

Anne NICOLLE\*

## **L'expérimentation et l'intelligence artificielle**

Le rôle de l'expérimentation en intelligence artificielle et l'usage, pour l'expérimentation dans les sciences cognitives, des techniques et outils de l'informatique et de l'intelligence artificielle sont analysés pour inviter à approfondir les recherches interdisciplinaires.

*Mots-clés* : expérimentation, intelligence artificielle.

**Experimentation and Artificial Intelligence.** The role of experimentation in artificial intelligence and the use for experimentation in cognitive science of technics and tools developed by computer science and artificial intelligence are analysed to invite to go further into interdisciplinary research.

*Key Words* : experimentation, artificial intelligence.

L'intelligence artificielle expérimentale est une discipline scientifique qui a pour rôle d'augmenter les connaissances, dans différents domaines du fonctionnement mental et social (mémoire, langage, perception, conception, diagnostic...), en utilisant les ordinateurs comme outils pour expérimenter des modèles. Les modèles proposés par les chercheurs d'autres disciplines, comme la linguistique ou la psychologie cognitive, sont des modèles descriptifs, qui permettent une analyse rétroactive des traces d'une activité. Il n'est pas possible de les tester directement sur des machines car, contrairement à l'homme, la machine ne comprend pas avant d'analyser ; elle a besoin d'un modèle pour agir, pas seulement pour décrire. Il faut donc concevoir des modèles plus profonds, des modèles de la production et de la compréhension des systèmes de signes et de leurs expressions. L'implantation en machine de ces modèles qu'on appelle des logiciels

---

\* GREYC, URA CNRS 1526, Université de Caen, 14032 Caen Cedex, E-mail : Anne.Nicolle@info.unicaen.fr

d'étude permet de construire des modèles effectifs, qui ont une activité, et permet d'analyser les points de ressemblance et de dissemblance avec le fonctionnement du phénomène modélisé. Un exemple bien connu de logiciel d'étude est AM (Lenat, 1982) qui a suscité beaucoup de réflexions intéressantes sur les processus de recherche scientifique. Mais en l'absence de relais interdisciplinaires dans les sciences de la cognition (épistémologie, métaphysique, psychologie cognitive), il est difficile de produire des résultats cumulatifs. Construire des modèles effectifs, ou proactifs, est un objectif qui donne un fondement solide à l'interdisciplinarité dans les sciences de la cognition. L'ensemble des articles de ce dossier en est l'illustration : il relate différentes expériences de recherche interdisciplinaire développées sous des formes variées et avec des présupposés différents au sein et autour du pôle interdisciplinaire "Modélisation en sciences cognitives" de la Maison de la Recherche en Sciences de l'Homme de Caen.

L'examen et l'illustration de deux questions vont faire le lien entre les différents articles présentés dans ce dossier :

- Comment l'informatique et l'intelligence artificielle modifient-elles la démarche expérimentale dans les sciences cognitives ?
- Quel est le rôle de l'expérimentation en Intelligence Artificielle et comment les autres sciences cognitives amènent-elles l'intelligence artificielle à construire une démarche expérimentale ?

Avant d'entrer dans le vif de ces questions, proposons une analyse pragmatique du processus de modélisation et d'expérimentation afin de préciser certains termes.

### **1. LA DEMARCHE EXPERIMENTALE**

Elle s'appuie sur trois activités complémentaires, observer, analyser et modéliser sans qu'on puisse dire laquelle est première. Elle fait appel à trois sortes d'activités dont le chercheur a plus ou moins conscience : une activité de l'ordre de la métaphysique pour le choix ou la construction des théories qui guident l'analyse des phénomènes, une activité de l'ordre des mathématiques pour la construction d'un méta-modèle, qui sera instancié comme modèle à vérifier, et une activité de l'ordre des sciences de l'ingénieur, qui va permettre de construire des situations d'observation pertinentes en créant des instruments de plus en plus fins. L'ordinateur est un instrument particulier puisqu'il peut se programmer et donc il est plus un générateur d'instruments variés qu'un instrument au sens où l'est un microscope. Il ouvre donc un champ nouveau à la recherche expérimentale.

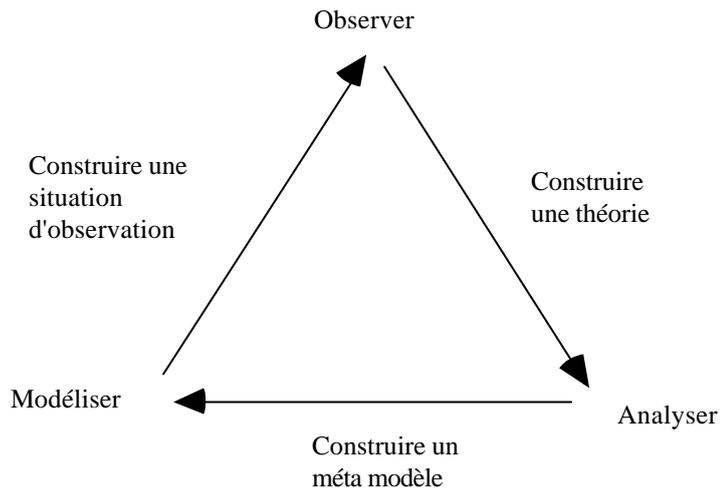


figure 1 : les trois stades de la démarche expérimentale et les processus de transition

La méthode expérimentale dans les sciences de la matière est la plus ancienne et définit le paradigme auquel les autres sciences font référence. La modélisation consiste à trouver des formules mathématiques qui soient de bonnes approximations des phénomènes observés. Elle est fondée sur une théorie de la mesure objective et sur des techniques de construction d'appareils de plus en plus perfectionnés. Le modèle obtenu est valide s'il est prédictif. Lorsque deux hypothèses théoriques contradictoires sont proposées, on cherche à construire une expérience dont le résultat permettra de faire la différence entre les deux hypothèses.

Dans les sciences cognitives, la méthode expérimentale peut être fondée sur trois théories : le behaviorisme, le structuralisme et le constructivisme. Les deux premières se rattachent au mythe de l'objectivité et la troisième au mythe expérientialiste. Une analyse des mythes de l'objectivité et de la subjectivité est proposée dans Lakoff et Johnson, 1980, comme leur remplacement par le mythe expérientialiste. Ce mythe, qui conserve les critères de rationalité de l'objectivité, mais qui tient compte du sujet épistémique, permet de fonder plus solidement une méthode scientifique en sciences cognitives. L'objet de ce dossier est aussi de discuter et d'illustrer cette assertion, en particulier pour ce qui concerne la place du sujet dans les modèles (que ce soit comme sujet-type ou comme sujet-occurrence). L'article de Jean Lassègue introduira cette question.

Trois méta-modèles sont utilisés dans les sciences cognitives, et s'il y a une certaine indépendance entre une théorie et un méta-modèle, ils sont le plus souvent associés comme suit :

*Théorie 1 — Le behaviorisme + méta-modèle 1 - les statistiques*

Hypothèse fondatrice du behaviorisme : on ne peut pas comprendre les processus cognitifs, il faut chercher des régularités entre les entrées et les sorties des processus.

Le méta-modèle associé au behaviorisme est statistique. Il est descriptif "en moyenne" et la prédiction individuelle n'a pas de sens. Les tests d'hypothèses proposés dans le méta-modèle permettent d'invalider des hypothèses, jamais de prouver leur validité. Son principal intérêt est d'avoir fait sortir la psychologie du subjectivisme et de l'auto-observation. La principale critique qu'on puisse lui faire est qu'il n'est ni explicatif, ni prédictif des phénomènes individuels. Il explique les tendances et non les différences.

*Théorie 2 — Le structuralisme + méta-modèle 2 - logique et algèbre.*

Hypothèse fondatrice du structuralisme : il existe des structures régulières dans les phénomènes sociaux ou cognitifs et l'explication consiste à les mettre à jour.

Le méta-modèle associé au structuralisme est un schéma logique (théorie des ensembles, calcul des prédicats) et algébrique (théorie des graphes, des groupes) qui décrit un état des choses. Il est statique et explicatif car la place de chaque objet est définie par rapport aux autres et à la structure globale. L'invalidation des modèles est à base de contre-exemples. La principale critique que l'on peut faire au structuralisme est que les modèles construits sont statiques : ils décrivent un état des choses et non les évolutions entre les états. Étudier les structures est une bonne chose mais il faut décrire en outre les formes que les structures construisent et les processus qui les instancient et les remplissent. Le principal intérêt du structuralisme est d'avoir mis en évidence les relations complexes entre le tout et ses parties. Au lieu de s'intéresser uniquement à décrire les objets, on s'intéresse à leurs relations et aux invariants de structure. Dans la description des objets, on cherche à les différencier et non plus seulement à les caractériser.

*Théorie 3 — Le constructivisme et la systémique + méta-modèle 3 - fonctionnel.*

Hypothèse fondatrice du constructivisme : l'observateur ne peut pas être évacué des modèles scientifiques qu'il construit. C'est son activité qui crée son rapport au monde et c'est ce rapport du sujet au monde dont il est question dans les modèles scientifiques, et non d'un monde "objectif".

Hypothèse fondatrice de la systémique : il faut analyser les systèmes complexes dans leurs structures et dans leurs formes, et les processus qui font évoluer les structures et les formes. Il y a plusieurs niveaux d'analyse correspondant à des points de vue différents qui peuvent être dans des rapports complexes, non compositionnels.

En partant de ces deux hypothèses, l'observation n'est pas un problème trivial, car elle résulte des choix du sujet relativement à son projet (Mugur-Schächter, 1989). La reconnaissance de la relativité de l'analyse amène à s'intéresser de plus près aux différences, car les représentations sont par nature incomplètes (Davis & *al.*, 1993]. La modélisation consiste à construire, à partir des phénomènes perçus, des systèmes ayant un intérieur et un extérieur. L'intérieur peut avoir plusieurs niveaux qui se définissent les uns les autres. L'interprétation d'un niveau dans les termes du niveau précédent n'est pas compositionnelle mais émergente. Le plus petit exemple de système est sans doute le jeu de GO (intérieur : le plateau, les pierres, les joueurs, les règles du jeu) où les pierres et les territoires forment deux niveaux d'analyse distincts, en relation non compositionnelle. Chaque niveau est un point de vue du sujet modélisateur, mais pas en tant qu'individu, en tant qu'humain ayant un rapport d'intériorité-extériorité aux choses qu'il décrit, car lui-même fait partie du monde dans tous ses aspects : physico-chimique, biologique, psychologique et social. Les modèles proposés en partant de ces deux hypothèses sont effectifs, parce qu'ils prennent en compte simultanément les objets, les structures et les processus. On peut les simuler sur machine et observer les différences avec les phénomènes réels pour les ajuster (invalidation par la simulation). Le méta-modèle doit être fonctionnel, ce peut être le lambda-calcul, la géométrie différentielle ou d'autres manières de mettre en place une dynamique, mais cette dynamique est indispensable, elle permet de passer des modèles conceptuels, systèmes de symboles qui sont de l'ordre des idées, aux modèles effectifs, systèmes de symboles physiques qui sont de l'ordre des faits.

Dans ce paradigme, le schéma précédent peut être développé de la manière suivante :

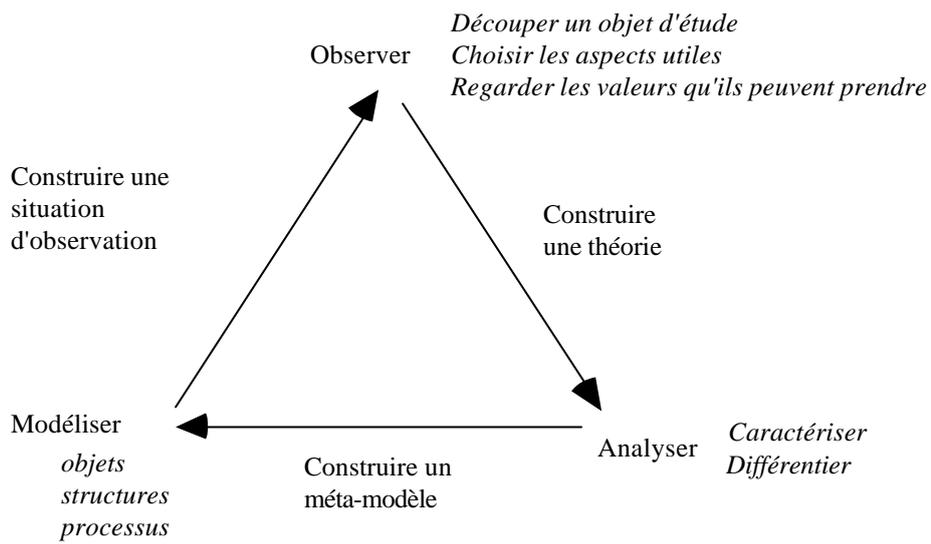


figure 2 : La méthode expérimentale du point de vue systémique.

*Remarque 1* : les théories scientifiques sont provisoires, les changements de modèles et les changements de méta-modèle sont fréquents et ils ne soulèvent pas de controverse, seulement des discussions. Par contre, changer de théorie est très difficile, Le passage du Moyen Age à la Renaissance en offre de nombreux exemples, mais à notre époque la relativité, la double nature de la lumière, le modèle probabiliste de la matière à l'échelle de l'atome se sont heurtés à de grandes résistances parce qu'ils mettaient en cause les théories bien établies sur le monde.

*Remarque 2* : confondre méta-modèle et théorie est une déviation fréquente en intelligence artificielle où le recours à la logique (systèmes experts), aux réseaux connexionnistes (comme solution miracle) ont occulté à la fois les choix théoriques, les modèles conceptuels et la pratique expérimentale dont nous voudrions rendre compte ici.

## 2. QU'EST-CE QUI VALIDE/INVALIDE CHACUN DES ASPECTS D'UNE DEMARCHE EXPERIMENTALE ?

Dans le champ scientifique, la validité des modèles proposés est la question philosophique majeure, celle qui fait la différence avec d'autres productions de l'esprit comme l'alchimie ou la magie. Qu'est ce qui fait qu'un modèle est meilleur qu'un autre, qu'une théorie est meilleure qu'une autre, qu'une expérience est réussie ou sans intérêt ? Il faut définir des critères objectifs, partagés par une communauté pour en

décider. Avant d'en débattre dans ce dossier, proposons quelques jalons :

Le choix d'une théorie est un choix métaphysique qui engage un sujet ayant un projet ; sa validation interne réside dans la cohérence des hypothèses sur le rapport du sujet au réel. Mais en fin de compte, c'est l'ensemble du processus expérimental qui fait qu'une théorie se maintient, doit évoluer ou changer. Ce qui est intéressant et curieux, c'est de retrouver à travers l'évolution des sciences la permanence des oppositions théoriques fondamentales (comme entre Platon et Aristote) sous des formes variées, à diverses époques.

Le choix du méta-modèle est de nature mathématique et sa validation interne réside dans sa rationalité. Il provient d'une abstraction des régularités des phénomènes observés.

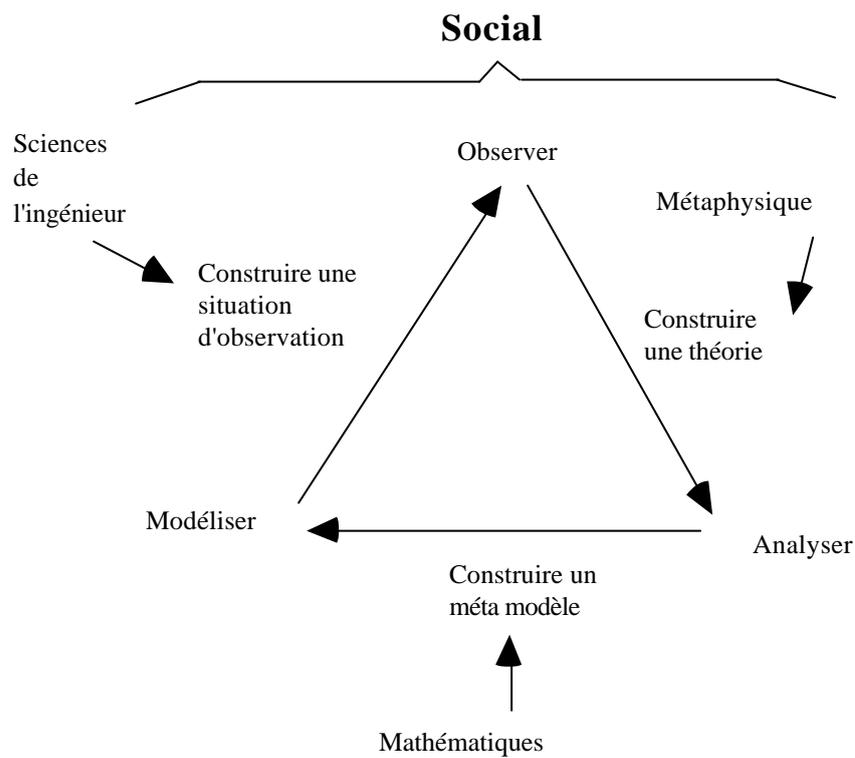


figure 3 : D'où provient la validité de la démarche scientifique

La construction des outils et des expériences relève des sciences de l'ingénieur et l'habileté mentale et artisanale du chercheur va y jouer un rôle fondamental. L'état des techniques joue un rôle majeur dans les expériences réalisables à une certaine époque, mais n'empêche pas de les projeter de manière imaginaire (Léonard de Vinci, Jules Verne,

Babbage...). Jean Lassègue reviendra sur ce point et montrera comment l'intelligence artificielle a transformé les expériences imaginaires en leur donnant un statut objectif.

Enfin l'ensemble du processus de recherche, avec ces trois niveaux, est un système ayant sa cohérence propre, mais qui est plongé dans le social. Le consensus d'une communauté scientifique détermine souvent ce qui émerge comme essentiel des activités de recherche individuelles. Dans le même ordre d'idées, il faut aussi tenir compte des aspects politiques et économiques de soutien aux programmes de recherche.

### **3. QU'EST-CE QUE L'INFORMATIQUE CHANGE DANS LA METHODE EXPERIMENTALE ?**

Citons un certain nombre de points où il est évident que l'informatique a apporté des outils nouveaux :

— L'échelle de grandeur des calculs effectuels, la taille des données statistiques par exemple, est le point positif le plus évident en faveur de l'informatique.

— L'amélioration de la rigueur dans le prélèvement et le traitement des données, en particulier en améliorant la précision des observations. Le contrôle d'un des pôles de l'interaction dans les expériences de psychologie en est un exemple.

— La nécessaire formalisation des modèles : tout modèle implanté en machine doit utiliser un langage dont la syntaxe et la sémantique sont définies complètement. Les règles représentant l'activité du modèle doivent être toutes décrites, en détail et complètement, et leur application automatique évite des interprétations supplémentaires à celles qui sont décrites par la sémantique. L'article de Laurent Gosselin en donne un bon exemple.

— Le test de l'effectivité des modèles : peuvent-ils ou non être implantés sur un ordinateur et tourner dans un temps et un espace raisonnable relativement au processus qu'ils représentent ?

— La simulation : l'ordinateur est un outil d'observation de la cohérence interne des modèles et permet de faire des expériences de pensée "objectives". Ce point de vue est introduit par Jean Lassègue et développé dans l'article de Gaio Mauro.

De ce point de vue, l'informatique fonctionne comme science de l'ingénieur, au service des sciences cognitives, pour leur apporter de nouveaux instruments.

### **4. QU'EST-CE QUE L'INTELLIGENCE ARTIFICIELLE CHANGE DANS LA METHODE EXPERIMENTALE ?**

L'intelligence artificielle peut être vue comme une science de l'ingénieur, une spécialisation de l'informatique dont l'objet est la résolution des problèmes complexes, et en tant que science de l'ingénieur elle apporte des outils logiciels (Lisp, Prolog, Smalltalk, systèmes experts, réseaux connexionnistes...) aux sciences cognitives. Elle est aussi à l'origine des machines actuelles avec souris, écran graphique et fenêtres, dont le premier exemple a été la machine Lisp de Xerox.

Mais l'intelligence artificielle forte ne se limite pas à la création d'outils et de modèles, même si cet aspect est très important. Elle se donne pour but de concevoir et de construire des machines intelligentes, des machines qui pensent. Il s'agit de développer des analogues-machine des fonctions de l'esprit : mémoire, jugement, raisonnement, décision, conscience. Elle se place donc de ce point de vue dans les sciences cognitives. Elle y apporte une théorie sur la possibilité de faire des analogues-machine des principales fonctions de l'esprit : la résolution de problèmes, la mémoire, le langage, l'apprentissage, sans pour cela être astreinte à faire de la simulation du comportement humain. Concevoir des modèles effectifs des facultés mentales met en avant les théories constructivistes et systémiques.

La conception de ces logiciels ne peut pas suivre les étapes classiques : spécification, analyse, programmation et test, car le problème n'est pas entièrement connu à l'avance et c'est l'activité de conception et d'expérimentation qui va permettre de terminer la spécification. Ce phénomène est bien connu dans d'autres domaines comme la conception architecturale (Guéna et Nicolle, 1990) où le travail de conception fait par l'architecte termine la spécification donnée par le commanditaire. La méthode de conception que nous proposons pour les logiciels d'étude est la méthode de conception par amorce (Nicolle, 1994) qui consiste à construire d'abord très vite un logiciel très simple dans le domaine qu'on veut étudier, à l'expérimenter pour observer les différences entre le fonctionnement du logiciel réalisé et celui qu'on pourrait espérer, pour déduire de ces différences un modèle plus complet du phénomène qui servira de base à un autre logiciel, à peine plus complexe que le premier, et qui sera donc vite réalisé et expérimenté. Cette démarche de conception est illustrée par l'article de Denis Jacquet, Anne Nicolle et Marc Andrès. Développer des logiciels provisoires et expérimentaux permet d'observer ce qui est facile et difficile pour une machine sans idée préconçue, de tester des parties de logiciel qui seront réutilisées dans les projets ultérieurs, d'étudier les réactions des utilisateurs.

La programmation par objets, en facilitant la réutilisation et le développement incrémental du logiciel, est un bon support pour cette démarche de conception du logiciel par amorce. Mais il est important de mener des expériences pour confronter nos idées à la réalité à toutes les étapes de la conception de cette machine, y compris pendant l'étape de modélisation. Notre démarche consiste à utiliser une machine pour amorcer — au sens de la construction par amorce, cf. Pitrat, 1984 — cette amorce de machine intelligente entrant dans le dialogue pour se construire. La démarche de conception par amorce correspond à un déplacement de la question *Les machines peuvent-elles le faire ?* vers la question *Pouvons nous concevoir et réaliser des machines qui le feront ?*, car les machines ne sont pas des objets naturels qu'il faudrait observer et modéliser, mais des artefacts qu'il faut concevoir et construire. Il devient alors difficile d'imposer des limites *a priori* à ce que l'homme peut concevoir et construire.

Le problème de l'amorce d'une machine intelligente se trouve ainsi posé ainsi au niveau du processus de conception lui-même, dans sa dynamique. La machine devant pouvoir se modifier elle-même par sa propre expérience, il est nécessaire que le modèle soit réflexif, c'est-à-dire contienne une description de son propre comportement, et que cette description soit en liaison de dépendance causale avec le modèle. L'expérimentation, clé du voûte du processus, ne correspond pas vraiment aux étapes de test et de validation bien connues des informaticiens. Elle se place à un autre niveau, celui de la validation du modèle, ou plutôt de la mise en évidence de ses forces et de ses faiblesses, et suppose, pour être menée dans des conditions scientifiques, la collaboration avec d'autres disciplines expérimentales.

##### **5. QU'EST-CE QUE L'INTELLIGENCE ARTIFICIELLE PEUT APPRENDRE DES AUTRES SCIENCES COGNITIVES EN CE QUI CONCERNE L'EXPERIMENTATION ?**

Au cours des trente dernières années, l'interaction majeure entre l'intelligence artificielle et les autres sciences cognitives a été l'emprunt de concepts et de modèles. Ils passent d'un domaine à l'autre et reviennent ensuite à leur point de départ, sans qu'on sache qui en a vraiment la paternité : réseaux sémantiques, *frames*, systèmes de production, en sont des exemples. Dans d'autres situations, l'une des disciplines joue un rôle de service par rapport à l'autre.

Le projet Compèrobot est un exemple d'interaction beaucoup plus large, avec la construction commune d'expériences, de théories, de modèles, d'outils, à travers lesquels chacun fait le travail propre à sa

discipline (Nicolle & Vivier, 1996). Les apports sont réciproques et les rétroactions dans les deux sens. Jean Vivier décrira plus en détail le processus qui s'est mis en place et les avancées qui en ont résulté du point de vue de la psychologie. Du point de vue de l'intelligence artificielle, en me fondant sur cette expérience de travail interdisciplinaire, je pense que les points principaux à apprendre des sciences ayant une longue pratique de l'expérimentation, comme la psychologie cognitive, sont l'importance et la difficulté de l'observation, l'organisation rationnelle de l'expérimentation, la prudence dans les conclusions. Ces points sont illustrés dans l'article Denis Jacquet, Anne Nicolle et Marc Andrès sur le projet Compèrobot.

Sans doute reste-t-il beaucoup de choses à dire sur les rôles réciproques de l'intelligence artificielle et des autres sciences cognitives dans l'expérimentation. Notre but n'est pas l'exhaustivité mais, en rendant compte d'expériences interdisciplinaires, d'amorcer un processus de dialogue sur ces questions entre les chercheurs des différentes disciplines pour faire varier et améliorer leurs interactions.

### ***Bibliographie***

- Davis R., Shrobe H., Szolovits P. (1993) What is a knowledge representation ? *AI Magazine*, 1.
- Guéna F., Nicolle A. (1990) Réutilisation de solutions génériques pour résoudre des problèmes de conception. *Sur la modélisation des processus de conception créative*. Paris, Europa, p. 29-48.
- Lakoff G. et Mark J. (1980) *Les métaphores dans la vie quotidienne*. Paris Éditions de Minuit, 1985, pour la traduction française.
- Lenat D. (1982) AM: An Artificial Intelligence Approach to Discovery in Mathematics as Heuristic Search, *Knowledge-Based Systems in Artificial Intelligence*, New York McGraw-Hill.
- Mugur-Schächter M. (1989) Esquisse d'une méthode générale de conceptualisation relativisée, *Arguments pour une méthode, autour de Edgar Morin*, Paris Seuil.
- Nicolle A. (1994) Conception du logiciel par objets : réification, réflexivité et amorce, Aix en Provence, Actes de la Journée d'étude du GRASCE, *De l'Objet au Système*.
- Pitrat J. (1984) *Maciste : un système qui utilise des connaissances pour utiliser des connaissances*, Paris, publication n° 49 du GR 22 du CNRS.