

En période de « fake news »
Comment évaluer les performances d'un logiciel ?
Quel crédit accorder aux démonstrations ?

Les vraies démonstrations d'IA faites en toute rigueur scientifique sont évidemment indiscutables.

Malheureusement, en période de grandes mutations technologiques où l'IA est à la mode, où l'argent peut donc être levé plus facilement, la malhonnêteté intellectuelle n'épargne pas les scientifiques affairistes.

Epoque de « **fake news** » et de compromissions oblige, d'une part les démonstrations scientifiques sont rares, et d'autre part on a tendance à appeler IA un peu n'importe quoi, notamment ce qu'autrefois on appelait plus modestement « *aide scientifique à la décision* ».

J'ai acheté une brosse à dent électrique qui s'arrête automatiquement lorsqu'on appuie trop fort.

Mention sur la boîte : une réalisation de l'« Intelligence Artificielle » !

Le danger existe donc d'utiliser sans vergogne de fausses « démonstrations » soit pour des raisons commerciales, soit pour valider de fausses théories, soit pour masquer un échec, soit pour des objectifs financiers peu avouables, soit encore pour en faire une opération de Communication et s'attirer les faveurs d'une Presse inexpérimentée ou complice.

Déjà les célèbres « **avions renifleurs** » des années 70 nous avaient alertés du danger des démonstrations non scientifiques destinées à valider des théories fantaisistes pour lever des capitaux, et plus tard en 1997 la supercherie du match d'Échecs **Deep Blue/Kasparov** du danger des « démonstrations » complaisantes, si l'on en croit l'unanimité des experts.

Les impératifs pour qu'une démonstration de bridge soit recevable

Pour qu'elle soit scientifique, et donc recevable, toute démonstration destinée à montrer le niveau de performance d'un logiciel de bridge par exemple, doit impérativement respecter les conditions suivantes :

1. Il ne doit pas y avoir plusieurs jeux stockés sur la machine du logiciel à tester (on l'appellera « Sud ») en dehors du sien, et de celui du mort pour le jeu de la carte.

Ainsi seul sur sa machine, le logiciel à tester ne doit connaître que son seul jeu pour les enchères et, au jeu de la carte, son propre jeu et celui du mort.
(*Si le mort est Sud, on fait une substitution circulaire de 2 positions de façon à ce que Sud récupère le jeu de Nord, Nord celui de Sud, Est celui d'Ouest etc.*.)

2. La machine de Sud ne doit pouvoir communiquer avec aucune autre machine, ni avec un quelconque humain, en dehors des enchères et des cartes que l'on aurait à lui faire connaître.

3. Les donnes jouées peuvent être choisies ou aléatoires, mais elles doivent obligatoirement être de vraies donnes de bridge, complètes, variées, avec enchères, entames et jeu de la carte, à jouer en défense et en déclarant, à Sans-Atout et à la Couleur.

4. Les règles du jeu et le score en vigueur doivent être rigoureusement ceux du bridge, sans restriction, ni modification.

Pas question d'inventer un quelconque score idoine, bricolé pour la circonstance en vue d'arriver au résultat souhaité.

5. Enfin il est obligatoire pour une démonstration scientifique de conclure par un rapport mentionnant et expliquant en détail les succès et les éventuels échecs.

Toute « démonstration » qui ne respecterait pas strictement ces 5 impératifs ne saurait être qualifiée de « scientifique » et ne pourrait donc être considérée autrement que comme une tentative de supercherie.

De fait, il faut bien reconnaître qu'en 2023 aucune démonstration scientifique de bridge n'a encore eu lieu et que la façon la plus simple et la plus probante de démontrer la qualité de jeu d'un logiciel de bridge est bien d'expliquer en détail comment les difficultés ont été résolues.

Quiconque peut alors étudier et évaluer, de bonne foi, sans arrière-pensées, la pertinence de la méthode, la réalité et la portée des arguments.

N'est-ce pas ainsi, par des explications et des calculs plutôt que par une exploration hasardeuse, que Le Verrier a prouvé l'existence de Neptune ?

A une époque où hélas dans tous les domaines triomphent les sophistes, où florissent les « fake news » et les compromissions financières, la richesse technologique actuelle est bien tentante d'oublier morale et souci scientifique.

Voici ce qui pourrait bien se passer, dans quelque domaine que ce soit, pour ceux qui préféreraient les arrière-pensées inavouables à la rigueur scientifique...

Masquer les échecs et tenter de justifier les erreurs de méthode

A propos du danger signalé plus haut des démonstrations destinées à masquer les échecs et imposer des erreurs de méthode, il serait facile d'imaginer par exemple une fable de ce genre :

La distance de la Terre à la Lune est d'environ 400 000 km et l'Homme a réussi à parcourir cette distance le 20 juillet 1969 pour se poser sur la Lune.

50 ans plus tard, après plusieurs années d'échecs à construire une fusée, un savant Cosinus en mal de reconnaissance scientifique et qui cherchait à se donner célébrité (ou tout simplement à lever des capitaux), feignant d'ignorer que le problème avait été résolu depuis belle lurette, annonça tout benoîtement :

- J'ai une méthode qui permettrait d'aller sur la Lune ...

*et, goguenard, il s'empressa d'en informer son entourage.
Il réussit même à appâter quelques banquiers de sa connaissance, tout acquis à l'idée de découvertes importantes qui pourraient bien faire leur fortune.*

Il monta alors sur une chaise et, gagnant ainsi 1 mètre, déclara radieux :

**- « Vous voyez, c'est la bonne méthode : on a progressé.
Convoquons la Presse et faisons une démonstration... »**

*La démonstration évidemment ne démontra rien, mais peu importe.
Il y avait du beau monde, l'ambiance était courtoise et sympathique
La Presse, sans trop comprendre ce qui s'était tramé et dans quel piège elle
était tombée, publia les rapports dithyrambiques qu'on lui avait donnés.
Le Ministre salua comme il se doit ce nouvel exploit de l'Humanité...
Et banquiers d'applaudir.*

Attention même aux mises en situation réelle

Paradoxalement les mises en situation réelles, objets de démonstrations à *priori* indiscutables, pourraient être trompeuses.

Rappelez-vous :

il a fallu plusieurs années aux organisateurs du Tour de France pour démasquer la tricherie du coureur Lance Armstrong.

Les manipulations de la Presse

Dans tous les domaines, manipuler la Presse pourrait aussi devenir si l'on n'y prend garde, le danger des « fausses démonstrations ».

Imaginons par exemple cette fiction :

Il y a plusieurs décennies, une société Sud-Américaine avait été créée pour réaliser un robot capable de jouer au football.

Après plusieurs années de recherche et d'importants investissements, elle dû admettre son échec.

Seule petite consolation, elle avait toutefois connu un succès très modeste et très partiel en réussissant à faire faire à son robot des « touches » plus longues que celles que faisaient les meilleurs joueurs.

L'entreprise décida alors de faire une « démonstration ».

Elle convoqua des journalistes, de préférence ignorants en matière de football, et organisa un concours de touches où son robot était opposé à des footballeurs Sud-Américains champions du monde.

Pour l'occasion elle inventa un score idoine.

Le robot gagna donc le concours et la Presse publia sans vergogne les "éléments de langage" fournis :

« Une IA est championne du monde de football »

Les fausses démonstrations à grand spectacle

Parfois le danger pourrait venir du fait que des sociétés décident de dépenser une débauche de luxe pour organiser des « démonstrations » à grand spectacle pour mieux convaincre une Presse inexpérimentée alors qu'il existe des méthodes scientifiques beaucoup plus simples et plus probantes, ... mais beaucoup moins manipulables.

Exemple :

Pour savoir si un coureur de 100m est de niveau olympique, il n'est pas nécessaire par exemple d'organiser à grands frais, pendant des journées entières, des courses avec pléthore de champions du monde plus ou moins complices :

il suffit de le chronométrer, tout simplement.

Pour évaluer les performances d'un logiciel de bridge il en est de même.

Pour connaître très simplement et sans conteste son niveau de jeu, il suffit de le confronter à des données du plus haut niveau de difficulté (tirées par exemple des célèbres ouvrages de **Hugh Kelsey** : « *Advanced Play at Bridge* » pour le jeu du déclarant et de « *Killing defense at Bridge* » pour le jeu en flanc).

Une telle confrontation réalisée par d'authentiques Champions du Monde et d'Europe, véritable démonstration déjà indiscutable en soi, sera d'autant plus probante si, de surcroît, il est expliqué en détail comment et selon quelle méthode le logiciel est capable de jouer et d'expliquer, à la demande, son raisonnement :

« Pourquoi telle enchère ? Pourquoi pas telle autre ?

Pourquoi telle carte et pas telle autre ? etc. etc. »

grâce à d'importantes **avancées en Intelligence Artificielle**.

C'est la stratégie qu'avait choisie Will-Bridge à la fin des années 80 pour témoigner d'un succès sans conteste dans la résolution complète de la problématique du bridge avec des données les plus complexes de Kelsey :

Confronter et expliquer en détail quelles ont été les difficultés et comment elles ont été résolues, est la méthode la plus simple et la plus probante.

*

En 2023, c'est-à-dire pendant plus de 30 ans, aucune démonstration scientifique de bridge au monde n'a eu lieu qui aurait pu contredire les performances du logiciel de Will-Bridge et la méthode utilisée
(IA analogique et IA hybride).

Ce logiciel « expert » est aujourd'hui le seul à proposer un niveau de jeu humain « champion du monde ».

*.