

Devoir surveillé d'« Intelligence artificielle »

José MARTINEZ

22 octobre 2024

Important. Vous rendrez les deux parties sur deux copies séparées. **Documents autorisés.** Notes manuscrites personnelles de cours et de TD sur une page A4 recto-verso.

Partie I

A ► CROISSANCE DES LLM

● ● ○ ○ ○

Les grands modèles de langage (ou LLM pour *Large Language Models* en anglais) ont été et sont encore entraînés à partir de sources textuelles disponibles sur la Toile et plus généralement tout système de communication électronique tels que sites HTML, dépôts git, forums, blocs-notes (*blogs*), messageries, etc.

Maintenant, ils sont utilisés pour générer à leur tour du contenu !

Afin de pouvoir différencier l'origine des informations, les grands acteurs du domaine proposent de créer une identité numérique. Elle permettrait de garantir que le producteur d'une information est un être humain ou, au moins, que le contenu généré a été vérifié, voire amendé, par un être humain.¹

Remarque. Nous ne nous intéressons pas ici aux problèmes techniques (sécurité, protection de la vie privée) ni même aux buts *annoncés* (lutte contre la fraude et la désinformation, éthique). Ils s'agit là de régler des problèmes créés par les acteurs eux-mêmes.

- 1 ► Analysez brièvement la raison d'être profonde de cette proposition de « certification d'humanité » par les grands acteurs des LLM ?

B ► UNE RECHERCHE ITÉRÉE

● ● ● ○ ○

Soit A un algorithme de recherche de *profondeur bornée* et renvoyant sinon une solution du moins un meilleur état trouvé.

Construisons un algorithme B qui consiste à effectuer des itérations de A . L'itération initiale part de l'état initial i . Les itérations suivantes partent du meilleur sommet trouvé à l'étape précédente. L'algorithme B s'arrête soit lorsqu'une solution a été trouvée, soit après un nombre maximal d'itérations.

- 1 ► Quels sont :
 - (i) les avantages de cet algorithme ;
 - (ii) ses inconvénients ;
 - (iii) les améliorations possibles ?

1. Louis DREYFUS, Jérôme FENOGLIO, *Intelligence artificielle : un accord de partenariat entre « Le Monde » et OpenAI*, Le Monde, 13 mars 2024 à 18 h 30, 4 avril 2024 à 10 h 11

Partie II

C ► DE L'EXPLORATION

- 1 ► Expliquez le concept du dilemme exploration-exploitation dans le parcours d'un arbre de recherche. Vous illustrerez votre explication avec un exemple.
- 2 ► L'algorithme A^* repose sur l'utilisation d'une heuristique.
Quelle stratégie, et quels conseils, peut-on donner lorsque l'on veut définir cette heuristique sur un problème ?

D ► JEU DES BÂTONNETS

Le jeu des bâtonnets est un jeu simple à deux joueurs, dont les règles sont les suivantes. On démarre la partie avec une série de n allumettes alignées. Lors de son tour, un joueur enlève 1, 2 ou 3 allumettes de cette série. Le joueur enlevant la dernière allumette est désigné perdant.

- 1 ► Quels sont les états et les actions associés à ce problème ?
- 2 ► Dessiner l'arbre de recherche d'une partie pour $n = 7$.
- 3 ► Un joueur en particulier peut-il être assuré de gagner ?
- 4 ► Une variante possible est d'imposer que les allumettes enlevées soient contiguës.
Quelles sont les conséquences de cette modification dans :
 - (i) la modélisation ;
 - (ii) la résolution ?