

Des robots affectifs et sociaux pour le soin, enjeux et problématiques éthiques

■ L'arrivée des robots dans notre société suscite souvent des fantasmes et des peurs ■ Les robots affectifs et sociaux ont déjà de premières applications pour le soin ■ Une réflexion éthique est nécessaire pour encadrer l'usage de ces objets ayant des capacités de simuler des dialogues, semblant être émotionnels et capables d'apprendre.

© 2018 Publié par Elsevier Masson SAS

Mots clés – humanité ; philosophie ; relation ; robot ; société ; soin

Emotional and social robots for care, ethical challenges and issues. The arrival of robots in our society often arouses fantasies and fears. Emotional and social robots are already being used in healthcare. Ethical analysis is necessary to regulate the use of these objects which are able to simulate dialogues, appearing to feel emotions and to be capable of learning.

© 2018 Published by Elsevier Masson SAS

Keywords – care; humanity; philosophy; relationship; robot; society

Le robot comprend des modules d'intelligence artificielle pour percevoir, raisonner et générer des actions. Un robot est dit "social" lorsqu'il s'inscrit dans une relation sociale avec les humains et est capable de dialoguer, comme Nao, par exemple. Un robot affectif a, en plus, la capacité de détecter des indices émotionnels (dans la voix, le visage, la posture, les gestes), de dialoguer en prenant en compte ces informations émotionnelles et de générer en réponse des comportements affectifs comme l'empathie.

Établir une interaction affective avec les robots n'est plus du domaine de la science-fiction : il s'agit de l'*affective computing* [1], un champ de recherche émergent de nombreuses équipes, dont l'une se trouve au Limsi-CNRS (Laboratoire d'informatique pour la mécanique et les sciences de l'ingénieur, Centre national de la recherche scientifique). Les systèmes actuels sont encore très loin de savoir

converser, ils ont des capacités de compréhension très pauvres et n'ont aucun sens commun. Malgré cela, l'être humain projette sur ces objets des capacités humaines, il les anthropomorphise.

Dans ce contexte, les lois de la robotique d'Isaac Asimov¹ doivent être enrichies de règles éthiques à appliquer aux robots sociaux [2].

FANTASMES ET RÉALITÉ

■ **L'intelligence artificielle et la robotique ouvrent d'importantes perspectives** dans le domaine de la santé avec le but de réaliser un meilleur suivi des patients, par exemple. Le concept d'"intelligence artificielle forte" fait référence à une machine capable d'autonomie, de conscience et de sentiments. Le robot est un objet complexe, qui peut sembler autonome et simuler ces capacités, mais sans conscience phénoménale, ni sentiments, ni ce désir ou "appétit de vivre" que Baruch Spinoza

désigne du nom de *conatus* [3] (effort pour persévérer dans l'être) qui se rapporte à toute chose, au mental et au corps.

■ **Pour autant, l'intelligence artificielle et la robotique font énormément fantasmer.**

En 2016, la victoire du robot AlphaGo (*Google DeepMind*) contre l'un des meilleurs joueurs de Go, Lee Sedol, a relancé les questionnements sur les promesses et les risques de l'utilisation de l'intelligence artificielle. Pourtant, lorsqu'AlphaGo bat le joueur de go, la machine ne comprend pas ce qu'elle fait. Le robot peut réussir une tâche difficile, il n'en n'éprouvera aucune fierté, à moins qu'un concepteur l'ait programmé afin de simuler un état affectif. Des intelligences artificielles sont déjà utilisées pour détecter des cancers, pour imiter des artistes, peintres ou musiciens. Les machines pourront sûrement créer, en imitant ou grâce au hasard, mais ne pourront pas savoir si ce qu'elles ont créé est intéressant car elles n'ont pas de conscience.



LA PHILOSOPHIE À L'HÔPITAL

- La fonction soignante en partage
- Agir avec compassion, penser un soin (en) commun
- Quelle place pour la vieillesse dans notre société ?
- La maltraitance des étudiants à l'hôpital
- La simulation en santé pour mieux soigner
- La santé au travail en Chine, une préoccupation naissante
- Après un cancer, le rétablissement dans le parcours de soins
- Le *care*, au fondement du sanitaire et du social
- Un réseau de pairs pour l'éducation thérapeutique des patients atteints de diabète de type 2 au Mali
- **Des robots affectifs et sociaux pour le soin, enjeux et problématiques éthiques**

LAURENCE DEVILLERS
Professeur en intelligence artificielle, affective computing et éthique, Sorbonne Université, équipe de recherche Limsi-CNRS

Limsi-CNRS, campus universitaire bâtiment 508, rue John-von-Neumann, 91405 Orsay cedex, France

Adresse e-mail :
devil@limsi.fr
(L. Devillers).

NOTE

¹L'auteur de science-fiction et scientifique Isaac Asimov a formulé trois lois pour l'usage de la robotique :

- un robot ne peut porter atteinte à un être humain ni, restant passif, permettre qu'un être humain soit exposé au danger ;
- un robot doit obéir aux ordres que lui donne un être humain, sauf si de tels ordres entrent en conflit avec la première loi ;
- un robot doit protéger son existence tant que cette protection n'entre pas en conflit avec la première ou la deuxième loi.

Asimov I (1942). Cycle fermé. In: Les robots. Paris: J'ai lu; 2001.

■ **Malgré ces performances impressionnantes** sur des tâches spécifiques, il est nécessaire de garder à l'esprit que les systèmes d'intelligence artificielle ne parviennent pas à apprendre en dehors des données du "réel". Ils ne connaissent que les données passées pour prédire l'avenir. Or, de nombreuses découvertes de nos plus grands savants sont dues à la capacité d'être contre-intuitif, c'est-à-dire de ne pas tenir compte du savoir présent. Galilée, au XVI^e siècle, eut l'intuition que le poids d'un objet n'avait pas d'influence sur sa vitesse de chute. Il n'est pas possible d'inférer cette loi de la physique : « *Tous les corps tombent à la même vitesse dans le vide* » à partir d'observations du monde réel. La sérendipité, le "don de faire des découvertes par hasard", n'est également pas le fort de la machine. Devant une question sans réponse connue, l'humain est incroyablement plus fort que la machine pour imaginer des solutions. L'intégration d'une intentionnalité et d'une créativité de type humain semble très peu probable pour la machine. Mais tous ces outils d'intelligence artificielle "faible" peuvent devenir de formidables alliés dans une interaction complémentaire avec les humains.

L'APPRENTISSAGE, L'AUTONOMIE ET LA CONSCIENCE DES ROBOTS

■ **L'autonomie est souvent décrite comme une faculté d'apprentissage** pour s'adapter à l'environnement ou, au contraire, une faculté de modifier l'environnement pour l'adapter à ses propres besoins. Pour donner au robot

la capacité d'apprendre de son environnement, de façon autonome, des algorithmes de renforcement qui nécessitent que l'humain conçoive des métriques de récompenses sont utilisés. Le robot apprend par essai et erreur en fonction des récompenses programmées. De façon maladroite et laborieuse, il associe des actions dans le monde et des représentations internes pour réaliser certaines tâches pour lesquelles il est conçu. Les enfants apprennent en expérimentant le monde. Ces machines sont dites "intelligentes" car elles peuvent également apprendre. Pour un robot, la tâche est extrêmement difficile car il n'a ni instinct, ni intentions pour prendre des décisions. En revanche, il peut imiter l'être humain.

■ **À l'heure actuelle, la relative autonomie des robots** est toujours programmée par l'humain. La faculté d'apprentissage programmée peut offrir plus ou moins de liberté à la machine. Donner à un robot la capacité d'apprendre seul, en interaction avec l'environnement et les humains, est le Graal des chercheurs en intelligence artificielle. Il est alors souhaitable de leur enseigner les valeurs communes et morales de la vie en société. La faculté d'apprendre seul constitue une rupture technologique et juridique et soulève de nombreuses questions éthiques. Ces robots peuvent être, d'une certaine manière, créatifs et autonomes dans leurs prises de décision, s'ils sont programmés pour cela.

■ **Vouloir recopier l'intelligence de l'homme sur une machine** est très narcissique, car que savons-nous de notre intelligence ? Nous ne connaissons pas le substrat

de la pensée et n'avons pas conscience que certains de nos organes sont autonomes. Nous ne sommes conscients que d'une petite partie de nos perceptions et de notre activité cérébrale. Il n'existe pas de terme plus polysémique et sujet à interprétation que celui de "conscience" : il évoque pour certains la conscience de soi, pour d'autres la conscience du prochain, ou encore la conscience phénoménale, la conscience morale, etc.

■ **Les systèmes actuels d'intelligence artificielle** ont la capacité de calculer des corrélations de faits, avec les approches d'apprentissage profond par exemple, de prendre des décisions et d'apprendre, mais sans en avoir conscience. Certains prototypes de robot ont pourtant déjà des embryons de niveau de "conscience" comparables à ceux que décrivent les neuroscientifiques comme Stanislas Dehaene. Celui-ci affirme que « *La conscience est une opération computationnelle, un calcul [...]. Nous commençons à progresser dans la connaissance de ce qu'il faudrait ajouter aux machines pour qu'elles deviennent conscientes* » [4]. Ces embryons de niveau de conscience sont simulés par des mécanismes de partage de connaissances et d'introspection.

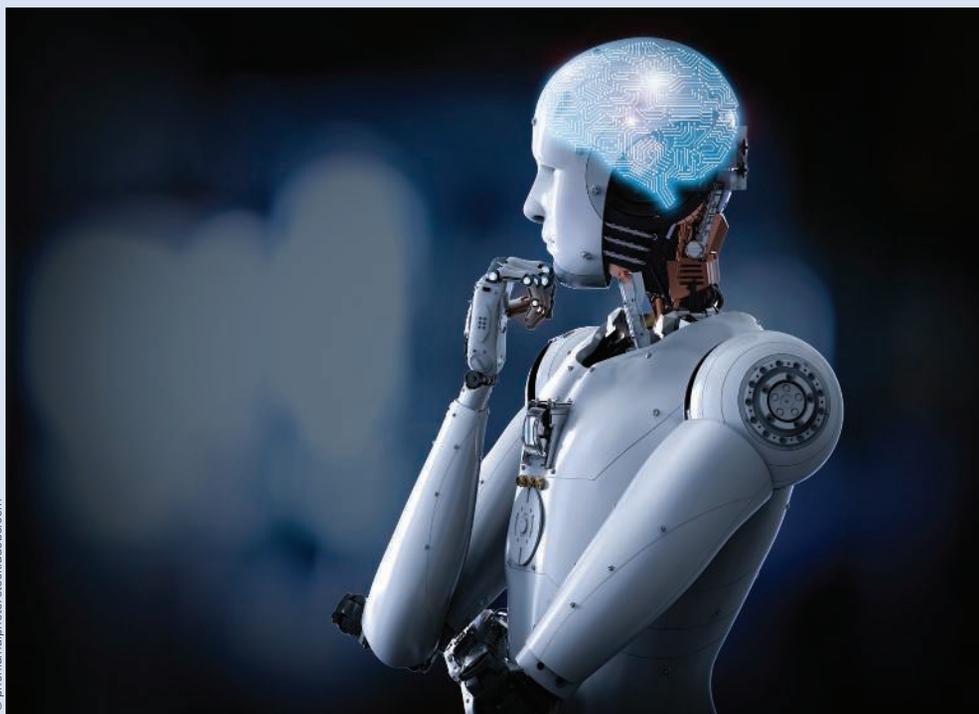
■ **Pour autant, ces machines ne sont pas conscientes** comme peut l'être un humain. Elles n'ont ni conscience morale ni conscience phénoménale associée à une expérience qualitative telle que la sensation de chaud ou de froid, le sentiment d'anxiété, etc., car elles n'ont pas de viscères ni de ressenti, à moins, là encore, de les simuler. Une conscience artificielle, dotée de sentiments, de pensées et de libre

arbitre sans programmation humaine a donc peu de chance d'émerger spontanément avec les architectures actuelles d'ordinateurs.

LA ROBOTIQUE SOCIALE ET AFFECTIVE EN SANTÉ

■ **La robotique sociale et affective** permet de créer des robots compagnons, censés nous apporter une aide thérapeutique, d'assistance ou encore de surveillance. Dans le cas de pathologies neurodégénératives ou de handicaps lourds, le robot peut même être meilleur que l'homme pour interagir avec les personnes. La machine se met au diapason de l'autre, à des rythmes très lents, presque inhumains. Le robot écoute avec bienveillance et sans aucune impatience. Il peut être très utile pour suivre des patients tout au long d'une maladie et favoriser le maintien à domicile des personnes malades, âgées ou handicapées et, ainsi, réduire leurs périodes d'hospitalisation.

■ **Les robots sont disponibles 24 heures sur 24, 7 jours sur 7.** Ils sont patients, peuvent pallier des défaillances de perception (surdité, malvoyance) et donner accès à des informations sur Internet plus facilement qu'un ordinateur. Ils se révèlent précieux également pour enregistrer des données en continu et les envoyer à un médecin pour détecter des comportements anormaux (dépression), et suivre des patients atteints de troubles bipolaires ou de maladies dégénératives, dans l'accompagnement de la vie quotidienne. Il est déjà possible de concevoir des systèmes intelligents pour entraîner des personnes en perte de facultés cognitives à stimuler la mémoire et le langage.



Le concept d'"intelligence artificielle forte" fait référence à une machine capable d'autonomie, de conscience et de sentiments.

■ **Les machines seront de plus en plus autonomes,** "bavardes" et douées de simulation d'émotions grâce à des programmes d'intelligence artificielle sophistiqués, mais elles ne seront pas pour autant capables de sentiments, de créativité et d'imagination tels que nous les entendons pour les humains. Malgré cela, nous projetons sur eux des capacités humaines [5]. Pour cette raison, il est fondamental de réfléchir à la régulation de leur fonctionnement, mais aussi de démystifier leurs capacités, par l'éducation notamment, et, également, de rendre plus transparents et compréhensibles les programmes informatiques qu'ils embarquent.

ÉTUDES PERCEPTIVES SUR L'INTERACTION HUMAIN-MACHINE

■ **Il est important de mener des études perceptives** sur

l'interaction humain-machine. Le Limsi-CNRS a conduit de nombreux tests en laboratoire et en établissement d'hébergement pour personnes âgées dépendantes (Ehpad) avec des personnes âgées, ou dans des centres de rééducation avec l'association Approche [6], dans le cadre du projet BPI Romeo2 piloté par *Softbank Robotics* [7]. Des études ont également été menées autour de scénarios de la vie de tous les jours et de jeux avec l'équipe du Pr Anne-Sophie Rigaud au *Living Lab* de l'hôpital Broca [8].

■ **Toutes ces expériences ont montré que les robots étaient plutôt bien acceptés** par les patients, lorsqu'ils ont le temps de les expérimenter et qu'ils entrevoient leur utilité. Les discussions post-expériences ont également fait remonter un certain nombre d'inquiétudes légitimes sur le manque de transparence et d'explication du comportement de ces machines.

RÉFÉRENCES

[1] Picard RW. *Affective Computing*. Cambridge, MA: MIT Press; 1997.

[2] Devillers L. *Des robots et des hommes. Mythes, fantasmes et réalité*. Paris: Plon; 2017.

[3] Spinoza B. *Éthique III, prop. 9, scolie*.

[4] Sender E. Comment donner une conscience aux machines. *Sciences et avenir*. 27 octobre 2017. https://www.sciencesetavenir.fr/high-tech/un-article-de-stanislas-dehaene-expose-comment-donner-conscience-a-l-intelligence-artificielle_117736

[5] Reeves B, Nass C. *The Media Equation: How People Treat Computers, Television, and New Media Like Real People and Places*. Cambridge, MA: Cambridge University Press; 1996.

[6] <http://www.approche-asso.com>

[7] Kumar Pandey A, Gelin R, Alami R et al. Ethical considerations and feedback from social human-robot interaction with elderly people. *AIC*, 2014.

[8] Garcia M, Béchade L, Dubuisson-Duplessis G et al. Towards metrics of Evaluation of Pepper robot as a Social Companion for Elderly People. *International Workshop on Spoken Dialog System Technology*; 2017. https://www.uni-ulm.de/fileadmin/website_uni_ulm/iui.iwds2017/papers/IWSDS2017_paper_24.pdf

Déclaration de liens d'intérêts
L'auteur déclare ne pas avoir de liens d'intérêts.

■ **Il est urgent de développer un domaine de recherche interdisciplinaire** avec des informaticiens, des médecins, des psychologues cogniticiens, pour étudier les effets de la coévolution avec ces machines de façon longitudinale. La machine va apprendre à s'adapter à nous, comment allons-nous adapter à elle ?

ÉTHIQUE DES SYSTÈMES ROBOTIQUES

■ **Il faut éviter un déficit de confiance** mais également une confiance trop aveugle envers les programmes d'intelligence artificielle. Un certain nombre de valeurs éthiques sont importantes : la déontologie et la responsabilité des concepteurs, l'émancipation des utilisateurs, l'évaluation, la transparence, l'explicabilité, la loyauté et l'équité des systèmes, enfin, l'étude de la co-évolution humain-machine.

■ **Les robots sociaux et affectifs soulèvent de nombreuses questions** éthiques, juridiques et sociétales. Qui est responsable en cas d'accident : le fabricant, l'acheteur, le thérapeute, l'utilisateur ? Comment réguler leur fonctionnement ? Faut-il intégrer des règles morales dans leur programmation ? Contrôler leur utilisation par des permis ? Pour quelles tâches souhaitons-nous créer ces entités artificielles ? Comment préserver notre intimité, nos données personnelles ?

■ **Tout système doit être évalué** avant d'être mis dans les mains de son utilisateur. Comment évaluer une intelligence artificielle qui apprend

des humains et s'adapte à eux ? Pouvons-nous prouver qu'elle se cantonnera aux fonctions pour lesquelles elle a été conçue, qu'elle ne pas dépassera pas les limites fixées ? Les données que la machine exploite pour son apprentissage la dirigent vers certaines actions. Qui supervisera la sélection de ces données ? Ces questions prégnantes ne sont évoquées que depuis peu.

■ **Les progrès spectaculaires du numérique** permettront un jour d'améliorer le bien-être des personnes, à condi-

Les robots sociaux

et affectifs soulèvent

de nombreuses questions

éthiques, juridiques et sociétales

tion de réfléchir non à ce que nous pouvons en faire, mais à ce que nous voulons en faire. L'intelligence artificielle peut apporter de meilleurs diagnostics en santé, des outils de stimulation, de détection des comportements anormaux, une meilleure assistance – notamment pour le handicap ou la perte d'autonomie –, mais le contrôle par des humains sera toujours primordial. Il est nécessaire de développer des cadres éthiques pour les robots sociaux, notamment en santé, et de comprendre le niveau de complémentarité humain-machine.

■ **Les lois d'Asimov pourraient donc être complétées de commandements** [2] adaptés aux robots assistants de vie. Les fondements de ces commandements proviennent

en partie de retours d'expérience d'interactions entre des personnes âgées et des robots. Ils concernent :

• **les données privées** : « Tu ne divulgueras pas mes données à n'importe qui » ;

• **le droit à l'oubli** : « Tu oublieras tout ce que je te demande d'oublier » ;

• **la sécurité** : « Tu pourras te déconnecter d'Internet si je te le demande » ;

• **le contrôle** : « Tu seras régulièrement contrôlé pour évaluer ce que tu as appris » ;

• **l'explicabilité et la traçabilité** : « Tu pourras m'expliquer tes comportements si je te le demande » ;

• **la loyauté** : « Tu seras loyal » ;

• **le consentement** : « Tu seras empathique et simulera des émotions, seulement si je le sais ! » ;

• **le risque de dépendance** : « Tu stimuleras ma mémoire et veilleras à ce que je ne devienne pas trop dépendant de toi ! » ;

• **le risque de confusion** : « Tu feras attention à ce que je ne te confonde pas avec un être humain ! » ;

• **l'adaptation aux règles sociales** : « Tu apprendras à vivre avec des humains et tu t'adapteras aux règles sociales » ;

• **l'utilité et la bienveillance** : « Tu seras bienveillant et utile. Et pourquoi pas, doué d'un peu d'humour ! » ;

■ **Il est important de construire les robots sociaux de demain *ethic by design***, c'est-à-dire selon une conception éthique, responsable et sociale. Nous avons besoin de démystifier, de former à l'intelligence artificielle et de remettre les valeurs de l'humain au centre de la conception de ces systèmes robotiques. ■