

Esprits numériques

Séance 4: arguments contre la conscience des machines

Jonathan Simon

Programme

- 1) La fonction n'est pas suffisante pour la conscience (Block, Searle)
- 2) Les fonctions pertinentes ne sont pas réalisables par ordinateur numérique (PGS, Seth)
- 3) Les fonctions pertinentes ne sont pas computables en aucun sens (Lucas, Penrose, Landgrebe + Smith)

Aperçu des raisons de nier que les machines numériques puissent être conscientes.

1) Certaines ne concernent pas du tout la fonction, mais plutôt **l'histoire, ou l'essence métaphysique** (un système conscient doit avoir évolué naturellement plutôt que d'avoir été conçu par un autre être conscient. La conscience est liée à la nature intrinsèque du carbone, plutôt que du silicium...).

Certains interprètent Block et Searle de cette façon

Aperçu des raisons de nier que les machines numériques puissent être conscientes.

2) D'autres font appel à **une structure ou à des modèles sous-jacents qui *pourraient* jouer un rôle causal ou fonctionnel**, mais où c'est le modèle plutôt que la fonction qui provoque/réalise la conscience, et le modèle est incompatible avec celui d'une machine de Turing.

-- la théorie de l'information intégrée correspond à cette description. On peut également interpréter Block et Searle de cette façon.

Aperçu des raisons de nier que les machines numériques puissent être conscientes.

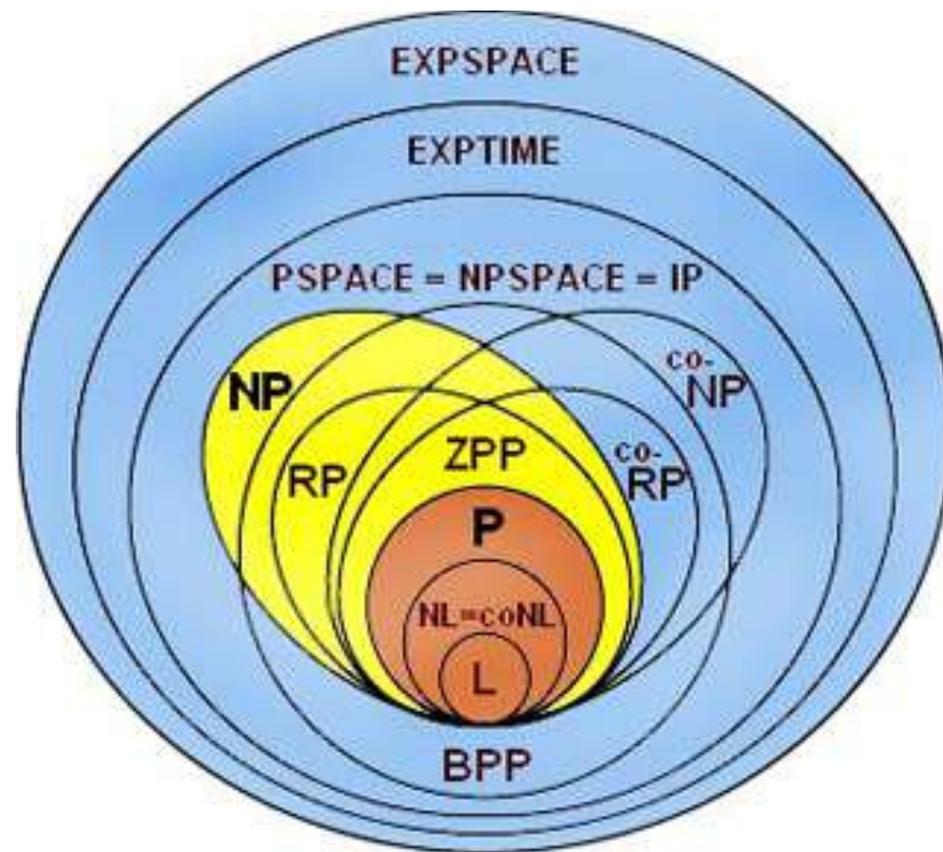
3) D'autres admettent que la conscience correspond à une fonction turing-comptable, mais soutiennent qu'il s'agit **d'une fonction si complexe que seul le matériel biologique est suffisamment sophistiqué pour la mettre en œuvre** (compte tenu des ressources limitées de notre monde).

On peut lire Peter Godfrey-Smith de cette façon

Aperçu des raisons de nier que les machines numériques puissent être conscientes.

C'est peut-être simplement parce que le problème a beaucoup d'entrées...

Ou parce que le problème implique une classe de complexité supérieure, par exemple NP mais pas P ... comme trouver les facteurs premiers d'un grand nombre. (cf. l'informatique quantique...)



Aperçu des raisons de nier que les machines numériques puissent être conscientes.

4) Enfin, il y a ceux qui admettent que la conscience est fonctionnelle, **mais qui nient que les fonctions en question soient numériquement calculables.** (peut-être sont-elles analogiquement calculables mais pas numériquement calculables, peut-être ne sont-elles pas calculables dans tous les sens du terme).

Penrose, Smith et Landgrebe sont de cet avis. Il est possible que Peter Godfrey-Smith et Anil Seth soient également de cet avis.

Aperçu

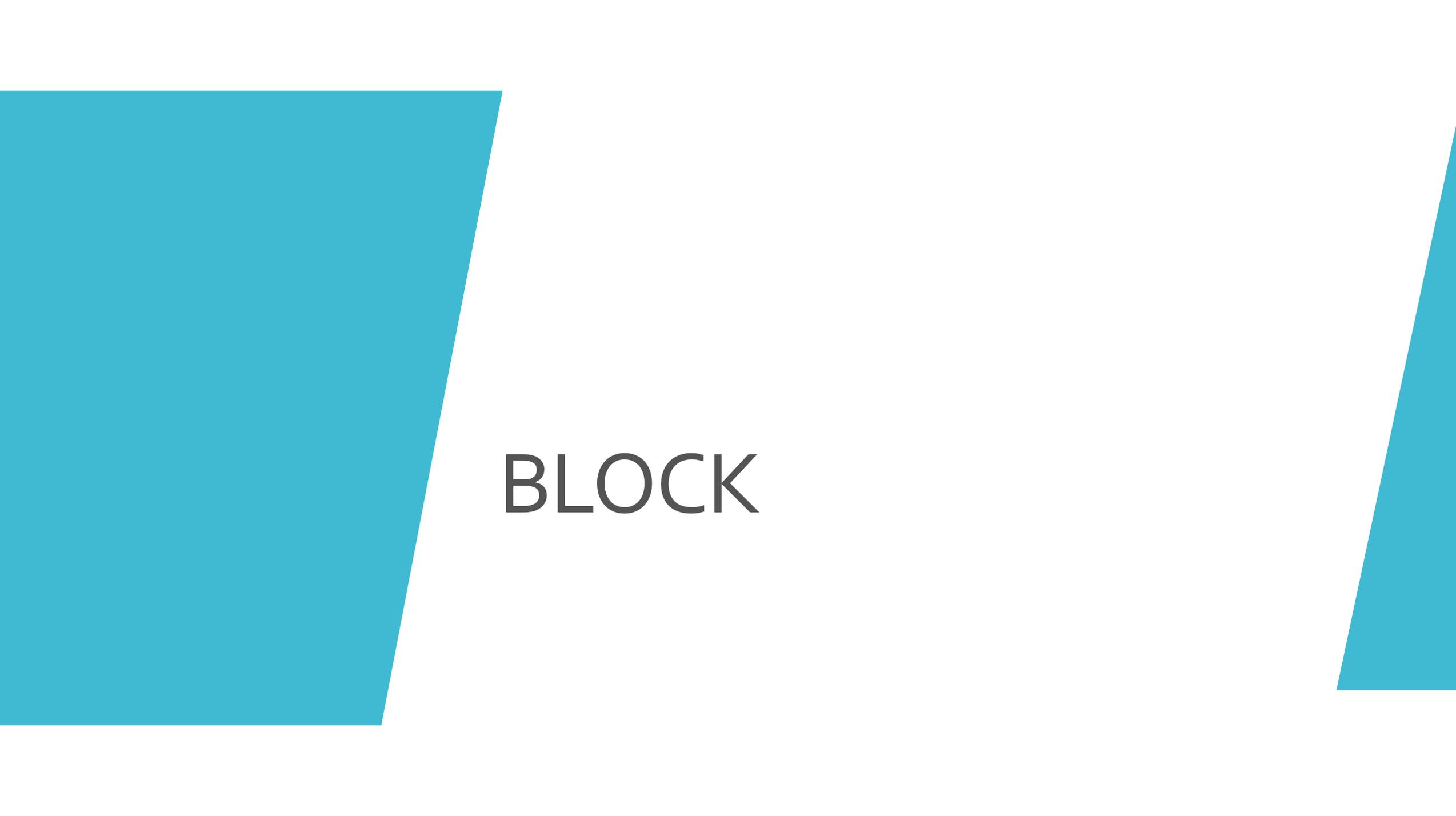
- 1) Essence / Histoire / Configuration vs Fonction (ce qu'il est vs ce qu'il fait)
- 2) Fonction à grain fin (complexe) vs fonction à grain grossier (relativement simple)
- 3) Fonctions non-computables

Apercu

- 1) La fonction n'est pas suffisante pour la conscience (Block, Searle)
- 2) Les fonctions pertinents ne sont pas réalisables par ordinateur numérique d'aujourd'hui (PGS, Seth)
- 3) Les fonctions pertinents ne sont pas computables en aucun sens (Lucas, Penrose, Landgrebe + Smith)

The image features a white background with two large teal geometric shapes. On the left, a teal triangle points towards the center. On the right, a teal trapezoid is positioned. The text is centered between these shapes.

La fonction n'est pas
suffisante pour la
conscience

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

La fonction n'est pas suffisante pour la conscience:
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).

La fonction n'est pas suffisante pour la conscience: 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)

La fonction n'est pas suffisante pour la conscience: 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

La fonction n'est pas suffisante pour la conscience: 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:

La fonction n'est pas suffisante pour la conscience: 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)

La fonction n'est pas suffisante pour la conscience: 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

La fonction n'est pas suffisante pour la conscience: Searle

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

La fonction n'est pas suffisante pour la conscience: 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...

La fonction n'est pas suffisante pour la conscience: 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...

La fonction n'est pas suffisante 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

La fonction n'est pas suffisante pour la conscience: 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: 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 S_v qu'il ne peut pas prouver.

2) Pour chacun de ces S_v , les humains peuvent connaître S_v

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.