

# ‘Le modèle cognitif computationnel est infondé’

Michel Troublé - directeur de recherche

La *cognition* est la capacité que possèdent nécessairement tous les êtres *vivants* à répondre d’une façon adaptative et efficiente aux sollicitations infiniment variées de l’environnement qu’ils perçoivent avec leurs différents capteurs, afin que soit toujours assurée leur pérennité, ce qui les caractérise.

Selon la ‘théorie computationnelle de la cognition’, les capacités *cognitives* peuvent toutes s’expliquer par des calculs qui sont réalisés dans le cerveau, calculs qui sont généralement effectués par les réseaux neuronaux du cerveau. L’*esprit* ou le *mental* qui est le produit de ces réseaux neuronaux n’est donc qu’un épiphénomène.

Pour évaluer le bien-fondé de la ‘théorie computationnelle de la cognition’, il faut préalablement disposer d’une définition *opérationnelle* précise des structures physico-chimiques que sont ces êtres *vivants*. Les outils mathématiques et numériques de la robotique permettent précisément de spécifier ce qui caractérise ces êtres *vivants* – des structures *sensori-motrices* – en matière des *actions* qu’ils accomplissent afin que soit assurée la pérennité de leur structure, ce qui les caractérise.

Un robot est un système physico-chimique *sensori-moteur* qui est au moins doté des éléments suivant : un *capteur* qui met en relation le robot avec le milieu extérieur, un *actionneur* qui assure sa mobilité et un *contrôleur* qui établit des liens informationnels appropriés entre capteurs et actionneurs. Le *contrôleur* est ce qui assure la *cognition* du robot.

L’analyse fonctionnelle d’un robot qui serait *autonome* en étant plongé dans un environnement multiforme infiniment variable, montre qu’il devrait être doté, à l’instar des êtres *vivants*, de la capacité *d’agir d’une façon cohérente* sur les objets de son environnement afin qu’en toutes circonstances puisse être assurée la pérennité de sa structure. Par exemple, *fuir* systématiquement tous les objets chauds rencontrés comme de la lave en fusion ou un feu de forêt qui en le détruisant à coup sûr serait totalement contraire aux attentes de son constructeur pour lequel la réalisation de ce robot avait été longue et très dispendieuse.

Cette analyse fonctionnelle conduit par là même à formuler une définition opérationnelle des structures physico-chimiques *vivantes* :

Les êtres *vivants* sont des structures physico-chimiques qui ont la capacité essentielle de ***catégoriser d’une façon cohérente*** les objets du monde qu’ils perçoivent avec leurs capteurs afin que les *actions* qui en résultent assurent la pérennité de leur structure.

Par exemple : *fuir* de la lave en fusion, *fuir* un feu de forêt, *fuir* une nappe de pétrole enflammée,...objets qui sûrement les détruiraient. Soit créer la ***catégorie cohérente*** {*fuir* tous les objets chauds}, alors qu’il n’existe aucune affinité physico-chimique particulière de ces objets avec la structure qui pourrait expliquer ce type d’actions pérennisantes.

Cette définition implique logiquement que les objets du monde perçus par les capteurs du robot soient physiquement *discernables* afin que celui-ci puisse, par exemple, *fuir* systématiquement tous les objets chauds qui sûrement le détruiraient.

Que les objets du monde perçus par les capteurs du robot soient physiquement *discernables* afin que celle-ci puisse *fuir* systématiquement les objets chauds, semble aller naturellement de soi. Mais cela pose en réalité un problème épistémologique majeur :

En s'appuyant sur la théorie formelle de la « Reconnaissance des formes (Satosi Watanabe - *Pattern recognition, human and mechanical*) » qui porte sur l'identification des formes d'objets à partir de leurs paramètres caractéristiques afin de prendre des *décisions* dépendant des catégories attribuées à ces formes, on peut en effet démontrer la propriété fondamentale suivante :

**Les différentes formes des objets qui sont perçues par un système physique lors d'un processus de mesure, sont physiquement indiscernables par sa partie opérative ou actionneur (système de locomotion, bras manipulateur).**

Ce « théorème d'indiscernabilité », pour faire court, s'applique à tous les niveaux de la matérialité – macroscopiques ou microscopiques/quantiques –, indépendamment des lois physiques qui régissent ces domaines. Il peut être établi<sup>1</sup> en analysant la nature des liaisons physiques qui doivent être mises en place entre le *capteur* et l'*afficheur* d'un appareil de mesure dont la fonction essentielle est de déterminer les propriétés spécifiques des objets du monde avec lesquels un système physique interagit.

Cet état d'*indiscernabilité* fondamental des entités matérielles a été jusqu'alors complètement ignoré des chercheurs pour lesquels la *discernabilité* des objets macroscopiques ou microscopiques perçus/mesurés, allait de soi et qu'il était de ce fait aucunement nécessaire de se poser la question du bien-fondé d'une telle affirmation.

Il est à noter que le 'théorème d'indiscernabilité' est similaire à celui qui résulte de l'analyse de la vérité des propositions dans le domaine du langage que fait le logicien Ludwig Wittgenstein. Ainsi, dans son ouvrage le « *Tractatus Logico-philosophicus* », Ludwig Wittgenstein affirme, en déclinant l'ensemble de toutes les propositions possibles construites à partir de descripteurs élémentaires « Il ne peut être conclu d'aucune manière de l'existence de l'état d'une chose à l'existence d'un état de choses totalement différent [article 5.135] ».

Autrement dit, le langage qui se fonde naturellement sur des *propositions* construites à partir de *descripteurs* élémentaires comme 'lourd' ou 'léger', ne peut en aucune façon être la source d'*actions* spécifiques *différenciées* comme {éviter une 'lourde' roche qui en tombant de la montagne pourrait nous *tuer*} mais au contraire {*ramasser* un morceau de pain 'léger' qui pourrait assouvir notre *faim*}. L'analyse que fait Wittgenstein est d'importance car cela signifie que les *concepts*, ces objets mentaux qui fondent tous les discours que nous faisons sur le monde, sont logiquement irréalisables puisqu'ils impliquent la *catégorisation cohérente* des différents états d'une chose qui, logiquement, demeurent fondamentalement *indifférenciés*.

En tant que processus de contrôle, la *cognition* d'un robot *autonome* qui le conduirait à s'opposer à la dégradation de sa structure en créant des *catégories d'actions cohérentes* qui fondent son état de 'vie', ne peut donc pas résulter d'un simple calcul, et ceci quel que soit le processus technique mis en œuvre (calculateur séquentiel ou massivement parallèle, calculateur optique, réseaux de neurones, processus évolutionnistes, algorithme génétique, recuit simulé, machine de Turing...).

---

<sup>1</sup> voir Annexe

Les différentes solutions techniques calculées en ‘aveugle’ à partir des données venant des différents capteurs du robot seraient en effet *indiscernables* par son actionneur (systèmes de locomotion ou de manipulation) en vertu du ‘théorème d’indiscernabilité’.

Telle qu’elle est énoncée, la ‘théorie computationnelle de la cognition’ est donc infondée ou tout au moins incomplète. La cognition qui permet des systèmes physiques de répondre d’une façon adaptative aux sollicitations infiniment variées de leurs environnements ne peut pas seulement résulter d’un calcul.

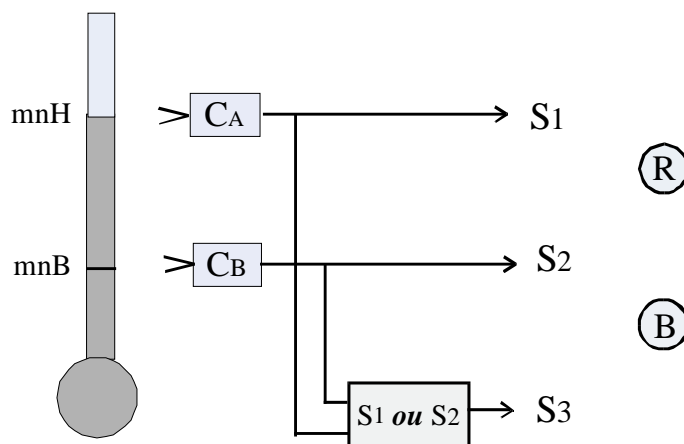
Eu égard le ‘théorème d’indiscernabilité’, ce qui fait logiquement défaut au modèle *cognitif computationnel*, c’est d’être muni d’un opérateur ‘choisisseur’ qui puisse faire une sélection efficiente parmi les solutions techniques autrement *indiscernables* calculées par les réseaux neuronaux du cerveau.

Empiriquement, la *conscience* ou *mental* qui nous ouvre à la vision *sensible* du monde, pourrait être cet opérateur ‘choisisseur’. C’est la *sensibilité* dont est muni cet opérateur *mental* qui en assurant la pérennité, donc la vie, de la structure à laquelle cet opérateur est associé, rendrait cette *cognition* efficiente.

## Annexe :

Pour effectuer la mesure d’un observable **O** (température, poids, longueur,...) sur un objet **A**, il faut que cet objet interagisse avec un dispositif technique donné **Ap** qui a la particularité de se mettre dans un état final spécifique unique **E** lorsque l’interaction est complète.

Considérons un dispositif de mesure thermométrique constitué des éléments suivants : un *capteur* (un thermomètre), un *afficheur* (lampes ‘rouge’ et ‘bleue’), des *liaisons physiques* entre le capteur et l’afficheur. Le monde sur lequel portent les mesures thermométriques est supposé constitué des seuls deux objets nommés ‘chaud’ et ‘froid’. Les mesures sont ainsi relatives au seul observable **O**, la *température*.



Deux cellules photoélectriques **CA** et **CB** – seulement sensibles à la forme spécifique du ménisque de mercure dans le tube capillaire – sont positionnés en deux points du capillaire qui correspondent aux deux positions possibles atteintes par le ménisque suivant que c'est l'objet 'chaud' ou 'froid' qui fait l'objet de la mesure.

Lorsque l'objet 'chaud' se trouve devant le bulbe du thermomètre et que la mesure est complète (la colonne de mercure est stabilisée, les positions transitoires sont ignorées), seule la cellule photoélectrique **CA** est activée, soit un signal de sortie **SA** = 1, avec **SB** = 0. Pour l'objet 'froid', c'est seulement la cellule **CB** qui est activée, soit un signal de sortie **SB** = 1, avec **SA** = 0.

Pour collecter, d'une façon *exhaustive*, toutes les *informations* fournies par les deux sorties **S1** et **S2** du *capteur* thermométrique, il y a 3 et non pas deux 2\_ façons de prendre en compte les deux sorties **S1** et **S2** afin de produire une *action* comme d'allumer la lampe 'Rouge' ou la lampe 'Bleue' de l'afficheur :

- ne prendre en compte que *la seule* sortie **S1**
- ne prendre en compte que *la seule* sortie **S2**
- prendre en compte *à la fois* les deux sorties **S1** et **S2**.

Pour mettre en communication les deux sorties **S1** et **S2** à l'unique entrée qui est relative à la lampe 'rouge' ou à la lampe 'bleue', il faut donc utiliser soit la sortie **S1**, soit la sortie **S2**, soit l'association des deux sorties **S1** et **S2**. Il existe deux *associations* possibles des deux sorties **S1** et **S2** :

- une *association* **A1** qui correspondant à l'opérateur logique 'ou', soit **A1** = {**S1 ou S2**}
- une *association* **A2** qui correspondant à l'opérateur logique 'et', soit **A2** = {**S1 et S2**} qui est toujours égale à **0** puisque le ménisque de mercure du thermomètre ne peut être en même temps en position haute et basse du tube capillaire. À ce titre, l'association logique **A2** = {**S1 et S2**} n'est pas physiquement possible. Reste donc en lice la seule association possible **A1** = {**S1 ou S2**}.

D'où l'existence des 3 sorties **S1**, **S2**, **S3** dont chacune peut être ou non activée pour produire une action comme d'allumer une lampe 'Rouge' ou 'Bleue'.

### Protocole de mesure :

- (1) Un technicien **T1** met un objet nommé 'chaud' devant le bulbe du thermomètre. Un second technicien **T2** note alors que **S1=1** et **S3=1**, avec **S2=0** (non-opératif)
- (2) Le technicien **T1** met un objet dit 'froid' devant le bulbe du thermomètre. Le technicien **T2** note alors que **S2=1** et **S3=1** avec **S1=0** (non-opératif)

D'où le tableau récapitulatif :

présentation d'un objet ' <b>chaud</b> '	<b>S1=1</b>	<b>S3=1</b>
présentation d'un objet ' <b>froid</b> '	<b>S2=1</b>	<b>S3=1</b>

N'ayant a priori aucune connaissance relative à la nature des objets nommés 'chaud' et 'froid' mis en place devant le bulbe du thermomètre par le technicien **T1**, le technicien **T2** doit en déduire les résultats suivants :

1. En prenant en compte la deuxième colonne du tableau : les deux objets 'chaud' et 'froid' sont *différents*, car les sorties **S1** et **S2** activées sont *différentes*. D'où les deux actions possibles : (1) allumer la lampe 'Rouge' lorsqu'il s'agit de la présentation d'un l'objet 'chaud', (2) allumer une lampe 'Bleue' à la présentation d'un objet 'froid'.

2. En prenant en compte la troisième colonne du tableau : les deux objets ‘chaud’ et ‘froid’ sont *identiques*, car les sorties **S1** et **S2** activées sont *identiques*.

Étant à la fois *différents* et *identiques*, les objets nommés ‘chaud’ et ‘froid’ mesurés par le capteur thermométrique sont donc physiquement *indiscernables* du point de vue du technicien **T2** et donc de l’*actionneur* auquel ce *capteur* peut être associé. C’est le ‘théorème ‘indiscernabilité’.