

Éthique et robotique humanoïde

En quête d'un nouveau pacte anthropomorphique

Denis Vidal (IRD, Paris)

Contrairement à ce qui se passe dans les débats qui portent sur les implications éthiques de la plupart des découvertes scientifiques ou des innovations technologiques, celui qui concerne les robots humanoïdes précède depuis toujours l'existence même des machines qui pourraient lui donner un semblant de réalité. De tous temps, en effet - comme en témoignent de très nombreux mythes et récits religieux - on s'est interrogé sur les implications éthiques qu'aurait pour l'humanité l'existence de créatures artificielles ressemblant de près ou de loin à des humains. Plus récemment aussi des myriades d'ouvrages et de films de science-fiction ont pratiquement décliné toutes les conséquences imaginables d'une telle invention. Aussi, est-ce cette dimension largement spéculative du débat qui explique, peut-être, le scepticisme ou l'indifférence de nombreux roboticiens à ce sujet. Mais si la majorité d'entre eux semblent estimer que les questions éthiques que posera peut-être un jour la robotique humanoïde, ne sont pas encore à l'ordre du jour, il y en a aussi qui considèrent - comme je l'ai montré - qu'un tel débat est déjà dépassé (R. Brooks, par exemple) ou encore qu'il repose sur des présupposés philosophiques largement erronés.

Aussi, plutôt que de m'interroger, à mon tour, sur les implications de la robotique dans un futur hypothétique, je proposais plutôt dans ce papier de mettre l'accent sur des considérations d'ordre éthique qui lui sont également associées mais qui sont plus directement d'actualité : Tel est le cas, en particulier, de celles qui découlent aujourd'hui, de la manière dont les roboticiens conçoivent le fonctionnement de leurs machines et dont ils cherchent concrètement à en définir l'usage. Partant d'une perspective ethnographique, j'ai choisi de mettre plus particulièrement l'accent, à l'occasion de cette réunion, sur le rôle joué par l'anthropomorphisme dans la conception des robots humanoïdes.

Aujourd'hui, en effet, une majorité des roboticiens qui conçoivent des robots humanoïdes cherchent à appliquer des recettes dont l'efficacité est plus ou moins vérifiée sur le plan empirique, pour susciter des réactions d'ordre anthropomorphique chez les utilisateurs de ces derniers. Dans la mesure où de telles réactions sont censées être largement inconscientes, le risque est cependant de voir ces dernières apparentées à des formes de manipulation psychologique qui peuvent susciter légitimement des inquiétudes sur le plan éthique. Or il n'est pas évident que de telles ambiguïtés se justifient, même sur un plan strictement méthodologique.

On doit laisser, en effet, aux psychologues, le soin de décider dans quelle mesure nous réagissons vis-à-vis de certains artefacts comme s'il s'agissait plus ou moins d'êtres humains. Mais il n'y a pas de raisons de supposer que l'importance centrale donnée à l'anthropomorphisme dans toutes les cultures humaines doive s'expliquer seulement - ou même, principalement - sur la base d'une logique sous-jacente liée à notre inconscient ou à notre système cognitif. Si l'on considère, en effet, les implications épistémologiques des recherches les plus récentes en psychologie dans ce domaine, le fait le plus significatif n'est pas tant que nous ayons tendance à confondre implicitement des artefacts avec des êtres humains dans certaines circonstances ; c'est plutôt le constat opposé : à savoir notre capacité très précoce - sinon véritablement innée - à opérer une telle distinction. Or un tel constat ne

contredit nullement l'importance culturelle de l'anthropomorphisme ou de son universalité ; mais il nous invite plutôt à le considérer sous un jour différent. Car le fait de ne pas confondre un objet ou un animal avec un être humain n'empêche nullement de se comporter à leur égard comme si on les considérait comme tels. C'est ce que font les jeunes enfants, quand ils jouent avec leurs jouets préférés. Et ce sont les implications sociales et culturelles d'une telle tendance dont il faut tenir compte, d'avantage, à mon sens, si l'on veut examiner les implications éthiques de l'anthropomorphisme dans le domaine de la robotique. Aussi, plutôt que de concevoir l'anthropomorphisme à la manière d'une sorte d' « inconscient » dont on pourrait faire usage pour susciter l'attachement des utilisateurs vis-à-vis des robots androïdes, de manière discutable moralement parce que fondamentalement 'asymétrique', je me demande s'il ne serait pas préférable - à la fois sur un plan éthique et d'un point de vue simplement méthodologique - d'envisager plutôt l'anthropomorphisme comme une sorte ' de **'pacte anthropomorphique'** que nous pourrions passer éventuellement avec des robots , à la manière dont nous le passons déjà, sans en faire d'ailleurs grand cas, avec toutes sortes d'objets ou d'animaux familiers.

La notion de 'pacte' permet d'abord, à mon sens d'insister sur le fait que – le plus souvent - le fait de traiter un objet de manière anthropomorphique ne résulte pas seulement d'un réflexe inconscient mais aussi d'un choix tout à fait délibéré. C'est ainsi que lorsqu'un enfant joue avec une poupée ou un animal en peluche, il ne s'illusionne nullement sur la nature véritable de ces derniers ; mais cela ne l'empêche pas pour autant de traiter provisoirement ce jouet à la manière d'un être vivant. Or, l'on pourrait décrire toutes sortes d'autres contextes sociaux et culturels où une même tendance est à l'oeuvre. J'en ai donné plusieurs exemples dans ma communication.

D'autre part , le fait d'employer le terme de 'pacte' permet également d'insister sur le fait que lorsque nous traitons un animal ou un objet de manière anthropomorphique, il s'agit d'une forme d'interaction qui ne relève pas simplement de la psychologie individuelle ou d'un réflexe cognitif inconscient. Nous nous comportons alors, en effet, le plus souvent comme si l'objet ou la créature en question ne sauraient être le réceptacle passif et indifférent de ce que nous projetons sur elle, et rester totalement indifférente à un tel traitement. Mais nous attendons aussi, bien souvent, de notre entourage qu'il partage jusqu'à un certain point notre manière de voir.

Aussi peut-on multiplier les exemples qui montrent que notre tendance à anthropomorphiser des artefacts peut bien découler dans une certaine mesure, de leurs caractéristiques propres ; mais qu'elle dépend aussi, dans une large proportion, de la sorte de 'pactes anthropomorphiques' que nous passons avec ces derniers, sur des bases qui peuvent être individuelles mais qui sont le plus souvent aussi collectives et d'ordre culturel.

Dés lors que l'on envisage d'avantage le rôle joué par l'anthropomorphisme dans une telle perspective, il devient plus aisé, à mon sens, de lever certaines des ambiguïtés qui persisteront tant qu'on l'envisage plutôt à la manière d'un réflexe purement inconscient dont il s'agirait seulement de jouer pour manipuler les réactions des utilisateurs de robots à son gré.

bibliographie

- Breazeal C. 2002. Designing sociable robots , Harvard, MIT Press.
- Brooks R. 2002. Robot: The future of Flesh and Machines. London. Allen, The Penguin Press.
- Goetz, J., Kiesler S., Powers, A. 2003. Matching Robot appearance and Behavior to tasks to Improve Human-Robot cooperation. In Proceedings of the 12th IEEE Workshop on Robot and Human Interactive Communication. ROMAN 2003, Oct. 31, 55-60.
- Freedberg D. (1989) , *The power of images, Studies in the History and theory of Response*, Chicago, the university of Chicago Press.
- Guthrie, S.E. (1993) *Faces in the clouds, a new theory of religion*, New-York, Oxford University Press
- Kidd C. D., Breazeal C. 2004, Effect of a robot on user perceptions. In 2004 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS 2004), Sendai, Japan, 2004 (available on-line: web.media.mit.edu/~coryk/publications.html)
- Mori, M. 1970. The uncanny valley. In *Energy*, 7(4), 33-35 (tr. K. F. MacDorman, T Minato
- Reeves, B., Nass, C. 1996. *The Media equation: How people treat computers, Television and new Media like real People and Places*. Chicago: Chicago university Press.

