

## 282. La linguistique et l'intelligence artificielle

1. Avènement du courant symbolique en IA
2. IA symbolique et grammaire générative: le malentendu
3. Sémantique et représentation des connaissances
4. Analyseurs et modèles formels
5. Etiqueteurs et linguistique de corpus
6. Vers une linguistique expérimentale?
7. Bibliographie

En tant que projet scientifique, l'intelligence artificielle (IA) se caractérise par l'ambition de développer un modèle complet de l'homme, corps et pensée (*hard et soft*), représentation et action. En tant que technologie, elle développe des outils de simulation et de représentation pour remplacer ou assister l'homme dans certaines de ses activités intellectuelles. Sous ces deux aspects, l'IA s'intéresse dès ses débuts au langage et à sa compréhension; les langages *artificiels* qui servent à programmer les ordinateurs constituent des hypothèses sur le langage humain, dit – par opposition – *naturel*.

Dans sa déjà longue histoire, la linguistique a plusieurs fois redéfini son projet scientifique; une redéfinition importante a eu lieu au milieu du siècle, et il est probable qu'une autre soit en cours. Dans la transition entre ces deux moments, les relations entre la linguistique et l'IA – deux sciences de l'homme – jouent un rôle majeur.

Dans les sciences de l'homme, le modèle de rigueur qui s'est imposé au XIXe siècle dans les sciences de la matière devient dominant à partir des années 1940; il décrit tout système comme un mécanisme d'horlogerie, le décompose en parties aux fonctions bien délimitées, en axiomes, en *modules*, qui permettent d'en décrire les états successifs; sa version la plus achevée est l'axiomatique hypothético-déductive. C'est l'avènement du structuralisme puis de la grammaire générative, et de l'IA symbolique.

1. Avènement du courant symbolique en IA

Dès ses débuts, l'IA se partage en deux courants: le courant symbolique, qui développe des modèles mécanistes (algorithmiques) de la pensée et du raisonnement, à la suite de Von Neumann, et le courant connexionniste, qui développe des modèles du substrat neuro-

nal pour doter la machine de moyens de perception, à la suite des travaux du neurophysiologue McCulloch et du logicien Pitts. Budgétairement favorisé en raison d'intérêts stratégiques liés à la guerre froide, le courant connexionniste élabore des ordinateurs analogiques constitués de multiples processeurs interconnectés en réseau, ou *réseaux neuromimétiques*, spécialisés dans la reconnaissance de formes. Le premier réseau neuromimétique, la perception (Rosenblatt 1958), capable de reconnaître des images simples, suscita l'espoir et l'enthousiasme de nombreux chercheurs; Widrow (1959) implémente la notion de seuil.

En 1969, Minsky et Papert, deux chercheurs du courant symbolique, démontrent l'incapacité des perceptrons à calculer certaines fonctions simples, et en concluent à leur incapacité à effectuer les tâches requises pour la perception et l'intelligence. Cette critique s'appliquait au modèle à une couche, mais non aux modèles multi-couches au comportement difficile à caractériser par la logique formelle qu'on réussira à construire 13 ans plus tard. Mais à cette époque cette critique, utilisée comme preuve de la non-viabilité du connexionnisme, tombe bien: les recherches connexionnistes, moins stratégiques que par le passé, coûtent cher; la branche symbolique draine près de 80% des chercheurs et des crédits.

Les quelques chercheurs qui continuèrent dans cette voie, reconnus aujourd'hui comme des sommités en IA (Grossberg, Kohonen, Fukushima), furent marginalisés et ne trouvèrent à publier que dans des revues et colloques de psychologie, de neurophysiologie, d'acoustique. Il faudra attendre les années 1980, le succès du modèle de Hopfield, l'implémentation du modèle multi-couche et les développements technologiques favorisés par l'IA symbolique pour que le connexionnisme retrouve un souffle nouveau, dont témoigne le retour de chercheurs longtemps abrités par le courant symbolique (Rumelhart, Fahlman).

2. IA symbolique et grammaire générative: le malentendu

Dans les années 1960 la linguistique formelle, surtout la grammaire générative, devient le courant dominant en linguistique. L'IA