

Synthèse IAC : *Commonsense Problem*

Contexte

Le common sense regroupe des savoirs implicites, souvent intuitifs, qui aident un individu à agir et à comprendre le monde sans passer par un raisonnement formel. Bien que ces connaissances paraissent évidentes, tel que le feu brûle, elles reposent sur une compréhension fine du monde physique, social et causal, acquise à travers l'expérience.

Chez l'humain, ce bon sens se développe à travers l'expérience, mêlant perception sensorielle, actions motrices et échanges sociaux. En effet, dès l'enfance, nos sens, nos actions et nos émotions façonnent notre compréhension intuitive de la réalité. Ce savoir est dit incarné puisqu'il provient du corps, du vécu et des interactions, plutôt que d'une réflexion abstraite.

Cependant, pour une intelligence artificielle, cette capacité est extrêmement difficile à reproduire. Une IA n'a pas de corps pour expérimenter, pas de perceptions sensorielles pour relier les symboles aux objets réels, et pas de vie sociale pour saisir les intentions, les émotions ou les implicites des interactions humaines. Elle apprend à partir de données désincarnées en détectant des régularités statistiques, sans compréhension du monde qu'elles représentent. Ainsi, une IA peut prédire correctement des faits, sans pour autant comprendre la relation de cause à effet, ni le contexte associé. Doter une IA de bon sens ne consiste pas à lui fournir davantage de données, mais à lui permettre de comprendre et d'interagir avec le monde comme un être vivant, à travers l'expérience, la perception et le contexte.

Les approches explorées pour construire un “bon sens artificiel”

Les premières tentatives pour doter les machines d'un bon sens reposaient sur une approche symbolique, où l'on cherchait à représenter explicitement le savoir sous forme de règles logiques. Le projet CYC illustre cette idée. En effet, il visait à encoder des millions de faits du quotidien comme « les humains ont deux jambes » ou « les objets tombent quand on les lâche ». Cette méthode permettait à la machine de raisonner logiquement, mais elle s'est rapidement heurtée à ses limites, car le bon sens ne se résume pas à une accumulation de faits.

Pour pallier cette rigidité, les chercheurs sont passés aux graphes de connaissances, tels que ConceptNet, qui relient les concepts entre eux afin de représenter les associations d'idées humaines. Cette approche offrait plus de souplesse, mais restait incapable de généraliser, ça ne “devine” pas comme un humain.

Avec l'émergence du deep learning, l'idée a alors été d'enrichir les représentations du monde à travers des approches multimodales, combinant texte, image, son ou vidéo, pour donner à l'IA une compréhension plus “incarnée”. En parallèle, les approches interactives, notamment l'apprentissage par renforcement, ont cherché à permettre à la machine d'apprendre par l'action, comme un être humain le ferait en expérimentant. Cependant, ces apprentissages se déroulent encore dans des environnements simulés, bien loin de la complexité du monde réel.

C'est pourquoi des approches plus récentes, dites neuro-symboliques, tentent aujourd'hui de combiner le raisonnement logique et la puissance d'apprentissage des réseaux de neurones afin de tirer parti des avantages des deux modèles. L'objectif est alors de donner à la machine la capacité d'apprendre comme un réseau de neurones tout en raisonnant comme un système symbolique mais cet équilibre reste difficile à atteindre et la compréhension du contexte est encore superficielle/limitée.

Enjeux, limites et applications

Le vrai défi, c'est de donner aux IA un bon sens réel, c'est-à-dire une compréhension du monde qui tienne compte du contexte et de l'expérience, pas seulement des liens statistiques entre les données. Les IA voient bien passer du contexte dans leurs bases d'apprentissage, mais elles ne le vivent pas étant donné qu'elles manipulent des symboles sans percevoir ce qu'ils représentent. C'est pour ça

qu'elles ont du mal à raisonner quand la situation sort du cadre prévu. Si une IA pouvait relier ce qu'elle apprend à une forme d'expérience du monde, elle éviterait certaines erreurs. On imagine par exemple un assistant médical capable d'adapter son diagnostic au ressenti du patient, ou un tuteur virtuel ajustant son explication selon la réaction d'un élève. Mais tant qu'une machine n'aura ni perception ni vécu, ce bon sens restera une illusion, une imitation du nôtre, limitée à des contextes simples.

Débat

Ces discussions ont mis en lumière la complexité à la fois conceptuelle et technique du sujet, mais aussi la diversité des visions parmi les étudiants. Le débat a principalement porté sur deux questions à noter : peut-on réellement formaliser scientifiquement le bon sens, ou est-ce une notion trop humaine ? Et faut-il vraiment chercher à doter les intelligences artificielles de cette faculté ?

Concernant la première question, plusieurs étudiants ont souligné que le bon sens humain reste encore mal compris, ce qui est d'ailleurs vrai même au sein des sciences cognitives. Si nous ne savons pas encore précisément comment le cerveau transforme l'expérience vécue en connaissances de bon sens, il paraît difficile, voire impossible, de le reproduire dans une machine. De plus, il n'existe pas un bon sens universel. Celui-ci dépend de la culture, de l'expérience, de l'environnement et du contexte social. Certains ont mis en garde contre le risque qu'impliquerait une tentative de formalisation, qui reviendrait à imposer une vision particulière du monde. D'autres ont rappelé que le bon sens repose sur de nombreuses connaissances implicites, que nous utilisons sans pouvoir toujours les formuler. Transformer ce type de savoir tacite en données exploitables pour une IA reste donc un véritable défi.

C'est justement cette difficulté qu'a illustrée la professeure lorsqu'elle a dit "je ne vois pas", la plupart d'entre nous avons immédiatement cherché une explication logique, comme un problème de vue, alors qu'en réalité, elle faisait simplement référence à un éblouissement dû au soleil. Cet exemple montrait bien que le bon sens, c'est avant tout la capacité à comprendre le sens caché d'une situation, à relier les indices contextuels sans raisonnement explicite, une compétence que les machines peinent encore à reproduire.

La deuxième question portait sur l'intérêt même de vouloir donner du bon sens aux IA. Deux points de vue se sont opposés. Les partisans du "oui" ont avancé que le bon sens permettrait de rendre les IA plus sûres, plus cohérentes et plus proches du raisonnement humain. Une IA capable de comprendre le contexte éviterait des erreurs absurdes et renforcerait la confiance des utilisateurs dans ses décisions. Les opposants estiment au contraire que le bon sens n'est pas indispensable : les IA actuelles sont déjà performantes dans leurs domaines sans "comprendre" le monde. Selon eux, ce qui compte, ce n'est pas la réflexion interne de la machine, mais la pertinence de son résultat final. Doter une IA de bon sens pourrait donc être une quête coûteuse et inutile si ses performances sont déjà suffisantes pour la tâche visée.

Le débat a aussi mis en évidence une confusion fréquente entre bon sens et éthique. Certains exemples relevaient davantage de la morale que du bon sens. Ce dernier concerne plutôt ce que l'on apprend implicitement par l'expérience donc à une forme d'intelligence pratique et intuitive, et non à une réflexion sur les valeurs.

Conclusion

Pour conclure, le bon sens ne se résume pas à un savoir ou à un ensemble de règles. C'est une forme d'intelligence contextuelle, incarnée et intuitive, qui permet à l'humain de relier la perception, la mémoire, les émotions et la raison pour agir de manière adaptée. Les IA, elles, savent traiter de grandes quantités de données, mais tant qu'elles n'auront ni corps ni véritable expérience du monde, leur bon sens restera une imitation.