

# Synthèse atelier débat

The frame problem - Le problème du cadre

**BESSY Axel** 11910329

**BOUAMRA Yasmine** 12212284

**LERAY Nicolas** 11809852

IAC - Intelligence Artificielle et Cognition



Université Claude Bernard  Lyon 1

October 30, 2023

# 1 Mise en contexte : le problème du cadre

## 1.1 Introduction

Le travail présenté dans cette synthèse se base sur trois grandes sources de documentation : L'encyclopédie de Stanford de philosophie ; L'article Cognitive Wheels : The frame problem of AI de Daniel C. DENNET et l'article The Frame Problem in Artificial Intelligence and Philosophy de Jarek Gryz.

Le Problème du cadre, tel qu'on l'introduit ci-dessous, est un sujet qui a fait beaucoup débat en intelligence artificielle et en philosophie. C'est une question complexe qui a poussé à la réflexion et au questionnement dans la communauté scientifique en informatique et qui a pu impulser de l'innovation dans le domaine.

## 1.2 Le problème du cadre

En IA, le problème du cadre concerne la définition des règles, des limites et des paramètres selon lesquels un système d'IA fonctionne, et de la définition du monde extérieure pour une intelligence artificielle. Comment définir la pertinence des informations à modéliser et comment modéliser cette propriété même de pertinence. Il s'agit également de spécifier ce que le système peut faire et ne pas faire, et définir sa politique de prise de décisions et de traitement de données. C'est là toute la complexité du problème du cadre : un cadre bien défini est essentiel pour garantir que l'IA fonctionne de manière prévisible, éthique et sûre, mais comment définit-on un cadre approprié ?

En philosophie, le problème du cadre est lié à la question de conscience, de pertinence et de notre capacité à modéliser notre processus de réflexion. On s'intéresse à deux approches en particulier. Daniel Dennett propose une approche de l'intelligence artificielle basée sur le holisme épistémologique. Selon lui, au lieu de se concentrer sur la manipulation computationnelle de l'information, on devrait considérer les connaissances comme un réseau interconnecté. Le contexte est essentiel pour déterminer la pertinence des informations. Pour que l'IA fonctionne efficacement dans le monde réel, elle doit être capable de stocker et de réorganiser ses connaissances de manière explicite, en ne maintenant que les informations pertinentes au moment de l'action. Hubert Dreyfus propose l'approche Nihiliste. Il met en avant trois aspects de l'intelligence humaine que l'IA traditionnelle négligeait : nos mécanismes de réflexion inconsciente, le rôle de nos stimuli physiques et émotionnels ainsi que nos valeurs et sentiments. La méthode Nihiliste implique que la solution réside dans un changement de perspective pour mieux aborder et formaliser le problème du cadre.

Dans l'ensemble, le problème du cadre en IA et en philosophie a été un enjeu crucial sur l'évolution et le développement du modèle et a mis la lumière sur des défauts fondamentaux des modèles d'IA de nos jours.

## 2 Débat

### 2.1 Est-ce nécessaire de résoudre le problème du cadre, ou le contourner est-il suffisant ?

Les avis étaient partagés, selon certains il est nécessaire de le résoudre pour pouvoir avoir des IA réellement compétentes, pour d'autres, le postulat de contourner le problème a déjà permis des avancées. Il a été soulevé à plusieurs reprises que ça peut-être nécessaire mais pas suffisant pour avoir des IA compétentes, car ce n'est qu'une petite brique parmi tant d'autres problèmes qui freignent le développement du domaine.

La question de l'introduction de contexte a été étudiée et les avis qui en sont ressortis étaient que le contexte n'était pas suffisant à la résolution du problème. Le débat a dérivé sur le cadrage du contexte et de l'IA de manière générale et sur le besoin qu'on peut avoir de laisser un IA se développer dans un environnement fermé et de la monitorer jusqu'à ce qu'une certaine confiance soit acquise.

### 2.2 Faut-il résoudre le problème du cadre pour atteindre une IA forte ?

Les avis rejoignent le premier point : ça peut être nécessaire pas suffisant. Beaucoup de problèmes existent au delà du problème du cadre qui nous empêchent de déployer des IA et beaucoup de rapprochements avec les voitures autonomes ont put être fait.

Le problème de modélisation de sens commun a été notamment abordé sur cette question, et des biais que la modélisation d'un sens commun peut engendrer dans une IA et les conséquences que cela peut avoir dans notre société et sur l'acceptation même des IA dans notre quotidien.

### 2.3 Doit-on encoder notre propre sens commun ou l'IA doit-elle développer son propre sens commun ?

Encoder notre propre sens commun est problématique car on retrouve le problème de représentation et d'inclusivité et de biais qui est souvent étudié en IA. Qui peut prétendre à encoder son sens commun ? Quel sens commun encoder ? A-t-on tous le même sens commun ? Serait-il plus judicieux de laisser la machine développer son sens commun ? Doit-on lui donner des informations de bases pour qu'elle puisse le former ? Ces informations vont-elles créer des biais à leur tour ? Peut-on corriger des biais sans en induire d'autres ? La question a été abordé sous divers aspects et la conversation a beaucoup penché sur la question de confiance en ces systèmes et sur le processus d'apprentissage à mettre en place pour qu'une IA puisse avoir un sens commun.

### 2.4 La résolution de ce problème causerait-elle un boom de l'IA ?

Beaucoup d'avis semblait dire que non. Bien que ce problème a été qualifié comme étant aussi dur que le problème  $P = NP$ , il a été jugé que sa résolution ne serait pas un poitn chamboulant pour le domaine, dû aux divers arguments énoncés plus haut. Certains avis

pensaient que oui, de part le manque d'humilité des chercheurs en IA qui pensent avoir trouvé la solution à tous leurs problèmes dès qu'il y a une avancée dans le domaine.