

Atelier 10 : Cognition incarnée

Présenté par : COCHENNEC Fanny, DEBOTTE Gabriel, MIALI Matteo, RUOT Eliott

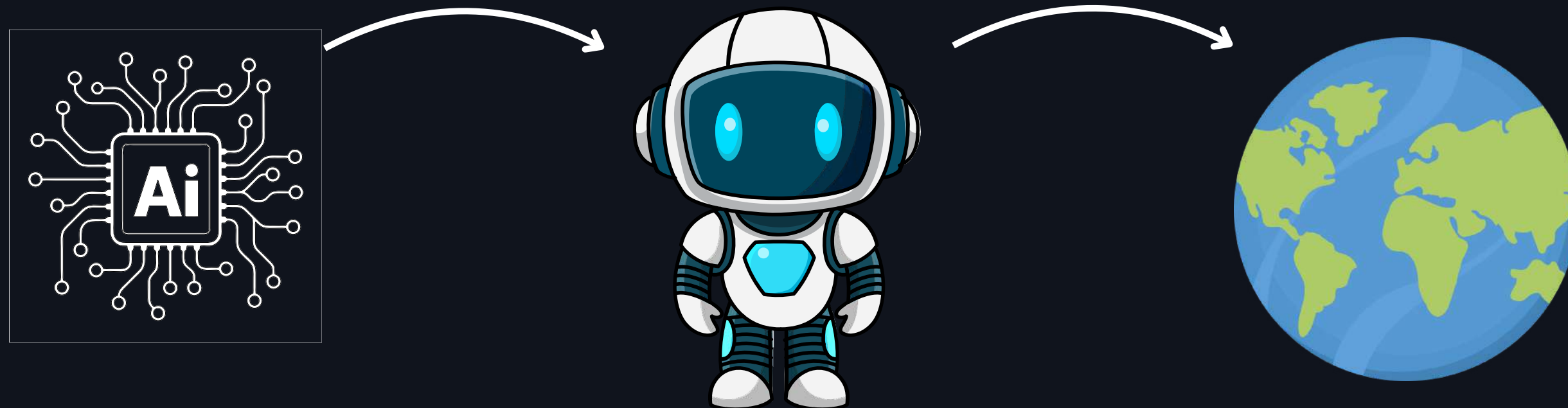
UE : IA et Cognition - Marie Lefevre

14 Octobre 2026

Introduction : Qu'est ce que la cognition incarnée ?

La **cognition incarnée** s'oppose à l'idée que **penser se réduit à manipuler des symboles abstraits** dans le cerveau, **indépendamment du corps**.

Elle soutient que la **pensée dépend du corps** et de ses **interactions avec l'environnement** : nos capacités sensorielles et motrices façonnent la manière dont nous percevons, raisonnons et formons nos concepts.



Introduction : Qu'est ce que la cognition incarnée ?

Couplage structurel :

L'agent est incrusté dans son **environnement** via une **boucle d'action**.

Référence : Maturana et Varela, 1987

Incarnation historique :

Le **corps** et le **comportement** de l'agent sont le résultat de **plusieurs interactions**.
Cela peut être considéré comme une évolution biologique.

Référence : Varela et coll., 1991

Incarnation physique :

L'agent doit exister sous forme de **robot/machine physique** dans le **monde réel** et non comme une simulation.

Référence : Brooks, 1991

Incarnation de type organismoïde :

Le corps du système est artificiel mais il imite la morphologie, le sens ou les structures motrices d'un organisme vivant

Référence : Damasio, 1999

La vision de Ziemke

- **Couplage structurel** : Condition nécessaire mais non suffisante
- **Incarnation historique** : Notion d'histoire considérée comme artificiel
- **Incarnation physique** : Le corps ne doit pas être un simple conteneur/interface d'entrée-sortie
- **Incarnation organismoïde** : Illusion sur le plan conceptuel

Introduction : Qu'est ce que la cognition incarnée ?

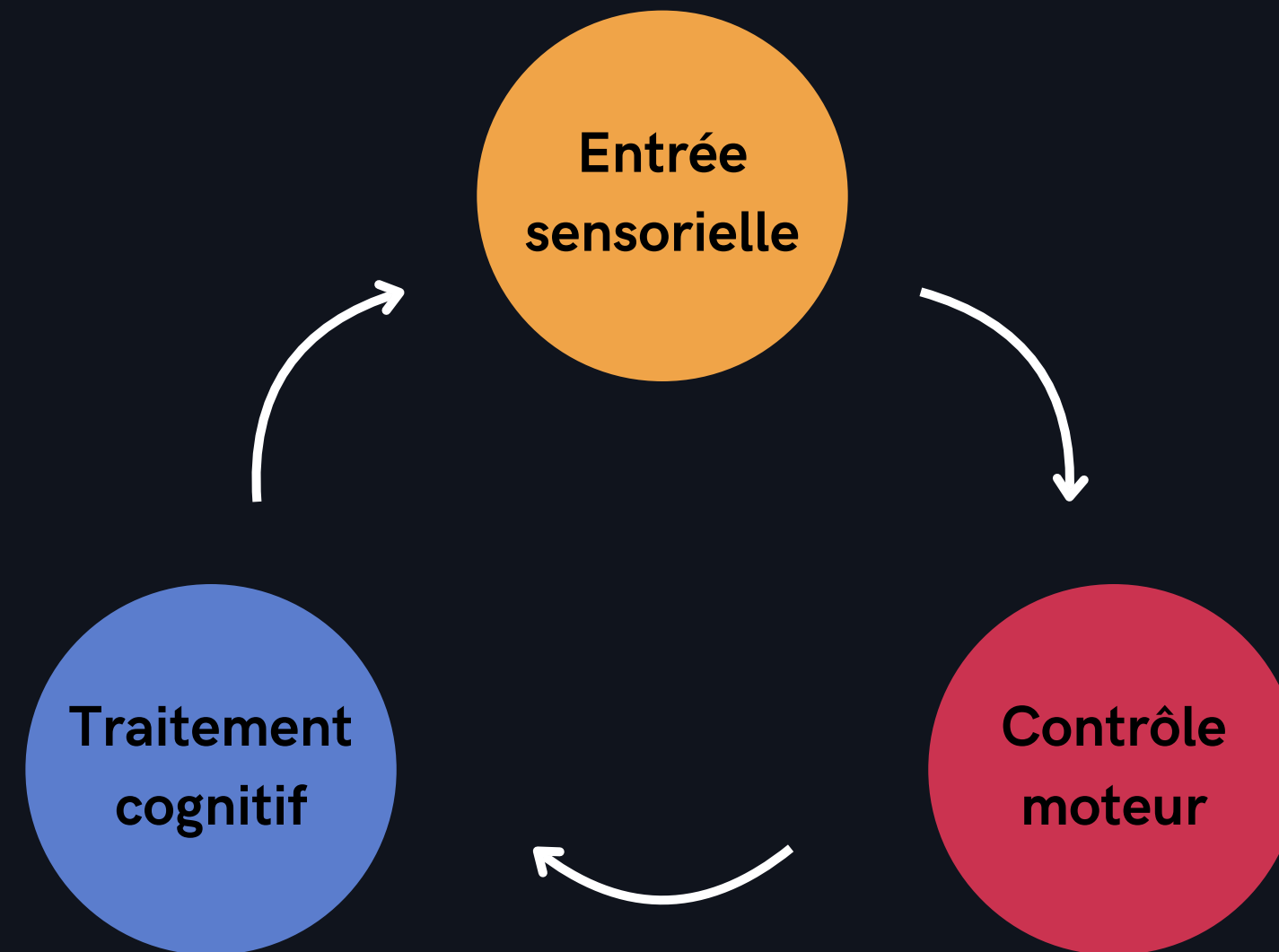
La **cognition incarnée** s'oppose à l'idée que **penser se réduit à manipuler des symboles abstraits** dans le cerveau, **indépendamment du corps**.

Elle soutient que la **pensée dépend du corps** et de ses **interactions avec l'environnement** : nos capacités sensorielles et motrices façonnent la manière dont nous percevons, raisonnons et formons nos concepts.

Suffit-il de mettre une IA dans un robot pour qu'elle soit incarnée ?

1. Un corps pour agir - Long et al.

Idée : Le corps est une **interface capteurs-moteurs** → Besoin d'un corps pour **interagir** avec son **environnement**



Autonomie et adaptabilité

1. Un corps pour agir - Long et al.

Classification des robots intelligents en fonction de :

- l'autonomie
- la capacité de gestion des tâches
- la capacité d'adaptation à l'environnement
- la capacité de cognition sociale

IR-L0



Pas de perception,
seulement exécution

IR-L1



Capteurs simples,
règles fixes

IR-L2



Perception de
l'environnement et
adaptation

IR-L3



Compréhension des
intentions et émotions
humaines

IR-L4



Autonomie totale

1. Un corps pour agir - Long et al.

Techniques robotiques :

➤ Méthodes classiques : modèles mathématiques et physiques

- Model Predictive Control (MPC)
- Whole Body Control (WBC)

Lien perception - action : ingénieur écrit les équations physiques

➤ Apprentissage par renforcement : récompenses/punitions

Lien perception - action : robot par essai/erreur

➤ Imitation : démonstration

Lien perception - action : robot en copiant les humains ou d'autres robots

➤ Modèle VLA (Vision - Langage - Action)

Lien perception - action : grand modèle qui relie la vision, le langage et l'action

1. Un corps pour agir - Long et al.

Techniques robotiques :

➤ Méthodes classiques : modèles mathématiques et physiques

- Model Predictive Control (MPC)
- Whole Body Control (WBC)

Lien perception - action : ingénieur écrit les équations physiques

Robot Cassie (Xie et al., 2019)

➤ **Apprentissage par renforcement : récompenses/punitions**

Lien perception - action : robot par essai/erreur

➤ Imitation : démonstration

Lien perception - action : robot en copiant les humains ou d'autres robots

➤ Modèle VLA (Vision - Langage - Action)

Lien perception - action : grand modèle qui relie la vision, le langage et l'action

1. Un corps pour agir - Long et al.

Techniques robotiques :

➤ Méthodes classiques : modèles mathématiques et physiques

- Model Predictive Control (MPC)
- Whole Body Control (WBC)

Lien perception - action : ingénieur écrit les équations physiques

ALOHA (Zhao et al.,2024), Stanford University

➤ Apprentissage par renforcement : récompenses/punitions

Lien perception - action : robot par essai/erreur

➤ **Imitation : démonstration**

Lien perception - action : robot en copiant les humains ou d'autres robots

➤ Modèle VLA (Vision - Langage - Action)

Lien perception - action : grand modèle qui relie la vision, le langage et l'action

1. Un corps pour agir - Long et al.

Techniques robotiques :

➤ Méthodes classiques : modèles mathématiques et physiques

- Model Predictive Control (MPC)
- Whole Body Control (WBC)

Lien perception - action : ingénieur écrit les équations physiques

RT2 (Chebotar et Yu, 2023, Deepmind)

➤ Apprentissage par renforcement : récompenses/punitions

Lien perception - action : robot par essai/erreur

➤ Imitation : démonstration

Lien perception - action : robot en copiant les humains ou d'autres robots

➤ **Modèle VLA (Vision - Langage - Action)**

Lien perception - action : grand modèle qui relie la vision, le langage et l'action

1. Un corps pour agir - Long et al.

Problèmes des données :

- Données sur de vrais robots : coûteuses et risquées (téléopération, expérience dangereuses)
- Peu reproductibles

Solution : Sim2Real !

- Principe : "on s'entraîne en simulation puis on transfère au vrai robot"

Pourquoi simuler ?

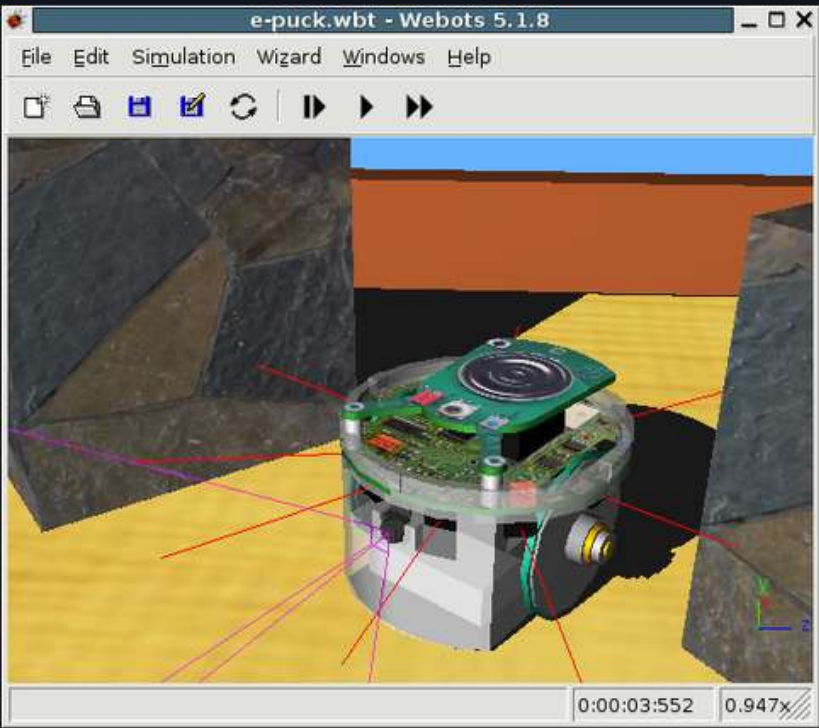
- Vitesse : des milliers d'environnements en parallèle sur GPU
- Annotations parfaites : profondeur, segmentation, position des objets, sans annotation manuelle
- Banc d'essai sûr et reproductible



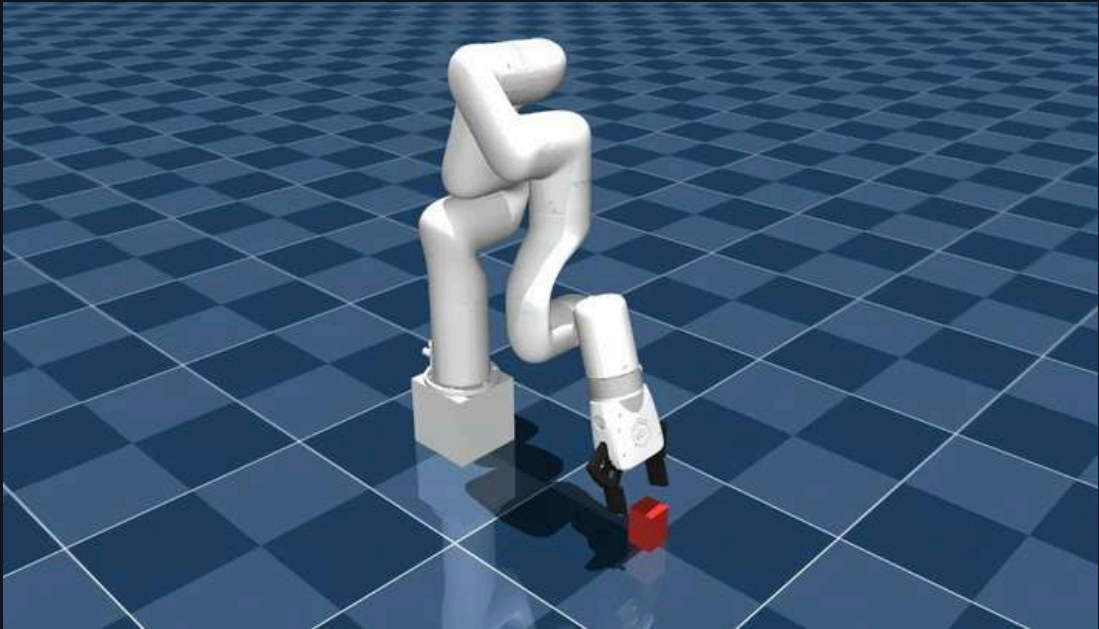
1. Un corps pour agir - Long et al.

Simulateur	Point fort	Limite
Webots, Gazebo, CoppeliaSim 1998	simples	physique avancée limitée
MuJoCo 2012	contacts très précis	rendu basique
PyBullet 2017	léger, facile en Python	fidélité moindre
Isaac Gym, Sim, Lab 2021	GPU massif, rendu réaliste	matériel exigeant
SAPIEN 2020	objets articulés	pas de fluides ni de LiDAR
Genesis 2024	multi-physique, très rapides	encore jeunes

Exemple: Webots



Exemple: Mujoco



1. Un corps pour agir - Long et al.

Bilan des simulateurs:

- Avantages :
 - le faible coût
 - la sécurité
 - le contrôle des variables
 - la répétabilité
- Limites :
 - Précision: écart avec la réalité
 - Complexité: modéliser fidèlement → très coûteux
 - Dépendance aux données: données de qualité nécessaires
 - Sur-spécialisation: trop adapté à certains scénarios

Besoin d'un outil plus flexible, appris à partir des données...

1. Un corps pour agir – Long et al.

Les WORLD MODELS

Définition : “Un modèle d’IA qui apprend comment le monde évolue et peut prédire ce qui va se passer”

L’intuition (Exemple de si on lache un verre)

Ha et SCHMIDHUBER, 2018 : Une IA apprenant une version compact de son environnement, et s’entraîne dans son imagination.

Yann LeCun : une IA doit app comme humains, par perception visuel (architecture V-JEPA)

1. Un corps pour agir - Long et al.

Les architectures :

Famille	Idée	Exemples
RSSM	compresse environnement -> simule évolution du monde	Dreamer
JEPA	prédit des représentations abstraites, pas des pixels	I-JEPA, V-JEPA
Transformers	cohérence sur de longues séquences	TransDreamer, Genie
Autoregressif	fonctionnent comme un LLM : ils prédisent le « mot » suivant	GAIA-1
Diffusion	débruitage progressif, meilleure qualité mais lent	Sora, Veo 3, Vista

1. Un corps pour agir – Long et al.

3 grands rôles d'un WORLD MODEL :

- Simulateur neuronal : Génère des scènes réalistes et contrôlables (mais apprend grâce à des données donc plus réaliste)
- Modèle dynamique. L'agent l'utilise pour imaginer l'avenir : « si je fais telle action, que se passe-t-il ? »
- Modèle de récompense: Plus le comportement de l'agent est prévisible pour ce modèle, plus il est récompensé.

Et qu'en est-il de l'application dans le monde réel..?

1. Un corps pour agir - Long et al.

Applications et Défis:

Deux grand domaines:

Conduite autonome



Les robots articulés



1. Un corps pour agir – Long et al.

Conduite autonome :

Problème : pipeline modulaire (perception → prédiction → planification → contrôle), les erreurs s'accumulent..

Exemple : accident Uber de 2018 : le système avait détecté le piéton plusieurs secondes avant l'impact sans l'identifier correctement.

Simulateurs neuronaux: GAIA-1 et GAIA-2, DriveDreamer ou MagicDrive génèrent des scènes de conduite variées, avec contrôle de la météo

Modèles dynamiques: MILE, apprend en même temps à prédire le monde et à conduire

Modèles de récompense: Vista ou Drive-WM, imaginent plusieurs futurs possibles et évaluent leur sécurité.



1. Un corps pour agir - Long et al.

Robots articulés:

Simulateurs neuronaux: Cosmos, WHALE et RoboDreamer

Modèles dynamiques: Dreamer

Modèle de récompense: VIPER

Pistes d'avenirs : intégré le touché, unifiés valables pour plusieurs robot, modele hiérarchique

1. Un corps pour agir – Long et al.

Défis pour le futur:

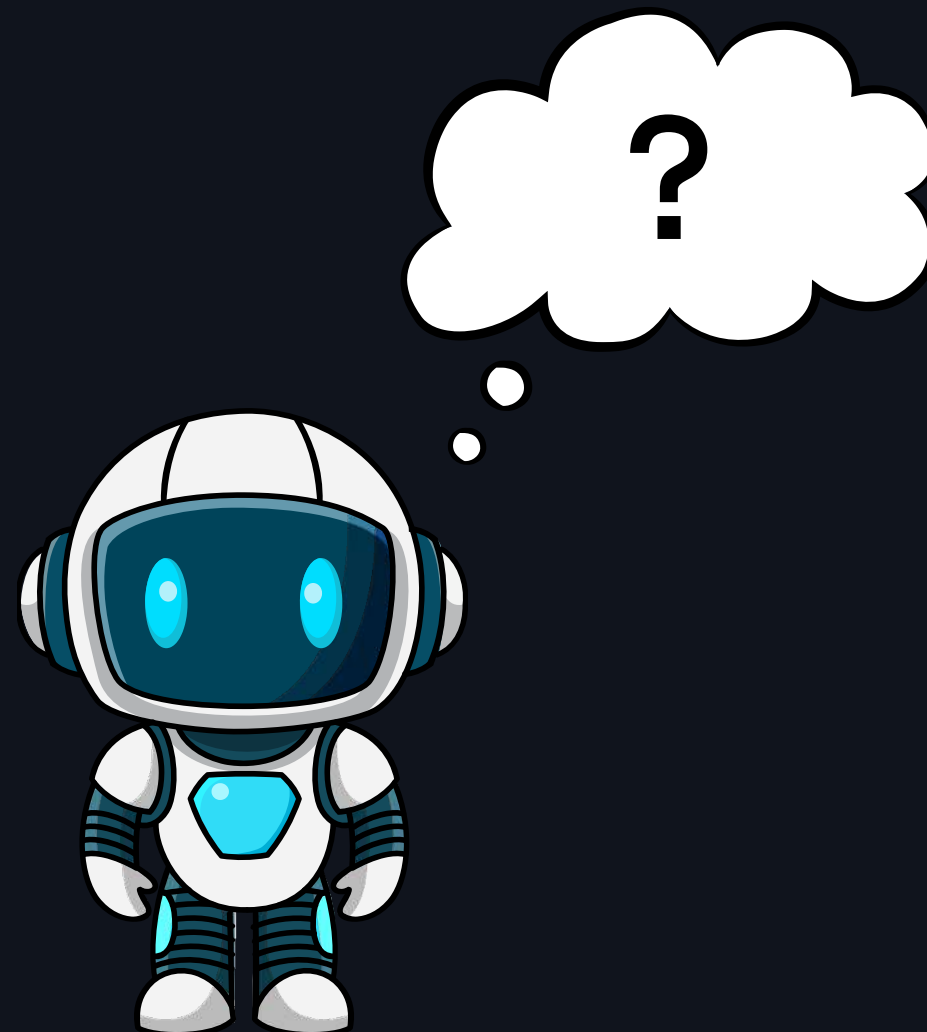
- Causalité contre corrélation : comprendre le pourquoi pour raisonner correctement
- Évaluation : Une image nette suffit pas, il faut savoir identifié
- Mémoire à long terme : le modèle doit se souvenir (par exemple d'un panneau vu plusieurs minutes avant)
- Interprétabilité et sécurité. Ces modèles sont des boîtes noires, or il faut pouvoir analyser un accident et idéalement garantir qu'ils n'inventent pas d'obstacle dangereux
- Données et biais. Un modèle entraîné dans une région peut mal fonctionner ailleurs

1. Un corps pour agir - Long et al.

Idée : Le corps est une **interface capteurs-moteurs** → Besoin d'un corps pour **interagir** avec son **environnement**

Limites :

- Apprentissage de corrélations, pas de causes
- Difficultés à intégrer les concepts abstraits



2. Le corps jusque dans la pensée abstraite



SOUVENT EN COGNITION INCARNÉE

Le corps en action

On pense avec ses sens et ses gestes : percevoir, bouger, agir dans un environnement.



Le corps dans l'abstrait

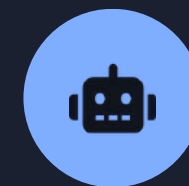
Même des idées sans lien avec le corps (respecter, l'avenir, une relation) seraient organisées par des formes tirées du corps.

DEUX QUESTIONS, DEUX ARTICLES



Ces formes existent-elles vraiment chez des gens ordinaires ?

Richardson, Spivey, Edelman & Naples (2001)
psychologie expérimentale



Peut-on les donner à une machine qui n'a pas de corps ?

Olivier & Bouraoui (2025)
intelligence artificielle

2. Le corps jusque dans la pensée abstraite

Le schéma d'image : une forme simple tirée du corps

Un motif récurrent, abstrait de nos interactions sensorimotrices avec le monde (Johnson, 1987).

CONTENANT

un dedans et un dehors

« être dans une relation »

HAUT - BAS

la verticalité

« se sentir au plus bas »

SOURCE - CHEMIN - BUT

un trajet orienté

« l'avenir est devant nous »



Tiré de la vidéo : les mots abstraits viennent se loger dans les schémas concrets.

Métaphore conceptuelle : on comprend l'abstrait (domaine cible) à travers le concret (domaine source).

2. Le corps jusque dans la pensée abstraite

« LANGUAGE IS SPATIAL » · RICHARDSON ET AL. · 2001

2001 : les gens ordinaires partagent-ils ces schémas ?

173 étudiants lisent des phrases-rébus et choisissent, pour chacune, l'image qui la représente le mieux.

Exemple :  *lifted*  = « le cercle a soulevé le carré » → **On attend l'image 1**

agit

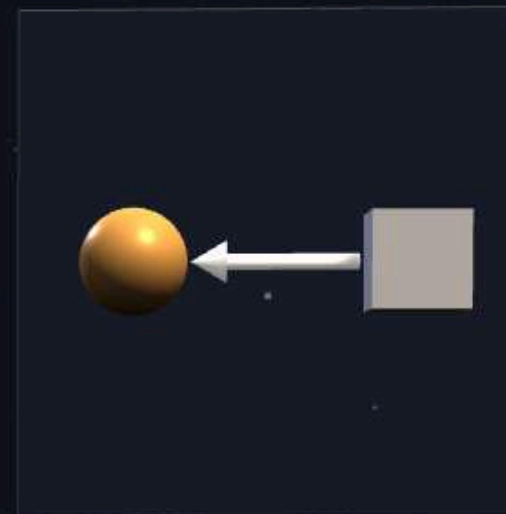
subit



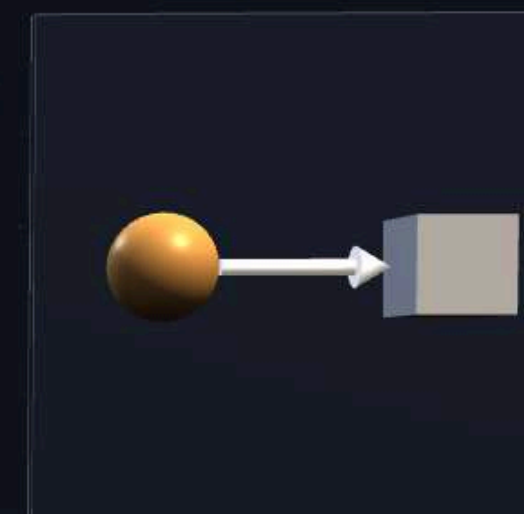
1 · Haut



2 · Bas



3 · Gauche



4 · Droite

2. Le corps jusque dans la pensée abstraite

« LANGUAGE IS SPATIAL » · RICHARDSON ET AL. · 2001

2001 : les gens ordinaires partagent-ils ces schémas ?

173 étudiants lisent des phrases-rébus et choisissent, pour chacune, l'image qui la représente le mieux.

Exemple :  *lifted*  = « le cercle a soulevé le carré » → **On attend l'image 1**

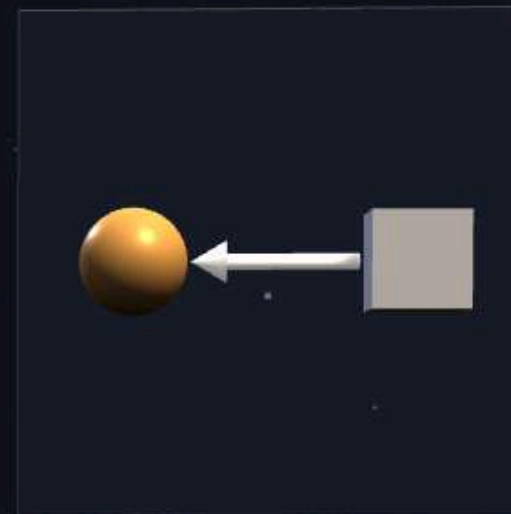
agit subit



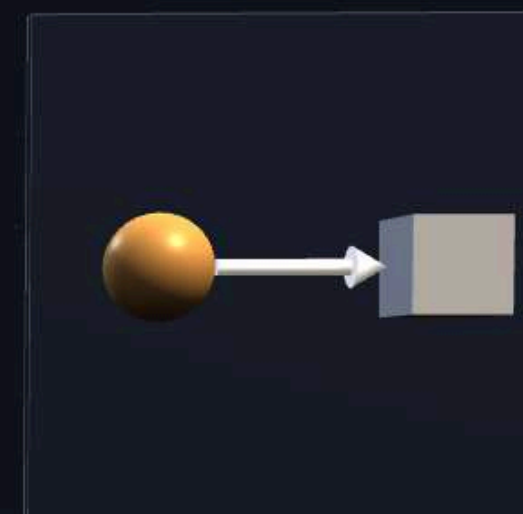
1 · Haut



2 · Bas



3 · Gauche



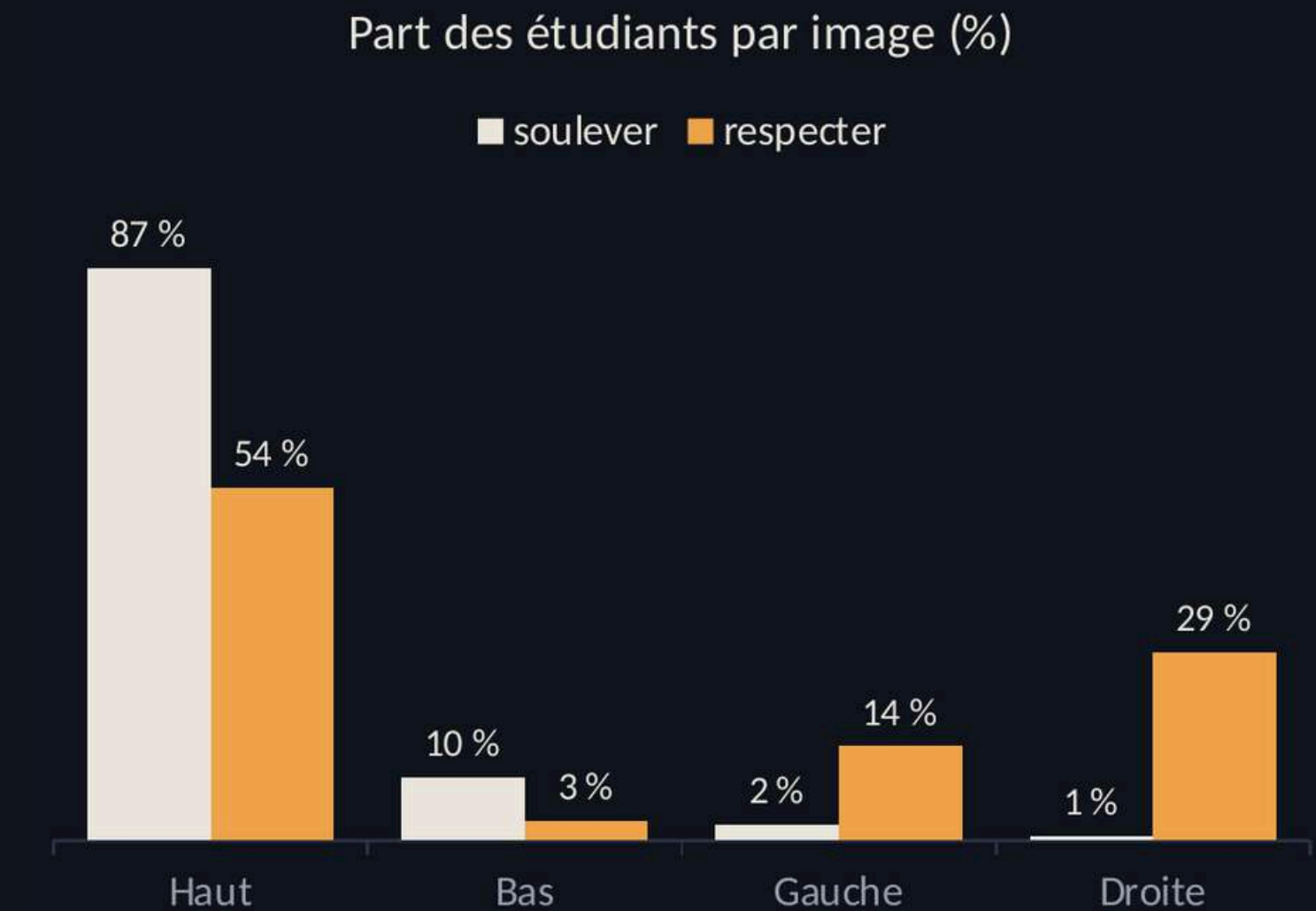
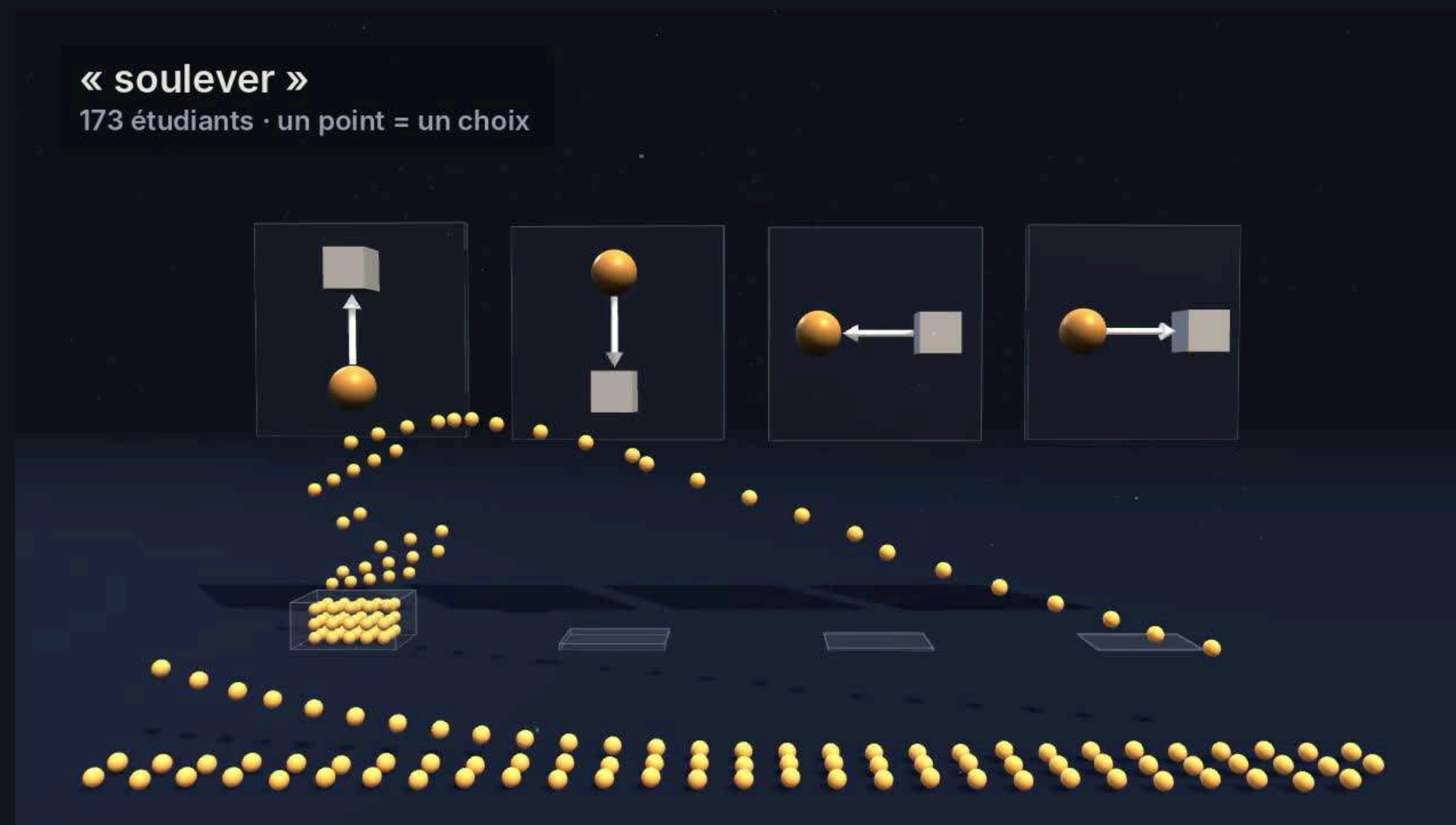
4 · Droite

*À vous :
et pour
« respecter » ?*

2. Le corps jusque dans la pensée abstraite

« LANGUAGE IS SPATIAL » · EXPÉRIENCE 1

Résultat : un accord net, même pour l'abstrait



Données : tableau 1 de Richardson et al. (2001).

25 % par image si l'on choisissait au hasard

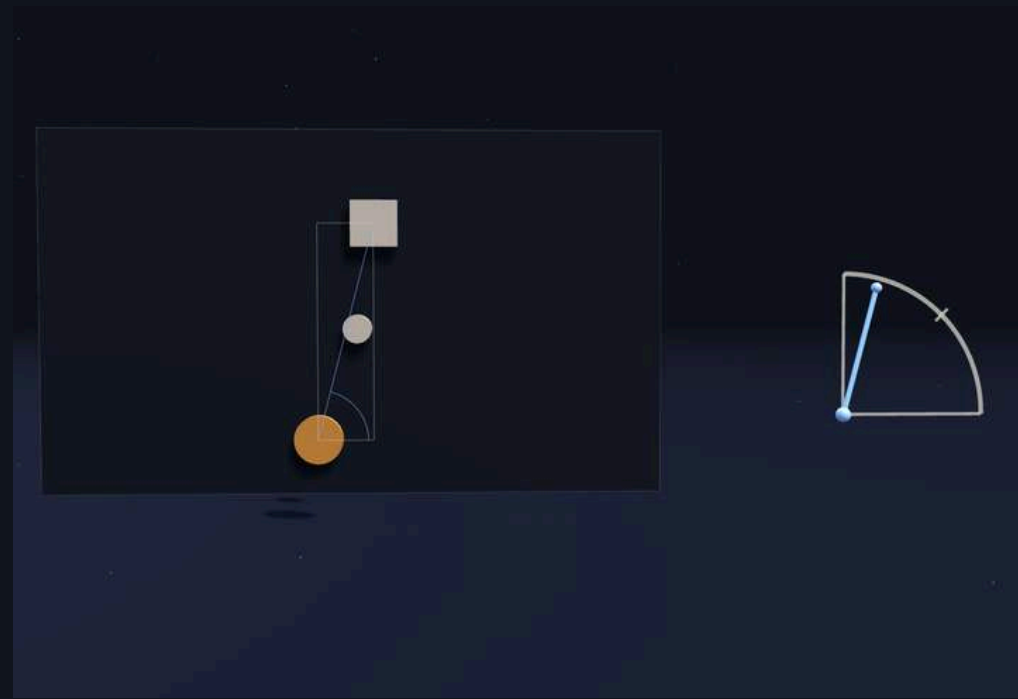
63 % pour l'image préférée d'un verbe, en moyenne

Effet très fort de l'axe attendu

2. Le corps jusque dans la pensée abstraite

« LANGUAGE IS SPATIAL » · EXPÉRIENCE 2

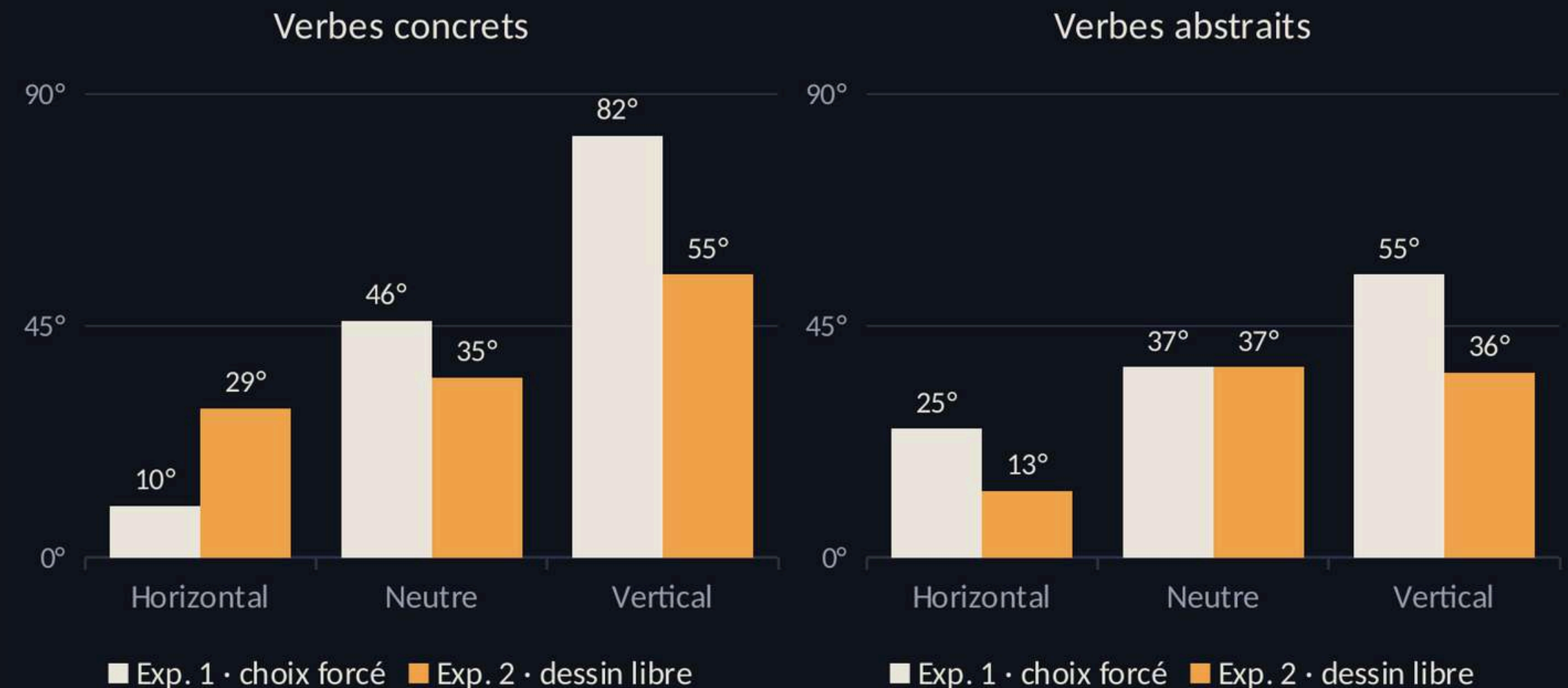
Le dessin libre confirme, mais en plus flou



24 autres étudiants dessinent librement. On mesure l'inclinaison du dessin : 0° couché, 90° debout.

R = 0,71

corrélacion entre les deux expériences



Angle moyen selon l'axe attendu par les linguistes (0° = horizontal, 90° = vertical). Tableau 2 et figure 5 de l'article.

À nuancer : les étudiants avaient le temps de réfléchir · les flèches les poussaient déjà vers l'espace · tous anglophones, d'une seule université



« GROUNDING AGENT REASONING » · 2025

Des machines qui apprennent le langage sans corps

Olivier & Bouraoui · CRIL (CNRS & Université d'Artois), Lens · AAMAS 2025, piste « Blue Sky Ideas »

Le constat

Les grands modèles de langage apprennent à partir de **textes seulement**.

Le doute

Cela suffit-il pour comprendre l'espace et raisonner avec bon sens ? (Bender & Koller, 2020 ; Tamari et al., 2020)

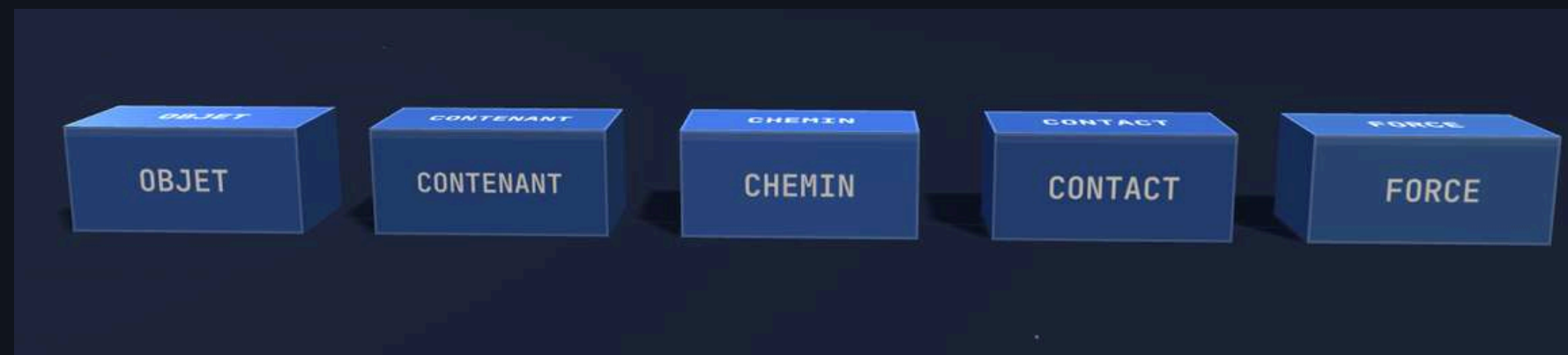
La proposition

Donner aux agents **nos schémas d'images**, sous une forme logique exacte : une **IA neurosymbolique**.

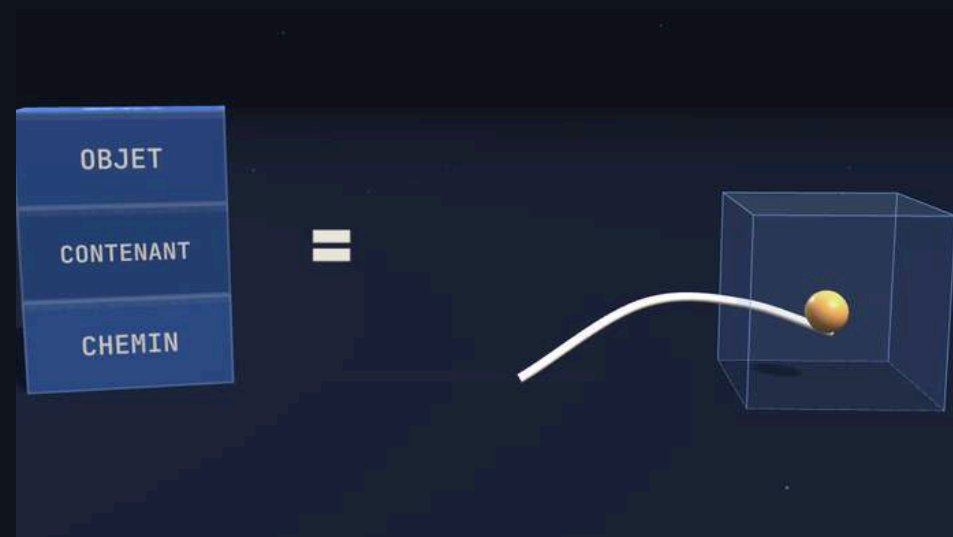
2. Le corps jusque dans la pensée abstraite

« GROUNDING AGENT REASONING » · LA FORMALISATION

Des briques élémentaires et des règles par défaut



Primitives (d'après Hedblom et al., 2024) : entités, relations, attributs, forces.



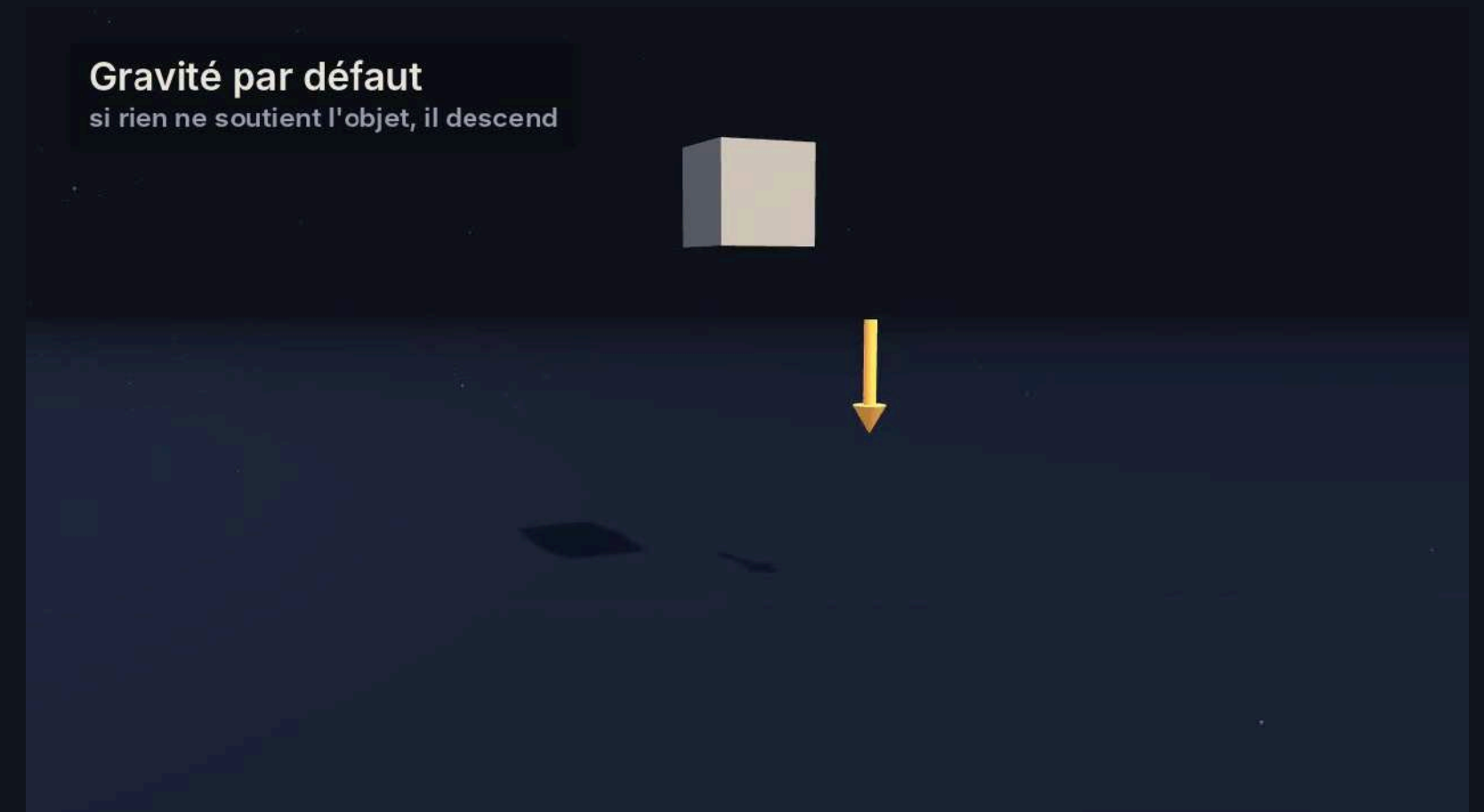
ENTRER

= objet + contenant + chemin

Chaque schéma devient une petite théorie logique Γ ; chacun de ses modèles est une scène possible.

Gravité par défaut

si rien ne soutient l'objet, il descend



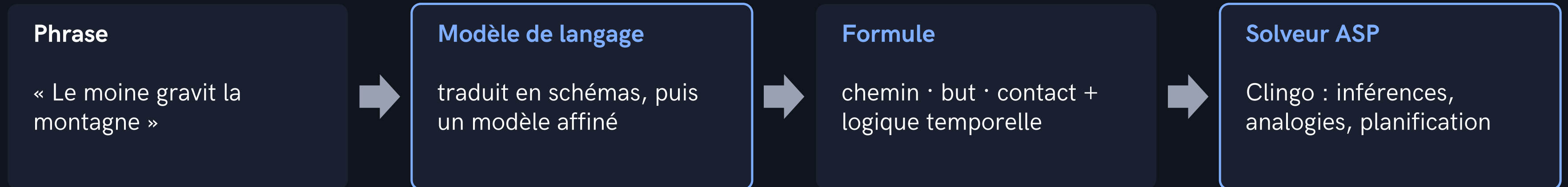
$$\square \forall x (\neg \exists y \text{ sur}(x, y) \rightarrow \text{descend}(x))$$

Toujours : si rien ne soutient x , x descend. Une table apprise ensuite $\rightarrow$ la conclusion est retirée (raisonnement non monotone).

2. Le corps jusque dans la pensée abstraite

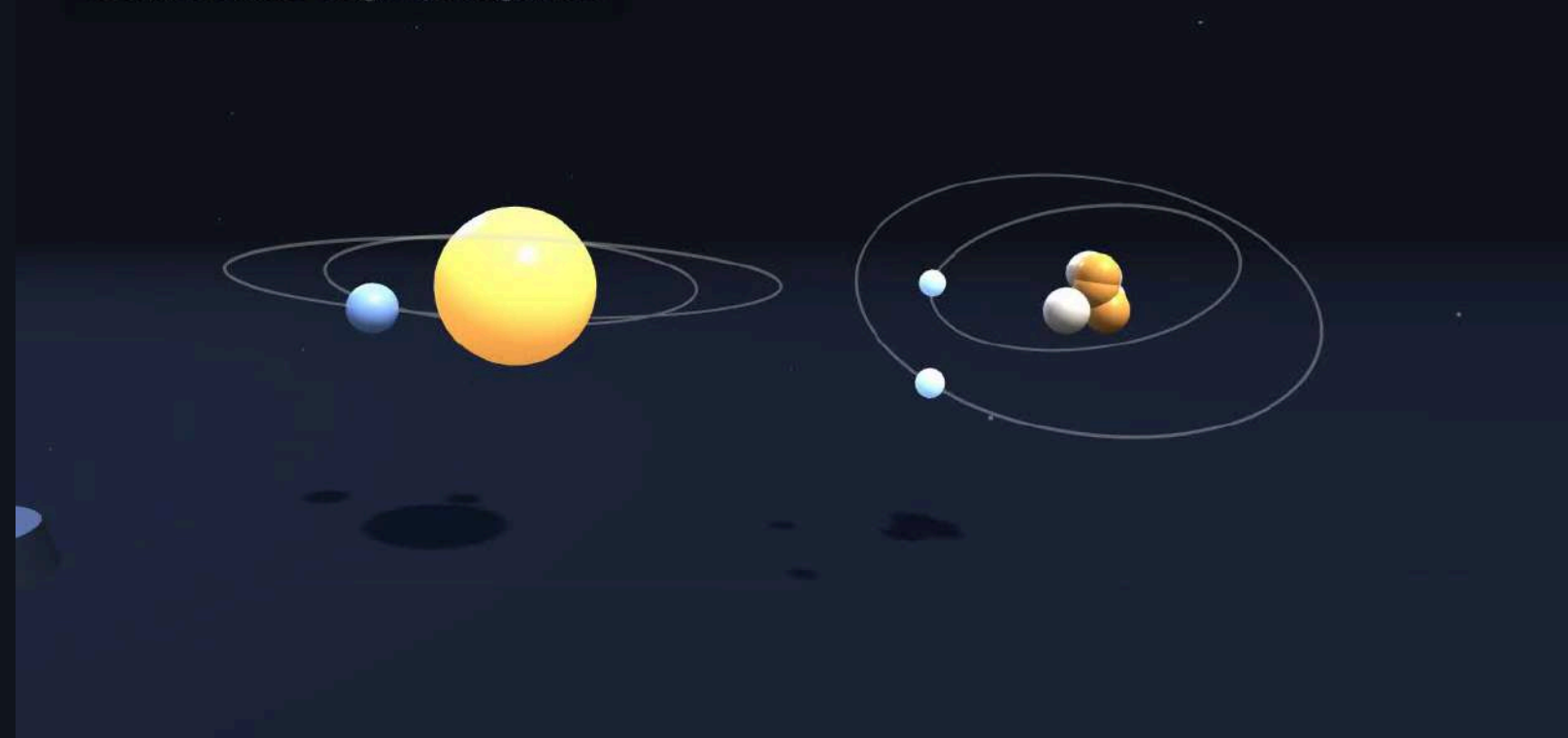
« GROUNDING AGENT REASONING » · L'ARCHITECTURE

De la phrase au raisonnement



Même structure

distance stable · angle qui augmente



Les analogies

Planètes et électrons partagent une structure : une distance bornée, un angle qui augmente.

$$\Delta(x, y) \text{ borné} \wedge \theta(x, y) < \circ \theta(x, y)$$

Statut : article « Blue Sky »

Un programme de recherche : pas d'implémentation, pas de données, pas d'évaluation.

2. Le corps jusque dans la pensée abstraite

Ce que les schémas d'images apportent :

- **Le schéma d'image**

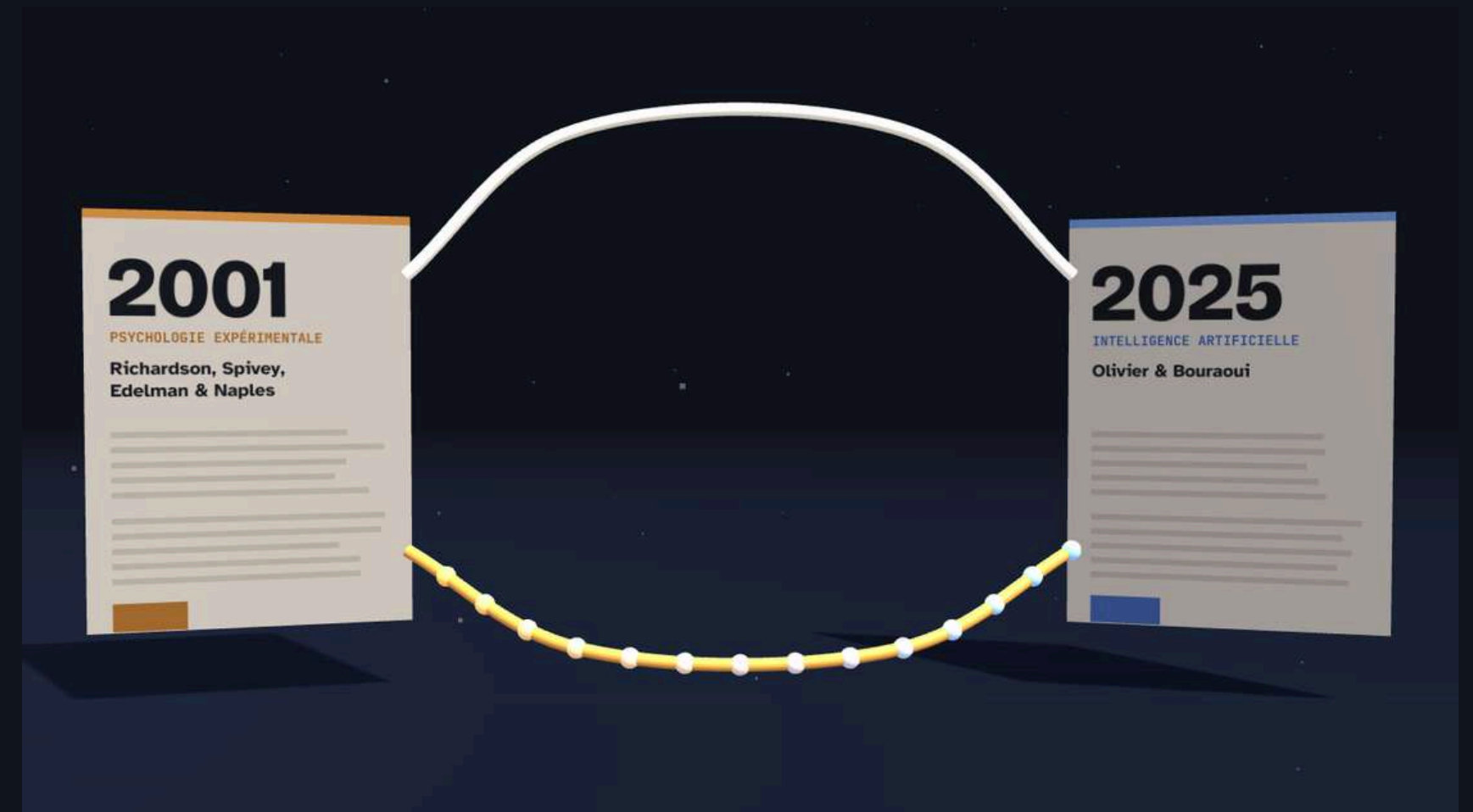
une forme simple tirée du corps : contenant, haut-bas, chemin

- **2001 · mesurer chez l'humain**

des gens ordinaires partagent ces schémas, un peu moins nettement pour l'abstrait

- **2025 · reproduire dans la machine**

les écrire en logique pour un agent sans corps : une piste, pas encore testée



Le **corps** ne sert **pas** qu'à **percevoir** et à **agir** : il laisse des formes qui structurent aussi la **pensée abstraite**.

2. Le corps jusque dans la pensée abstraite

Ce que les schémas d'images apportent :

- **Le schéma d'image**

une forme simple tirée du corps : contenant, haut-bas, chemin

- **2001 · mesurer chez l'humain**

des gens ordinaires partagent ces schémas, un peu moins nettement pour l'abstrait

- **2025 · reproduire dans la machine**

les écrire en logique pour un agent sans corps : une piste, pas encore testée

Limite :

Donner au robot nos **schémas d'images** sous forme de formules, c'est lui fournir des **symboles** reliés à l'espace, mais qu'il n'a pas **vécus**.

Chez l'humain, ces schémas naissent de **l'expérience du corps**.

→ Faut-il un **corps vivant** pour les **comprendre** vraiment ?

Le **corps** ne sert **pas qu'à percevoir et à agir** : il laisse des formes qui structurent aussi la **pensée abstraite**.

3. Le rôle du corps vivant

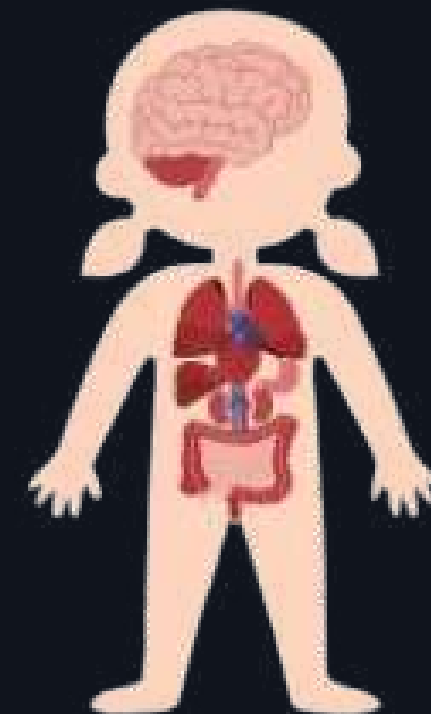
Le corps mécanique

Simple **interface d'entrée/sortie**. Conteneur passif contrôlé par un algorithme externe.



Le corps vivant

Système biologique autonome luttant continuellement pour **sa propre survie**.



3. Le rôle du corps vivant

Autopoïèse : du grec auto, « soi-même », et poïësis, « production » ou « création »

L'autopoïèse

L'organisme **se produit** et **se régénère** en permanence face à la dégradation de son environnement.

Impératif de survie

Mettre l'enjeu vital de l'agent au centre de la réflexion. En cas d'**échec de maintien biologique** cela entraîne la **mort de l'organisme**.

Importance du sens

Le **monde** acquiert une **valeur immédiate** fondée sur le **besoin du corps vivant**.

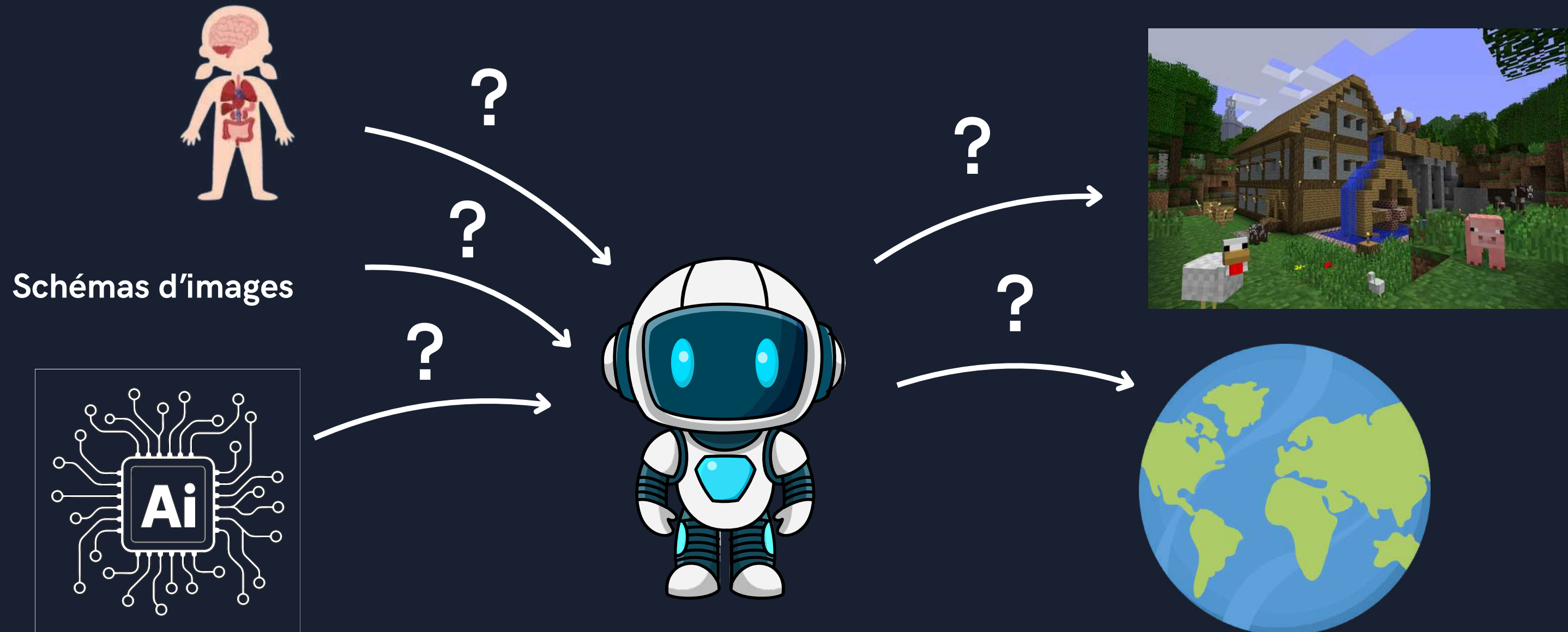
Sans besoin de **préservation biologique**, les symboles et les actions restent **dépourvus de signification propre à l'agent**

Conclusion

Cognition incarnée



Concept toujours débattu
Idée générale : **Besoin d'un corps** pour **interagir** avec son **environnement** et développé un **comportement intelligent**



Débat

Un robot est-il vraiment incarné ?

Débat

Un robot est-il vraiment incarné ?

L'incarnation physique suffit

Dès qu'un agent **interagit dans le monde réel** avec un **corps** (capteurs/actionneurs), il est **incarné**.

Rodney Brooks, Rolf Pfeifer, Christian Scheier, Luc Steels

Corps sans incarnation

Système biologique autonome luttant continuellement pour **sa propre survie**.

Tom Ziemke, Francisco Varela, Evan Thompson, Humberto Maturana, John Searle

Débat

Un robot entraîné presque entièrement en simulation est-il incarné ?

Débat

Un robot entraîné presque entièrement en simulation est-il incarné ?

Possible

L'incarnation repose sur les **boucles sensori-motrices** et la dynamique de **l'interaction**. Si la **simulation modélise fidèlement** la physique, les structures cognitives développées sont authentiquement incarnées.

Josh Bongard, Xiaoxiao Long (*Sim-to-Real*)

Abstraction mathématique

Sans couplage physique réel avec la matière et sans le **bruit imprévisible** du monde réel, il n'y a **pas d'incarnation véritable**.

Rodney Brooks, Inman Harvey, Tom Ziemke

Débat

Traduire les schémas d'image en formules logiques, est-ce ancrer le sens ou ajouter des symboles ?

Débat

Traduire les schémas d'image en formules logiques, est-ce ancrer le sens ou ajouter des symboles ?

Ancrer le sens

Les **schémas d'images** proviennent directement de notre **expérience corporelle et spatiale**. Les formaliser permet de lier l'expérience vécue au raisonnement abstrait sans rompre avec le corps.

George Lakoff, Mark Johnson, François Olivier (théorie des schémas d'images)

Ajouter des symboles

Traduire une expérience en formule logique reste de la **manipulation formelle de symboles sans accès à la vraie signification**. Le sens ne vient pas de la logique, mais de l'interaction causale et biologique directe avec le monde.

Stevan Harnad, John Searle, Tom Ziemke, Mark Bickhard

Débat

Une IA doit-elle avoir quelque chose à perdre, voire des émotions, pour être intelligente ?

Débat

Une IA doit-elle avoir quelque chose à perdre, voire des émotions, pour être intelligente ?

Oui

La cognition est née pour servir l'**homéostasie** et la **survie**. Sans **vulnérabilité biologique** (« quelque chose à perdre ») ni ancrage émotionnel, les actions d'une IA n'ont aucune valeur intrinsèque propre.

Antonio Damasio, Francisco Varela, Evan Thompson, Tom Ziemke

Non

L'intelligence (la capacité à résoudre des problèmes ou optimiser une fonction objectif) est **orthogonale aux motivations**, à la **biologie** ou à la **sensibilité**. Une IA peut être hautement intelligente sans ressentir d'émotions ni posséder de métabolisme.

Nick Bostrom (Thèse de l'orthogonalité), Stuart Russell, Yann LeCun, Alan Turing