

INTELLIGENCE ARTIFICIELLE ET FORMATION

Claire Terlon

«L'Intelligence Artificielle est un champ renommé pour son manque de consensus sur les questions fondamentales» dit Wilensky (1983), directeur du Département d'Intelligence Artificielle à l'Université de Californie (Berkeley). Et au seuil des années quatre-vingt dix, le périodique «Byte» (Janvier 1991) -revue de référence en informatique- ouvre un dossier sur la «crise d'identité» que traverserait l'Intelligence Artificielle.

Si l'observation de Robert Wilensky atteste de la diversité des champs de recherche et des approches théoriques développés en Intelligence Artificielle (IA), et si l'article de Byte examine les résultats obtenus en regard d'attentes peut-être excessives suscitées ces vingt dernières années par l'aventure de l'Intelligence Artificielle, cela signifie également qu'il est assez périlleux d'essayer de fournir des repères sur les enjeux et les acquis de cette nouvelle discipline à l'intention de formateurs, dans une perspective où ils se sentent concernés. Cependant il nous paraît urgent de le faire, car dès à présent cette jeune «science», qui fait travailler ensemble des informaticiens, des linguistes, des spécialistes des sciences cognitives et des neurosciences, des psychologues, des philosophes, a produit des résultats -et même ses échecs sont instructifs- qui questionnent et souvent renouvellent un certain nombre de nos connaissances sur la communication langagière, sur le raisonnement (en particulier le raisonnement «commun» plus souvent utilisé que le raisonnement logique

Repères bibliographiques

Perspectives documentaires en éducation, n° 22, 1991

traditionnellement étudié), sur l'acquisition et l'utilisation des connaissances, et plus largement sur les phénomènes d'apprentissage. En outre, l'IA développe des outils pratiques, en particulier sous la forme d'un enseignement «intelligemment» assisté par ordinateur (EIAO), qui commencent à être utilisables dans des situations pédagogiques variées.

L'Intelligence Artificielle s'est d'abord développée comme un sous-domaine de l'informatique, mais a rapidement reconnu la nécessité de s'allier les compétences d'autres spécialités. C'est que, même si de nombreux chercheurs partagent le point de vue de Patrick Wilson, directeur du laboratoire d'IA au MIT (Massachusetts Institute of Technology), pour qui «le but premier de l'IA est de faire des *machines plus intelligentes*» (1987), la quête fondamentale de l'IA -qui revêt des aspects variés, marqués par les échecs ou les succès des orientations de recherche du moment - est d'élucider la *nature de l'intelligence*. Terry Winograd (Stanford University) le dit indirectement en exprimant l'idée partagée par beaucoup, qu'une des sources majeures des difficultés rencontrées et des limites aux réalisations de l'IA tient à ce que «la conception de l'intelligence qui sous-tend l'IA est superficielle et inadéquate». Les multiples chantiers de recherche ouverts par Roger Schank (Yale University), qui inspirent d'innombrables autres équipes, attestent très directement de l'objectif ambitieux de l'IA concernant l'intelligence (voir ci-dessous : 2.1).

Ceci nous introduit au coeur d'un débat très profond de l'IA, mais aussi d'autres sciences : les machines -les ordinateurs- doivent-elles chercher à imiter la nature, ici, l'intelligence humaine, ses capacités d'apprentissages variés, quitte à s'en éloigner à un certain moment (sinon les avions auraient des plumes !). Le récent développement des *réseaux neuro-mimétiques* illustre cette démarche. Si les premiers ordinateurs avaient été métaphoriquement qualifiés de «cerveaux électroniques», ce sont les difficultés rencontrées dans la réalisation de certaines tâches (en particulier en matière d'apprentissage, de reconnaissance de formes en temps réel...) qui ont fait s'interroger sur les limites résultant de l'architecture de la quasi-totalité des ordinateurs actuels (architecture de type von Neumann) imposant un traitement séquentiel de l'information. Dans le cerveau au contraire, la richesse des connexions inter-neurales assure un traitement coopératif, en parallèle, de l'information. D'où les tentatives (voir 4.) de ces toutes dernières années de simuler le fonctionnement des neurones : avec des réseaux artificiels comportant un nombre dérisoire d'éléments de traitement (quelques dizaines à quelques milliers) par rapport aux quelques 10¹¹

neurones du cerveau humain, reliés par des connexions également très schématiques par rapport aux connexions synaptiques cérébrales, les performances déjà obtenues sont radicalement différentes de celles des ordinateurs classiques ; ces réseaux permettent déjà, malgré le caractère rudimentaire de cette approximation d'un cerveau biologique, d'observer déjà des conduites émergentes qui jusqu'ici paraissaient le propre des systèmes vivants.

Cette démarche, qui va chercher des idées dans le monde naturel, est encore illustrée par le développement d'«*algorithmes génétiques*» qui effectuent des optimisations en s'inspirant de mécanismes de sélection naturelle de type Darwinien au niveau du génome. Le couplage d'algorithmes génétiques avec un réseau neuro-mimétique, inauguré par David Jefferson (Université de Californie à Los-Angeles), produit des résultats d'un extraordinaire intérêt, non seulement au niveau de l'informatique, mais également au niveau de l'éclairage qu'il offre en retour sur le phénomène de l'évolution des espèces... Ces recherches toutes récentes (Jefferson, 1990) restent du domaine du laboratoire. Mais qui n'a entendu dans le grand public, parler de «*virus informatiques*» ? Ces virus qui se multiplient à partir d'un fichier où le «*virus*» a été dissimulé par un programmeur (sous forme de quelques lignes de code), et infectent progressivement d'autres logiciels, s'ils présentent un comportement difficile à qualifier d'intelligent -même au sens de l'IA !- imitent cependant assez bien les propriétés d'une infection virale naturelle...

Cette brève évocation de développements très récents vise seulement à donner une idée du foisonnement des recherches -trop souvent l'Intelligence Artificielle est identifiée, dans les médias de grande audience, aux systèmes-experts, lesquels ne sont qu'une application très particulière des méthodes de l'IA... Cependant, si les réseaux neuromimétiques par exemple offrent de nouvelles possibilités sur la voie de la réalisation de machines intelligentes, l'essentiel des apports théoriques et pratiques de l'IA à la formation est pour l'instant à situer dans les domaines plus «*traditionnels*» de l'IA tels qu'ils se sont développés depuis une trentaine d'années.

Apports théoriques de l'IA à la formation

Dans ces domaines, il est possible de repérer deux grandes orientations autour desquelles s'organisent les travaux qui s'y effectuent : les uns participent d'une recherche à visée essentiellement théorique, les autres

relèvent d'une approche technologique de science appliquée avec l'objectif de développer des produits utilisables. (La distinction entre ces deux types de recherche est toutefois moins tranchée ici que dans d'autres disciplines : la validation d'un modèle théorique est la réalisation d'un programme (ou d'une machine) qui produit les résultats attendus de la théorie... et ce programme peut servir de première étape à la mise en oeuvre d'applications pratiques). Ces deux orientations, IA théorique et IA appliquée ont chacune une contribution à apporter au domaine de la formation.

Dans un article influent, Ira Goldstein et Seymour Papert (1977), du MIT, posent que «le but fondamental de l'IA est la compréhension des processus intelligents indépendamment de leur réalisation physique particulière». Cette position prenait acte de l'échec relatif de 20 ans de recherches informatiques orientées vers la résolution de problèmes généraux à l'aide des outils de la logique formelle. Mais cette prise de position s'inscrivait aussi, plus positivement, dans un renouvellement du questionnement sur l'intelligence qui intervenait dans les sciences humaines, les neuro-sciences fournissant de nouvelles méthodes d'investigation et de nouveaux modèles du fonctionnement cérébral. De même, sur l'autre côte des Etats-Unis, à Berkeley, l'objectif assigné à l'IA théorique au début des années 80 est très explicitement d'explorer «l'architecture de la cognition» (Wilensky, 1983), à la fois directement - les sciences humaines n'offrent pas de modèles d'activités cognitives complexes assez fins pour être représentés dans des programmes informatiques satisfaisants, et en testant les modèles élaborés sous forme de programmes informatiques que l'on fait tourner sur machine.

Ce que l'IA a ainsi fortement contribué à mettre en évidence, c'est le rôle crucial des connaissances dans toute activité cognitive humaine (comme aussi dans toute communication inter-personnelle). Ce rôle est reconnu par l'IA comme critique dans le fonctionnement du *raisonnement naturel* (ou commun), lequel, fonctionnant de façon non monotone, opérant des réajustements sous l'effet de connaissances nouvelles ou contradictoires, n'est pas réductible au maniement d'une logique formelle, et est pratiqué dans les circonstances les plus fréquentes de la vie courante (aussi par les scientifiques quand ils abordent un nouveau problème) : or cette modalité majeure de raisonnement - majeure aussi pour l'apprenant !- n'avait jusqu'à il y a peu, guère paru mériter une investigation scientifique car jugé irrationnel, inconstant... (Au lecteur curieux de ce raisonnement, nous voudrions suggérer le passionnant petit livre d'un professeur de

philosophie de Princeton University : «Change in View : Principles of Reasoning», par G. Hartman).

Le rôle des connaissances apparaît critique également dans le *fonctionnement de la langue*, et aussi dans *l'acquisition de nouvelles connaissances*. La complexité du fonctionnement d'un langage naturel (par rapport aux langages artificiels de programmation créés par les informaticiens) ne s'est trouvée que progressivement mise en évidence - que l'on songe aux tentatives, jugées aujourd'hui simplistes, de traduction automatique des années cinquante... Les modèles d'analyse linguistique existants se sont révélés dramatiquement insuffisants pour construire des programmes permettant à l'ordinateur de manipuler la langue. Il est apparu nécessaire de faire appel, pour rendre compte du fonctionnement de la langue comme outil de communication, à des processus extra-linguistiques (tels que raisonnements inférentiels) et surtout à d'énormes quantités de connaissances sur le monde. Si on veut pouvoir rendre la machine capable d'utiliser efficacement le langage naturel, il faut lui donner un moyen d'acquérir ces connaissances, et d'en assurer un stockage dynamique (faire qu'un contenu mémorisé soit mis à jour par l'arrivée d'une nouvelle information).

Ainsi l'IA se trouve-t-elle amenée à accorder au stockage mnémorique et à *l'architecture de la mémoire* une attention que les théories classiques de l'apprentissage n'explicitaient pas autant. C'est que dans toute activité cognitive et dans tout apprentissage, et pas seulement celui de la langue, l'IA montre comment l'acquisition de nouvelles connaissances est conditionné chez l'apprenant par le «background» de connaissances qu'il a déjà - sur lequel il s'agit d'articuler des stratégies pédagogiques adaptées (et pour l'informaticien de mettre en oeuvre une technique de représentation de connaissance pertinente). En fait cette importance des connaissances est telle qu'une machine ne pourra fonctionner de façon «intelligente» qu'à partir d'un stock considérable de connaissances... qu'il est impossible de fournir «à la main». C'est pourquoi un problème majeur de l'IA est de trouver des méthodes pour automatiser l'acquisition des connaissances par la machine. D'où une nécessité pour l'IA de ré-examiner les processus d'apprentissage humain pour en formaliser les différents types d'une façon utilisable par la machine. Ce faisant, l'IA enrichit considérablement nos connaissances sur les activités cognitives de l'apprenant.

Les recherches théoriques amenées en IA ne sont pas seules à contribuer à une meilleure connaissance des apprentissages. Des produits de l'IA appliquée telle que la technologie des systèmes-experts (1) a conduit à

élucider, mieux qu'on ne l'avait fait jusqu'ici, des caractéristiques du *fonctionnement de l'expert* (humain) en situation de résolution de problèmes de son domaine ; la distinction des stratégies de *l'expert* par rapport à celles du *novice* doit beaucoup à l'IA, qui fournit des indications directement utilisables par le formateur. Ici encore l'IA montre le rôle déterminant des connaissances, en mettant en évidence que ce sont les connaissances (dont des méta-connaissances) plus que le raisonnement qui différencient les comportements en matière de résolution de problème.

Des outils pratiques pour la formation

Enfin, un type d'applications directement utilisables par le formateur est constitué par les *systèmes d'EIAO* (enseignement «intelligemment» assisté par ordinateur). Un système complet d'EIAO doit comporter en principe trois composantes : une expertise du domaine, contenant les connaissances à faire acquérir à l'étudiant ; un «modèle de l'étudiant» indiquant ce que l'étudiant sait et ne sait pas encore, ses faiblesses particulières, et si possible ses modalités préférentielles de fonctionnement cognitif ; enfin, des stratégies tutorielles distinctes, c'est-à-dire non enfouies dans les connaissances du domaine enseigné, qui tendent à présenter des situations d'apprentissage adaptées au profil de l'étudiant (donné par le modèle de l'étudiant). L'idée fondamentale est qu'un système censé enseigner dans un domaine doit être un expert du domaine, donc posséder une représentation interne du domaine et être capable de raisonner sur ses connaissances -à la différence des programmes d'EAO traditionnels qui stockent des contenus sous la forme selon laquelle ils seront présentés et sur lesquels ils ne peuvent pas raisonner (les réponses aux questions attendues de la part de l'étudiant doivent donc être stockées préalablement à la mise à disposition du programme pour l'utilisateur).

Dans l'état actuel du développement de tels systèmes, c'est le module expert qui est le plus fréquemment développé (parce que le plus facile à faire) ; le modèle de l'étudiant fait l'objet de diverses approches (voir un exemple plus bas : programme Buggy) ; le module tutoriel est le plus délicat à réaliser : un tutorat intelligent implique que la stratégie pédagogique proposée ne soit pas uniquement fonction d'un contenu de connaissances à faire acquérir, mais prenne en compte le modèle (évolutif) de l'étudiant.

Les réalisations (voir : 3.2) restent limitées, ce qui traduit les limites des techniques informatiques utilisables dans le domaine de l'apprentissage, et aussi qu'une théorie de la transmission des connaissances reste à faire. Mais il faut souligner dans le même temps que les techniques d'IA offrent de nouvelles perspectives à la pédagogie, à la fois par les outils qu'elles rendent disponibles et par le moyen qu'elles offrent de tester les théories de l'apprentissage et de la cognition humaine.

Notes de lecture

Les ouvrages qui vont être brièvement évoqués ont été choisis en excluant les livres techniques nécessitant des connaissances approfondies d'informatique ou de mathématiques.

1. *Qu'est-ce que l'Intelligence Artificielle (IA) ?*

Deux ouvrages écrits par d'éminents informaticiens pour un large public, peuvent être proposés ici : ne requérant pas de connaissances mathématiques ou informatiques spécialisées, tout en étant d'une grande rigueur, ces deux ouvrages introduisent aux concepts et aux enjeux les plus importants du domaine.

- 1.1. Bonnet, Alain. *Intelligence Artificielle*. Inter-Editions, 1984. 242 p.

Cette excellente introduction à l'IA permet au lecteur d'acquérir un certain nombre de données de base, et les moyens de les situer dans une problématique plus vaste. Ce livre décrit les problèmes-clés de la représentation des connaissances, et donne une idée des modèles actuellement disponibles. En illustrant son exposé par des exemples concrets, l'auteur rend intuitifs des aspects délicats de certains formalismes, et leurs limites. De même se trouve clarifiée la notion d'*objets structurés*, permettant ainsi au lecteur non «spécialiste» de tirer profit de lectures ultérieures qui utilisent des appellations très diverses pouvant déconcerter (d'autant plus que les distinctions entre les caractéristiques des objets ainsi dénommées ne sont pas toujours spécifiées, quand elles existent). Le formateur y trouvera (chap. 18) un chapitre -un peu optimiste- sur l'EIAO (Enseignement intelligemment assisté par ordinateur), et d'intéressantes notations sur «les programmes capables d'apprendre» (chap. 19). Une présentation de quelques programmes qui ont marqué la courte histoire de l'IA complète heureusement ce panorama.

- 1.2. Arzac, Jacques. *Les machines à penser*. Seuil, 1987. 249 p.

L'auteur, pionnier de l'informatique en France, a joué un rôle essentiel pour l'introduction de l'informatique dans l'enseignement secondaire - et d'innombrables enseignants de tous horizons disciplinaires ont bénéficié, à travers «Premières leçons de programmation» (Cedic-Nathan, 1980) et «La construction de programmes structurés» (Dunod, 1977), d'un talent pédagogique que l'auteur n'a pas consacré uniquement à ses étudiants de Paris VI. Ce talent est à nouveau à l'oeuvre dans le présent ouvrage : même si l'Intelligence Artificielle ne représente environ qu'un tiers du livre, le formateur ne regrettera pas le temps qu'il consacrera à la lecture des pages limpides que l'auteur consacre à présenter l'informatique, comme il savourera la partie traitant de l'informatique à l'école - laquelle démolit allègrement pas mal d'idées reçues.

Moins «technique» que le précédent, ce livre veut engager son lecteur à une réflexion sur les outils et les méthodes de l'IA, et plus profondément sur la *question du sens*, c'est-à-dire, ultimement, de la possibilité d'une intelligence artificielle (qui manie autre chose que du formel ou du combinatoire).

- 1.3. Rich, Elaine. *Intelligence Artificielle*. Masson, 1987. 439 p.

est la traduction française d'un ouvrage publié aux Etats-Unis en 1983 (1ère édition) et qui est un des manuels parmi les plus utilisés par les étudiants commençant l'intelligence artificielle. Il s'agit d'un cours, très progressivement construit, avec le souci d'être accessible à des non-informaticiens : puisque l'IA est un domaine qui s'enrichit de nombreux apports disciplinaires, il importe de donner, à des étudiants se spécialisant dans des domaines autres que l'informatique, des éléments qui pour certains auront une dimension essentiellement culturelle et pour d'autres seront un point de départ à un approfondissement ultérieur.

Ce livre a sa place ici pour deux raisons. D'une part, le lecteur intéressé par les perspectives ouvertes par les ouvrages précédents peut avoir envie d'une information plus précise sur certains points ; d'autre part, un formateur peut être tenté d'utiliser un outil de développement de système-expert (ou «shell») parmi les dizaines de produits commercialement disponibles, pour créer un petit système-expert d'aide tutorielle dans le domaine de son choix : s'il s'est essayé à utiliser un langage-auteur pour construire un programme d'EAO (enseignement assisté par ordinateur), il verra qu'avec un «shell» il n'a pas davantage besoin de connaissances spécialisées en informatique, mais surtout qu'il peut introduire «en vrac» -de façon déclarative- des connaissances qui seront gérées par le logiciel

pour répondre à des questions de l'étudiant, sans avoir à prévoir ces questions dans le détail pour stocker les réponses par avance, ce qui est une des gageures de la construction de programmes d'EAO. Cependant, pour faire un choix adapté à ses objectifs parmi la diversité des produits offerts, et ensuite pour utiliser le plus efficacement possible l'outil de développement dont il a fait l'acquisition (et tirer le meilleur parti de la documentation accompagnant le logiciel), le formateur aura intérêt à avoir quelque connaissance des *mécanismes de raisonnement* et des *modes de représentation* des connaissances utilisées en IA, qu'il pourra trouver dans le livre de Rich. Le linguiste pourra être intéressé par les importants développements que l'auteur consacre à la «*compréhension du langage naturel*» (par la machine) ; les exemples qui servent à illustrer le fonctionnement des méthodes d'analyse et de représentation des structures linguistiques sont les exemples anglais traduits en français, parfois au prix d'acrobaties (la traduction du livre est globalement assez bonne à quelques erreurs près) : le professeur d'anglais peut trouver matière à réflexion sur le fonctionnement de la langue qu'il enseigne (surtout s'il peut avoir accès à l'ouvrage américain «*Artificial Intelligence*» McGraw Hill, source de la traduction ; nous aimerions aussi lui suggérer une excursion dans le superbe «*Language as a Cognitive Process*» de Terry Winograd, Addison-Wisley, 1983).

2. IA et sciences cognitives

L'IA questionne et enrichit les sciences cognitives par les modélisations de processus cognitifs qu'elle fait «tourner» sur machine. L'IA contribue également à renouveler certaines approches des processus d'apprentissage, qu'elle étudie pour progresser dans le difficile problème de l'automatisation de l'acquisition des connaissances par la machine - problème crucial pour ce qui est de la réalisation de machines «intelligentes».

- 2.1. Schank, Roger C.
- 2.1.1. *The Cognitive Computer*. Addison-Wesley, 1984. 194 p.
- 2.1.2. *Reading and Understanding ; Teaching from the Perspective of Artificial Intelligence*. Lawrence Erlbaum Associates, 1982. 196 p.
- 2.1.3. Schank, R.C. and Abelson, R.P. *Scripts, Plans, Goals and Understanding (An Inquiry into Human Knowledge Structures)*. Lawrence Erlbaum Associates, 1977. 248 p.

Schank est un des plus brillants et des plus influents avocats d'une conception de l'IA qui assigne à celle-ci l'objectif fondamental de mener

une investigation des activités cognitives humaines pour construire des modélisations qui, tout en utilisant des procédés distincts des processus cognitifs dont elles s'inspirent explicitement, voient leur validité testée sur ordinateur (plutôt qu'éprouvée au regard d'une théorie de la cognition humaine -qui pour l'instant n'existe pas). Au lieu de passer du temps sur comment faire réaliser par l'ordinateur des tâches étroitement définies (du jeu d'échecs aux systèmes-experts), les chercheurs doivent, soutient Schank, s'affronter à la vraie question de la nature de l'intelligence, en s'efforçant de définir avec précision en quoi consiste pour les humains le fait de comprendre, d'apprendre, d'avoir des idées qui se modifient en fonction des expériences... (p.30-31). Alors seulement pourront-ils essayer de construire des machines réellement utiles, c'est-à-dire capables d'apporter une aide intelligente dans une grande diversité de domaines, à un utilisateur qui interagira avec la machine de la façon qui lui est la plus naturelle, c'est-à-dire en usant de la langue naturelle.

D'où l'intérêt de Schank pour l'étude de la communication langagière, qui le conduit directement au problème de la représentation des connaissances (les gens comprennent plus que le contenu lexical d'une expression : interviennent des connaissances qui permettent de gérer l'implicite du discours, de générer des attentes (expectations) sur ce qui peut arriver dans tel contexte, de faire des inférences plausibles pour lever des ambiguïtés, etc...). Pour représenter les connaissances nécessaires, Schank propose le modèle de la «Dépendance Conceptuelle», dont la première forme est exposée dans l'ouvrage cité 2.1.3. (on notera le sous-titre, qui dès 1977, fixe clairement l'objectif de Schank); ce modèle sera ensuite sans cesse modifié et enrichi, par l'auteur lui-même et par ses nombreux étudiants, dont Wilensky déjà cité. Un intérêt supplémentaire de la démarche schankienne est de reconnaître que le problème central de la compréhension du langage est dans le même temps celui de l'organisation de la mémoire (qui ne stocke pas littéralement toutes les propositions rencontrées, mais probablement doit user de formes canoniques -du genre de celles que Schank crée comme opérateurs dans sa Dépendance Conceptuelle : en l'état actuel des recherches, les structures complexes créées à partir de ces opérateurs permettent une première approche de l'organisation d'une mémoire dynamique (c'est-à-dire capable de réorganisation en fonction d'expériences nouvelles), ce qui donne déjà des programmes d'application assez stupéfiants (en particulier : compréhension de textes permettant des réponses élaborées à des questions non triviales). L'ouvrage cité (2.1.1.) en donne une idée, certes déjà dépassée par de plus récentes publications de recherche, mais sous

une forme très accessible. Dans ce livre également, on trouvera exposées (avec la clarté des vulgarisations scientifiques anglo-saxonnes) les grandes idées schankiennes, et une formulation des principaux concepts de la théorie de la Dépendance Conceptuelle pour un large public.

Le deuxième livre proposé (2.1.2.) présente une application des concepts schankiens à l'enseignement de la lecture, sous-tendue par les idées-clés de l'IA sur le rôle crucial des connaissances dans la compréhension du langage et de processus inférentiels pour faire sens de formulations recouvrant une large part d'implicite (et pour permettre au lecteur d'anticiper ce qui va suivre : mot, unité syntaxique, déroulement d'événements). Un des intérêts du livre est de montrer comment les concepts schankiens permettent de rendre compte des démarches -et de les formaliser- par lesquels l'enfant chemine pour acquérir les moyens d'une active construction du sens qui soutient son intérêt, accroît ses habiletés cognitives, et devient ainsi un bon lecteur -qui aime la lecture. Ce faisant, et en montrant comment les processus d'analyse et d'inférences engagés par l'enfant ne peuvent être étudiés isolément de la mémoire - on ne comprend quelque chose qu'en relation avec ce qu'on a déjà mémorisé en d'autres circonstances- Schank est amené à formuler des recommandations très pratiques à l'intention des auteurs de livres visant à l'apprentissage de la lecture, et à l'intention des acteurs de la relation pédagogique, qui ne peuvent laisser le formateur indifférent.

- 2.2. Michalski, Ryszard S., Carbonnel, Jaime G., Mitchell, Tom M. *Machine Learning*. vol. II. Morgan Kaufmann Publishers, 1986. 739 p.

Ce livre s'ouvre sur un chapitre intitulé «Comprendre la nature de l'apprentissage» qui fait un bilan des connaissances de l'IA en la matière, des problèmes non résolus, et argumente la nécessité de mieux analyser les divers types d'apprentissage (sur lesquels sont fournies, avec des définitions précises, une excellente bibliographie qui est un véritable outil de recherche). Ce n'est qu'à ce prix que l'on pourra faire des progrès dans l'automatisation de l'acquisition des connaissances par une machine, si l'on veut que celle-ci puisse fonctionner comme système-expert, ou manipuler la langue naturelle de façon non superficielle, ou rendre les services d'un tutoriel intelligent, ou, plus largement, avoir un «comportement intelligent».

Pour tirer le meilleur parti de ce livre, il est nécessaire d'avoir de bonnes connaissances informatiques. Mais si ce livre est proposé ici, c'est, qu'outre les chapitres déjà mentionnés qui sont une riche source d'informations pour un formateur non-informaticien - que celui-ci pourra trouver, dans les quatre grandes sections (l'apprentissage à partir

d'exemples, l'apprentissage par analogie, l'apprentissage par l'observation et la découverte, les aspects cognitifs de l'apprentissage) - il se trouve des exemples et des analyses qui pourraient figurer tels quels dans un ouvrage de psychologie cognitive (c'est seulement l'aspect «implémentation» informatique des modèles proposés qui fait appel à des connaissances techniques spécifiques à l'IA). Ainsi l'analyse par Forbus et Gentner (p.311-346), et les modèles qu'ils proposent, des étapes par lesquelles passe l'apprenant pour construire sa compréhension d'un phénomène physique jusqu'au modèle du physicien, est-elle très intéressante non seulement pour un professeur de physique, mais aussi pour un lecteur curieux de voir mis en oeuvre des concepts d'IA en matière d'apprentissage pour éclairer certaines étapes des acquisitions de l'apprenant. D'une façon analogue, les difficultés rencontrées dans la réalisation de programmes capables d'apprendre par analogie, permettent en retour une meilleure attention aux problèmes rencontrés par un élève dans le maniement de ce type de raisonnement, et d'en préciser les conditions pour un usage efficace (et donc les limites de l'utilisation).

- 2.3. Anderson, John R. *The Architecture of Cognition*. Harvard University Press, 1983. 345 p.

A partir d'une modélisation de la mémoire long-terme (modèle HAM : Human Associative Memory), l'auteur a élaboré une théorie de l'apprentissage et de la mémoire : en ajoutant une mémoire de travail (court-terme) à son modèle précédent, il bâtit le système ACT (Adaptative Control of Thought) qui, empruntant à l'IA comme aux neurosciences, fournit un cadre conceptuel de référence pour de nombreuses recherches actuelles en psychologie cognitive.

Ce modèle, qui vise à fonder une théorie unitaire des fonctions cognitives, décrit l'acquisition d'une habileté cognitive (cognitive skill) comme comportant essentiellement deux étapes : une étape déclarative, dans laquelle les faits du domaine de cette habileté sont interprétés (et codés sous forme d'un réseau de propositions dans le modèle ACT), et une étape procédurale dans lesquelles les connaissances du domaine se trouvent condensées sous forme de procédures (codées dans ACT sous forme de règles de production). Un des intérêts du modèle est qu'il interprète un certain nombre de conditions d'apprentissage de façon à permettre des interventions pédagogiques adaptées : ce modèle montre, par exemple, pourquoi un feed-back immédiat fourni à l'étudiant en situation de résolution de problème, est beaucoup plus efficace qu'une évaluation différée de son travail ; comment par exemple certaines erreurs de l'apprenant peuvent résulter d'un excès de charge imposé à la

mémoire de travail (où sont mises en relation les conditions des règles des connaissances procédurales, avec les données du problème) etc...

(Divers apprentissages (de la programmation en LISP, de connaissances de géométrie...) ont été modélisés comme cas particuliers du système ACT, et les simulations ont été comparées avec succès aux données collectées par observation d'apprenants. Un tutoriel de géométrie élémentaire a également été développé comme application de la théorie d'Anderson. Ces développements qui n'apparaissent pas dans le livre de 1983, mais dans des publications ultérieures, attestent de la fécondité du modèle).

3. IA et Enseignement assisté par ordinateur.

- 3.1. Nicaud, Jean-François et Vivet, Martial. *Les tuteurs intelligents* : réalisations et tendances de recherches in *Technique et Science Informatique*, vol. 7, n° 1, 1988. p. 21-45.

Cet article décrit les composantes d'un système d'EIAO en s'appuyant sur des exemples qui ont été développés aux Etats-Unis (Guidon/ Mycin ; West ; Scholar ; Geometry Tutor) et en France (Proust, système capable de détecter des erreurs dans les programmes en Pascal écrits par des débutants ; Amalia, destiné à enseigner le calcul algébrique au niveau du 1er cycle des universités ; Aplusix, projet de tuteur sur l'enseignement de manipulations algébriques). Les auteurs fournissent un panorama général des tuteurs intelligents existant ou en cours de réalisation, et donnent des indications sur ce qui leur paraît être les tendances actuelles dans le domaine.

- 3.2. Sleeman, D. et Brown, J.S. *Intelligent Tutoring Systems*. Academic Press, 1982. 344 p.

Cet ouvrage reste une référence pour tout développeur de modules tutoriels. On y trouve à la fois une analyse critique des techniques d'IA utilisables pour construire ces tutoriels, et surtout -par rapport à la perspective qui est la nôtre ici- une analyse approfondie de quelques programmes qui ont fait date dans l'histoire de l'EIAO. Cette analyse concerne les objectifs des auteurs de chacun de ces projets, les données de l'observation d'étudiants utilisant le tutoriel ; les limites qui ont été rencontrées sont interprétées : limites liées à la technique informatique particulière d'implémentation du projet ; limites venant des connaissances encore très partielles sur le fonctionnement cognitif des apprenants et sur les interventions pédagogiques pertinentes aux difficultés de divers types pouvant surgir dans une situation de résolution de problème. La

portée de ces analyses dépasse largement le cas particulier de réalisation tutorielle à partir duquel elles ont été développées.

Les programmes présentés sont très variés, du diagnostic de pannes dans un circuit électronique (Sophie) à l'apprentissage de techniques de calcul intégral, en passant par un système qui analyse les explications données en langue naturelle par l'étudiant (Ace), ou un système (Buggy, Debuggy) qui cherche à inférer le cheminement erroné de l'élève aux prises avec des opérations arithmétiques simples -l'hypothèse étant que si certaines erreurs sont le fruit du hasard, beaucoup peuvent être interprétées comme le résultat de l'application de règles erronées (qu'il s'agit de détecter), pour une intervention pédagogique adaptée au vrai problème de l'élève (Debuggy).

(Un logiciel, FRACT, a été développé à l'INRP avec des équipes universitaires associées, qui met en oeuvre ces idées dans le domaine des fractions : un corpus d'erreurs a été constitué, des «mal-règles» ont été proposées pour rendre compte d'erreurs systématiques ; un logiciel les utilise pour aider l'enseignant à repérer un cheminement plausible de l'élève -la difficulté étant la superposition éventuelle de plusieurs mal-règles concourant à la production d'un résultat erroné (le logiciel peut faire intervenir jusqu'à trois mauvaises règles susceptibles d'avoir été utilisées dans l'addition de deux fractions)).

4. Influence des architectures

- Caudill, Maureen and Butler, Charles. *Naturally Intelligent Systems*, The MIT Press, 1990. 304 p.

Ce livre nous paraît une des meilleures introductions aux réseaux neuro-mimétiques auxquels un lecteur cultivé peut avoir accès : celui-ci y trouve présentés avec simplicité et élégance les concepts fondamentaux, un panorama détaillé des nouvelles architectures, une réflexion sur la nature et la portée de ces simulations, des éléments prospectifs sur les applications.

Un réseau neuro-mimétique s'inspire du fonctionnement du réseau des neurones cérébraux : des éléments de traitement qui simulent les neurones (et pour lesquels Caudill crée le néologisme ingénieux de «neurode», condensation de l'anglais neuron et node, soit : neurone, et noeud -terme général qui désigne les intersections d'un réseau), sont assemblés dans une structure où ils sont reliés les uns aux autres pour former en général des couches qui assurent des fonctions différentes. A

la manière d'un neurone, un neurode combine des signaux d'entrée et émet, si le niveau d'activation résultant de cette combinaison est supérieur à un seuil, un signal de sortie qui peut être éventuellement envoyé comme signal d'entrée à un ou plusieurs autres neurodes.

Toutes les architectures (c'est-à-dire les modes d'assemblage des neurodes) ont en commun de mettre en oeuvre la vieille loi de Hebb (2) (1949) : si deux neurones connectés entre eux sont activés en même temps, alors la connexion qui les relie est renforcée (dans le cas contraire, la connexion n'est pas modifiée). Dans un réseau neuro-mimétique les connexions entre neurodes sont affectées de poids modifiables (modulant le signal qui suit ce trajet) : l'apprentissage est possible par une modification des poids de ces connexions (le rôle majeur de ces connexions est à l'origine du terme «connexionisme» pour désigner les développements concernant les réseaux neuro-mimétiques). Certaines architectures mettent en jeu des formes d'apprentissage supervisé (le système reçoit une information sur ses erreurs), d'autres un apprentissage autonome (réseaux de T. Kohonen, Université Technologique d'Helsinki, Finlande ; réseaux ART, basés sur la théorie de la résonance adaptative développée par S. Grossberg, Boston University). Le plus grand nombre d'architectures actuellement expérimentées utilisent l'apprentissage supervisé, avec des approches différentes de la façon de rétro-propager l'erreur observée en sortie par rapport à la réponse désirée (la difficulté majeure est en effet : à quels neurodes imputer quelle part de cette erreur) ; ce sont aussi ces architectures qui actuellement ont le plus d'applications pratiques (décisions financières, distinction de mines immergées de rochers de formes voisines, modélisation de la circulation sur autoroute urbaine, apprentissage par un bras de robot de mouvements variés sous diverses charges etc...). Dans tous ces apprentissages on ne programme pas à l'avance les actions à exécuter ou les décisions à prendre, mais on présente au système une grande variété d'exemples à partir duquel le système apprend avec l'aide d'un feed-back (de plus l'opérateur humain qui apprend en robot à utiliser son bras n'a pas besoin de comprendre les propriétés du robot pour l'entraîner...)

Avant d'en arriver là, la recherche a parcouru de curieux zig-zags, et l'un des intérêts du livre de Caudill - outre ses qualités d'excellente vulgarisation, à la fois scientifiquement exigeante et passionnante pour un non spécialiste - est de mettre en perspective des contributions dispersées qui ont finalement concouru à la genèse des nouvelles conceptions. Ainsi de la première tentative de réseaux visant à simuler le fonctionnement cortical nécessaire à la perception visuelle : c'est le

célèbre perceptron des années «50», réseau très simple à deux couches de cellules, capable de réaliser un certain nombre de fonctions logiques lui permettant des apprentissages limités. Ce sont précisément ces limites, analysées par S. Papert et A. Minsky (1968-69) comme définitives, qui mettent un terme, pour longtemps, aux recherches dans le domaine des réseaux. Or dès la fin des années 50, un théorème avait été établi par le mathématicien soviétique Kolmogorov, qui montre qu'un réseau à trois couches permet de surmonter les limites indûment extrapolées du perceptron à des réseaux plus complexes : il faudra attendre 20 ans pour voir appliquer ce théorème à la théorie des réseaux de neurones artificiels...

En parallèle avec le livre de Caudill on pourra consulter «*Des réseaux de neurones*» de F. Davals et P. Naïm (Eyrolles, 1990) pour avoir une idée du modèle mathématique sous-jacent à chaque type de réseau. Il paraît difficile de conseiller la lecture de ce livre seul, qui juxtapose des descriptions de modèles sans aider le lecteur à l'exploration du domaine –en l'éclairant sur les raisons de telle ou telle orientation, la signification de tel ou tel développement, de tel ou tel échec..., c'est-à-dire en l'engageant à la réflexion.

Claire Terlon

Enseignant-chercheur, Université Paris 2

- (1) Un système-expert est une application de techniques d'IA à la simulation de l'expertise humaine dans un domaine limité de connaissances approfondies : il permet la résolution de problèmes dans un domaine spécifique en utilisant des connaissances acquises auprès des experts du domaine, et un mécanisme de raisonnement sur ces connaissances.
- (2) avec des modulations : tel par exemple le modèle néo-hebbien de S. Grossberg.

Références additionnelles

- Byte, Jan. *AI's Identity Crisis*. B. Ryan, 1991. pp. 239-246.
- Goldstein, I. ; Papert, S. *Artificial Intelligence, language, and the study of knowledge*, Cognitive Science, 1977, 1, p. 84-123.
- Harman, G. *Change in View : Principles of Reasoning*. MIT Press, 1986. 147 p.
- Jefferson, D.R. «*The Genesys System : Evolution as a Theme in Artificial Life*», in : *Proceedings of the 2nd Artificial Life Workshop*. Langton, C. ed. Addison-Wesley, 1990.
- Wilensky, R. *Planning and Understanding*. Addison-Wesley, 1983. 168 p.

- Winograd, T. *Thinking Machines : Can There Be ? Are we ?* in : *Understanding Computers and Cognition : A New Foundation for Design*. Ablex, Norwood, NJ, 1986. 212 p.
- Wilson, P.H. *Artificial Intelligence : a perspective*, in : *Artificial Intelligence in the 1980s and Beyond*. Edited by W. Eric, L. Grimson, and R.S. Patil. Cambridge, MA : MIT Press, 1987.

