



**POLYTECH<sup>®</sup>**  
NANTES

Par :

CATRY Olivier  
MEUNIER Vincent

Enseignante responsable :  
Véronique Guy

# Robot : être-vivant ?

18 Décembre 2009





**POLYTECH<sup>®</sup>**  
NANTES

Par :  
CATRY Olivier  
MEUNIER Vincent

Enseignante responsable :  
Véronique Guy

# Robot : être-vivant ?

18 Décembre 2009

## Table des matières

Introduction.....	5
I- Le robot actuel.....	6
A Le robot exécute ce qui lui est ordonné.....	6
B Un schéma toujours vérifié aujourd'hui.....	6
C Approche du robot pensant par la science cognitive.....	6
II- Le robot pensant.....	7
A Le robot pensant selon deux théories.....	7
B Définition de la pensée.....	8
C Construire le robot sur la pensée platonicienne ?.....	8
D Le robot concevant sa propre vision de l'univers.....	8
III- Le robot conscient.....	9
A La perception de soi, chez le robot.....	9
B Approche de l'émotion robotique.....	9
C Le robot et ses émotions.....	9
D Des émotions qui doivent être créées par le robot lui-même.....	10
E Approche de la morale.....	11
F La morale inculquée artificiellement à la machine.....	12
G La morale apprise par le robot.....	12
Conclusion.....	13
Bibliographie.....	14

# Introduction

## *Machine*

**Appareil** ou ensemble d'appareils capable d'effectuer un certain travail ou de **remplir une certaine fonction**, soit **sous la conduite** d'un opérateur, soit **d'une manière autonome**.

*Dictionnaire Larousse*

Le tournevis est devenu le complexe ordinateur. La roue est désormais la rapide voiture. La lance s'est transformée en redoutable missile. Il ne fait aucun doute, l'outil a suivi une évolution parallèle à celle de l'homme. Cette évolution ne semble pas avoir de limite aujourd'hui, si l'on en croit les différents pôles de recherche actuels en physique, en chimie, ou encore en mathématiques. Mais c'est en informatique que se pose actuellement de nombreux problèmes d'ordre philosophique : cet outil anodin qu'est l'ordinateur pourrait-il, un jour, dépasser l'homme dans le domaine qui a fait de lui le roi des animaux ? La machine pourrait-elle un jour devenir un être-vivant intelligent, conscient et acquérir des droits ? La déclaration universelle des droits de l'homme et de la machine serait-elle l'aboutissement future de cette évolution sans-cesse croissante ? Projétons-nous alors dans un futur possible ...

**A quel point l'existence d'une machine pensante et consciente remettrait-elle en cause la définition d'être-vivant ?**

Il n'est que trop dangereux de répondre à cette interrogation au pied de la lettre, car la définition même d'être vivant est à l'origine de nombreux débats sur l'ensemble du globe et de l'histoire de l'homme. Or, d'un accord universel, l'homme se définit lui-même « être-vivant ». Ainsi, sera-t-on plus enclin à apporter un schéma de réponse en rapprochant la machine de l'homme.

Pour cela, il faudra présenter le robot à son dernier stade d'évolution actuel, puis imaginer une évolution possible de la machine et comparer ses caractéristiques à celles de l'homme, d'abord sur le plan de la pensée, puis sur le plan de la conscience. Cette projection de l'avenir sera étudiée par rapport à celle de nombreux ingénieurs et chercheurs en informatique, mais aussi d'écrivains et de personnes éloignées de ce domaine.

## I- Le robot actuel

### A *Le robot exécute ce qui lui est ordonné*

L'intelligence du robot actuel se résume à une série d'instructions. L'intelligence artificielle telle qu'elle existe aujourd'hui est en effet le résultat de procédures écrites à l'avance par l'homme ou la machine. Si cette dernière parvient à retenir une leçon de ses agissements, alors cette capacité d'apprentissage a été écrite auparavant par l'homme. Il n'est pas encore possible qu'une machine décide par elle-même d'étudier un résultat donné, c'est à dire, sans que cette réaction ne soit une conséquence directe ou indirecte d'un travail de l'homme. C'est pourquoi le terme utilisé aujourd'hui pour décrire cette aptitude mécanique qu'ont nos outils à réaliser des calculs d'un si haut niveau qu'ils semblent ne pas avoir été prévus par l'homme est l'intelligence dite « artificielle ».

### B *Un schéma toujours vérifié aujourd'hui*

Cet état de l'avancée en intelligence artificielle est loin des prévisions que l'on en faisait par le passé. Les plus grands penseurs imaginaient le robot capable de communiquer avec l'homme, avant de parvenir à se confronter à lui sur des problèmes de raisonnement. L'exemple le plus célèbre est HAL 9000, l'ordinateur de bord de la station du film *2001 : L'odyssée de l'espace*, de Stanley Kubrick, capable d'entretenir un dialogue complexe avec l'homme. Pourtant, il est à l'origine de mauvaises décisions liées au bon fonctionnement de la station (mort d'un astronaute accidentelle).

Will Wright dans la préface de *Robots, genèse d'un peuple artificiel* de Daniel Ichbiah

Je ne pouvais pas croire que HAL pouvait battre un humain aux échecs, mais on était persuadé qu'une machine pensante devait réussir à soutenir facilement une conversation. Depuis, nous avons découvert que c'est l'inverse.

*Le cas « HAL », une projection futuriste de l'I.A.*

### C *Approche du robot pensant par la science cognitive*

Les sciences cognitives sont les disciplines scientifiques qui visent à étudier et comprendre les mécanismes de la pensée. Depuis le milieu du XXe siècle, ces sciences se sont développées et deux notions importantes sont nées : le **cognitivisme** et le **connexionnisme**. Selon le premier, l'esprit est étroitement lié à une machine, capable de traiter l'information à l'aide d'un jeu complexe d'entrées et sorties. Mais, il implique que l'information soit traitée à l'aide de symboles, et non de leur signification. Le **connexionnisme** s'oppose à cette seconde hypothèse : ce ne sont pas des symboles qui permettent de traiter l'information mais une entité parallèle formée de groupes neuronaux qui, à un instant donné, représentent l'information au cerveau.

## II- Le robot pensant

### A *Le robot pensant selon deux théories*

Ainsi, selon le dogme du cognitivisme, le robot, selon l'architecture électronique qui nous lui donnons aujourd'hui, pourra avoir un esprit et donc être capable de penser. Selon le connexionnisme, il faut revoir l'architecture du robot, dans le but de créer non pas une machine, comme nous la définissons à présent, mais un esprit pensant.

#### Hypothèse du système des symboles physiques émise par Simon et Allen Newell

"un système physique est capable d'un comportement intelligent (intelligent étant pris dans le sens humainement intelligent) s'il s'agit d'un système de symboles physiques."

"un système de symbole physique est une machine qui produit dans le temps un assemblage évolutif de structures de symboles"

*Une approche cognitiviste de la machine pensante.*

#### Intelligence artificielle et informatique théorique (auteurs : cf bibliographie)

Croire en la possibilité de machines intelligentes ou raisonnant "comme un être humain" relève de l'acte de foi et ne repose sur aucun fondement scientifique. La thèse contraire n'est d'ailleurs pas plus démontrable, tout au plus peut-on penser que si jamais on parvient à réaliser une forme d'intelligence artificielle, elle s'appuiera sur un matériel et sur des techniques fondamentalement différentes de celles que nous connaissons aujourd'hui.

*Une approche connexionniste de la machine pensante.*

## **B Définition de la pensée**

Pour comprendre une telle démarche, il faut tenter une approche de la définition de la pensée. C'est durant l'antiquité que la recherche a été effectuée en opposant deux conceptions différentes :

- L'aristotélicienne, moniste, dont le raisonnement est le suivant : l'homme a une perception du réel à partir de l'unique réel et de ses lois.
- La platonicienne, dualiste, dont le raisonnement est le suivant : l'univers est une conception d'idées universelles auxquelles l'homme accède par la pensée.

La conception de la machine pensante doit alors s'accorder à l'une ou l'autre de ces pensées. Ainsi, le robot doit-il penser à partir des connaissances uniques qui lui auront été suggérées, ou doit-il, à partir de rien, acquérir sa propre vision de l'univers ?

## **C Construire le robot sur la pensée platonicienne ?**

Arrivera-t-on alors à concevoir une telle machine pensante selon la théorie Platonicienne ? Selon les experts en intelligence artificielle, la machine saura penser lorsqu'elle sera capable d'apprendre une langue et de soutenir une conversation. Actuellement, les programmes de dialogue sont basés sur les connaissances et la vision humaines de l'univers, et sont capables de manier la langue selon des principes que l'homme est parvenu à lui inculquer. Mais, la machine ne comprend pas le sens de la conversation qu'elle tient. Alain Cardon dans son livre -- s'interroge alors sur ce fait : "Comment et sur quoi se réalise le processus qui conduit un système à générer une représentation de quelque chose, à engendrer une forme d'idée qui fait sens pour le système lui-même ?". C'est toute la recherche de nos jours en intelligence artificielle : amener la machine à s'interroger sur elle-même.

Will Wright dans la préface de *Robots, genèse d'un peuple artificiel*, de Daniel Ichbiah

Le robot doit concevoir par lui-même sa propre représentation du monde.

*La philosophie platonicienne de Will Wright sur le robot pensant.*

## **D Le robot concevant sa propre vision de l'univers.**

Le fait de concevoir sa propre vision du monde est à l'origine de l'esprit. La perception du monde propre à la machine l'oblige à reconnaître des notions dont elle ne peut percevoir la constitution, mais qu'elle peut diriger ou subir. C'est sur ce raisonnement qu'Alain Cardon explique l'existence possible de la machine consciente. La mémoire organisationnelle permet au robot de comprendre des notions complexes qu'il ne peut s'expliquer. L'esprit du robot n'est alors pas si loin de celle de l'homme, à condition que le robot soit un jour capable de concevoir son propre univers.

*De l'intelligence humaine à l'intelligence artificielle, Hugues Bersini*

la perception et, de manière générale, l'ensemble des processus sensori-moteurs (même les plus primitifs) se rajoutent comme contrées mentales à découvrir pour ces explorateurs de la pensée machine.

*Le robot concevant sa vision du monde, un intérêt particulier de la part des pionniers de l'intelligence artificielle.*

### **III- Le robot conscient**

#### **A *La perception de soi, chez le robot.***

Si le robot conçoit sa propre vision du monde, il devient alors capable de se percevoir lui-même. C'est un état de fait chez une entité qui est consciente de son univers. Faisant elle-même partie de cet univers, elle se perçoit elle-même et comprend alors que si elle disparaît, c'est sa perception du monde qui disparaît également. Le robot sait alors qu'il existe, car il pense de lui-même et n'obéit plus à des règles programmées sur la conduite de son existence, comme le ferait l'instinct de l'animal. L'émotion peut être la définition que l'on donnerait alors à cet état d'esprit et cette connaissance de lui-même.

#### **B *Approche de l'émotion robotique.***

Si le robot peut s'interroger sur lui-même, s'il est capable de ressentir sa propre existence, de comprendre le sens des célèbres paroles de Descartes "Je pense donc je suis", il se pose alors un nouveau problème : la machine devient-elle consciente ?

#### **C *Le robot et ses émotions***

Cet état d'esprit implique alors l'existence d'émotions chez la machine. En effet, la volonté de concevoir sa vision de l'univers est le résultat de nombreuses émotions, dont particulièrement la curiosité. Cette émotion est l'une des plus primitives. Le fait d'être conscient de son existence entraîne la volonté de vouloir vivre. La machine pensante pourrait donc exprimer de la colère à l'idée d'être détruite. Il est à noter l'usage du conditionnel, car de nombreux penseurs s'opposent à cette idée, sur lesquels nous reviendrons plus loin. A priori, rien n'empêche la possibilité de présence d'émotions chez la machine. Il reste à savoir quelles en sont les limites.

## **D Des émotions qui doivent être créées par le robot lui-même.**

Pourtant, sans intervention extérieure, certaines émotions ne seront pas présentes chez la machine. Ce sont ces émotions humaines que nous ne savons pas décrire, et que, tant que cet état de fait sera vérifié, nous ne pourrons programmer à une machine. Selon Rabelais : "Le rire est le propre de l'Homme". Si nous ne comprenons pas nous même la raison qui nous pousse à rire, si ce n'est une situation ou une idée que l'on définit abstraitement d'"humoristique", comment concevoir le rire chez la machine pensante et consciente ? Si nous poussons l'idée plus loin, il est tout à fait probable que la machine, ayant acquis un état de conscience suffisant, crée des émotions nouvelles, indescriptibles et inimaginables a priori par l'homme.

Lieutenant Commander Data



Le Lieutenant Commander Data est un androïde (un robot à l'apparence humaine) de la série télévisée *Star Trek : The Next Generation* de Gene Roddenbery. De nombreux épisodes s'interrogent sur son appartenance, ou non, aux êtres-vivants. Par ailleurs, Data doit prouver à mainte reprise son état de conscience. Il est extrêmement curieux et fasciné par les coutumes humaines et extra-terrestres. Mais un fait marquant est que Data ne parviendra jamais à saisir la notion d'humour.

*Data, un robot fondé sur les théories connexioniste et platonicienne.*

### Conscience artificielle et systèmes adaptatifs Alain Cardon

Ce sont des suscitations, des tendances vers l'expression des états conscients qui représentent l'acquis temporel, la mémoire du système. Le système a donc une mémoire organisationnelle qui se construit par son usage.

*Une approche de l'état de conscience pour la machine par les formes filaires et la mémoire*

## **E Approche de la morale.**

L'intelligence et la perception de soi justifient en partie la conscience. Mais, cet état de conscience n'est toujours pas complet. Un animal relativement évolué (animal de compagnie), peut comprendre le monde, et a probablement une perception de lui-même et de son existence au moment présent. Pourtant, il n'est pas un être doté de conscience. Lorsqu'il fait du mal à un autre être vivant, l'animal ne se rendra pas compte du mal qu'il aura occasionné, ou, du moins, n'en éprouvera aucun remord ultérieurement. Ce qui n'est pas le cas de l'homme. Les humains sont dotés d'un sens moral, qu'ils décident de s'expliquer par leur bonne ou mauvaise "conscience", par la religion, ou encore, qu'ils décident de ne pas s'expliquer du tout. Ainsi, la machine consciente est une machine dotée d'une morale, mais comment doit-elle être implémentée ?

### Le complexe de Frankenstein

Le complexe de Frankenstein est né en même temps que l'entrée de l'informatique dans notre vie de tous les jours. Il se traduit par deux craintes liées aux machines :

- La peur d'une erreur critique, d'un imprévu, qui entraînerait un dysfonctionnement du système.
- La peur d'une faille dans le système, dont la fonction serait alors détournée par un pirate ou toute entité agissant dans un but de nuisance.

Ce complexe a été étendu aux robots par de nombreuses œuvres de science-fiction. Il s'agit principalement du robot maléfique qui se retourne contre son créateur ainsi que l'espèce humaine.

Pour le chimiste Isaac Asimov, un robot est conçu par des ingénieurs, qui ne sont pas « ces savants-fous ou autres bidouilleurs » de machine. Le programme du robot est le résultat d'une profonde réflexion et de recherches en mathématiques extrêmement rigoureuses. C'est pourquoi il lui semble improbable qu'une machine puisse décider de se retourner contre un humain, puisque cela n'aurait alors rien de logique.

*La vision négative d'un robot dépourvu de morale.*

## **F La morale inculquée artificiellement à la machine**

Ne perdons pas de vue que la machine, aussi consciente soit-elle, est à la base conçue, directement ou indirectement par l'homme. Dans sa conception initiale, et récursivement, l'homme peut intégrer l'interdiction formelle de certaines concessions. Par exemple, peut-il interdire à la machine de faire du mal à l'humanité ? Si tel était le cas, cette machine ne pourrait pas aider à la conception d'une machine consciente capable d'éprouver de la haine envers l'humain. Le chimiste Isaac Asimov a exploité cette théorie au travers d'un grand nombre de romans, basés sur les trois lois de la robotique, qui sont programmées dans chaque machine pensante. Cette limitation est alors un frein à l'épanouissement de la conscience du robot car il ne peut concevoir lui-même la notion de bien et mal. Est-ce pour autant qu'il s'éloigne de l'humain ? Ce dernier, lui-même, ne parvient pas à savoir si sa conception du bien et du mal est le fruit d'une réflexion universelle ou une notion supérieure, inaccessible et régissant l'ensemble des êtres biologiques. Alors, sommes-nous en état d'affirmer que la manipulation des émotions de la machine la rendrait vraiment différente de l'ensemble des êtres-vivants ?

## **G La morale apprise par le robot**

Dans une autre mesure, le robot peut également concevoir sa propre morale. En effet, comme nous avons pu le voir, le robot conscient est capable d'apprendre, de comprendre, de retenir, et d'exprimer des émotions. Avec l'ensemble de ces caractéristiques, si le robot ne peut distinguer le bien du mal, il est appelé littéralement un "psychopathe". Pourtant, l'être humain, aux débuts de son évolution, ne devait pas être apte à distinguer clairement le bien du mal. Il est cependant inutile de s'extraire d'un contexte religieux pour corroborer cette distinction qui a dû être le résultat d'un long apprentissage, que nous pourrions même ne pas avoir achevé. De même pour le robot conscient, les premiers modèles peuvent ne pas être dotés du sens moral. Ils seraient donc un danger pour certaines espèces du globe, tout comme nous l'avons été. Puis, à mesure d'expériences et d'apprentissages, sur une longue durée, le robot conscient pourrait apprendre les principaux concepts du sens moral et en être affecté. N'est-ce pas l'évolution naturelle de l'être vivant conscient que nous sommes ?

Les trois lois de la robotique, *Manuel de la robotique - 58<sup>e</sup> édition*, Isaac Asimov

Première Loi : Un robot ne peut blesser un être humain ni, par son inaction, permettre qu'un humain soit blessé.

Deuxième Loi : Un robot doit obéir aux ordres donnés par les êtres humains, sauf si de tels ordres sont en contradiction avec la Première Loi.

Troisième Loi : Un robot doit protéger sa propre existence aussi longtemps qu'une telle protection n'est pas en contradiction avec la Première et/ou la Deuxième Loi.

*Les trois lois de la robotique, un exemple de la morale inculquée artificiellement à la machine.*

## Conclusion

Il est certain que l'intelligence artificielle n'a pas réalisé le bond que l'on escomptait il y a cinquante ans. De toute évidence, l'architecture même que l'on a donnée à la machine ne permet pas d'affirmer qu'il sera possible un jour qu'elle échappe au contrôle direct ou indirect de l'homme. La théorie connexionniste s'impose alors comme l'achèvement d'une architecture informatique permettant d'instaurer les bases de l'intelligence de la machine.

Cette architecture est sub-symbolique, elle se veut proche de celle du cerveau animal.

Par interpolation, il devient alors possible que la machine parvienne un jour à penser, tel que la théorie platonicienne l'entend. Le robot sera capable de saisir sa propre existence et de réaliser des tâches pour lesquelles il n'aura pas été préalablement programmé.

La pensée est un domaine propre à de nombreux animaux, dont l'homme.

Cette pensée, fort évoluée, est étroitement à l'origine de la conscience. Le robot, conscient qu'il existe et qu'il vit, conscient de l'essence du moment, éprouvera alors des émotions et deviendra une entité sensible dotée d'un sens moral. Il pourra alors agir dans son intérêt et dans celui des autres, sans obéir uniquement à ses sous-programmes.

La conscience est l'un des stades les plus hauts de l'évolution de l'homme.

Nous sommes sans-cesse à la recherche de nouvelles formes de vie, dans les profondeurs océaniques, ou à l'écoute de l'espace, sans pour autant pouvoir définir clairement cette notion de vie hors du schéma que nous connaissons. Admettons alors une telle projection du futur, une telle machine, si son architecture était proche du cerveau animal, si elle était capable de penser et de comprendre, si elle acquérait sa propre conscience.

Serait-ce un être vivant comme nous le définissons ?

Il s'agit a priori d'une nouvelle forme de vie.

Nous avons classé les êtres-vivants parmi les végétaux et les animaux. Peut-être est-ce une erreur de notre part. Peut-être devons-nous, dans le futur, les classer parmi les cellulaires et les électroniques. Cette perspective pourrait alors être chamboulée par une nouvelle interrogation proche de notre problématique initiale :

**A quel point l'existence d'une machine ayant les mêmes propriétés qu'une cellule remettrait-elle en cause la définition d'être-vivant ?**

## Bibliographie

BERSINI, Hugues - (Directeur du laboratoire d'intelligence artificielle IRIDA)

*De l'intelligence humaine à l'intelligence artificielle*

éd : Ellipses, 2006, ISBN 2-7298-2813-3

ALLIOT Jean-Marc - (Docteur en informatique)

SCHIEX Thomas - (Docteur en intelligence artificielle)

BRISSET Pascal - (Docteur en informatique)

GARCIA Frédéric - (Docteur en intelligence artificielle)

*Intelligence artificielle et informatique théorique (2e éd.)*

éd : Cepadues, 2002, ISBN : 2-85428-578-6

VARENNE Franck - (Docteur en histoire et philosophie des sciences)

*Qu'est ce que l'informatique ?*

éd : Vrin, 2009, ISBN10 2-7116-2178-2

CARDON Alain – (Professeur d'informatique à l'université du Havre)

*Conscience artificielle et systèmes adaptatifs*

éd : Eyrolles, 1999, ISBN 2-212-09124-9

CARDON Alain

*Modéliser et concevoir une machine pensante*

éd : Vuibert, 2004, ISBN10 2-7117-7152-0

ICHBIAH Daniel – (écrivain, journaliste, musicien)

*Robots : Genèse d'un peuple artificiel*

éd : Minerva, 2005, ISBN 2830707893

ASIMOV Isaac – (biochimiste, romancier)

*Le cycle des robots*

éd : J'ai lu, 2009, (dépôt 1976), ISBN 978-2-290-34248-0

RODDENBERRY Eugene Wesley – (scénariste, producteur)

SNODGRASS Melinda M. – (écrivaine)

*The Measure Of A Man, Star Trek : The next generation*, 1989