

L'ÉXPÉRIENCE DE PENSÉE DE LA CHAMBRE CHINOISE

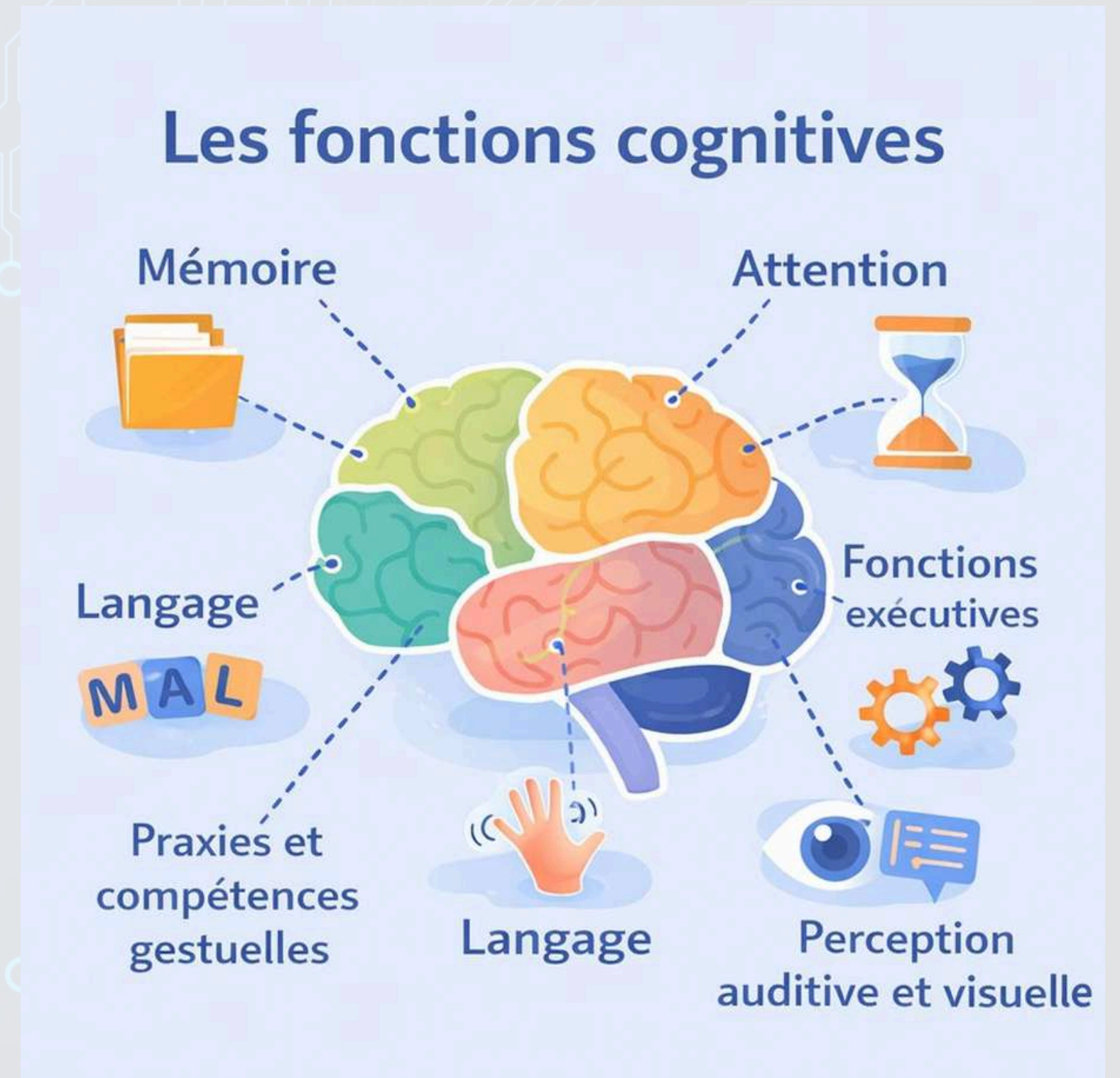
JOHN SEARLE (1980)

Encadrante: Mme Marie lefevre

Paul, Ali, Oscar, Ryan

LA CHAMBRE CHINOISE: EST-CE QUE L'IA POSSÈDE UNE CAPACITÉ COGNITIVE?

Nous allons traiter ce problème d'un point de vu philosophique selon John Searle l'auteur de cette expérience.



SOMMAIRE

1. Introduction

- Contexte de la conscience des machines.
- L'expérience de la chambre chinoise
- L'exécution d'un programme informatique suffit-elle à générer un esprit conscient ?

2.L'expérience de pensée de la Chambre chinoise

- expérience: l'homme, les règles en anglais, les symboles chinois.
- Critique de l'IA forte selon Searle
- la distinction fondamentale entre syntaxe et sémantique.

3. Les objections et les réponses de Searle

- La réponse du système (System Reply) :
- La réponse du robot (Robot Reply)
- L'expérience : un aimant dans le noir

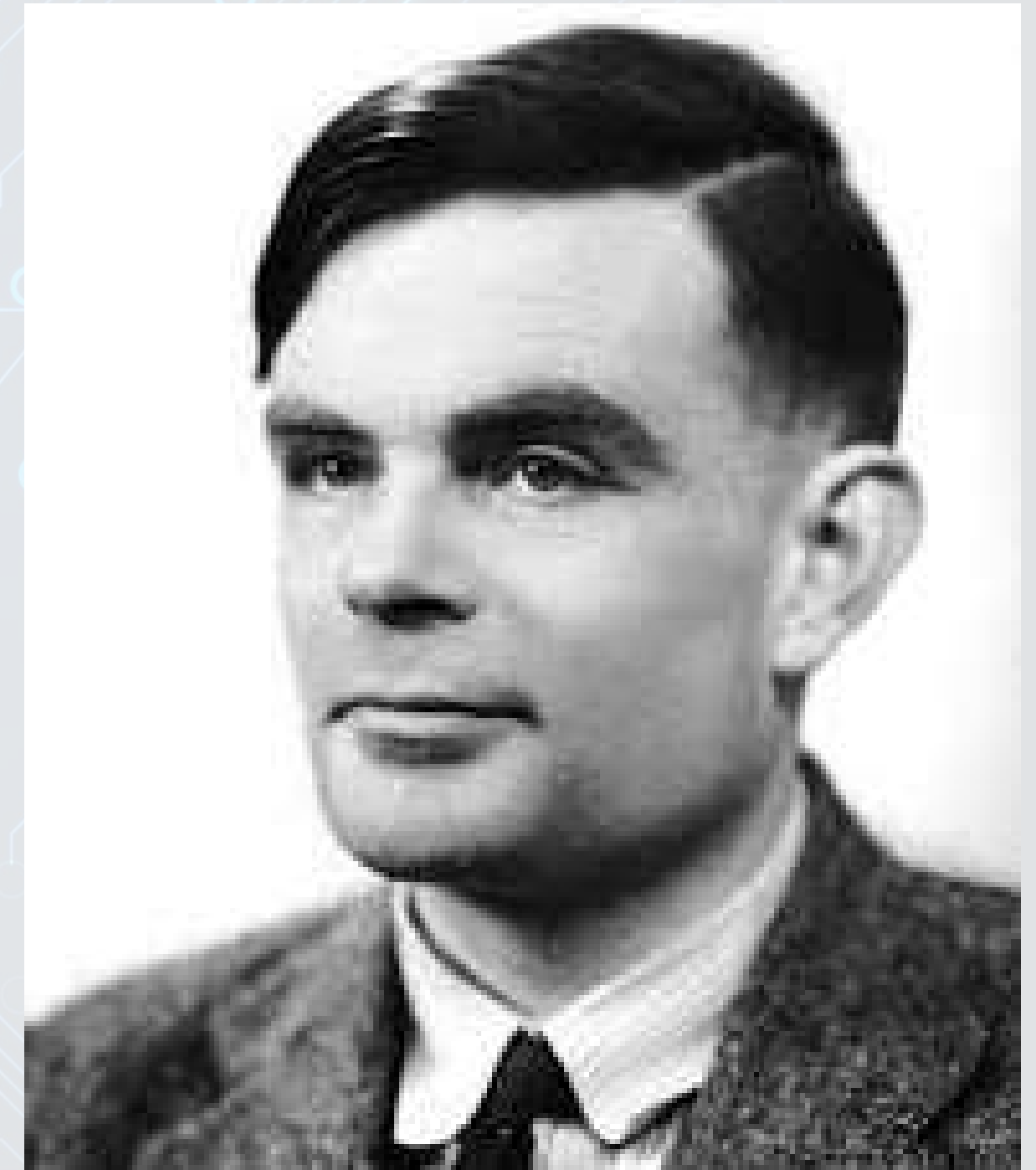
4. Portée contemporaine, limites et débats actuels

- Les Points de vues philosophiques
- Les LLM modernes qu'on connaît
- Impact sur la philosophie de l'esprit moderne

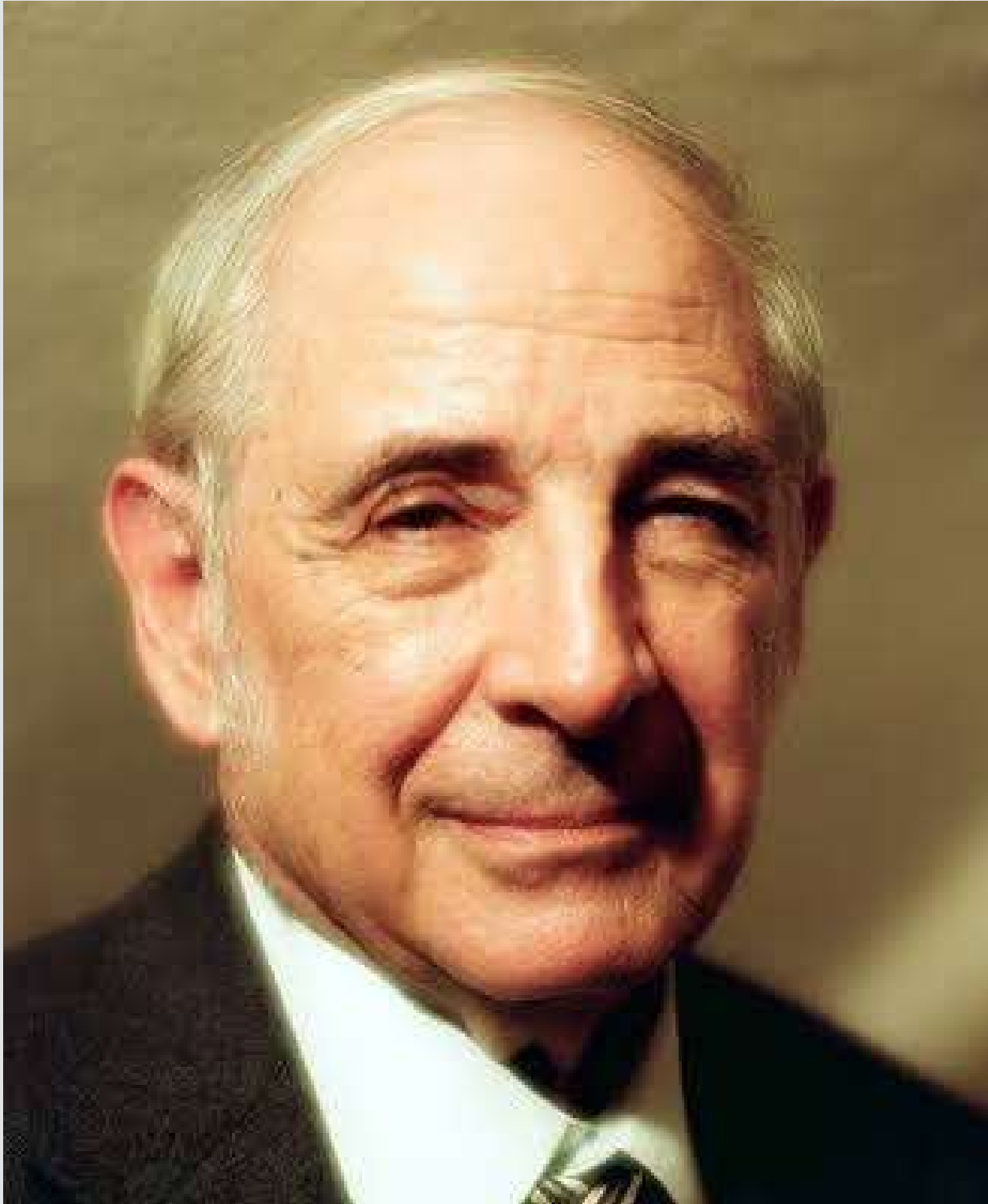
5. Conclusion

CONTEXTE HISTORIQUE

- EN 1950 l'idée célèbre d'Alan Turing si une machine pense ou pas "COMPUTING MACHINERY AND INTELLIGENCE"
- L'intelligence artificielle symbolique est le courant informatique le plus connue à l'époque
- Système expert (1965 ,Dendral) utiliser pour déterminer des structures moléculaires



JOHN SEARLE



- John Searle philosophe américain né en 1932 et décède l'an dernier
- Connue notamment pour ses travaux concernant la philosophie de l'esprit et notamment la chambre chinoise.
- Un point de vue différent d'autres scientifiques

LA PROBLÉMATIQUE ESSENTIEL

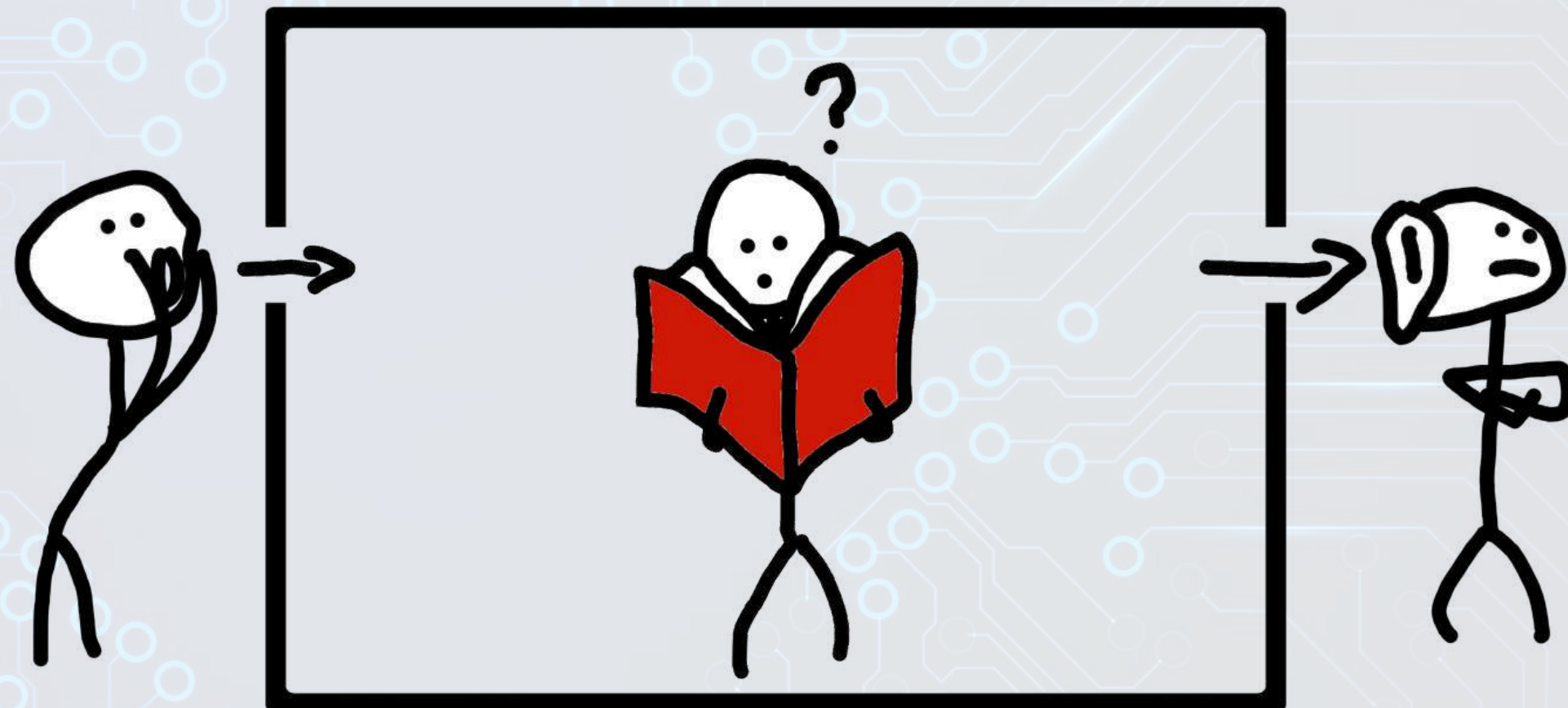
PENSÉE HUMAINE = SUIVRES DES RÈGLES?

- Les étudiants utilisent toujours des LLM pendant leurs études.
- Manipuler des symboles et exécuter des programmes suffit pour dire si une machine comprend?

Partie 2

L'ÉXPÉRIENCE DE PENSÉE DE LA CHAMBRE CHINOISE

JOHN SEARLE (1980)



(schéma dessiné par nos soins)



Paul

小鐵買車

住家箱

Français

Chinois

Paul



Français

~~Chinois~~

Paul



Paul

小鐵櫃

住家箱

Français

Paul

小鐵買車

住家箱

Français
Chinois → Chinois
Paul

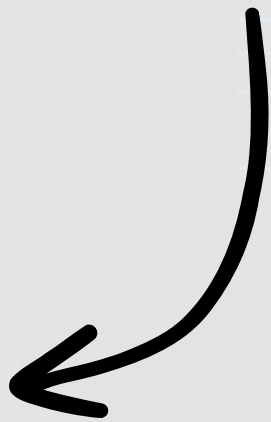


Paul

小鐵買車

住家箱

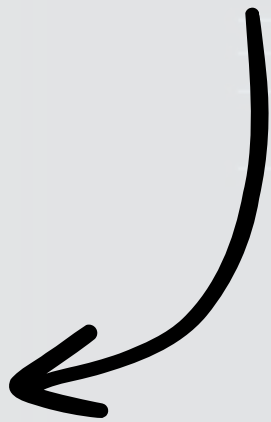
Ali



Paul



Ali



你好吗？

Paul



“Si on te dit 你好吗？, réponds 好的”

Paul

Ali

你好吗？



“Si on te dit 你好吗 ?, réponds 好的 “

Ali

Paul

你好吗 ?

小鐵櫃

住家箱

Boite

“Si on te dit 你好吗？, réponds 好的”

Ali

Paul

你好吗？

好的

Boite

“Si on te dit 你好吗 ?, réponds 好的 “

Ali

Paul

你好吗 ?

好的



“Si on te dit 你好吗？, réponds 好的”

Ali

Ryan

Paul

小鐵櫃

住家箱

你好吗？

好的

“Si on te dit 你好吗？, réponds 好的”

Ali

Paul

Ryan

你好吗？

好的



小鐵櫃

住家箱

Es-ce que Paul a compris ce qu'il vient de renvoyer ?

Partie 2

L'ANALOGIE INFORMATIQUE

La chambre reproduit exactement l'architecture d'un ordinateur.



L'homme enfermé (Paul) = Le processeur (CPU)



Le manuel en français = Le programme informatique



Les boites de symboles = La base de données

LA CRITIQUE DE J. SEARLE DE L'IA FORTE



Le Test de Turing

RÉUSSI HAUT LA MAIN

Illusion parfaite d'une conscience à l'extérieur.



La Faille (Paul)

ZÉRO COMPRÉHENSION

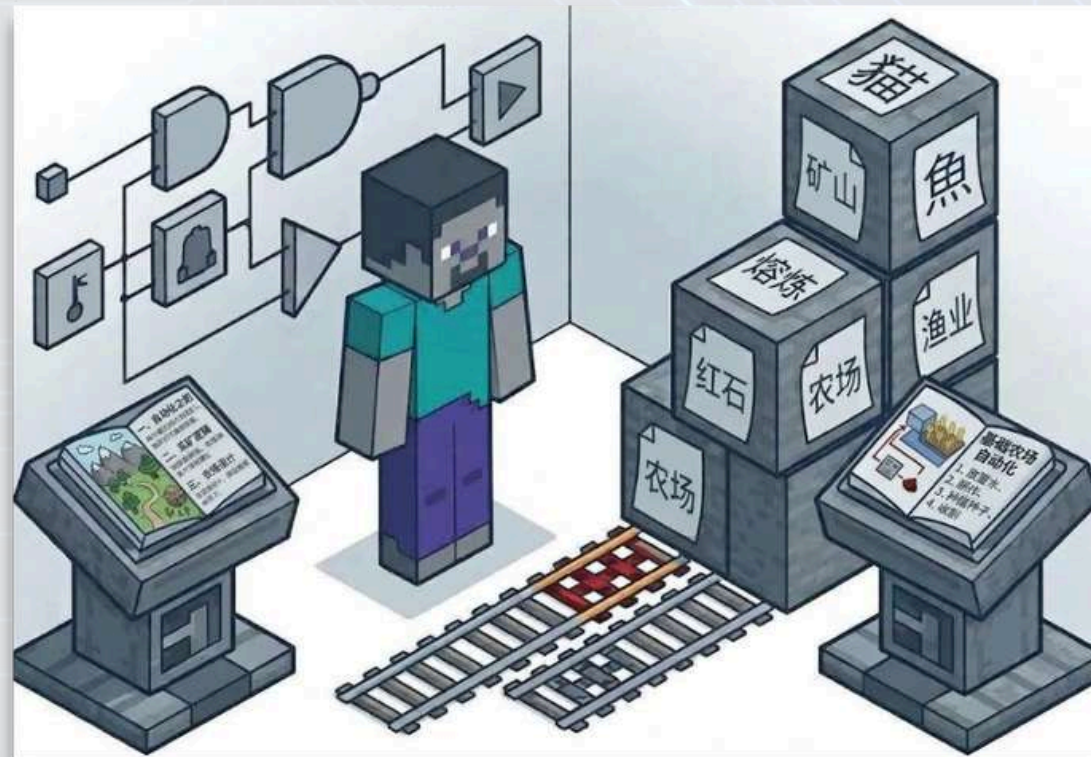
Aucune conscience des mots manipulés à l'intérieur.

Simuler l'intelligence n'est pas la dupliquer.

SYNTAXE

VS

SÉMANTIQUE



Manipulation mécanique aveugle



Compréhension du sens et conscience

« La syntaxe ne suffira jamais à créer la sémantique. »

J. R. Searle

Partie 3

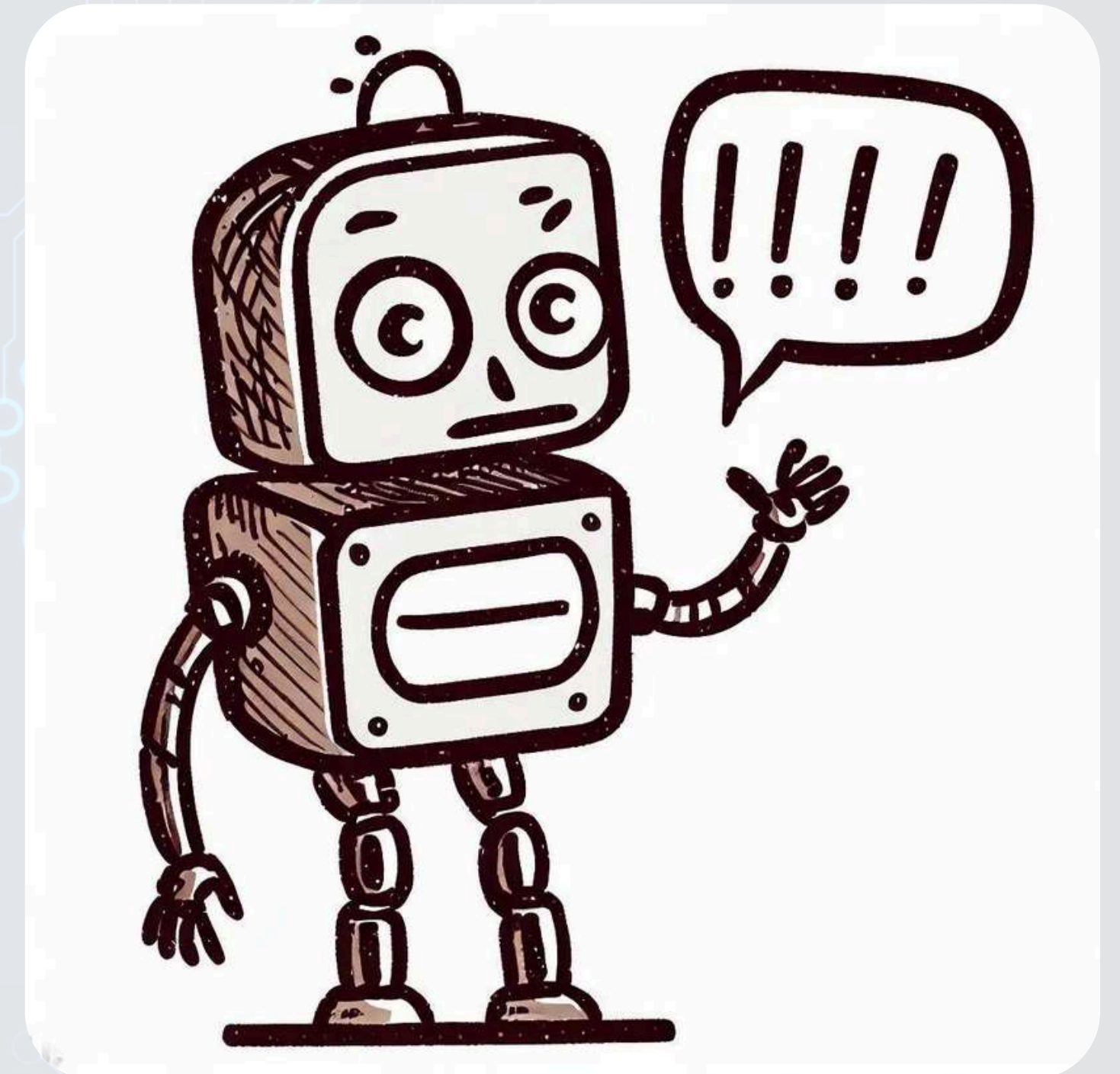
LES OBJECTIONS ET LES RÉPONSES DE SEARLE

1980 : l'article paraît, avec 27 réponses critiques

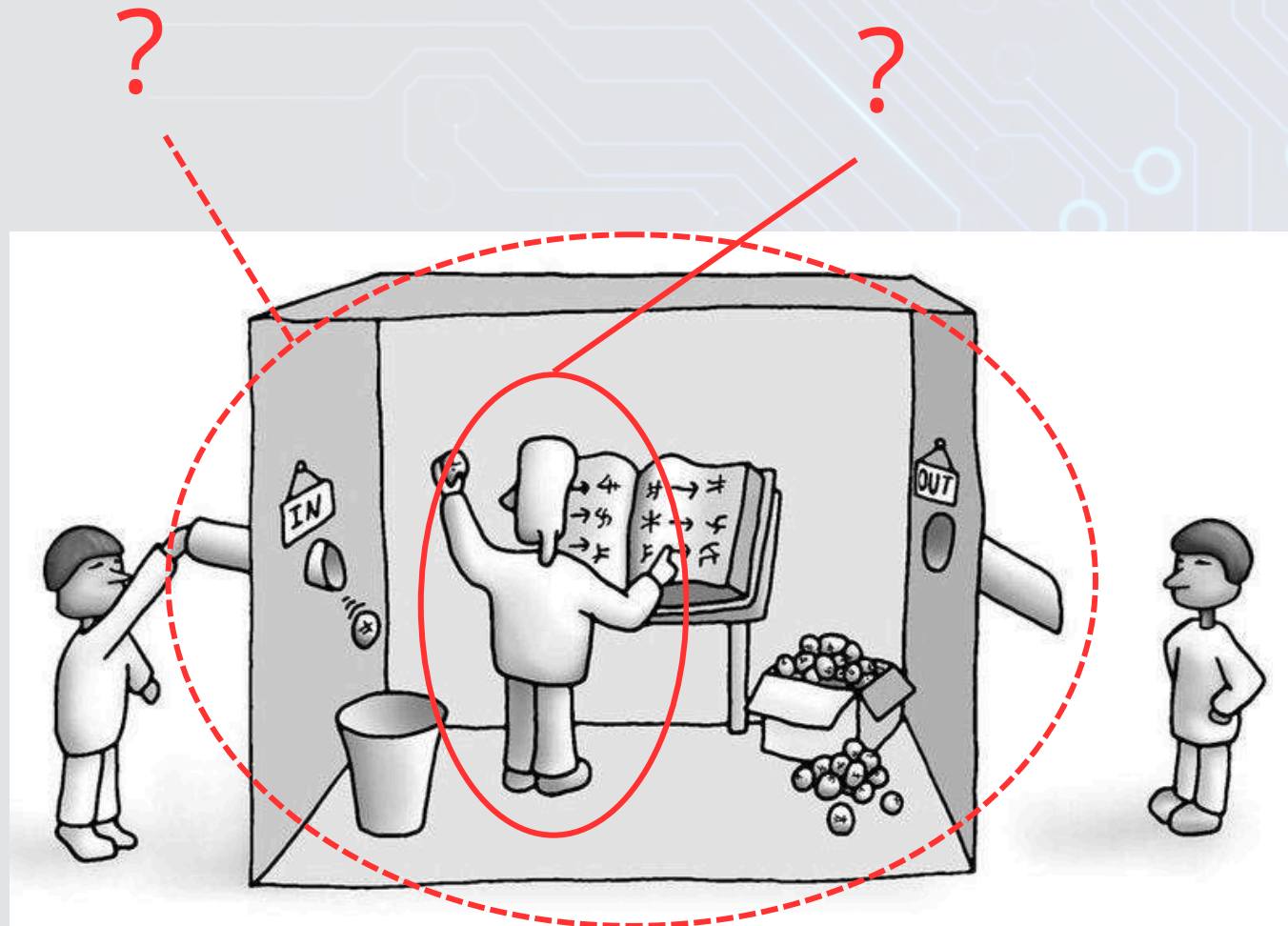
La syntaxe ne suffira jamais à créer la sémantique.

J. R. Searle

aujourd'hui : plus de 45 ans de débat



L'OBJECTION : C'EST LE SYSTÈME QUI COMPREND



L'objection la plus célèbre, encore débattue aujourd'hui

CITATION

« Certes, l'individu enfermé dans la chambre ne comprend pas l'histoire, mais il n'est qu'une partie d'un système global, et le système, lui, comprend l'histoire. »

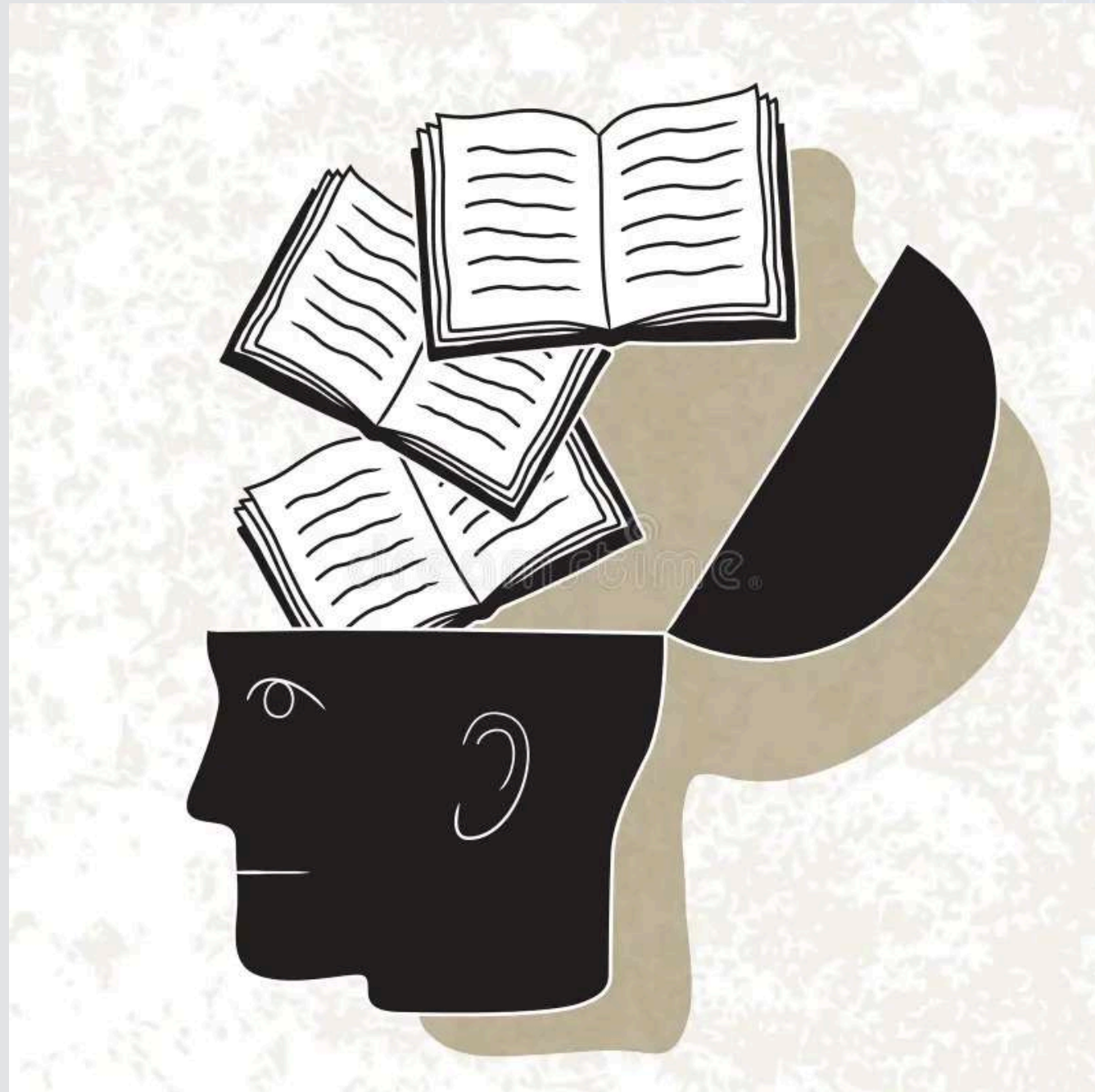
Chercheurs de Berkeley, cités par John R. Searle, « Minds, Brains and Programs », 1980

Pour faire simple:

L'homme n'est qu'une pièce du mécanisme, L'ensemble lui comprends

Exemple: Un neurone seul / les fourmies

LA RÉPONSE DE SEARLE : TOUT APPRENDRE PAR CŒUR



CITATION

« Que l'individu intériorise tous les éléments du système. [...] Il mémorise les règles du registre et les banques de symboles chinois [...] et il fait tous les calculs dans sa tête. [...] Il ne comprend rien au chinois, et a fortiori le système non plus. »

Chercheurs de Berkeley, cités par John R. Searle, « Minds, Brains and Programs », 1980

Pour faire simple:

Si tout le système est dans la tête de Searle et qu'il ne comprend toujours pas, alors le système ne comprend pas non plus.

Probleme ? Hypothese de base devient fausse ...

L'OBJECTION : RELIER LES SYMBOLES AU MONDE



CITATION

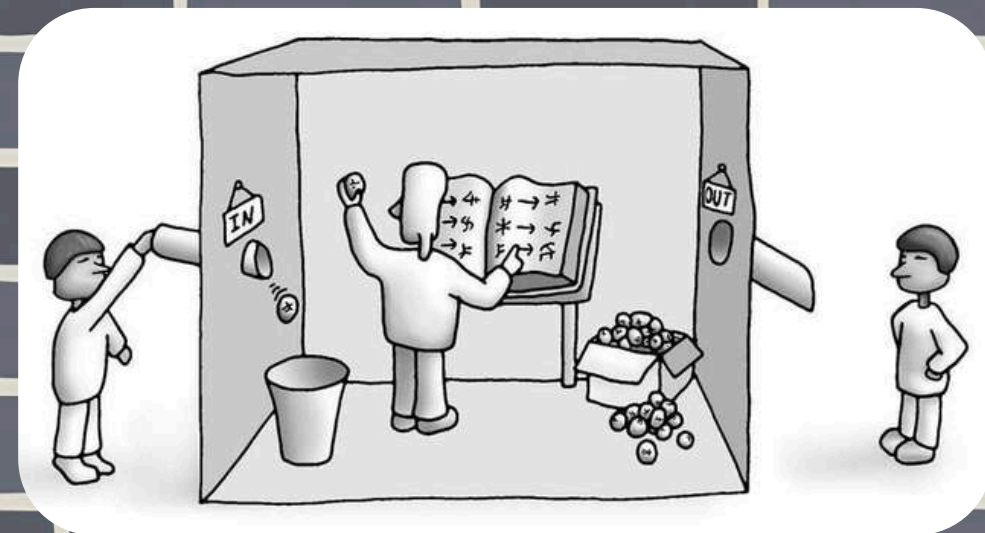
« Supposons qu'on place un ordinateur dans un robot, et que cet ordinateur ne se contente pas de recevoir et de renvoyer des symboles formels, mais fasse réellement fonctionner le robot [...] Un tel robot aurait, contrairement à l'ordinateur de Schank, une véritable compréhension et d'autres états mentaux. »

Chercheurs de Yale, cités par John R. Searle (1980)

Pour faire simple:

La chambre est coupée du monde. Avec des caméras et des bras, les mots seraient reliés aux choses.

Exemple : Chat reconnu par vision



LA RÉPONSE DE SEARLE : ENCORE DES SYMBOLES



CITATION

« La première chose à remarquer, c'est que la réponse du robot concède implicitement que la cognition n'est pas seulement une affaire de manipulation formelle de symboles [...] l'ajout de telles capacités “perceptives” et “motrices” n'apporte rien en matière de compréhension. »

John R. Searle, « Minds, Brains and Programs », 1980

Pour faire simple:

les caméras lui envoient encore des symboles. Mais plus de symboles = pas plus de sens.

```
1110111011000110110101100010110111101100
1010011011110011001100010111101101111000
0011000100000101010111010011011011010101
1010010111111011111111000110110101001000
0001000111101101000110001101111111101000
0010111001111111011000101101011100101100
01001000101011011101100001001111101000001
0011111100010101011100000100111001111000
0110010011110001111010001101100001011100
1111001100101010001110000001010011111000
1010001011000011001011111101100101110011
11010110110111111111110000000110000111000
1010011001101001011010011100111010100101
00111010010100111111110000010110110101110
0101111111101110110001101101011000101101
0100001110100110111100110011000101111011
1011010000110001000001010101110100110110
0010100010100101111111011111110001101101
001001010001000111101101000110001101111
```

L'EXPÉRIENCE : UN AIMANT DANS LE NOIR



CITATION

« Imaginez une pièce sombre où un homme tient un aimant ou un objet chargé. [...] Il est inconcevable de produire de la vraie lumière simplement en déplaçant des forces ! »

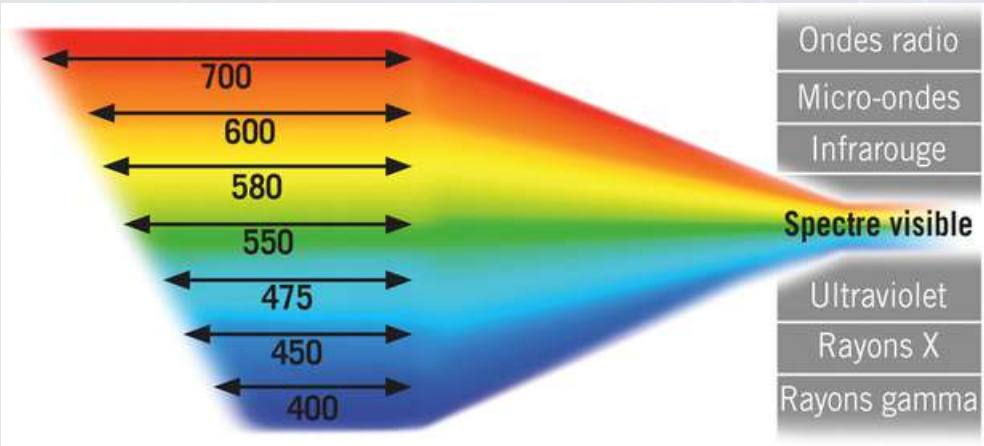
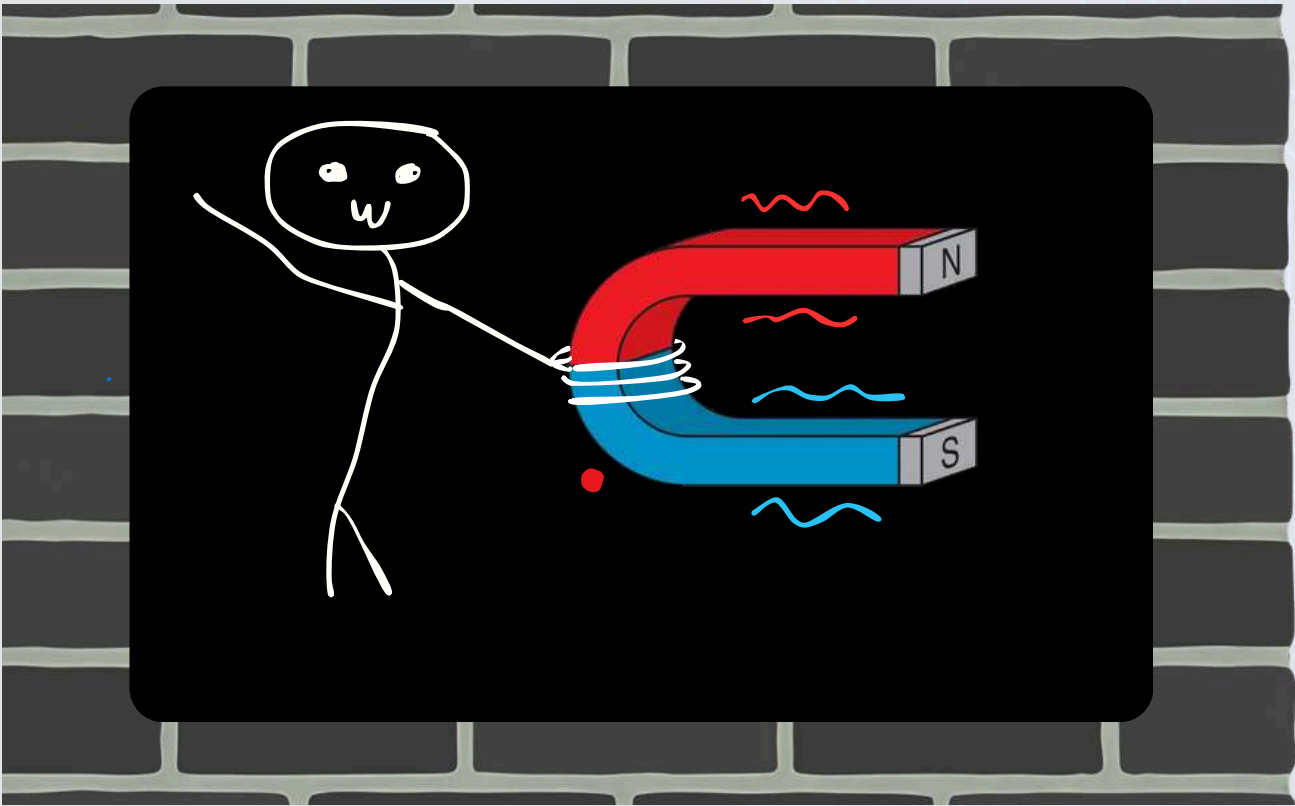
Paul M. Churchland & Patricia S. Churchland, « Could a Machine Think? », Scientific American, 1990

Pour faire simple:

C'est une parodie de Searle : Pour les Churchland, Searle fait la même erreur que l'homme à l'aimant. (une version minuscule et ultra lente du processus)

Exemple : L'homme à l'aimant / La difference entre un "La" et un "toc"

Info: réponse de Searle: « L'explication de la lumière par le rayonnement électromagnétique est une histoire causale de bout en bout. Mais l'analogie avec les symboles formels échoue, car les symboles formels n'ont aucun pouvoir physique, causal. »



Do	Ré	Mi	Fa	Sol	La	Si
261.6	293.7	329.6	349.2	392.0	440.0	493.9

MISE EN PERSPECTIVE, LIMITES ET DÉBATS ACTUELS



De Searle aux LLM, un débat
toujours ouvert

POINTS DE VUE PHILOSOPHIQUES



John Searle
(1932-2025)



Daniel Denett
(1942-2024)



David Chalmers
(1966)

Traiter des symboles ne suffit pas à comprendre.

La compréhension vient du cerveau biologique =
naturalisme biologique

Si un système se comporte comme s'il comprenait, on peut
parler de compréhension fonctionnelle.

La compréhension est une propriété émergente du système
global.

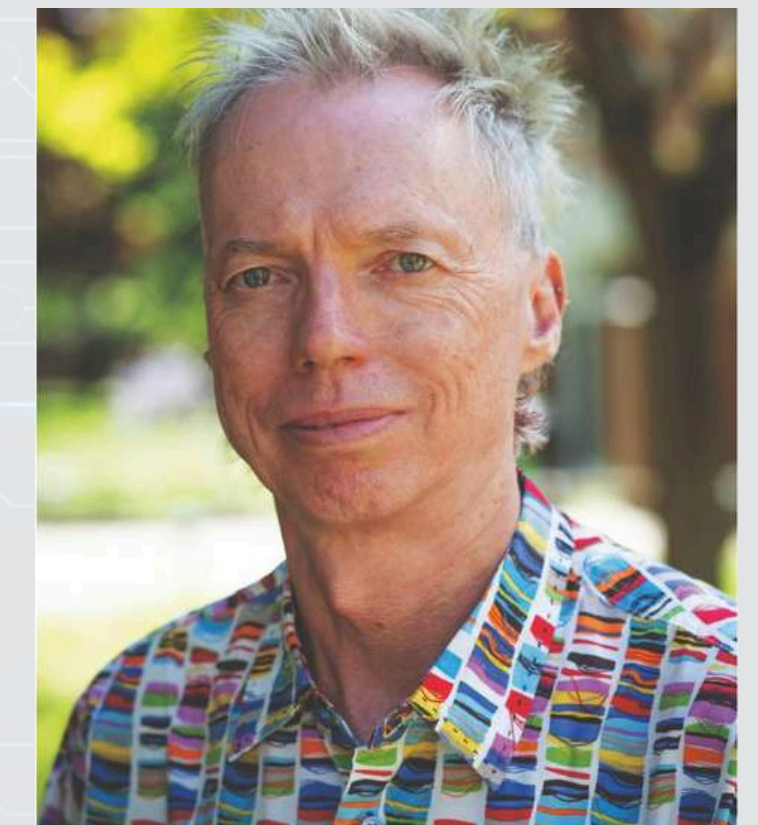
ASPECT COGNITIVE

Harnad & Clark et la cognition incarnée :

- Le sens naît de l'interaction avec le monde, pas des symboles pris isolément
- Harnad : le "symbol grounding problem" → sans ancrage dans le réel, les symboles ne renvoient qu'à d'autres symboles, jamais au sens
- Clark : la pensée s'étend au-delà du cerveau → elle inclut le corps et l'environnement (thèse de "l'esprit étendu")
- Comprendre suppose de percevoir et d'agir, pas seulement de manipuler des symboles



Stevan Harnad (1945)



Andy Clark (1957)

ASPECT COGNITIVE

McCulloch — position neurocognitive :

- L'esprit émerge des circuits neuronaux, pas d'un programme abstrait
- Dès 1943, avec Walter Pitts, il propose le premier modèle mathématique du **neurone artificiel** → à l'origine des réseaux de neurones actuels
- Formule clé : "The mind is in the head" → **l'esprit** est ancré dans un substrat physique réel
- La compréhension vient d'un système physique complexe, pas de la seule manipulation symbolique



Warren Sturgis McCulloch (1898–1969)

MISE EN PERSPECTIVE : LES LLM (CÔTÉ SEARLE)



Comprends-tu ce que tu dis ?

Tout dépend de ce qu'on entend par « comprendre » : je manipule le sens des mots de façon cohérente et utile, mais on ne sait pas si cela s'accompagne d'une compréhension vécue au sens humain, ce qui est précisément la question posée par l'argument de la Chambre chinoise de Searle.



CITATION

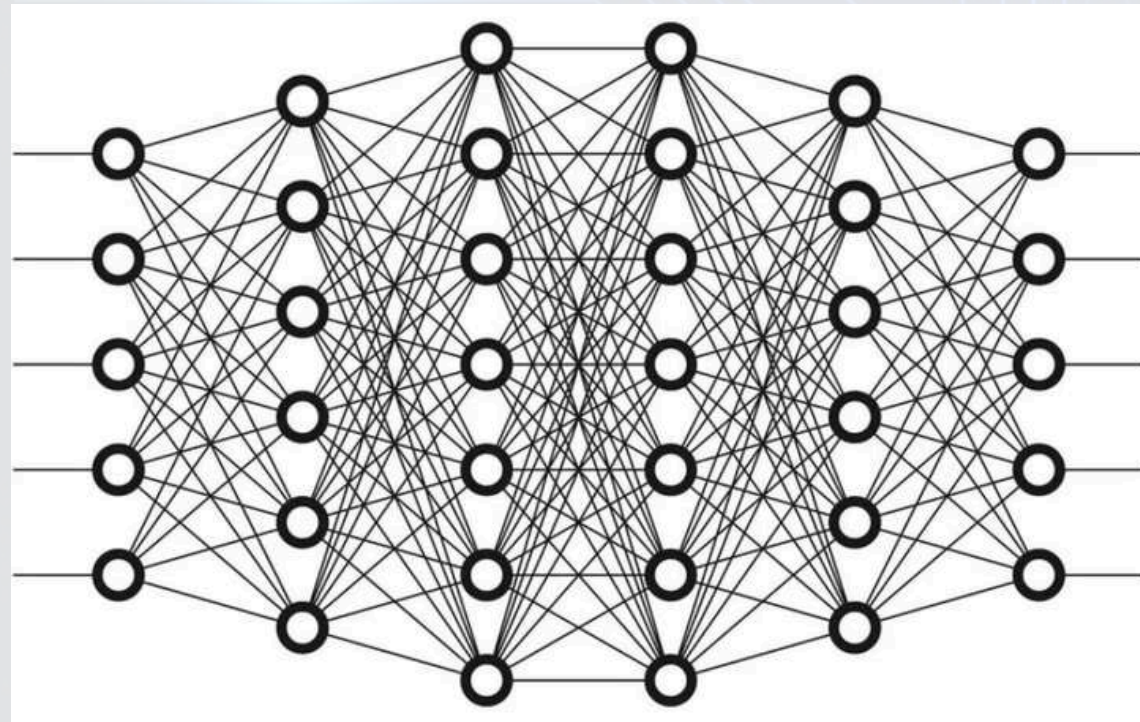
Francais: « un modèle de langage est un système qui assemble au hasard des suites de formes linguistiques observées dans ses immenses données d'entraînement, selon des probabilités sur leur combinaison, mais sans aucune référence au sens : un perroquet stochastique. »

Emily M. Bender, Timnit Gebru, Angelina McMillan-Major & Shmargaret Shmitchell (2021)

Pour faire simple:

Un LLM serait une chambre chinoise géante : il prédit le mot suivant, sans savoir de quoi il parle.

MISE EN PERSPECTIVE: LES LLM (UN SENS QUI ÉMERGE ?)



CITATION

« Le succès des grands modèles de langage a suscité le scepticisme : possèdent-ils quoi que ce soit comme des concepts ou du sens humains ? Contrairement à l'idée que les LLM n'ont aucun sens, nous soutenons qu'ils capturent probablement des aspects importants du sens. »

Steven T. Piantadosi & Felix Hill, « Meaning without reference in large language models », 2022



Pour faire simple:

C'est la suite "du système" version 2022 : le sens viendrait des relations entre les concepts que le modèle a appris.

Exemple : Fourmies = neurones

L'ÉTHIQUE SUR L'IA

Une IA peut produire des réponses cohérentes sans rien "comprendre". Cela pose des problèmes de responsabilité, de confiance, de biais.

- Une **IA ne comprend pas moralement ses actes** → la responsabilité reste portée par les concepteurs et les utilisateurs
- Il distingue **l'agentivité** (agir, produire des effets) de **l'intentionnalité morale** (vouloir, être conscient de la portée de ses actes) → une IA aurait la première sans la seconde
- Risque d'illusion de compréhension : hallucinations, décisions automatisées qui semblent réfléchies
- Plus un système est autonome et opaque, plus la chaîne de responsabilité devient difficile à tracer : qui est responsable? Le concepteur, l'utilisateur, ou l'IA elle-même ?



Luciano Floridi (1964)

Conclusion

A retenir :

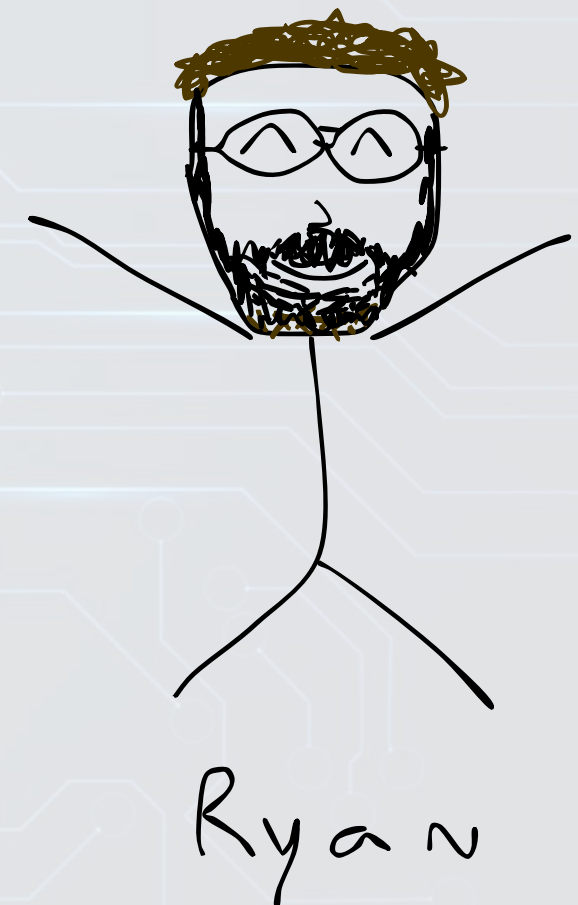
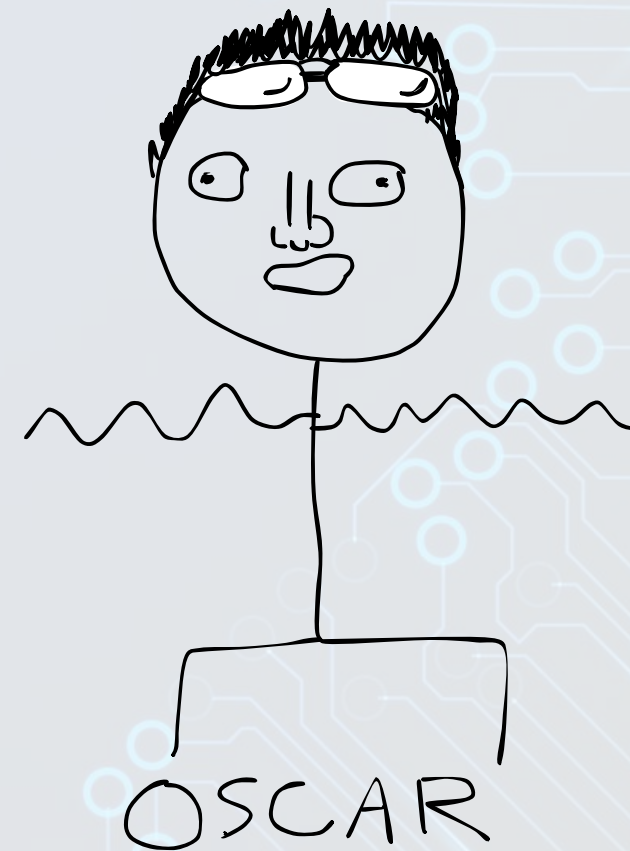
- Apparition de l'idée de la Conscience des machines
- Critique principale de Searl :La Syntaxe diffère de la sémantique
- Plusieurs objections et un désaccord sera toujours présent
- Lien avec les travaux en intelligence artificielle jusqu'à nos jours

Références:

Point de vue:philosophique:<https://plato.stanford.edu/entries/chinese-room/>

D'autres références:<https://www.ababord.org/La-chambre-chinoise>

MERCI POUR VOTRE ATTENTION!





Débat sur la chambre chinoise

Be heard, collaborate, and share ideas—make meetings and classes more engaging with real conversations.