

Synthèse de l'atelier - Le problème du cadre

Introduction : Le *frame problem*, introduit par McCarthy et Hayes en 1969, questionne comment un système intelligent peut décider sans être submergé par des informations inutiles. Ce défi, crucial en IA et en philosophie de l'esprit, interroge la nature de l'intelligence humaine. Dennett y voit une question clé : comment le raisonnement humain reste-t-il pertinent et rapide dans un monde changeant ? Dreyfus, critique des approches classiques, soutient que l'intelligence humaine repose davantage sur l'intuition et l'immersion dans le monde réel.

De nombreuses approches ont été proposées pour résoudre ce problème. Bien qu'aucune d'entre elles ne représente une solution complète, nous pouvons citer parmi ces approches les suivantes :

- **Axiomes de cadre :** cette approche repose sur la déclaration de l'état de chaque élément de l'environnement à travers des énoncés logiques appelés axiomes de cadre. Dans l'exemple du "World of Blocks", nous avons vu que cette méthode est limitée, car elle entraîne une explosion combinatoire en générant de nombreux axiomes pour chaque action, ce qui peut également être monotone et sujet à des erreurs.
- **Explanation Closure (EC) :** contrairement à l'approche précédente, l'EC réduit le nombre d'axiomes en se basant sur les axiomes de clôture, couvrant uniquement les changements dans l'environnement et supposant que le reste demeure inchangé. Cette approche suit une logique monotone qui suppose que, une fois qu'une conclusion est tirée d'une base de connaissances initiale, cette conclusion reste vraie même si la base change, ce qui peut créer des contradictions entre les conclusions. Bien que cette méthode réduise le nombre d'axiomes, dans un système d'information suffisamment grand, générer des explications qui tiennent compte de toutes les données pertinentes peut entraîner une surcharge d'information, rendant difficile l'extraction d'insights utiles pour l'utilisateur.
- **Logique non-monotone:** La logique non-monotone est conçue pour capturer et représenter le raisonnement révisable, qui est une forme de raisonnement où les conclusions peuvent être rétractées à la lumière de nouvelles informations, et où nous faisons souvent des hypothèses basées sur des informations incomplètes et révisons nos croyances à mesure que nous en apprenons davantage, même si les prémisses originales sont conservées.
- **Approches probabilistes et statistiques:** Ces approches utilisent des modèles mathématiques basés sur les théories des probabilités et des statistiques pour gérer l'incertitude, réaliser des prédictions et apprendre à partir des données (Réseaux Bayésiens, Programmation Probabiliste)

Comparaison avec le problème de l'induction et les courants de la cognition :

Lien avec le problème de l'induction : L'induction cherche à généraliser des observations passées, comme déduire que le soleil se lèvera demain parce qu'il l'a toujours fait. Cependant, même si une déduction correcte est possible, il reste complexe d'identifier les connaissances pertinentes pour un contexte donné. Le problème du cadre ajoute à cette difficulté en nécessitant un tri efficace des connaissances adaptées au cas actuel, sous peine de recherche interminable.

Lien avec les visions de la cognition :

- **Cognitivism** : Se base sur la représentation symbolique des connaissances, mais pose la question de quelles informations sélectionner et comment gérer les exceptions. Un agent cognitiviste pourrait se perdre dans des règles rigides, manquant de flexibilité pour réagir en temps réel.
- **Connexionnisme** : Utilise des réseaux neuronaux qui identifient les informations pertinentes par apprentissage, mais peinent à s'adapter aux changements de contexte, car chaque modification exige un réapprentissage.

- **Énactivisme** : Ici, la cognition émerge de l'interaction dynamique avec l'environnement, sans modélisation exhaustive. La pertinence des actions découle du contexte, mais l'agent nécessite une base de connaissances initiale avant de pouvoir agir efficacement.

Points clés du « frame problem » :

- **Pertinence et filtrage** : L'agent doit identifier les informations utiles et ignorer le reste, une tâche complexe pour les machines comparée à l'intuition humaine.
- **Mise à jour des croyances** : L'agent doit ajuster son modèle face à de nouvelles informations sans recalculer tout son environnement, ce qui est particulièrement difficile dans des environnements dynamiques.
- **Raisonnement implicite** : La capacité à mobiliser des connaissances familières et à effectuer des raccourcis contextuels est essentielle mais difficile à programmer. Les approches d'apprentissage peuvent imiter cette capacité, bien que leurs limites d'adaptation restent problématiques.
- **Prise de décision rapide et flexible** : Le cerveau traite les informations en parallèle, une flexibilité difficile à reproduire par les agents qui fonctionnent souvent en série.
- **Délibération et projection** : Face à des situations nouvelles, l'humain sait analyser et planifier, en mobilisant les connaissances appropriées et en anticipant les conséquences de ses actions, une capacité encore limitée dans l'IA actuelle. En bref, le **problème du cadre** concerne l'acquisition de connaissances et le filtrage de celles-ci pour ne retenir que l'essentiel face à un environnement dynamique et complexe.

Débat:

Lors du débat, plusieurs points intéressants ont émergé autour du "frame problem" et de son avenir face aux progrès en intelligence artificielle. Sur la question de sa persistance comme défi fondamental, la majorité s'est mise d'accord qu'il n'y a pas de solution pour le moment mais dans un avenir plus ou moins proche il se peut que l'on trouve une solution grâce au progrès technologique. Pourtant, bien qu'il ne soit pas nécessaire de résoudre ce problème pour créer de l'IA, il reste crucial pour atteindre l'IA forte, capable d'imiter au plus près voir égaler l'intelligence humaine.

Pour répondre à la question de la manière dont les systèmes d'IA pourraient s'inspirer de la cognition humaine pour surmonter le "frame problem", deux points ont été privilégiés. D'une part, l'amélioration des croyances (l'apprentissage de nouvelles connaissances) est considérée comme essentielle, car, combinée au problème de pertinence, elle permettrait de déterminer plus facilement la meilleure décision. D'autre part, il ne suffirait pas de se baser uniquement sur l'expérience ; il serait également crucial d'intégrer l'instinct et le sens commun, propres à l'Homme.

Finalement, une solution proposée lors du débat consiste à trouver un modèle composé de plusieurs approches différentes, comme le cognitivisme et le connexionnisme. Avec les progrès dans divers domaines, il se pourrait que ces approches convergent pour n'en former qu'une seule, un "super modèle" ou méthode universelle capable de prendre le dessus sur toutes les autres et de potentiellement résoudre le "frame problem".