

Depuis les années 1980, des chercheurs ont développé des systèmes perfectionnés de reconnaissance de la parole, de sorte que cette technologie atteint sa maturité. Pour autant, il y a loin entre transcrire un discours et le comprendre ! Songez à la complexité de la chaîne d'événements cognitifs, neurologiques et musculaires qui se combinent quand deux personnes se parlent. L'un des interlocuteurs formule ses pensées, choisit des mots puis articule, tandis que l'autre reçoit les sons et les autres signaux envoyés, avant d'y déchiffrer le message qui y est enfoui.

La chaîne qui s'établit lors d'un dialogue entre un humain et une machine est plutôt la suivante : les ondes sonores associées à la parole sont d'abord numérisées, puis de nombreux paramètres caractérisant le discours sont extraits de la série numérique obtenue ; ensuite, un logiciel de reconnaissance de la parole convertit ces paramètres en mots et un décodeur sémantique transforme la série de mots ainsi extraite en sens.

Au début de nos recherches sur les robots empathiques, nous nous sommes aperçus que des algorithmes similaires à ceux conçus pour analyser les émotions exprimées par des utilisateurs dans leurs commentaires en ligne pourraient nous aider. Ces algorithmes d'apprentissage automatique traquent la présence, dans leur contenu, de signaux révélateurs de la présence d'émotions. Des mots-clés tels que « chagrin » et « crainte » suggèrent la solitude. L'utilisation répétée d'expressions familières (« *Let's go !* », par exemple) révèle qu'une chanson est énergique. L'information donnée par le style du discours est aussi à prendre en compte : les réponses d'une personne sont-elles assurées et claires, ou hésitantes, parsemées de silences et de formulations non compromettantes ? Élaborées et détaillées, ou brèves et sèches ?

Au cours de nos recherches sur la reconnaissance de l'humeur dans la musique, nous avons entraîné des algorithmes à repérer des signaux émotionnels dans les paroles. Au lieu d'extraire des signatures sonores de chaque morceau de musique, nous avons tiré, à partir des paroles, des séries de mots – des *n*-grammes – dont nous nous sommes servis pour entraîner 14 classificateurs, chacun responsable de la détection d'une des 14 humeurs définies précédemment.

« Nous avons obtenu une machine capable d'extraire l'humeur de n'importe quel morceau de musique »

Nous avons aussi défini ce que nous nommons des « étiquettes de partie de discours », que nous avons prises en compte dans nos définitions des signatures de telle ou telle humeur. Sur la base d'une collection suffisante de jeux de ces paramètres extraits de discours humains, un ordinateur peut inférer des approximations statistiques des règles grammaticales de n'importe quelle langue. Ce sont ces règles élaborées au cours d'un entraînement qui permettent à des programmes tels *Siri* de comprendre à leur manière un discours ou encore à un programme tel *Google Translate* de traduire un texte.

Dès qu'une machine peut comprendre le contenu d'un discours, elle peut comparer ce contenu avec la façon dont il est émis. Un algorithme pourra détecter la contradiction entre le sens d'un discours et les signaux émotionnels qui l'accompagnent. C'est le cas, par exemple, quand la phrase « Je suis si heureux de passer le week-end à travailler » est accompagnée d'un soupir. Dans un tel cas, l'ordinateur calculera la probabilité que la déclaration humaine soit en réalité un sarcasme.

Intentions complexes

Afin de détecter les intentions plus complexes, une machine capable d'appréhender le contenu d'un discours et les émotions qui le baignent pourra aussi associer les informations recueillies à des données extérieures susceptibles d'aider à son interprétation. Si quelqu'un s'exclame « J'ai faim ! », elle tiendra compte, pour déterminer la meilleure réponse à apporter, de l'heure ou du lieu ainsi que de sa mémoire des usages et préférences de la personne. Au milieu du Sahara, elle s'abstiendra de proposer la recherche d'une « pizzeria dans le quartier » ; à la maison à l'heure du déjeuner, en revanche, elle pourrait répondre : « Souhaitez-vous que je vous prépare un en-cas ? »

Début 2015, des étudiants et des chercheurs postdoctoraux de mon laboratoire ont commencé à réunir nos différents modules de reconnaissance de la parole et de reconnaissance de l'émotion pour créer une machine empathique prototype que nous appelons *Zara la Supergirl*. Il a fallu des centaines d'heures de données pour entraîner *Zara*, mais aujourd'hui, ce programme tourne

Quels obstacles à l'insertion de robots en société ?

Membre de l'équipe Larsen (de l'anglais *Life-long Autonomy and Interaction Skills for Robots in a Sensing Environment*) à Nancy, Serena Ivaldi travaille sur les interactions humains-robots. Elle nous parle des principaux obstacles déjà identifiés s'opposant à l'acceptation des robots par les humains.



Serena IVALDI, roboticienne, travaille chez Inria dans l'équipe Larsen, à Nancy.

POUR LA SCIENCE

Pensez-vous que les robots vont se multiplier dans la société ?

Serena IVALDI : Il est certain que ce phénomène est pour bientôt. La situation est comparable à celle qui a précédé l'introduction des ordinateurs et des téléphones portables : leur usage par chacun a paru superflu jusqu'à ce qu'il s'impose de façon explosive, et aujourd'hui plus personne ne s'interroge sur l'utilité de ces technologies.

PLS

Comment testez-vous l'acceptation des robots par les humains ?

S.I. : Nous étudions d'abord les interactions entre robots et humains au sein de notre laboratoire en faisant appel à des personnes ordinaires, puis nous recommençons dans le milieu où les robots doivent être insérés, où les gens ont des particularités. Par exemple, nous étudions l'insertion de robots aides soignants dans un établissement d'hébergement pour personnes âgées dépendantes.

PLS

Quels sont les principaux obstacles à l'insertion de robots dans la société ?

S.I. : Il y en a de plusieurs sortes. À ce stade, le principal obstacle est simplement le coût prohibitif des robots ! Sinon, le problème principal est qu'il nous faut apprendre à bien organiser les interactions des personnes avec les robots que l'on va introduire. Un robot peut être vu comme un ordinateur doté d'un corps. Toutefois, quand un ordinateur est personnifié par un corps, les interactions humains-robots deviennent comparables aux interactions humaines, qui sont multimodales et complexes. En particulier, un corps c'est physique, et cela entraîne des interactions physiques !

PLS

Quels éléments interviennent au cours d'une interaction humain-robot ?

S.I. : Quand un humain est en présence d'un robot – surtout humanoïde – maints détails influencent profondément la façon dont il va tenter de communiquer avec lui : sa distance par rapport à lui ; l'apparence de la machine ; son « regard » ; les attentes de l'humain, notamment sur le rôle de la machine ; les forces développées par chacun d'eux dans la situation ; etc.

PLS

De quelles forces parlez-vous ?

S.I. : Traditionnellement, les robots industriels étaient séparés des ouvriers, car ils développent des forces considérables ; aujourd'hui on tolère davantage de proximité, afin de rendre possibles des gestes, voire du travail collaboratif. De là l'importance de parfaitement contrôler les forces en jeu.

PLS

Pouvez-vous décrire une expérience que vous avez menée pour caractériser les difficultés de la communication humain-robot ?

S.I. : À l'Isir, l'Institut de robotique de l'université Pierre-et-Marie-Curie, nous avons fait interagir iCub – un petit robot en forme d'enfant – avec une cinquantaine d'adultes non versés en robotique. Nous leur avons demandé par exemple de collaborer avec iCub pour désigner le plus lourd de deux objets ou encore l'objet le plus approprié (le plus utile par exemple) dans une certaine situation. Nous avons constaté que les humains adoptent avec le robot des attitudes variables – détachées, pédagogiques, parentales... –, qui traduisent à quel degré ils sont prêts à le considérer comme un véritable agent social.

PLS

Qu'est-ce qui détermine l'attitude de l'humain ?

S.I. : Soit un humain tend à considérer un robot comme « un porteur de capteurs », auquel cas il tend à accepter ses décisions comme on accepte le verdict du thermomètre. Soit quelque chose fait qu'il considère plus ou moins inconsciemment le robot comme un agent social et, dans ce cas, il tend à penser que, la machine ne pouvant avoir son expérience sociale, ses décisions sont invalides. Par exemple, si l'on demande à une personne et à un robot : « Que faut-il mettre dans son sac pour aller à la piscine ? », l'humain, qui tendra à ressentir que son compagnon « n'allant pas à la piscine, ne peut pas savoir », rejettera sa décision. Son attitude change quand il a l'impression d'enseigner au robot quel comportement adéquat avoir en société ou avec lui... Nous avons même constaté avec stupéfaction que certaines personnes se mettent à essayer de convaincre le robot, comme s'il n'avait pas encore appris...

PLS

Ainsi, la question essentielle semble être celle de la confiance ?

S.I. : Oui, beaucoup se méfient encore des machines qui jouent un rôle social.

PLS

Dès lors, comment leur faire mieux accepter les robots dans la société ?

S.I. : En leur donnant assez de contrôle sur la façon dont les robots se comporteront autour d'eux, tout particulièrement si ces derniers jouent un rôle social et encore plus s'ils sont humanoïdes !

Propos recueillis par François Savatier

sur un ordinateur de bureau. Pour l'instant, Zara est un robot virtuel, représenté sur un écran par un personnage de dessin animé.

Au début de toute conversation, elle demande toujours : « Patientez s'il vous plaît pendant que j'analyse votre visage. » Les algorithmes de Zara exploitent en effet les images prises par une caméra afin de déterminer le sexe et l'origine ethnique de son interlocuteur. Elle tente d'en déduire sa langue (Zara comprend l'anglais et le mandarin, et apprend le français) afin de l'interroger dans celle-ci : « Quel est votre plus ancien souvenir ? Parlez-moi de votre mère. Comment se sont passées vos dernières vacances ? Racontez-moi une histoire mettant en scène une femme, un chien et un arbre. » Exploitant les expressions du visage de son interlocuteur, les caractéristiques acoustiques de sa voix et le contenu de ses réponses, Zara évolue ainsi vers une forme d'empathie artificielle à l'égard de son interlocuteur.

Après cinq minutes de conversation, elle jauge à sa manière la personnalité de son interlocuteur, qu'elle interroge ensuite sur son attitude envers les machines empathiques.

Zara n'est qu'un prototype, mais comme elle est constituée d'algorithmes d'apprentissage automatique, elle ne peut que devenir plus « intelligente » et donc plus empathique à mesure que ses interactions avec de plus en plus de personnes l'entraîneront. Pour l'heure, sa « connaissance » des humains découle du seul entraînement que représentent ses interactions avec les étudiants de troisième cycle de mon laboratoire. Cette année, nous prévoyons de donner un corps à Zara en l'installant dans un robot humanoïde.

Il serait prématuré de dire que nous entrons dans l'ère des robots empathiques. Même quand les premiers descendants de Zara seront sur le marché, ils seront loin d'être parfaits. Du reste, j'en suis venue à penser que se concentrer sur la fabrication de machines parfaites nous éloignerait du but. L'important est de rendre nos robots plus humains, même s'ils sont imparfaits, ce qui, après tout, est très humain ! Si nous y parvenons, les robots empathiques ne deviendront pas les machines dominatrices que certains redoutent. Ils seront au contraire nos soignants, nos enseignants et une nouvelle sorte d'amis. ■

■ BIBLIOGRAPHIE

D. Su, P. Fung et N. Auguin, **Multimodal music emotion classification using Adaboost with decision stumps**, présentation à la conférence ICASSP 2013 (International Conference on Acoustics, Speech and Signal Processing), Vancouver, 26 au 31 mai 2013.

X. Zuo, T. Li et P. Fung, **A multilingual database of natural stress emotion**, présentation à LREC 2012 (8th International Conference on Language Resources and Evaluation), Istanbul, 21 au 27 mai 2012.

M. Anderson et S. Leigh Anderson, **Robot be good**, *Scientific American*, n° 303, octobre 2010.

Dans l'interêt de la science

mathieu vidard | la tête au carré 14:05-15:00

france inter venez
franceinter.fr