on, ici et maintenant.

Une deuxième limite provient du fait qu'un modèle par nature, et comme déjà discuté plus haut, ne retient que peu d'aspects du domaine sur lequel il porte. En principe donc pour le même domaine D de la réalité, de multiples modèles, D_{mi} , sont envisageables¹⁰, et là également, il convient de bien suivre duquel il s'agit, dans le contexte courant.

Exemple

Considérons par exemple la tâche d'assurer qu'une pièce soit posée du côté pile.

Ce domaine peut se représenter par deux associations, correspondant 1. au cas d'une pièce initialement posée

¹⁰ Une alternative radicale à l'utilisation de modèle consiste, comme Parménide l'a proposé, à laisser l'observateur face au domaine D lui-même : « ce qui est, est ; ce qui n'est pas, n'est pas ». L'expérience, le vécu semblent pour l'humain un préalable difficilement contournable à toute modélisation ; il vaudrait pourtant la peine de tenter cela, ou tout au moins de contenir cette étape au minimum, ne serait-ce que comme préparation au passage à « l'automatisé ».

du côté pile, et 2. au cas d'une pièce initialement posée du côté face.

Dans le premier cas, rien n'est à faire, et dans le deuxième, il convient de retourner la pièce.

Une représentation possible de ce domaine cognitif est par exemple la suivante :

$$D_m : \left\{ (pile, rien - à - faire) , (face, retourner - la - pièce) \right\} \quad (4)$$

Théorème 5 – La réalité ne peut pas s'approcher directement

Preuve: Un domaine de la réalité, D, n'est pas directement gérable dans la théorie MCS. Une certaine représentation, D_m , est nécessaire, comme défini à l'équation 3.

Discussion : On hérite ici des restrictions déjà mises en évidence par les théorèmes 1 à 4, relatifs aux notions classiques d'information et de modèle.

En cohérence avec l'exigence de continuité, qui impose que la nouvelle théorie se base notamment sur l'acquis que constitue la théorie de l'information, il faut constater que nous sommes emprisonnés dans nos modèles. Ceux-ci ont néanmoins au moins le double mérite M1 : d'une certaine finitude, et, surtout, M2 : d'une certaine capacité à aider l'agent cognitif à atteindre ses buts.

Dans l'exemple ci-dessus, on ne retient qu'un bit d'information relatif à la pièce dont il est question : son état pile ou face. Dans la réalité, cette pièce a forcément une infinité de caractères additionnels qui seront éventuellement critiques dans certaines situations : nature de la matière dont elle est faite, épaisseur, type de monnaie, défauts de surface, identité du propriétaire, etc., etc. .

5.3 Information (2a)

Le concept d'information a été présenté de façon détaillée au §2.1 et illustré sur la figure 1. Rappelons brièvement ici les éléments principaux relatifs à l'information: l'information est convoyée par des messages ; un système cognitif (SC) se fait une représentation interne simplifiée (un modèle) de la réalité ; les messages sont attendus avec une certaine probabilité.

Définition :

L'information, c'est ce qui permet à un système cognitif de former ou mettre à jour son modèle interne. Et la *quantité* d'information reçue est une fonction de la *probabilité* des messages reçus.

Discussion :

Dans la définition, on s'est attaché au cas d'un système qui *reçoit* de l'information. Mais il peut arriver qu'un système en émette, et, en fait la situation se gère alors de façon analogue : Si une certaine probabilité caractérise l'émission d'un message, la quantité d'information correspondante est donnée par la même équation 1.

Dans la théorie MCS, une attention particulière est portée à l'information entrant dans un système cognitif, ainsi qu'à l'information sortant du système, c'est-à-dire générée par lui.

L'équation 1 donnait la quantité d'information contenue dans un message, Q , c'est-à-dire qu'on visait l'essentiel. Sauf cas d'école, plusieurs messages sont toujours possibles. Et pour chacun d'eux l'équation 1 est bien sûr applicable, donnant notamment pour le $i^{\text{ème}}$ message la quantité d'information Q_i . Il vaut la peine

d'ajouter ici qu'en pratique c'est ensuite souvent la quantité *moyenne* d'information contenue dans un message, Q_m , qui devient alors intéressante, car cela permet de se rendre facilement compte de l'effet global¹¹ :

$$Q_m = \sum_{i=1}^N p_i Q_i = \sum_{i=1}^N p_i \log_2 \left(\frac{1}{p_i} \right) \quad [\text{bit}] \quad (5)$$

C'est donc cette équation qu'on appliquera pour estimer les quantités d'information entrant et, respectivement, sortant d'un système cognitif ; n_{in} et, respectivement, n_{out} [bit].

5.4 Message (2b)

Définition :

Un message est un élément d'information. Essentiellement, la quantité d'information qu'il transporte est déterminée par sa probabilité d'occurrence.

Discussion :

Comme déjà discuté, la probabilité d'occurrence déterminante est celle que le récepteur du message estime. Et c'est tout ce qu'on retient du message pour en estimer la quantité d'information convoyée.

¹¹ Notons qu'il s'agit de la moyenne sur tous les messages reçus. On peut aussi simplement estimer cette grandeur sous la forme d'une moyenne sur les messages possibles, pondérés par leurs probabilités respectives.

Il apparaît une sorte de contradiction qui n'est pas sans intérêt lorsqu'on souhaite simplifier les choses: alors qu'individuellement un message très peu probable transporte par définition une grande quantité d'information, dans la moyenne par contre, sa contribution par rapport à des messages beaucoup moins improbables est négligeable : son poids est aussi faible que sa faible probabilité d'occurrence.

Aujourd'hui, beaucoup de gens ont l'expérience de systèmes informatiques et sont souvent très influencés par les représentations numérisées¹² qu'ils rencontrent. Il est vrai que la façon dont un message a été codé sous forme informatique peut s'avérer *a posteriori* un indice significatif de la quantité d'information qu'il contient. Il n'en reste pas moins que c'est toujours l'équation 1 qui est déterminante. En cas de conflit, deux situations peuvent schématiquement se produire. S1. Soit le code utilisé semble utiliser plus de bits que nécessaires, et donc contenir, par définition, de la redondance (chose pouvant s'avérer utile, par exemple dans le cas de communication en présence de perturbations). S2. Soit au contraire le code peut sembler utiliser moins de bits que nécessaire pour un tel message, mais il faut alors savoir que ceci n'est pas possible d'après la théorie (l'analyse doit alors sans doute se refaire plus minutieusement afin de localiser une erreur certaine).

En principe il existe des techniques simples et efficaces pour coder des messages avec le minimum d'information déterminé par l'équation 1 (notamment : code de Huffman). L'équation 1 définit aussi la limite ultime pour toute tentative de compression d'information sans erreur.

5.5 Complexité (3a)

Définition :

La complexité, c'est la propriété d'un objet qui requiert une grande quantité d'information pour être décrit de

¹² équivalent du mot anglais « digital »

façon exhaustive. L'unité de mesure est celle de l'information, c'est-à-dire le [bit]¹³.

Discussion :

De façon très directe, cette définition reporte sur la complexité les propriétés de l'information elle-même: non-stationnarité et subjectivité (cf. théorèmes 1 et 2).

Une équation supplémentaire utile pour quantifier l'information, c'est celle qui donne la quantité moyenne d'un message, applicable en cas d'équiprobabilité des messages. Cela permet une estimation rapide des quantités d'informations et donne dans tous les cas une borne supérieure. Voici cette équation, dans le cas de N messages différents possibles, tous étant équiprobables :

$$Q_m = \sum_{i=1}^N p_i \log_2 \left(\frac{1}{p_i} \right) = \sum_{i=1}^N \frac{1}{N} \log_2 \left(\frac{1}{\left(\frac{1}{N} \right)} \right) = \log_2(N) \quad [\text{bit}] \quad (6)$$

Le moyen le plus sûr de limiter la complexité consiste à se contraindre autant que possible à de petits contextes ; à un nombre aussi limité que possible de messages différents.

Des méthodes classiques pour approcher cet idéal incluent les représentations hiérarchisées¹⁴, ou encore

¹³ Des définitions pour la complexité ont été proposées par de nombreux auteurs. La définition la plus proche de celle de la présente théorie MCS, c'est peut-être celle de Chaitin-Kolmogorov. En substance, la complexité d'une chaîne de caractères y est estimée par la longueur du programme permettant de générer la chaîne de caractères. Voir aussi sous « Réductibilité ».

les mécanismes de tri. L'effort est alors reporté de la gestion directe d'un contexte complexe vers l'optimisation des techniques d'adressage et de navigation d'un contexte à l'autre.

Il est évident aujourd'hui que l'adressage par contenu et les hyperliens que l'informatique et les réseaux assurent, permettent à l'individu de considérer valablement un très grand nombre de contextes différents, et chacun d'eux reste de complexité gérable. Globalement, la complexité de nos mondes sur ordinateur se compte aujourd'hui en téraoctets (1 octet=8 bit) à l'échelle du poste individuel, et probablement de l'ordre de 10^{10} fois plus à l'échelle d'internet.

Il apparaît, en particulier, que la présente définition de la complexité résout un paradoxe classique en intelligence artificielle, le paradoxe des experts : plus on en sait, plus on peut facilement en apprendre encore¹⁵. Pour l'expert, les messages sont largement prévisibles ; par définition, ils contiennent moins d'information pour lui et sont donc ainsi, par définition également, moins complexes. La suite n'est plus paradoxale: il est logique qu'on apprenne plus facilement ce qui est moins complexe.

¹⁴ Dans un rapport par exemple, on aura notamment 1. dans le contexte du résumé, une présentation de l'ensemble du rapport de façon relativement abstraite, 2. par ailleurs à l'intérieur du rapport, un contexte particulier, un point spécifiquement dévolu aux détails concernant un aspect singulier, ou encore 3. un titre très résumé permettant de transférer sans excès d'information le contenu du rapport dans le contexte plus général d'une bibliothèque ou d'une bibliographie.

¹⁵ C'est le contraire des phénomènes dont on a l'habitude dans le monde physique. Par exemple l'augmentation de la pression dans un pneu : plus on pompe, plus il devient difficile de pomper.

Exemple:

La définition de la complexité, dans la théorie MCS, tient en 2 lignes et demi. Ceci représente environ 30 mots, soit environ 300 bit¹⁶. Sur cette base, la complexité de la définition peut s'estimer à environ 300 [bit].

5.6 Abstraction (3b)

Définition :

L'abstraction, c'est la propriété d'un système qui délivre en sortie une quantité d'information inférieure à celle qu'il reçoit en entrée. Quantitativement, dans la théorie MCS, un indice d'abstraction, i_{Abs} , est défini comme étant le rapport de la quantité d'information en entrée, n_{in} , sur la quantité d'information en sortie, n_{out} :

$$i_{Abs} = \frac{n_{in}}{n_{out}} \quad (7)$$

Discussion :

Alors que les quantités d'informations en entrées et en sorties se mesurent en [bit], l'indice d'abstraction est lui sans unité.

Typiquement, l'indice d'abstraction est supérieur ou égal à 1. Si au contraire la quantité d'information en sortie est plus grande qu'en entrée, il ne s'agit alors pas, à proprement parler, d'abstraction, mais plutôt de concrétisation (voir point suivant).

De façon générale, l'abstraction est typique des processus perceptifs, de la reconnaissance de formes et

¹⁶ Admettons qu'en français nous avons 1000 mots (ce qui est sans doute sous-évalué) et qu'ils sont équiprobables (ce qui constitue par contre manifestement une surestimation) ; dès lors nous avons environ 10 bit par mot, sur la base l'Equ. 6.

de façon générale tend à caractériser la compréhension, et la démarche scientifique. A l'évidence, cela caractérise la modélisation.

5.7 Concrétisation (3c)

Définition :

La concrétisation, c'est la propriété d'un système qui délivre en sortie une quantité d'information supérieure à celle qu'il reçoit en entrée. Quantitativement, dans la théorie MCS, un indice de concrétisation, i_{Concr} , est défini comme étant le rapport de la quantité d'information en sortie, n_{out} , sur la quantité d'information incidente, n_{in} :

$$i_{Concr} = \frac{n_{out}}{n_{in}} \quad (8)$$

Discussion :

Alors que les quantités d'informations en entrées et en sorties se mesurent en [bit], l'indice de concrétisation est, lui, sans unité.

Typiquement, l'indice de concrétisation est supérieur ou égal à 1. Si au contraire la quantité d'information en sortie est plus petite qu'en entrée, il ne s'agit alors pas, à proprement parler, de concrétisation, mais plutôt d'abstraction (voir point précédent).

Les processus de concrétisation sont typiques en synthèse, en réalisation. Ils sont créatifs et, à l'évidence, générateurs d'information. La génération d'information, et plus particulièrement la concrétisation, sont synonymes d'art (cf. artisans et artistes).

5.8 Connaissance (4a)

Définition :

La connaissance, c'est une propriété essentielle des systèmes cognitifs. Elle leur permet de délivrer en sortie l'information pertinente ; ou bien de façon proactive, ou alors en réaction aux messages incidents.

Quantitativement, dans la théorie MCS, la quantité de connaissance, K , d'un système cognitif, est définie comme étant une fonction des quantités d'information incidente, n_{in} , et délivrée, n_{out} :

$$K = \log_2(n_{out} \cdot 2^{n_{in}} + 1) \quad (9)$$

L'unité, c'est le [lin] (Logarithm of Information).

Discussion :

La connaissance, c'est ce qui permet aux systèmes cognitifs « de faire juste ».

La connaissance, c'est la propriété la plus évidemment liée aux systèmes cognitifs ; voire à la cognition en général, comme la racine de ces mots l'indique.

Considérons dans un premier temps l'information telle que traditionnellement générée par l'humain dans le passé.

A cette époque, l'utilisateur qui avait besoin d'une information devait soit s'adresser à un autre humain, soit reprendre à partir d'une forme ou l'autre de mémoire, le résultat passé de processus cognitifs propres à l'humain également.

Voici deux exemples : 1. le « livret » de multiplication appris par cœur (de 1x1, jusqu'à 12x12), ou de façon plus complexe ; 2. la « table des logarithmes et des fonctions trigonométriques ».

La révolution, due aux systèmes cognitifs artificiels, c'est que tout à coup l'information souhaitée par les humains a pu s'élaborer à la demande, sans avoir été mémorisée, sans tabulation préalable. Cet aspect est bien rendu par le nom d'« agent » cognitif parfois aussi utilisé pour désigner un système cognitif.

Reprenons nos deux exemples : les premières calculettes datent des années 70, et depuis, elles sont capables par exemple de rapidement indiquer le résultat d'un calcul tel que 13×13 ou $\log(5)$ sans recourir à une mémoire contenant de façon explicite les réponses.

Considérons, comme troisième exemple, le cas plus actuel des traductions automatiques sur internet. Là également, pour la traduction de phrases, on ne recourt pas à une mémoire contenant de façon explicite toutes les réponses possibles, tabulées au préalable.

Dans une grande mesure, les systèmes cognitifs dispensent donc de l'élaboration préalable d'une grande quantité d'information.

Par contre alors que l'information mémorisée est statique, disponible virtuellement sans délai, la cognition implique par nature un processus, une « action » et donc forcément un certain temps.

Ce qu'on a d'essentiel c'est, on l'a dit, d'une part la dispense de devoir pré-élaborer les informations dont on aura éventuellement besoin par la suite; mais d'autre part c'est aussi la nécessité nouvelle de mettre en œuvre un processus.

Cette interchangeabilité des rôles, entre information pré-établie, et processus capable de la générer à la demande, inspire, dans la théorie MCS, la façon d'évaluer les quantités de connaissances.

L'idée fondamentale est d'évaluer la quantité de connaissance d'un système cognitif en fonction de la quantité d'information préétablie que cela permet de remplacer.

Au cœur de l'équation 9, le terme $M = n_{out} \cdot 2^{n_{in}}$ vise ce but. On peut imaginer une mémoire, une table dont la « largeur » serait de n_{out} bit, et le nombre de lignes correspondrait à tous les messages possibles en entrée ($2^{n_{in}}$ lignes, dans l'hypothèse d'un code optimal, ce qui serait donc un minimum)¹⁷.

$$M = n_{out} \cdot 2^{n_{in}} \quad (10)$$

La présence de la quantité d'information en exposant du 2 doit nous alerter. La croissance d'une telle fonction est très rapide. Avec 10 à l'exposant, la base 2 nous donne environ mille; avec 30, un milliard.

Effectivement il apparaît qu'en général dans les processus cognitifs humains relevant de la perception, n_{in} se compte par milliers et bien plus. En conséquence la taille de cette table virtuelle, M , est tellement énorme qu'il serait tout à fait impossible de la réaliser physiquement. Afin d'alléger quelque peu les notations, c'est donc la fonction logarithmique¹⁸ qui est ajoutée à cette estimation¹⁹, c'est-à-dire dans l'équation 9.

¹⁷ Il n'y a pas de différence fondamentale entre la taille de la table virtuelle M , la complexité du domaine cognitif pour lequel un système cognitif est fait, D_m , et la quantité d'information nécessaire à décrire toutes les associations possibles, $A_1, A_2, \dots A_N$, dans ce contexte.

¹⁸ Une démarche similaire s'est faite dans le domaine des puissances sonores par exemple, ce qui a conduit au Bell.

¹⁹ Dans le cas d'école, un peu absurde, où l'on n'aurait aucune information stockée ($n_{out}=0$ [bit]), il serait souhaitable que l'équation donne 0 [lin] de connaissances également ; pour cette raison le terme

Il faut encore rappeler que l'objectif est simplement ici d'élaborer un standard pour quantifier les connaissances des systèmes cognitifs, et non pas d'éventuellement simuler leur fonctionnement à l'aide d'une table de réponses préétablies (même si ceci est envisageable pour des systèmes très simples, dans le cas général cela ne peut pas être la solution).

Exemple :

Dans le cas du lancer de pièce à pile ou face déjà mentionné plus haut, D_m , nous avons 2 messages équiprobables possibles en entrée, « pile » ou « face », donc 1 bit d'information.

Pour les actions correspondantes en sortie, « rien – à – faire » et « retourner – la – pièce », nous avons également deux messages équiprobables, et donc 1 bit d'information également.

Dès lors la quantité de connaissance d'un système cognitif capable d'assumer cette tâche est la suivante :

$$K = \log_2(n_{out} \cdot 2^{n_{in}} + 1) = \log_2(3) \approx 1.6 \text{ [lin]} \quad (11)$$

En principe tout est dit : nous avons ici 1.6 [lin] de connaissance. Allons néanmoins dans le détail afin de sentir un peu, intuitivement, l'analogie citée plus haut avec une table virtuelle qui contiendrait l'ensemble des solutions. Pour tabuler ce domaine, il nous faudrait $2^{n_{in}} = 2^1 = 2$ lignes ; la première ligne pourrait correspondre au cas « pile », et la seconde au cas « face ». Nous avons déjà vu que les actions correspondantes en sortie, « rien – à – faire » et « retourner – la – pièce », contiennent également 1 bit

« +1 » est ajouté dans l'Equ. 9 ; mais en général cette quantité est négligeable et ne change rien de significatif au résultat.

d'information. On pourrait coder respectivement par « 0 » et « 1 » ces deux actions et il apparaît bien que dans ce cas que la taille de cette table serait la suivante : $M = n_{out} \cdot 2^{n_{in}} = 2$ [bit], valeur que nous avons déjà dans l'équation 11.

Théorème 6 – La quantité de connaissance d'un système cognitif est beaucoup plus liée à la quantité d'information qu'il peut percevoir qu'à la quantité d'information qu'il peut synthétiser.

Preuve: Observons le cœur de l'équation 9 permettant l'estimation de la connaissance, K, soit aussi le terme M selon l'équation 10. La quantité moyenne d'information synthétisée, n_{out} , y figure comme facteur ; alors que la quantité d'information incidente, n_{in} , y participe comme exposant.

A valeur égale, n , n_{out} apporte typiquement une contribution négligeable à K par rapport à celle de n_{in} .

Si n est nul, la situation est discutable, car c'est le terme constant 1 qui s'impose dans l'équation. Mais un système cognitif ne délivrant pas d'information en sortie n'a guère de sens. Et pour toute valeur de n égale ou supérieure à 1, $2^n > n$, voire $2^n \gg n$. A valeur égale en entrée et en sortie, n , l'effet de n_{in} , en exposant, est donc bien plus grand que celui de n_{out} , qui est un simple facteur dans l'équation.

Discussion :

L'équation de base établissant la mesure de l'information, l'équation 1, comprend un logarithme pour passer de la probabilité (et donc du nombre de messages possibles en cas d'équiprobabilité) vers la

quantité d'information²⁰. Dès lors la fonction inverse, c'est-à-dire la fonction permettant d'aller dans le sens quantité d'information vers probabilité (et nombre de messages possibles), comprend une exponentielle. C'est cet effet qui affecte principalement la contribution de n_{in} .

Par contre la même fonction logarithmique entraîne que si l'on considère plusieurs messages simultanément, leurs quantités d'information correspondantes s'additionnent simplement. C'est cet effet qui affecte principalement la contribution de n_{out} .

L'analogie de la table virtuelle, comme développée dans l'exemple précédent, permet bien de constater la dissymétrie des effets entre entrées et sorties.

Voici encore deux exemples de nature plus intuitive :

1. Un lecteur de CD ordinaire fait très simplement le passage du numéro de morceau ($n_{in} \cong 4$ [bit]) à la génération du son musical (10 minutes de musique, soit environ $n_{out} \cong 10 \cdot 60 \cdot 44'000 \cdot 20 = 528 \text{ Mbit}$), alors qu'un système capable de reconnaître de tels morceaux de musique ($n_{in} \cong 528 [\text{Mbit}]$; $n_{out} \cong 4$ [bit]) n'est guère envisageable²¹.
2. Il est beaucoup plus facile aujourd'hui de faire de la synthèse vocale (transformation dans le sens du texte vers la voix) que de la reconnaissance vocale

²⁰ L'équation 6 rend encore plus évidente la relation logarithmique qui lie quantité d'information et nombre de messages.

²¹ Notons que ceci est très différent du type de reconnaissance d'air musicaux qui commence à se faire actuellement sur internet, dans le style « karaoke ». Une rapide évaluation quantitative le démontre. Et même dans le domaine très largement simplifié, nous sommes encore loin de la symétrie.

(transformation dans le sens de la voix vers le texte). Et il s'avère bien que la quantité d'information correspondant à un message sous forme vocale est nettement plus grande que le même message sous la forme textuelle habituelle.

Le terme « connaissance » défini ici peut très bien prendre comme synonyme, dans le langage courant, le terme « savoir ».

Théorème 7 – Les systèmes cognitifs courants ne peuvent pas être réalisés comme de simple lecteurs de réponses pré-calculées

Preuve: Il est banal aujourd'hui d'additionner trois nombres en quadruple précision, tant par ordinateur que directement à la main. Or ceci signifie que l'on sait additionner trois nombres à 34 chiffres caractéristiques chacun. Une table contenant toutes les possibilités de cette tâche devrait comporter $10^{3 \times 34}$ lignes, chacune comportant un nombre à 35 chiffres. Or ce nombre de lignes est plus d'un milliard de milliards de fois le nombre de particules (électrons, protons, neutrons) se trouvant dans tout l'univers connu à ce jour (10^{80}). C'est dire qu'il est tout à fait exclus aujourd'hui de mémoriser a priori toutes les réponses possibles dans ce domaine.

Discussion : Dans l'exemple ci-dessus, la quantité de connaissance, K, est d'environ 350 lin. Et pourtant c'est une quantité de connaissance plutôt petite dans le monde des systèmes cognitifs.

Il est par exemple courant que des systèmes automatisés tentent de reconnaître des mots parlés, ou des images.

1. Une seconde de parole en qualité « CD » correspond à environ $44'000 \times 8$ bit . Il faudrait ici une table à $10^{100'000}$

entrées pour tabuler tous les sons possibles durant cette seconde ; il s'agit d'une quantité de connaissance égale à environ 350'000 lin.

2. Imaginons, comme on l'a entendu dans la presse, qu'un contrôle automatisé, visuel, veuille se faire en frontière. Une image de qualité « télévision » traditionnelle comporte environ 600 lignes et l'équivalent d'environ 800 colonnes. Même si on néglige la couleur et qu'on ne retient que les niveaux saturés noir et blanc pour chaque élément d'image, cela donne environ déjà 500'000 bit d'information. Il faudrait ici une table à $10^{145'000}$ entrées pour stocker a priori la réponse appropriée pour chacune d'entre elles. Dans de tels contextes, et s'il s'agit de quantifier la connaissance, la quantité d'information en sortie peut généralement se négliger (cf. théorème 6). Si un seul bit d'information est stocké par ligne (par exemple : accepté ou refusé), donc éventuellement délivré en sortie, la quantité de connaissance, K, est la suivante : $K=500'000$ lin ; si par ailleurs c'est l'équivalent de tout un dictionnaire qui est stocké sur chaque ligne, soit 1 Mbit par ligne (ce pourrait par exemple être un curriculum extrêmement détaillé de la personne éventuellement reconnue, ou dans les autres cas une copie largement commentée de l'image), nous avons alors $K'=500'020$ lin.

Dans le monde de la cognition, les nombres peuvent décidément être très grands. Ces deux exemples apparemment banals mobilisent le gogol à la puissance 1000 ; et plus.

Théorème 8 – La quantité de connaissance d'un générateur parfaitement aléatoire est infinie

Preuve: Au minimum, un générateur d'information parfaitement aléatoire n'a pas d'entrée ; $n_{in}=0$ [bit].

Considérons comme minimal un message individuel, en sortie, comportant un seul bit, par exemple « 0 » ou « 1 ». En pratique on ne se satisfait pas d'une seule réponse. On attend d'un tel générateur, parfait, qu'il fournisse de multiples messages de sortie, de nombreux « 0 » et « 1 » successifs ; sans aucune régularité ; à l'infini. Dès lors $n_{out} = \infty$ [bit].

La quantité de connaissance d'un tel système cognitif est la suivante :

$$K = \log_2(n_{out} \cdot 2^{n_{in}} + 1) = \log_2(\infty \cdot 2^0 + 1) \approx \log_2(\infty) = \infty \text{ [lin]} \quad (12)$$

Discussion : Voilà un surprise. On a traditionnellement l'intuition qu'une faible quantité de connaissance est nécessaire pour générer un signal aléatoire.

La mémoire équivalente contenant l'information du domaine doit au moins être aussi grande que l'ensemble des messages de sorties. Si la séquence attendue n'est pas globalement de longueur limitée, la mémoire doit elle aussi, être de taille infinie.

En pratique, il est courant, en électronique ou sur ordinateur, que l'on ait des générateurs de nombres non pas aléatoires, mais *pseudo*-aléatoires. Typiquement les séquences générées sont *cycliques* et la quantité de connaissance correspondante y est de l'ordre de la dizaine de [lin] . On est évidemment très loin d'un vrai générateur aléatoire et ce n'est pas sur cette base que la loi autoriserait des jeux de loterie et de tombola !

Une autre approche courante pour un système cognitif consiste à « transférer » en sortie une information reprise d'une source extérieure, elle-même jugée aléatoire ; par exemple le centième de seconde de l'heure instantanée, selon l'horloge de bord : si l'on souhaite que son ordinateur dise de façon aléatoire et

équiprobable « Bonjour » ou « Hello » lorsqu'on l'allume, on pourrait lui demander d'aller lire l'heure courante à l'enclenchement (cette fois-ci ce serait par exemple 9h 10min 12s et 20 centièmes) et si le centième est pair (ici pour « 20 », c'est le cas) alors il dit « Bonjour », sinon il dit « Hello ». En principe, la méthode fonctionne ; même probablement de façon satisfaisante pour cette application. Mais sans doute aussi que l'analyse précise d'un tel cas est difficile, et que ceci ne constitue pas en soi une garantie de qualité pour le caractère aléatoire du résultat.

5.9 Expérience(4b)

Définition :

L'expérience, R , c'est la propriété d'un système cognitif qui a été exposé à un certain domaine cognitif. Quantitativement, l'expérience est traditionnellement évaluée en termes de durée, T , c'est-à-dire en fonction du temps. L'unité est alors la seconde, [s].

$$R_t = T \quad [s] \quad (13)$$

Une autre vue, plus fondamentale dans la théorie MCS, consiste à mesurer l'expérience, R , sur la base des quantités d'information des associations « message d'entrée - message de sortie », A_i , auxquelles le système cognitif a été exposé :

$$R_i = \sum_{i=1}^N (n_{in i} + n_{out i}) \quad [bit] \quad (13b)$$

où $n_{in i}$ et $n_{out i}$ correspondent aux quantités d'information du message d'entrée et du message de sortie de la $i^{ième}$ association.

Discussion :

Bien que l'expérience se compte traditionnellement en unités de temps, il est clair que la nature de l'environnement n'est pas sans importance.

Les deux définitions se rejoignent si l'on qualifie un environnement donné par le débit moyen d'associations concernant le domaine cognitif concerné qu'on y rencontre.

Exemple:

Un système cognitif a pu assister à une centaine de lancers de pièces, suivis du retournement éventuel sur le côté pile, soit une centaine d'exécutions de la tâche décrite plus haut, et correspondant au domaine D_m . Dès lors nous avons pour ce système la quantité d'expérience suivante :

$$R_i = \sum_{i=1}^N (n_{ini} + n_{out i}) = \sum_{i=1}^{100} (1 + 1) = 200 \quad [\text{bit}] \quad (14)$$

Ceci pourrait aussi se mesurer comme une expérience d'un jour, en adoptant la définition traditionnelle. Mais alors, il conviendrait de préciser que dans ce contexte on peut observer en moyenne 100 associations A_i relevant du domaine D_m par jour.

5.10 Fluidité (4c)

Définition :

La fluidité, c'est la propriété d'un système cognitif qui délivre rapidement ses messages de sortie ; c'est en particulier aussi l'agilité du système à réagir aux messages d'entrée.

Quantitativement, dans la théorie MCS, la fluidité, f , s'estime comme l'inverse du temps nécessaire à

l'élaboration du message de sortie. L'unité est alors l'inverse de la seconde : $[1/s]$.

Discussion :

La fluidité peut se voir comme la vitesse de traitement du système cognitif.

Exemples:

1. Un système cognitif peut donner sa réponse 0.5 seconde après avoir reçu chaque message d'entrée. Sa fluidité dans le domaine considéré est alors de 2 $[1/s]$.
2. Un générateur de nombres aléatoires, tel que défini plus haut, délivre ses messages (à 1 bit) à chaque 0.1 seconde. Sa fluidité est alors de 10.

5.11 Simplicité (4d)

Définition :

Dans la théorie MCS, la simplicité est définie comme la propriété d'un objet qui requiert peu d'information pour être exhaustivement décrit. L'unité de mesure pour la simplicité, c'est l'inverse du bit : $[1/bit]$.

Discussion :

La simplicité, c'est l'inverse de la complexité. Ces deux grandeurs peuvent s'estimer très directement sur la base des quantités d'information mises en jeu.

Rappelons que l'on hérite ici des propriétés de l'information telle que définie classiquement ; et notamment, selon théorèmes 1 et 2, de ses caractères fluctuant au cours du temps et très liés au récepteur considéré.

Exemple:

La définition de la simplicité, dans la théorie MCS, tient en 2 lignes et demi. Ceci représente environ 30 mots, soit environ 300 bit. Sur cette base, la simplicité de la définition est donc d'environ $1/300=0.0033$ [1/bit].

5.12 Expertise (5a)

Définition :

Dans la théorie MCS, l'expertise, c'est la propriété d'un système cognitif qui délivre *rapidement* l'information *pertinente*. Quantitativement, l'expertise, E, se mesure comme le produit de la connaissance, K, et de la fluidité, f :

$$E = K \cdot f \quad [\text{lin/s}] \quad (15)$$

L'unité de mesure pour l'expertise est le lin par seconde, [lin/s] .

Discussion :

L'expertise, c'est ce qui permet aux systèmes cognitifs de faire « juste et vite ».

L'expertise c'est la propriété la plus fascinante des systèmes cognitifs²². Dans le langage général, l'expertise a de nombreux synonymes : savoir-faire, compétence, habileté, adresse, excellence, etc.

Si un système est décomposé en sous-systèmes ou réciproquement des systèmes multiples sont intégrés pour former un système plus global, la plupart des caractéristiques cognitives correspondantes, y compris l'expertise, ne peuvent pas se combiner de façon directe et standard. L'analyse de chaque cas doit se faire dans le détail, mettant en évidence pour chaque sous-

²² Cf. un « B-Prize » à définir

système les informations d'entrées et de sorties correspondantes, ainsi que la fluidité. Ce qui est sûr, c'est que les quantités d'information entrant, et respectivement sortant du système à la plus grande échelle restent constantes, pour un domaine cognitif donné. Mais par contre, suivant la structure de (dé-)composition adoptée, il peut apparaître, à l'intérieur du système global, un grand nombre de variantes, et donc d'interfaces différents entre sous-systèmes; cela se reflète évidemment aussi sur les grandeurs cognitives respectives de ces sous-systèmes.

Exemples:

E1. Dans l'exemple vu plus haut, portant sur le domaine D_m , si le système est capable d'assurer l'état pile de la pièce dans la demi-seconde qui suit le lancer aléatoire, la quantité d'expertise est la suivante :

$$E = K \cdot f = 1.6 \cdot 2 = 3.2 \quad [\text{lin/s}] \quad (16)$$

E2. Un robot, tel RH3-Y à Suzhou, durant le championnat mondial Robocup, en juillet 2008, est capable de suivre une personne sans contact, dans un contexte « ouvert », c'est-à-dire déambulant dans le contexte d'un groupe de personnes comprenant du public, à la vitesse d'environ 1 m/s, avec des mouvements d'une grande douceur assurés par le robot. De combien d'expertise fait-il preuve dans ce domaine ?

La réponse peut bien sûr se donner de façon globale ; nous allons le faire au point E2.4 de l'exemple. Mais en fait pour des raisons d'optimisation au stade de la réalisation, il vaut la peine de décomposer le système cognitif global en trois sous-systèmes organisés de façon séquentielle : Perception-Décision-Locomotion. Et pour répondre de façon précise à la question ci-dessus,

l'analyse sera d'abord faite pour chacun de ces sous-systèmes. Finalement le point E2.5 met en perspective les différents résultats.

E2.1 Décision:

En entrée du sous-système de décision, le système tient compte de la position relative de l'humain qui guide le robot. Considérant qu'une précision de l'ordre de 1% sur les coordonnées spatiales du guide est suffisante (2 coordonnées, par exemple de distance et de direction angulaire), nous avons environ 15 bit à l'entrée de ce sous-système:

$$n_{in} = 2 \cdot \log_2(100) \approx 15 \quad [\text{bit}] \quad (17)$$

Se basant sur des considérations internes, notamment d'accélération et de décélération, mais aussi du protocole de commande et de variables d'état internes, ce sous-système transmet finalement au système suivant des consignes de vitesses instantanées tangentielle à la trajectoire (avant/arrière) et de rotation (gauche/droite). Une précision de 1% est largement suffisante:

$$n_{out} = 2 \cdot \log_2(100) \approx 15 \quad [\text{bit}] \quad (18)$$

La fluidité optimale dans ce sous-système est de 1/(25 ms), c'est-à-dire de 40 [1/s]. Dès lors voici la quantité d'expertise du sous-système de décision:

$$E_{\text{Décision}} = K \cdot f = \log_2(n_{out} \cdot 2^{n_{in}} + 1) \cdot f = \log_2(15 \cdot 2^{15} + 1) \cdot 40 \approx 800 \quad [\text{lin/s}] \quad (19)$$

E2.2 Perception:

Mais comment connaître la position du guide ? La solution adoptée pour RH3-Y se base sur l'estimation par un laser à balayage plan d'environ 700 distances radiales, acquises chacune avec une précision de l'ordre du pourcent également ;

$$n_{in} = 700 \cdot \log_2(100) \approx 4650 \quad [\text{bit}] \quad (20)$$

La quantité d'information sortant du sous-système de perception est ici identique à celle qui entre dans le sous-système suivant, de décision (cf. équation 17). Dès lors : $n_{out} \approx 15$ [bit]. Une fluidité de $1/(0.1 \text{ s})$, soit de 10 [1/s], est largement suffisante. Dès lors la quantité d'expertise pour le sous-système de perception est la suivante :

$$E_{Perception} = K \cdot f = \log_2(n_{out} \cdot 2^{n_{in}} + 1) \cdot f = \log_2(15 \cdot 2^{4650} + 1) \cdot 10 \approx 46500 [\text{lin/s}] \quad (21)$$

E2.3 Locomotion:

Le sous-système de locomotion tient compte non seulement des consignes de vitesse instantanées reçues du sous-système de décision mais aussi des perturbations éventuelles sur les roues, principalement dues aux inégalités du sol. Considérant qu'une précision de l'ordre de 1% sur les consignes de vitesse et sur les coordonnées angulaires des deux roues motrices suffit, nous avons environ 30 bit à l'entrée de ce sous-système:

$$n_{in} = 4 \cdot \log_2(100) \approx 30 \quad [\text{bit}] \quad (22)$$

En sortie, le système impose un courant approprié dans chacune des deux roues motrices. Une précision de 1% est largement suffisante:

$$n_{out} = 2 \cdot \log_2(100) \approx 15 \quad [\text{bit}] \quad (23)$$

La fluidité doit être de $1/(1\text{ms})$, c'est-à-dire 1000 [1/s] pour la régulation des moteurs. Dès lors voici la quantité d'expertise du système:

$$E_{Locomotion} = K \cdot f = \log_2(n_{out} \cdot 2^{n_{in}} + 1) \cdot f = \log_2(15 \cdot 2^{30} + 1) \cdot 1000 \approx 34000 [\text{lin/s}] \quad (24)$$

E2.4 Evaluation globale:

A titre d'illustration, il est aussi possible d'évaluer globalement la quantité d'expertise du robot.

Globalement, nous avons en entrée les 700 distances radiales et la position des roues motrices, soit 702 valeurs connues avec une précision de l'ordre du pourcent :

$$n_{in} = 702 \cdot \log_2(100) = 4664 \quad [\text{bit}] \quad (25)$$

Globalement aussi, nous avons en sortie le courant dans les deux roues motrices ; considérons aussi à nouveau une précision de l'ordre du pourcent :

$$n_{out} = 2 \cdot \log_2(100) \cong 15 \quad [\text{bit}] \quad (26)$$

Par ailleurs, il est toujours nécessaire de gérer avec une bonne fluidité les courants moteurs, $f = 1000$ [1/s]. Dès lors voici la quantité d'expertise du système considéré globalement:

$$E_{Global} = K \cdot f = \log_2(n_{out} \cdot 2^{n_{in}} + 1) \cdot f = \log_2(15 \cdot 2^{4664} + 1) \cdot 1000 \cong 4668000 [\text{lin/s}] \quad (27)$$

E2.5 Commentaire général:

L'exemple détaillé fait apparaître plusieurs points intéressants, très représentatifs. La quantité d'expertise requise au niveau global est très grande ; cela peut se voir comme une limite supérieure, a priori exigée par la tâche. Des ressources, du travail et de l'ingéniosité sont nécessaires pour trouver une solution. Structurer la tâche globale de façon à pouvoir la répartir judicieusement sur trois sous-systèmes, respectivement de perception, de décision et d'action constitue une contribution importante, qui permet de réduire les quantités intégrées d'expertise finalement nécessaires. On note des quantités d'expertise plutôt grandes pour les tâches de perception et d'action ; dans le premier cas, c'est-à-dire pour les tâches de perception, les grandes valeurs proviennent de la quantité d'information

à traiter ; dans le second, pour l'action, c'est la grande fluidité requise qui se répercute sur l'expertise requise. La quantité d'expertise nécessaire pour la tâche de décision est comparativement petite.

5.13 Réductibilité (5b)

Définition :

La réductibilité, c'est la caractéristique d'un objet dont il est plus simple de décrire « comment en faire la description » plutôt que de le décrire directement.

Dans la théorie MCS, la réductibilité s'évalue quantitativement comme le rapport de la complexité de l'objet sur la complexité de sa description indirecte.

$$\text{Réductibilité} = \frac{\text{Complexité}(\text{Objet})}{\text{Complexité}(\text{Description}(\text{Objet}))} \quad (28)$$

La réductibilité est un rapport, donc sans unité.

Discussion :

Décrire « directement » l'objet, c'est ce qui détermine la complexité de l'objet. Le décrire « indirectement » signifie par contre qu'on donne les instructions nécessaires afin d'élaborer par ailleurs la description directe de l'objet ; qu'on décrit comment faire la description de l'objet.

La définition classique de Chaitin-Kolmogorov (CK) pour la complexité se focalise sur ce qu'on appelle ici la description indirecte : en substance, selon CK, la complexité d'une chaîne de caractères (tout objet peut se ramener à ce type de représentation), c'est la longueur du programme permettant de générer la chaîne.

Dans la théorie MCS, il est jugé utile de faire la différence entre objets intrinsèquement simples, et

objets réductibles. Il est aussi utile de faire la différence entre description directe, et description indirecte. Dans ce dernier cas, on a forcément besoin de ressources et de temps pour finalement obtenir, si tout va bien, la description de l'objet.

Examples :

E1. Considérons la chaîne de caractères suivante (il s'agit d'un "z" répété 50 fois):

[illegible]

On peut aussi décrire cette chaîne de caractères de la façon suivante :

« Répéter le caractère z 50 fois. »

Dans cet exemple, l'objet est réductible ; et quantitativement, la réductibilité vaut environ 2 :

- La complexité de la chaîne de caractère est de $50 \times 8 = 400$ bit, considérant que chaque « z » est représenté par un octet en code ASCII.
- La complexité du programme est de 31 caractères soit environ $31 \times 8 = 248$ bit, avec le même code ASCII (Notons que pour disposer concrètement de la chaîne de caractère en question, il nous faut encore trouver un ordinateur bien préparé, ou un volontaire humain, pour lire l'instruction et faire le travail).

$$- \text{Réductibilité} = \frac{\text{Complexité}(\text{Objet})}{\text{Complexité}(\text{Description}(\text{Objet}))} = \frac{400}{248} = 1,6 \quad (28b)$$

E2. Considérons le nombre π , sous trois formes.

E2.1 Forme explicite. Le nombre π nécessite une infinité de caractères pour s'exprimer directement,

explicitement. Ce n'est pas possible en pratique. Complexité = ∞ [bit].

E2.2 On peut aussi décrire le nombre π sous les formes suivantes :

- « Le nombre π est le rapport entre circonférence et diamètre du cercle »,
- Ou encore : « $\pi \cong 4 \cdot \left(1 - \frac{1}{3} + \frac{1}{5} - \frac{1}{7} \text{etc.}\right)$ » (cette suite converge vers la valeur exacte).

La complexité de chacune de ces descriptions est ici de l'ordre de la centaine de bit, ce qui conduit à une réductibilité infinie.

E2.3 On peut aussi décrire le nombre π de façon approximative sous la forme suivante : $\pi \cong 3.1416$. La complexité de cette description est alors d'environ 20 [bit]. Sur cette base, la réductibilité du nombre π est à nouveau infinie, mais vu son caractère approximatif, le cas est discutable (quoi qu'il en soit, peut-on en pratique espérer mieux qu'une approximation pour un objet dont la complexité est infinie ?)

Deux conclusions principales s'imposent . C1 : les points E2.1 et E2.2 entraînent que la réductibilité du nombre π est infinie. C2 : Bien que le point E2.3 nous donne une approximation assez grossière du nombre π , on préfère généralement cette expression, directe, immédiate, à la formulation E2.2 qui est elle-même simple et précise, mais nécessite d'être mise en œuvre pour qu'on obtienne concrètement le nombre π .

E3. Considérons l'addition de deux nombres entiers à n chiffres. Pour $n > 1$, le domaine est réductible.

E3.1 La complexité, C , du domaine correspondant à l'addition de deux nombres entiers à n chiffres est la suivante :

$$C \equiv (n + 1) \cdot 3.3 \cdot 10^{2 \cdot n} \quad [\text{bit}] \quad (28c)$$

considérant que l'on a environ 3.3 bit par chiffre, que la table (virtuelle) des solutions contient $10^{2 \cdot n}$ lignes, et que chaque ligne comporte un résultat à $n+1$ chiffres²³.

E3.2 Décrivons la règle d'addition : « Le nombre résultant s'obtient en additionnant le chiffre le plus à droite de la première opérande au chiffre similaire de la seconde opérande ; puis cela se répète vers la gauche, chiffre après chiffre, jusqu'à épuisement. Au cas où le résultat pour une position est plus grand que 9, on lui soustrait 10 et on ajoute 1 (la retenue) à la position suivante.» La complexité, C , de cette règle est de 358 caractères, c'est-à-dire environ 2'900 [bit] (considérant qu'un caractère peut se coder sur 8 bit).

En conclusion de ce troisième exemple, on pourrait penser que, pour de grands entiers, la complexité du domaine semble intuitivement surestimée selon la formule indiquée en E3.1, c'est-à-dire selon la définition de la complexité proposée dans la théorie MCS. Néanmoins il faut bien voir qu'en fait la description simple selon E3.2, indirecte, ne nous donne pas concrètement le résultat ! Le travail reste à faire, avec des ressources et le temps ad hoc à définir et à trouver, ce qui devient même pratiquement impossible pour des valeurs n arbitrairement grandes. La complexité

²³ Dans le cas de l'addition, les nombres résultants ne sont pas équiprobables, à strictement parler. Mais il ne vaut pas la peine ici de faire la différence car cela ne change qualitativement rien d'essentiel à l'exemple.

éventuelle d'un objet ne peut s'ignorer. Sa réductibilité éventuelle n'est pas égale à simplicité. C'est pourquoi les définitions proposées dans la théorie MCS sont utiles.

5.14 Apprentissage (6)

Définition :

Apprendre, c'est la capacité d'un système cognitif à élever son niveau d'expertise au cours du temps (ou selon la définition alternative, avec l'expérience).

Selon la théorie MCS, la quantité résultant de l'apprentissage, L , s'estime comme une différence d'expertise, et l'unité de mesure est donc la même que pour celle-ci, c'est-à-dire le lin par seconde, [lin/s]:

$$L = E(t_1) - E(t_0) \quad [\text{lin/s}] \quad (29)$$

$$L = E(r_1) - E(r_0) \quad [\text{lin/s}] \quad (30)$$

Discussion :

Un système cognitif qui apprend est capable d'augmenter son niveau d'expertise, c'est-à-dire qu'il fait progressivement plus juste (plus de connaissance), ou plus vite (plus de fluidité).

Il peut aussi arriver qu'un système « désapprenne », dans le sens où la quantité d'expertise diminue dans certains cas, au cours du temps et de l'expérience.

Exemples:

1. Un robot explore un labyrinthe, dans un premier temps. Par la suite, il atteint plus vite la sortie car il a mémorisé les impasses et ne s'y engage plus. La quantité de connaissance reste la même ; par contre la fluidité s'améliore.

2. A Atlanta, pour Robocup 2007, dans la ligue « At-Home » le robot RH2-Y apprenait les mouvements à exécuter simplement par l'observation de mouvements exécutés par l'homme à titre d'exemple (test « CopyCat »). Avant l'observation, il n'y avait pas de connaissance ; après observation, dans le cas le plus simple démontré, la quantité apprise était environ la suivante :

$$L = K \cdot f = \log_2(n_{out} \cdot 2^{n_{in}} + 1) \cdot f = \log_2(2 \cdot 2^{320 \cdot 240 \cdot 24} + 1) \cdot \frac{1}{240} \cong 7680 [\text{lin/s}] \quad (31)$$

5.15 Intelligence (7)

Définition :

L'intelligence, c'est la propriété qui rend un système cognitif capable d'apprentissage²⁴.

Dans la théorie MCS, l'intelligence peut s'estimer quantitativement comme un indice, comme le rapport entre quantité apprise, L, et expérience, R.

De la même façon que deux formes d'expérience ont été définies – l'une plus conventionnelle et intuitive, par rapport au temps, et l'autre plus rigoureuse, par rapport à l'information observée -, nous avons ici deux formes pour l'indice d'intelligence :

$$i_{I_t} = \frac{L}{R_T} \left[\frac{\text{lin}}{s^2} \right] \quad (32)$$

$$i_{I_i} = \frac{L}{R_i} \left[\frac{\text{lin}}{s \cdot \text{bit}} \right] \quad (32b)$$

²⁴ La théorie MCS vise à la clarté et à l'essentiel ; de façon générale, aussi bien que dans le cas particulier de la présente définition. Il faut néanmoins reconnaître que dans le langage courant, le mot intelligence est plus ambigu, recouvrant parfois aussi divers autres sens, tels que « information », « connaissance », « renseignement » ou « compréhension ».

Discussion :

L'intelligence peut ainsi se voir comme la dérivée de l'expertise par rapport au temps ; ou mieux, par rapport à l'expérience.

Si possible, il semble préférable pour un système cognitif d'avoir immédiatement un haut niveau d'expertise, c'est-à-dire de n'avoir pas à apprendre. Même si, donc, le fait que l'indice d'intelligence donnerait « 0 » dans ce cas peut sembler paradoxal.

La définition de l'intelligence proposée ici, dans la théorie MCS, nous paraît la meilleure ; elle est applicable aussi bien à l'humain qu'à des systèmes non humains. Néanmoins vu l'importance de cette notion, voici d'autres définitions, brièvement présentées et commentées ici :

D1 – Selon Alan Turing, l'intelligence, pour une machine, c'est la capacité de chatter comme un humain. Cette définition est la plus classique.

D2 – Une approche factuelle consiste à décrire ce qui se passe dans les laboratoires et les bureaux étiquetés comme étant relatifs à l'intelligence (par ex. capacité à jouer aux échecs, à traduire des textes, à émuler des réseaux neuronaux, etc.)

D3.a – Malgré leurs dénégations éventuelles, beaucoup d'humains pensent que l'intelligence est une propriété exclusive de l'humain, et qu'ainsi tout ce qui peut se faire par machine, par définition, ne peut relever de l'intelligence.

D3.b - Une approche plus subtile que la précédente mais menant aussi potentiellement à l'impasse pour les agents non-humains consiste à réserver l'intelligence au

traitement de tâches pour lesquelles « il n'existe pas d'autre solution ».

D4 – Pour ce qui concerne le langage courant, le lecteur est renvoyé aux dictionnaires classiques. On y observe que, suivant le contexte, le sens du mot peut varier de façon substantielle. Notamment, bien que la définition selon la théorie MCS corresponde largement au sens général du mot, il est aussi vrai que parfois le mot intelligence s'utilise couramment dans le public à la place d'autres concepts définis précisément et de façon plus nuancée dans la théorie MCS: connaissance, expertise, etc.

5.16 Correct (Right) (8a)

Définition :

« Correct », c'est la qualité d'une information qui est compatible avec une certaine loi. Il s'agit typiquement d'une grandeur booléenne.

Discussion :

Dans la théorie MCS, « correct » (en anglais « right ») c'est le contraire de « erroné » (en anglais, « wrong »). Dans le langage courant, un synonyme fréquemment utilisé, y compris dans cet ouvrage , c'est « juste»; mais en fait l'antonyme de ce dernier, « injuste », indique clairement qu'il s'agit principalement d'une autre notion.

Exemple:

Si l'assertion à laquelle on se réfère est par exemple la suivante : « Les éléments A et B appartiennent au groupe C», alors il est correct que « A appartient à C ».

5.17 Erroné (Wrong) (8b)

Définition :

« Erroné », c'est la qualité d'une information qui est contraire à une certaine loi. Il s'agit typiquement d'une grandeur booléenne.

Discussion :

Dans la théorie MCS, « erroné » (en anglais « wrong ») c'est le contraire de « correct » (en anglais, « right »).

Dans le langage courant, un synonyme serait semble-t-il « faux » ; mais dans la théorie MCS de dernier terme correspond à une autre notion (voir ci-après sous « faux » - en anglais, « false »).

Exemple:

Si les assertions auxquelles on se réfère c'est par exemple « Les éléments A et B appartiennent au groupe C », et « l'élément D appartient au groupe E », il est erroné que « A appartient à E ».

5.18 Vrai (True) (9a)

Définition :

« Vrai » peut se définir sur la base de « correct ». « Vrai » est égal à « correct », lorsque la loi considérée est la suivante : « correspondance à la réalité ». Il s'agit typiquement d'une grandeur booléenne.

Discussion :

Dans la théorie MCS, « vrai » (en anglais « true ») c'est le contraire de « faux » (en anglais, « false »).

La définition se veut claire, mais comme indiqué au point 2.2, il n'est peut-être guère possible d'appréhender la réalité.

Exemple:

Par exemple, on peut typiquement considérer comme « vrai » que « freiner réduit la vitesse ».

5.19 Faux (False) (9b)

Définition :

« Faux » peut se définir sur la base de « erroné ». « Faux » c'est égal à « erroné », lorsque la loi considérée c'est la suivante : « correspondance à la réalité ». Il s'agit typiquement d'une grandeur booléenne.

Discussion :

Dans la théorie MCS, « faux » (en anglais « false ») c'est le contraire de « vrai » (en anglais, « true »).

La définition se veut claire, mais comme indiqué au point 2.2, il n'est peut-être guère possible d'appréhender la réalité.

Notons que le mot « faux » est particulièrement ambigu en français, étant généralement aussi bien proposé comme traduction du mot « false » que du mot « wrong ».

Exemple:

Par exemple, on peut typiquement considérer comme « faux » que « freiner augmente la vitesse ».

5.20 Bon (9c)

Définition :

« Bon » peut se définir sur la base de « correct ». « Bon » c'est égal à « correct », lorsque la loi considérée est la

suivante : « progression vers un certain but ». Il s'agit typiquement d'une grandeur booléenne.

Discussion :

Dans la théorie MCS, « bon » (en anglais « good ») est le contraire de « mauvais » (en anglais, « Bad »).

Exemple:

Par exemple, on peut considérer comme « bon » d'enclencher les circuits de puissance si le but est qu'un robot bouge.

5.21 Mauvais (9d)

Définition :

« Mauvais » peut se définir sur la base de « erroné ». « Mauvais » c'est égal à « erroné », lorsque la loi considérée c'est la suivante : « progression vers un certain but ». Il s'agit typiquement d'une grandeur booléenne.

Discussion :

Dans la théorie MCS, « mauvais » c'est le contraire de « bon ».

Exemple:

Par exemple, on peut considérer comme « mauvais », si le but est « qu'un robot bouge », que « l'on déclenche les circuits de puissance ».

5.22 Sagesse (10)

Définition :

La sagesse, c'est la propriété d'un système cognitif expert en termes de bonnes décisions, c'est-à-dire qui délivre l'information permettant d'atteindre un certain but.

Dans la théorie MCS, et sur un plan quantitatif, la sagesse s'estime de façon booléenne : vrai ou faux, suivant que le but visé est atteint ou non.

Discussion :

Sans être essentielle, une caractéristique habituelle de la sagesse, c'est d'être liée à des situations complexes, et à des buts d'importance majeure : survivre, gagner le jeu, obtenir une place au tableau d'honneur.

5.23 Sapience (11)

Définition :

La sapience, c'est la propriété essentielle d'un système cognitif, c'est-à-dire d'une structure capable de cognition. La sapience se révèle par un ou plusieurs signes, comme la connaissance, l'expertise, ou l'intelligence (grandeurs déjà définies et rendues mesurables dans la théorie MCS). Sur le plan quantitatif, la sapience peut se caractériser par un indice, i_{sapience} , évalué par référence à l'homme (« homo sapiens »). Il s'agit d'un rapport ; et reste donc typiquement sans unité.

Discussion :

Les performances de l'homme devraient s'évaluer sur les mêmes bases que pour les systèmes artificiels, ne serait-ce que pour quantifier la sapience.

Les premières évaluations semblent indiquer qu'intuitivement, on a typiquement tendance à surévaluer les capacités de l'humain.

Une difficulté pratique c'est par ailleurs que les caractéristiques de l'humain sont des variables stochastiques.

6. Estimation de connaissance en présence d'erreurs

Dans le point précédent, les concepts principaux de la théorie MCS ont été présentés pour les cas de base. En particulier, la connaissance a été définie pour des systèmes cognitifs qui génèrent une information correcte ; éventuellement limitée à un petit domaine, mais néanmoins toujours correcte (cf. Equ. 9).

Une extension très utile pour évaluer la quantité de connaissance d'un système cognitif concerne les cas où le flux d'information généré n'est pas complètement sans erreur.

Dans un tel cas, le système ne connaît pas parfaitement le domaine, D_e . Un message de sortie particulier du système, d_{osj} , diffère alors éventuellement du message correct correspondant dans le domaine, d_{oj} . L'équation 9 est encore applicable en principe :

$$K = \log_2(n_{out} \cdot 2^{n_{in}} + 1) \quad \text{rappel (9)}$$

Mais la part d'information générée de façon incorrecte, c'est-à-dire qui ne correspond pas au domaine D_e (appelons cette part le « bruit », ou l' « erreur ») ne doit pas contribuer au résultat . La quantité d'information correcte délivrée, n_{osc} , peut s'évaluer comme suit :

$$n_{osc} = \sum_{j=1}^n p(d_{osj}) \cdot p(d_{osj} = d_{oj}) \cdot \log_2((p(d_{osj}))^{-1}) \quad [\text{bit}] \quad (33)$$

où $p(d_{osj})$ est la probabilité que le message d_{osj} soit généré par le système, et d_{oj} est le résultat correct

correspondant, c'est-à-dire que le résultat appartient au domaine cognitif considéré, notamment en réponse à un message spécifique entrant le système, d_{ij} . Le terme $p(d_{osj} = d_{oj})$ est la probabilité que le j ème message du système soit correct. L'idée de base est ici que la quantité d'information des messages de sortie soit pondérée par leur probabilité d'être correcte. Si les réponses du système sont toutes correctes, le deuxième terme à la droite de l'Equ. 33 n'a pas vraiment d'effet (facteur égal à 1), et en conséquence les deux quantités d'information n_{osc} et n_{out} auront la même valeur ; de même d'ailleurs que la quantité d'information brute délivrée par le système, n_o . Dans le cas extrême inverse, si les messages de sorties ne sont pas liés au domaine de connaissance, ou, pour dire brièvement, les réponses sont erronées, n_{osc} vaudra zéro, conduisant à zéro [lin] de connaissance, et cela même si n_o est bien plus grand que n_{out} .

En cas d'erreurs éventuelles, c'est alors n_{osc} qui doit remplacer la quantité d'information en sortie d'un système sans erreur dans le domaine, n_{out} , pour l'équation permettant l'estimation de connaissance :

$$K = \log_2(n_{osc} \cdot 2^{n_{in}} + 1) \text{ [lin]} \quad (34)$$

7. Conclusion

Après avoir progressivement assisté ou remplacé l'humain dans ses activités de nature physique, telles la production d'énergie ou le travail mécanique, les systèmes automatisés ont maintenant commencé à envahir le domaine cognitif.

Les sciences cognitives ont sous une forme ou sous une autre commencé leur développement il y a des millénaires, et c'est surtout les notions de modèle et d'information qui sont reprises ici en préalable à la nouvelle définition des termes cognitifs principaux, éléments centraux de la théorie « MCS ».

Les philosophes classiques l'ont bien dit, et l'estimation quantitative que la théorie de l'information permet de le confirmer : la réalité n'est perceptible que de façon infinitésimale. Si l'objectif de la modélisation était la représentation du réel, l'échec serait quasi-total. Mais l'objectif est ailleurs : un bon modèle a le grand mérite de permettre aux systèmes cognitifs, humains ou artificiels, d'atteindre les buts pour lesquels on les a définis.

La théorie MCS définit de façon rigoureuse les grandeurs cognitives essentielles, dont notamment la complexité, la connaissance, l'abstraction et la concrétisation, l'expertise, les notions d'apprentissage et d'intelligence. Elle apporte en plus une métrique pour ces concepts. La prise en compte d'erreurs aléatoires éventuelles est aussi proposée.

La théorie MCS et ses techniques quantitatives apportent une grande clarté au domaine essentiellement rationnel de la cognition ; elles font aussi apparaître de

grandes limites. Il faut aller au-delà de la cognition et constater que d'autres fonctions existent : l'intuition pour toute phase de modélisation afin de faire le pont avec la réalité ; l'éthique, fondamentale en tant que guidant le choix des buts à atteindre ; et les émotions qui assurent le passage du domaine de la cognition à celui de l'action en tant que mise en œuvre physique, qui change le monde.

Annexes

Annexe A. Tableau récapitulatif des concepts présentés – par ordre logique

L'ordre logique adopté ci-dessous assure que les définitions peuvent « s'emboîter », c'est-à-dire qu'au moment où chacune d'elles s'énonce, elle peut faire référence à des termes déjà définis au préalable.

0 Cognitique

1 Modèle

2a Information, 2b Message

3a Complexité, 3b Abstraction, 3c Concrétisation

4a Connaissance, 4b Expérience, 4c Fluidité, 4d Simplicité

5a Expertise 5b Réductibilité

6 Apprentissage

7 Intelligence

8a Correct , 8b Erroné (Wrong)

9a Vrai, 9b Faux (False), 9c Bon, 9d Mauvais

10 Sagesse

11 Sapience

Annexe B. Tableau récapitulatif des concepts présentés – par ordre alphabétique

Concept (français)	Ordre logique des définitions	Table des matières	Concept (English)
Abstraction	3b	5.7	Abstraction
Apprentissage	6	5.15	Learning
Bon	9c	5.20	Good
Cognitique	0	5.1	Cognitics
Complexité	3a	3.3, 5.5	Complexity
Concrétisation	3c	5.7	Concretization
Connaissance	4a	3.1, 5.8	Knowledge
Correct	8a	5.16	Right
Erroné	8b	5.17	Wrong
Expérience	4b	5.9	Experience
Expertise	5a	3.2, 5.12	Expertise
Faux	9b	5.19	False
Fluidité	4c	5.10	Fluency
Information	2a	2.1, 5.3	Information
Intelligence	7	5.15	Intelligence
Mauvais	9d	5.21	Bad
Mémoire		2.3	Memory
Message	2b	5.4	Message
Modèle	1	2.2, 5.2	Model
Réductibilité	5b	3.3, 5.13	Reductibility

Sagesse	10	5.22	Wisdom
Sapience	11	5.23	Sapience
Simplicité	4d	5.11	Simplicity
Vrai	9a	5.18	True

Annexe C: Tableau de valeurs probabilités-logarithmes-information

Tableau comportant quelques valeurs représentatives de probabilités, logarithmes, et de quantités d'information.

<i>Probabilité, p</i>	<i>Probabilité, p</i>	<i>1/p</i>	<i>Log₁₀ : dit</i>	<i>Log₂ : bit</i>
1	1	1	0	0
1/2	0.5	2	0.3	1
1/3	0.33	3	0.5	1.6
1/4	0.25	4	0.6	2
1/8	0.125	8	0.9	3
1/10	0.1	10	1	3.3
1/100	0.01	100	2	6.6
1/1000	0.001	1000	3	10.0
1/1000000	0.000001	1000000	6	19.9

Annexe D. Pour en savoir plus...

Cette annexe comporte schématiquement trois parties. La première se concentre sur les références exploitées de Shannon, Chaitin-Kolmogorov, et Magritte, la deuxième sur les publications faites par l'auteur et d'autres coauteurs, en relation avec la théorie MCS (la plupart peuvent se télécharger, tous droits restant réservés, du site <http://lara.populus.org/rub/3>). Enfin la troisième partie renvoie à deux références, l'une externe additionnelle, française, et l'autre pointant vers une évolution probable du texte au-delà de sa présente forme.

Les articles de l'auteur cités ci-dessous contiennent à leur tour de nombreux renvois à d'autres auteurs travaillant dans les domaines associés ainsi que des remerciements à de nombreuses personnes ayant par leurs discussions contribué à certains de nos travaux. Pour le présent ouvrage, les contributions de Claire pour la couverture, et de Christophe pour de nombreuses remarques et suggestions, notamment relatives à la clarté des explications et à la structure de l'ouvrage, sont à signaler spécialement.

- [1] Claude Elwood Shannon (1948). A mathematical theory of communication. Bell System Technical Journal, 27, 379--423, 623--656.

« La » référence concernant la théorie fondamentale de l'information

- [2] G.J.Chaitin, Scientific American 232(1975)47-52.

- [3] A. N. Kolmogorov, Information theory and the theory of algorithms , ed. by A. N. Shirayev, transl. from the Russian by A. B. Sossinsky Dordrecht (etc.) (Kluwer Academic Publishers, 1993)

Deux références très significatives pour la définition classique de la complexité selon Chaitin et Kolmogorov

- [4] http://www.magritte.com/5_1.cfm?img=04
Site web où les reproductions de certaines œuvres de Magritte, dont le fameux tableau de la pipe, peuvent se commander
- [5] Jean-Daniel Dessimoz, Pierre-François Gauthey, "RH3-Y – Toward A Cooperating Robot for Home Applications", Robocup-at-Home League, Proceedings Robocup08 Symposium and World Competition, Suzhou, China, 14-20 July 2008.
Informations complémentaires pour divers exemples concernant les robots coopératifs pour assistance domestique, et bref résumé de la théorie MCS en anglais. Discussions avec pairs.
- [6] Jean-Daniel Dessimoz, "Cognitique: la cognition automatisée; yc modération de deux tables rondes concernant les aspects scientifiques, économiques et industriels, ainsi que les applications en robotique des services", Forum RobotX organisé par Robot-CH, Y-Parc, Rosam-Diff, et al., Parc scientifique et technologique, Yverdon-les-Bains, 18 juin 2008.
Informations complémentaires disponibles sur internet pour des exemples d'institutions de référence en cognitique (cf. Institut de Cognitique de l'Université de Bordeaux) et de grands projets (cf. projet européen Cogniron, LAAS Toulouse). Discussion avec pairs.
- [7] Jean-Daniel Dessimoz, "Ten Years of Experience with Eurobot ; Achievements, Lessons Learned and General Comments", Proc. Eurobot Conference 2008, Heidelberg, Germany, 22-24 May 2008, ISBN: 978-80-7378-042-5. Paper(.pdf, ca 3.6 Mb), Slides(.pdf, ca 7.7 Mb)
Application et démonstration des concepts de la cognitique quantitative en contexte de robots autonomes. Discussions avec pairs.
- [8] Jean-Daniel Dessimoz and Pierre-François Gauthey, "Quantitative Cognitics and Agility Requirements in the Design of Cooperating Autonomous Robots", Proc. Eurobot Conference 2008, Heidelberg, Germany, 22-24 May 2008, ISBN: 978-80-7378-042-5. Paper(.pdf, ca 2.2 Mb), Slides(.pdf, ca 5.6 Mb)
Application et démonstration des concepts de la cognitique quantitative en contexte de robots autonomes.
- [9] Jean-Daniel Dessimoz, Pierre-François Gauthey, "RH2-Y – Toward A Cooperating Robot for Home Applications", Robocup-at-Home League, Proceedings Robocup07 Symposium and World Competition, Georgia Tech, Atlanta, USA, 30 June-10 July 2007.
Application et démonstration des concepts de la cognitique quantitative en contexte de robots autonomes coopérant avec l'humain. Discussions avec pairs.

- [10] Jean-Daniel Dessimoz, "La robotique appliquée, et son rôle entre science et société", "La robotique en mouvement - opportunités pour l'industrie", org. GESO, EPFL, Robot-CH, HEIG-VD et d'autres, EPFL, Lausanne, 28 juin 2007.
Cognitive appliquée, en contexte de robotique, d'éducation et de recherche. Discussion avec pairs.
- [11] Jean-Daniel Dessimoz, Pierre-François Gauthey and Carl Kjeldsen, "Ontology for Cognitics, Closed-Loop Agility Constraint , and Case Study in Embedded Autonomous Systems – a Mobile Robot with Industrial-Grade Components", Proc. Conf. INDIN '06 on Industrial Informatics, IEEE, Singapore, Aug.14-17, 2006, pp6. Paper (pdf, 778 kB)
Actualisation de la théorie MCS, avec extension à la notion de groupe, et notions associées, notamment culture. Pont établi avec l'automatique, notamment pour en qui concerne des contraintes de temps et d'implémentation.
- [12] Jean-Daniel Dessimoz, Pierre-François Gauthey and Carl Kjeldsen, "Interest of Ludic Competitions for Robotic Education and Research", Workshop on Educational Robotics 2006, org. University of Catania, Eurobot, and IEEE, with support of the European Commission-Directorate General for Research, Science and Society Program , Acireale (Catania), Italy, June 1st, 2006, pp10. Paper (pdf, 1.6 Mb)
Cognitive appliquée, en contexte de robotique, d'éducation et de recherche. Discussion avec pairs.
- [13] Jean-Daniel Dessimoz, "MCS Model for Cognitics and for Addressing «Complex, Human, and Impossible» Issues", New Emerging Science and Technology Event, Brussels, 6 December 2005, 38sl. Presentation: Slides (.pdf, 1.1Mb) Event site: NEST
Présentation synthétique de la théorie MCS, avec des discussions avec pairs sur les notions de complexité, d'impossibilité et de caractère humain. Nécessité d'une cognitive quantitative.
- [14] Jean-Daniel Dessimoz, "About the Necessary Move from Cognitics to Ethics; Additional Definitions, and Contributions to Metrics in MCS ", « DARH-2005 - 1st International Conference on Dextrous Autonomous Robots and Humanoids », with sponsorship Eurobot, IEEE, CLAWAR, and CTI, HESSO-HEIG (West Switzerland University of Applied Sciences), Yverdon-les-Bains, Switzerland, May 19-22, 2005. (re: DARH-2005)
Variante actualisée de la théorie MCS. Extension notamment aux notions d'éthique, de sagesse et de sagesse. Discussion avec pairs.
- [15] Nicolas Uebelhart, Florian Glardon and Pierre-François Gauthey, "Lomu, an Autonomous Mobile Robot with Robust Architecture and

- Components", « DARH-2005 - 1st International Conference on Dextrous Autonomous Robots and Humanoids», with sponsorship Eurobot, IEEE, CLAWAR, and CTI, HESSO-HEIG (West Switzerland University of Applied Sciences), Yverdon-les-Bains, Switzerland, May 19-22, 2005. (re: DARH-2005) AbstractWithFig (1.3Mb)
- [16] A. Trad, P.-F. Gauthey, C. Luethi, J.-D. Dessimoz "Architectural Proposal for Cognitive Processing of Multimodal Images in Large Databases Environments with fast Access" Workshop national IPLnet 2004 ; Automation and Logistics for Manufacturing in Switzerland», Morat, Sept. 6-8, Réseau national de compétences en Productique et Logistique Intégrales IPLnet,. ISBN: 2-8399-0056-4 (voir aussi sur le web. cf : www.iplnet.ch)
- [17] J.-D. Dessimoz, « Knowledge Management and Cognitics; Some Fundamental Aspects" Proceedings « Workshop IPLnet 2002 ; From Research to Application », Kongresszentrum Saas-Fee, Sept. 9-11, 2002, IPLnet National Network of Excellence in Integral Production Automation and Logistics, J.D. Dessimoz, P.-F. Gauthey and J. Moerschel Editors, École d'Ingénieurs du canton de Vaud (EIVD), HES-SO, Yverdon-les-Bains, Febr. 2003. Paper (.pdf, ca 180 kb)
- [18] Jean-Daniel Dessimoz, « Gestion des connaissances et cognitique quelques aspects fondamentaux », Visions, Revue scientifique de l'École d'Ingénieurs du canton de Vaud (EIVD), HES-SO, No 7, Yverdon, CH, novembre 2002
- [19] J.-D. Dessimoz, « ICT for Integral Manufacturing Automation and Logistics » Moderator, Panel Discussion with international Experts : K. Preiss, G. Merli, R. Dillman, and C. Boër ; Workshop du réseau national de compétences en Productique et Logistique Intégrales IPLnet, Locarno/Ascona/Monte Verità, Switzerland. Sept 10-12. 2001 *Application de la cognitique en contexte de productique. Discussion avec pairs.*
- [20] J.-D. Dessimoz et al., « Eurobot », Coupe d'Europe de Robotique, Participation au concours de robots mobiles autonomes, Classe de microtechnique Hesso-Eivd-MI5, La Ferté-Bernard, France (4ième place Prix du Design) May 24-28, 2001
- [21] J.-D. Dessimoz, "Beyond information era : cognition and cognitics for managing complexity; the case of "enterprise", from a holistic perspective", Proc. ICIMS-NOE (Intelligent Control and Integrated Manufacturing Systems- Network of Excellence), ASI - 2000 (Advanced Summer Institute) and Annual Conference, CNRS-ENSERB-Université de Bordeaux, Bordeaux, France , 18-20 Sept. 2000, pp.164-170. Paper (pdf, ca. 227 kB)
Présentation minimale de la théorie MCS et extension à la systémique et à l'entreprise intégrée en réseau.

- [22] J.-D. Dessimoz, "Is a Robot that can Autonomously Play Soccer Intelligent?", Proc. Intern. Conf. on Intelligent Autonomous Systems, No 6, E. Pagello et al (Eds), IOS Press, co-sponsored by IEEE, Venice, Italy, July 2000, pp. 951-958
Variante actualisée de la théorie MCS et mise en perspective avec l'initiative Robocup, intégrant intelligence artificielle et robotique. Discussion avec pairs.
- [23] J.-D. Dessimoz, Pierre-François Gauthey, Michel Etique, Bernard Saugy et Andrea Vezzini, "Serpentine - un Système Intelligent de Transport Urbain pour Passagers, avec des Propriétés de Robots Mobiles Autonomes", Conf. et Actes du Congrès IEEE Francophone pour l'Automatique, CIFA-IEEE, Lille, France, Juillet 2000
Application de la cognitique dans le domaine des transports très automatisés, avec phase de réalisation avancée
- [24] J.-D. Dessimoz, "Robots mobiles: systèmes autonomes intelligents et supports didactiques motivants", Rencontre des enseignants en réglage automatique, SGA/ASSPA - Association suisse pour l'automatique, Ecole d'Ingénieurs du Canton de Vaud Yverdon (VD), 13 mars 2000, pp. 17
Application de la cognitique en contexte didactique
- [25] J.-D. Dessimoz, "Is a Robot that can Autonomously Play Soccer Intelligent?", Rencontre des enseignants en réglage automatique, SGA/ASSPA - Association suisse pour l'automatique, Ecole d'Ingénieurs du Canton de Vaud Yverdon (VD), 13 mars 2000, pp. 5
- [26] J.-D. Dessimoz, "Estimation quantitative et évaluation économique de l'information et de la connaissance", invited author in La Societe Informationnelle: enjeux sociaux et approches économiques, éd. et coord. Anne Mayere, préface : Blaise Cronin, Professor and Dean, University of Indiana, L'Harmattan, Paris, mai 97.
Présentation minimale de la théorie MCS et extension dans le domaine économique.
- [27] J.-D. Dessimoz , A. Beran, S. Ernst, O. Olmo and L. Venries, "Diverses approches pour automatiser les tâches d'assemblage; application à la prise et mise en pile de boîtes de films", Visions, Revue scientifique de l'École d'Ingénieurs de l'État de Vaud (EINEV), No 1, Yverdon, CH, pp35-43, nov. 1995
Similaire à [27] et [28], en version française.
- [28] J.-D. Dessimoz, "Quantitative Assessment and Economical Evaluation of Information and Knowledge", Actes du Colloque International Economie de l'Information, ENSSIB (Ecole Nationale Supérieure des Sciences de l'Information et des Bibliothèques), Lyon, France, 10-12 Mai, 1995, pp. 185-200.

Présentation minimale de la théorie MCS et extension dans le domaine économique. Discussion avec les pairs.

- [29] J.-D. Dessimoz, A. Beran, S. Ernst, O. Olmo and L. Venries, "Two Antinomical Methods For Automatically Solving Assembly Tasks; Case of Picking and Stacking Film Cans", Proc. 1994 Latsis Conference: From Perception to Action (Per'Ac), at EPFL, Lausanne, Switzerland, Sept. 4-9, published by IEEE Comp. Soc., New York, 1994 (winner of Robot Contest).

Traitement d'une tâche cognitive très exigeante sous trois formes complémentaires : théorique avec métrique MCS, et appliquées dans un cas avec système cognitif «classique » : robot industriel et vision par ordinateur ; et dans l'autre avec un système « câblé », c'est-à-dire sans ordinateur ni processeur numérique – Rendu théorique sur papier.

- [30] J.-D. Dessimoz, S. Ernst, and L. Venries, "Picking and Stacking Film Cans; Case A: Robot with vision", Videotape (10 min.) and Accompanying document for Videotape, 1994 Latsis Conference: From Perception to Action (Per'Ac), at EPFL, Lausanne, Switzerland, Sept. 4-9, published by EPFL/LAMI, 1994.

Traitement d'une tâche cognitive très exigeante sous trois formes complémentaires : théorique avec métrique MCS, et appliquées dans un cas avec système cognitif «classique » : robot industriel et vision par ordinateur ; et dans l'autre avec un système « câblé », c'est-à-dire sans ordinateur ni processeur numérique – Illustration expérimentale.

- [31] J.-D. Dessimoz and Giovanni Mele, "Performance assessment of cognitive systems; case of elementary mobile robots", Proc. ECAI 94, 11th European Conference on Artificial Intelligence, Amsterdam, 7-12 Aug., A. Cohn. ed., John Wiley & Sons, New York, pp. 689-693, 1994
Conférence où les fondements de la théorie MCS ont pu formellement se discuter entre pairs. Exemples détaillés pour robots mobiles. Version anglaise de la théorie MCS.

- [32] J.-D. Dessimoz, "Proposition de définitions et d'une métrique pour les sciences cognitives", Actes du Congrès bisannuel de l'AFCET (Assoc. française des sciences et technol. de l'inform. et des syst.), Versailles, France, 8-10 Juin 1993.

Conférence où les fondements de la théorie MCS ont pu formellement se discuter entre pairs, ayant donné lieu à une publication plus développée qu'en [32].

- [33] J.-D. Dessimoz, "Towards a Theory of Knowledge", Proc.3rd Ann. Conf. of the Swiss Group for Artificial Intelligence and Cognitive Sciences (SGAICO'91), Inst. for Informatics and Applied

- Mathematics, Univ. of Bern, Switzerland, Vol. IAM-91-004, pp. 60-71, March 1992.
Variante plus élaborée que [33] contenant les fondements de la théorie MCS .
- [34] J.-D. Dessimoz, "Towards a Theory of Knowledge", SGAICO Conf., Swiss Group for Artificial Intelligence and Cognitive Science, Bienne, Oct. 10-11, 1991.
Premier forum où les fondements de la théorie MCS ont pu formellement se discuter entre pairs.
- [35] J.-D. Dessimoz, "Knowledge in formulas", SGAICO Newsletter, SI Information, Soc. Suisse des Informaticiens, Zürich, Aug. 1991.
Première publication contenant les fondements de la théorie MCS ; sous forme très concise
- [36] Bernard Claverie, « Cognitique - Science et pratique des relations à la machine à penser », L'Harmattan, ISBN : 2-7475-9135-2, nov. 2005, pp.150.
Point de vue général et commentaires concernant les applications et la situation française
- [37] <http://cognitics.org> et <http://cognitique.populus.ch>
Points d'entrée (dès l'an 2000) pour sites en cognitique. Y c. présentation brève de la théorie MCS pour les sciences cognitives, références, et accès à publications mise-à-jour

Table des matières

Préface	5
1. Introduction.....	7
2. Terrains connus ?	9
2.1 L'information	9
2.2 La notion de modèle.....	15
2.3 Mémoire	19
3. Grandes étapes de développement de la théorie MCS. ...	21
3.1 <i>Vers le concept de connaissance</i>	21
3.2 <i>Vers le concept d'expertise</i>	23
3.3 <i>Autres développements</i>	24
4. Caractéristiques générales du modèle proposé	25
4.1 <i>Modèle comportemental</i>	26
4.2 <i>Modèle applicable à tout niveau de détail</i> <i>(granularité)</i>	27
4.3 <i>Modèle en principe indépendant de la nature</i> <i>physique du support d'implémentation</i>	27
4.4 <i>Exemples de systèmes cognitifs</i>	28
4.5 <i>But de la théorie MCS</i>	29
5. Concepts principaux de la théorie MCS.....	30
5.1 La cognitive (0).....	31
5.2 Modèle (1).....	32
5.3 Information (2a)	35

5.4 Message (2b).....	36
5.5 Complexité (3a)	37
5.6 Abstraction (3b)	40
5.7 Concrétisation (3c)	41
5.8 Connaissance (4a)	42
5.9 Expérience(4b)	51
5.10 Fluidité (4c).....	52
5.11 Simplicité (4d).....	53
5.12 Expertise (5a)	54
5.13 Réductibilité (5b)	59
5.14 Apprentissage (6).....	63
5.15 Intelligence (7).....	64
5.16 Correct (Right) (8a)	66
5.17 Erroné (Wrong) (8b).....	67
5.18 Vrai (True) (9a).....	67
5.19 Faux (False) (9b).....	68
5.20 Bon (9c)	68
5.21 Mauvais (9d).....	69
5.22 Sagesse (10)	69
5.23 Sapience (11)	70
6. Estimation de connaissance en présence d'erreurs.....	71
7. Conclusion.....	73
Annexes.....	75
Annexe A. Tableau récapitulatif des concepts présentés – par ordre logique	75
Annexe B. Tableau récapitulatif des concepts présentés – par ordre alphabétique	76
Annexe C: Tableau de valeurs probabilités-logarithmes- information	77
Annexe D. Pour en savoir plus.....	78

Achevé d'imprimer la première édition : août 2008

Dépôt légal: prévu de suite