

Justice aveugle

DENNIS SHASHA

Pour trancher certaines affaires, un juge utilise les probabilités, mais aussi la bonne foi des deux parties opposées.

Un juge versé en mathématiques utilise un étrange moyen pour régler les litiges financiers. Avant de défendre leur cause devant le magistrat, le plaignant et l'accusé estiment le verdict : le plaignant inscrit sur un papier la somme P qu'il pense être en droit de recevoir, puis cache le papier dans une enveloppe scellée ; de la même façon, l'accusé décide de la somme A dont il se sent redevable.

Le juge ignore P et A . Après les plaidoiries de chacun, il rend son verdict : «L'accusé doit la somme J au plaignant à titre de dommages et intérêts!» Puis, il compare cette somme à P et à A . Quand J est plus proche de P , l'accusé doit P . En revanche, lorsque J est plus proche de A , l'accusé paie la somme A . Par exemple, le plaignant a écrit 18 millions

d'euros, tandis que l'accusé pense ne rien devoir. Le juge estime le différend à huit millions d'euros. Puisque 8 est plus proche de 0 que de 18, le plaignant ne touche rien. Ainsi, chacun a-t-il intérêt à être le plus juste dans ses appréciations.

Quelle est la meilleure stratégie du plaignant sachant que le verdict du juge sera compris entre trois et dix millions d'euros, avec une probabilité équivalente pour toutes les sommes? En d'autres termes, combien le plaignant doit-il revendiquer pour optimiser son indemnité? Doit-il changer de stratégie s'il suspecte l'accusé de pouvoir lire la somme P à travers l'enveloppe?

La solution de l'énigme sera donnée le mois prochain, et sera bientôt disponible sur le site Internet : www.pourlascience.com



Solution du problème d'énigmaphys du numéro 296

La femme F frappe D , mais celle-ci ne frappe personne, elle est donc entrée dans la hutte de F . Cette observation est cohérente avec les deux phrases : « C parle à F » et « A parle à F ».

On remarque que F ne frappe personne d'autre que D : C et A ont toutes les deux accusé D .

E frappe A . A est-elle coupable d'intrusion dans la hutte de E , alors qu'elle frappe B qui a parlé à E ? En d'autres termes, B aurait-elle accusé A à tort d'être entrée dans la hutte de E ?

A frappe C qui, elle, ne frappe personne. Aussi, C est-elle entrée dans la hutte de A et a été dénoncée par F .

A frappe B , car B a fait un faux témoignage. C'est la réponse à notre précédente question.

B a frappé A pour une seule raison : E a accusé A d'intrusion. Puisque A frappe ensuite E , A montre qu'il s'agissait d'une fausse accusation.

Les deux dernières bastonnades résultent de la fausse déclaration de A selon laquelle C est entrée dans la hutte de B .

En résumé, voici ce qui s'est passé dans le village :

C est entrée dans la hutte de A .

D est entrée dans la hutte de F .

C dit à F que D est entrée dans sa hutte.

B dit à E que A est entrée dans sa hutte.

A dit à F que D est entrée dans sa hutte.

F dit à A que C est entrée dans sa hutte.

E dit à B que A est entrée dans sa hutte.

A dit à B que C est entrée dans sa hutte.



Les machines pensent-elles?



JEAN-PAUL DELAHAYE

Faut-il comprendre pour imiter? La question est posée par les programmes qui simulent (et stimulent) l'intelligence humaine.

Définir l'intelligence n'est pas chose facile et il n'est pas certain que l'on y parvienne un jour de façon satisfaisante. Cette difficulté poussa le mathématicien britannique Alan Turing, il y a un peu plus de 50 ans, à proposer «le jeu de l'imitation» qui fixait un objectif précis à la science naissante des ordinateurs que l'on n'appelait pas encore informatique.

Le jeu de l'imitation consiste à mettre au point une machine qu'il serait impossible de distinguer d'un être humain. Turing suggérait qu'un juge *J* échange des messages dactylographiés avec, d'une part, un être humain *H* et, d'autre part, une machine *M*, ces messages pouvant porter sur toutes sortes de sujets. Le juge *J* ne sait pas lequel de ses deux interlocuteurs, désignés par *A* et *B*, est la machine *M* et lequel est l'humain *H*. Après une série d'échanges, le juge *J* doit deviner qui est la machine et qui est l'être humain.

Turing pensait que si nous réussissions un jour à mettre au point des machines pour lesquelles l'identification correcte est impossible (c'est-à-dire que le juge *J* se trompe dans 50 pour cent des cas, pourcentage identique à ce que donnerait une réponse au hasard), alors nous pourrions affirmer que nous avons construit une machine intelligente ou – ce que nous considérerons comme équivalent – une machine qui pense. Sa procédure aujourd'hui dénommée *test de Turing* a donné lieu à de nombreuses discussions souvent intéressantes et a suscité une série de réalisations informatiques concrètes mises en compétition annuellement lors d'un prix organisé par Hugh Loebner. Les logiciels participant au jeu de l'imitation progressent chaque année.

Alan Turing (1912-1954)

Face aux prétentions parfois déraisonnables de certains chercheurs en Intelligence artificielle et aux doutes non moins étonnants de ceux qui affirment péremptoirement qu'on ne réussira jamais à créer des programmes intelligents, nous proposons un bilan de 50 ans de réflexion et de travail autour de ce fameux test de Turing.

LE TEST DE TURING ET L'INTELLIGENCE

Avant tout examen, mesurons combien la réalisation d'un programme qui passe le test de Turing est difficile. Mettons-nous dans la peau du juge *J* dialoguant par le biais d'un terminal d'ordinateur avec les deux interlocuteurs, *A* et *B*.

Une première idée pour reconnaître la machine, consiste à poser une question du type : «Combien vaut 429 à la puissance 3?» Si *A* répond 78 953 589 au bout d'une seconde et que *B* refuse de répondre ou attend trois minutes pour proposer un résultat, il ne fera pas de doute que *A* est la machine et *B* l'humain. Évidemment, les spécialistes qui conçoivent les programmes pour passer le test de Turing ont prévu cette ruse grossière. Leur programme, bien que capable de calculer 429³ en une fraction de seconde, refusera de répondre, ou demandera dix minutes avant de fournir le résultat, ou même proposera une réponse erronée. Tout ce qu'un humain ne sait pas très bien faire et qu'un ordinateur réussit sans mal (la mémorisation d'une suite de 100 chiffres aléatoires est un autre exemple) sera traité de la même façon : l'ordinateur fera semblant de ne pas réussir. Pour reconnaître l'être humain, le juge doit donc s'appuyer sur des tâches que les humains traitent sans mal et sur lesquelles les ordinateurs butent.

L'humour est un point d'attaque possible : racontez une histoire drôle à vos interlocuteurs *A* et *B* et demandez-leur

d'expliquer où il faut rire et pourquoi? Les questions d'actualité sur lesquelles nous sommes tous informés peuvent aussi servir de base à une tentative d'identification. L'association des deux difficultés sera obtenue par exemple en demandant un commentaire explicatif du titre de la première page du dernier numéro du *Canard Enchaîné*. Un grand nombre d'autres stratégies envisageables devraient, sans trop de mal, vous conduire à identifier qui est l'homme *H* et qui est la machine *M*. Faire tenir à un ordinateur une conversation traitant de toutes sortes de sujets (de science, d'histoire, d'art, de spectacles, de musique, de gastronomie, etc.) est un objectif extrêmement difficile et l'idée de Turing semble donc bonne : si l'on réussit vraiment à tromper un juge avec un programme, cela signifie indubitablement que l'on a mis dans l'ordinateur quelque chose ressemblant à ce que nous appelons de l'intelligence.

Disons-le tout de suite, nous sommes loin d'avoir atteint le but et ce qui est obtenu avec les meilleurs programmes actuels est un simulacre de conversation qui, malgré des progrès récents, ne réussit pas à entretenir l'illusion plus de quelques minutes.

QUE PROUVERAIT LA RÉUSSITE?

Malgré l'évidente difficulté du problème et notre échec relatif, certains philosophes, dont John Searle, pensent que même si nous réalisons un programme passant le test de Turing cela ne prouvera pas que nous avons mis de l'intelligence dans notre ordinateur. Le raisonnement de J. Searle est le suivant :

- les programmes informatiques sont syntaxiques (ils sont formels, ce ne sont que des outils à manipuler des symboles) ;
- les pensées humaines ont un contenu sémantique (un sens est attaché aux mots que nous utilisons, sens