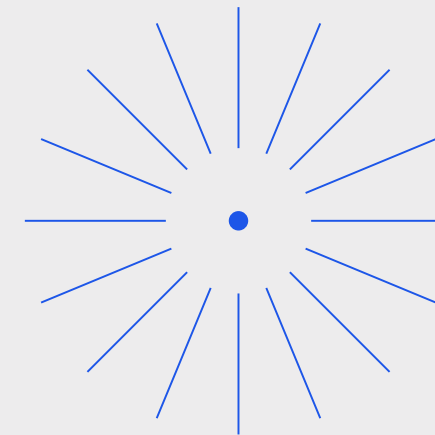
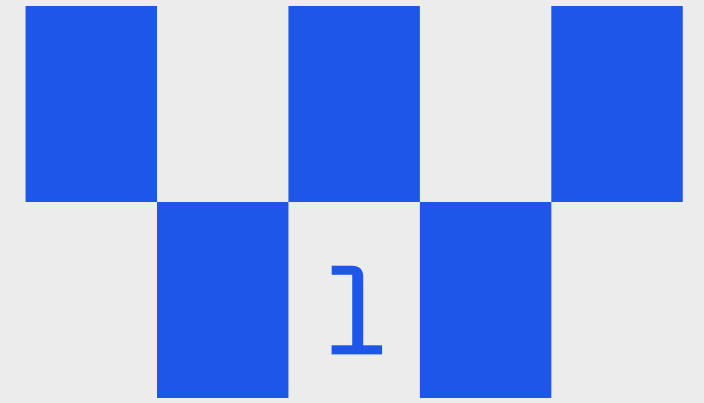
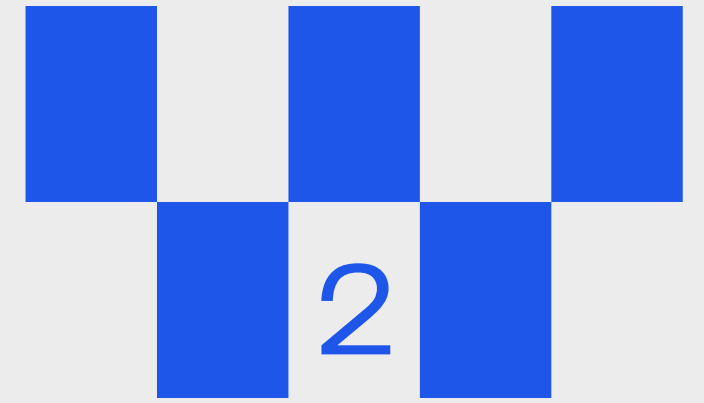


Embodied Cognition

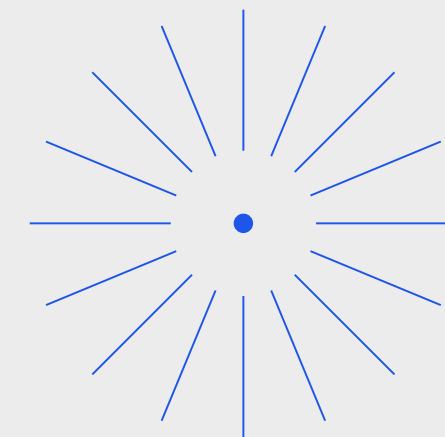


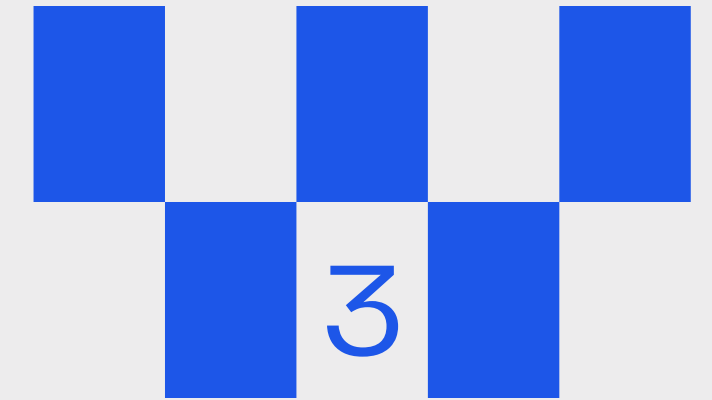
L'IA a-t-elle besoin d'un corps ?



Cognition désincarnée, c'est quoi ?

- La cognition désincarnée, c'est une manière de voir l'intelligence et la pensée comme quelque chose de purement mental, **séparé** du corps et du monde physique.
- Intelligence indépendant du corps
- Manipulation de symboles et de données



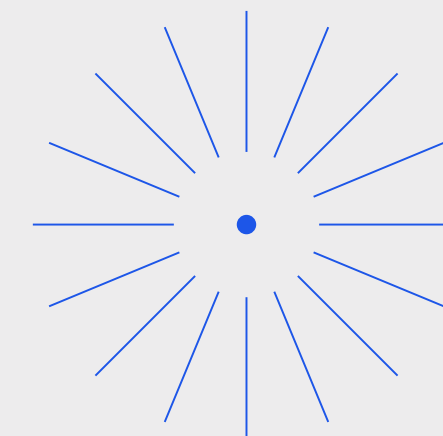


Critiques des philosophes sur la cognition désincarnée

- **Searle** : Il montre que manipuler des symboles (comme le fait un ordinateur) ne suffit pas pour comprendre leur sens
- **Dreyfus** : Selon lui, l'intelligence humaine dépend du **corps**, de **l'expérience** et du **contexte**, pas seulement de règles logiques.
- L'esprit est étroitement lié à nos expériences dans le monde

Travaux de Lakoff et Johnson sur les concepts abstraits

- On comprend des concepts abstraits en nous appuyant sur nos expériences physiques
- “L'avenir est devant nous”

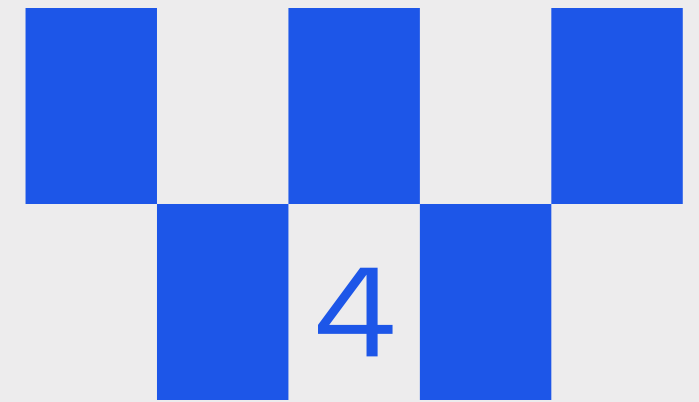
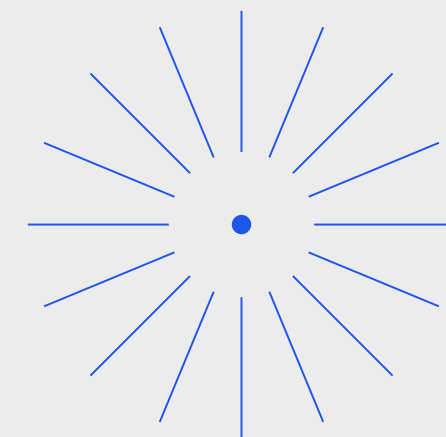


1 – Le problème de l'ancrage des symboles

- Comment lier le raisonnement abstrait d'un système IA à la compréhension et à l'action dans le monde réel ?
- Besoin de sens et de raisonnement

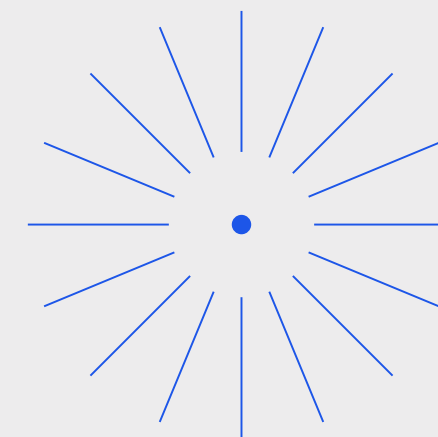
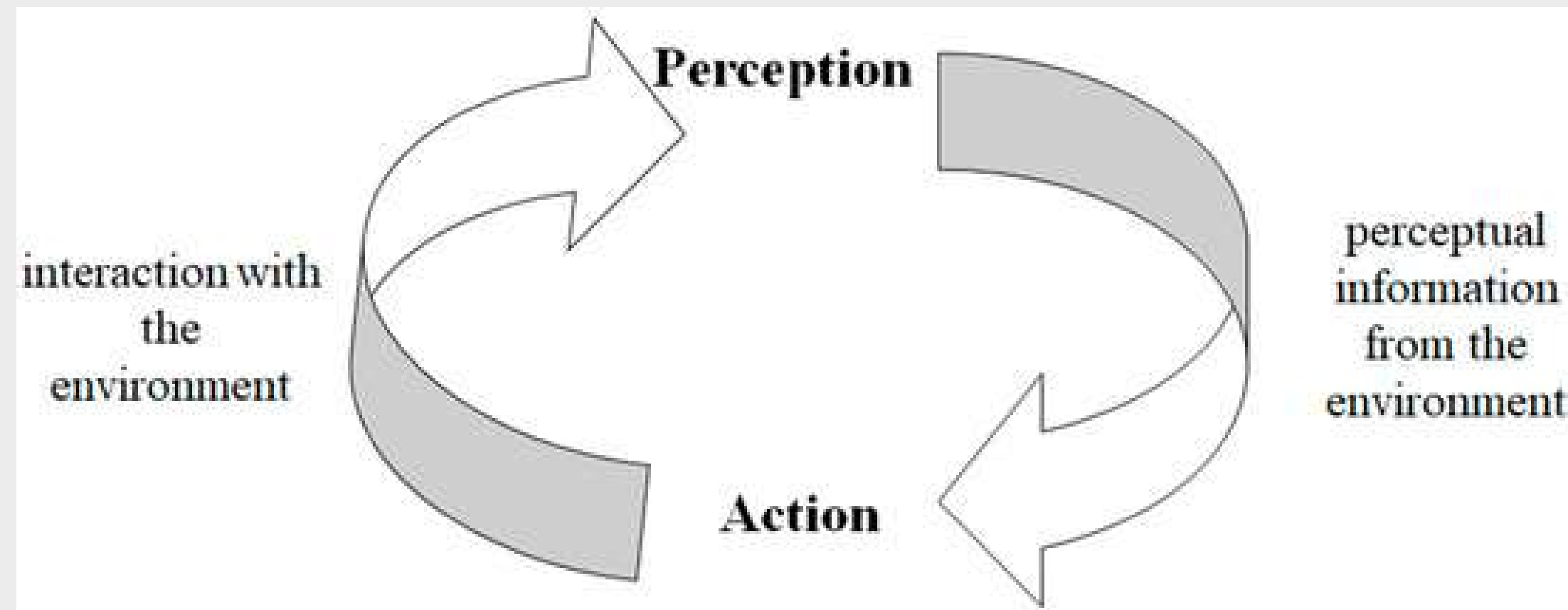
2 – Créer de l'intelligence à partir d'expériences

Intelligence Incarnée



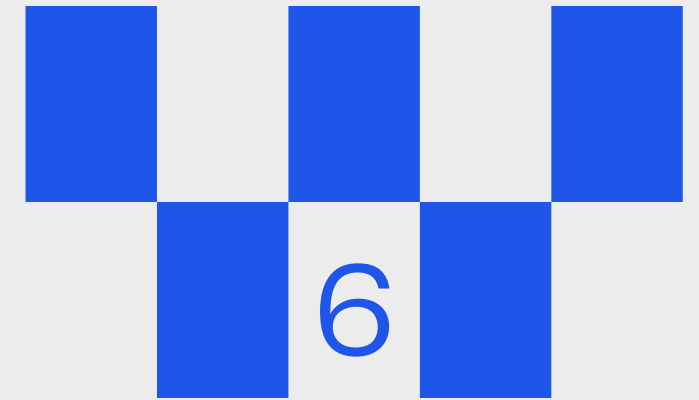
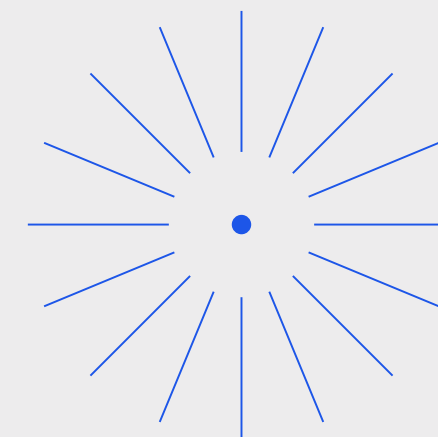
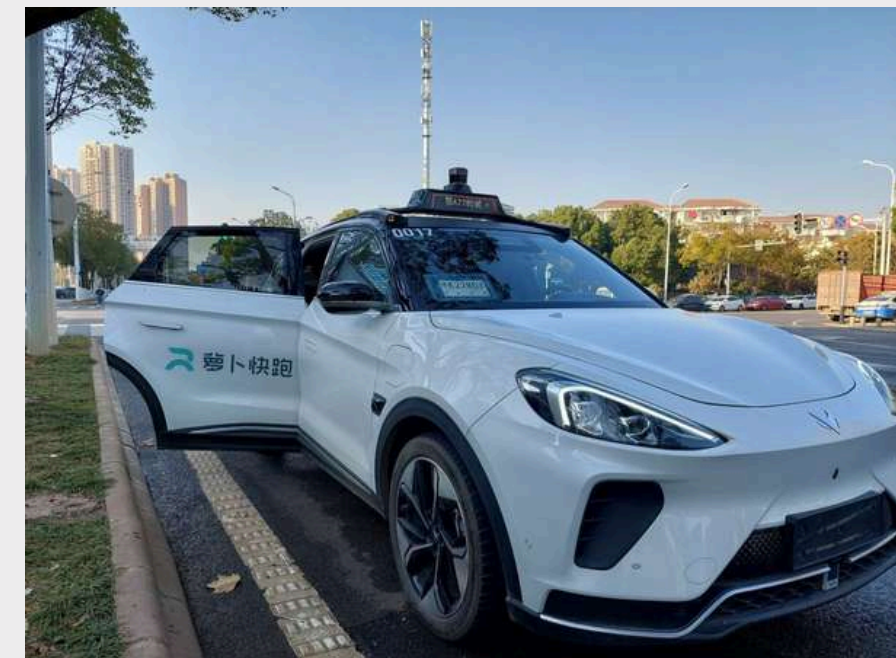
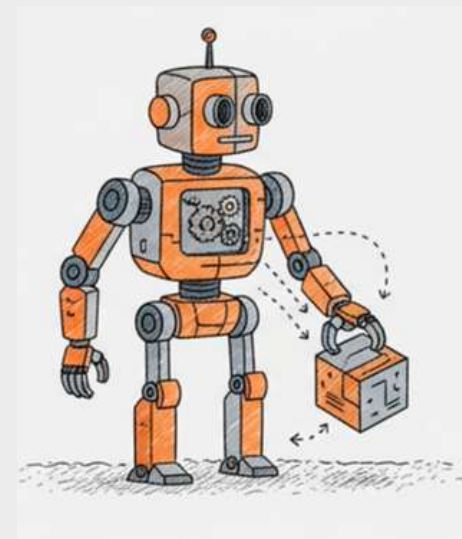
Lien entre l'intelligence et le corps

- La boucle Perception-Action



Pourquoi incarner l'IA ?

- Plus d'autonomie
- Voiture autonome
- Robots sauveteurs (Incendie, Séisme)
- Se rapprocher de l'AGI



Comment évaluer l'intelligence d'un robot

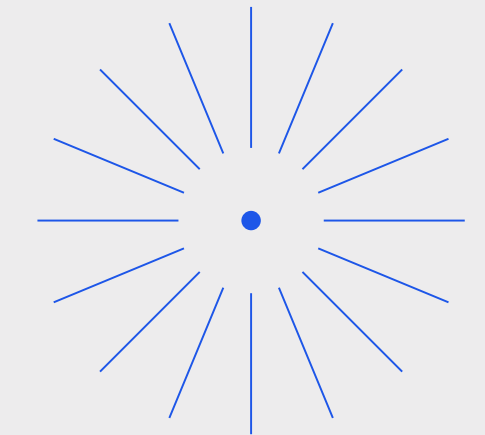
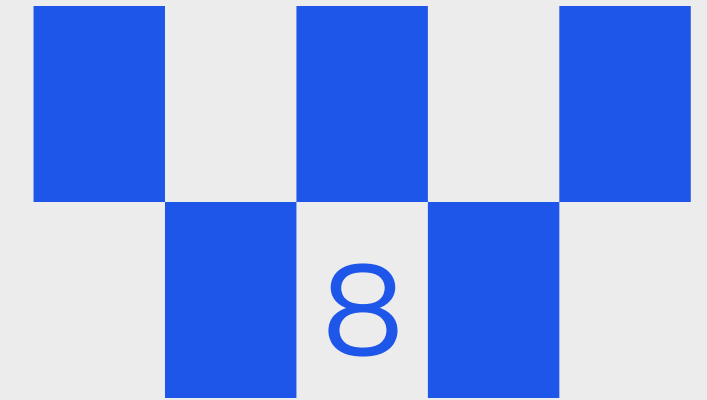
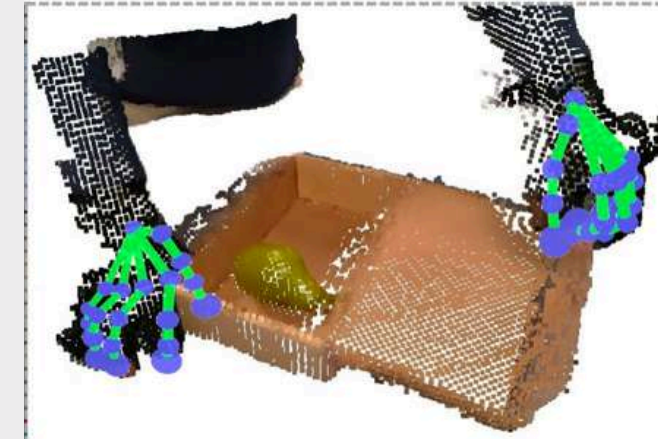
- Test de Turing ?
- Niveaux IR-L (Intelligent Robot Level) :



Niveau	Nom du Niveau	Autonomie	Gestion des Tâches	Adaptabilité	Cognition Sociétale	Exemples
IR-L0	Exécution de base <i>(Basic Execution)</i>	Contrôle humain	Tâches simples	Aucune	Aucune cognition sociale	Robots industriels
IR-L1	Réponse programmatique <i>(Programmatic Response)</i>	Supervisé par l'humain	Navigation complexe	Environnements prévisibles	Reconnaissance de base	Robots de service simples
IR-L2	Perception et adaptation de base <i>(Basic Perception and Adaptation)</i>	Assisté par l'humain	Collaboration dynamique	Apprentissage adaptatif	Interaction simple	Robots de service généraux
IR-L3	Cognition et collaboration humanoïde <i>(Humanoid Cognition and Collaboration)</i>	Autonomie conditionnelle	Multitâche	Adaptation dynamique	Intelligence émotionnelle	Robots compagnons
IR-L4	Totalement autonome <i>(Fully Autonomous)</i>	Autonomie totale	Innovation	Flexibilité universelle	Intelligence sociale avancée	Robots de science-fiction

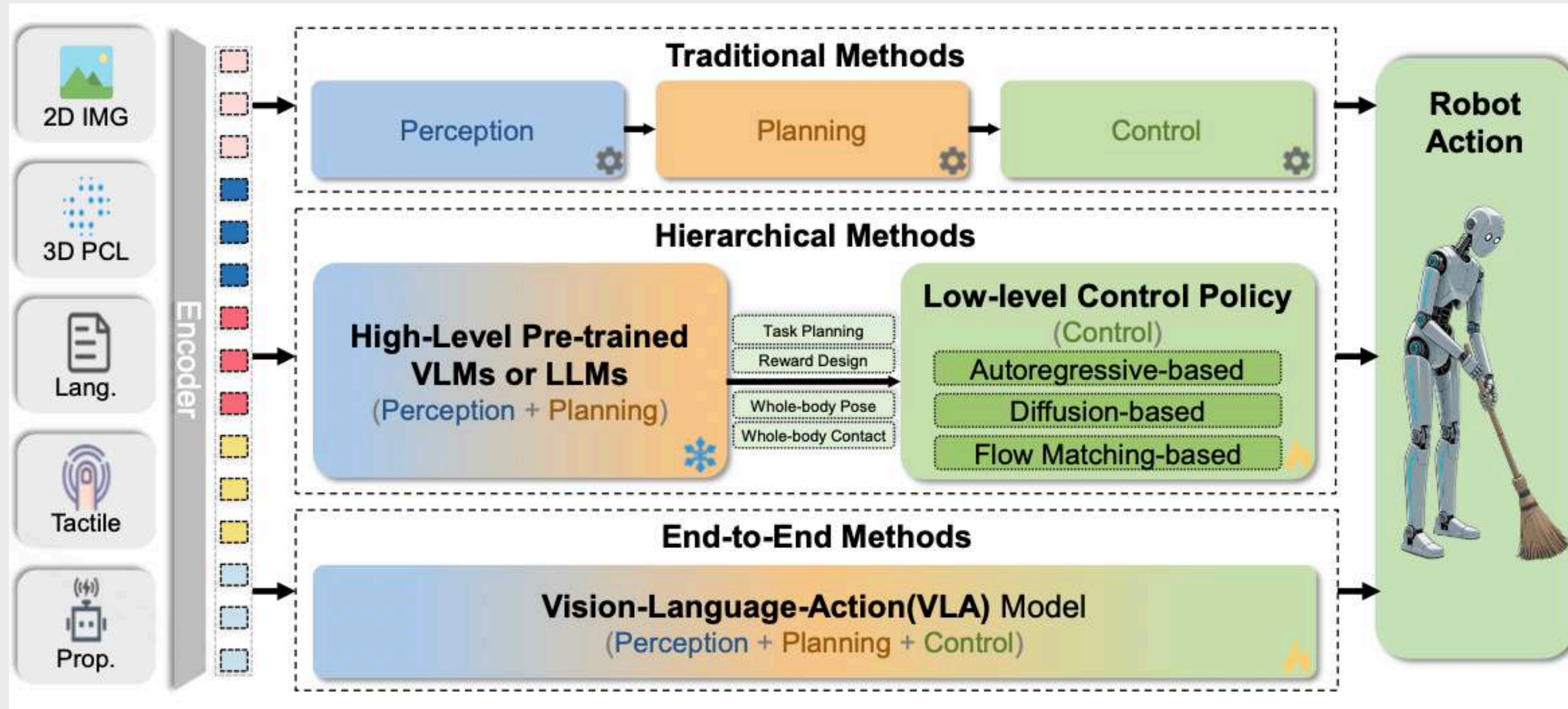
Incarner l'IA : Solutions techniques

- Model Predictive Control
 - Whole Body Control
- } Modèles physiques/mathématiques
-
- Reinforcement Learning → Récompenses
 - Imitation Learning → Demonstrations
-
- Multimodal Foundation Models :
 - Language
 - Vision
 - Actions



Solutions techniques

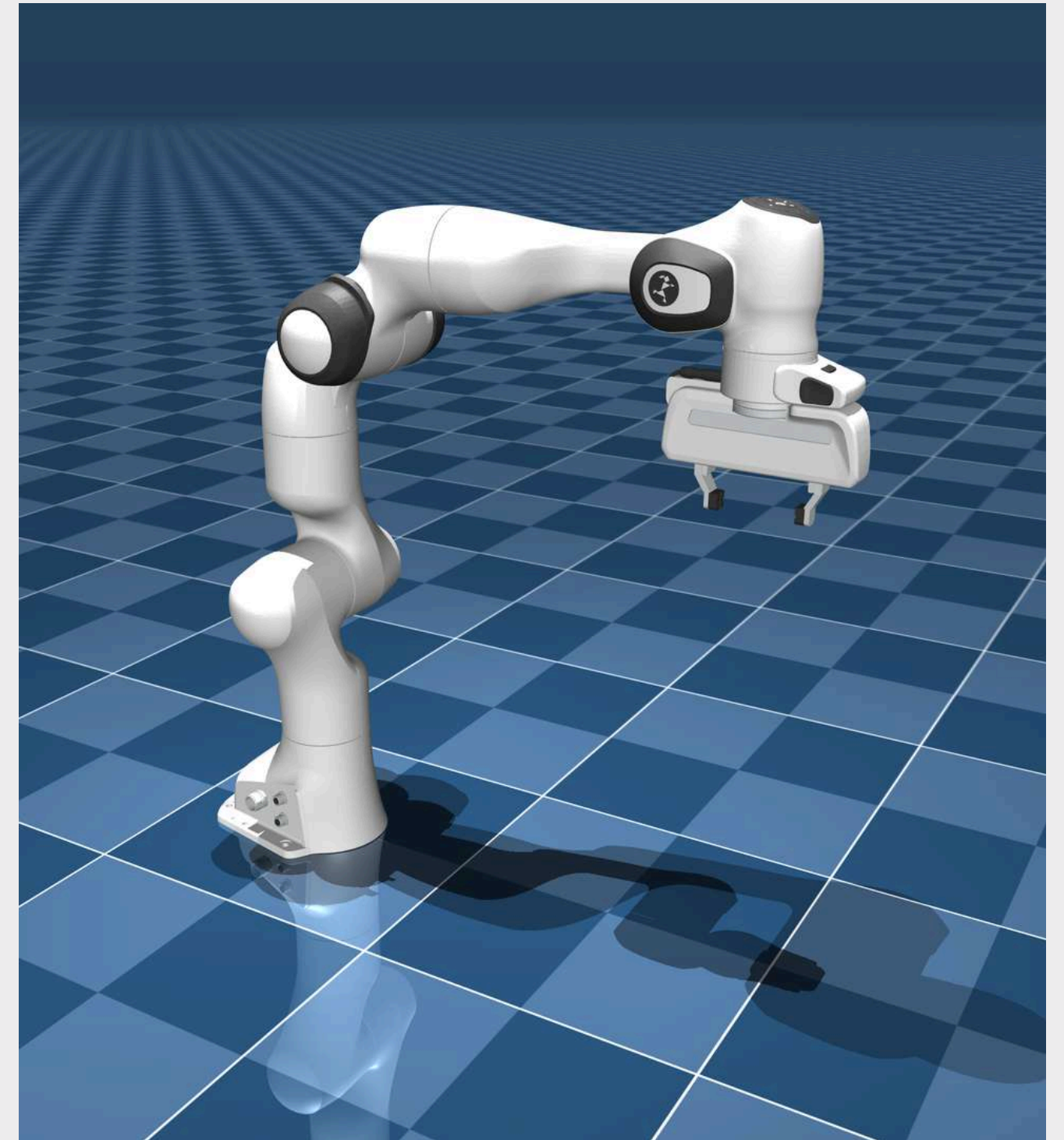
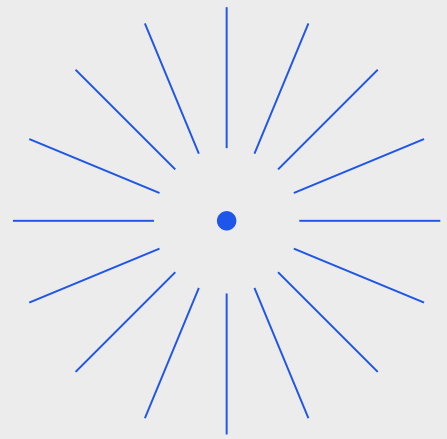
9



Apprentissage par interaction avec l'environnement

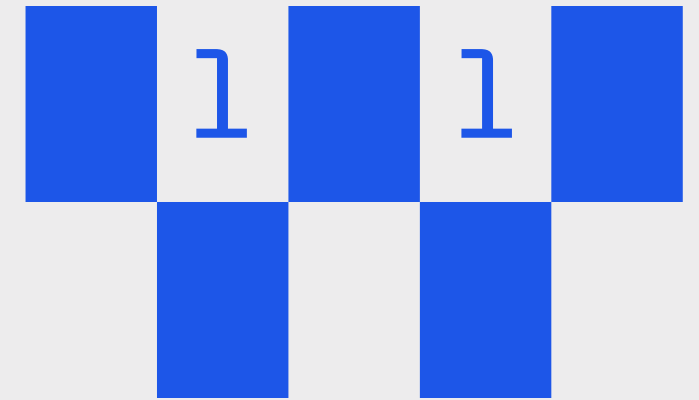
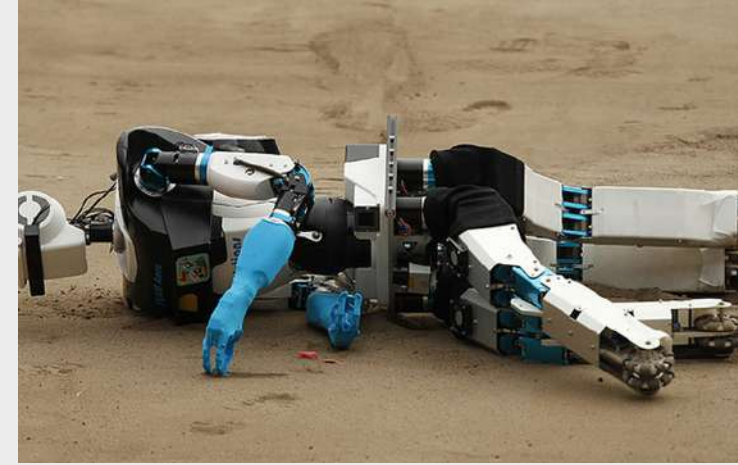
“A Survey: Learning Embodied Intelligence from Physical Simulators and World Models”, 3 Sep 2025,
X. Long, Q. Zhao et al., arXiv:2507.00917v3 [cs.RO]

Les simulateurs physiques



P O U R Q U O I E N T R A I N E R U N R O B O T D A N S
U N S I M U L A T E U R ?

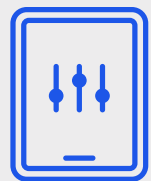
Les nombreux avantages du simulateur



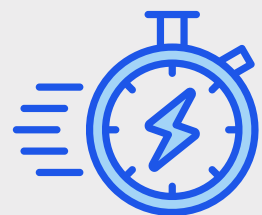
Entraîner un robot dans le monde réel userait sans cesse le matériel. Avec un simulateur, l'entraînement est virtuel et seul le coût des serveurs compte.



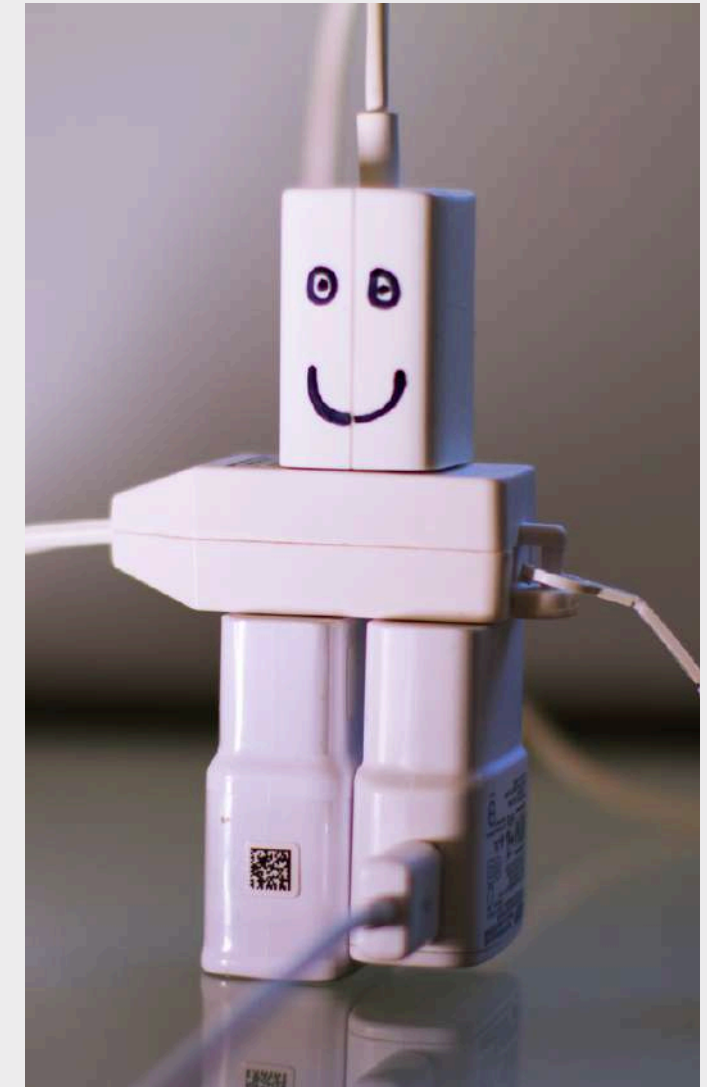
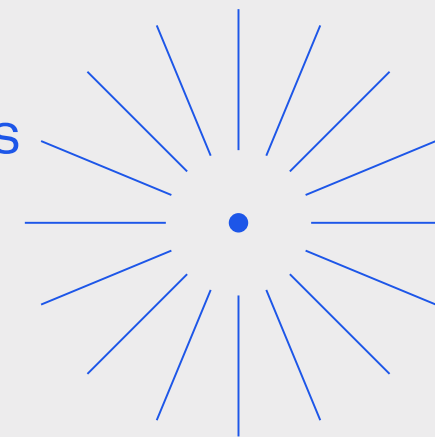
Tester un robot dans le monde réel comporte des risques pour lui et son entourage. Le simulateur permet, lui, d'expérimenter ces situations sans danger.



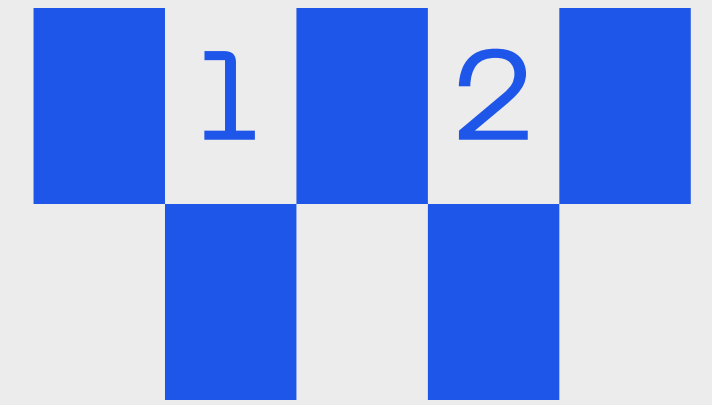
Le simulateur offre un contrôle total, ce qui permet d'élargir le champ d'action du futur robot.



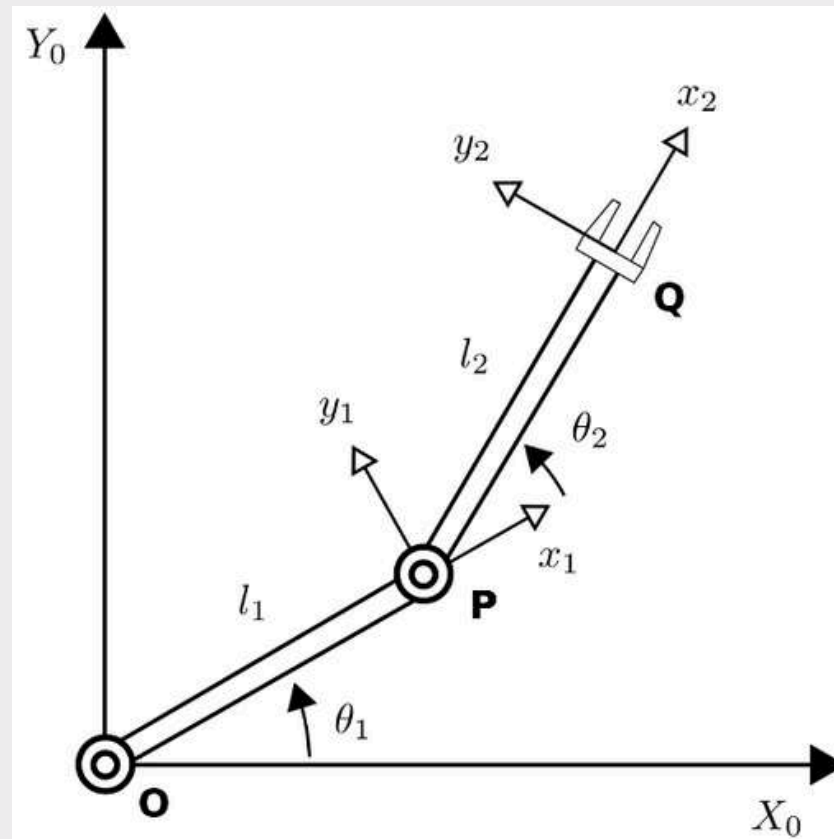
Entraîner un robot dans le monde réel serait long et contraignant. Le simulateur permet, lui, de gagner un temps précieux.



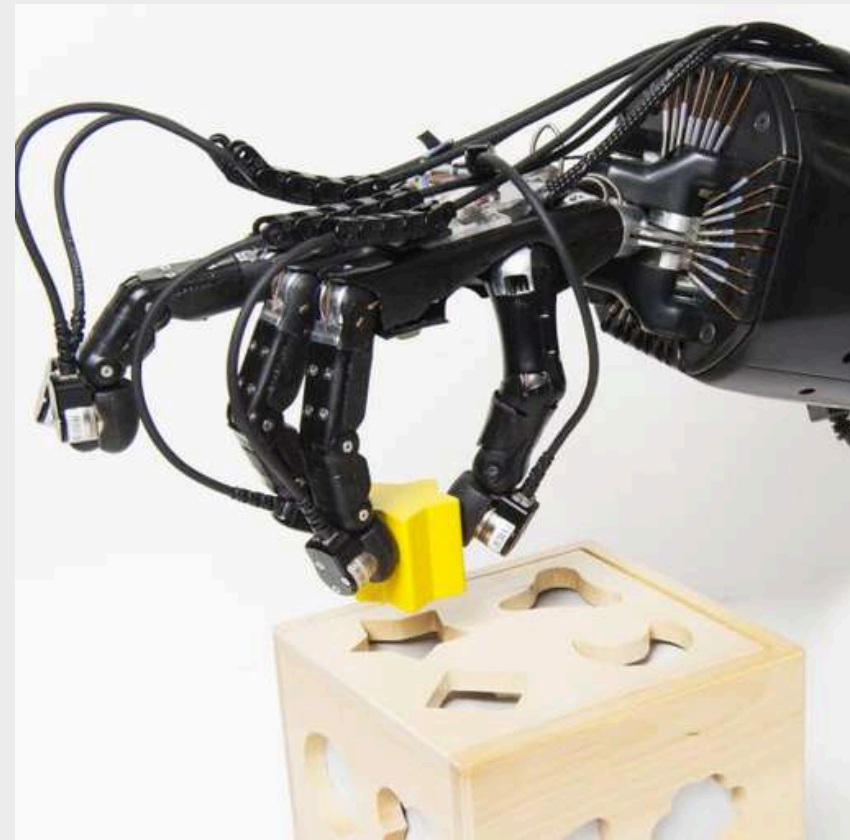
Qu'imite réellement une simulation physique ?



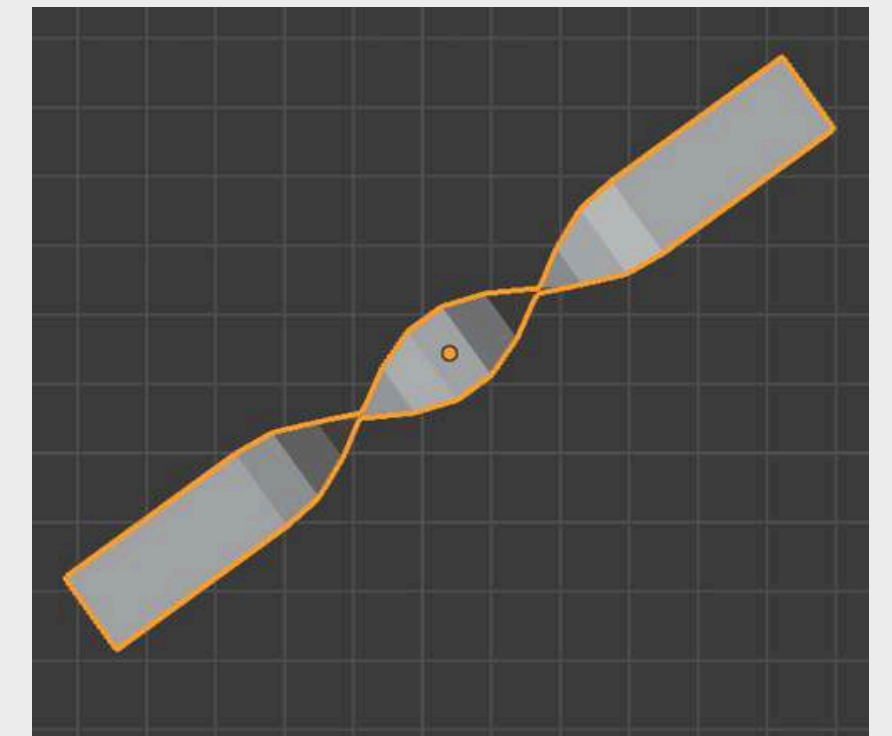
Dynamique des corps rigides



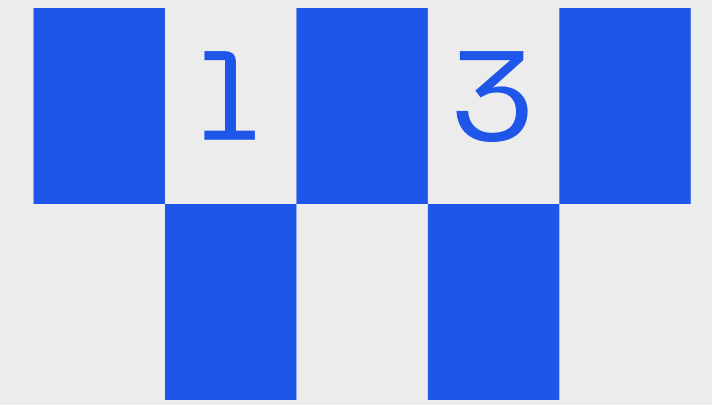
Modélisation des contacts



Dynamique des corps pliables



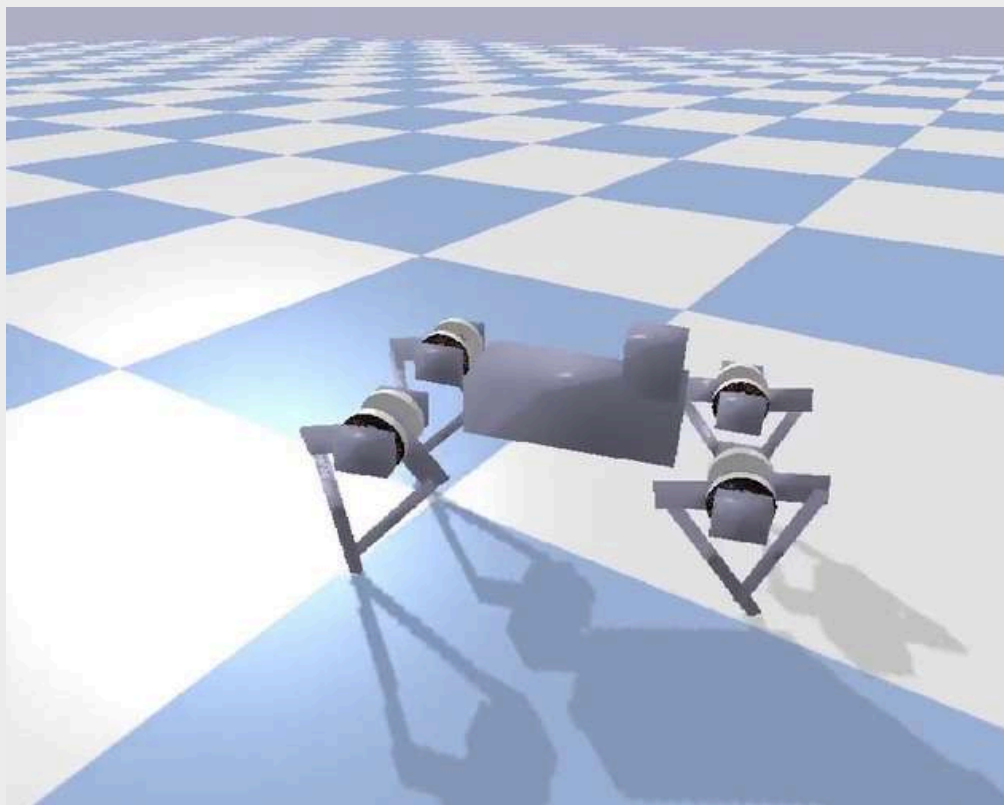
Les simulateurs physique de référence



PyBullet

Crée en 2015

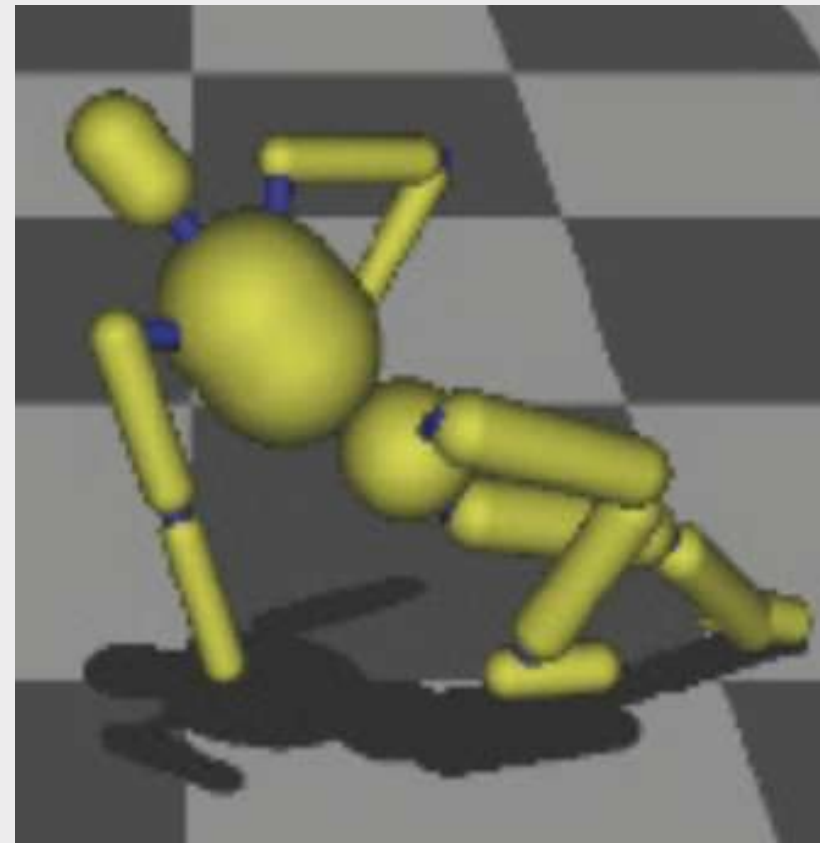
Idéal pour les projet académique et petit projet



MuJoCo

Crée en 2016

Idéal pour la recherche scientifique et pour du RL



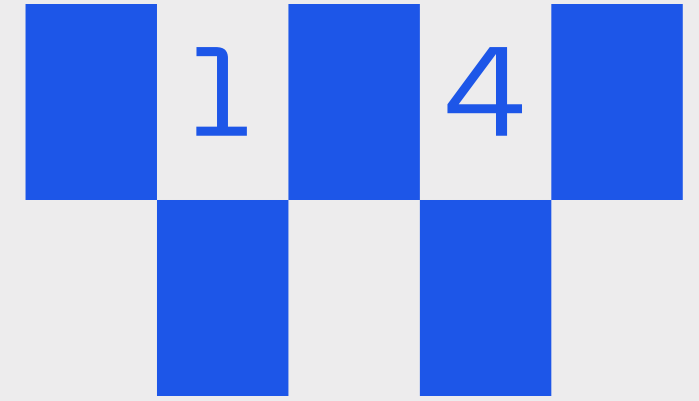
Isaac Sim

Crée en 2020

Idéal pour les robots du monde réel et les tests industriels



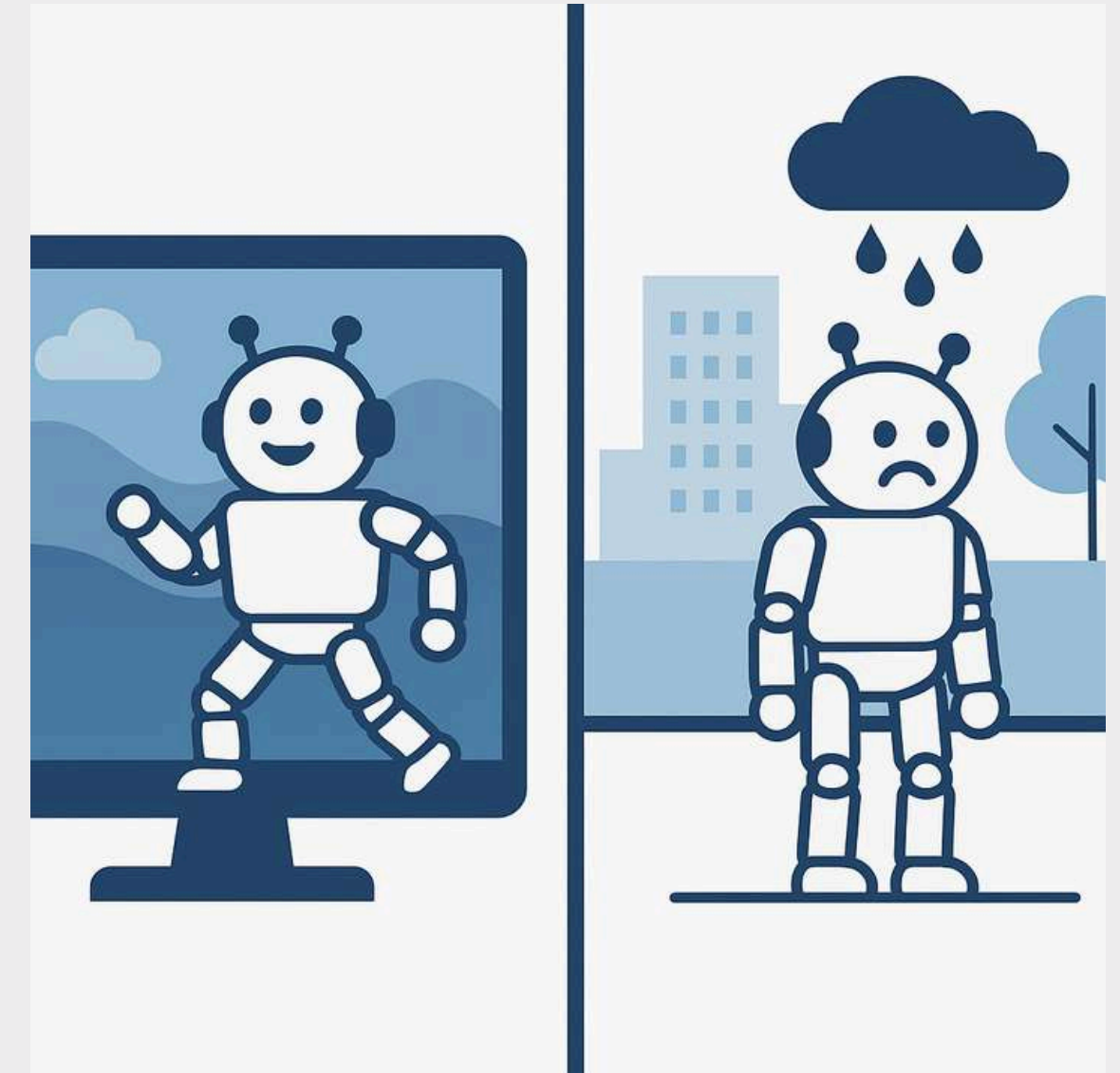
Le problème : le Sim-to-Real Gap



- Les robots apprenent des “hack” propres à leurs simulateurs. Exemple de la voiture autonome qui n’apprend que dans des parfaitement plat, sans bosse ni poussière.
- Même les modèles les plus élaborés échouent à imiter la complexité du monde réel : trop coûteux à calculer, et fondés sur des lois physiques encore incomplètes.

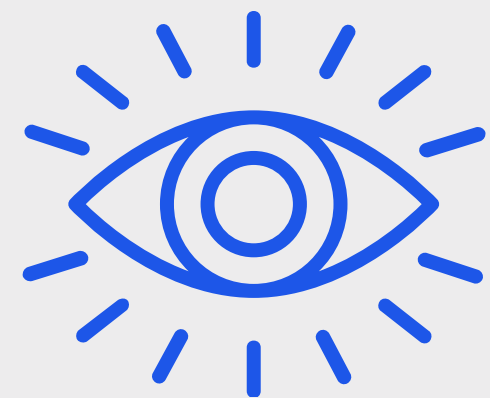
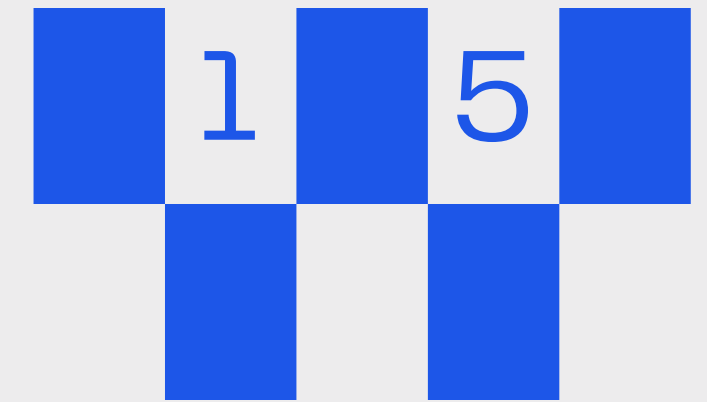


**Les robots manquent de capacité
d'adaptation**



La solution : les modèles du mondes

- Les humains imaginent les conséquences de leurs actions.
Exemple : balle de baseball qui vient à 100 km/h, que faire ?
- Ces modèles du mondes permettent alors à une IA d'apprendre une simulation interne du monde.
- Concept établi dans l'article fondateur de David Ha et Jurgen Schmidhuber (2018):



La vision

encode l'observation



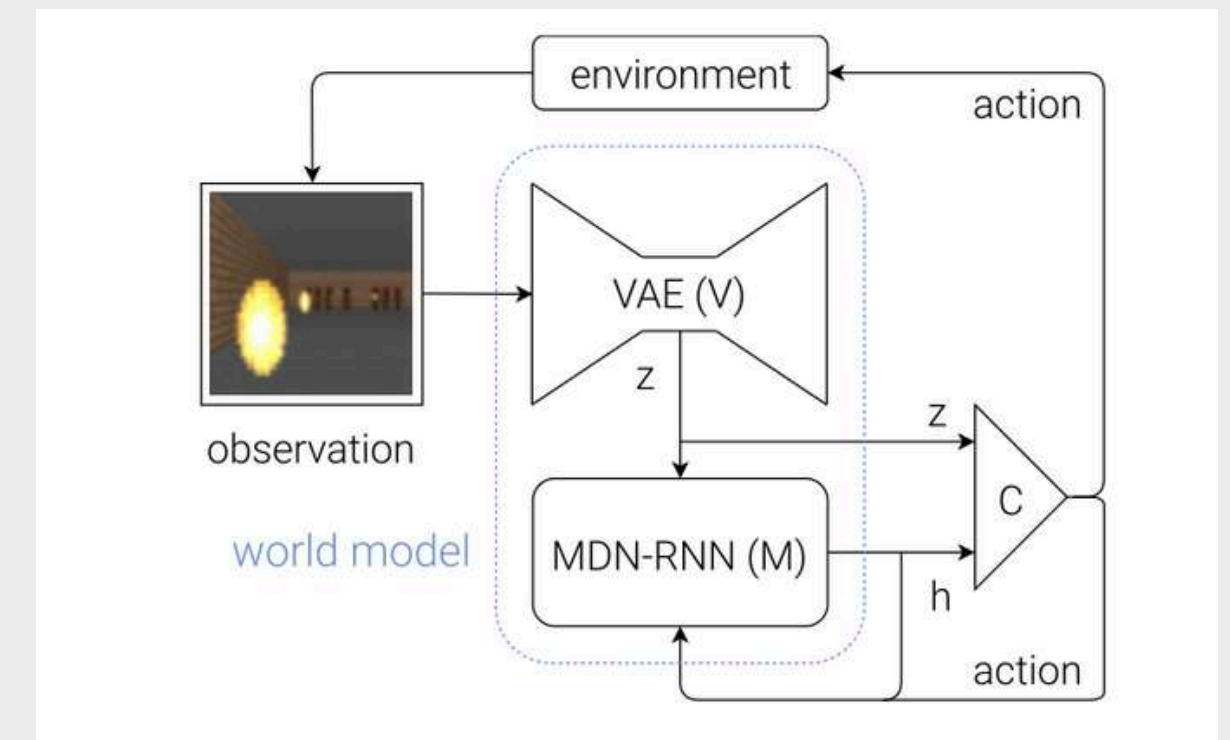
La mémoire

prédit les états à $t+1$



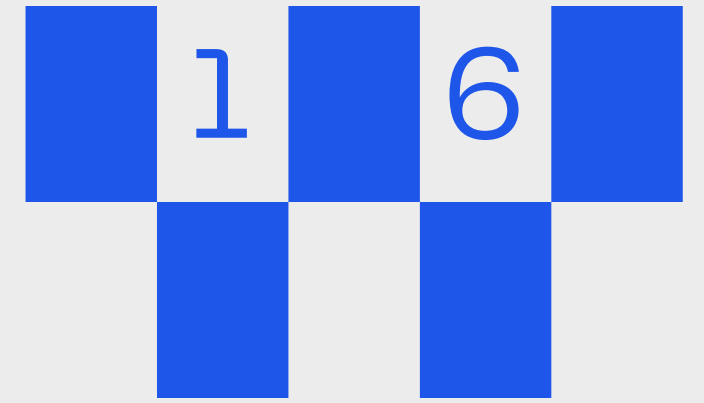
Le contrôleur

Choisit la meilleure action

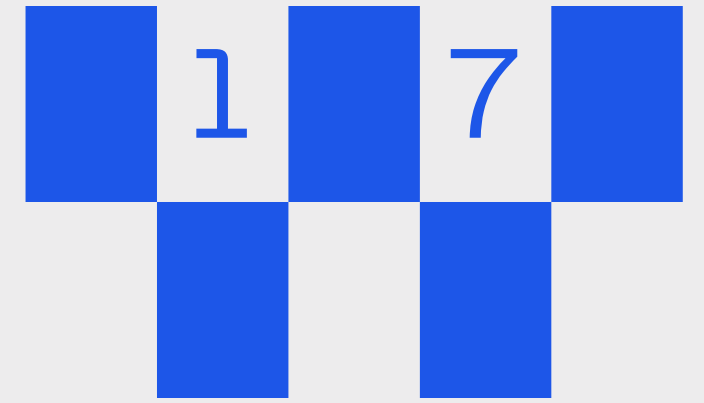


Les simulateurs internes : générer le futur

- Modèles auto-régressif : prèdisent image par image
Exemple : GAIA 1 (modèle du monde pour la conduite apparue en 2023)
- Modèles de diffusion : permettent de générer des scènes complètes.
Exemple : Midjourney (2022), Sora (2024), GAIA 2 (2025)

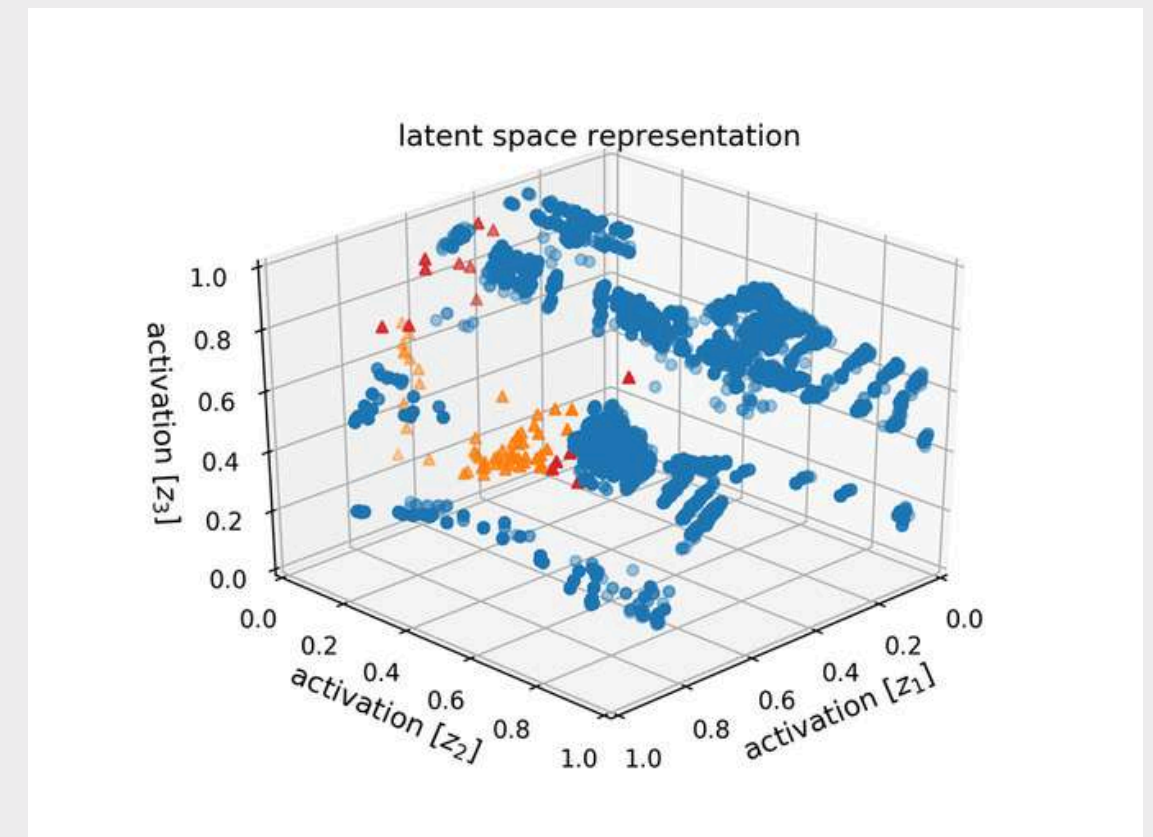


Les modèles du monde dynamique

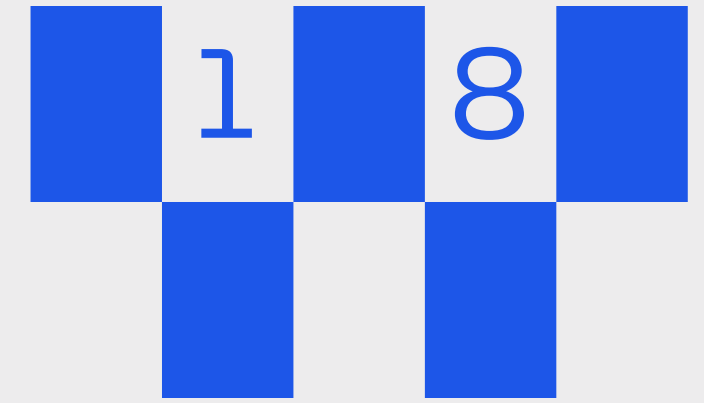


- A partir de vidéo apprenant la physique et la logique de l'environnement
- Ces modèles travaillent sur un espace latent compact.
- Exemple: Dreamer regarde des images d'un environnement (ex: une balle qui tombe ou un bras qui bouge). Ce modèle apprend ensuite dans l'espace latent.

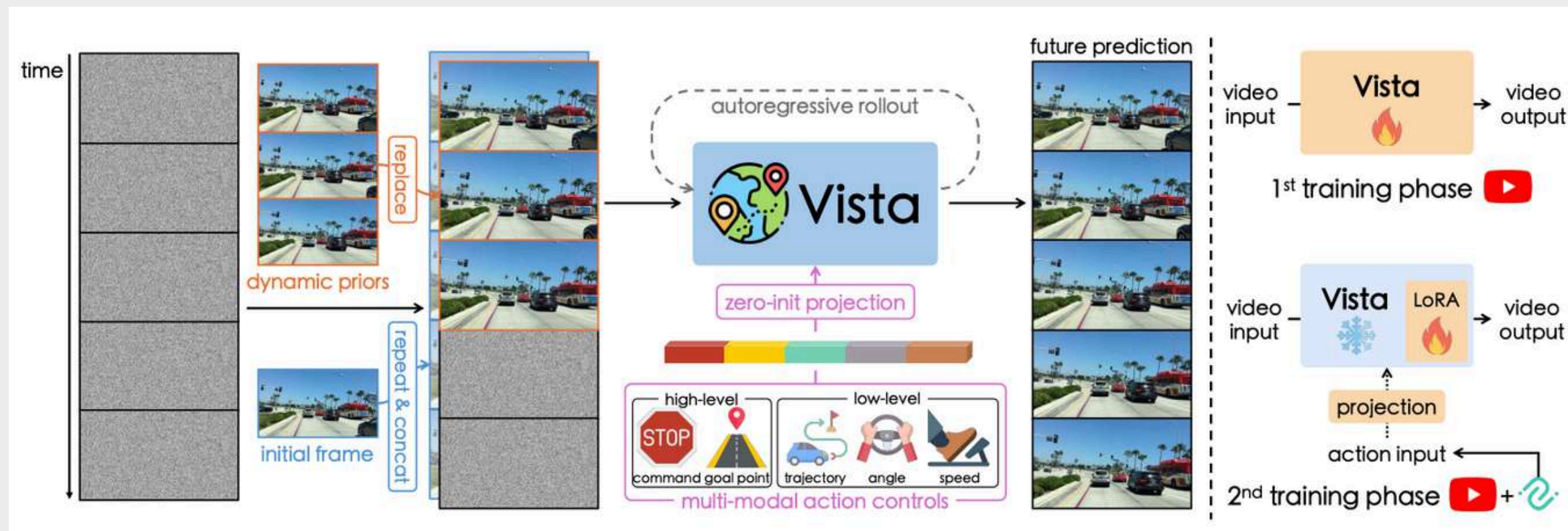
Si je fais telle action, qu'est ce qui va se passer ensuite ?



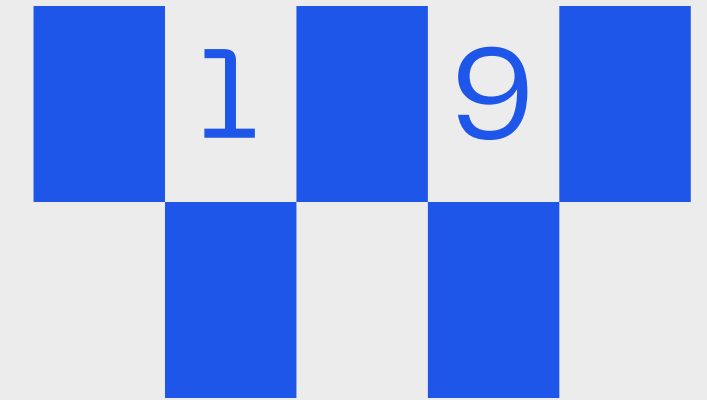
Quand l'IA apprend qu'est ce qu'un “bon” comportement



- Apprentissage sans fonction de récompense explicite
- Exemple: Vista

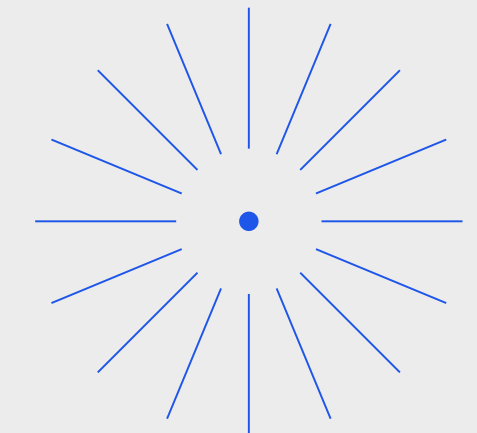


Vers l'IA générale ?

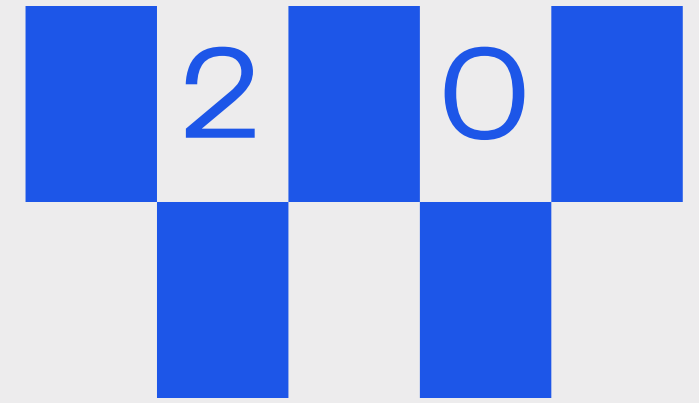


Les "World Models" apprennent une simulation interne du monde, inspirée de notre cerveau.

- Apprentissage Autonome :
Ils construisent leur compréhension à partir de l'expérience (vidéos, interactions), sans programmation explicite des règles du monde
- Ancrage dans l'Incarnation :
Comme l'intelligence humaine, ils relient la cognition abstraite à l'action physique via la boucle perception-action.
- Prédiction & Planification :
Permettent à l'IA d'imaginer le futur et de choisir les meilleures actions sans tout tester réellement.



Les limites de l'IA incarnée

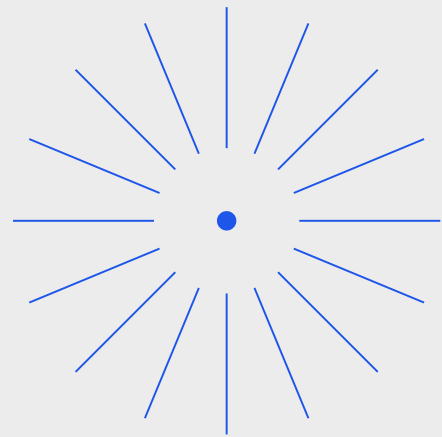


Limitations actuelles des world models

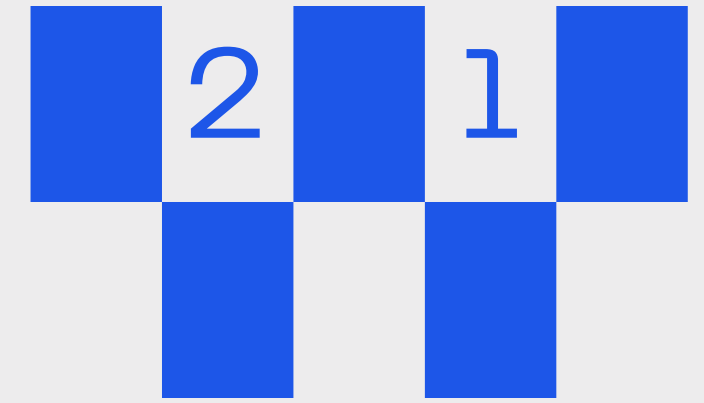
- Correlation $\neq$ Causalité
- Mémoire à court terme
- Interprétabilité

Mais aussi ...

- Alignement
- Hallucinations
- Guardrails



Débat



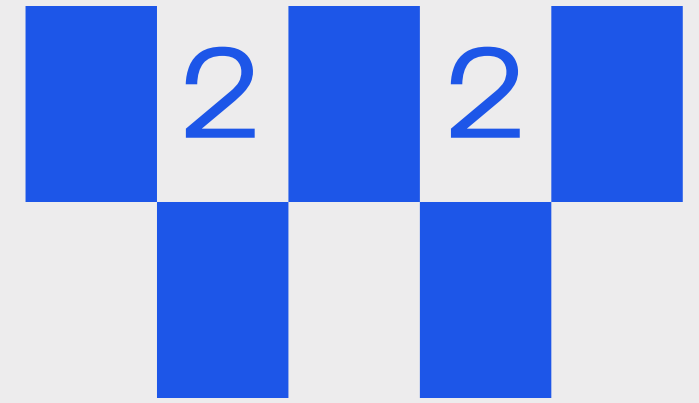
Unitree H1 Malfunction



Système relativement simple, mal fonctionnement capteur

→ **Quels pourraient être les risques avec des systèmes plus complexes et autonomes ?**

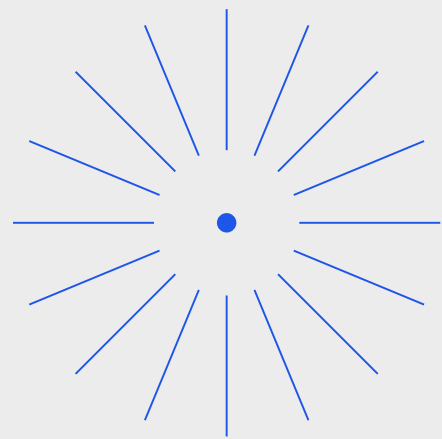
Débat



<https://deepmind.google/models/gemini-robotics/>

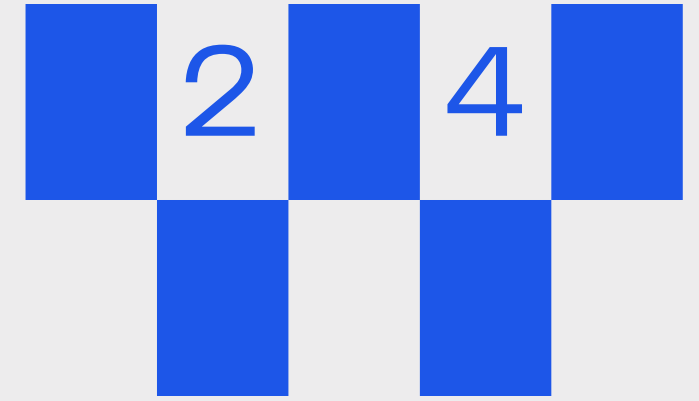


Based on CA trash sorting rules, here is the step-by-step procedure: First, pick up the red coke can ; since metal beverage cans are recyclable, place it in the light



Face à ces capacités, **Quelle approche vous semble la plus prometteuse pour développer une IA incarnée plus forte et adaptable ?** : VLA/VLM, World Models, Combinaisons...

Débat

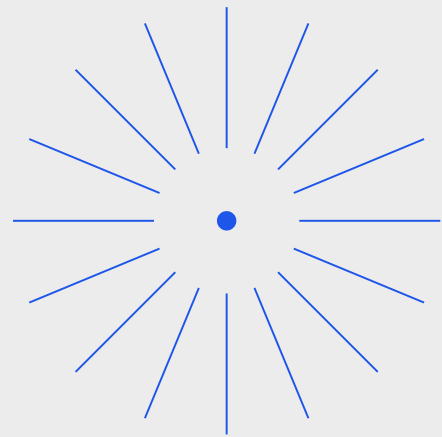


Les modèles actuels sont souvent entraînés avec des données ou objectifs définis par l'humain.

Que se passerait-il si on laissait une IA avec un "world model" apprendre avec beaucoup **plus de liberté**, juste guidée par un objectif très large comme "**explorer**" ou "**survivre**" ?

Quels types de représentations du monde et quels comportements pourraient émerger ?

Est-ce une piste intéressante ou un **risque** ?



Biblio

- A Survey: Learning Embodied Intelligence from Physical Simulators and World Models
- Mastering Diverse Domains through World Models
- World Models
- Grounding Agent Reasoning in Image Schemas: A Neurosymbolic Approach to Embodied Cognition
- The body of knowledge: On the role of the living body in grounding embodied cognition
- Learning Latent Dynamics for Planning from Pixels
- Isaac Gym: High Performance GPU-Based Physics Simulation For Robot Learning
- MuJoCo: A physics engine for model-based control
- RT-2: Vision-Language-Action Models Transfer Web Knowledge to Robotic Control
- ViZDoom: A Doom-based AI Research Platform for Visual Reinforcement Learning
- Gemini Robotics 1.5: Pushing the Frontier of Generalist Robots with Advanced Embodied Reasoning, Thinking, and Motion Transfer

