



PSYCAUSE
Revue scientifique étudiante de
l'École de psychologie de l'Université Laval



UNIVERSITÉ
LAVAL

Faculté des sciences sociales
École de psychologie

revues.ulaval.ca/ojs/index.php/psycause

MAI 2019 – VOL. 9 N° 1

HUMANISATION DES ROBOTS : IMPLANTER L'EMPATHIE DANS L'INTELLIGENCE ARTIFICIELLE

Marc-Antoine LINTEAU^{1,*}

¹ École de psychologie, Université Laval, Québec, Canada

* marc-antoine.linteau.1@ulaval.ca

Pour citer l'article

Linteau, M.-A. (2019). Humanisation des robots : implanter l'empathie dans l'intelligence artificielle. *Psycause: Revue scientifique étudiante de l'École de psychologie de l'Université Laval*, 9(1), 56-63.

Droits d'auteur

© 2019 Linteau. Cet article est distribué en libre accès selon les termes d'une licence Creative Commons Attribution 4.0 International (de type CC-BY 4.0) qui permet l'utilisation du contenu des articles publiés de façon libre, tant que chaque auteur ou autrice du document original à la publication de l'article soit cité(e) et référencé(e) de façon appropriée.

ISSN: 2562-4385

HUMANISATION DES ROBOTS : IMPLANTER L'EMPATHIE DANS L'INTELLIGENCE ARTIFICIELLE

Marc-Antoine LINTEAU

École de psychologie, Université Laval, Québec, Canada

Résumé

Depuis l'avènement de l'intelligence artificielle, plusieurs domaines de recherche, dont l'informatique affective (*affective computing*), se sont intéressés à implanter celle-ci dans différentes sphères de la société. L'informatique affective se penche surtout sur les interactions sociales entre les humains et l'intelligence artificielle. Avec les nombreux projets d'insertion de cette nouveauté dans notre vie, une meilleure compréhension des interactions humaines qui vont en découler est donc d'autant plus importante. Par le biais de ce domaine de recherche, l'empathie semble jouer un rôle de première instance dans l'intégration de l'intelligence artificielle à notre quotidien. Étant l'une des habiletés essentielles chez l'humain, elle contribue à nos interactions et sert de base à la différenciation avec les interactions animales. L'empathie artificielle est donc l'application de certaines caractéristiques de l'empathie à une intelligence artificielle, mais plus particulièrement à des robots, soit une entité physique, ou à un système d'avatar. La présence physique de l'empathie (par le biais d'un robot) s'avère primordiale à l'insertion de l'intelligence artificielle dans nos vies courantes, puisqu'elle a une incidence sur l'établissement d'une relation empathique. La présence de capacités mnésiques est également essentielle au développement de relations à long terme. Un robot universel signifie une application de l'empathie qui permet d'éviter l'impression d'interactions inauthentiques, mais surtout un robot pouvant s'appliquer, s'adapter et pouvant interagir avec l'ensemble de la société. Cette recension des écrits sert à éclairer les lecteurs sur l'état actuel et sur l'avenir de l'informatique affective. Ultimement, l'objectif est d'implanter la capacité holistique de l'empathie pour satisfaire les nombreuses clientèles spécifiques comme les TSA, les personnes âgées ou les enfants, ainsi que l'unicité de chaque individu représentant la race humaine.

Mots-clés : Robotique sociale, empathie, intelligence artificielle, émotions, *affective computing*

Abstract

Since the advent of artificial intelligence, several fields of research, including affective computing, have been interested in implementing it in different spheres of society. Emotional computing focuses on social interactions between humans and artificial intelligence. There are many projects wanting to integrate artificial intelligence into our lives and a better understanding of the resulting human interactions is therefore important. Throughout this area of research, empathy plays a primary role in the insertion of artificial intelligence into our lives. As one of the essential skills in humans, empathy contributes to our interactions and serves to differentiate human's and animal's interactions. Artificial empathy is therefore the application of certain characteristics of empathy to an artificial intelligence, but more particularly to robots, representing a physical entity. The physical presence of empathy (through a robot) is essential to insert artificial intelligence into our daily lives, since it has an impact on the establishment of an empathic relationship. The presence of memory skills is also essential for the development of long-term relationships. A universal robot means an application of empathy that avoids the impression of inauthentic interactions, but above all a robot that can adapt, interact and be applied to the whole society. This literature review serves to inform readers about the current state and future of emotional computing. Ultimately, the objective is to implement the holistic capacity of empathy to satisfy the many specific clienteles, such as ASDs, the elderly, the children, and the uniqueness of each individual among the human race.

Keywords : Social robotic, empathy, artificial intelligence, emotions, *affective computing*

Au cours de l'évolution, l'empathie s'est développée pour maintenant se distinguer entre les animaux

(p. ex., les autres mammifères) et les humains. Elle est aujourd'hui définie comme étant « l'expérience

subjective naturelle de similitude entre les sentiments exprimés par soi-même et les autres, sans perdre de vue à qui appartiennent les sentiments» (Decety & Jackson, 2004). Cette habileté sociale suscite de nombreuses interrogations auprès de nombreux philosophes, biologistes et primatologues et ce, depuis le 18^e siècle. Entre autres, comment l'empathie s'est-elle développée au cours de l'évolution des espèces, des premiers mammifères à l'humain? De plus, comment permet-elle de favoriser les interactions sociales? Le concept fut notamment traité par Frans de Waal, qui amena le modèle de la poupée russe pour démontrer les différentes formes d'empathie chez les animaux et les êtres humains (de Waal, 2008). Ce modèle se compose de trois poupées emboîtées représentant des niveaux de complexité d'empathie croissants. La première poupée contient des comportements primitifs ou biologiques présents chez les animaux, soit le mimétisme moteur (c.-à-d., l'imitation d'un comportement observé) et la contagion émotionnelle (c.-à-d., le partage d'un état émotionnel; Asada, 2015). La seconde poupée contient des comportements plus complexes observés chez certains types d'animaux, soit la coordination vers un but commun et la préoccupation sympathique (c.-à-d., la consolation). La troisième est donc la plus caractéristique chez les humains, soit le stade le plus évolué d'empathie. Elle est composée de la prise de perspective (aide ciblée) et de l'imitation véritable, à ne pas confondre avec le mimétisme moteur renvoyant plutôt à une caractéristique biologique se produisant automatiquement chez de nombreuses espèces, telle que le réflexe de bâiller en voyant quelqu'un d'autre bâiller. L'imitation véritable renvoie à comprendre ce que l'autre fait et à reconnaître qu'il doit faire ainsi pour le mieux de tous, comme dormir en même temps et chasser en groupe (de Waal, 2008). Le chimpanzé posséderait également certaines caractéristiques moins développées de la troisième poupée comme la prise de conscience introspective, soit la capacité de comprendre la présence de son corps et de pouvoir le distinguer de celui des autres (de Waal, 2008). Les études du primatologue Frans de Waal ont démontré la complexité du phénomène et son importance pour le développement de bonnes interactions sociales. C'est dans le but d'optimiser nos interactions sociales que la recherche sur l'empathie est actuellement très active. Ce relevé de littérature sert donc à synthétiser la recherche actuelle sur l'empathie et à montrer les avenues futures qu'elle pourra

prendre, mais particulièrement d'un aspect robotique, donc la présence d'une entité physique.

L'empathie aujourd'hui

L'empathie est un concept bien ancré dans la recherche qui a pris une nouvelle direction depuis les vingt dernières années. Auparavant, les recherches sur l'empathie étaient presque exclusivement axées sur la compréhension du phénomène, soit par l'étude des comportements ou par les corrélats physiologiques et neurophysiologiques. Aujourd'hui, on tente en plus de reproduire et d'introduire l'empathie humaine dans des systèmes d'informatique grâce, entre autres, à l'intelligence artificielle. Ainsi, certains chercheurs, tels qu'Asada (2014) ainsi que Kirby, Forlizzi et Simmons (2009), tentent d'implanter l'empathie dans des machines, soit une entité réelle et physique pouvant accomplir une tâche, mais plus particulièrement dans des robots remplissant des tâches caractéristiques de l'humain.

Pour pouvoir implanter l'empathie à une intelligence artificielle dans de nombreux domaines de la société, il faut connaître les caractéristiques holistiques de l'empathie. Certains auteurs mentionnent que cette dernière, étant à la base de la coopération sociale et du comportement prosocial, serait essentielle dans la création de robots pouvant maintenir des interactions positives avec des humains (Leite, Pereira, Mascarenhas, Martinho, Prada & Paiva, 2013). L'empathie permet de partager l'expérience, les besoins et les désirs, et ce, grâce à un lien émotionnel avec autrui permettant de se mettre dans la perspective de l'autre tout en étant en mesure de distinguer nos émotions et celles des autres. En effet, il a été démontré qu'un manque d'empathie nuit fortement aux relations humaines (Riess, 2017). Cela étant dit, pour qu'une intelligence artificielle empathique fonctionne, celle-ci se doit de devenir mutuelle en prenant en compte les deux partis de l'échange (Lim, Aylett & Jones, 2005), et elle se doit d'être coopérative, donc de communiquer son état affectif pour favoriser une bonne interaction (Ratcliffe, 2017). Elle se doit aussi de prendre en compte les normes sociales et les aspects culturels (Niculescu, van Dijk, Nijholt & Haizhou, 2014). Ainsi, pour les futurs développements des robots pour lesquels des capacités empathiques sont introduites, l'empathie se doit d'être authentique,

donc avec la forme la plus humaine possible, pour qu'elle puisse être introduite dans certaines sphères sociétales, telles que la médecine ou l'éducation.

Deux objets d'étude caractérisent la recherche sur l'implantation de l'empathie chez les robots: les expressions émotionnelles externes et les expressions émotionnelles internes (Damiano, Dumouchel & Lehmann, 2015). L'aspect externe est associé à la sociabilité, donc à l'interaction entre les humains et les robots. Au contraire, l'aspect interne se concentre sur le développement d'agents empathiques ayant leur comportement influencé par des processus endogènes, qui sont modulés par des mécanismes de régulation émotionnelle. L'aspect externe est donc associé à l'expression d'émotions sans que le robot puisse ressentir ces émotions, alors que l'aspect interne tente de réellement créer des émotions artificielles afin de rendre le robot empathique (Damiano, Dumouchel & Lehmann, 2015). En somme, ces objets d'étude de l'informatique affective sont mutuellement exclusifs, donc ils partagent différents points de vue et façons de procéder pour implanter l'empathie, mais ils partagent un but commun.

L'expression de l'empathie

D'une part, les chercheurs qui abordent l'aspect interne prônent le fait que l'intelligence artificielle se doit d'être la plus humaine possible, donc mutuelle, coopérative (Ratcliffe, 2017) et avec des émotions qui varient au fil du temps (Kirby, Forlizzi & Simmons, 2009) pour que celle-ci fonctionne dans une situation d'interaction humaine. L'étude du robot réceptionniste de Kirby, Forlizzi et Simmons (2009) est un bon exemple d'étude utilisant l'aspect interne. Cette étude a tenté de créer un modèle de robot pouvant développer une humeur générale au robot et non une réponse émotionnelle immédiate ou automatique. Le but derrière l'étude était de favoriser des relations à long terme les plus authentiques possibles. D'autre part, les études qui abordent l'aspect externe promeuvent l'opérationnalisation d'une intelligence artificielle efficace dans les interactions humaines puisqu'elles font référence à la communication et à l'expression non verbale de l'empathie, donc les mouvements, les gestes, la posture et les traits faciaux (Stephan, 2015). Selon la théorie de l'*Uncanny Valley* de Mashiro Mori, la réaction d'une personne envers

un robot ressemblant à un être humain passerait de la répulsion à un sentiment positif à mesure que le robot s'approcherait d'une apparence humaine, pour soudainement tomber drastiquement et provoquer une réaction négative lorsque le robot ressemble presque à un humain, mais pas tout à fait (Mori, MacDorman & Kageki, 2012). Cette hypothèse a amené certains chercheurs à tester cette théorie tracée par Mori. Par exemple, l'*Affective AutoTutor*, un avatar de tutorat intelligent, a démontré que les caractéristiques physiques sont importantes pour améliorer l'acceptation sociale des robots et l'interaction efficace avec ces derniers (Stephan, 2015). Cette expérience produite à l'aide d'un avatar n'était peut-être pas assez réaliste pour bien reproduire une interaction avec une entité physique. Ainsi, si la théorie de Mori est prise en compte, le faible réalisme de l'expérience amènerait le sentiment positif initial. Il y aurait donc une forte importance reliée au fait que les caractéristiques externes ne produisent pas automatiquement un sentiment empathique et qu'il faut plus que l'apparence pour obtenir une relation empathique, comme la présence de capacités mnésiques.

Ce manque de consensus dans l'aspect externe de l'empathie prouve comment il est difficile de reproduire de bonnes interactions sociales, même sur une base purement physique. Ainsi, l'aspect externe des expressions émotionnelles semble faire référence à la première poupée du modèle de Frans de Waal, soit la contagion émotionnelle et le mimétisme moteur, qui sont aussi deux traits primitifs présents chez une grande partie des animaux. C'est pourquoi l'imitation des expressions émotionnelles externes est capitale dans l'application de l'empathie artificielle.

Exemple d'un système de tutorat empathique

Le défi est de rendre les systèmes empathiques plus performants dans leurs interactions. Puisque de nombreuses études ont été produites dans des situations spécifiques, les limites dans la capacité des robots à interagir (Asada et al., 2009) étaient perçues comme une grande lacune de l'authenticité (Asada, 2014). Un exemple de cette affirmation est le système de tutorat *Affective AutoTutor*. Ce dernier est un système d'avatars qui fut testé dans un contexte d'apprentissage pour démontrer comment les interactions entre individus

et agents virtuels peuvent s'améliorer grâce aux capacités empathiques de l'agent. Le système est capable d'être sensible à l'état affectif de l'utilisateur grâce à ses expressions faciales, à son langage corporel et à des indices dans la conversation, tels que la frustration, l'ennui ou la prosodie. Toutefois, si le système détecte de la frustration chez un élève et qu'il tente de lui démontrer de l'empathie, celui-ci possède des capacités trop limitées pour pouvoir réellement interagir. Il peut seulement lui donner quelques phrases d'encouragement. Par exemple, si un élève démontre de la frustration à cause du divorce de ses parents, le système ne peut pas être empathique envers la situation de l'élève en lui demandant s'il désire en discuter ou simplement arrêter l'apprentissage. L'empathie doit pouvoir s'appliquer à toute situation pour établir réellement une relation positive (Stephan, 2015). Le concept de mutualité, vu auparavant, est essentiel pour fonder une interaction saine avec une intelligence artificielle. C'est donc sa faible capacité de pouvoir communiquer spontanément à l'aide du langage parlé qui nuit le plus à l'*AutoTutor* dans l'établissement d'un bon lien. Cette capacité de communication verbale est donc d'une importance primordiale pour introduire pleinement l'empathie dans les interactions sociales et favoriser de saines relations.

Robots et TSA

Certaines études ont tenté d'introduire l'intelligence artificielle dans des situations d'apprentissage très précises (Asada, 2015). En effet, il peut y avoir certaines applications spécifiques, notamment pour l'entraînement des habiletés sociales chez les individus atteints du trouble du spectre de l'autisme (TSA). Les TSA possèdent une triade de symptômes communs : une difficulté dans l'interaction, une difficulté dans la communication sociale et des comportements, intérêts ou activités limitées et répétitives (American Psychiatric Association, 2013). Bien que les TSA aient ces trois types de déficits en commun, le niveau de sévérité de chacun d'entre eux et les différences individuelles rendent l'utilisation d'un robot générique très peu probable. De ce fait, les TSA forment un groupe très hétérogène (Dautenhahn & Werry, 2004). Au vu de leurs particularités, les chercheurs du projet *Aurora* (*Autonomous Robotic platform as a Remedial tool for children with Autism*) ont tenté de développer un robot autonome pour encourager les enfants atteints de

TSA à s'engager dans des interactions humaines par le biais, entre autres, du contact visuel, de l'attention conjointe, de l'approche, de l'évitement et des jeux d'imitation. L'utilisation d'agents sociaux comme les robots a beaucoup de potentiel pour aider les enfants, puisque si l'on se fie à la définition du DSM-V (American Psychiatry Association, 2013), les individus atteints du TSA aiment ce qui est prévisible, répétitif et monotone, donc des aspects qu'une intelligence artificielle pourraient procurer. Une limite peut alors être soulevée dans le développement de robots : l'intelligence artificielle ne doit pas développer des comportements imprévisibles pouvant générer des crises chez les TSA (Dautenhahn & Werry, 2004). L'imprévisibilité étant un trait typiquement humain, il semble contre-intuitif de vouloir le retirer de la relation. Toutefois, celui-ci ne serait pas un aspect favorable au développement d'une bonne relation avec un TSA. Dépendamment de l'objectif de l'utilisation des robots avec les TSA (p. ex., offrir un compagnon ou développer des compétences sociales favorisant l'intégration dans la société), l'implantation de l'empathie serait une caractéristique à considérer ou non dans la création d'intelligence artificielle.

Comparativement à l'*Affective AutoTutor*, qui possédait une lacune dans l'aspect de mutualité, le robot autonome du projet *Aurora* manque de prise de perspective ou d'aide ciblée, qui a été identifiée préalablement comme l'une des deux caractéristiques de la troisième poupée du modèle de Frans de Waal (2008). Cette capacité permet à l'être humain de pouvoir se placer dans la peau d'autrui et donc de comprendre ses comportements. Dans le cas présent, la prise de perspective inclut l'état psychologique différent d'un TSA, ses déficits d'interaction et de communication. Cette capacité cognitive qu'est la prise de perspective se voit donc essentielle pour s'adapter à n'importe quelle interaction.

Développement d'une relation à long terme avec un robot

Le développement de relations à long terme pour déterminer l'impact réel de l'empathie (Dautenhahn & Werry, 2004) et la nécessité du développement d'un modèle riche d'affects (Kirby, Forlizzi & Simmons, 2009) sont d'une grande importance. Pour réussir l'implantation d'empathie artificielle dans des systèmes

informatiques, Kirby, Forlizzi et Simmons (2009) ont tenté de créer un modèle social-affectif de robot. Ils ont tenté d'améliorer les interactions en amenant le robot à afficher des humeurs changeantes et des émotions pour imiter le plus naturellement possible le comportement humain et pour fournir un soutien psychologique afin de favoriser une interaction positive (Kirby, Forlizzi & Simmons, 2009). Le robot est composé d'un écran LCD affichant un visage humain et un clavier relié à cet écran pour permettre aux individus d'interagir avec lui. À l'entrée d'un pavillon universitaire, le robot avait comme rôle d'accueillir et de donner des indications aux étudiants. Il possédait deux histoires, soit celle de Valérie et de Tank. Ces deux histoires inventées étaient destinées à susciter un intérêt et à encourager les relations à long terme avec le robot. Les conclusions qui ont été tirées à la suite de la recherche étaient que les individus sont capables de comprendre les émotions et les expressions du robot, et que ces divergences d'émotions et d'expressions influencent comment un individu interagit et perçoit ce dernier. Les gens évitaient le robot lorsque celui-ci présentait des émotions négatives, et étaient attirés à ce dernier lorsqu'il présentait des émotions positives, ce qui reflète bien le comportement humain normal. De plus, le robot, n'ayant aucune mémoire des individus rencontrés auparavant, réagissait de la même façon aux nouveaux venus qu'aux gens déjà rencontrés, ce qui affecte grandement une relation (Kirby, Forlizzi & Simmons, 2009). Pour éviter cette limite, une solution pourrait être d'ajouter une technologie de reconnaissance faciale et une banque de données de visages aux robots pour ainsi leur créer une sorte de « mémoire ». En plus de l'empathie, la présence de capacités mnésiques semble donc essentielle pour le développement d'une relation à long terme avec des robots.

La dimension interactionniste

La recherche basée sur l'aspect interne de l'empathie artificielle stipule que les émotions sont privées (à l'intérieur de l'individu) et que bien qu'elles soient exprimées aux autres, elles ne restent comprises que par le sujet (Damiano, Dumouchel & Lehmann, 2015). Ainsi, l'expression de l'émotion est uniquement secondaire. Au contraire, l'étude de l'aspect externe considère que, pour tirer des conclusions sur les états émotionnels de l'autre, une analyse rationnelle des

comportements et des émotions observables chez un sujet est uniquement nécessaire. Les deux aspects de la recherche sur l'empathie artificielle sont donc en accord sur le fait que les processus empathiques se situent dans l'espace de la relation entre le sujet et son environnement (Damiano, Dumouchel & Lehmann, 2015), ce qui est appelé la dimension interactionniste. Celle-ci est de plus en plus présente dans les études actuelles qui tentent d'engager les sujets davantage dans la relation pour qu'ils se sentent plus impliqués avec le système. De cette manière, les sujets développent de plus en plus une « présence sociale » envers le système, soit le degré d'accessibilité à l'intelligence, aux intentions et aux impressions sensorielles du système (Leite, Martinho, Pereira & Paiva, 2009). Cette relation entre le sujet et l'environnement du système est communément appelée l'*affective loop* et fait référence à l'établissement d'une relation empathique à long terme qui devient mutuelle (Lim, Aylett & Jones, 2005) et coopérative pour bien se bâtir (Ratcliffe, 2017).

Un des grands problèmes de la recherche sur l'aspect externe de l'empathie artificielle provient du fait que les robots qui ont été créés à ce jour possèdent un nombre restreint de comportements et d'expressions pouvant être produits, ce qui empêche l'adaptation nécessaire à des relations à long terme. La relation interactionniste tente de palier à ce problème en amenant le robot à devenir plus autonome en développant une certaine habileté d'apprentissage pour développer de nouveaux comportements et expressions (Vircikova, Magyar & Sincak, 2015). Selon le modèle de l'*affective loop* proposé par Kristina Höök (2008), l'utilisateur débute par exprimer une émotion à l'aide de mouvements physiques. Ensuite, le système produit des expressions affectives en utilisant de la couleur, des animations ou des interactions incluant le toucher. Ces stimuli produits par le système vont générer une réponse émotionnelle chez l'utilisateur. Ainsi, ce dernier va se sentir de plus en plus impliqué dans son interaction avec le système. Le système tente d'engager les utilisateurs dans des processus émotionnels et physiques sans tenter d'interférer dans ses états émotionnels, et donc, le choix revient à l'individu de participer ou non dans l'interaction. La décision de s'engager favoriserait la façon d'appréhender la relation (Höök, 2008). Ce modèle fut testé dans une recherche avec des enfants en

milieu scolaire (Vircikova, Magyar & Sincak, 2015). Le robot utilisé était capable de percevoir les états émotionnels grâce à un système de reconnaissance émotionnelle. Il pouvait aussi réguler des émotions artificielles grâce à la construction d'un modèle émotionnel. En plus, le robot avait la capacité d'apprendre des expressions émotionnelles par l'expérience et la planification de ses réponses. L'objectif de ce projet était que les élèves de cinq à sept ans puissent apprendre des nouveaux mots en anglais. Lors de l'expérience, le robot régulait ses émotions en analysant les réactions émotionnelles des enfants. Par exemple, il démontrait de la joie si l'enfant était perçu comme content lorsqu'il se rappelait un mot à apprendre. Par contre, si l'enfant ne se rappelait pas du mot, le robot allait exprimer de la tristesse. L'étude de Vircikova, Magyar et Sincak (2015) a donc permis de démontrer que l'adaptation de la réponse mène à un besoin et à une efficacité dans l'établissement d'une relation plus persistante dans le temps. En somme, l'adaptation à des environnements variés et la personnalisation de l'interaction en fonction des individus sont deux caractéristiques nécessaires pour le développement d'une relation à long terme (Vircikova, Magyar & Sincak, 2015).

Lien de confiance avec un robot

En reprenant la théorie de l'*Uncanny Valley*, une étude de Moore (2012) a entrepris de déterminer quelles sont les caractéristiques qui suscitent un sentiment de confiance envers un robot. L'objectif de ce projet de recherche était de savoir si des individus pouvaient percevoir une différence dans l'interaction avec un robot et un humain, et la différence entre une haute intelligence émotionnelle et une faible intelligence émotionnelle chez le robot. C'est à l'aide du robot personnel PR2, qui peut produire divers mouvements physiques personnalisés, que les chercheurs ont mené leur expérience. Pour ce faire, des vidéos représentant soit le robot PR2 ou un humain agissant avec ou sans empathie ont été présentés aux participants. Ceux-ci devaient noter la confiance qu'ils avaient envers le robot ou envers l'acteur ainsi que le niveau d'intelligence émotionnelle perçue. Certaines facettes de la communication comme la posture et le ton de voix, ont également été variées pour quantifier leur effet sur les interactions. Aucune différence significative dans la façon dont l'humain et le robot

étaient perçus n'a été trouvée. Ils ont prouvé que les robots peuvent être perçus comme démontrant une intelligence émotionnelle similaire à l'humain dans les interactions sociales. Peu importe si le stimulus était un humain ou un robot, les participants ont toujours été en mesure de discerner les différences dans l'intelligence émotionnelle. Par exemple, lorsque le robot ou l'acteur humain utilisait une gestuelle d'ouverture, les participants discernaient facilement la forte intelligence émotionnelle. Au contraire, lorsque le robot ou l'acteur humain croisait les bras en signe de fermeture et n'interagissait pas, les participants discernaient la faible intelligence émotionnelle. Les chercheurs ont toutefois décelé que les facteurs de communication non-verbale, tels que le ton de la voix et la posture, influençaient la perception de l'intelligence émotionnelle chez l'humain, mais pas chez le robot. Ceci suggère que les facteurs de communication, autant verbaux que non verbaux, sont les plus déterminants dans la perception de l'intelligence émotionnelle. Il faut donc créer un modèle de robot pouvant communiquer le plus naturellement possible pour favoriser un plus grand lien de confiance (Fan, Scheutz, Lohani, McCoy & Stokes, 2017).

Conclusion

L'empathie est une habileté sociale présente chez plusieurs espèces, dont l'humain (de Waal, 2010). De façon évolutive, l'empathie s'est développée en s'encrant dans deux concepts clés: le mimétisme moteur et la contagion émotionnelle. Les aspects les plus évolués de l'empathie, selon le modèle de la poupée russe de Frans de Waal, sont la prise de perspective et l'imitation véritable. Cette dernière est à différencier du mimétisme moteur par le fait qu'elle ne se produit pas de façon automatique ou biologique (de Waal, 2008). Depuis maintenant quelques années, ces aspects de l'empathie tentent d'être intégrés dans des intelligences artificielles. Ce domaine de recherche est divisé en objets d'étude, soit l'aspect externe de l'interaction robot-humain représentant la communication des émotions, et l'aspect interne réunissant les processus endogènes des émotions induits par l'intelligence artificielle.

La division du domaine provoque certaines divergences dans l'application de l'empathie artificielle. En effet, certains chercheurs, comme Stephan, Moore,

Dautenhahn et Werry, l'appliquent à des problèmes spécifiques, comme en milieu scolaire dans une optique d'apprentissage ou même dans la thérapie pour les TSA, et d'autres se questionnent vers la possibilité d'une application holistique de l'empathie dans un but de produire une intelligence artificielle universelle la plus humaine possible. Le développement de l'aspect interne passe donc par celui de l'aspect externe. L'application en contexte spécifique de l'empathie artificielle stimule de nouvelles connaissances favorisant le développement de l'intelligence artificielle universelle. Cependant, le développement d'un robot universel pouvant s'appliquer à toutes les situations et à chaque individu est loin d'être une réalité (Vircikova, Magyar & Sincak, 2015). L'avènement d'un robot pouvant produire de nombreuses capacités humaines, comme une interaction plus empathique, est d'autant plus probable. Toutefois, le cerveau humain n'étant pas totalement compris, il semble illogique de vouloir appliquer une capacité holistique d'une chose que l'on ne comprend pas dans son ensemble. La conscience morale est un exemple de facultés humaines qui s'avèrent, à ce jour, impossibles à implanter dans une intelligence artificielle. De plus, s'il devient possible d'implanter ces facultés à une intelligence artificielle, il faut se demander si celle-ci n'est pas biaisée par les caractéristiques et les valeurs de la personne qui programme (p. ex, ce qu'elle considère comme bien et mal). Ainsi, malgré les avantages indéniables de l'implantation de l'empathie artificielle à des robots, il reste toutefois à peaufiner les applications spécifiques du domaine avant de se lancer dans une application holistique de l'empathie.

Références

- American Psychiatric Association. (2013). *Diagnostic and Statistical Manual of Mental Disorders – Fifth edition*. Washington, DC: Author.
- Asada, M. (2015). Towards artificial empathy. *International Journal of Social Robotics*, 7(1), 19-33.
- Asada, M., Hosoda, K., Kuniyoshi, Y., Ishiguro, H., Inui, T., Yoshikawa, Y. ... Yoshida, C. (2009). Cognitive developmental robotics: A survey. *IEEE transactions on autonomous mental development*, 1(1), 12-34.
- Baron-Cohen, S. & Wheelwright, S. (2004). The empathy quotient: An investigation of adults with Asperger syndrome or high functioning autism, and normal sex differences. *Journal of Autism and Developmental Disorders*, 34(2), 163-175.
- Damiano, L., Dumouchel, P. & Lehmann, H. (2015). Towards human-robot affective co-evolution overcoming oppositions in constructing emotions and empathy. *International Journal of Social Robotics*, 7(1), 7-18.
- Dautenhahn, K. & Werry, I. (2004). Towards interactive robots in autism therapy: Background, motivation and challenges. *Pragmatics & Cognition*, 12(1), 1-35.
- Decety, J., Jackson, P. & Brunet, E. (2007). The cognitive neuropsychology of empathy. In T. Farrow & P. Woodruff (Eds.), *Empathy in Mental Illness* (239-260). Cambridge: Cambridge University Press.
- De Waal, F. B. (2008). Putting the altruism back into altruism: the evolution of empathy. *Annual Review of Psychology*, 59, 279-300.
- De Waal, F. B. (2010). *The age of empathy: Nature's lessons for a kinder society*. Toronto: Emblem Editions.
- Fan, L., Scheutz, M., Lohani, M., McCoy, M. & Stokes, C. (2017, Août). Do We Need Emotionally Intelligent Artificial Agents? First Results of Human Perceptions of Emotional Intelligence in Humans Compared to Robots. In *International Conference on Intelligent Virtual Agents* (pp. 129-141). Springer, Cham.
- Höök, K. (2008, June). Affective loop experiences—what are they? In *International Conference on Persuasive Technology* (pp. 1-12). Springer, Berlin, Heidelberg.
- Kirby, R., Forlizzi, J. & Simmons, R. (2010). Affective social robots. *Robotics and Autonomous Systems*, 58(3), 322-332.
- Leite, I., Castellano, G., Pereira, A., Martinho, C. & Paiva, A. (2014). Empathic robots for long-term interaction. *International Journal of Social Robotics*, 6(3), 329-341.
- Leite, I., Martinho, C., Pereira, A. & Paiva, A. (2009, Septembre). As time goes by: Long-term evaluation of social presence in robotic companions. In *The 18th IEEE International Symposium on Robot and Human Interactive Communication RO-MAN 09, 2009*. (pp. 669-674). IEEE.

- Lim, M. Y., Aylett, R. & Jones, C. M. (2005, Avril). Empathic interaction with a virtual guide. In *Proceeding of the Joint Symposium on Virtual Social Agents, AISB* (Vol. 5, pp. 122-129).
- Mori, M., MacDorman, K. F. & Kageki, N. (2012). The uncanny valley. *IEEE Robotics and Automation Magazine*, 19(2), 98–100.
- Moore, R. K. (2012). A Bayesian explanation of the ‘Uncanny Valley’ effect and related psychological phenomena. *Scientific Reports*, 2, 864.
- Niculescu, A., van Dijk, B., Nijholt, A., Li, H. & See, S. L. (2013). Making social robots more attractive: the effects of voice pitch, humor and empathy. *International Journal of Social Robotics*, 5(2), 171-191.
- Ratcliffe, M. (2017). Empathy Without Simulation. Dans M. Summa, T. Fuchs & L. Vanzago (Eds.), *Imagination and social perspectives: Approaches from phenomenology and psychopathology* (pp.199-220). New York : Routledge
- Riess, H. (2017). The science of empathy. *Journal of Patient Experience*, 4(2), 74-77.
- Stephan, A. (2015). Empathy for artificial agents. *International Journal of Social Robotics*, 7(1), 111-116.
- Vircikova, M., Magyar, G. & Sincak, P. (2015). The affective loop: A tool for autonomous and adaptive emotional human-robot interaction. In J.-H. Kim et al. (Eds.), *Robot Intelligence Technology and Applications 3* (247-254). Springer, Cham.

Pour citer l'article

Linteau, M.-A. (2019). Humanisation des robots : implanter l'empathie dans l'intelligence artificielle. *Psycause : Revue scientifique étudiante de l'École de psychologie de l'Université Laval*, 9(1), 56-63.

Droits d'auteur

© 2019 Linteau. Cet article est distribué en libre accès selon les termes d'une licence Creative Commons Attribution 4.0 International (de type CC-BY 4.0) qui permet l'utilisation du contenu des articles publiés de façon libre, tant que chaque auteur ou autrice du document original à la publication de l'article soit cité(e) et référencé(e) de façon appropriée.