# Pour une approche complémentaire de l'IA traditionnelle :
## LE MANIFESTE ANIMAT

**Jean-Arcady Meyer**

*Directeur de recherche CNRS,*
*AnimatLab, Département de Biologie*
*Ecole Normale Supérieure*
*46, rue d'Ulm — 75230 Paris Cedex 05*
*Email : meyer@wotan.ens.fr*

L'approche animat [Meyer, 1995, 1996a, 1996b] vise à synthétiser des animaux simulés, ou des robots, dont les lois de fonctionnement sont inspirées de celles des animaux réels. Son objectif à court terme est de comprendre les mécanismes qui permettent aux animaux de s'adapter et de survivre dans des environnements plus ou moins changeants et imprévisibles, à des fins à la fois fondamentales et appliquées. Son objectif à plus long terme est de contribuer à l'avancement des sciences cognitives en recherchant en quoi l'intelligence de l'homme s'explique à partir des comportements adaptatifs les plus simples hérités de l'animal, dans une perspective à la fois évolutionniste et ascendante relativement négligée à ce jour. Pour atteindre ces objectifs, l'approche animat postule que les mécanismes sous-tendant les comportements adaptatifs doivent autant que possible être étudiés et conçus en situation, c'est-à-dire en tenant compte des buts que l'animal ou le robot essaie d'atteindre, et des opportunités que l'environnement lui offre pour réaliser ces buts. De ce fait, l'approche animat est une entreprise essentiellement interdisciplinaire, à l'interface entre les neurosciences, les sciences cognitives, l'éthologie, et l'écologie, d'une part, l'informatique et la robotique, d'autre part.

Les animats sont des animaux simulés sur ordinateur, ou des robots réels, qui s'avèrent capables de "survivre" ou d'assurer leur mission dans un environnement éventuellement imprévisible ou menaçant. Pour ce faire, les animats doivent maintenir dans leur zone de viabilité les variables essentielles à cette survie ou à cette mission [Ashby, 1952]. Ils doivent donc être équipés de senseurs et d'effecteurs adaptés aux problèmes que leur pose l'environnement, ainsi que d'une structure de contrôle qui leur permette de choisir quelle action accomplir à quel moment (Figure 1), en fonction des buts immédiats et lointains qu'ils essaient d'atteindre.

Dans ces conditions, le principal défi de l'approche animat est celui de décider quels senseurs et quels actuateurs doivent être reliés par quelle architecture de contrôle en vue de résoudre quel problème de survie et dans quel environnement.

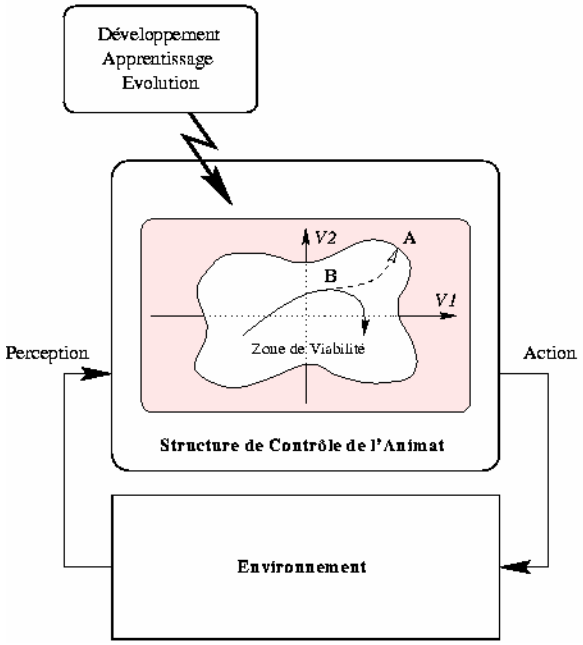


*Figure 1 — Un animat est "situé" dans un environnement réel ou simulé. Il a des senseurs qui captent des informations de l'environnement externe et/ou interne, des effecteurs qui lui permettent d'agir dans l'environnement, et une structure de contrôle qui coordonne ses perceptions et ses actions. Le comportement d'un animat peut être qualifié d'adaptatif tant que sa structure de contrôle permet de maintenir ses variables essentielles (par exemple, V1 et V2) dans leur zone de viabilité. Ici, une action correctrice a été accomplie au point B, de façon à éviter de quitter la zone de viabilité au point A. Dans la mesure où cette structure de contrôle sert aussi à choisir les buts successifs que l'animat cherche à atteindre ou à arbitrer entre des buts conflictuels, elle joue le rôle d'un système motivationnel. L'organisation de la structure de contrôle peut être modifiée par des processus de développement, d'apprentissage ou d'évolution.*

Il devrait être clair à ce stade que les objectifs fondamentaux et appliqués de l'approche animat recoupent ceux de l'Intelligence Artificielle (IA) traditionnelle [Barr & Feigenbaum, 1981] en bien des points. Cependant, dans leur tentative commune

d'appréhender les ressorts de l'intelligence humaine, l'approche animat et l'IA se distinguent par de nombreux aspects stratégiques ou méthodologiques que résume le Tableau 1 (page suivante).

En particulier, il apparaît que la plupart des travaux en IA ont trait à la simulation des facultés les plus élaborées du cerveau humain - résolution de problèmes, compréhension du langage naturel, raisonnement logique, par exemple. Ces simulations mettent en œuvre des programmes informatiques qui implémentent une approche computationnelle de la cognition [Newell & Simon, 1976]. Une telle approche implique le traitement de "physical symbol tokens", c'est-à-dire de divers objets ou patterns physiques (comme, par exemple, l'assemblage des moments magnétiques des diverses particules constituant la mémoire d'un ordinateur) qui représentent quelque chose d'autre qu'eux-mêmes (des objets dans l'environnement, des concepts, des désirs, des émotions, etc.), au moyen de règles syntaxiques opérant sur les caractéristiques physiques de ces symboles.

*Tableau 1 — Différences méthodologiques et stratégiques entre l'Approche Animat et l'IA traditionnelle*.

| **Approche Animat** | **Intelligence Artificielle traditionnelle** |
|---|---|
| *Approche ascendante* — Etude des capacités adaptatives basiques des animaux et des robots (e.g. navigation, recherche de nourriture, évitement des prédateurs). | *Approche descendante* — Etude des capacités cognitives élaborées propres à l'homme (e.g. résolution de problèmes, raisonnement logique, compréhension du langage naturel). |
| *Systèmes qui agissent* pour résoudre un problème donné. | *Systèmes qui utilisent leurs connaissances* pour résoudre un problème donné. |
| Application à des *problèmes généraux* et éventuellement *mal définis.* | Application à des *problèmes limités* et *bien définis* |
| Structures *décentralisées* et calculs *parallèles* | Contrôle *centralisé* et calculs *séquentiels* |
| Fonctionnalités *émergentes* | Fonctionnalités *programmées* |
| *Systèmes ouverts* — Agents autonomes qui interagissent en temps réel avec un environnement imprévisible, sans intervention humaine. Approche *située* | *Systèmes fermés* — Environnements contrôlés par l'expérimentateur qui généralement fournit les données d'entrée. |
| Intérêt majeur pour les problèmes de *coordination perceptions / actions* | Intérêt limité pour la *perception* et l'*action* |
| Intérêt majeur pour les mécanismes adaptatifs liés à la *maturation*, au *développement*, à l'*apprentissage* et à l'*évolution*. | Intérêt limité pour l'*apprentissage* |
| Intérêt majeur pour les causes *"ultimales*" (le pourquoi) et les causes "*proximales*" (le comment) des *comportements adaptatifs*. | Intérêt majeur pour les causes "*proximales*" (le comment) des *comportements intelligents*. |

Le postulat fondamental de l'Intelligence Artificielle est qu'un tel système symbolique physique est à la fois nécessaire et suffisant pour produire des comportements intelligents. Dans cette conception, les interactions d'un programme avec son environnement sont assurées par des symboles d'entrée, qui sont fournis par l'utilisateur du programme, puis traités par un système centralisé de règles qui raisonne sur ces symboles indépendamment de leur signification, et enfin transformés en symboles de sortie, dont l'interprétation est laissée à l'utilisateur du programme.

Les simulations de l'IA se caractérisent de plus par le fait qu'elles portent sur des données soigneusement sélectionnées par le programmeur et sur des domaines d'application très restreints dont l'hétérogénéité et l'imprévisibilité sont aussi rigoureusement contrôlées que possible. Ainsi, le système expert MYCIN [Shortliffe, 1976], spécialisé dans le diagnostic des maladies infectieuses du sang, est bien incapable de s'apercevoir que tel malade va mourir parce qu'une de ses artères a été tranchée. De même, le générateur de plans GPS [Newell & Simon, 1972] - qui a pourtant vocation à être un "General Problem Solver" - ne s'avère capable de résoudre un problème que si ce problème lui est posé par un humain dans un formalisme qui favorise la découverte de la solution correspondante.

Malgré la production de certains programmes aux performances remarquables, il paraît maintenant clair [Putnam, 1988] que l'IA n'a pas autant fait avancer notre compréhension de l'intelligence naturelle que d'aucuns l'ont prétendu ou espéré. Cela est probablement dû au fait que ces programmes ont été centrés sur la production de compétences isolées, négligeant ainsi le fait que les êtres vivants sont toujours situés dans un environnement, avec lequel ils interagissent en permanence par l'intermédiaire de leurs senseurs et de leurs effecteurs. Cela est probablement dû aussi au fait que l'accent mis par l'IA sur les processus élaborés mis en œuvre dans l'exploration de graphe ou le raisonnement logique, par exemple, a conduit à négliger les processus de perception, de catégorisation et d'adaptation qui sont primordiaux dans une approche située. Il s'en est suivi divers problèmes théoriques ou pratiques, tels que le "symbol grounding problem" (comment les symboles manipulés par le programme sont-ils liés aux stimuli externes? [Harnad, 1990], le "frame problem" (comment distinguer à tout moment les stimuli externes importants de ceux qui ne le sont pas? [McCarthy & Hayes, 1969], ou le "brittleness problem" (comment éviter la fragilité d'un programme lorsqu'il est utilisé dans un champ d'application plus vaste que celui pour lequel il a été conçu? [Duda & Shortliffe, 1983].

Plutôt que d'immerger dans des environnements simplifiés à l'extrême des programmes qui n'exhibent qu'une compétence avancée particulière et qui utilisent des données qui préjugent les solutions à découvrir, l'approche animat vise à modéliser des organismes simples, mais capables d'interagir avec un environnement aussi proche du réel que possible, et dans lequel ces organismes doivent poursuivre plusieurs buts parfois contradictoires, comme se nourrir, se reproduire, échapper à leurs prédateurs, etc. Pour survivre, ces organismes ne peuvent utiliser que les informations que leur fournissent leurs senseurs, informations qu'ils doivent parfois activement rechercher et qu'ils doivent toujours interpréter en fonction de l'environnement rencontré et du but poursuivi. Ainsi, leurs structures de contrôle mettent à l'épreuve les conceptions de James, selon lesquelles

> *le cerveau n'est pas construit pour penser de façon abstraite, il est construit pour assurer notre survie dans le monde* [James, 1890].

Souvent, de plus, les animats peuvent améliorer leurs facultés d'adaptation par un apprentissage individuel, ou par suite d'un processus évolutif impliquant plusieurs générations successives. En ces divers points, l'approche animat s'inspire des travaux récents sur les capacités cognitives de l'animal [Roitblat & Meyer, 1995] et met en œuvre des outils informatiques inspirés de la biologie - tels que les réseaux de neurones [Durbin et al., 1989] ou les algorithmes génétiques [Holland, 1975] - ou inspirés des sciences cognitives - comme les systèmes de classeurs [Holland et al., 1986].

Le postulat fondamental de l'approche animat est qu'il est possible d'aborder l'étude de la cognition humaine par une approche ascendante - ou bottom-up - qui, partant d'architectures de contrôle minimales et d'environnements simples, consiste à complexifier progressivement ces architectures et ces environnements. Si on prend soin de n'ajouter à ces architectures que des mécanismes indispensables qui préservent leurs capacités de base à percevoir, à catégoriser et à poursuivre des buts générés de manière autonome, il devrait être possible de résoudre des problèmes de survie de plus en plus complexes, sans perdre pour autant l'aptitude à en résoudre de plus simples. A long terme, on peut espérer ainsi comprendre quel est le rôle adaptatif de la cognition, en précisant à partir de quels problèmes et pour quels environnements le recours à des processus cognitifs - et notamment la manipulation de symboles - s'avère indispensable à la survie. Plus généralement, on peut espérer ainsi comprendre par quels mécanismes évolutifs les capacités cognitives de l'homme découlent des capacités adaptatives d'animaux plus primitifs et pourquoi il a fallu beaucoup plus de temps aux systèmes vivants sur terre pour apprendre comment survivre dans un environnement difficilement prévisible que pour apprendre à manipuler des symboles. Comme l'écrit Brooks,

> *cela suggère que la résolution de problèmes, le langage, l'acquisition et l'utilisation de connaissances, sont toutes des facultés relativement simples à acquérir sitôt que celles d'exister et de réagir ont été mises en place* [Brooks, 1991].

Plusieurs synthèses décrivant les résultats de l'approche animat sont disponibles [Meyer & Guillot, 1991; Meyer & Wilson, 1991; Meyer et al., 1993; Cliff et al., 1994; Meyer & Guillot, 1994; Maes et al., 1996). Elles démontrent que, si certains animats se contentent de mettre en œuvre de simples réactions stimulus - réponse pour résoudre des problèmes de survie relativement triviaux, d'autres s'avèrent capables de rechercher activement toutes informations utiles sur les objets et les relations causales qui caractérisent leur environnement, puis d'utiliser ces représentations internes de façon flexible et intelligente. Elles démontrent également que les animats sont capables de se déplacer dans leur environnement, d'éviter les obstacles ou de saisir des objets qui s'y trouvent, et qu'ils peuvent naviguer vers un but donné. Les animats peuvent aussi interagir, et même communiquer, afin de résoudre collectivement des problèmes complexes. Ils peuvent même évoluer, se développer, apprendre, mémoriser, et générer des plans.

## Références bibliographiques

**[Ashby, 1952]** Ashby, W. (1952). *Design for a Brain*. Chapman and Hall.

**[Barr & Feigenbaum, 1981]** Barr, A. and Feigenbaum, E. (1981).*The Handbook of Artificial Intelligence*. Pitman. Brooks, R. (1991). Intelligence without representation. *Artificial Intelligence*, 47:139-159.

**[Cliff et al., 1994]** Cliff, D., Husbands, P., Meyer, J.A. and Wilson, S. (1994). *From Animals to Animats 3. Proceedings of theThird International Conference on Simulation of Adaptive Behavior*. The MIT Press/Bradford Books.

**[Duda & Shortliffe, 1983]** Duda, R. and Shortliffe, E. (1983).Expert systems research. *Science,* 220:261-268. Durbin, R., Miall, C., and Mitchison, G. (1989).*The Computing Neuron*. Addison Wesley. Harnad, S. (1990).The symbol grounding problem. *Physica D*., 42:335-346.

**[Holland, 1975]** Holland, J. (1975). *Adaptation in Natural and Artificial Systems*.The University of Michigan Press.

**[Holland et al., 1986]** Holland, J., Holyoak, K., Nisbett, R., and Thagard, P. (1986). *Induction. Processes of Inference, Learning and Discovery*.The MIT Press.

**[James, 1890]** James, W. (1890). *Psychology (briefer course).* Holt.

**[Mc Carthy & Hayes, 1969]** McCarthy, J. and Hayes, P. (1969). Some philosophical problems from the standpoint of artificial intelligence. In Michie, D. and Meltzer, B. (Eds.), *Machine Intelligence 4*. Edinburgh University Press.

**[Maes et al., 1996]** Maes, P., Mataric, M., Pollack, J., Meyer, J.A. and Wilson, S. (1996). *From Animals to Animats 4. Proceedings of the Fourth International Conference on Simulation of Adaptive Behavior*. The MIT Press/Bradford Books.

**[Meyer, 1995]** Meyer, J.A. (1995). The Animat Approach to Cognitive Science. In Roitblat, H. and Meyer, J.A. (Eds.). *Comparative Approaches to Cognitive Science.* The MIT Press/Bradford Books.

**[Meyer, 1996a]** Meyer, J.A. (1996a). Artificial Life and the Animat Approach to artificial Intelligence. In Boden, M. (Ed.). *Artificial Intelligence.* Academic Press.

**[Meyer, 1996b]** Meyer, J.A. (1996b). From Animals to Animats: Adaptive Behavior in Biomimetic Artifacts. *Robotics and Autonomous Systems.* Sous presse.

**[Meyer & Guillot, 1991]** Meyer, J.A. and Guillot, A. (1991). Simulation of adaptive behavior in animats: review and prospect. In Meyer, J.A. and Wilson, S. (Eds.) >*From Animals to Animats. Proceedings of the First International Conference on Simulation of Adaptive Behavior.* The MIT Press/Bradford Books.

**[Meyer & Guillot, 1994]** Meyer, J.A. and Guillot, A. (1994). From SAB90 to SAB94: Four years of animat research. In Cliff, D., Husbands, P., Meyer, J.A. and Wilson, S. (Eds). *From Animals to Animats 3. Proceedings of theThird International Conference on Simulation of Adaptive Behavior.* The MIT Press/Bradford Books.

**[Meyer et. al, 1993]** Meyer, J.A., Roitblat, H., and Wilson, S. (Eds.) (1993). *From Animals to Animats 2. Proceedings of the Second International Conference on Simulation of Adaptive Behavior.* The MIT Press/Bradford Books.

**[Meyer & Wilson, 1991]** Meyer, J.A. and Wilson, S. (Eds.) (1991). *From Animals to Animats. Proceedings of the First International Conference on Simulation of Adaptive Behavior.* The MIT Press/Bradford Books.

**[Newell & Simon, 1972]** Newell, A. and Simon, H. (1972). *Human Problem Solving.* Prentice-Hall.

**[Newell & Simon, 1976]** Newell, A. and Simon, H. (1976). Computer science as empirical enquiry. In Haugeland, J. (Ed.), *Mind design.* The MIT Press.

**[Putnam, 1988]** Putnam, H. (1988). Much ado about not very much. In Graubard, S. (Ed.), *The Artificial Intelligence Debate. False Starts, Real Foundations.* The MIT Press.

**[Roiblat & Meyer, 1995]** Roitblat, H. and Meyer, J.A. (Eds.) (1995). *Comparative Approaches to Cognitive Science.* The MIT Press/Bradford Books.

**[Shortliffe, 1976]** Shortliffe, E. (1976). *Computer Based Medical Consultations: MYCIN.* Elsevier.