

international humanitaire, fixé principalement par les conventions de Genève de 1949, constitue la référence essentielle dans le débat et qu'on a donc pas à interdire les SALA capables de le respecter... ce qui est totalement illusoire et procède de la même erreur que celle commise par Asimov en croyant qu'on pourrait dire à nos robots : « Voici les règles que vous allez appliquer ».

Dans les romans et films de science-fiction, un thème revient régulièrement : la révolte des robots. C'est sans doute en pensant à elle qu'Asimov a souhaité introduire ses fameuses lois : il faut que les esclaves mécaniques restent sages et soumis.

Bien que les robots d'aujourd'hui soient totalement incapables de se rebeller, le problème de ces éventuelles futures révoltes est intéressant. Il est cependant beaucoup plus délicat qu'on ne le présente en général. D'abord se pose le problème des *bugs*. On le sait, il est quasiment impossible de fabriquer des systèmes informatiques sans y laisser des erreurs. Il est donc à prévoir qu'il y en aura dans les robots, même quand nous aurons été beaucoup plus loin qu'aujourd'hui.

## Des robots rebelles ou défectueux

Dès lors, comment éviter un *bug* qui conduirait un robot capable de tuer à s'attaquer systématiquement à tous les humains qu'il rencontre ? Un tel *bug*, par exemple dû à une ligne de programme effacée malencontreusement, doit-il être considéré comme une révolte ?

Sans doute pas. La question est en fait analogue à celle de la responsabilité chez les humains. Un fou n'est pas responsable, même quand il tue. Il faudrait donc définir ce qu'est la folie d'un robot. Sans doute qu'entre *bug* et révolte (au sens usuel), tout un continuum de situations seront possibles. Le véritable problème de la révolte des robots sera avant tout celui de cette distinction, plus que délicate. Avant qu'une armée de robots organisés et décidés à arracher le pouvoir aux humains se lève et nous attaque, nous aurons à faire face à

une multitude de situations embrouillées où le dysfonctionnement d'une machine aura engendré des blessures ou des morts, voire de graves catastrophes dont nous ne distinguerons pas la nature : panne stupide ou rébellion délibérée ?

Susan Anderson, de l'université du Connecticut, spécialiste reconnue des problèmes d'éthique des robots, se demande si une telle révolte en fait ne serait pas justifiée. Pour elle, le grand philosophe de la morale Emmanuel Kant avait abordé indirectement le problème des robots intelligents. Elle jette un regard nouveau sur les lois de la robotique d'Asimov : si des « robots » avaient les capacités à comprendre et à appliquer les lois de la robotique, aurions-nous le droit de les leur imposer ou d'en faire nos esclaves ?

Susan Anderson remarque que pour Kant, et pour notre société aussi, nous devons un certain respect aux animaux parce qu'en ayant ce respect, nous l'aurons les uns envers les autres. De la même façon, d'après Susan Anderson, si Kant avait réfléchi au problème des robots, il aurait défendu l'idée que le jour où ils existeront et seront capables de comprendre les lois de la robotique, nous devrons avoir du respect pour eux, ce qui nous interdira de leur imposer les lois de la robotique auxquelles il faudra donc renoncer.

L'idée n'est pas loin de celle défendue par le roboticien Hans Moravec, de l'université Carnegie-Mellon, selon laquelle les robots seront notre descendance et même s'ils prennent notre place, cela se fera avec notre accord, comme lorsqu'un vieillard confie à ses enfants la gestion de ses affaires.

Entre ces problèmes lointains traités avec délices par la science-fiction et que la science et la philosophie nous obligent à reconsidérer, et ceux immédiats de la juste évaluation de ce que sont vraiment nos robots actuels (tueurs ou non) et de ce qu'on doit en faire, se dessine une longue série de délicates questions éthiques et philosophiques dont certaines exigent un traitement urgent. Philosophes, éthiciens, logiciens, militaires et informaticiens doivent se parler et s'entendre.

## L'AUTEUR



J.-P. DELAHAYE est professeur émérite à l'université de Lille et chercheur

au Centre de recherche en informatique, signal et automatique de Lille (LIFL).

## BIBLIOGRAPHIE

S. Hawking *et al.*, *Autonomous weapons : An open letter from AI & robotics researchers*, 2015 : [http://futureoflife.org/AI/open\\_letter\\_autonomous\\_weapons](http://futureoflife.org/AI/open_letter_autonomous_weapons)

J. Galliot, *Military Robots, Mapping the Moral Landscape*, Ashgate, 2015.

R. Doaré *et al.*, *Drones et killer robots : faut-il les interdire ?* Presses universitaires de Rennes, 2015.

M. Englert *et al.*, *Logical limitations to machine ethics with consequences to lethal autonomous weapons*, 2014 [prépublication arXiv:1411.2842].

A. Krishnan, *Killer robots : Legality and ethicality of autonomous weapons*, Ashgate, 2009.

S. L. Anderson, *Asimov's "three laws of robotics" and machine metaethics*, *AI & Society*, vol. 22(4), pp. 477-493, 2008.



Retrouvez la rubrique  
Logique & calcul sur  
[www.pourlascience.fr](http://www.pourlascience.fr)