



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-02444894v1

HAL Id: hal-02444894

<https://hal.science/hal-02444894v1>

Preprint submitted on 19 Jan 2020 (v1), last revised 18 Dec 2020 (v2)

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

Entre raisonnement, conscience et émotion

Arnaud Kohler

Résumé Les outils issus à ce jour de l'IA symbolique et de l'IA connexionniste relèvent de l'IA faible. Il manque aux intelligences artificielles la jonction entre raisonnement (logique ou statistique), conscience et émotion. C'est ce que je propose de combler, en modélisant la notion de pensée dans le langage formel.

Abstract The tools to date from symbolic AI and connectionist AI fall under weak AI. Artificial intelligence lacks the junction between reasoning (logic or statistics), consciousness and emotion. This is what I propose to fill, by modeling the notion of thought in formal language.

Préambule

Après avoir vécu une période glorieuse, l'Intelligence Artificielle symbolique est entrée, dans le courant des années 1990, dans un hiver gris et froid. Ses chercheurs, autrefois insolents, se cachent depuis et sont à leur tour 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).

La condamnation est rude, et se veut sans appel : il y a injonction d'aveu. Les logiciens qui reconnaissent leur faute sont pardonnés, et intègrent en récompense le camp des *vraiment modernes*. Ils laissent derrière eux les têtus irréductibles, qui n'ont pas compris que le temps est venu de *La revanche des neurones* (Dominique Cardon, Jean-Philippe Cointet et Antoine Mazières, 2018).

L'IA symbolique propose d'exploiter les données par l'application de règles logiques, exprimant des relations entre les objets. Elle affirmait la venue des robots autonomes, capables de remplacer l'humain dans l'ensemble de ses tâches les plus complexes. Péchés d'orgueil. On retient essentiellement de l'aventure un ensemble de promesses non tenues. Elle n'a produit « que » les systèmes experts, utilisés pour quelques applications tangibles, mais

circonscrites et sans reconnaissance publique. Ils se dissimulent aujourd'hui derrière le sigle barbare de BRMS, pour *Business Rules Management System*.

L'IA connexionniste (ou neuronale) exploite les données par l'outil statistique. 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... Ses réussites sont flamboyantes, et elle bénéficie depuis environ 10 ans d'une communication d'envergure et d'un soutien financier conséquent.

Cet enthousiasme devrait toutefois être tempéré, sous peine de reproduire la même faute que les logiciens *non modernes*. En effet, toutes les applications réalisées à ce jour, qu'elles soient symbolistes ou connexionnistes, ont une limite commune : elles relèvent de l'IA dite faible. Les systèmes réagissent et répondent, dans le contexte environnemental pour lesquels ils ont été conçus, aux questions qui leur sont posées. Ils reproduisent mécaniquement les processus plus ou moins complexes qui leur ont été inculqués, s'adaptant à ce pour quoi il était prévu qu'ils s'adaptent.

La statistique apporte une capacité d'induction (généralisation d'un cas particulier) qui manque à la logique, mais avec des limites d'application. Une forte probabilité d'occurrence ne fait pas une vérité, et peut générer des erreurs d'interprétation, voire de décision. Par exemple, avec les outils actuels, l'absorption par une IA connexionniste de plusieurs milliers d'images de chat ne garantit toujours pas la reconnaissance d'un chat par la machine avec une certitude de 100%. Dans le même temps, l'être humain obtient des résultats au moins équivalents à partir de quelques photos – ou de quelques *rencontres* devrait-on peut-être plutôt dire.

En fait, l'ensemble de ces systèmes relève de la très belle orfèvrerie, mais avec les limites intrinsèques à un système mécanique. Ils ne prognoient pas, ils ne se posent pas de questions. En bref, ils ne bénéficient pas, dans leurs processus d'interaction et de décision, de la stimulation qu'apportent les émotions. Ils n'ont pas la conscience de l'action qui incite à l'action. Cela, ce sera l'IA forte, appelée aussi parfois IA générale. Elle sera symbolique, connexionniste, ou les deux. Ou autre chose peut-être, la solution reste à découvrir. Certains prédisent qu'elle sera quantique. Ils font le pari de la puissance, mais l'expérience montre que

la réponse n'est pas toujours dans l'usage de la force. Parfois, plus simplement, la réponse adéquate n'est pas encore imaginée.

Dans ce cadre, je vais tenter de montrer que l'IA symbolique a encore à apporter à la réflexion. Sa condamnation a été trop hâtive. De nombreux sentiers de recherche sont toujours à explorer. Je vous propose de parcourir l'un d'entre eux.

Je commencerai par présenter quelques défis techniques qui ont mis l'IA symbolique en échec. J'emprunterai ensuite une nouvelle piste de réflexion. Elle consiste à intégrer la notion de pensée dans le langage. L'exercice apporte au formalisme (la logique contextuelle, notée L_c) deux propriétés qui le singularisent fondamentalement des autres langages formels : il bénéficie d'une syntaxe monotone et d'une sémantique non monotone d'une part, et il supporte la production de pensées créatives d'autre part. L'IA symbolique sera revisitée au travers de ces propriétés.

Dans la suite, j'éviterai autant que possible l'emploi de formules dans ma rédaction. Mais lorsqu'il me semblera que le signe remplace avantageusement un long discours, j'utiliserai, dans leur sens classique, 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 – c'est-à-dire mécaniquement – de a ») et d'interprétation sémantique ($a \models b$, pour « b hérite d'une valorisation sémantique de a »).

Les limites techniques qui ont mis l'IA symbolique en échec

Il est un avis communément 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. 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. Je vous propose de visiter les 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 logiques : 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 $20^{\text{ième}}$ 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, l'explication la plus souvent avancée est que le raisonnement humain ne suit pas les mécaniques d'un langage formel, mais d'un autre mode – dont chaque langage formel connu ne serait qu'une approche partielle et simplificatrice. En fait, l'intelligence humaine relèverait d'une *règle d'inférence non démonstrative* (voir par exemple Jean-Luc Roulin, 2006), qui échapperait à la logique formelle et qui reste à définir.

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). Le refus de l'incohérence est donc inscrit dans les gènes des langages formels. De nombreuses tentatives ont été faites pour y échapper (voir les logiques para-cohérentes, multivaluées ou floues par exemple), mais sans succès définitif.

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

Le problème d'un langage formel est que la présence d'une incohérence syntaxique lui permet d'inférer très exactement tout et son contraire : depuis $a \wedge \neg a$, je peux produire b quel que soit b . C'est la propriété dite d'explosion, et elle n'est évidemment pas acceptable.

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

L'appropriation et l'exploitation de nouvelles informations relèvent d'une mécanique naturelle à l'être humain. Ceux sont des 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. En IA symbolique, les nombreuses études en la matière sont portées par la discipline connue sous le nom de *Révision des Croyances* (la traduction anglaise *Belief Revision* est plus fréquemment utilisée). Aucune n'a abouti.

L'énoncé du problème est pourtant d'une simplicité confondante : face à deux informations incohérentes, laquelle retenir ? L'être humain doit faire face tous les jours, et à de multiples occasions, à des situations de ce genre. Comme vu au paragraphe précédent, cette situation paralyse un système logique, et il n'a pas d'autre choix que d'y remédier. La solution a été recherchée au travers d'un nouveau concept, celui de *l'enracinement épistémique* (voir par exemple Romy Sauvayre, 2012).

Il suppose l'existence d'une sorte de relation d'ordres entre les connaissances. Propre à chaque individu selon Peter Gärdenfors et David Makinson (1991), elle permettrait de sélectionner les informations en fonction de leurs niveaux de pertinences. Cette solution n'est malheureusement pas satisfaisante : si l'enracinement épistémique relève d'un formalisme, le problème boucle sur lui-même : que faire lorsque deux informations incohérentes ont le même niveau de pertinence ?

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 fonction 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 l'impossibilité d'obtenir l'exhaustivité des formules qui décriraient *le monde* dans sa complétude.

Imaginer la base des connaissances formelles qui décrivent 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, et c'est notamment de là qu'il tire sa richesse et sa puissance d'expressivité.

5) *Les langages formels interdisent toute créativité*

Les langages formels n'ont qu'une seule règle de production, qui est le *modus ponens*. C'est une 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.

Cela s'explique par deux raisons. D'une part, le principe originel de la logique est de raisonner justement et de ne produire que des vérités – c'était le but recherché par Aristote. D'autre part, et c'est une conséquence immédiate, un langage formel ne peut pas prendre le risque d'accepter une règle qui, potentiellement, peut générer de l'incohérence (voir l'argument 2).

L'IA connexionniste n'a pas la même limite. En tant qu'outil statistique, elle permet de généraliser des faits constatés, ce qui lui donne une forme de capacité d'induction.

Remarque : on parle parfois de raisonnement inductif (généralisation de cas particuliers) en mathématique à propos des raisonnements récursifs. Selon plusieurs auteurs, dont je partage l'avis, 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 de formules logiques (c'est-à-dire calculer l'ensemble des formules déductibles) nécessite un temps de calcul très long. Or la saturation est nécessaire, par exemple pour l'opération de révision. Un langage formel a besoin de s'assurer de l'existence ou non d'une incohérence syntaxique (voir l'argument 2). Ne pas le faire, c'est prendre le risque d'avoir des informations incohérentes dans la base, avec les risques d'explosion induits.

7) *Et de nombreux autres arguments*

Les arguments que je viens de présenter rapidement sont de complexités inégales. L'exercice ne se veut ni encyclopédique ni exhaustif. Je renvoie 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. Je ne les aborde 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 cet article : je souhaite le cantonner aux seuls problèmes techniques, précisément identifiés et formalisés, rencontrés par l'IA symbolique.

La logique contextuelle, ou l'intégration de la pensée dans le langage

De ces constats, il ressort un sentiment de doute, voir pour les plus pessimistes la certitude d'un échec annoncé. Mais ce serait une erreur d'abandonner aussi rapidement. Échec ne vaut pas preuve, et il est prudent, dans une démarche de recherche fondamentale, de maintenir l'ouvrage sur le métier tant que la démonstration définitive n'est pas faite. Après tout, les réflexions autour de l'IA n'ont qu'une petite centaine d'année, ce qui est extrêmement court.

Puisque la quête porte sur l'IA forte, et que celle-ci est caractérisée par la notion d'émotion, l'approche que je propose de tester 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 ne relève pas de la nouveauté, voir Oswald Ducrot (1971) par exemple. La question philosophique est très ancienne. Je ne rentrerai pas dans la réflexion. Mon approche consiste, en prenant le parti d'une pensée supposée réalité en soit, originellement autonome du langage, d'analyser les conséquences de cette hypothèse sur les formalismes logiques.

Je suppose donc que, quel que soit le formalisme utilisé, une formule, une phrase, notée *f*, n'a en elle-même aucune signification : c'est un ensemble de signes qui ne dit rien. Le sens n'est

porté que par la pensée, symbolisée par c , que f décrit. Cela conduit à devoir modéliser une relation entre c et f . Je propose :

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, 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, j'entends que cette relation est intériorisée dans la syntaxe du langage. Je fais par ailleurs les hypothèses que, pour toute phrase f , il existe une pensée c qui porte sa sémantique, et que c est un signe qui ne se dit pas.

L'application de ces assertions à la logique propositionnelle conduit à l'élaboration d'un nouveau langage, appelé la logique contextuelle, et noté L_c . Il conserve les règles syntaxiques de la logique classique et le respect des trois principes d'Aristote. Son apport porte sur la sémantique, qui est une sémantique des croyances. La fonction d'interprétation de L_c :

- exprime qu'une croyance (c'est-à-dire une formule, ou une phrase) peut prendre le statut sémantique « non connu ». Dans le cas contraire, elle est crédible ou non crédible, et uniquement relativement à un ensemble de pensées (appelé un contexte),
- autorise d'interpréter simultanément selon plusieurs contextes différents (les contextes d'interprétation) éventuellement incohérents.

Le langage formel L_c bénéficie de deux propriétés comportementales qui le singularisent fondamentalement des autres formalismes :

- une base de connaissances contextuelles est un catalogue de description de pensées sous la forme $c \models f$. Elle est toujours syntaxiquement cohérente : un agent intelligent peut penser tout et n'importe quoi. La difficulté est alors de déterminer quelles pensées mobiliser et interpréter pour agir de la façon la plus adéquate : selon L_c , la réponse à l'IA n'est pas dans la cohérence syntaxique ou sémantique des pensées, mais dans la manière dont elles sont exploitées.

- 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 monotone (ce qui est produit reste produit) et non monotone (ce qui est crédible à un instant t peut ne plus l'être à l'instant $t+1$).

Je ne présenterai 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 formés aux concepts très cartésiens 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).

Cela provient, selon moi, d'un quiproquo. L'origine du langage fait toujours l'objet de nombreuses réflexions (voir par exemple Louis van Haecht, 1947). J'adopte pour ma part l'idée selon laquelle la raison d'être d'un langage *naturel* n'est pas de simuler un raisonnement juste pour produire des vérités, mais, plus prosaïquement, d'exprimer les pensées – les émotions – de celui qui l'emploie. Si le but qu'on vise est d'utiliser un langage formel comme moyen d'expression d'une intelligence artificielle, alors il est primordial qu'il retrouve intrinsèquement cette utilité. C'est ce que modélise l'expression $c \models f$, qui peut se comprendre comme : « f exprime ma pensée c ».

Les réponses de L_c aux difficultés techniques rencontrées par l'IA symbolique

Je propose de reprendre ici les arguments techniques considérés jusqu'à présent comme rédhibitoire à l'IA symbolique.

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

Nous avons vu précédemment qu'une question est de savoir s'il existe un langage fédérateur de l'ensemble des langages logiques. La logique contextuelle se pose en candidate. La preuve n'est pas finalisée, mais des ponts intéressants ont été établis notamment avec les logiques para-cohérentes, les logiques des défauts et les logiques modales (voir Arnaud Kohler, 1995). Et quand bien même le résultat serait négatif, L_c apporte la preuve qu'il est possible d'étendre

les capacités d'expression des formalismes sans enrichissement de la syntaxe classique et en préservant les principes d'Aristote. Il y a donc encore des choses à étudier et à découvrir.

Le logicien qui se lancerait dans cette étude se rendrait rapidement compte que la solution se trouve dans la définition des contextes d'interprétation : pour chaque logique classique, il existe une fonction de sélection des pensées « crédibles » qui capture sa sémantique. La logique contextuelle est donc une mécanique qui capture différents modes de raisonnement. Mais, seule, elle ne dit rien : elle n'apporte pas de solution pour décider, c'est-à-dire pour identifier la bonne fonction à appliquer.

On retombe donc sur un problème de choix multiple, qui accrédite l'idée de l'inférence non démonstrative. Elle échappe à l'IA symbolique. Je rejoins de nombreuses voix qui affirment aujourd'hui que la réponse à l'IA forte se trouvera dans la fusion de plusieurs disciplines. Les sciences cognitives, et plus précisément les études sur les modèles cognitifs, proposent des éléments de réponse à cette question de fonction de décision (voir hal-02120285). L'IA connexionniste pourrait dans ce cadre apporter sa capacité à traiter le très grand volume des pensées.

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

Nous l'avons vu, le sujet de l'incohérence syntaxique disparaît dans L_c . 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 vrai et faux) 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 logique peut être vraie *par défaut*, tant qu'elle n'est pas face à une situation pour laquelle elle est explicitement fausse.

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

Dans L_c , une base de connaissances est toujours syntaxiquement cohérente : la nouvelle information prend une forme $c \models f$, qui est descriptive (elle dit que f décrit c) et non sémantiquement valorisante (elle ne se prononce pas sur la validité de c ou de f). Dans ce cadre, la révision des croyances devient un faux problème : la nouvelle connaissance est simplement enregistrée comme la description d'une nouvelle pensée.

L'étape suivante porte sur l'exploitation de cette nouvelle information, c'est-à-dire sur la sélection des pensées pertinentes permettant de réagir au stimulus. La logique contextuelle adopte le principe de l'enracinement épistémique, en permettant, comme les logiques modales, de modéliser des connaissances aléthiques. Cependant, elle apporte une réponse à la difficulté originelle : un système contextuel est capable de supporter un paradoxe. S'il n'y a pas de possibilité de choix entre 2 pensées contradictoires, c'est tout simplement parce qu'on ne le sait pas, voire parce qu'elles sont peut-être vraies (ou fausses) l'une et l'autre.

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 – pris isolément, aRb est un ensemble de signes qui ne dit rien. L'ensemble des propositions élémentaires de L_c contient les pensées et les propositions atomiques du langage. L_c distingue :

- les propositions atomiques, c'est-à-dire les mots élémentaires du langage *qui se disent*,
- les pensées atomiques, c'est-à-dire les mots élémentaires *qui ne se disent pas* et s'expriment par des propositions atomiques,
- les pensées élémentaires, qui s'expriment par des relations logiques (des formules composées, ou des phrases) entre les propositions atomiques,
- et les méta-pensées, qui s'expriment par des relations logiques entre les pensées et les propositions atomiques. Cette définition permet d'établir une relation d'ordre entre les pensées (Arnaud Kohler, 1995).

Ces distinctions permettent d'associer une proposition atomique (ou une formule) à plusieurs pensées différentes – et donc une proposition atomique à plusieurs sens, puisque la sémantique est portée par les pensées.

Pratiquement, la modélisation $c \models f$ dissocie la pensée des proposition atomiques. Cela autorise les combinaisons les plus complexes entre les propositions atomiques. Celles-ci

peuvent identifier des émotions – qui peuvent alors être associées à des mécanismes de réactions : avec L_c , il devient possible, sans risque d'intégrité syntaxique, de concevoir un système qui réagit non pas par rapport au contenu rationnel de ses pensées, mais par rapport aux émotions qu'elles génèrent.

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 $C \models f \wedge g$. Elle est donc sensible à la modélisation et à l'historique de la connaissance.

La fonction d'interprétation sémantique de L_c est mathématique et stricte. Elle est toutefois assujettie à la notion de contexte (les 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), rejoindre le domaine de la multi valuation (voir hal-02320997) : la notion d'incertitude est intrinsèque au formalisme contextuel, et il modélise des niveaux d'incertitudes par l'intermédiaire des méta-pensées.

5) *Les langages formels interdisent toute créativité*

Rappelons-nous qu'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 les plus incohérentes. Cela lui ouvre la porte de la créativité. Pour rendre L_c créatrice, il « suffit donc » de l'étendre, en lui intégrant une règle de production créative. Cette notion reste à définir. J'en ai proposé une (voir hal-02120285), il faudrait tester d'autres possibilités.

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

La lenteur des algorithmes des formalismes classiques engendre différents problèmes :

- la gestion de l'incohérence syntaxique. Ce n'est plus un problème,

- 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 ne me semble donc pas un pré requis nécessaire à l'IA.

L'IA forte : pour une approche pluri disciplinaire

Comme nous venons de le constater, la logique contextuelle propose des solutions à différents points bloquants rencontrés jusqu'à présent en IA symbolique. Mais elle n'est pas la solution : en fermant des sujets, elle en ouvre d'autres. Pour avancer, il me semble nécessaire de continuer la réflexion en réunissant plusieurs disciplines, et *a minima* :

- la logique. Elle apporte les méthodologies de modélisation et de raisonnement. L_c propose des solutions pour exploiter les informations incohérentes, aléthiques, incomplètes ou incertaines. Mais il reste encore au moins un sujet majeur à traiter, qui est celui de la modélisation de la temporalité. Il a été approché par les logiques modales. L'enrichissement de la syntaxe est éventuellement une solution, ce n'est toutefois pas évident. La modalité permet de modéliser les propriétés du temps, elle n'est peut-être pas la mieux adaptée pour exprimer la pensée du temps,
- les sciences cognitives. Elles apportent, par l'intermédiaire notamment des modèles cognitifs, des éclairages sur deux notions fondamentales au mécanisme de décision : la fonction de sélection du contexte d'interprétation d'une part, et de l'enracinement épistémique d'autre part. Par ailleurs, la logique classique est mono pensée (Arnaud Kohler, 1995). L_c est multi pensées, et suppose que la pensée précède le langage (et que le langage enrichit la pensée). Son application nécessite une juste définition de la notion de pensée. Par exemples : une formule produite par déduction est-elle une nouvelle pensée ? Une nouvelle information est-elle une pensée en tout ou un ensemble de pensées ? Les réponses impacteront fortement le résultat de l'interprétation sémantique. Elles pourraient peut-être s'obtenir par tâtonnement (en testant différentes solutions). Une réflexion cognitive permettra de mieux diriger les essais,

- la psychologie. L_c apporte une solution pour intégrer les émotions dans le langage. Il nous faut maintenant définir ce qu'est une émotion, et qu'elles sont leurs impacts sur l'agent intelligent - un exemple trivial : la peur entraîne automatiquement le réflexe de tenter de sortir de la situation en cours. La psychologie apporte la connaissance des émotions. Celles-ci sont très liées au corporel et aux motivations primaires de l'agent intelligent. Il conviendra donc d'éviter le piège de l'anthropomorphisme ou celui du mythe de Frankenstein, et de définir les émotions de l'IA par rapport à ce qu'il est,
- l'ingénierie robotique : l'IA a besoin d'un corps, pour passer de la raison à la sensation et entrer dans le domaine de la pro-action,
- l'IA connexionniste. L'IA symbolique a toute sa place pour exploiter les relations logiques entre les objets, mais elle n'est manifestement pas adaptée pour certaines opérations, la préanalyse des sons et des images par exemple, et plus généralement de tout ce qui relève des sens. L'IA connexionniste apporte les outils d'analyses de données en masse nécessaires à ce type d'opération. Elle me semble par ailleurs une candidate intéressante pour modéliser la fonction de sélection des contextes d'interprétation, qui nécessite la manipulation d'un volume conséquent de pensées,
- et, évidemment, l'informatique, qui apporte l'outil de mise en œuvre.

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
- WITTGENSTEIN, Ludwig, *Tractatus logico-philosophicus*, traduction Gilles Gaston Granger, éd. Gallimard Tel (1993), 1921