

Un algorithme coopératif et communicant

CHRONIQUE

PAR PAUL SEABRIGHT

Professeur de sciences économiques
à l'Institut d'études avancées de
Toulouse

Dans sa chronique, l'économiste Paul Seabright s'appuie sur une étude récente qui démontre que les algorithmes les plus efficaces sont ceux à qui l'on donne, en plus de la capacité à apprendre, celle de communiquer.

Temps de lecture : 3 min

Recherches. Les prouesses de l'intelligence artificielle sont louées quotidiennement dans les médias, et l'on pourrait croire que les algorithmes essaient de dépasser les êtres humains dans tous les domaines de nos activités. En fait, la majorité des recherches en intelligence artificielle essaient de reproduire et de perfectionner des compétences que l'on peut mettre directement au service de l'être humain (comme avec la robotique) ou qui servent dans les combats à somme nulle (comme les échecs, le poker ou le go).

Rares sont les projets qui travaillent sur les compétences utiles pour les tâches qui mêlent le conflit et la coopération, où il y a des bénéfices à la coopération, mais des conflits sur la répartition de ces bénéfices. Pourtant ces tâches sont au cœur de la société humaine, notamment dans le domaine de l'échange économique.

Une étude qui vient de paraître s'adresse directement à ce défi passionnant (« Cooperating with machines », par Jacob Crandall, Mayada Oudah, Jean Tennom, Fatimah Ishowo-Oloko, Sherief Abdallah, Jean-François Bonnefon, Manuel Cebrian, Azim Shariff, Michael A. Goodrich et Iyad Rahwan, *Nature Communication* 9/233, 2018, [lien vers PDF en anglais](#)).

Expériences économiques

Les auteurs ont cherché à construire des algorithmes capables d'agir de manière coopérative autant avec d'autres algorithmes qu'avec des êtres humains. Leur tâche a été compliquée par le fait que la coopération humaine est facilitée par d'autres qualités que l'intelligence purement cognitive – surtout l'empathie, l'intuition, la sensibilité aux normes et à la communication.

Lors des expériences, les sujets humains arrivent plus facilement à coopérer lorsqu'ils ont la possibilité de communiquer entre eux, même si cette communication ne les engage en rien. En serait-il autant pour les algorithmes ? La réponse a été positive.

Les chercheurs ont mis en place des expériences économiques sous forme de jeux répétés aléatoires de divers types, comme le dilemme du prisonnier. Ils y ont fait jouer un grand nombre d'algorithmes, entre eux et avec des sujets humains – les auteurs n'en parlent pas mais on peut supposer que la déontologie professionnelle leur a permis de rémunérer les algorithmes bien moins cher que les sujets humains ; pour l'instant, les algorithmes ne siègent pas encore aux comités d'éthique...

Devancer des humains

Pour réussir l'expérience, un algorithme devait montrer des compétences généralistes – c'est-à-dire être efficace dans plusieurs types de jeu et avec plusieurs types de partenaire. Mais l'algorithme qui réussit de loin le mieux combine deux caractéristiques supplémentaires : il apprend en observant le comportement d'autres algorithmes déjà efficaces, et il communique, en envoyant aux autres sujets des messages destinés à influencer leur comportement. La capacité de communication le rend bien plus fort que tout autre algorithme qui ne fait qu'apprendre.

L'algorithme vainqueur devance même la plupart des sujets humains, pour deux raisons. La première est qu'une fois la coopération établie avec un partenaire, l'algorithme ne cède jamais à la tentation de le trahir, contrairement à beaucoup de sujets humains... La deuxième est que l'algorithme tient toujours parole, ce qui rend ses communications plus fiables que les leurs.

Ceux qui craignent le remplacement des humains par des robots peuvent au moins se consoler en pensant que nous pourrions résister en devenant plus loyaux et plus fiables que nous ne le sommes actuellement, au moins lorsqu'il s'agit du type de jeu qui figure dans ces expériences économiques. Pour « Game of Thrones », ce serait sans doute autre chose.

En outre, nous sommes à Toulouse particulièrement conscients que le jeu qui compte vraiment dans la vie est le rugby, et il n'est pour l'instant pas prévu que des équipes de robots puissent jouer mieux que des équipes humaines. Ce qui est vraiment à craindre, c'est que le jour où cela arrivera il n'y ait que des robots parmi les spectateurs, car les êtres humains ne pourront plus se payer l'entrée du stade.

Par PAUL SEABRIGHT Professeur de sciences économiques à l'Institut d'études avancées de Toulouse

Publié Il y a 25 minutes