

Commonsense Problems

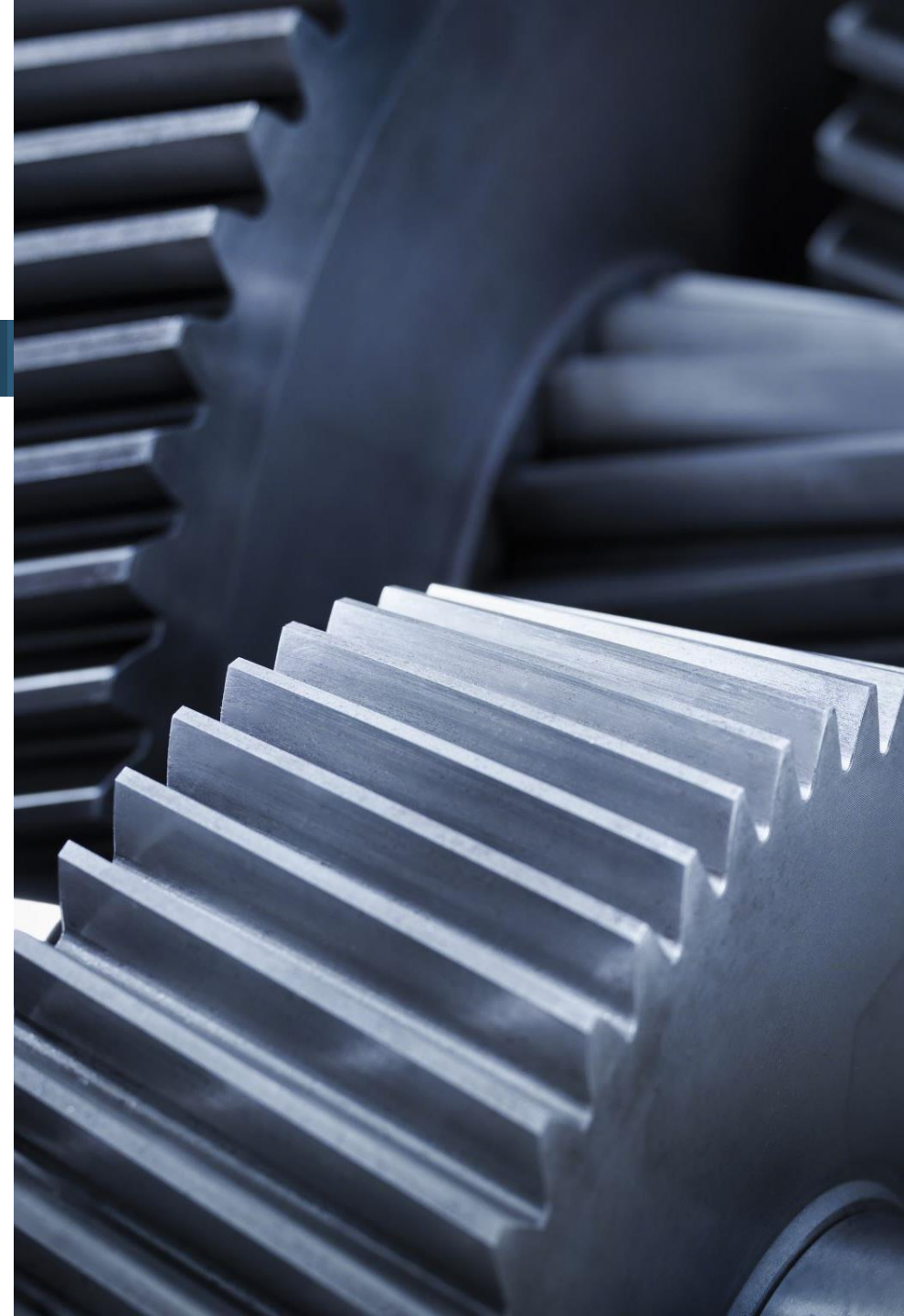
Pourquoi le bon sens
manque aux machines ?

UE IA & Cognition — Marie Lefèvre

DE OLIVEIRA Anthony

QUIBLIER Tom

CHOUATI Linda



Le bon sens... vraiment si simple ?

“Jean pose une assiette sur la table.”

“La table est renversée.”

Où est l'assiette ??

→ Et nous, savons-nous vraiment expliquer ce qu'est
avoir du bon sens ?



Qu'est ce que le bon sens = commonsense ?

- Définition du Larousse : Capacité de bien juger, sans passion, d'une situation ou de ce qu'il est raisonnable de faire, compte tenu des circonstances.

Qu'est ce que le bon sens = commonsense ?

Le bon sens est une **forme de connaissance pratique** qui ne dépend pas de connaissance pointue ni d'un raisonnement complexe. C'est ce que la plupart des gens « savent » ou « devinent » spontanément.

Tel que , le feu ça brule ou encore l'eau ça mouille.

inné ou acquis ?



Une part innée : certains réflexes ou intuitions semblent naturels, liés à nos sens et à l'évolution (ex. avoir peur du feu, éviter le vide).



Une grande part acquise : le bon sens se construit avec l'expérience, l'observation, la culture et l'éducation. Un enfant ne naît pas en sachant que « l'électricité peut tuer », il l'apprend par l'environnement.

Pourquoi le bon sens est simple chez l'Homme ?



L'évolution : Le bon sens étant primordial pour la survie d'une espèce , toutes les espèces sur terre possèdent un bon sens plus ou moins développé



Le langage : Contrairement aux autres êtres vivant le langage nous a permis d'échanger des connaissances efficacement

Pourquoi vouloir doter une IA de bon sens ?



Anticiper : Le bon sens nous permet d'anticiper avec efficacité les conséquences d'une action et de comprendre ce qui va probablement se produire ensuite.



Rapidité : Grâce au bon sens, nous réagissons instinctivement sans devoir raisonner longuement, ce qui rend nos décisions plus fluides et naturelles.



Sens et cohérence : Le bon sens nous aide à relier les informations entre elles, à comprendre le contexte et à interpréter le monde de manière logique et cohérente.



Adaptation : Face à une situation nouvelle ou imprévue, le bon sens nous permet de trouver spontanément une solution appropriée à partir de nos expériences passées.



Intention et emotion : Le bon sens nous permet de percevoir intuitivement les intentions, les émotions et les non-dits, pour réagir de manière empathique et adaptée.

Pourquoi le « common sense » est un défi pour les IA

Les IA excellent à prédire les bons mots, mais peinent à raisonner comme les humains.

Différence entre **apprentissage expérientiel humain** et **apprentissage statistique des IA**.

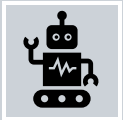
Les humains vivent et ressentent le monde, les IA n'en ont qu'une représentation abstraite, tirée de données.

Défi fondamental : relier le **monde des symboles** et le **monde de l'expérience**

Pourquoi le bon sens échappe aux IA ?



Le bon sens humain repose sur une **cognition incarnée**
(*Basarlou, 2008*)



Les IA fonctionnent sur une **cognition désincarnée** : elles traitent des symboles sans vécu sensoriel



Les **compétences implicites** échappent complètement aux modèles d'IA
(*Polanyi, 1966*)

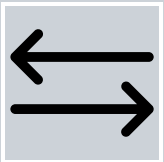


Résultat : elles produisent du langage cohérent sans véritable compréhension du monde réel.

Les trois barrières du bon sens artificiel



Absence de perception incarnée → pas d'apprentissage sensorimoteur.



Absence de raisonnement causal → apprentissage corrélatif sans compréhension des mécanismes. *(Pearl, 2019)*



Absence de contextualisation → incapacité à intégrer le cadre social, émotionnel et pragmatique. *(Tomasello, 2019)*

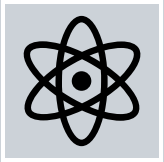


Comment doter l'IA de bon sens ?

Constat : la compréhension des IA est désincarnée, fragmentée et décontextualisée

Objectif : créer des approches capables de lui apprendre le "bon sens du monde réel".

Approche symbolique : première grande tentative pour doter l'IA de bon sens



Principe : basée sur la **manipulation de symboles** et la **représentation explicite des connaissances** du monde



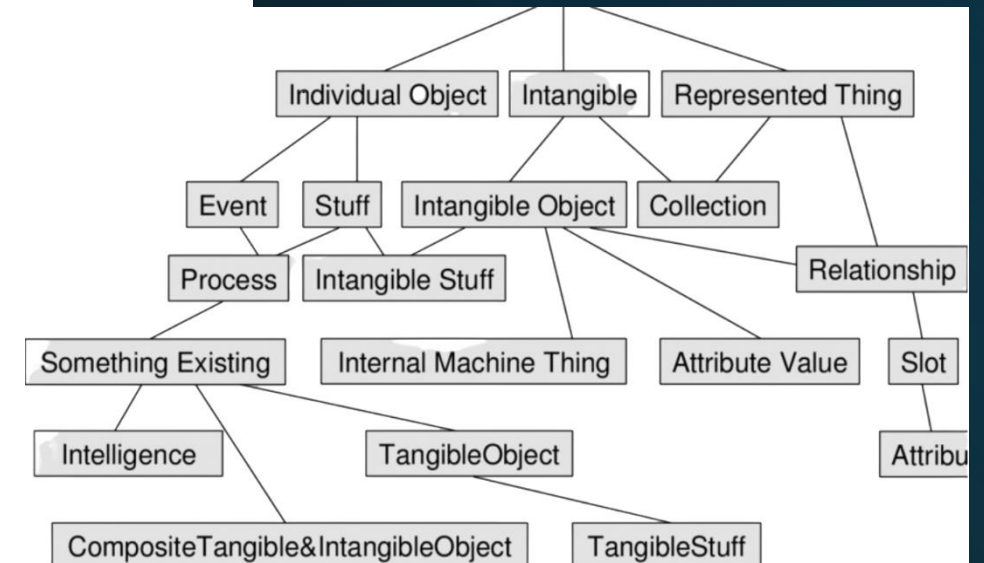
Objectif : faire **raisonner la machine** comme un humain rationnel à partir de règles logiques



suppose que **le savoir peut être entièrement formalisé**, comme un langage universel de la pensée

Le projet CYC : enseigner le bon sens à la machine

- **Effort colossal** : 10^5 concepts et 10^6 axiomes de bon sens
- Principe : construire une **base de connaissances universelle**, composée de faits et de règles
- Objectif : **permettre à la machine de raisonner** à partir de connaissances générales sur le monde
- Exemples : « Les humains ont deux jambes », « Les oiseaux volent », « Les objets tombent quand on les lâche »
- Avantage : raisonnement **explicable** et **structuré**
- Limite : impossible de tout formaliser -> **manque de contexte** et **d'expérience vécue**



Graphes de connaissances : relier les concepts pour mieux représenter le monde

Nés de l'idée de relier les **concepts entre eux** plutôt que d'écrire des règles isolées

Construisent des **réseaux de relations**, comme une **carte mentale géante** du savoir

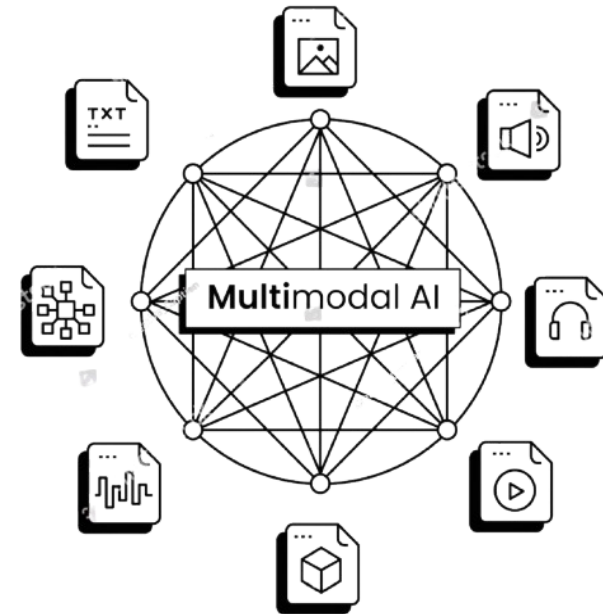
Problème : imaginent rien, généralisent pas.

Apprentissage multimodal : vers une IA qui perçoit le monde comme nous

Fusion de plusieurs modalités : texte, image, son, parfois même vidéo

Objectif : rendre la compréhension de l'IA plus contextuelle et incarnée

Limite : ces modèles associent les modalités sans construire une **représentation unifiée** du monde



Apprentissage par interaction : une IA qui apprend en agissant

- **Principe** : apprentissage **par renforcement** → l'agent agit, observe les conséquences et ajuste son comportement selon les **récompenses ou erreurs**
- **Objectif** : apprendre les **liens entre action, conséquence et contexte**, base du “**bon sens pratique**”
- **Limite** : ces IA évoluent dans des **mondes simulés**, loin de la **complexité du monde réel**
 - **apprennent à réagir**, mais pas encore à **comprendre**

Approche hybride : unir logique et apprentissage

- **Principe** : combinaison de deux paradigmes
 - **Raisonnement symbolique** (logique, explicable)
 - **Apprentissage neuronal** (adaptatif, empirique)
- **Objectif** : allier la **rigueur de la logique** à l'intuition du **machine learning**
- **Intérêt** : vers des IA capables de **prendre des décisions** et de **les expliquer**
- **Limite** : équilibre encore fragile entre **raisonnement formel** et **généralisation empirique**



L'illusion de la compréhension : les limites des approches actuelles

→ Les IA actuelles **manipulent des symboles, des images et des sons**, mais **sans expérience du monde réel**

→ **imitent** des aspects de **notre cognition**, sans parvenir à construire une compréhension globale

→ Le véritable défi : passer de l'**imitation** à la **compréhension**



Les enjeux du bon sens artificiel

- Le bon sens est essentiel pour rendre l'IA **plus fiable, interprétable et sûre.**
- Il permet une **compréhension contextuelle** du monde et des interactions humaines.
- Défi scientifique : **reproduire une cognition incarnée**, non pas seulement statistique.
- Défi sociétal : **aligner le raisonnement des IA sur nos valeurs humaines.**

Les limites actuelles

- **Surconfiance** : les IA paraissent intelligentes, mais leurs raisonnements restent superficiels.
- **Manque de transparence** : raisonnement interne souvent opaque (*black box*).
- **Biais de données** : le “bon sens” appris reflète nos stéréotypes culturels.
- **Absence d'intentionnalité** : pas de compréhension réelle des buts ou des valeurs humaines.

Applications du bon sens artificiel

- **Robotique sociale** : robots capables d'interpréter les situations humaines du quotidien.
- **Santé** : IA comprenant le contexte émotionnel ou comportemental d'un patient.
- **Éducation** : assistants pédagogiques adaptant leur discours au niveau de compréhension.
- **Sécurité** : IA capable d'anticiper les comportements anormaux dans un environnement.

Conclusion

- Le bon sens a un immense potentiel dans divers domaines et apporterait plusieurs solutions à des problèmes compliqués mais c'est un domaine très complexe que l'on ne comprend pas encore totalement chez l'Homme
- Avant de pouvoir l'intégrer pleinement dans une machine, il faut d'abord apprendre à mieux comprendre notre propre façon de raisonner, de ressentir et d'agir avec bon sens.

Références

- Barsalou, L. (2008). Grounded Cognition.
- Spelke, E.S., & Kinzler, K.D. (2007). Core Knowledge.
- Polanyi, M. (1966). The Tacit Dimension.
- Lenat, D.B. (1995). CYC: A Large-Scale Investment in Knowledge Infrastructure.
- Speer et al. (2017). ConceptNet 5.5.
- Sap et al. (2019). ATOMIC: An Atlas of Machine Commonsense
- Pearl, J. (2019). The Book of Why: The New Science of Cause and Effect.
- Tomasello, M. (2019). Becoming Human: A Theory of Ontogeny.
- Bian et al. (2023). ChatGPT is a Knowledgeable but Inexperienced Solver.

Références

- Harnad, S. (1990). The Symbol Grounding Problem.
- Davis, E. (2023) Benchmarks for Automated Commonsense Reasoning: A Survey.
- Nawaz et al. (2025). A review of neuro-symbolic AI
- Marcus, G. (2018). Deep Learning: A Critical Appraisal
- Lake, B.M. (2017). Building machines that learn and think like people
- Tenenbaum, J.B. (2011). The building blocks of cognition
- OpenAi (2023). GPT-4 Technical Report (section multimodale).

Réflexions

- Peut-on formaliser scientifiquement le bon sens, ou est-ce une notion trop humaine ?
- Le bon sens dépend-il plus de la logique, de l'expérience, ou d'une combinaison des deux ?
- Pensez-vous qu'on puisse parler de "cognition artificielle" si l'IA ne possède pas de corps ni de sens?
- Faut-il vraiment chercher à doter les IA de bon sens ?
- Si une IA développait du bon sens proche du nôtre, serait-elle encore un outil ou deviendrait-elle un partenaire ?