

La danse de la conscience dans un cerveau synthétique

Un système artificiel complexe pourrait-il avoir une conscience ? La question est fascinante. Un tel système ne sera pas construit dans les prochaines décennies, ni dans les prochains siècles, mais une simple expérience de pensée montre qu'un cerveau artificiel correctement organisé aurait le même type de conscience qu'un être humain.

Considérons un système constitué de puces en silicium qui ont la même organisation et le même fonctionnement que les neurones de notre cerveau. Chaque puce fait exactement ce que fait son analogue naturel, tout en étant reliée de la même manière aux éléments environnants. Le comportement de ce système artificiel sera exactement le même que le nôtre. Sera-t-il conscient de la même manière que nous ?

Supposons qu'il ne le soit pas : soit il a une conscience différente (il ressent du bleu lorsque nous ressentons du rouge), soit il n'a pas de conscience du tout. Nous allons examiner le premier cas (le raisonnement est le même dans les deux cas).

Comme les puces et les neurones ont les mêmes fonctions, ils sont interchangeables, à condition d'avoir une interface adaptée. Les puces peuvent donc remplacer les neurones, et l'on produit un continuum de situations où une proportion croissante de neurones est remplacée par des puces. Dans ce continuum, la conscience du système change aussi. Nous pourrions remplacer

tous les neurones de notre cortex visuel par des puces. Le cerveau résultant, dont le cortex visuel est artificiel, a une conscience différente de l'original : là où précédemment nous voyions du rouge, nous verrions du violet.

Imaginons alors que nous relions les deux cortex visuels au cerveau, par l'intermédiaire d'un commutateur à deux positions. Lorsque le commutateur est dans une position, nous utilisons le cortex visuel naturel ; dans l'autre position, c'est le cortex artificiel qui est activé. Notre sensation changerait alors du rouge au violet, ou réciproquement. Lorsqu'on actionne le commutateur de manière répétitive, nos sensations "dansent" entre les deux états conscients différents (rouge et violet).

Comme l'organisation de notre cerveau ne change pas, il ne peut y avoir de changement comportemental quand on bascule le commutateur. Par conséquent, lorsqu'on nous demande ce que nous voyons, nous disons que rien n'a changé : nous ne voyons que du rouge, bien que les deux couleurs dansent sous nos yeux. Cette conclusion est absurde : l'hypothèse qu'un système artificiel, dont l'organisation et le fonctionnement sont les mêmes que ceux d'un cerveau neuronal, a des sensations conscientes différentes est donc absurde. Les systèmes qui ont la même organisation ont donc la même conscience.



Si un cerveau synthétique en tout point semblable au nôtre avait une conscience différente, une pomme pourrait nous apparaître alternativement rouge ou bleue.

L'information, ou du moins certaines informations, posséderait donc deux aspects de base : l'un physique et l'autre intime. Cette hypothèse est un principe fondamental, sur lequel reposeraient les relations qui unissent les opérations physiques et la conscience. Partout où nous trouvons de la conscience, elle est l'un des aspects d'un état d'information, dont l'autre aspect est codé dans un processus physique à l'intérieur du cerveau. Il faudrait étoffer cette proposition pour en faire une théorie satisfaisante, mais elle remplit convenablement les principes précédemment mentionnés (des systèmes qui ont la même organisation matérialisent la même information, par exemple) et elle expliquerait de nombreux aspects de la conscience.

L'idée est au moins compatible avec plusieurs autres, telle la proposition du physicien John Wheeler, selon laquelle l'information est un fondement de la physique de l'Univers. Les lois de la physique pourraient être reformulées en termes d'information : les lois physiques et psychophysiques se rejoindraient alors. Les théories de la physique et de la conscience pourraient même être réunies en une super théorie de l'information.

La généralité excessive du concept d'information pose toutefois un problème : un thermostat, qui contient de l'information, a-t-il une conscience ? Il y a au moins deux réponses possibles. Nous pourrions inclure dans les lois fondamentales le fait qu'une partie seulement de l'information a un aspect de conscience, peut-être en fonction du traitement physique qu'elle subit. Nous pourrions, au contraire, admettre que toute information a un aspect de conscience : lorsqu'il y a un traitement complexe de l'information, il y a une conscience complexe, et lorsqu'il y a un traitement simple de l'information, il y a une conscience simple. Dans ce cas, même un thermostat pourrait avoir une conscience, mais elle serait beaucoup plus simple que même la conscience fondamentale que nous avons d'une couleur, et il n'aurait aucune émotion ni pensée associée. Une telle idée semble bizarre, mais si la conscience est vraiment fondamentale, elle doit être largement répandue. De ces deux possibilités, nous devrions choisir celle qui s'intégrerait dans la théorie la plus puissante.

Ces idées produiront-elles une théorie plus vaste, qui prédira la structure précise de notre conscience à partir des

mécanismes physiques de notre cerveau ? Si le projet échoue et que mes idées se révèlent fausses, d'autres idées seront explorées, et d'autres théories fondamentales seront sans doute élaborées. Nous connaissons peut-être un jour le plus grand mystère de l'esprit.

David CHALMERS est membre du département de philosophie de l'Université de Santa Cruz.

Daniel DENNETT, *La conscience expliquée*, Éditions Odile Jacob, 1993.

Roger PENROSE, *Les ombres de l'esprit*, InterÉditions, 1995.

David CHALMERS, *Absent Qualia, Fading Qualia, Dancing Qualia*, in *Conscious Experience*, sous la direction de Thomas Metzinger, Ferdinand Schöningh, 1995.

Explaining Consciousness : The «Hard Problem», numéro spécial du *Journal of Consciousness Studies*, vol. 2, n° 3, automne 1995.

Daniel PINKAS, *La matérialité de l'esprit*, Éditions La Découverte, 1995.

The Nature of Consciousness : Philosophical and Scientific Debates, sous la direction de Ned Block, Owen Flanagan et Güven Güzeldere, MIT Press (à paraître).
