

Esprits numériques

Séance 4: arguments contre la conscience des machines

Jonathan Simon

Programme

- 1) Une taxonomie en quatre parties des objections (à la possibilité d'une conscience numérique)
- 2) Consideration de chacun de ses parties

Taxonomie

- 1) Les correspondances entrée/sortie que nous associons à la conscience ne correspondent pas à ce que les ordinateurs peuvent faire.*
- 2) Bien que la conscience puisse être un processus de calcul, elle peut être différente de tout processus de calcul qui peut fonctionner sur une « vraie » machine de Turing (l'émulation n'est pas la mise en œuvre).*
- 3) La conscience concerne des détails de l'organisation (causale) qui sont orthogonaux aux calculs effectués par un système.*
- 4) La conscience concerne la nature intrinsèque et non la fonction.*

Taxonomie

- 1) Les correspondances entrée/sortie que nous associons à la conscience ne correspondent pas à ce que les ordinateurs peuvent faire.*

Lucas, Penrose, Landgrebe + Smith, Dreyfus (?)

Taxonomie

- 1) *Les correspondances entrée/sortie que nous associons à la conscience ne correspondent pas à ce que les ordinateurs peuvent faire.*

Ex: La conscience concerne la résolution de problèmes mathématiques qui ne sont pas calculables

Taxonomie

2) Bien que la conscience puisse être un processus de calcul, elle peut être différente de tout processus de calcul qui peut fonctionner sur une « vraie » machine de Turing (l'émulation n'est pas la mise en œuvre).

Piccinini, Block (?), Godfrey-Smith (?), Maudlin (?), Les énéactivistes (?)

Taxonomie

2) Bien que la conscience puisse être un processus de calcul, elle peut être différente de tout processus de calcul qui peut fonctionner sur une « vraie » machine de Turing (l'émulation n'est pas la mise en œuvre).

Un système doit être analogique ou neuromorphique pour calculer de la bonne manière pour la conscience.

Taxonomie

3) La conscience concerne des détails de l'organisation (causale) qui sont orthogonaux aux calculs effectués par un système.

Searle, Tononi, Block (?), Godfrey-Smith (?), Maudlin (?), Les énéactivistes (?), Dreyfus (?)

Taxonomie

3) La conscience concerne des détails de l'organisation (causale) qui sont orthogonaux aux calculs effectués par un système.

ex : La conscience a à voir avec l'organisation fonctionnelle pertinente pour la vie, mais cette organisation n'est pas vraiment computationnelle

Taxonomie

4) La conscience concerne la nature intrinsèque et non la fonction.

Ex: theories quantiques, certains versions de monism russelien...

Taxonomie

Compare la taxonomie de Marr :

- 1) Niveau computationnel*
- 2) Niveau algorithmique*
- 3) Niveau de l'implémentation*
- 4) (autre : niveau non fonctionnel)*

Taxonomie

À noter: Les objections à la conscience numérique ne sont pas les mêmes que les objections au fonctionnalisme computationnel

Taxonomie

... en principe, on peut approuver les objections à la conscience numérique des groupes 1) et 2) tout en étant un fonctionnaliste computationnel.

... Et d'un autre côté, on peut rejeter le fonctionnalisme computationnel pour des raisons indépendantes, mais aussi résister à toutes les objections à la conscience numérique

Taxonomie

À noter aussi:

Pour chaque catégorie, on peut distinguer les objections de principe et les objections de pratique : les premières soutiennent que quoi qu'il arrive, la conscience numérique est impossible, les secondes soutiennent que nous ne savons pas comment y parvenir / qu'elle est loin d'être atteinte.

Taxonomie

Ex: 1) Les correspondances entrée/sortie que nous associons à la conscience ne correspondent pas à ce que les ordinateurs peuvent faire.

Taxonomie

Objection de principe: Penrose: La conscience signifie que nous pouvons résoudre des fonctions non calculables (vérités de Godel, etc.).

Objection de pratique: Walsh: La conscience implique la résolution d'un problème très difficile que personne ne sait encore résoudre avec un ordinateur (par exemple, l'accomplissement de la tâche du prix ARC).

Taxonomie

À noter aussi:

Certaines objections sont ambiguës ou difficiles à classer. Godfrey-Smith soutient-il simplement que la fonction que la conscience calcule est très compliquée (niveau 1), ou qu'elle implique une sorte spécifique de forme dynamique et analogique de calcul (niveau 2), ou qu'elle implique une organisation causale qui est orthogonale à l'organisation computationnelle (niveau 3) ?

Taxonomie

À noter aussi:

- Block soutient-il simplement que l'émulation n'est pas la mise en œuvre, et que la nation chinoise ne fait qu'émuler (niveau 2), ou que ce qui importe est la structure causale orthogonale à la structure de calcul (niveau 3), ou que le carbone est magique (niveau 4) ?

The image features a white background with two teal-colored geometric shapes. On the left, there is a large teal trapezoid that tapers towards the right. On the right side, there is a smaller teal triangle that tapers towards the left. The word "BLOCK" is centered between these two shapes.

BLOCK

Block

L'argument de la nation chinoise :

Version neuroscience computationnelle :

chaque personne joue le rôle d'un neurone,
envoie des SMS aux personnes de son réseau...
(où le neurone associé se déclencherait).

Block

L'argument de la nation chinoise :

Version machine de Turing: Certaines personnes brandissent des signes disant «0» ou «1». Un directeur lit et tourne les pages du livre de la machine-table, et demande aux personnes d'invertir leurs signes (cf. « Le problème de trois corps » de Liu Cixin)







Block

- 1) Si le fonctionnalisme est vrai, un tel système (qui me reproduit à un niveau fonctionnel) serait conscient
- 2) Un tel système n'est pas conscient
- 3) Par conséquent, le fonctionnalisme n'est pas vrai

Block

Question : quelles limites cela impose-t-il, exactement, aux systèmes qui sont conscients ?

Remarque : certains présentent Block comme soutenant que la conscience n'est pas du tout une question de fonctionnalité, mais plutôt qu'il y a quelque chose de magique dans le carbone par rapport au silicium.

Ce n'est pas son point de vue. Autres interprétations:

Block

Option 1) Seules quelques contraintes supplémentaires sur la fonction :

elle doit fonctionner dans certaines contraintes d'espace et de temps, **ou peut-être**,

elle ne peut pas utiliser d'autres systèmes, mettant en œuvre la même fonction, comme composants.

(Pas de systèmes conscients qui fonctionnent beaucoup plus lentement que nos cerveaux, ou qui sont spatialement beaucoup plus grands, ou qui sont composés d'autres systèmes conscients)

Block

Option 2) La conscience nécessite des calculs analogiques (analog computation) qui ne correspondent pas à des calculs numériques (digital computation)

Option 3) La conscience nécessite des fonctions non informatiques (non-computable functions)

SEARLE

Searle

Searle s'intéresse à la nature de l'intelligence plutôt qu'à la conscience, mais l'argument peut être adapté facilement.

Searle

Cf : la réponse aux objections de trivialité, et les théories mécanistes vs sémantiques de l'informatique.

Searle

La chambre chinoise

Un homme dans une chambre avec un livre. Des phrases en chinois sont passées sous la porte. Le livre a une entrée pour chaque phrase possible, indiquant à l'homme ce qu'il doit écrire, et repasse sous la porte.

Le livre est la table de la machine de Turing pour un chatbot chinois. Mais l'homme ne parle pas chinois...

Searle

La chambre chinoise: edition conscience

Un homme dans une pièce avec un livre. Des phrases d'entrée sont passées sous la porte. Le livre a une entrée pour chaque phrase possible, indiquant à l'homme ce qu'il doit écrire, et repasse sous la porte. Le système les convertit en commandes comportementales, mettant en œuvre la machine de Turing pour la conscience...

Searle

- 1) si le fonctionnalisme est vrai un tel système parle chinois / est conscient
- 2) Un tel système ne parle pas chinois / n'est pas conscient (en tant que système)
- 3) Par conséquent, le fonctionnalisme est faux

Searle

Réponse du système : peut-être que l'homme ne parle pas chinois mais la pièce oui...

Les fonctions pertinents
ne sont pas réalisables
par ordinateur
numérique
(d'aujourd'hui)

Peter Godfrey-
Smith

Les fonctions pertinents ne sont pas réalisables par ordinateur numérique: PGS

Les liens profonds entre la vie, la proto-cognition, la cognition et la conscience

Les fonctions pertinents ne sont pas réalisables par ordinateur numérique: PGS

Le cerveau n'est pas une machine à pousser et à tirer, comme l'envisageait Leibniz, mais une série de mécanismes qui fonctionnent en biaisant les tendances dans une tempête moléculaire (qui ne peut vraiment s'autoalimenter que dans un certain régime : aqueux, petit et chaud...).

Les fonctions pertinents ne sont pas réalisables par ordinateur numérique: PGS

Réponse de PGS à l'argument des qualia dansants de Chalmers

Les fonctions pertinents ne sont pas réalisables par ordinateur numérique: PGS

- 1) Il est possible qu'il existe un robot, fonctionnellement isomorphe à moi, Z, sans qualia
- 2) Si 1), alors il est possible qu'il y ait une série de cas, chacun avec un neurone remplacé par une puce en silicium, entre moi et Z
- 3) soit le passage de mon expérience à aucune expérience est soudain, sur une seule étape, soit il est progressif (qualia évanescents)..

Les fonctions pertinents ne sont pas réalisables par ordinateur numérique: PGS

- 1) Il est possible qu'il existe un robot, fonctionnellement isomorphe à moi, Z, sans qualia
- 2) Si 1), alors il est possible qu'il y ait une série de cas, chacun avec un neurone remplacé par une puce en silicium, entre moi et Z

Les fonctions pertinents ne sont pas réalisables par ordinateur numérique: Seth

Beast Theory

La conscience est un moyen spécifique de réaliser les impératifs généraux (biologiques) de régulation et d'auto-entretien (par le biais du codage prédictif et de la modélisation générative).

Les fonctions pertinentes ne sont pas réalisables par ordinateur numérique: Seth

«... Cet impératif n'est pas limité à un seul niveau. L'auto-entretien des systèmes vivants (métabolisme) descend jusqu'au niveau des cellules. Chaque cellule de ton corps régénère continuellement les conditions nécessaires à sa propre intégrité au fil du temps. On ne peut pas en dire autant d'une machine actuelle ou d'un futur proche..... Les processus de régulation physiologique qui sous-tendent la conscience sont issus des processus fondamentaux de la vie...»

Note d'interprétation

PGS et Seth soutiennent-ils que 1) la conscience est intrinsèquement biologique, même si une machine de Turing peut «faire» toutes les mêmes choses ?

Ou bien affirment-ils que 2) il serait très difficile pour une machine de Turing de faire les choses en question, étant donné les complexités impliquées (le niveau des ions plutôt que celui des neurones, en tant que nœuds, signifie des quadrillions de paramètres plutôt que des trillions)?

OU bien que 3) les fonctions en question ne sont pas du tout calculables ?

Les fonctions pertinents ne
sont pas computables en
aucun sens (Lucas, Penrose,
Landgrebe + Smith)

Lucas et Penrose

Lucas + Penrose:

1) Pour chaque système formel (système de calcul) S , il existe une vérité $G(S)$ qu'il ne peut pas prouver.

2) Pour chacun de ces $G(S)$, les humains peuvent connaître $G(S)$

Conc) Par conséquent, aucun système formel ne décrit l'intelligence humaine

Landgrebe et Smith (et Dreyfus et les enactivistes)

- 1) Les systèmes conscients font face à («**cope with**») des systèmes complexes
 - 2) Les systèmes informatiques ne peuvent faire face aux systèmes complexes que s'ils peuvent les modéliser de manière adéquate.
 - 3) Les systèmes informatiques ne peuvent pas modéliser correctement les systèmes complexes (il s'agit des problèmes non-computables)
- Conc) Par conséquent, les systèmes conscients ne sont pas des systèmes informatiques.