

# Ethique et robots

Erhard Taverna



Un logiciel est appelé agent intelligent ou robot logiciel quand le programme répond à plusieurs obligations, par exemple dans une application médicale. Le système doit avoir un fonctionnement autonome, cognitif, adaptatif, proactif (actif de sa propre initiative), réactif, communicatif et social (en liaison avec d'autres systèmes). En 1959, Alan Turing avait imaginé une situation test, mise en œuvre en 1966 par Weizenbaum, avec ELIZA, dans laquelle la plupart des sujets étaient restés toute une partie d'un entretien incapables de distinguer un partenaire réel, dans ce cas un psychothérapeute, d'une machine. Depuis 1991, le prix Loebner est décerné à tout programme capable de satisfaire aux critères du test de Turing pendant 25 minutes. Pour l'instant, nous nous sommes arrêtés au bronze, l'or n'ayant jamais été gagné. Ici comme ailleurs, la guerre est à l'origine de toute chose. Les progrès les plus avancés concernent actuellement les *Unmanned Combat Air Vehicles* armés, communément appelés drones. Les paroles de Markus Christen, chercheur au centre d'éthique de l'université de Zurich (Ethik-Zentrum der Universität Zürich, EZEN) ont été récemment citées dans le *Tagesanzeiger*: «A l'avenir, les drones dotés d'un équipement éthique pourront faire la guerre.» Cette hypothèse provocante est intéressante si on suit son raisonnement au fil de l'interview. Il déclare en effet que nous percevons de plus en plus le monde à travers des machines, ce qui modifie notre intuition morale et notre pensée éthique. Aux Etats-Unis, le débat est plus avancé, pour des raisons faciles à comprendre. L'un des spécialistes de l'éthique le plus souvent cité là-bas est Ronald C. Arkin, professeur au College of Computing de l'université de Géorgie. Il est considéré comme un des meilleurs experts en robots militaires et participe depuis des années à différents projets de recherche. Son ouvrage, *Governing Lethal Behavior in Autonomous Robots*, est une référence, pour les partisans comme pour les détracteurs, quand il s'agit d'équiper des systèmes d'une conscience artificielle (*artificial conscience*). Ses arguments sont précurseurs de l'évolution à long terme des robots intelligents, qui agiraient «avec plus d'éthique que les soldats humains sur le champ de bataille» (*behaving more ethically than human soldiers on the battlefield*). En Europe, on continue de débattre âprement pour savoir si les drones ne sont que des machines infernales de plus, contraires au droit de la guerre applicable, ou si au contraire ils humanisent les combats. Mais dans les pays producteurs comme les Etats-Unis, la Grande Bretagne ou Israël, cela fait long-

temps qu'on a tranché. Plus ou moins précis, à savoir générant beaucoup de «dégâts collatéraux», ces engins de la mort stimulent les applications civiles dans les systèmes de santé, la police, la protection de la nature et contre les catastrophes naturelles, les finances, l'agriculture et les expéditions diverses et variées.

Manifestement, l'évolution de l'intelligence artificielle IA est avancée au point de faire entrer la morale en jeu. La question intéressante est la suivante: un robot pourrait-il prendre une bonne décision, adaptée à la situation, indépendamment de nous? Notamment Amy Michelle DeBaets, professeur de bioéthique à l'université de médecine et de biologie de Kansas City, s'est exprimée sur cette question. Dans son travail, elle définit les concepts de base permettant de comprendre la morale de l'intelligence artificielle (*Artificial Moral Agency*). Elle consigne les conditions techniques et pose comme postulat une conscience rudimentaire de l'empathie nécessaire: «*Empathy, and not merely rationality, is critical for developing and evaluating moral choices; just as emotion is inherent to human rationality, it is necessary for machine morality.*» Elle met en garde contre l'idée selon laquelle les robots pourraient un jour agir de façon plus morale que les hommes. Elle nous amène à prendre conscience que personne n'est en mesure de prévoir les conséquences sociales à long terme, comme pour Internet. Malgré toutes ces objections fondamentales, elle pense que l'évolution de systèmes artificiels intelligents va se poursuivre. Elle espère que nous réfléchirons à temps aux conséquences du transfert de plus en plus fréquent de nos décisions à ces machines. Des systèmes d'agents intelligents toujours plus complexes marquent déjà notre quotidien, des places financières au déblaiement des zones contaminées par le nucléaire, en passant par le déminage. Les robots travaillent dans les usines et les laboratoires, servent de partenaire sexuel ou de gardien. Ils facilitent notre quotidien et, à long terme, menacent notre identité. Dans vingt ou trente ans, nous leur confierons peut-être aussi le soin de penser et d'avoir une morale à notre place. Pourtant, cela nous fait peur. Du moins pour l'instant.

## Références

- Christen M. *Tagesanzeiger*. Wissen. 26 novembre 2014.
- Arkin C. *Governing Lethal Behavior in Autonomous Robots*. Chapman and Hall, 2009.
- DeBaets AM. Can a Robot Pursue the Good? Exploring Artificial Moral Agency. *Journal of Evolution & Technology*. Sept. 2014. [http://jetpress.org/v24\\_3/DeBaets.htm](http://jetpress.org/v24_3/DeBaets.htm)