

## Intelligence artificielle : une machine peut-elle ressentir de l'émotion ?

Certains programmes savent déjà les imiter à la perfection et même influencer sur nos propres émotions.

Le Monde.fr | 12.10.2015 à 14h59 • Mis à jour le 12.10.2015 à 16h20 | Par Morgane Tual ([journaliste/morgane-tual](#))

*Cet article fait partie d'une série consacrée à l'état des lieux de l'intelligence artificielle ([intelligence-artificielle](#)).*

La science-fiction nous abreuve de robots autonomes, si perfectionnés qu'ils disposent d'une conscience, d'émotions et nous inspirent même de l'empathie, à l'instar des êtres humains. Mais loin de la littérature et d'Hollywood, qu'en est-il aujourd'hui dans les laboratoires du monde entier ? L'émotion, et la conscience, apparaissent comme des éléments essentiels pour fabriquer une machine à l'image de l'être humain. Mais elles font aussi partie des plus difficiles à conceptualiser .

### « La vie intérieure, on ne sait pas ce que c'est »

Et c'est là que réside la plus grande difficulté : comment définir – et donc reproduire – l'émotion et la conscience ? « On ne sait pas ce que c'est que la conscience, on n'en connaît pas les fondements. On n'est donc pas capables de créer une machine consciente », tranche Jean-Gabriel Ganascia, chercheur au laboratoire d'informatique de Paris -VI et auteur de *L'Intelligence artificielle* (Editions Le Cavalier Bleu, collection *Idées reçues*, 2007). « Pour cela, il faudrait que la machine perçoive comme nous : la douleur, le plaisir... Et quand bien même, elle ne les percevra pas de la même manière que nous. »

Une analyse partagée par Jean-Michel Besnier, professeur de philosophie à la Sorbonne et spécialiste de l'intelligence artificielle :



**MIT**  
MANAGEMENT  
EXECUTIVE EDUCATION

## MIT online AI program

Transform your organization into an innovative, efficient, and sustainable business

« La vie intérieure, on ne sait pas ce que c'est. L'intelligence artificielle a donc fini par dépouiller la notion de conscience de la notion d'intériorité, et l'a définie simplement en termes de comportement : on peut l'imiter chez les machines. »

A défaut d'être capables de ressentir, les machines peuvent néanmoins simuler, en apparence, des émotions et une conscience : c'est sur ce sujet que se concentrent aujourd'hui les chercheurs en intelligence artificielle ; un domaine de recherche intitulé « informatique affective ». Car finalement, « quelle est la différence entre ressentir et donner les signes extérieurs du ressenti ? », interroge Jean-Michel Besnier :

« Ce sont les signaux qui comptent. En toute rigueur, votre intériorité, je ne sais pas si elle existe... Elle se manifeste à moi par des signes extérieurs. Je ne sais pas ce que c'est que de ressentir une émotion chez un autre être humain. Je peux ressentir de l'empathie, mais je peux aussi en ressentir face à un acteur, qui simule. Je ne peux pas savoir si ces signes émanent d'une intériorité ou d'une simulation. C'est pourquoi pour ceux qui fabriquent des machines, si elles sont capables de simuler, ça suffira. »

## Simuler l'émotion

Catherine Pelachaud fait partie de ceux-là. Directrice de recherche au CNRS et à Télécom-ParisTech, elle fabrique depuis des années des « agents conversationnels », sortes d'avatars capables de discuter avec des êtres humains. Son champ de recherche concerne plus précisément les « comportements non verbaux », soit les signes extérieurs d'émotion transmis par l'avatar. « La machine ne ressent pas, mais elle peut transmettre, souligne la chercheuse. Le ressenti est du domaine de l'homme, et ça doit le rester ! Une machine est là pour pallier des besoins. Pour cela, la simulation peut suffire. »

Les « agents » qu'elle élabore accompagnent leurs paroles de gestes, de mouvements de la tête ou d'expressions du visage qui les rendent plus humains. Ils sont aussi capables de réagir aux émotions transmises par leur interlocuteur. « Dans la communication, le non verbal apporte énormément, il permet de mieux se comprendre. Sans ça, ce serait comme parler à un mur. Ça permet d'oublier qu'on parle à une machine. »

Et afficher une émotion est moins simple qu'il n'y paraît. « Ça peut aller jusqu'à des micro-expressions. Il y a plusieurs types de sourires : si vous pincez les lèvres, si vous plissez les yeux, cela aura différentes significations », explique Catherine Pelachaud.

La chercheuse en psychologie Sylwia Hyniewska a observé, par exemple, grâce à une expérience, que les propos d'agents incapables de soulever la partie externe de leurs sourcils étaient considérés par leurs interlocuteurs comme moins pertinents.

Malgré ce souci du détail, les avatars utilisés par les équipes de recherche en informatique affective ne semblent pas très réalistes. A l'heure où les entreprises d'effets spéciaux sont capables de réaliser des images de synthèses ultra-détaillées, pourquoi se contenter d'agents si schématiques ? « Contrairement au cinéma, qui a des animateurs pour peaufiner chaque expression, nos agents doivent être autonomes et réagir en temps réel », indique Catherine Pelachaud.

Mais surtout, si le réalisme est trop important, « on tombe dans la vallée de l'étrange », prévient-elle. Selon cette théorie du Japonais Masahiro Mori, les représentations très réalistes, mais toujours imparfaites, de l'homme, nous paraissent dérangeantes, voire monstrueuses. Nous serions, en revanche, beaucoup plus enclins à trouver sympathiques et à ressentir de l'empathie pour des représentations de l'humain bien plus schématiques.

Ce robot capable de simuler des émotions, empruntant ses traits à Albert Einstein et développé par l'entreprise Hanson Robotics, en est un bon exemple.

## Détecter les émotions

Mais les programmes développés par l'informatique affective ne se contentent pas de mimer les émotions. Ils doivent aussi être en mesure de détecter celles des humains, et de s'y adapter en temps réel. Pour cela, ils observent et analysent les expressions et les mouvements de leur interlocuteur : s'il regarde ailleurs, s'il montre qu'il n'a pas compris, s'il manifeste un désaccord. Et ce n'est pas simple. Car en plus des émotions « de base », définies par le psychologue américain Paul Ekman (tristesse, joie, peur, colère, surprise, dégoût), il existe des émotions plus complexes à déchiffrer pour un programme.

Comment, par exemple, distinguer la surprise de l'étonnement ou la tension de l'anxiété ? Interpréter un haussement de sourcils n'est pas non plus aisé. Veut-il dire bonjour ? Signifie-t-il la surprise ? L'emphase ? « Pour cela, il faut des informations sur le contexte, et c'est très difficile », souligne la chercheuse. Le comprendre est pourtant indispensable pour répondre à ces signaux de façon appropriée. Sinon, l'agent risque de créer un malentendu, une incompréhension, voire de couper le dialogue.

Car en imitant l'émotion, les programmes les plus avancés sont aussi en mesure... de **générer** de l'émotion chez les humains. Ainsi, le **projet** européen Semaine, auquel participait Catherine Pelachaud, a donné des résultats surprenants. Les agents développés étaient chacun dotés d'un état émotionnel fort, comme la tristesse, la colère ou la joie. Objectif : **amener** leur interlocuteur, humain, vers le même état émotionnel qu'eux. « *Il y a eu des interactions absolument incroyables* », se souvient la chercheuse, qui a mené ces expériences il y a cinq ans.

« *Face à un agent dépressif, le sujet montrait de l'empathie. Parfois, l'état émotionnel du sujet changeait au milieu de la conversation. J'étais surprise qu'il y ait des interactions aussi riches, aussi naturelles. Car nos personnages ne bougeaient pas beaucoup, le modèle était assez simple. Mais le fait que le comportement arrive avec le bon timing nous a montré que nous avons réussi à **obtenir** un élément essentiel de la communication non verbale.* »

Des résultats qui posent aussi des questions éthiques : si un programme est capable de détecter et d'influencer les émotions d'un être humain, que doit-il en faire ? « *Si la machine peut **faire changer** l'état émotionnel d'une personne, on peut **imaginer** qu'elle s'en serve à des fins commerciales par exemple* », prévient Catherine Pelachaud.

D'autres cherchent à **exploiter** cet aspect d'une autre façon. L'Institut pour les **technologies** créatives de l'université de Californie du Sud est l'un des plus avancés dans ce champ de recherche en intelligence artificielle. Les agents qu'il a réussi à fabriquer combinent le langage, la communication non verbale et une apparence réaliste. L'équipe de recherche emploie même le terme de « humains virtuels » pour les **désigner**. Parmi les différents modèles qu'ils ont conçus, **SimSensei** (<http://fict.usc.edu/prototypes/simsensei/>) est une sorte de psychologue, capable de **mener** une conversation avec un patient, en s'intéressant à ses émotions. Mais attention, **précise l'institut sur Youtube** (<https://www.youtube.com/watch?v=ejczMs6b1Q4>), « *SimSensei n'est pas conçu pour la thérapie ou de diagnostic médical, mais se veut un outil pour **appuyer** les équipes médicales* ».

Malgré les avancées dans le domaine de l'informatique affective, on est encore bien loin des prédictions de **Ray Kurzweil** ([/culture/article/2015/02/12/ray-kurzweil-bateleur-pour-google\\_4575483\\_3246.html](/culture/article/2015/02/12/ray-kurzweil-bateleur-pour-google_4575483_3246.html)), le « **pape** » du transhumanisme embauché par **Google** en 2012. Dans un **entretien au magazine américain Wired** (<http://www.wired.com/2013/04/kurzweil-google-ai/>) en avril 2013, il prévoit qu'en 2029, des programmes seront capables « *d'intelligence émotionnelle, d'être drôles, de comprendre des blagues, d'être sexy, aimants et de comprendre l'émotion humaine. (...) C'est ce qui sépare les ordinateurs et les humains aujourd'hui. Je crois que ce fossé va se refermer d'ici 2029.* »

**Lire aussi Ray Kurzweil, le salarié de Google qui veut terrasser la mort** ([/festival/article/2015/09/23/ray-kurzweil-le-salarie-de-google-qui-veut-terrasser-la-mort\\_4767845\\_4415198.html](/festival/article/2015/09/23/ray-kurzweil-le-salarie-de-google-qui-veut-terrasser-la-mort_4767845_4415198.html))

Une vision qui exaspère le philosophe Jean-Michel Besnier : « *Je suis inquiet de **voir** que l'intelligence artificielle impose un point de vue de plus en plus simplificateur sur l'être humain, qu'on ne peut pas **réduire** à ces signaux. Pour comprendre les émotions humaines, moi, je préfère me **plonger** dans la littérature !* »

## En bref :

### Ce dont l'intelligence artificielle est aujourd'hui capable :

- simuler des émotions
- détecter les émotions des humains et y répondre en temps réel
- influencer les émotions de son interlocuteur

### Ce qu'on ne sait pas faire :

- doter un programme de conscience
- un programme capable de ressentir des émotions

### Les progrès qu'il reste à faire :

- distinguer des émotions complexes
- créer des agents ou robots assez réalistes pour **dépasser** la « vallée de l'étrange »

- s'adapter au contexte

**Lire dans la même série** **Intelligence artificielle : une machine est-elle capable de philosopher ?** ([/pixels/article/2015/09/18/intelligence-artificielle-une-machine-est-elle-capable-de-philosopher\\_4763043\\_4408996.html](http://pixels/article/2015/09/18/intelligence-artificielle-une-machine-est-elle-capable-de-philosopher_4763043_4408996.html))

**Lire aussi** **Intelligence artificielle : quand la machine imite l'artiste** ([/pixels/article/2015/09/08/intelligence-artificielle-les-machines-peuvent-elles-etre-creatives\\_4749254\\_4408996.html](http://pixels/article/2015/09/08/intelligence-artificielle-les-machines-peuvent-elles-etre-creatives_4749254_4408996.html))