BELABBAS-BENGRAA Joubrane
BENZIANE Amir
CALMETTES Léo

# Synthèse sur *Embodied Cognition*

Nous avons entamé notre exposé par une mise en lumière de l'évolution de la conception de l'intelligence, en commençant par la Cognition Désincarnée qui postule une pensée purement mentale, séparée du corps et du monde physique, basée sur la manipulation de symboles.

Puis, nous avons présenté les critiques majeures des philosophes à l'encontre de cette approche. Nous avons cité Searle, qui démontre que la simple manipulation de symboles ne suffit pas à conférer le sens ou la compréhension , et Dreyfus, pour qui l'intelligence humaine est inextricablement liée au corps, à l'expérience et au contexte. Les travaux de Lakoff et Johnson ont renforcé ce point, expliquant que nous comprenons les concepts abstraits en nous appuyant sur nos expériences physiques.

Nous avons par la suite expliqué le problème de l'ancrage des symboles , c'est-à-dire comment lier le raisonnement abstrait à l'action et à la compréhension dans le monde réel. C'est ce problème là ainsi que l'envie de créer de l'intelligence à partir d'interactions qui ont mené à l'émergence de l'intelligence incarnée. Nous avons par la suite montré un concept qui permet de faire le lien entre l'intelligence et le corps à travers la boucle Perception-Action , où l'interaction avec l'environnement guide la perception, qui à son tour influence l'action, dans un cycle continu.

Nous avons ensuite exploré pourquoi incarner l'IA , citant l'objectif d'atteindre une plus grande autonomie pour les applications pratiques comme la voiture autonome ou encore les robots sauveteurs. Ces robots sont importants et notamment dans les catastrophes naturelles ou dans les incendies pour sauver des vies. Un des objectifs cités était aussi l'envie de se rapprocher de l'AGI (Intelligence Générale Artificielle) et que certains chercheurs pensent que pour atteindre cet objectif là, l'intelligence artificielle doit être physiquement dans notre monde et pas uniquement dans des serveurs. Pour mesurer les progrès, nous avons vu pourquoi le test de Turing ne pouvait pas s'appliquer dans le cas des robots comme la voiture autonome. C'est pourquoi nous avons introduit les Niveaux IR-L (Intelligent Robot Levels) comme grille d'évaluation de l'intelligence d'un robot.

L'exposé s'est poursuivi avec les solutions techniques pour incarner l'IA. Nous avons mentionné le Model Predictive Control et le Whole Body Control pour le contrôle physique , ainsi que le Reinforcement Learning et l'Imitation Learning pour l'apprentissage. Un point clé est l'émergence des Modèles Fondateurs Multimodaux intégrant le langage, la vision et l'action. Nous avons détaillé comment ces modèles s'intègrent dans trois architectures : les méthodes Traditionnelles (Perception → Planning → Control), les méthodes Hiérarchiques (VLMs/LLMs de haut niveau et politiques de contrôle de bas niveau) et les méthodes End-to-End avec les modèles VLA (Vision-Language-Action).

Nous avons ensuite consacré une partie à l'importance des simulateurs physiques , expliquant pourquoi entraîner un robot dans un simulateur est avantageux en termes de réduction des coûts de matériel, de sécurité, de contrôle total, et de gain de temps. Cependant, nous avons rapidement identifié la limite majeure : le Sim-to-Real Gap. Ce

problème survient car les robots apprennent des choses propres au simulateur, et même les modèles physiques les plus élaborés échouent à imiter la complexité du monde réel, menant à un manque de capacité d'adaptation des robots.

Pour surmonter ce fossé, nous avons introduit les modèles du monde (World Models) , une solution inspirée de la capacité humaine à imaginer les conséquences de nos actions. Nous avons détaillé l'architecture de base d'un World Model, comprenant la vision (encode l'observation), la mémoire (prédit les états futurs), et le contrôleur (choisit la meilleure action). Nous avons aussi donné des exemples comme GAIA et Dreamer.

En conclusion, nous avons réaffirmé que les World Models sont une voie prometteuse vers l'IA générale. Ils permettent un apprentissage autonome à partir de l'expérience , ancrent la cognition abstraite dans l'incarnation , et offrent des capacités de prédiction et planification. Enfin, nous avons conclu en soulevant les limites de l'IA incarnée et des World Models actuels, notamment les problèmes de causalité, de mémoire à court terme, d'interprétabilité, ainsi que les risques plus larges liés à l'alignement et aux hallucinations dans des systèmes autonomes complexes.

Le débat qui a suivi l'exposé a permis d'aborder les implications éthiques et les défis conceptuels de l'Intelligence Artificielle Incarnée (Embodied AI), notamment en matière de sécurité et de définition de la cognition.

1. Sécurité et Contrôle des Systèmes Incarnés

Le débat a été lancé par la projection d'une vidéo montrant un robot aux comportements erratiques (une "crise"), soulevant immédiatement la question de la peur face à ces entités : « Avez-vous peur de ces robots ou de leur avenir ? » .La majorité des participants a répondu par l'affirmative, soulignant qu'une fois ces systèmes déployés dans notre environnement, l'imprévisibilité de leurs comportements et l'incapacité d'anticiper toutes les situations constituent un risque majeur.

- Solutions de Contrôle : Une solution proposée par un étudiant, inspirée de la vidéo, était d'intégrer un simple bouton Marche/Arrêt. Cette idée a été discutée : si elle peut résoudre certains problèmes ponctuels (comme la défaillance du robot montré), elle devient inapplicable dans des cas critiques comme celui des voitures autonomes, où l'arrêt brutal pourrait causer un accident.
- Limitation des Risques : D'autres propositions pour mitiger les dangers incluaient la restriction des robots à des environnements bien définis, ou l'imposition de tests de validation très stricts pour les systèmes dotés d'une intelligence élevée.

2. La Question de l'IA Incarnée dans les Voitures Autonomes

Une clarification conceptuelle a été nécessaire concernant les véhicules autonomes : s'agit-il réellement de cognition incarnée ?

- Argument « Oui » : Le système est incarné dans le sens où il prend des décisions et raisonne (par exemple, choisir de freiner ou d'accélérer) en fonction de sa perception de l'environnement (caméras, lidars, etc.) — illustrant la boucle Perception-Action mentionnée dans l'exposé.

- Argument « Non » : En revanche, il a été noté que la voiture autonome déployée n'apprend généralement pas en continu par elle-même et seulement avec des mises à jour.

Conclusion : On peut parler d'apprentissage incarné (car l'entraînement se fait à travers des interactions avec des données du monde réel, souvent en simulation) mais des doutes subsistent quant à la présence d'une véritable cognition incarnée au sens où l'on entend une capacité à l'adapter continuellement par l'expérience directe.

3. Lien entre Incarnation et Intelligence Générale Artificielle (AGI)

Le débat s'est conclu sur l'hypothèse soulevée précédemment dans l'exposé : l'IA incarnée est-elle essentielle pour atteindre l'AGI ?

Le consensus était en faveur d'un « Oui, probablement », car bien qu'il soit possible de créer des systèmes très intelligents de manière désincarnée, l'absence d'interaction physique avec le monde rend l'atteinte d'une intelligence véritablement générale (capable de s'adapter et de comprendre) incertaine.

Degré d'Incarnation : La discussion a soulevé la question du niveau d'incarnation requis :

Faut-il simplement des capteurs pour la perception ?

Ou est-il nécessaire que l'IA puisse également agir et interagir physiquement avec son environnement ?