



**HAL**  
open science

## Relation entre IA symbolique et IA forte

Arnaud Kohler

► **To cite this version:**

| Arnaud Kohler. Relation entre IA symbolique et IA forte. 2020. hal-02444894v2

**HAL Id: hal-02444894**

**<https://hal.science/hal-02444894v2>**

Preprint submitted on 18 Dec 2020

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

## Relation entre IA symbolique et IA forte

Arnaud Kohler

**Résumé** Intelligence Artificielle, ou IA. IA faible, IA forte, IA symbolique, IA connexionniste, IA neuronale, IA émotionnelle, IA générale... Cette profusion de termes, au-delà d'un effet de communication, masque une réalité (l'intelligence artificielle ne convainc pas en tant qu'intelligence) et une question ouverte (qu'est-ce que l'intelligence ?). Nous n'avons ni la compétence ni la prétention de tenter d'y apporter une réponse. Plus modestement, nous proposons d'enrichir la réflexion. Dans le cadre de l'IA dite symbolique, nous présentons une approche qui, à défaut de pouvoir apporter à date la preuve de sa pertinence, espère être originale dans la grande famille des langages formels. En intégrant la notion de pensée dans le formalisme, elle ouvre une voie de recherche nouvelle qu'il nous semble intéressant de parcourir.

**Abstract** Artificial Intelligence, or AI. Weak AI, strong AI, symbolic AI, connectionist AI, neural AI, emotional AI, general AI... This profusion of terms, beyond a communication effect, masks a reality (artificial intelligence does not convince as intelligence) and an open question (what is intelligence ?). We have neither the competence nor the pretension to try to provide an answer. More modestly, we propose to enrich the reflection by proposing, within the framework of the so-called symbolic AI, an approach which, failing to be able to provide proof of its relevance to date, hop to be original in the large family of formal languages. By integrating the notion of thought into formalism, it opens up a new avenue of research that we think is interesting to explore.

### Préambule

Après avoir vécu une période glorieuse, l'Intelligence Artificielle symbolique est entrée, dans les années 1990, dans un hiver gris et froid. Il perdure depuis. Ses chercheurs, autrefois présomptueux, se cachent aujourd'hui et sont à leurs tours moqués : « « *La* » logique, à supposer qu'il convienne encore de s'y intéresser, serait d'ailleurs, de l'aveu même des logiciens vraiment modernes, mal nommée : il n'y a que des logiques, en quantité indéfinie, qui sont comme les outils d'une trousse de bricoleur » (Daniel Andler, 2004). Le jugement se veut sans appel. Si la logique a été pendant quelques siècles une des disciplines majeures de la

philosophie, elle se révèle être finalement une distraction pour bricoleurs. Le temps est venu de *La revanche des neurones* (Dominique Cardon, Jean-Philippe Cointet et Antoine Mazières, 2018).

L'IA symbolique proposait d'exploiter les connaissances par l'application de relations logiques entre des propositions unitaires (« *si on est en hiver alors il fait froid dehors* » par exemple). Le postulat est que l'intelligence est une affaire de raisonnement. Le but est alors de découvrir et de modéliser les règles permettant de raisonner justement. Dans le dernier quart du XX<sup>ième</sup> siècle, elle prétendait la venue de robots autonomes qui remplaceraient l'humain dans ses tâches les plus complexes. La promesse n'a pas été tenue. Les « systèmes experts » qu'elle a générés sont bien encore utilisés pour quelques applications, mais ne bénéficient plus d'aucune reconnaissance publique.

L'IA connexionniste (ou neuronale) exploite les données grâce à une nouvelle génération d'outils statistiques (« *si on est en hiver alors il y a une probabilité de 80% qu'il fasse froid dehors* », par exemple). Elle fait de la reconnaissance de sons et d'images, de l'analyse et de la rédaction de textes, du diagnostic médical, conduit des voitures, gère des transactions boursières, joue au go, est championne d'échecs... Surfant sur le Big Data qui lui met à disposition des bases de données gigantesques, elle bénéficie depuis environ 10 ans d'une communication d'envergure et d'un soutien financier, peut-être insuffisant mais néanmoins réel (voir Cédric Villani, 2018). L'intelligence semble être une méga matrice statistique.

Cette conclusion est hâtive. Toutes les applications à ce jour, qu'elles soient symboliques ou connexionnistes, ont en effet une limite commune : les systèmes captent, raisonnent et interagissent, dans le contexte environnemental pour lesquels ils ont été conçus, aux questions que leurs concepteurs avaient anticipées. Ils répondent mécaniquement selon les processus plus ou moins complexes qui leur ont été inculqués. Ils réagissent à ce pour quoi, et de la manière dont, il était prévu qu'ils réagissent. Ils ne s'adaptent pas. Ils n'apprennent pas au-delà du périmètre pour lequel ils ont été programmés à apprendre.

Dans la mécanique de raisonnement, la statistique apporte à l'IA connexionniste une capacité d'induction qui fait défaut à l'IA symbolique. Et le Big Data lui ouvre une dimension vertigineuse. Mais l'intelligence résiste encore à la probabilité d'occurrence. Par exemple, avec les outils actuels, l'absorption par une IA connexionniste de plusieurs milliers d'images

de chats ne lui garantit pas la reconnaissance d'un chat. Dans le même temps, un enfant obtient des résultats au moins équivalents à partir de quelques photos.

En fait, si l'ensemble de ces systèmes relève de la très belle orfèvrerie, ils ne bénéficient pas, dans leurs processus d'interaction et décision, de la stimulation qu'apportent les émotions. Ils ne ressentent pas d'ennui, d'envie ou de plaisir. Ils ne ressentent pas l'émerveillement de l'enfant qui voit un chat. Cela, ce sera l'IA forte. Elle sera symbolique, connexionniste, ou les deux. Ou autre chose. Sera-t-elle quantique ? La réponse serait dans l'usage de la force, et peut-être découvrirons-nous un jour que le cerveau de nos lointains cousins les singes est « moins quantique » que le nôtre. Compte tenu des connaissances à date, ce serait une erreur d'écarter une voie de recherche qui n'aurait pas été complètement parcourue.

Notamment, l'IA symbolique a rencontré de nombreux échecs au XX<sup>ième</sup> siècle. Mais échec ne valant pas preuve, l'ouvrage doit être remis sur le métier. Si d'importants chantiers ont été menés, tous les sentiers n'ont pas été explorés. Nous proposons de parcourir l'un d'entre eux.

Dans ce document, nous éviterons autant que possible l'emploi de formules dans la rédaction. Mais lorsqu'il nous semblera que le signe logique remplace avantageusement un long discours, nous utiliserons les connecteurs de négation ( $\neg$ ), de disjonction ( $\vee$ ), de conjonction ( $\wedge$ ), d'implication ( $\rightarrow$ ) et d'équivalence ( $\leftrightarrow$ ), et les signes de production syntaxique ( $a \vdash b$ , qui signifie : «  $b$  est produit syntaxiquement de  $a$  ») et d'interprétation sémantique ( $a \models b$ , pour «  $b$  hérite d'un sens de  $a$  »).

### **Les causes de l'échec de l'IA symbolique**

Il est un avis souvent partagé : les langages formels obéissent à une syntaxe et à une sémantique toutes les deux strictes. Ils permettent d'exposer des énoncés précis et sans ambiguïté. Ils utilisent des règles qui transforment mécaniquement les énoncés, indépendamment de leurs significations, et qui garantissent leur valorisation sémantique (classiquement « vrai » ou « faux »). Ce sont des langages universels et techniques, peut-être adaptés aux disciplines scientifiques – mais en tous cas non naturels, et incapables de modéliser toutes les subtilités de la pensée humaine. Visitons quelques arguments qui confortent cette assertion.

### 1) *Le raisonnement humain ne suit pas la logique d'un raisonnement formel*

Il existe aujourd'hui une multitude de langages formels : la logique propositionnelle, du  $n^{\text{ième}}$  ordre, des défauts, de circonscription, paradoxale, para-cohérente, floue, multivaluée, temporelle, multi-agents, modale, multimodale, etc. Elles ont pour la plupart été formalisées dans la seconde moitié du  $XX^{\text{ième}}$  siècle. Chacune a des propriétés qui lui sont propres, et qui permettent de capturer un mode particulier de raisonnement. Mais elles échouent dans la tentative de modéliser, au travers d'un formalisme unique, tous les types de raisonnement.

Face à ce constat d'échec, une explication souvent avancée est que le raisonnement humain ne suit pas la mécanique d'un langage formel, mais d'un autre mode – dont chaque langage formel connu ne serait qu'une approche partielle et simplificatrice. L'intelligence humaine suivrait *une règle d'inférence non démonstrative* (voir par exemple Jean-Luc Roulin, 2006). Echappant à la logique formelle, elle resterait à découvrir.

### 2) *Les langages formels n'acceptent pas l'incohérence*

Les logiques formelles ont une origine commune, exprimée par les 3 principes d'Aristote : ce qui est est (principe d'identité), ne peut être que vrai ou faux (principe du tiers exclu), et ne peut pas être à la fois vrai et faux (principe de non-contradiction). L'objectif est d'identifier les règles qui garantirait un raisonnement juste. Le refus de l'incohérence est donc inscrit dans les gènes des langages formels.

Il n'est pas besoin d'épiloguer pour identifier les limites de cette approche : la pensée humaine est, c'est une évidence, pétrie de paradoxes et d'incohérences. Elles peuvent être immédiates : je crois  $a$  et  $\neg a$ . Et elles peuvent être obtenues indirectement par des raisonnements erronés : nombre de nos pensées sont fausses, alors même que nous avons, dans le même instant, les connaissances nécessaires pour penser justement.

Le problème d'un langage formel est que la présence d'une incohérence syntaxique lui permet d'affirmer tout et son contraire : depuis  $a \wedge \neg a$ , je peux syntaxiquement déduire  $b$  quel que soit  $b$ . C'est le principe dit d'explosion, et elle n'est évidemment pas acceptable. De nombreuses tentatives ont été faites pour y échapper (voir les logiques para-cohérentes, multivaluées ou floues par exemples), mais sans succès définitivement convaincant.

### 3) *Les langages formels ne savent pas gérer des nouvelles informations*

L'appropriation et l'exploitation de nouvelles informations est une mécanique naturelle à l'être humain. C'est un pré requis de l'intelligence – du moins dans une de ses définitions les plus répandues, qui est celle d'un processus permettant de s'adapter à des situations nouvelles (au sens non anticipables). En IA symbolique, les nombreuses études en la matière sont portées par la discipline connue sous le nom de *Révision des Croyances*. Aucune n'a abouti.

L'énoncé du problème est pourtant d'une extrême simplicité : face à deux informations incohérentes, laquelle retenir ? L'être humain fait face tous les jours, et à de multiples occasions, à des situations de ce genre. Le principe d'explosion paralyse un système logique, il n'a pas d'autre choix que de remédier à l'incohérence. Suite à l'échec des approches paracohérentes, Peter Gärdenfors et David Makinson (1991) ont proposé un nouveau concept, celui de *l'enracinement épistémique*.

Il suppose l'existence d'une sorte de relation d'ordres entre les connaissances. Propre à chaque individu, elle lui permettrait de sélectionner les informations en fonction des niveaux de pertinences qu'il lui prête. Cette solution n'est malheureusement pas satisfaisante. Que faire lorsque deux informations incohérentes ont, après application de l'enracinement épistémique, le même niveau de pertinence ? Il faudrait imaginer une « méta relation d'ordres », qui entraînerait inévitablement le besoin d'une « méta méta relation d'ordres », etc. Modéliser le problème revient à créer une boucle infinie. La solution serait donc en dehors du formalisme.

### 4) *Les langages formels ont une sémantique trop stricte*

La syntaxe et la sémantique d'un langage formel obéissent à des règles strictes et précises : chaque proposition atomique a un sens clairement défini et unique, associé à une valeur de vérité (classiquement « vrai » ou « faux ») précise et sans ambiguïté. Par une forme d'héritage, chaque phrase (ou formule) est vraie ou fausse. L'ensemble des formules définit les *mondes possibles* (l'expression est de Saul Aaron Kripke, 1971). Elles associent des valeurs de vérité à chaque proposition atomique qu'elles portent. S'il y a une limite aux

langages formels, elle proviendrait de la difficulté à obtenir l'exhaustivité des formules qui décriraient le monde dans sa complétude.

Imaginer la base des connaissances formelles qui décrit la totalité du monde relève en effet de la gageure. En face, le langage naturel n'a pas cette prétention. Et il assume ses imprécisions : un mot ou une phrase peuvent avoir plusieurs sens. Cela entraîne peut-être des difficultés dans la modélisation des matières scientifiques, mais c'est néanmoins une des sources de sa richesse et de sa puissance d'expression.

#### 5) *Les langages formels ne permettent aucune créativité*

Les langages formels n'ont qu'une règle de production, qui est le *modus ponens*. C'est la règle de déduction : si  $(a)$  et  $(a \rightarrow b)$  sont vrais, alors  $(b)$  est vrai. On remarque que la déduction est strictement incluse dans les prémisses. Dit autrement, les langages formels sont des mécaniques qui manquent cruellement de capacité créative.

Ce manque s'explique par deux raisons. D'une part, le principe originel de la logique est de raisonner justement – c'était le but recherché par Aristote. D'autre part, le principe d'explosion interdit toute incohérence syntaxique. Un langage formel ne peut donc produire que ce qui est vrai – c'est-à-dire ce qui est *déjà* vrai.

L'IA connexionniste n'a pas cette limite. En tant qu'outil statistique, elle admet une infinité d'états entre 0 (faux) et 1 (vrai). Et, plus largement, elle accepte une infinité d'états entre deux nombres  $x$  et  $y$  quels qu'ils soient pourvu qu'ils soient distincts. De surcroît, elle permet de généraliser des faits constatés, ce qui lui donne une forme de capacité d'induction. Elle peut donc, dans une certaine mesure, générer de la créativité depuis et entre plusieurs mondes possibles. Mais elle ne sait pas s'en écarter *perpendiculairement*.

Remarque : on parle parfois de raisonnement inductif (généralisation de cas particuliers) en mathématique à propos des raisonnements récursifs. Cela relève d'un abus de langage : le raisonnement récursif est entièrement déductif, la conclusion étant mécaniquement incluse dans les prémisses.

#### 6) *Les algorithmes des langages formels sont trop lents*

Les algorithmes de la logique classique sont de complexité exponentielle. Saturée une base contenant quelques centaines de formules logiques (c'est-à-dire calculer l'ensemble des formules déductibles) nécessite un temps de calcul de plusieurs secondes. C'est excessivement long. Or la saturation est nécessaire : un langage formel a besoin de s'assurer de l'existence ou non d'une incohérence syntaxique pour échapper au principe d'explosion.

### 7) *Et de nombreux autres arguments*

Les arguments que nous venons de présenter sont de complexités inégales. L'exercice ne se veut ni encyclopédique ni exhaustif. Nous renvoyons le lecteur qui souhaite creuser ces différents thèmes à la littérature, riche et foisonnante en la matière.

Et il est d'autres arguments, moins techniques et plus conceptuels, qui questionnent sur la possibilité même de l'IA. Nous ne les abordons pas, soit parce qu'ils relèvent encore du spéculatif compte tenu de l'état des recherches à ce jour, soit parce qu'ils nous feraient quitter le périmètre de ce document : nous souhaitons le cantonner aux seuls problèmes techniques, précisément identifiés et formalisés, rencontrés par l'IA symbolique.

### **Une proposition : intégrer la notion de pensée dans le langage**

Puisque notre quête porte sur l'IA forte, et que celle-ci est caractérisée par la notion d'émotion, l'approche que nous proposons d'emprunter est d'enrichir la syntaxe du langage formel par la notion de pensée – dans le sens défini par René Descartes (1641) : « *Par le nom de pensée, je comprends tout ce qui est tellement en nous que nous en sommes immédiatement connaissant ; ainsi toutes les opérations de la volonté, de l'entendement, de l'imagination et des sens sont des pensées* ».

Etudier les relations entre langage et pensée n'est pas une nouveauté, voir Oswald Ducrot (1971) par exemple. La question est ancienne, et nous n'avons pas la prétention de rentrer dans la réflexion. Notre démarche consiste à considérer comme vraie une des théories exprimées, et à analyser les conséquences de cette hypothèse sur les formalismes logiques.

Nous proposons de suivre Ludwig Wittgenstein (1921), en supposant qu'une pensée (notons-la  $c$ ), a une réalité en soit, originellement autonome du langage, et que la formule, la phrase (notons-la  $f$ ) qui l'exprime, n'a en elle-même aucune signification.  $f$  est un ensemble de signes qui ne dit rien, le sens n'est porté que par  $c$ . Cela conduit à devoir modéliser une relation entre ces deux signes. Nous proposons :

Soit  $L$  un langage muni de la règle d'interprétation sémantique  $\models$ . Soit  $c$  une pensée, et  $f$  la formule qui exprime  $c$  dans  $L$ . Alors  $c$  est une proposition atomique de  $L$ , et la relation dans  $L$  entre  $c$  et  $f$  est donnée par :  $c \models f$ .

Selon cet énoncé (appelons-le le postulat contextuel), l'agent ne mémorise ni la pensée  $c$  ni sa modélisation  $f$ , mais la *conscience* qu'il a d'exprimer  $c$  par  $f$ . Par conscience, nous traduisons le fait que cette relation est intériorisée dans la syntaxe du langage. Nous faisons par ailleurs les hypothèses que, pour toute phrase  $f$ , il existe une pensée  $c$  qui porte sa sémantique.

L'application de ce postulat à la logique propositionnelle conduit à l'élaboration d'un nouveau langage, appelé la logique contextuelle, et notée  $L_c$ . Elle conserve les règles syntaxiques de la logique classique et donc les trois principes d'Aristote. Son apport porte sur la sémantique, qui est une sémantique des croyances.  $L_c$  bénéficie de deux propriétés comportementales qui la singularisent des autres formalismes :

- une base de connaissances contextuelles est un ensemble de descriptions de pensées sous la forme  $c \models f$ . Dans  $L_c$ , un agent intelligent peut penser tout et n'importe quoi, la cohérence syntaxique est garantie. La difficulté est de déterminer, dans le puit des pensées, lesquelles mobiliser pour agir de la façon la plus adéquate face à un stimulus,
- l'adjonction d'une nouvelle connaissance dans une base contextuelle conserve l'ensemble des formules qui étaient précédemment produites, mais peut engendrer une modification de leur interprétation sémantique.  $L_c$  a donc la particularité d'être à la fois syntaxiquement monotone (ce qui est produit reste produit) et sémantiquement non monotone (ce qui était crédible à un instant  $t$  peut ne plus l'être à l'instant  $t+1$ ).

Nous ne présentons pas plus le formalisme de  $L_c$ . Le lecteur intéressé en trouvera une présentation détaillée dans hal-02120285. Ces deux propriétés sont nécessaires et suffisantes pour comprendre les propos du paragraphe suivant.

Elles pourraient perturber les lecteurs attachés aux concepts de l'adéquation (ce que produit le langage formel est vrai) et de la complétude (ce qui est vrai est produit par le langage formel). Les études en IA symbolique se sont jusqu'à présent attachées à reproduire les mécanismes de raisonnement. Ce n'est pas le but de  $L_c$ , qui, en appliquant  $c \models f$ , utilise le langage pour modéliser les pensées de celui qui s'exprime. En logique contextuelle, le raisonnement n'est pas l'objectif, il est une conséquence.

### **$L_c$ face aux difficultés rencontrées par l'IA symbolique**

#### *1) Le raisonnement humain ne suit pas la logique d'un raisonnement formel*

La question est de savoir s'il existe un langage fédérateur de l'ensemble des langages formels. Des ponts intéressants ont été établis entre la logique contextuelle et les logiques paracohérentes, la logique des défauts et les logiques modales (voir Arnaud Kohler, 1995). Il est possible d'étendre les capacités d'expression des formalismes classiques en respectant la syntaxe et les principes d'Aristote. Il serait donc prématuré de clôturer le dossier à date.

#### *2) Les langages formels n'acceptent pas l'incohérence*

Contournant les principes d'adéquation et de complétude,  $L_c$  maintient la cohérence syntaxique tout en autorisant dans le même temps l'incohérence sémantique : une croyance peut être simultanément vraie relativement à un contexte et fausse dans un autre. En fait, la possibilité de modéliser une incohérence (du type  $a$  est à la fois crédible et non crédible) est nécessaire au processus d'interprétation de  $L_c$ . Elle lui permet notamment de modéliser et d'exploiter les informations dites « avec exception » : une règle peut être vraie *par défaut*, tant qu'elle n'est pas soumise à une situation dans laquelle elle s'avère explicitement fausse.

#### *3) Les langages formels ne savent pas gérer des nouvelles informations*

Dans  $L_c$ , une nouvelle information prend une forme  $c \models f$ , qui est descriptive (elle dit que  $f$  décrit  $c$ ) et non sémantiquement valorisante (elle ne dit rien sur la validité de  $c$  ou de  $f$ ). Dans ce cadre, la révision des croyances devient un faux problème : une nouvelle connaissance est simplement enregistrée comme la description d'une nouvelle pensée. Nous retrouvons ici l'argument d'ontologie présenté contre la logique par le sophiste Gorgias dans son Traité du non-être (vers 440 avant J.C.) : « *Ce n'est pas l'être qui est l'objet de nos pensées* ». Dans  $L_c$ , le langage a sa propre loi, qui n'est pas celle des faits. La relation est portée par la pensée.

La difficulté rencontrée avec une base de connaissances contextuelles porte sur son exploitation, plus précisément sur la fonction de sélection des pensées pertinentes face à un stimulus. La logique contextuelle apporte une réponse sémantique à la difficulté soulevée par les paradoxes. S'il n'y a pas de possibilité de choix entre deux pensées contradictoires, c'est tout simplement parce qu'elles sont peut-être vraies (ou fausses) l'une et l'autre – ou peut-être parce que les connaissances en l'état sont insuffisantes pour décider. La sémantique de  $L_c$  propose les quatre possibilités (non interprétable, crédible, non crédible, ou les 2).

#### 4) *Les langages formels ont une sémantique trop stricte*

$L_c$  identifie les pensées comme étant le lieu du sens. Pour plagier Ludwig Wittgenstein (1921), la pensée « *a se trouve dans une certaine relation  $R$  avec  $b$*  » dit  $aRb$  – mais pris isolément,  $aRb$  est un ensemble de signes qui ne dit rien.  $L_c$  distingue donc les pensées (qui portent le lien entre le système intelligent et les faits, supposés ou réels, qu'il ressent) et les formules (qui portent son moyen d'expression).

Cette dichotomie génère une distinction entre la communication, les émotions, les ressentis et les actions. Ils forment l'ensemble des propositions atomiques du formalisme, qui modélise et exploite des relations logiques entre les différents objets. Le système intelligent qui en découle ne réagit pas uniquement par rapport au contenu rationnel de ses pensées. Ses actions sont aussi dépendantes des émotions auxquelles elles sont associées – et inversement.

D'autre part, à un instant  $t$ , l'ensemble des croyances d'un agent intelligent n'inclue pas systématiquement toutes les croyances que cet ensemble produit syntaxiquement. Une base contenant deux pensées  $A \models f$  et  $B \models g$  permet de croire  $f$  et  $g$ , et de ne pas croire  $h$  même si

$(f \wedge g)$  produit  $h$ . L'interprétation sémantique peut être différente lorsque la base contient la pensée  $C \models f \wedge g$ .  $L_c$  est donc sensible au contexte et à l'apprentissage.

La fonction d'interprétation sémantique de  $L_c$  est définie par une formule mathématique et stricte. Elle est toutefois assujettie à la notion de contexte (des ensembles de pensées). Cela lui apporte un niveau de valorisation qui quitte la dimension du {vrai, faux} pour, adossée aux méta-pensées (Jacques Pitrat, 1990), aller au-delà de la multi valuation. En fait, la notion d'incertitude est totalement intrinsèque et nécessaire au formalisme contextuel, qui est un langage faillibiliste et perspectiviste (voir hal-02120285).

#### 5) *Les langages formels interdisent toute créativité*

Une base de connaissances contextuelles est toujours syntaxiquement cohérente. De ce fait, contrairement aux langages formels classiques,  $L_c$  autorise la production de n'importe quelle connaissance, en autorisant toutes les pensées, même totalement folles, voire totalement incohérentes. Cela lui donne la capacité d'accepter la créativité : il « suffit » d'étendre le formalisme, en lui intégrant une règle de production créative. Cette notion reste à définir.

#### 6) *Les langages formels ont des algorithmes trop lents*

La lenteur des algorithmes des formalismes classiques engendre 2 problèmes principalement :

- la gestion de l'incohérence syntaxique. Ce n'est plus un problème, puisqu'une base contextuelle est toujours cohérente,
- le besoin de complétude du raisonnement, dans le but d'obtenir la réponse la plus juste compte tenu des informations disponibles. Ce n'est pas une réalité constatée chez l'être humain, qui commet constamment des erreurs de raisonnement. Cela n'est donc pas un pré requis nécessaire à l'IA.

### **Perspectives**

Des études de comparaisons techniques entre  $L_c$  et d'autres formalismes non classiques ont été réalisées et présentées dans divers documents déjà cités. Nous ne nous y attarderons donc

pas, en invitant le lecteur intéressé à s'en rapprocher. Nous proposons ci-dessous de parcourir rapidement quelques thèmes généraux que la logique contextuelle interroge.

### La sémantique générale

Le postulat contextuel et la sémantique générale proposée par Alfred Korzybski partagent un même constat initial : « *Nous vivons souvent en fonction des représentations des « faits » autant qu'en fonction des « faits ».* Avant de pouvoir agir, nous devons avoir une certaine forme de représentation dans notre cerveau. Avant de pouvoir vous asseoir sur une chaise, vous devez avoir une certaine forme de représentation de cette chaise dans votre cerveau et votre système nerveux. Vous agissez en fonction de ces représentations dans votre système nerveux. » - A. Korzybski, Séminaire de sémantique générale de 1937.

Face à ce constat, la sémantique générale conclue à une insuffisance de l'approche aristotélicienne.  $L_c$  adopte une démarche plus prudente : puisque qu'un agent intelligent a conscience de la distinction entre un fait et la représentation qu'il en a, sa proposition est d'intégrer ce constat dans le formalisme en tant que postulat complémentaire, et d'étudier l'impact de cette intégration. La conséquence est l'évolution d'un langage, initialement construit pour raisonner, vers un langage qui priorise la communication. Le postulat contextuel fait alors écho au mentalais (*le langage de la pensée*, thèse développée par Jerry Fodor en 1975), en concluant que la modélisation du raisonnement est une conséquence de la modélisation de la pensée.

### En mathématique

Un formalisme classique contextualisé respecte les trois principes d'Aristote, et est pourtant capable de modéliser et d'exploiter des informations aléthiques, incomplètes ou incohérentes. La contextualisation d'un langage du  $n^{\text{ième}}$  ordre sera une étape nécessaire pour enrichir l'expressivité du langage, et tester ses solutions de modélisation sur des cas plus complexes et représentatifs du raisonnement humain. A contrario de l'affirmation sur la trousse de bricoleur, la logique contextuelle ouvre des pistes de réflexions sur l'existence d'un formalisme fédérateur de l'ensemble des langages logiques.

### En philosophie

En assujettissant toute interprétation sémantique à un contexte,  $L_c$  contourne le dogme de l'adéquation et de la complétude. Le formalisme obtenu présente la singularité comportementale d'être dans l'incapacité de produire une certitude objective. Cela le contraint à adopter une approche faillibiliste (l'agent intelligent croit à l'instant  $t$  ce qui porte une justification à cet instant) et perspectiviste (à chaque instant, la croyance d'un agent intelligent se compose d'une somme de perspectives qu'elle a sur elle-même).

Ces caractéristiques enrichissent les réflexions sur les relations entre pensée et langage, et questionnent l'idée selon laquelle limiter la pensée à ce qui est logique l'appauvrirait.

### En sciences cognitives

En supposant l'ensemble exhaustif des connaissances, l'application du postulat contextuel conduit à considérer que toutes les pensées sont possiblement fausses, et que toute autre proposition élémentaire est improbable. Face à lui-même, un système cognitif appliquant un raisonnement contextuel est, dans l'absolu, condamné à douter de tout.

Les sciences cognitives proposent une réponse à ce problème. Elles distinguent la Mémoire à Long Terme (la MLT, contenant l'ensemble des connaissances) et la Mémoire à Court Terme (la MCT, contenant les connaissances utilisées à un instant  $t$  pour réagir à un stimulus). Des seuils cognitifs ont été identifiés. Ils bordent les connaissances utilisées par la MCT, en intégrant une fonction de sélection « pragmatique » des connaissances (privilégiant les pensées les plus récentes, par exemple). Le raisonnement ne se fait donc pas sur l'ensemble des connaissances, mais sur un sous-ensemble pré sélectionné par une mécanique qui lui est autonome.

### En Intelligence Artificielle

A la recherche de l'IA forte, la voie de recherche privilégiée actuellement consiste à tenter de simuler l'intelligence humaine au travers d'outils de plus en plus performants. Les limites calculatoires semblant atteintes, de grands espoirs sont fondés sur l'avènement, espéré proche, des calculateurs quantiques. Ils démultiplieront le nombre de données traitables. La

complexité calculatoire, associé au caractère aléatoire de la physique quantique, rendra impossible toute traçabilité.

Cette approche suppose que l'intelligence sera atteinte dès l'instant où son contraire ne pourra pas être démontré. On retrouve ici le « jeu de l'imitation » proposé en 1950 par Alan Turing. Il suggérait qu'un système est considéré comme intelligent à partir du moment où il devient impossible, en dialoguant avec lui (par l'intermédiaire d'un terminal), de discriminer si c'est une machine ou un être humain qui répond.

La logique contextuelle propose une approche complètement orthogonale. Elle apporte au système intelligent un formalisme qui modélise une notion de conscience, ce qui lui permet de raisonner par introspection. Il faut toutefois éviter le piège de l'anthropomorphisme ou celui du mythe de Frankenstein, en définissant ce qu'est une pensée ou une émotion dans le cadre de l'IA, et les réflexes auxquelles elle doit être associée : les réponses sont directement dépendantes du corporel de l'agent intelligent. Les émotions et les motivations que ressentira une IA forte seront, selon toute vraisemblance, très différentes de celle d'un être humain.

## **Conclusion**

Face aux difficultés rencontrées par les logiques classiques,  $L_c$  suppose que les propositions élémentaires d'un langage formel sont toutes égales face à la syntaxe, mais sont inégales face à la sémantique. Le formalisme obtenu respecte les trois principes d'Aristote, et est pourtant capable d'exploiter des informations aléthiques, incomplètes ou incohérentes, passant ainsi les premières difficultés face auxquelles l'IA symbolique a échoué. Il convient toutefois de rester prudent. L'étude a été essentiellement théorique. La confrontation de  $L_c$  aux constats empiriques fera émerger de nouvelles questions.

## **Références bibliographiques**

- ANDLER, Daniel, *Introduction aux sciences cognitives*, Gallimard, 2004
- CARDON, Dominique, COINTET Jean-Philippe, MAZIERES Antoine, *La revanche des neurones. L'invention des machines inductives et la controverse de l'intelligence artificielle*, Réseaux 2018/5 (n° 211), 2018

- DESCARTES, René, *Méditations métaphysiques. Objections et Réponses (I à VI)*, édition sous la direction de Jean-Marie Beyssade et de Denis Kambouchner, éd. Gallimard Tel (2018), 1641
- DUCROT, Oswald, *Langue et pensée formelle*, Langue française, n°12, Linguistique et mathématique, sous la direction de M.-C. Barbault et Oswald Ducrot, 1971
- FAUBLEE, Jacques, *L'origine du mot malgache harena*, dans *Tradition et dynamique sociale à Madagascar*, Collection IDERIC, 1978/9, 1978
- GÄRDENFORS, Peter & MAKINSON, David, *Relations between the logic of theory change and nonmonotonic logic*, in André Fuhrmann and Michael Morreau, eds. *The Logic of Theory Change*, LNAI-465, 1991
- KOHLER, Arnaud, *Proposition d'une structure de représentation de la connaissance pour les raisonnements non classiques*, thèse de doctorat en informatique, Université d'Aix-Marseille I, France, 1995
- KOHLER, Arnaud, *Les cultures et les langages formels*, LTML 2019, référence hal-02320997, 2019
- KOHLER, Arnaud, *La logique contextuelle*, CCSD HAL, référence hal-02120285, 2019
- KRIPKE, Saul Aaron, *Semantical considerations on modal logic, Reference and Modality*, L. linsky (éditeur), Oxford University Press, London, 1971
- LUKASZEWICZ, Jan, *Many-Valued Systems of Propositional Logic*, eds. S. Mc Call, *polish Logic*, Oxford University Press, Oxford, 1967
- MC CARTHY, John, *Circumscription - A Form of Non Monotonic Reasoning*, *Artificial Intelligence* 13, 1980
- NORMAN, Donald, *Models of human memory*, Academic Press, New York, 1970
- PITRAT, Jacques, *Méta-connaissance, Futur de l'Intelligence Artificielle*, Hermes, 1990
- RASTIER, François, *Sémiotique et science de la culture*, *Linx*, 44 | 2001, 2001
- REITER, Ray, *A logic for default reasoning*, *Artificial Intelligence* 13, 1980
- ROULIN, Jean-Luc, *Psychologie cognitive*, rédaction sous la coordination de Jean-Luc Roulin, édition Bréal, 2006
- SAUVAYRE, Romy, *Croire à l'incroyable*, Presses Universitaires de France, 2012
- THAYSE, André & co-auteurs, *Approche logique de l'intelligence artificielle*, Dunod informatique, 1993

- VAN HAECHT, Louis, *Le problème de l'origine du langage*, Revue Philosophique de Louvain, 1947
- VILLANI, Cédric & co-auteurs, *Donner un sens à l'Intelligence Artificielle – Pour une stratégie nationale et européenne*, mission parlementaire du 8 septembre 2017 au 8 mars 2018 confiée par le premier ministre Edouard Philippe, 2018
- WITTGENSTEIN, Ludwig, *Tractatus logico-philosophicus*, traduction Gilles Gaston Granger, éd. Gallimard Tel (1993), 1921