Ingénierie des connaissances et critique de la raison artificielle.

Gunnar Declerck, Jean Charlet

▶ **To cite this version:**

Gunnar Declerck, Jean Charlet. Ingénierie des connaissances et critique de la raison artificielle. : De la simulation de l'intelligence à sa supplémentation. Atelier Philosophie et Ingénierie. Le formel face à l'histoire, la technologie et la matérialité (IC2011), May 2011, Chambéry, France. pp.2. hal-00610788

HAL Id: hal-00610788

https://paris1.hal.science/hal-00610788

Submitted on 25 Jul 2011

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.

Ingénierie des connaissances et critique de la raison artificielle. De la simulation de l'intelligence à sa supplémentation

Gunnar Declerck[1], Jean Charlet [1]

[1] Laboratoire de Santé Publique et Informatique Médicale, INSERM
gunnar.declerck@crc.jussieu.fr

L'Ingénierie des Connaissances (IC) peut-elle, comme prétend parfois le faire l'Intelligence Artificielle (IA), nous renseigner sur les mécanismes de la connaissance ? Pour répondre à cette question, nous partirons des analyses que H.L. Dreyfus a consacrées à l'IA, en particulier dans son ouvrage *What computers can't do*, en essayant de déterminer si l'IC est en mesure de contribuer à la critique de la raison artificielle qu'il tenta de mettre en place. Nous chercherons à étayer l'idée que l'IC, bien que n'ayant pas comme l'IA pour projet de *reproduire* le comportement intelligent, peut néanmoins permettre de faire progresser notre compréhension des mécanismes de la connaissance, à travers l'entreprise d'*augmentation* ou de *supplémentation* de la cognition humaine qu'elle prétend réaliser.

Le principal caractère qui rend l'IA capable de nourrir, par ses réalisations, une critique de la raison artificielle est pour Dreyfus sa démarche de *reproduction* ou de *simulation* du comportement intelligent. L'IA, selon la définition qu'en propose M. Minsky, se donne en effet pour tâche de reproduire (imiter, simuler) les performances cognitives des systèmes intelligents, et en premier lieu de l'homme. Aussi l'échec ou la réussite de ses tentatives offrent-ils d'évaluer empiriquement *l'adéquation* des principes qu'elle adopte pour modéliser l'activité cognitive et réguler le comportement des machines aux mécanismes qui sous-tendent le comportement intelligent de l'être humain.

La situation est différente dans le cadre de l'IC, dont le but n'est pas de simuler des agents capables de faire usage de connaissances, mais de créer des dispositifs qui, intégrés dans le circuit des activités situées humaines, seront capables *d'assister* efficacement ces activités. L'IC n'est pas dans un paradigme de reproduction ou de simulation, mais plutôt dans un paradigme d'*augmentation* ou de *supplémentation*. Les dispositifs qu'elle élabore visent à prothétiser l'activité cognitive humaine, en palliant ses limites intrinsèques, non à la reproduire.

Néanmoins, il nous semble que l'IC pourrait nous renseigner sur les mécanismes de la connaissance humaine, et le cas échéant contribuer au projet dreyfusien de critique de la raison artificielle, autrement qu'en cherchant à reproduire ces mécanismes dans des machines. Différents courants en sciences cognitives défendent l'idée que les performances cognitives de l'homme proviennent avant tout des outils et techniques qu'il utilise, ses prothèses intellectuelles pour ainsi dire. Tout un axe de la critique contemporaine du paradigme cognitiviste consiste ainsi à marquer que le raisonnement, l'activité de calcul, ou plus généralement le comportement intelligent humain, sont absolument irréductibles à une manipulation interne de symboles, car ils prennent toujours appui sur des artéfacts, des inscriptions – de sorte que la pensée est toujours et de manière essentielle externalisée. C'est un point que Dreyfus passe manifestement sous silence dans son entreprise de déconstruction de la philosophie rationaliste et de critique de la raison artificielle. Sa critique gagnerait pourtant de ce côté un appui des plus consistants : l'approche cognitiviste de l'intelligence ne se caractérise en effet pas seulement par la prétention à réduire les performances des systèmes intelligents à un calcul sur des représentations symboliques (hypothèse de Newell & Simon), mais également par ce qu'on pourrait appeler le postulat d'une intelligence nue : l'intelligence du cognitivisme est celle d'un cerveau autosuffisant, le corps, ses artéfacts ou son environnement n'y remplissent jamais une fonction constitutive.

Or, il nous semble que l'IC pourrait avoir beaucoup à apporter à la critique dreyfusienne de la raison computationnelle une fois celle-ci replacée dans cet axe problématique. Comme le fait l'IA selon Dreyfus, elle pourrait offrir un terrain empirique rendant testables les partis-pris du cognitivisme, en particulier son parti-pris d'une intelligence nue. La supplémentation cognitive a ses lois et ses mécanismes propres, et il est permis de penser que ceux-ci décideront en partie de la capacité des dispositifs de manipulation des connaissances que conçoit l'IC à augmenter de manière effective l'intelligence. Ces dispositifs pourraient, s'ils sont conçus sur la base de principes qui les rendent incapables de s'inscrire dans les pratiques ordinaires d'usage de supports pour penser, rencontrer des échecs analogues à ceux qui ont frappé l'IA. En nous montrant quel type de dispositif, et le cas échéant quel type de mécanisme régissant le comportement des dispositifs, est effectivement en mesure de suppléer l'activité de connaissance humaine, les réussites et les échecs de l'IC pourraient ainsi nous éclairer sur les mécanismes de l'activité de connaissance humaine, en tant qu'elle s'appuie constitutivement sur des supports. L'IC pourrait par là contribuer à la critique de la raison computo-symbolique que Dreyfus appelait de ses vœux en marquant que la connaissance, ne se passant jamais à l'intérieur des esprits, mais toujours à même les structures du monde, n'est jamais réductible à une performance intellectuelle nue et individuelle.