Esprit

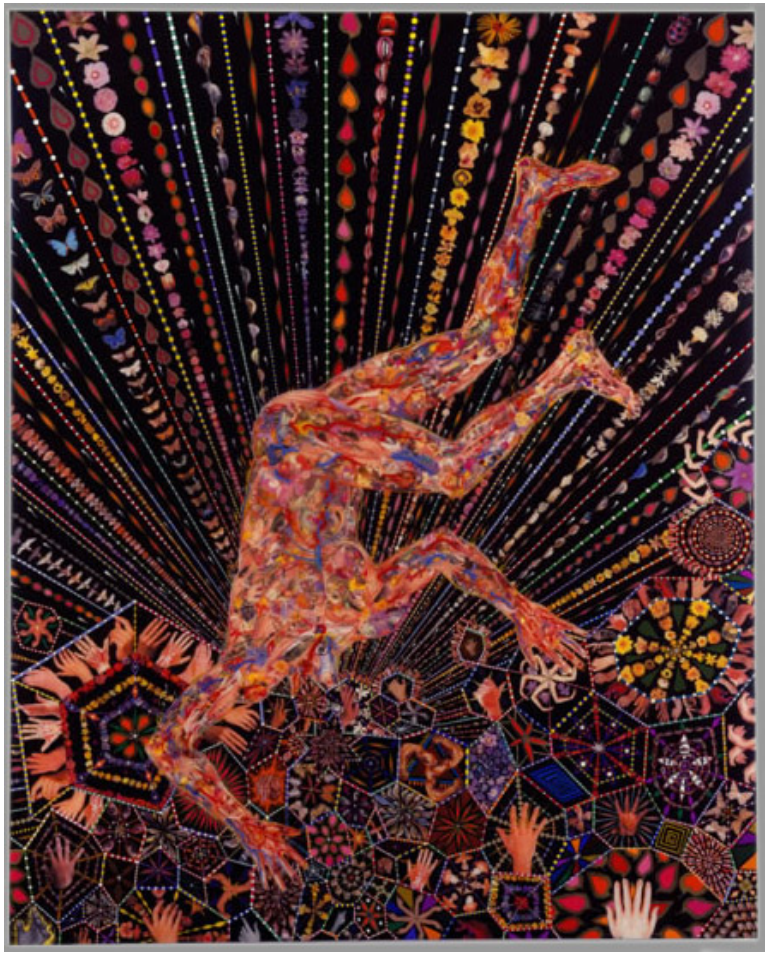
Christian Wüthrich

http://www.wuthrich.net/

BA3b Introduction à la métaphysique

Remerciements: Augustin Baas, Lorenzo Cocco

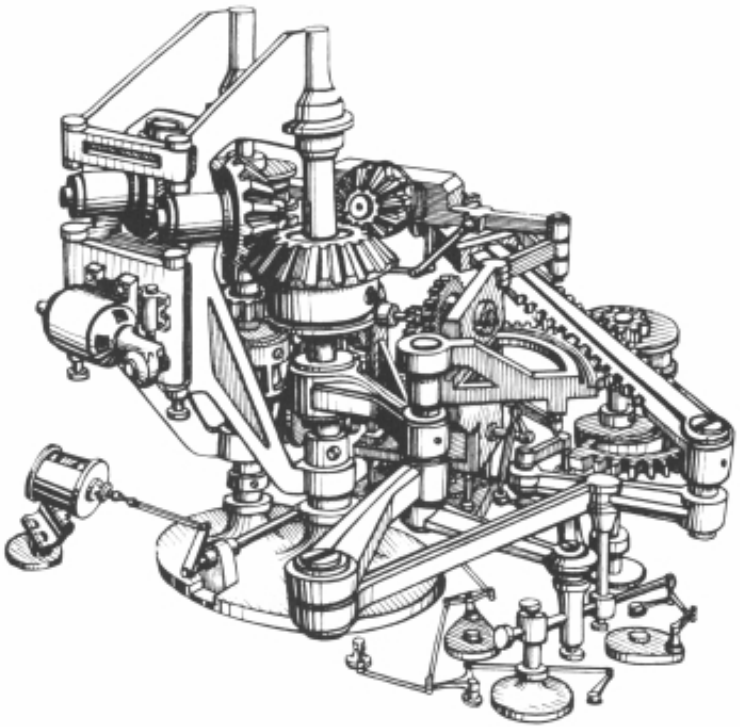
Le monde «biologique» ou «organique»



L'image «biologique» du monde:

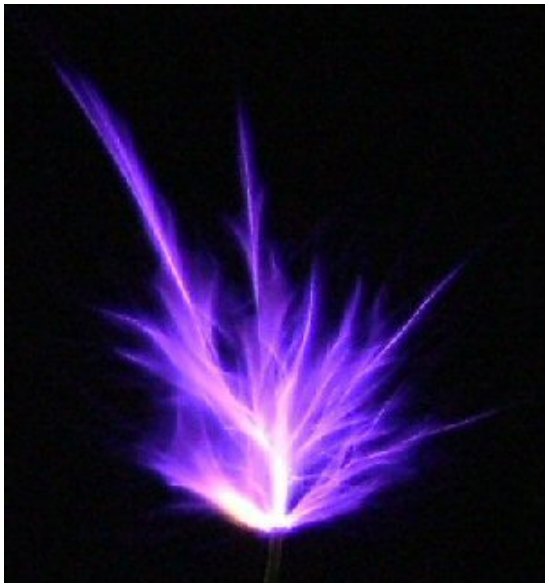
- Aristote
- téléologie
- vitalisme

L'image mécanique du monde



- Le *Novum Organum* de Bacon, la philosophie mécaniste de Descartes, Boyle, Leibniz, Newton
- Causes, lois de la Nature
- Non-téléogique

Le monde mécanique



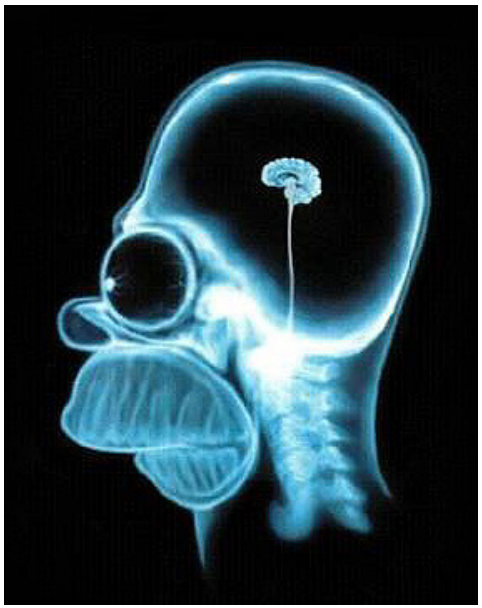
⇑
élan vital...

- Quelle est la place de la vie?
  - vitalisme: soutient qu'il existe dans tous les êtres vivants un facteur intrinsèque—insaisissable, inestimable et non mesurable—qui active la vie.
  - Vers les années 1940 environ, le vitalisme est mort.
  - Virus: comme nous, ils ont des caractéristiques d'hérédité, d'évolution, de reproduction, mais contrairement à nous, ils n'ont ni de métabolisme, ni de structure cellulaire, ni de homéostasie. Vivant ou mort?
- Où est la place de l'Esprit?

Quatre aspects du «problème corps-esprit»
(1) Conscience

John Searle (1984). *Minds, Brains and Science: The 1984 Reith Lectures*. Harvard University Press.

Dans ses conférences Reith de 1984, John Searle a dressé une liste de quatre problèmes auxquels fait face toute tentative de résoudre le «problème cerveau-esprit» (ou «problème corps-esprit»):



Problème (Conscience)

Comment la conscience est-elle possible? Comment un essaim de particules inconscientes peut-il devenir conscient?

Quatre aspects du «problème corps-esprit»
(2) Subjectivité

Problème (Subjectivité)

Comment rendre compte de la subjectivité des états mentaux dans une conception objective du monde? Notre vie mentale est privée: votre état de conscience actuel vous est accessible d'une manière qui ne m'est pas accessible (e.g. par imagerie par résonance magnétique).

Quatre aspects du «problème corps-esprit»
(3) Intentionnalité

Problème (Intentionnalité)

Comment des atomes dans le vide peuvent-ils avoir une intentionnalité? Comment peuvent-ils porter sur quelque chose? Comment peuvent-ils représenter les choses dans le monde comme nos pensées les représentent?

Définition

L'intentionnalité d'un état mental est son orientation vers une chose ou un objet, il est «à propos de» cette chose ou de ce objet.

Quatre aspects du «problème corps-esprit»
(4) Incorrigibilité

(Celui-ci est différent du quatrième problème de Searle.)

Problème (Incorrigibilité)

Il existe des états mentaux (comme la douleur) pour lesquels l'agent ne peut pas se tromper: ressentir la sensation de douleur est avoir de la douleur. Mais si les états mentaux sont simplement un produit de la matière en mouvement, pourquoi notre connaissance de ces états mentaux devrait-elle être différente de notre connaissance d'autres choses?

Le problème corps-esprit

Le problème corps-esprit est simplement la question de savoir comment le corps et l'esprit sont reliés entre eux. Nous croyons que les deux sont liés d'une manière ou d'une autre: chaque fois que nous buvons trop ou que nous nous blessons, des événements physiques affectent notre vie mentale; typiquement, lorsque nous pensons de «lever le bras», notre bras se lève et, dans ce cas, des événements mentaux affectent des événements physiques. Mais quel est leur lien exact? S'agit-il d'une seule et même chose, et si oui, dans quel sens? Ou les événements mentaux sont-ils radicalement différents des événements physiques, et si oui, les deux interagissent-ils vraiment?

Le problème corps-esprit: positions de base

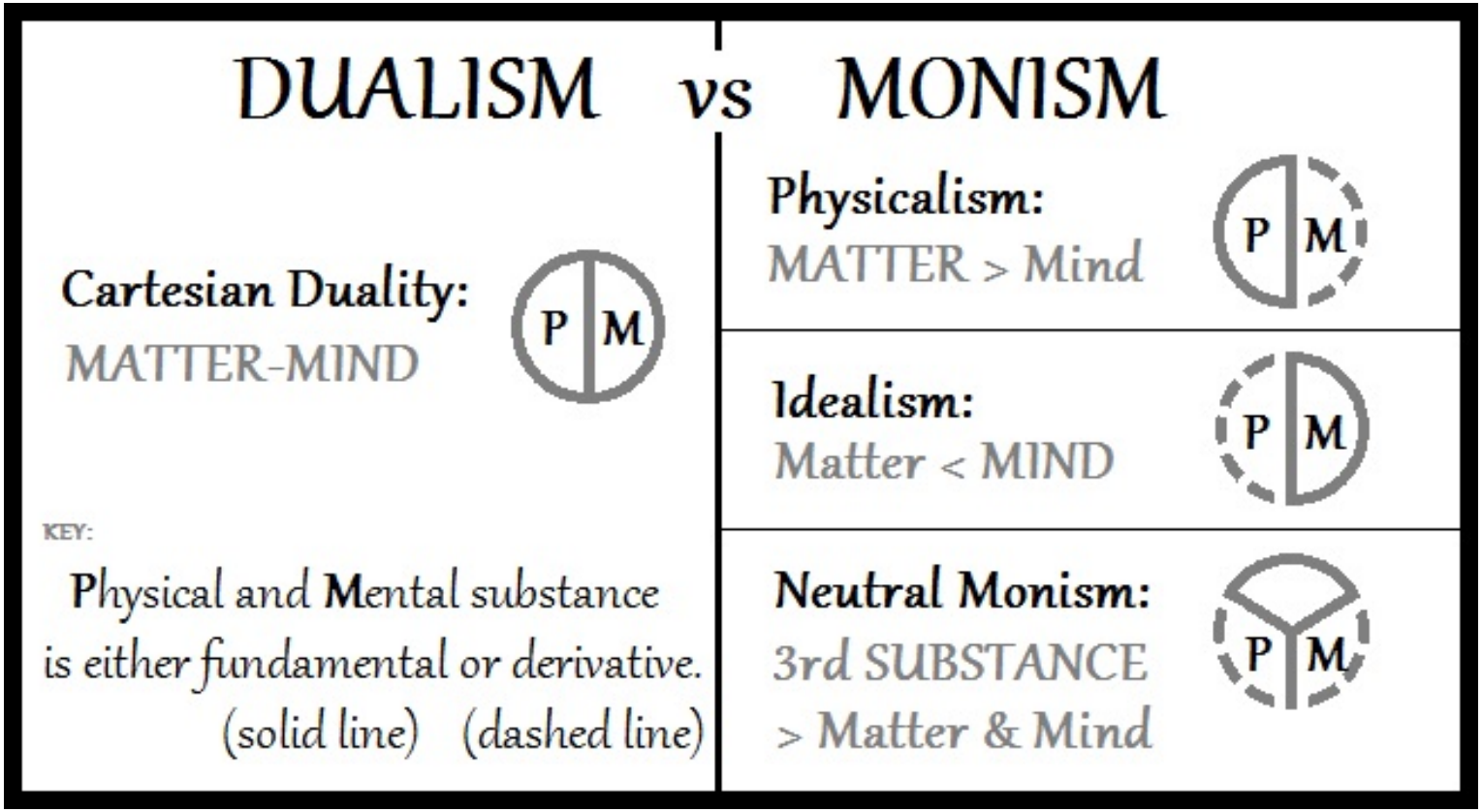
Position (Monisme)

Le monisme est l'idée qu'il n'existe qu'un seul type de substance élémentaire, généralement considérée comme la matière (matérialisme ou physicalisme).

Position (Dualisme)

Selon le dualisme, il y a deux types de substance élémentaire: l'esprit et le corps, distincts et séparables.

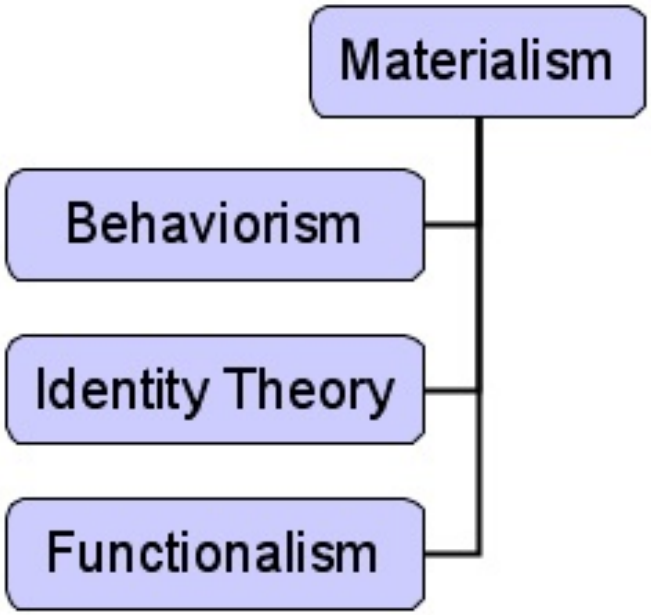
Monisme et dualisme

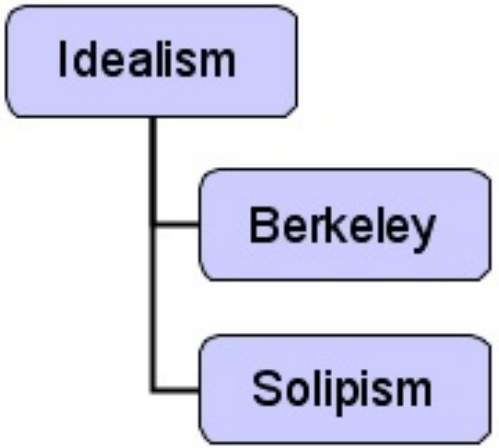


Source: Dustin Dewynne

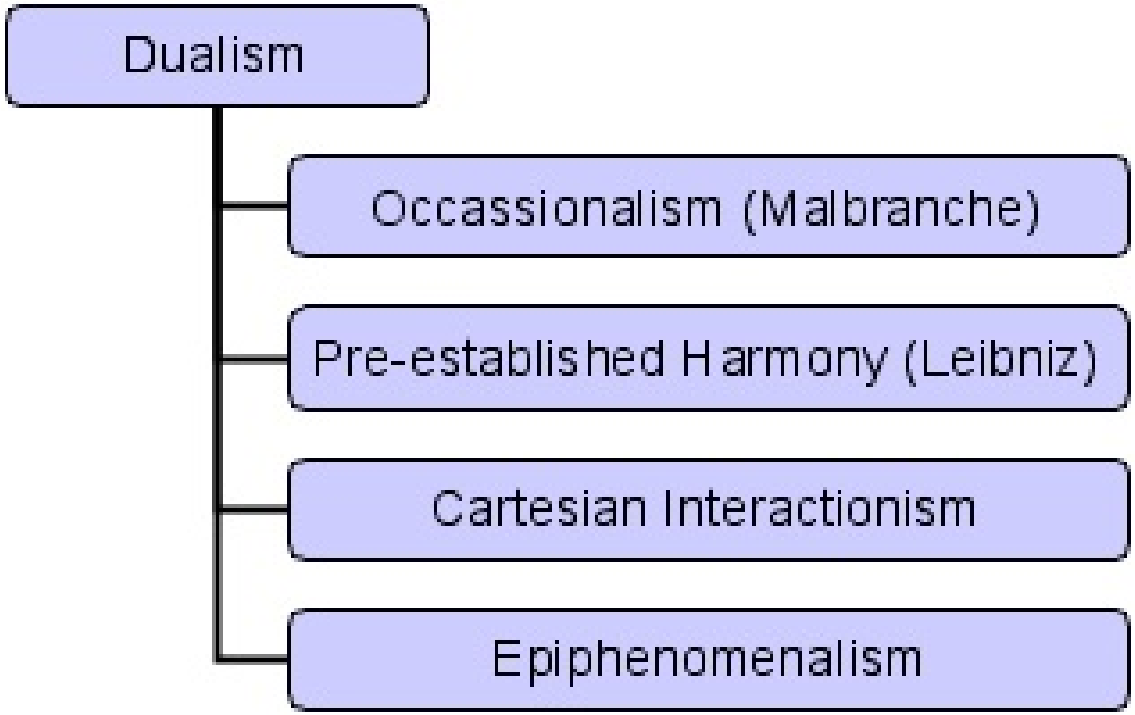
Outre cette division de base, il existe de nombreuses autres distinctions qui délimitent différentes théories.

Le monisme


Materialism
Behaviorism
Identity Theory
Functionalism

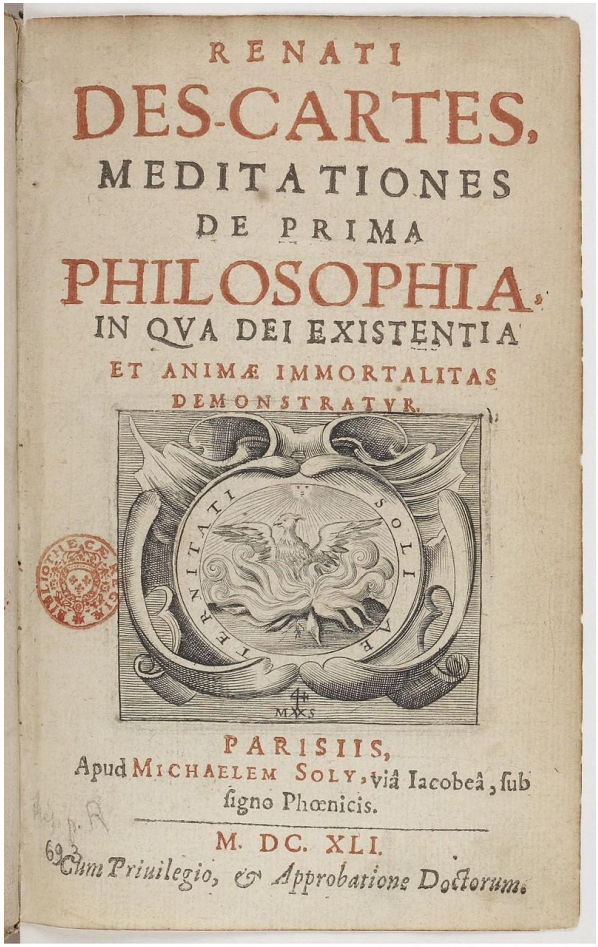

Idealism
Berkeley
Solipism

Le dualisme


Dualism
Occassionalism (Malbranche)
Pre-established Harmony (Leibniz)
Cartesian Interactionism
Epiphenomenalism

René Descartes (1596-1650) et le dualisme

René Descartes (1641). *Meditationes de Prima Philosophia, in qua Dei existentia et animæ immortalitas demonstratur*. Paris.


RENATI
DES-CARTES,
MEDITATIONES
DE PRIMA
PHILOSOPHIA,
IN QVA DEI EXISTENTIA
ET ANIMÆ IMMORTALITAS
DEMONSTRATVR.
PARISIIS,
Apud MICHAELEM SOLY, viâ Iacobeâ, sub
signo Phœnicis.
M. DC. XLI.
Cum Priuilegio, & Approbatione Doctorum.

Source gallica.bnf.fr / Bibliothèque nationale de France

La méthode: trouver une proposition p qui est absolument certaine....

1. Si j'ai une idée claire et distincte de moi-même sans mon corps, alors je ne suis pas mon corps.
2. C'est le cas.
3. Par conséquent, je ne suis pas mon corps.

Justification

- Justification de (1): Ce que je peux percevoir clairement et distinctement peut l'être par Dieu. (Dieu ne me tromperait pas quand il s'agit d'idées claires et distinctes.)
- Justification de (2): Je peux me considérer (clairement et distinctement) que comme une chose qui pense seulement.

Descartes: deux (ou trois) types de substances

Caractérisation (Ontologie cartésienne)

L'ontologie de Descartes est tripartite: Dieu, la res cogitans, et la res extensa. La res cogitans est la «substance pensante» immatérielle, l'essence non temporelle, non spatiale et libre d'une personne, celle qui doute, espère, croit et pense. La res extensa est la «substance étendue», c'est-à-dire les corps matériels déterminés dans l'espace et le temps, tels que les corps des humains et des animaux.

L'existence des choses matérielles

1. Les idées sensorielles sont causées indépendamment de ma pensée.
2. Je ne suis pas la cause de ces idées sensorielles.
3. Si c'est le cas, alors la cause est soit (a) des objets physiques, (b) Dieu, ou (c) une chose non physique créée.
4. Il n'est pas le cas que (b) ou (c).
5. Par conséquent, (a), les objets physiques existent.

Justification de (4):

- Si Dieu, alors Il est un trompeur.
- Si une chose non physique a été créée, alors Dieu est l'auteur d'un trompeur.

L'existence des choses pensantes

Paul Churchland (1984, third edition 2013). *Matter and Consciousness*. MIT Press.

(voir Churchland 1984, 32))

1. Mes états mentaux me sont introspectivement reconnaissables en tant qu'états de mon moi conscient.
2. Mon état cérébral ne l'est pas.
3. X et Y sont identiques ssi ils ont tous et seulement les mêmes propriétés.

---

4. Par conséquent, mes états mentaux ne sont pas identiques à mes états cérébraux.

Mais...

Superman = Clark Kent

L'argument de la connaissance pour le dualisme
(Originalement dans Jackson 1982; voir aussi Churchland, p. 33f)

Frank Jackson (1982). Epiphenomenal qualia. *Philosophical Quarterly* 32: 127-136.

L'argument de la connaissance: la mise en place

Stalnaker (2008, 24, trad. Augustin Baas):

> [*Mary*] *est une scientifique brillante qui a été confinée depuis sa naissance dans une chambre noire et blanche. Elle sait, pour avoir lu les livres en noir et blanc qui remplissent les étagères de sa chambre, tout ce qu'il faut savoir sur la physique de la couleur et la neurophysiologie de la vision des couleurs, mais elle n'a jamais eu l'occasion de voir des couleurs. Même si elle connaît toutes les sciences physiques et biologiques pertinentes, il y a encore quelque chose qu'elle ne sait pas, quelque chose qu'elle n'apprendra que lorsqu'elle sortira de sa chambre et qu'elle verra des choses colorées: elle ne sait pas ce que c'est de voir des couleurs.*

Robert Stalnaker (2008). *Our Knowledge of the Internal World*. Oxford.

L'argument de la connaissance: l'argument

1. «L'histoire semble laisser entendre qu'une personne pourrait connaître tous les faits physiques pertinents tout en demeurant ignorante de certains autres faits, c'est-à-dire des faits concernant le caractère qualitatif de notre expérience visuelle.
2. «Il doit donc y avoir des faits à connaître qui ne sont pas des faits physiques.
3. «Mais s'il y a des faits qui vont au-delà des faits physiques, alors le matérialisme—la thèse que tous les faits sont des faits physiques—est fausse.» (24f)

Arguments contre le dualisme
(Churchland 1984)

- Simplicité
- Impuissance explicative: «Comparez maintenant ce que le neuroscientifique peut nous dire sur le cerveau, et ce qu'il peut faire avec cette connaissance, avec ce que le dualiste peut nous dire sur la substance spirituelle, et ce qu'il peut faire avec ces hypothèses?»
- Interaction?
- Dépendance neuronale de tous les phénomènes connus
- Histoire évolutionnaire

Le béhaviorisme philosophique

Qu'est-ce vraiment que «la solubilité»? L'une des réponses est qu'il n'y a pas vraiment de solubilité; les affirmations sur la solubilité sont plutôt des affirmations sur les dispositions à s'engager dans divers types de comportements observables.

Entrées	**Sorties**
Mettre de l'aspirine dans l'eau	dissoudre
Mettre de l'aspirine dans un soda	dissoudre
⋮	⋮

Quel est vraiment l'état mental de la douleur? Une des réponses est qu'il n'y a pas vraiment quelque chose qui soit un état mental; ce sont plutôt des affirmations sur la douleur, etc. qui sont des affirmations sur les dispositions à adopter divers types de comportements observables.

Entrées	**Sorties**
Tester la pression du sang	plus haut que la norme
«Voulez-vous jouer aux échecs?»	«Non!»
«Est-ce que ÇA fait mal?»	«Arghh!»
«Vraiment, il n'y a rien tel que la douleurs»	Mettre un coup de poing dans le nez de l'orateur

Thèse (Le béhaviorisme philosophique)

Le béhaviorisme philosophique est la thèse selon laquelle tous les états mentaux doivent être analysés comme des dispositions à s'engager dans certains types de comportements observables. Les termes psychologiques peuvent être traduits en termes comportementaux non psychologiques.

Définition (Zombie philosophique)

Un zombie philosophique est un être hypothétique qui est physiquement indiscernable d'un personne normale, mais qui n'a pas des expériences conscientes, des qualia ou du ressenti.

$\Rightarrow$ Selon le béhaviorisme, il n'y a pas de différence entre des personnes normales et des zombies.

Problèmes

- L'exemple de l'acteur parfait?
- Observable?
- Liste finie d'évènements?
- Qualia?
- La principale: a besoin des états mentaux pour expliquer le comportement

$\Rightarrow$ signature d'un programme de recherche qui dégénère

Théorie de l'identité

J J C Smart (2007). The mind/brain identity theory. In E. Zalta (ed.), *Stanford Encyclopedia of Philosophy*. http://plato.stanford.edu/entries/mind-identity/

Position (Théorie de l'identité)

Mes états mentaux sont juste mes états cérébraux. Cela implique qu'il y a une correspondance 1:1 entre mes états cérébraux et mes états mentaux.

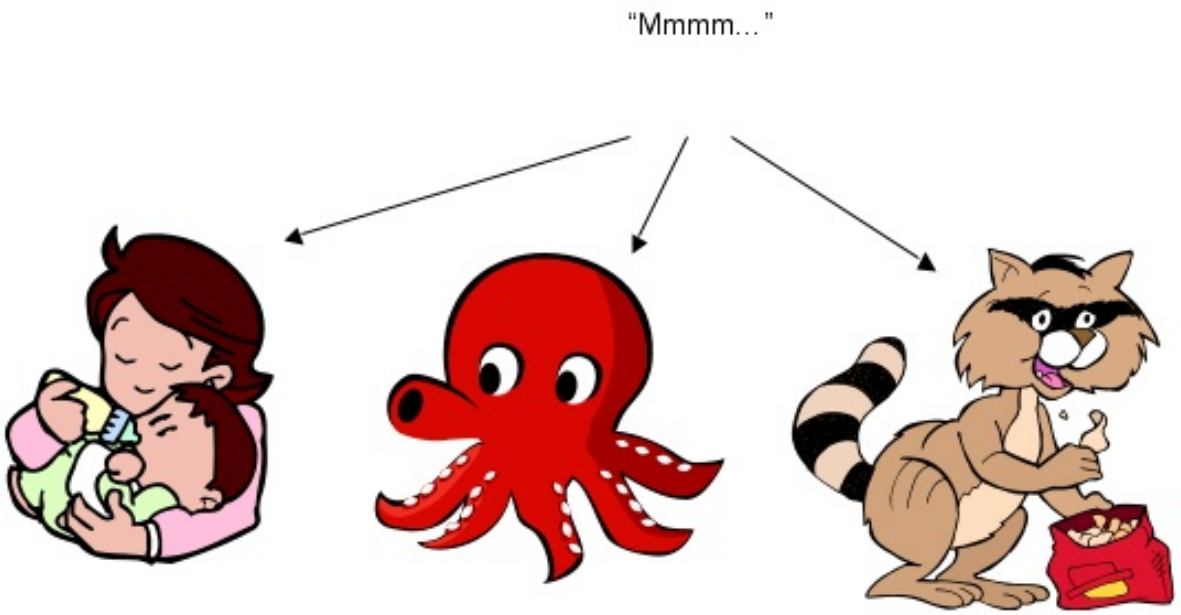
- De la même façon que l'eau est H_2O, ou la température est l'énergie cinétique moyenne, les états mentaux et les états cérébraux sont identiques, malgré leurs différences apparentes.

$\Rightarrow$ Il n'y a rien de mystérieux quant à l'interaction.

$\Rightarrow$ L'existence de «qualia» est niée.

Types et tokens

- Types versus tokens: distinction entre un type général (un «type») et ses instances concrètes particulières (un «token»)
- Exemple: ambiguïté de «Conduisent-ils la même voiture?»
- Théorie de l'identité type-type: Chaque type mental est l'équivalent d'un type physique, par exemple une douleur aiguë = un tir de fibres c à une certaine fréquence.
- Objection: réalisation multiple (bébés, pieuvres, ratons laveurs, ...)



Functionnalisme

«Software»		**«Hardware»**
	$\Leftrightarrow$	
Minecraft		iPhone, Samsung Galaxy

«Etat mental»		**«Etat physique»**
	$\Leftrightarrow$	
Douleur		Homme, Femme, Pieuvre

Functionnalisme

Janet Levin (2018). Functionalism. In E. Zalta (ed.), *Stanford Encyclopedia of Philosophy*.
http://plato.stanford.edu/entries/functionalism/

Position (Fonctionnalisme)

«Le fonctionnalisme dans la philosophie de l'esprit est la doctrine selon laquelle ce qui fait d'une chose un état mental d'un type particulier ne dépend pas de sa constitution interne, mais plutôt de la façon dont elle fonctionne, ou du rôle qu'elle joue, dans le système dont elle fait partie.» (Levin 2018)
Les états et processus mentaux sont définis par leur rôle causal (leurs fonctions). Et on remarque alors que les neurones instancient ces rôles.

Machine à café (le café coûte €1.00)

Etat Interne Courant	Pièce Entrée	Café Sortie	Pièce Sortie	Etat Interne Nouveau
E1	€0.50	Non	Non	E2
E1	€1.00	Oui	Non	E1
E2	€0.50	Oui	Non	E1
E2	€1.00	Oui	Oui	E1

Fonctionnalisme

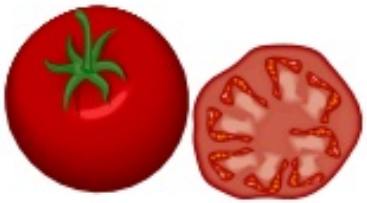
Lewis (1966, trad. Augustin Baas):

> *Mon argument est le suivant: La caractéristique définitive de toute (sorte d') expérience en tant que telle est son rôle causal, son syndrome des causes et des effets les plus typiques. Mais nous, matérialistes, croyons que ces rôles causaux qui appartiennent par nécessité analytique aux expériences, appartiennent en fait à certains états physiques. Puisque ces états physiques possèdent le caractère définitif des expériences, ils doivent être des expériences.*

David Lewis (1966). An argument for the identity theory. *Journal of Philosophy* 63: 17-25.

Contre le fonctionnalisme: le spectre inversé

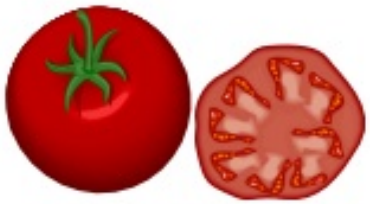
INPUTS → RED → OUTPUTS



"Look at that red tomato!"

"That tomato is the same color as a stop sign"

INPUTS → BLUE → OUTPUTS



"Look at that red tomato!"

"That tomato is the same color as a stop sign"

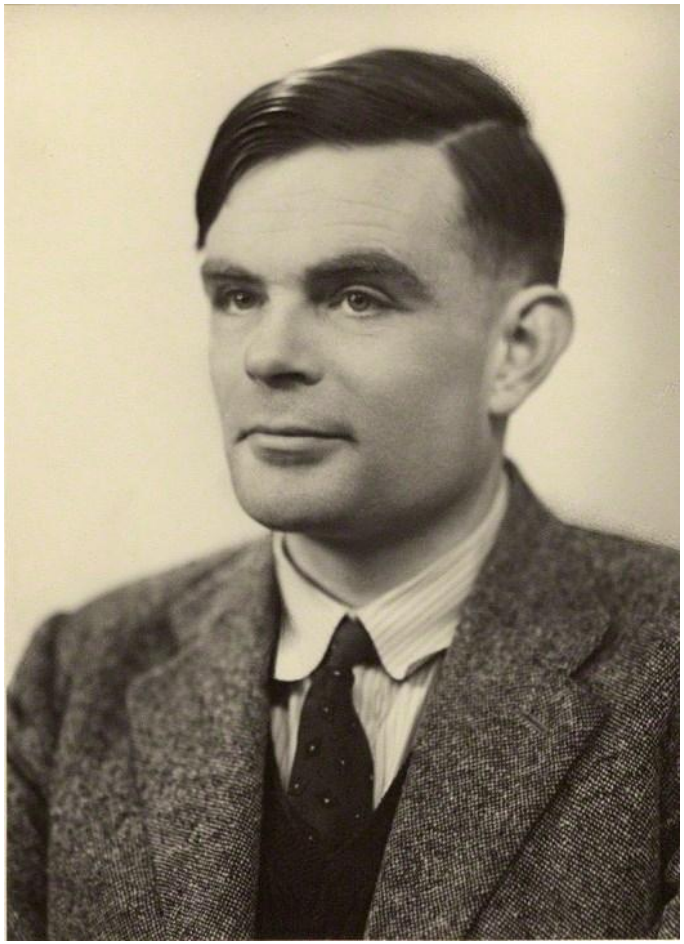
Réponses des fonctionnalistes

- Ce n'est pas possible: L'expérience des lunettes inversées, qui peut être interprétée comme soutenant l'idée que, à mesure que deux personnes deviennent plus semblables sur le plan fonctionnel, leurs expériences deviennent également plus semblables.

$\Rightarrow$ Fonctionnalisme incomplet (Searle)

Les ordinateurs peuvent-ils penser?

Alan Turing (1912-1954)



- «père» de l'informatique moderne
- machine de Turing, test de Turing
- contributions à l'intelligence artificielle
- pendant la seconde guerre mondiale: a travaillé à Bletchley Park (le centre de décodage du Royaume-Uni), où il a dirigé un département qui a aidé à casser la machine Enigma

Machines de Turing

de Mol (2018, trad. Augustin Baas):

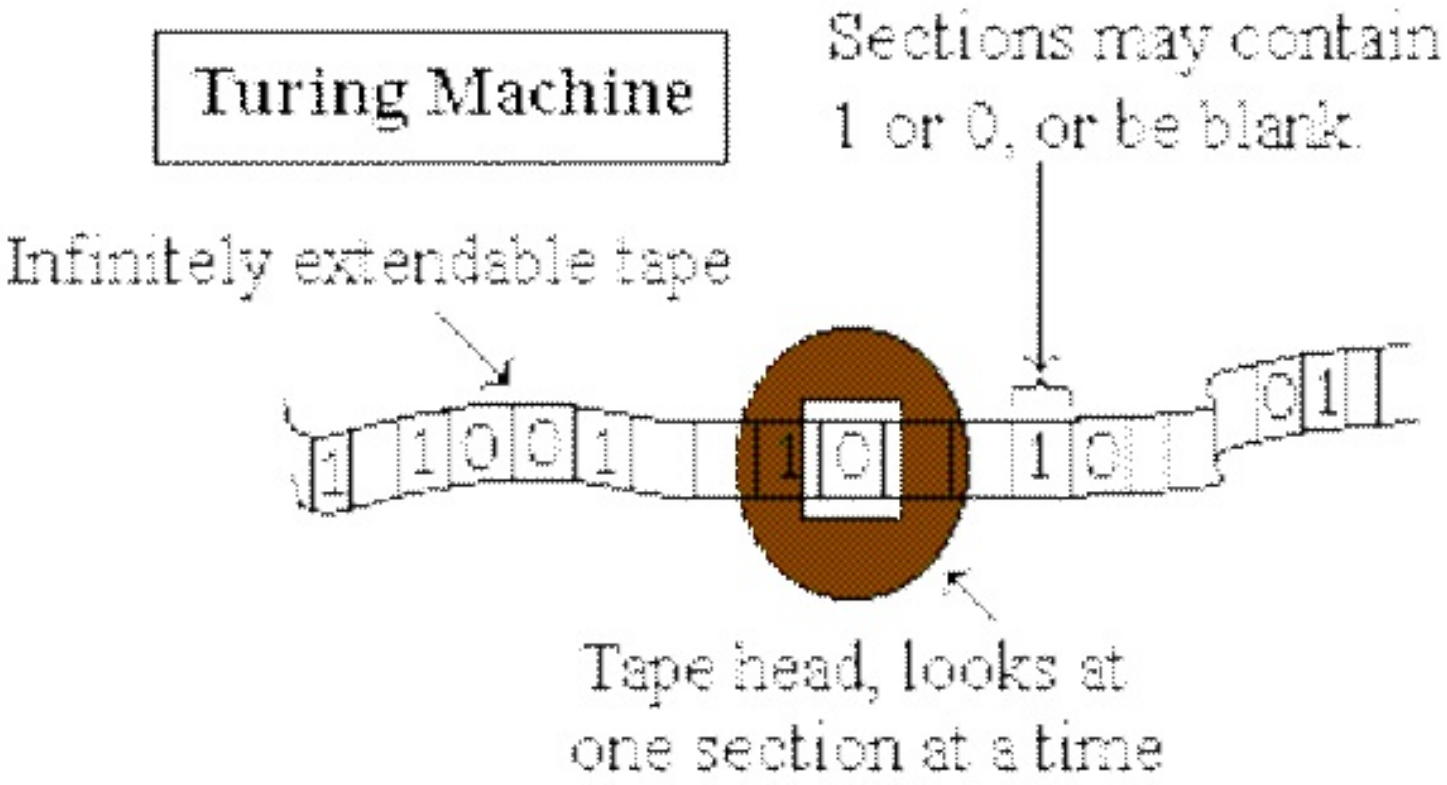
Une machine de Turing est une représentation abstraite d'un dispositif de calcul. Il se compose d'une tête de lecture/écriture qui scanne une bande unidimensionnelle (bidirectionnelle) (éventuellement infinie) divisée en carrés, chacun d'eux portant l'inscription 0 ou 1. Le calcul commence avec la machine dans un «état» donné, balayant un carré. Il efface ce qu'il y trouve, imprime un 0 ou un 1, se déplace sur une case adjacente, et entre dans un nouvel état. Ce comportement est complètement déterminé par trois paramètres: (1) l'état dans lequel se trouve la machine, (2) le numéro sur le carré qu'elle scanne, et (3) un tableau d'instructions. Le tableau d'instructions spécifie pour chaque état et chaque entrée binaire, ce que la machine doit écrire, dans quelle direction elle doit se déplacer et dans quel état elle doit passer. (Par exemple «Si dans l'état 1, si vous scannez un 0: imprimez 1, allez à gauche, et passez dans l'état 3».)

Liesbeth de Mol (2018). Turing machines. In E. Zalta (ed.), *Stanford Encyclopedia of Philosophy*.
https://plato.stanford.edu/entries/turing-machine/.

Machines de Turing (suite)

de Mol (2018, trad. Augustin Baas):

> *Le tableau ne peut passer en revue qu'un nombre limité d'états, chacun d'eux étant implicitement défini par le rôle qu'il joue dans le tableau des instructions. Ces états sont souvent appelés «états fonctionnels» de la machine. Une machine de Turing, par conséquent, est plus comme un programme informatique (software) qu'un ordinateur (hardware).... Les informaticiens et les logiciens ont montré que les machines de Turing—avec suffisamment de temps et de bande—peuvent calculer n'importe quelle fonction que n'importe quel ordinateur numérique conventionnel peut calculer.*


Turing Machine
Sections may contain
1 or 0, or be blank
Infinitely extendable tape
1 1 0 0 1 1 0 1 0 0 1
Tape head, looks at
one section at a time

Qu'est-ce qu'un ordinateur?

Current State	Current Symbol	New Symbol	Move	New State
A	space	space	L	A
A	0	1	L	B
A	1	0	L	B
B	0	1	L	B
B	1	0	L	B
B	space	halt	L	B

Machines de Turing

Thèse

La machine de Turing peut calculer n'importe quelle fonction qu'un ordinateur ordinaire peut calculer.

Définition

Une machine de Turing est une machine de Turing universelle si elle est capable de simuler toute autre machine de Turing.

Définition

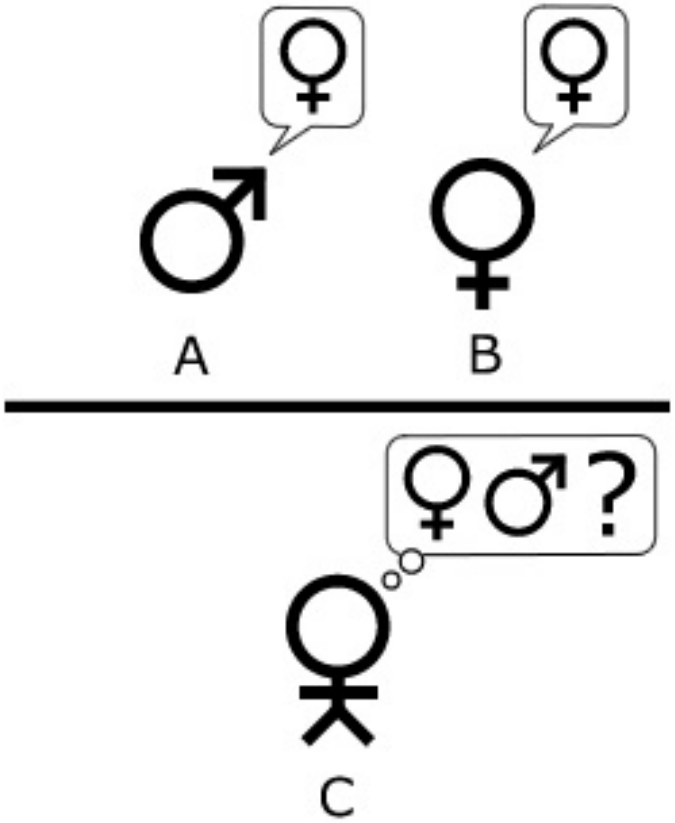
Tout ce qui est calculable avec une machine de Turing universelle est Turing calculable.

Thèse (Church-Turing)

Tout ce qui est calculable est Turing calculable.

Jeu d'imitation

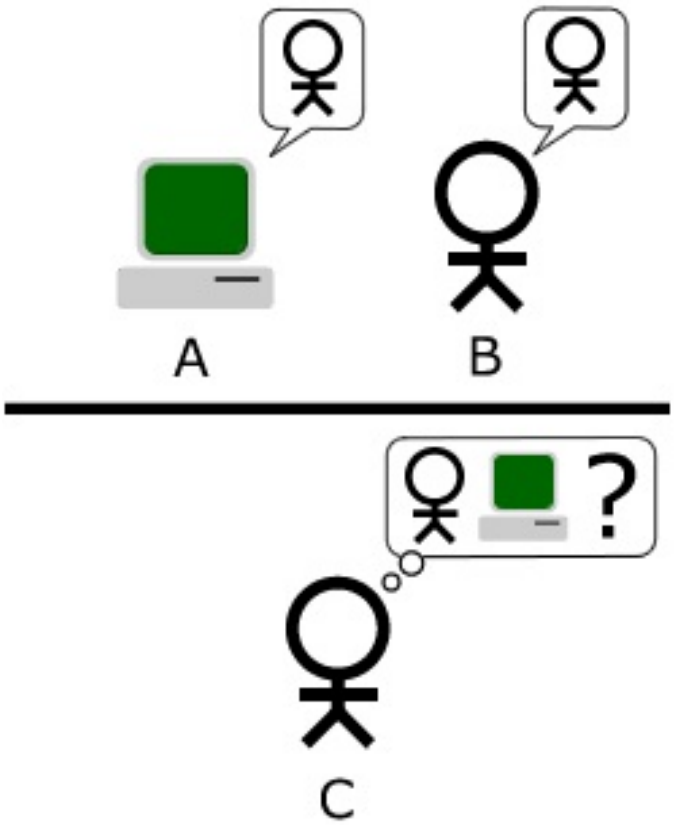
Alan Turing (1950). Computing machinery and intelligence. *Mind* 59: 433-460.



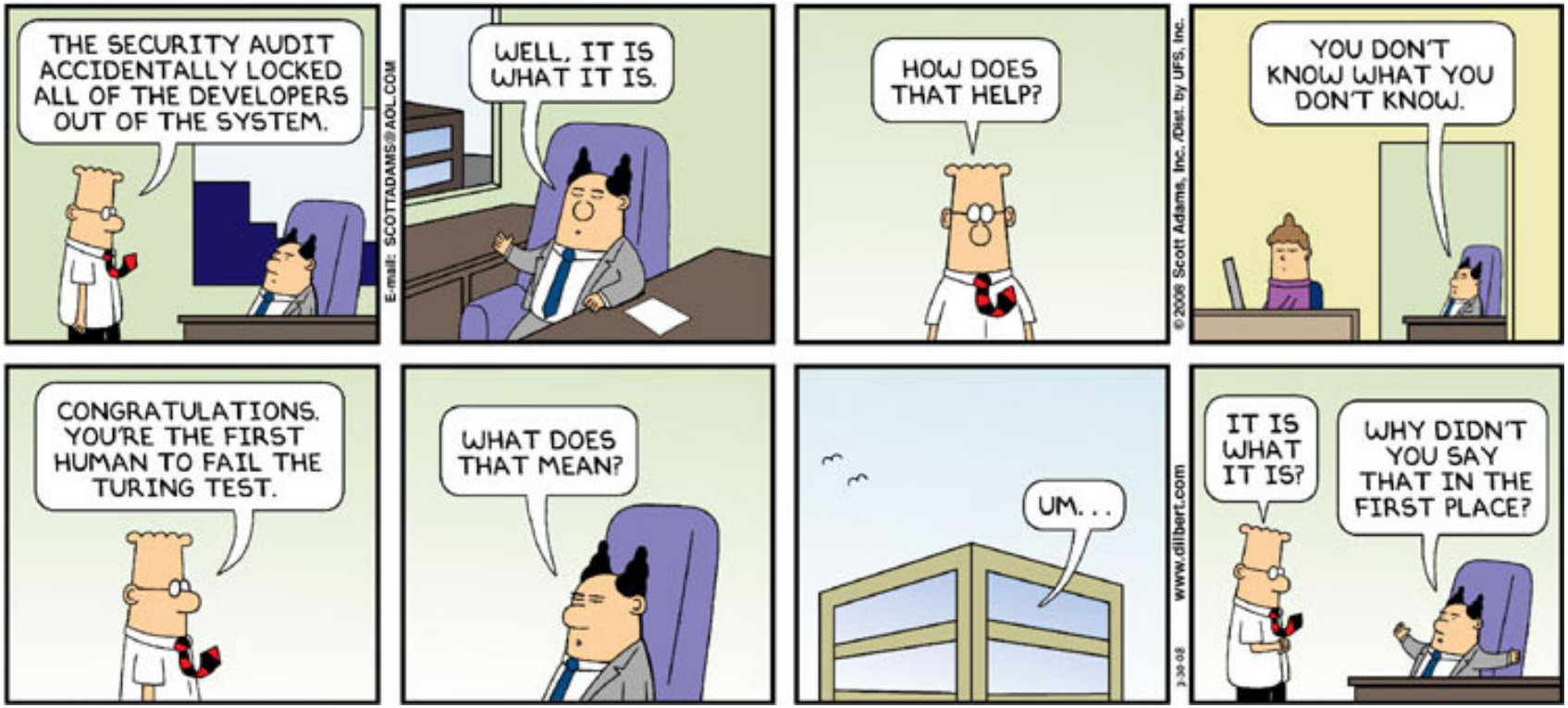
Le joueur C tente, au moyen d'une série de questions écrites, de déterminer lequel des deux autres joueurs est un homme et lequel des deux est la femme. Le joueur A, l'homme, essaie d'amener le joueur C à prendre la mauvaise décision, tandis que le joueur B essaie d'aider le joueur C.

Teste de Turing

Alan Turing (1950). Computing machinery and intelligence. *Mind* 59: 433-460.



Décrit par Alan Turing dans le même article, le test de Turing se déroule comme suit: un juge humain doit découvrir, au moyen d'une conversation dactylographiée avec un humain et une machine, dont chacun essaie d'apparaître humain, lequel est l'humain. Si le juge n'est pas en mesure de distinguer de façon fiable la machine de l'humain, on dit que la machine a réussi le test.


THE SECURITY AUDIT ACCIDENTALLY LOCKED ALL OF THE DEVELOPERS OUT OF THE SYSTEM.
WELL, IT IS WHAT IT IS.
E-mail: SCOTTADAMS@AOL.COM
HOW DOES THAT HELP?
YOU DON'T KNOW WHAT YOU DON'T KNOW.
© 2008 Scott Adams, Inc./Dist. by UFS, Inc.
CONGRATULATIONS. YOU'RE THE FIRST HUMAN TO FAIL THE TURING TEST.
WHAT DOES THAT MEAN?
UM. . .
IT IS WHAT IT IS?
WHY DIDN'T YOU SAY THAT IN THE FIRST PLACE?
www.dilbert.com

Un ordinateur ne peut pas (??)

- être gentil
- avoir de la ressource
- être beau
- être amical
- avoir de l'initiative
- avoir le sens de l'humour
- extraire le vrai du faux
- faire des erreurs
- tomber amoureux
- apprécier des fraises à la crème
- faire que quelqu'un tombe amoureux
- apprendre de son expérience
- utilise les mots correctement
- être le sujet de ses propres pensées
- avoir autant de diversité de comportement qu'un humain.
- faire quelque chose de vraiment nouveau

Intelligence artificielle

- Est-ce que réussir le test de Turing est nécessaire pour l'intelligence?
- Est-ce que réussir le test de Turing est suffisant pour l'intelligence?

Deux définitions

Simon Blackburn (1994). *Oxford Dictionary of Philosophy*. Oxford: Oxford University Press.

Définition (Syntaxe)

«La syntaxe d'une langue est sa grammaire, ou la façon dont ses expressions peuvent être assemblées pour former des phrases. Une étude syntaxique est une étude qui ne concerne pas le sens des phrases, mais les aspects purement formels de la combinaison de mots dans une langue.» (Blackburn 1994, trad. par Augustin Baas)

Définition (Sémantique)

La sémantique est «l'étude du sens des mots et de la relation des signes avec les objets auxquels ils s'appliquent. Dans les études formelles, une sémantique est fournie pour un langage formel lorsqu'une interprétation ou un modèle est spécifié.» (ibid.)

Deux thèses

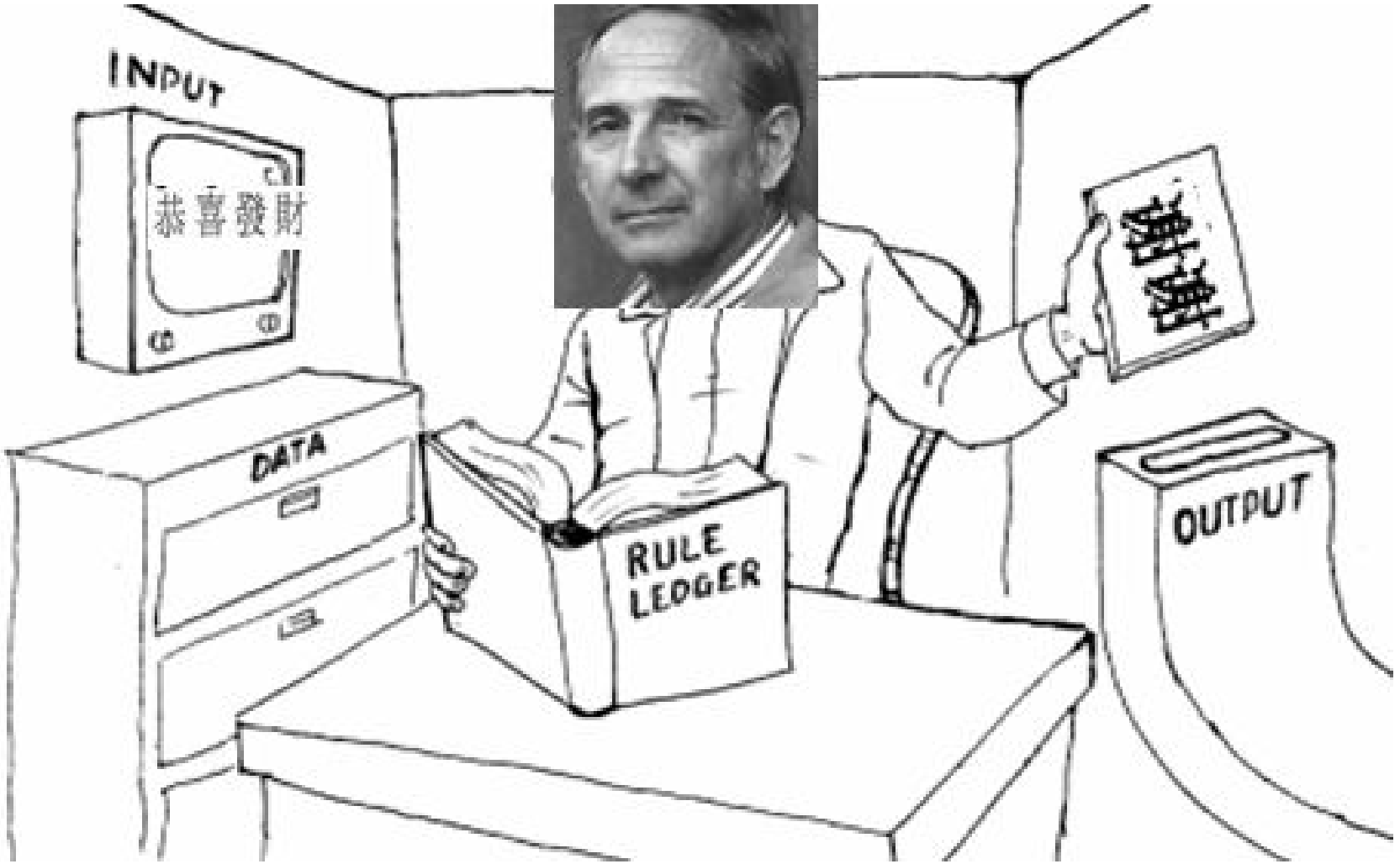
Thèse (IA forte)

«L'IA forte est la thèse philosophique selon laquelle les ordinateurs bien programmés ont de l'esprit tout comme nous en avons.» (ibid.)

Thèse (IA faible)

«L'IA faible est la croyance méthodologique que la meilleure façon d'explorer l'esprit est de procéder comme si [la thèse de l'IA forte] était vraie, sans commenter les héritages du dualisme qui mènent à mal la thèse forte.» (ibid.)

La chambre chinoise de Searle



Searle demande à son lecteur d'imaginer que, dans de nombreuses années, les gens auront construit un ordinateur qui se comporte comme s'il comprenait le chinois. Il utilise des caractères chinois en entrée et, à l'aide d'un programme informatique, produit d'autres caractères chinois qu'il présente en sortie. Supposons, dit Searle, que cet ordinateur accomplisse sa tâche de façon si convaincante qu'il réussisse le test de Turing: il convainc un locuteur chinois humain que le programme est lui-même un locuteur chinois humain. Toutes les questions que l'humain lui pose reçoivent des réponses appropriées, de sorte que le locuteur chinois est convaincu qu'il ou elle parle à un autre être humain parlant chinois. La plupart des partisans de l'intelligence artificielle tireraient la conclusion que l'ordinateur comprend le chinois, tout comme l'humain parlant chinois.

Searle demande alors au lecteur de supposer qu'il se trouve dans une pièce où il reçoit des caractères chinois, consulte un livre contenant une version anglaise dudit programme informatique et traite les caractères chinois selon ses instructions. Il ne comprend pas un mot de chinois; il manipule simplement ce qui, pour lui, sont des symboles dénués de sens, en utilisant le livre et tout autre équipement qui lui est disponible, comme du papier, des crayons, des gommes et des classeurs. Après avoir manipulé les symboles, il répond à une question chinoise donnée dans la même langue. Comme l'ordinateur a passé le test de Turing de cette façon, il est juste, dit Searle, de déduire qu'il l'a fait, lui aussi, simplement en exécutant le programme manuellement.: «Personne ne peut dire, rien qu'en regardant mes réponses, que je ne parle pas un mot de chinois», écrit-il. (extrait de Wikipedia)

La conclusion de Searle

John Searle (1980). Minds, brains and programs. *Behavioral and Brain Sciences* 3: 417-424. Citations suivant la traduction d'Eric Duyckaerts, paru dans *Quaderni* 1987/1: 65-96.

$\Rightarrow$ L'ordinateur ne comprend pas plus le chinois que lui.

$\Rightarrow$ Les ordinateurs n'ont pas de «compréhension» de ce qu'ils disent.

$\Rightarrow$ Les ordinateurs n'ont pas d'état mental conscient.

$\Rightarrow$ Réussir le test de Turing n'est pas une preuve d'avoir des états mentaux conscients, c'est-à-dire d'avoir un esprit.

La structure logique de l'argument de la chambre chinoise

P1: Les programmes sont formels. (syntactique)

P2: Les esprits ont des contenus. (sémantique)

P3: La syntaxe n'est pas suffisante pour la sémantique.

C: Les programmes n'ont pas d'esprit.

La structure logique de l'argument de la chambre chinoise

P1′: Les processus sont formels. (syntactique)

P2: Les esprits ont des contenus. (sémantique)

P3: La syntaxe n'est pas suffisante pour la sémantique.

C′: Les processus n'ont pas d'esprit.

(1) La réponse systèmique

Considérons quelques réponses à l'argument de Searle, discuté par lui-même:

Searle (1980, 419; 1987, 72 trad.)

«Bien qu'il soit vrai que la personne individuelle qui est enfermée dans la pièce ne comprenne pas l'histoire, elle n'est en fait qu'une partie d'un système global et celui-ci comprend vraiment l'histoire. La personne a sous les yeux un registre où les règles sont écrites, il dispose de beaucoup de papier de brouillon et de crayons pour faire ses calculs, il a des «banques de données» faites d'ensembles de symboles chinois. Aussi l'entendement n'est-il pas assigné au seul individu; au contraire, on l'assigne à tout le système dont il n'est qu'une partie.»

- Comparez avec la machine de Turing
- Searle: il est tout simplement ridicule de dire que, «alors qu'une personne ne comprend pas le chinois, d'une certaine façon, la conjonction de cette personne et de bouts de papiers pourrait le comprendre» (1980, 419; 1987, 73 trad.).

(2) La réponse robotique

- Ce qui empêche la personne dans la salle chinoise d'attacher des significations aux symboles chinois (et donc l'empêche de les comprendre), c'est la déconnexion sensorimotrice des symboles par rapport aux réalités qu'ils sont censés représenter—un lien causal est requis.
- Searle: raconter l'histoire à nouveau mais avec de l'espace dans la tête du robot
- Searle: la réponse admet tacitement qu'on a besoin d'un ensemble de relations causales avec le monde extérieur

(3) La réponse du simulateur de cerveau

Searle (1980, 420; 1987, 79 trad.)

«Supposez que nous concevions un programme qui ne représente pas notre information sur le monde... mais qui simule la séquence réelle des déclenchements neuronaux dans les synapses du cerveau d'un locuteur chinois d'origine, quand il comprend des histoires en chinois et y répond.»

- Pas de différence entre le programme du cerveau du locuteur natif chinois et le programme de la personne dans la salle.
- Searle: tuyaux d'eau

(4) La réponse composite

- Les systèmes, les robots et le simulateur de cerveau tous ensemble.
- Searle: $3 \times 0 = 0$

(5) La réponse des autres esprits

Searle (1980, 421; 1987, 83 trad.)

«Quand l'ordinateur passe des tests comportementaux aussi bien que des personnes (en principe), si vous attribuez de la cognition à celles-ci vous devez en principe faire de même pour celui là.»

- Searle: ce n'est pas «comment je sais que d'autres personnes ont des états cognitifs, mais plutôt ce que je leur attribue quand je leur attribue des états cognitifs. La pointe de mon argument est que ça ne peut se réduire à des processus de calcul ('computational') et leurs sorties, car les uns et les autres peuvent exister sans état cognitif.» (1980, 421f; 1987, 83 trad.)

- La syntaxe ne détermine pas la sémantique. Cependant, nous pensons des choses, et apparemment notre cerveau fonctionne en manipulant des représentations. Qu'est-ce qui pourrait bien donner à ces représentations le sens qu'elles ont?
- C'est difficile à concilier avec le «naturalisme biologique» de Searle.
- Est-ce la théorie de l'identité ou le dualisme? Searle dit ni l'un ni l'autre...

Le fonctionnalisme à nouveau

- Programmes bizarrement implémentés (ou peut-être que l'esprit de Searle est fait de roues à aubes, de gobelets d'étain, etc.): y a-t-il les bons ingrédients?
- Fonctionnalisme: les programmes peuvent être implémentés dans des choses bizarres
- Searle: une bonne programmation ne suffit pas; il faut plutôt les bons ingrédients