



Fonctions,
Algorithmes :
Niveaux
d'abstraction et
niveaux d'explication

PHI 6385 A22 SÉANCE 2
JONATHAN SIMON

Apercu

- 1) Behaviorisme, physicalisme, fonctionnalisme
- 2) Turing : Qu'est-ce qu'une machine de Turing (et pourrait-elle être intelligente) ?
- 3) Marr : Niveaux d'explication et de description dans les sciences cognitives
- 4) Maudlin, Chalmers, Klein, Bostrom : Quand un système compte-t-il comme mettant en œuvre un algorithme ?

1) Behaviorisme, physicalisme,
fonctionnalisme

Comportementalisme / Behaviorisme: les états mentaux sont des états comportementaux

Physicalisme: les états mentaux sont des états neurobiologiques (le matériel du cerveau)

Fonctionnalisme: les états mentaux sont des états computationnels (le logiciel du cerveau)

-
- états mentaux: mon état mental de ressentir une douleur (un événement ou un processus auquel mon cerveau participe)
 - propriétés mentales: ma propriété de ressentir une douleur / d'être dans la douleur (un type / universel / caractéristique qui me caractérise) (ou, la propriété de mon état...)

-
- « The Type-Token Distinction »: La distinction entre les types et les individus
 - Type Physicalism = les propriétés mentales sont des propriétés physiques
 - Token Physicalism = les choses (objets, états) ayant des propriétés mentales sont des choses ayant des propriétés physiques
 - (Le physicalisme de type est plus fort en tant que théorie réductrice)

Comportementalisme: les états/propriétés mentaux sont des états/propriétés comportementaux

Physicalisme: les états/propriétés mentaux sont des états/propriétés neurobiologiques (le matériel du cerveau)

Fonctionnalisme: les états/propriétés mentaux sont des états/propriétés computationnels (le logiciel du cerveau)

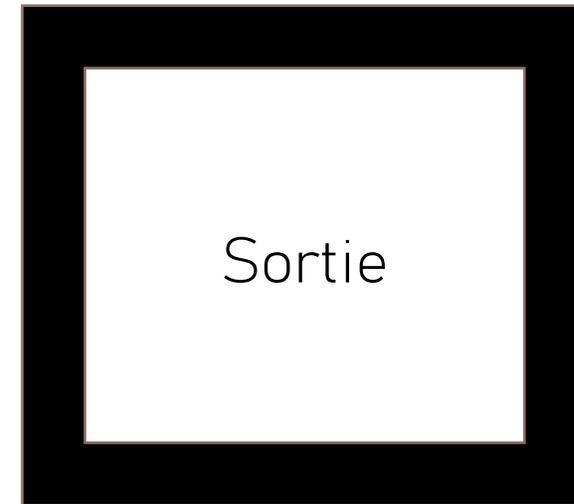
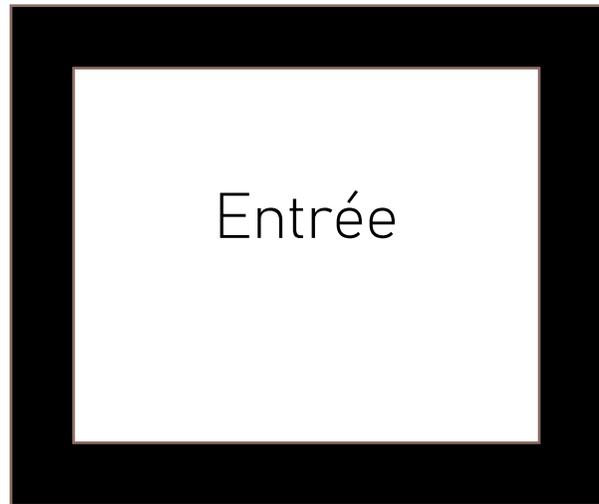
Comportementalisme: un état mental est un ensemble de dispositions/tendances pour se comporter de certaines manières, en fonction de certaines stimuli

- sentir une douleur dans la main, c'est retirer la main si on la touche, dire que ça fait mal si on le demande, etc.

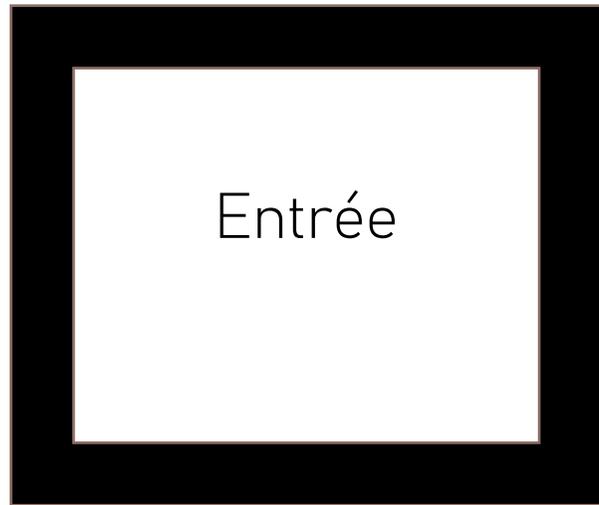
Comportementalisme:

- des fonctions d'entrée-sortie

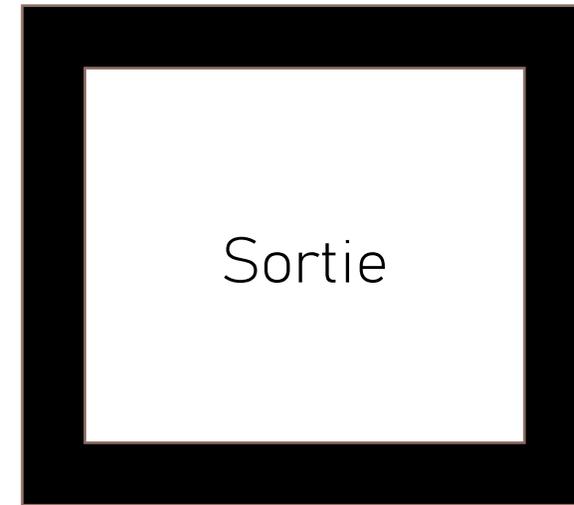
comportementalisme



comportementalisme

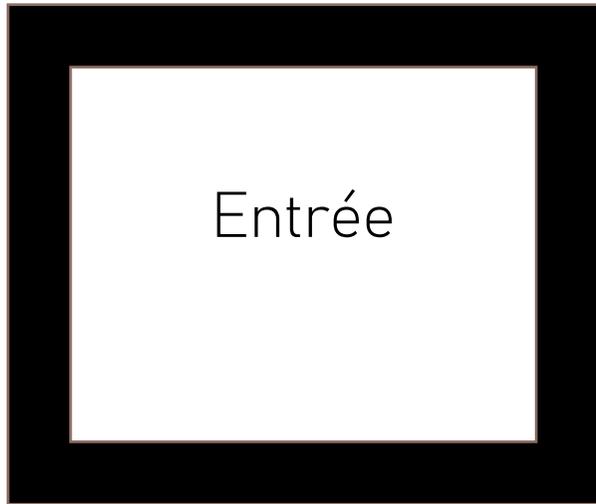


On touche votre main



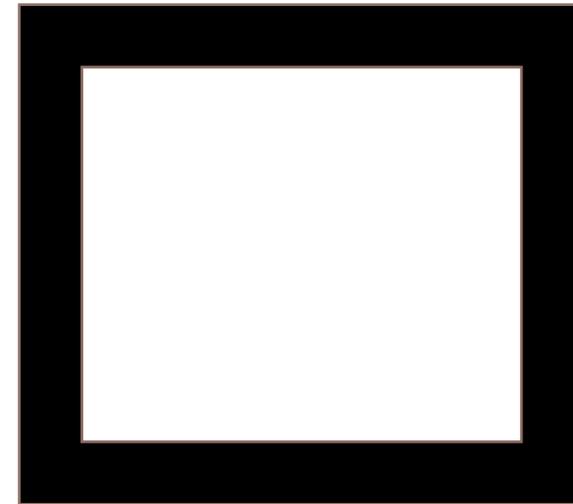
Vous retirez votre main

comportementalisme



On touche votre main

« Comment vous sentez-vous ? »



Vous retirez votre main

« Ça fait mal ! »

Pour contraster:

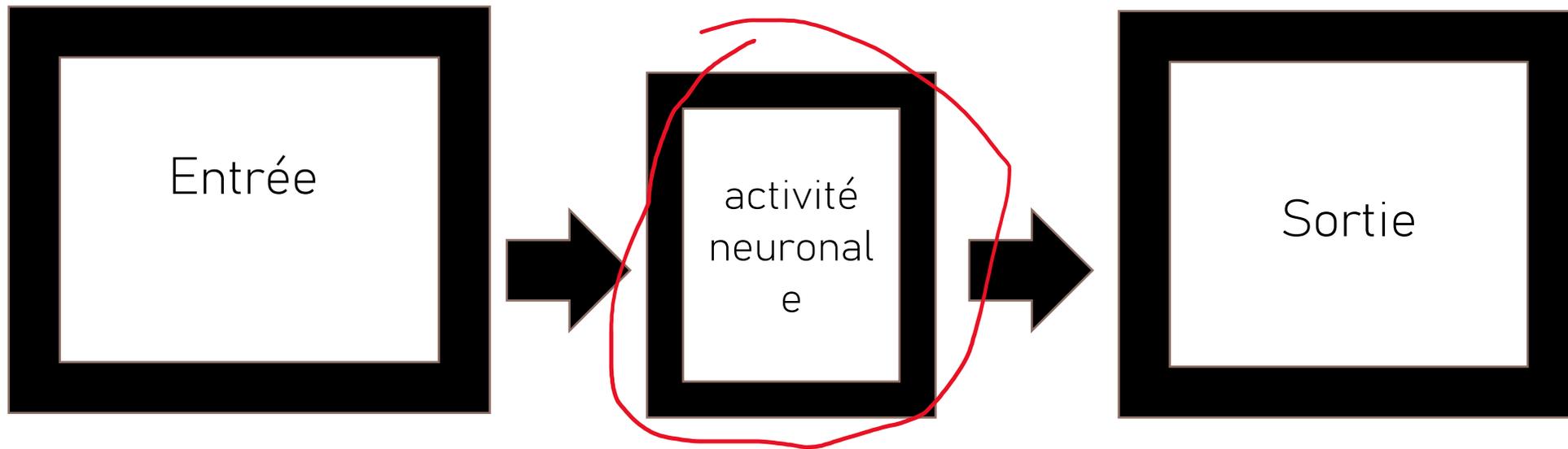
Le behaviorisme philosophique: les états internes sont des dispositions à se comporter...

Le behaviorisme méthodologique: les états internes sont impossibles à étudier, on doit plutôt étudier le comportement.

Physicalisme: un état mental est un état neurobiologique

- sentir une douleur dans la main est pour un certain de vos assemblages neuronaux de tirer selon un certain modèle (le tir des fibres C)

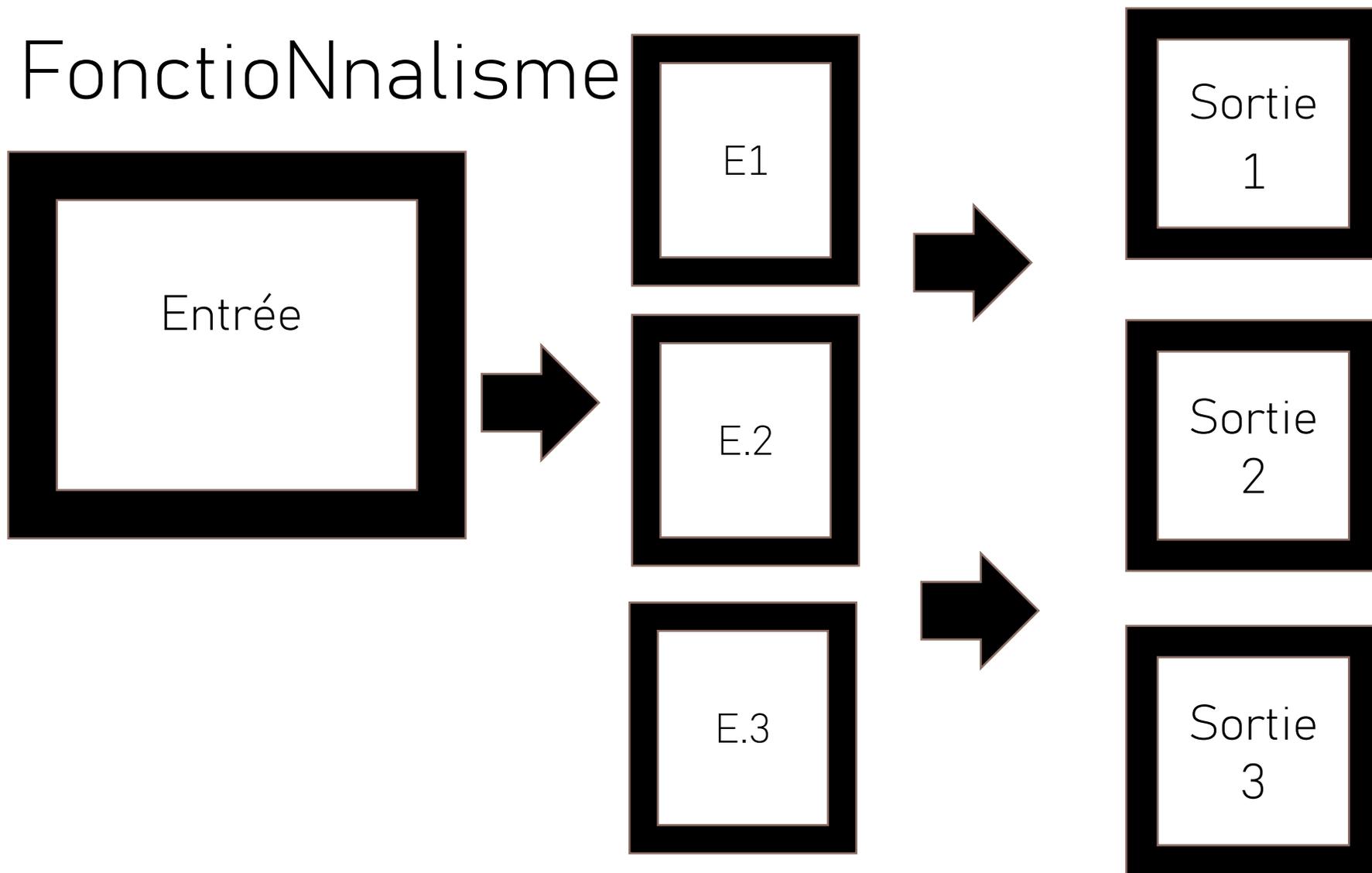
Physicalisme



Fonctionnalisme: un état mental est un état computationnel (ou fonctionnel)

- sentir la douleur dans votre main est pour vous d'être dans un certain état interne qui, *étant donné les autres états mentaux dans lesquels vous êtes*, a tendance à produire un comportement douloureux avec des stimuli appropriés

Fonctionnalisme



-
- Pour chaque équivalence théorique, trois degrés de force (logique):
 - 1) Version Type-A/Analytique/Logique/A Priori: Il n'y a pas de fossé épistémique : l'analyse conceptuelle révèle la vérité de l'équivalence
 - 2) Version Type-B/Synthétique/A Posteriori: Il y'a un fossé épistémique : l'équivalence est un vrai équivalence d'identité, mais il faut le découvrir empiriquement
 - 3) Version Type-D,E,F/Corrélatif: l'équivalence n'est pas un équivalence d'identité, c'est juste une correspondance : les deux choses se présentent ensemble, bien qu'elles ne soient pas les mêmes

Le fonctionnalisme

- Block (dans « Troubles with Functionalism») définit le fonctionnalisme ainsi:
- Fonctionnalisme: Chaque type d'état mental est un état qui, étant donné certaines entrées/stimuli sensoriels et certains autres états mentaux, consiste dans une disposition à agir d'une certaine façon et à avoir certains états mentaux

Le fonctionnalisme

- Pour contraster :
- Comportementalisme: Chaque type d'état mental est un état qui, étant donné certains inputs sensoriels et certains états mentaux, consiste dans une disposition à agir d'une certaine façon et à avoir certains états mentaux

Le fonctionnalisme

- Un premier défi pour le fonctionnalisme : la circularité
- Le fonctionnalisme devrait offrir une analyse des états mentaux (au moins la version de type A doit le faire). Mais une analyse ne devrait pas faire appel à ce qu'elle analyse (les états mentaux)

Le fonctionnalisme

- Deux approches à ce défi :
- 1) Le fonctionnalisme machinique (Putnam)
- 2) Fonctionnalisme causal-théorique (Lewis)

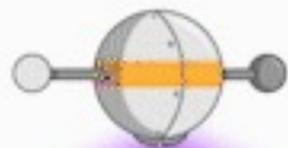
Le fonctionnalisme machinique

- Le fonctionnalisme machinique comprend les états mentaux comme des états d'un système de calcul (typiquement : une machine de Turing).

2) Turing

Le fonctionnalisme machinique

- Le fonctionnalisme machinique comprend les états mentaux comme des états d'un système de calcul (typiquement : une machine de Turing: De nombreuses définitions de la computation s'avèrent être équivalentes aux machines de Turing.).



State 2

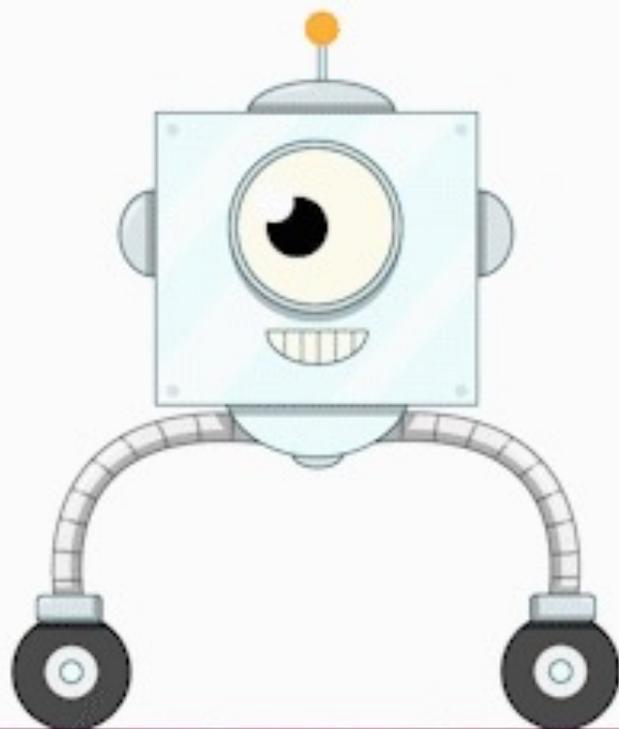
if 0

then erase

write 1

move right

go to state 5



1 1 0 1 0 0 1 0 1 1

Le fonctionnalisme machinique

- Une machine de Turing : une longueur de bande / papier (aussi longue que nécessaire) divisée en cellules.
- Chaque cellule porte un numéro écrit dessus.
- Un lecteur/enregistreur fonctionne sur une cellule à la fois.
- Il lit le numéro, puis le modifie selon son *état*, et selon un ensemble de règles, puis se déplace vers une autre cellule (et modifie son état) selon un ensemble de règles.

Le fonctionnalisme machinique

- Il est essentiel que les règles indiquent au scanner ce qu'il doit écrire étant donné ce qu'il lit *et dans quel état il se trouve*. Les règles disent aussi *dans quel état il doit aller*.

Le fonctionnalisme machinique

La liste des règles est appelée *une table de machine*. Voici un exemple de table de machine pour une machine à coke (prochaine page). Notez que la "signification" des états est définie implicitement (les étiquettes ne sont ici qu'à des fins heuristiques)

Example: Machine à Coke (Coke coute 10c)

TABLE 1

	S_1	S_2
Input : pièce de 5 cents	N'émettre aucun output Passer en S_2	Emettre un coca-cola Passer en S_1
Input : pièce de 10 cents	Emettre un coca-cola Rester en S_1	Emettre un coca-cola et une pièce de 5 cents Passer en S_1

Example: Machine à Coke (coke coute 15c)

Input	État	Prochain État	Output
5c	Désire 15c	Désire 10c	rien
5c	Désire 10c	Désire 5c	rien
5c	Désire 5c	Désire 15c	Coke!
10c	Désire 15c	Désire 5c	rien
10c	Désire 10c	Désire 15c	Coke!
10d	Désire 5c	Désire 10c	Coke!

Example: Machine à Coke (coke coute 15c)

Input	État	Prochain État	Output
5c	S_0	S_1	rien
5c	S_1	S_2	rien
5c	S_2	S_0	Coke!
10c	S_0	S_2	rien
10c	S_1	S_0	Coke!
10d	S_2	S_1	Coke!

Le fonctionnalisme machinique

Observations :

1) Seules les entrées et les sorties sont définies explicitement, les états internes sont définis implicitement

2) Holisme : les états ne sont définis qu'implicitement en référence les uns aux autres. Si vous modifiez l'un des états, vous modifiez chacun d'eux

Le Test de Turing

- Turing fait beaucoup de choses dans son article, entre autres, proposer son célèbre test pour savoir quand une machine est intelligente. Nous reviendrons sur ce sujet dans quelques semaines.

Marr: niveaux d'explication

Marr

- 3 Niveaux :
- Le niveau computationnel
- Le niveau algorithmique
- Le niveau d'implémentation

Marr

- Le niveau computationnel
- Le niveau général, en précisant le problème à résoudre :
p. ex.: «faire un gâteau»

Marr

- Le niveau algorithmique
- Le niveau spécifiant les éléments cruciaux du plan pour résoudre le problème, par exemple : «La recette de ma grand-mère pour faire un gâteau». (Mélange une poignée de farine, une cuillère de sucre...).

Marr

- Le niveau d'implémentation
- Le niveau spécifiant tous les détails de la façon dont on suit exactement la recette. Quelle quantité de farine exactement, et quelle sorte (raffinée ou non) ? Quelle quantité de sucre exactement, et quelle sorte (brun, blanc, canne) ? Quel type de bol à mélanger, en bois ou en métal ? Etc...

Marr

Note que cette distinction en 3 niveaux est arbitraire : en réalité, nous pouvons parler de degrés de spécification, de la quantité de zoom que nous faisons....

(malgré tout, la règle des 3 est convaincante, par exemple dans le domaine de la gestion, la même distinction apparaît : stratégique vs tactique vs opérationnel).

Marr

Note que le physicalisme considère naturellement la conscience comme une question de mise en œuvre, le fonctionnalisme, comme une question d'algorithme, et le behaviorisme, comme une question de problème de calcul.

(mais encore une fois, nous devons nous rappeler qu'il y a vraiment un spectre ici, et non pas 3 niveaux clairement distincts)

Marr

Note que le physicalisme considère naturellement la conscience comme une question de mise en œuvre, le fonctionnalisme, comme une question d'algorithme, et le behaviorisme, comme une question de problème de calcul.

(mais encore une fois, nous devons nous rappeler qu'il y a vraiment un spectre ici, et non pas 3 niveaux clairement distincts)

Chalmers, Maudlin, Klein, Bostrom:

Bonnes et mauvaises abstractions

Bonnes et mauvaises abstractions

- Si les machines de Turing (en tant que descriptions abstraites ou mathématiques) spécifient des algorithmes, comment dire qu'un système donné "est" (ou, implémente / met en oeuvre) une machine de Turing ?

Bonnes et mauvaises abstractions

- Premier problème : les descriptions des machines de Turing sont trop abstraites - même un rocher pourrait être "interprété" comme une machine de Turing
- Deuxième problème : les descriptions des machines de Turing sont trop exigeantes : même les ordinateurs ont des problèmes. Cela signifie-t-il que ce ne sont pas vraiment des ordinateurs ?

Bonnes et mauvaises abstractions

- Problème: les descriptions des machines de Turing sont trop abstraites - même un rocher pourrait être «interprété» comme une machine de Turing

Bonnes et mauvaises abstractions

- Problème: les descriptions des machines de Turing sont trop abstraites - même un rocher pourrait être «interprété» comme une machine de Turing

Example: Machine à Coke (coke coute 15c)

Input	État	Prochain État	Output
5c	Désire 15c	Désire 10c	rien
5c	Désire 10c	Désire 5c	rien
5c	Désire 5c	Désire 15c	Coke!
10c	Désire 15c	Désire 5c	rien
10c	Désire 10c	Désire 15c	Coke!
10d	Désire 5c	Désire 10c	Coke!

Example: Machine à Coke (coke coute 15c)

Input	État	Prochain État	Output
5c	S_0	S_1	rien
5c	S_1	S_2	rien
5c	S_2	S_0	Coke!
10c	S_0	S_2	rien
10c	S_1	S_0	Coke!
10d	S_2	S_1	Coke!

Example: Machine à Coke (coke coute 15c)

Input	État	Prochain État	Output
l_1	S_0	S_1	O_1
l_1	S_1	S_2	O_1
l_1	S_2	S_0	O_2
l_2	S_0	S_2	O_1
l_2	S_1	S_0	O_2
l_2	S_2	S_1	O_2

Bonnes et mauvaises abstractions

- Problème: les descriptions des machines de Turing sont trop abstraites - même un rocher pourrait être «interprété» comme une machine de Turing

Bonnes et mauvaises abstractions

- Problème:

Étiquette l'environnement initial de la roche comme I_1 , sa configuration initiale S_1 , sa configuration un moment plus tard S_2 , l'environnement comme O_1 : alors il correspond à une ligne sur ce tableau !

Bonnes et mauvaises abstractions

- Problème:
- Beaucoup répondent que a) les états d'entrée et de sortie sont fixés à l'avance
- b) nous avons besoin d'une sorte de contrainte contrefactuelle, ou d'isomorphisme entre table de machine et mécanisme, c'est-à-dire que si le système avait été dans un état initial différent ou avait reçu une entrée différente, il aurait quand même satisfait à la description en question...

Bonnes et mauvaises abstractions

- problèmes résiduels :
- 1) Est-ce trop fort ? Nous devons permettre que, par exemple, les ordinateurs puissent avoir des problèmes (tout en restant des ordinateurs). Quel degré de défaillance pouvons-nous autoriser ? À quel moment un système passe-t-il du statut d'ordinateur très glitchy ou peu fiable, à celui de ne plus être un ordinateur (Bostrom).

Bonnes et mauvaises abstractions

- problèmes résiduels :
- 2) Cause par omission : le fait de se fier à un critère contrefactuel peut impliquer que cela fait une différence, pour quelque chose qui ne se produit pas, si cela aurait pu se produire (l'alto silencieux ajoute à l'esthétique de la pièce).

Bonnes et mauvaises abstractions

- problèmes résiduels :
- 3) Ce n'est pas assez fort ? Nous pouvons décrire des systèmes qui respectent la condition contrefactuelle, mais qui intuitivement ne semblent toujours pas correspondre à la réalité (Maudlin, Klein).