le centre de masse), soit une énergie dix fois supérieure à celle qui fut accessible au CERN. Deux faisceaux de noyaux d'or sont accélérés en sens opposé sur une trajectoire annulaire de 3,83 kilomètres de longueur. Constitués de paquets successifs d'environ un milliard d'ions, ces deux faisceaux sont précipités l'un contre l'autre dès qu'ils ont atteint 0,99995 pour cent de la vitesse de la lumière. Au cours d'une collision, environ 10 000 particules sont émises dans toutes les directions.

Les expérimentateurs ont commencé par vérifier que la densité d'énergie atteinte dans la matière nucléaire est bien supérieure à la valeur critique. L'opération, complexe, nécessite de résoudre plusieurs problèmes inverses, c'est-à-dire de déduire des données enregistrées par les détecteurs (et qui correspondent aux particules formées lors de la collision), les propriétés physiques de la matière.

A-t-on obtenu d'autres preuves indirectes de la formation d'un plasma de quarks et de gluons? Les quelques centaines de physiciens de la collaboration STAR ont d'abord observé que les particules produites lors de la collision ne sont pas réparties uniformément dans l'espace : elles semblent s'écouler dans des directions privilégiées. Ce caractère anisotrope ne correspond pas à ce que l'on attend si l'on tient seulement compte de la géométrie de la collision. La réorientation des flux de particules sortantes – leur anisotropie – s'explique par les interactions fortes intenses qui se déroulent au sein du milieu, et qui sont sans doute dues à la formation d'un plasma en équilibre thermodynamique.

Les distributions spectrales obtenues au RHIC ont confirmé l'existence d'un équilibre thermodynamique. En effet, quand un gaz est en équilibre thermique, toutes ses particules ont un comportement identique, c'est-à-dire, notamment, que toutes les quantités de mouvement ont la même répartition. Les observations au RHIC confirment que c'est bien le cas.

Fragmentation et jets de particules

La collaboration PHENIX au RHIC a aussi mis en évidence un phénomène nouveau, sans doute dû à la présence du plasma de quarks et de gluons. On a observé que la production de particules ayant la plus grande quantité de mouvement transverse est atténuée dans les collisions frontales. Aux énergies mises en jeu au RHIC, une collision d'ions lourds est, en première approximation, un choc de quarks et de gluons. Lors de la collision, ces derniers sont diffusés à des angles d'autant plus grands que le choc est violent. La composante transverse de la quantité de mouvement (perpendiculaire à l'axe) est évaluée à partir des angles de diffusion. Une fois qu'il s'est échappé de la zone d'interaction (le volume de collision), un quark ou un gluon donne naissance, par un mécanisme de «fragmentation», à un jet de particules. En effet, les quarks qui quittent le volume de collision ne peuvent rester isolés (un quark n'est jamais observé seul) : ils «s'hadronisent», c'est-à-dire qu'ils donnent naissance à des hadrons. L'énergie des hadrons du jet reflète celle des quarks et des gluons initiaux, et elle est observée par les détecteurs. Cependant, avant de s'échapper de la zone d'interaction, les quarks et les gluons de grande impulsion rencontrent sur leur chemin la matière nucléaire chaude et dense (le plasma?) formée par la collision. Or, un quark qui traverse une telle matière perd d'autant plus d'énergie que le milieu traversé est coloré. En d'autres termes, il est bien plus freiné dans la matière de quarks que dans la matière hadronique. Dans les détecteurs, cela se traduit par un affaiblissement de l'énergie des particules sortantes : c'est précisément le phénomène nouveau que les physiciens ont observé.

RHIC, Brookhaven

5. L'ANNEAU DU COLLISIONNEUR du Laboratoire de Brookhaven, nommé RHIC, est enfoui dans un tunnel à quatre mètres de profondeur. Des paquets d'ions d'or s'y déplacent à une vitesse très proche de celle de la lumière, en suivant un anneau de près de quatre kilomètres de circonférence.

Le plasma de quarks et de gluons a-t-il, pour autant, été découvert? La matière déconfinée existe-t-elle? En 2001, les expériences du RHIC ont accumulé dix fois plus de données qu'en 2000, et la récolte continue. Les résultats du RHIC semblent indiquer que la matière confinée existe bel et bien, mais demandent à être renforcés. L'étape suivante de la quête du plasma de quarks et de gluons aura lieu au CERN. Un grand détecteur est d'ores et déjà en cours de réalisation, et sera employé sur le faisceau d'ions lourds du grand collisionneur LHC en construction au CERN. Les noyaux de plomb seront accélérés dans deux anneaux de 27 kilomètres de circonférence, qui permettront d'atteindre une vitesse égale à 0,99999993 pour cent de celle de la lumière!

Chacune des collisions libérera une énergie de 1 100 téraélectronvolts (1 téraélectronvolt vaut 10^{12} électronvolts), suffisante pour porter la matière nucléaire concentrée à quelque 1 000 milliards de kelvins. Cette énergie correspond à 0,2 millijoule, soit l'énergie d'un moustique en vol, mais elle est concentrée dans un volume de quelques milliers de fermis cubes seulement : elle est considérable. Le millier de physiciens et d'ingénieurs de 27 pays qui préparent cette nouvelle campagne expérimentale sera alors en mesure d'explorer le domaine d'existence du plasma de quarks et de gluons, d'autant plus qu'au grand collisionneur, le phénomène observable durera dix fois plus longtemps qu'au RHIC. Gageons que cela suffira pour révéler, sans contestation possible, le déconfinement et pour l'étudier sous toutes ses facettes.

Yves SCHUTZ et Hugues DELAGRANGE sont directeurs de recherche au Laboratoire SUBATECH de Nantes, une unité mixte de recherche du CNRS (IN2P3), de l'Université de Nantes et de l'École des mines de Nantes. Ils collaborent aux expériences PHENIX au RHIC et ALICE au CERN.

Madhusree MUKERJEE, *Un Big Bang miniature*, *Pour la Science*, n° 259, mai 1999.

K. H. ACKERMANN et col., *Elliptic Flow in Au+Au Collisions at sqrt(snn) = 130 GeV*, Collaboration STAR, in *Phys. Rev. Lett*. 86, 402, 2001.

K. ADCOX et col., *Suppression of Hadrons with Large Transverse Momentum in Central Au+Au Collisions at sqrt(snn) = 130 GeV*, Collaboration PHENIX, in *Phys. Rev. Lett*. 88, 022301, 2002.

Justice aveugle

DENNIS SHASHA

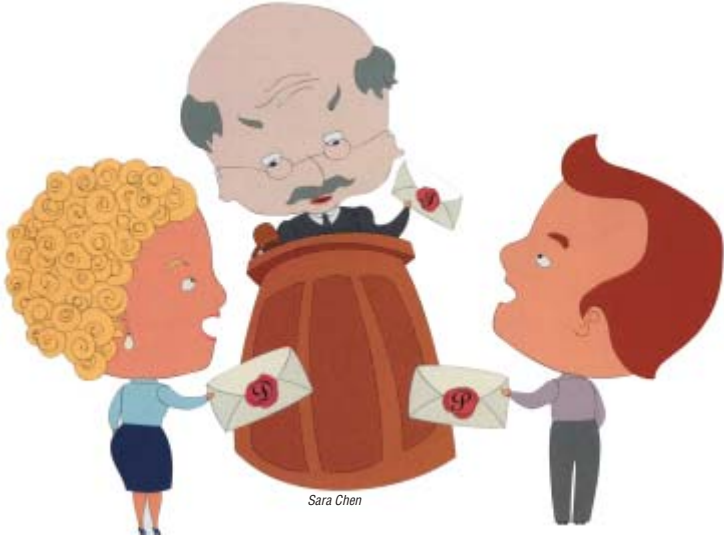
Pour trancher certaines affaires, un juge utilise les probabilités, mais aussi la bonne foi des deux parties opposées.

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



Sara Chen

Solution du problème d'énigmaths 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 réussissons 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 pourrons 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 429^3 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 buttent.

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

que l'on ne peut ramener à de la syntaxe formelle. Ce sens provient de propriétés biologiques particulières de nos neurones qu'un programme, de par sa nature, ne peut pas posséder) ;

– donc les programmes ne penseront jamais et ne seront jamais intelligents.

Cette façon de voir décourage toute tentative concrète. Pour J. Searle, la réussite au test de Turing n'est pas une condition suffisante pour qu'on puisse qualifier les ordinateurs d'intelligents.

Nous n'entrerons pas dans les controverses qu'engendre ce type d'analyses. L'article original où, en 1980, J. Searle exposait son point de vue, était accompagné de 26 commentaires critiques, et depuis une multitude de discussions se sont déroulées. Remarquons cependant qu'il paraît un peu étrange d'affirmer qu'un système, que rien ne distinguerait dans un dialogue d'un être humain, ne possède pas – l'intelligence – qu'on attribue aux humains. Il y aurait comme un privilège accordé par décret, à une catégorie particulière d'êtres – les humains – qu'on refuserait définitivement à une autre catégorie d'êtres, les ordinateurs. J. Searle manque de fair-play : si un jour les ordinateurs réussissent à nous imiter, alors nous devrons alors avoir l'élégance de les qualifier d'intelligents.

Cette critique de la position de J. Searle est qualifiée de *béhavioriste* : on propose de fonder un jugement uniquement sur les comportements sans entrer dans l'analyse même des processus internes (dans le cerveau et dans l'ordinateur) qui produisent les comportements. Cependant comme le font remarquer A. Saygin, I. Cicekli et V. Akman, cette accusation de béhaviorisme méritera d'être prise au sérieux le jour où, nous humains, disposerons d'autres moyens que l'analyse du comportement pour évaluer l'intelligence des êtres humains !

Le cri de la mouette

Une autre critique, sans doute plus subtile, a été formulée par Robert French à l'encontre du test de Turing. Alors que J. Searle considère que le test de Turing n'est pas une condition suffisante pour qu'on puisse affirmer qu'un ordinateur pense, R. French soutient lui que passer le test de Turing n'est pas une condition nécessaire d'intelligence.

R. French imagine un peuple qui ne connaîtrait qu'une seule espèce d'oiseaux : les mouettes. Ce peuple se pose le problème de réaliser une machine volante et pour savoir s'il a réussi utilise le *test de la mouette* : une machine est dite volante s'il est impossible de la distinguer d'une mouette dont le comportement est observé à l'aide d'un radar. Le radar limite la précision de la demande d'imitation comme le dialogue, par le biais d'échanges dactylographiés, limite la portée de l'imitation dans le test de Turing. Bien sûr les avions, les hélicoptères, les montgolfières et même les oiseaux autres que les mouettes ne réussiront pas le test de la mouette et ne seront donc pas considérés comme capables de voler. Est-ce bien raisonnable ? Certes non.

Pour R. French, le test de Turing est une condition suffisante d'intelligence, mais d'intelligence humaine. De plus, ce test est lié à la langue utilisée pour les dialogues, ce qui empêche de le considérer comme universel. R. French poursuit son analyse et affirme qu'il se pourrait bien que nombre de nos comportements dépendent fortement de la façon particulière dont notre cerveau traite l'information au niveau le plus profond (il qualifie ces processus de subcognitifs). Imposer une imitation servile du comportement humain c'est se protéger à trop bon compte du risque d'admettre que des machines intelligentes sont devenues nos concurrentes. Pour illustrer son idée, R. French imagine un exercice d'évaluation des néologismes. Il propose qu'avant d'appliquer le test de Turing, on fasse évaluer la qualité d'un ensemble de néologismes par un groupe de sujets humains de référence. Par exemple, on envisage les néologismes pour désigner de très longues voitures. Les sujets humains notent de 0 à 10 les néologismes : auto-fusée, fusauto, auto-basset, longivoiture, finauto, auto-aiguille, voiture-serpent.

Les notes attribuées constituent des données types et en comparant ensuite, lors du test de Turing, les notes proposées par *A* et *B* à ces données, on distinguera l'humain de la machine, car – c'est l'idée de R. French – la manière dont nous notons les néologismes est liée à la façon particulière dont notre cerveau traite le langage, manière qui a peu de chances d'être identique à celle qu'utiliserait l'ordinateur (de même : la façon dont un humain joue aux échecs est très différente de celle utilisée par un programme). L'évaluation des néologismes démasquerait l'ordinateur. Pour R. French, de même qu'il n'est pas nécessaire d'être indiscernable d'une mouette pour voler, il n'est pas nécessaire d'être indiscernable d'un humain pour penser ou avoir un comportement intelligent.

Cependant pour R. French et la majorité des gens qui ont réfléchi à ces questions, il est certain que même si la difficulté du test de Turing est très grande (trop sans doute), la réussite à ce test ne laisserait aucun doute sur le fait que nous avons construit des machines intelligentes. Il semble donc intéressant d'essayer. Qu'avons-nous obtenu ?

1. LE TEST DE TURING

Dans le jeu de l'imitation proposé par Alan Turing en 1950, le juge essaie de reconnaître, en dialoguant par écrit avec un interlocuteur, si celui-ci est une machine ou un être humain. Le juge ne voit pas son interlocuteur à qui il pose des questions et dont il obtient des réponses. La machine essaie de se faire passer pour un être humain et l'être humain essaie de se faire reconnaître comme tel. Pour Turing, si le juge se trompe dans 50 pour cent des cas, alors on doit considérer que la machine est intelligente et qu'elle pense. La photographie représente les juges dialoguant par ordinateur avec des hommes ou des programmes lors du prix Hugh Loebner.

Si la majorité des gens s'accordent pour dire qu'une machine passant le test de Turing doit être considérée comme intelligente et qu'elle dispose de la capacité de penser, Robert French, lui, considère que le test de Turing est trop difficile et ne caractérise pas l'intelligence en général, car il est trop lié au langage utilisé, à la façon dont fonctionne le cerveau humain au niveau le plus bas (subcognitif) et aux traits culturels particuliers de celui qui interroge.

R. French imagine ainsi un peuple qui, ne connaissant qu'un seul animal volant, la mouette, croit définir un critère absolu pour la propriété de voler en formulant le *test de la mouette* : on considérera qu'un objet est capable de voler, si, lorsqu'on observe son image à l'aide d'un radar, il est indiscernable d'une mouette. Il est clair que ce test manque son but, car si, incontestablement un objet passant le test de la mouette doit être considéré comme volant, en revanche aucun avion, ni aucune autre espèce d'oiseaux ou d'insectes ne réussiront à passer le test de la mouette qui est trop difficile. Il en va de même, pour French, de la réussite au test de Turing : c'est une condition suffisante d'intelligence, mais pas une condition nécessaire.

Richard Wallace *(à gauche)*, vainqueur avec son programme ALICE du concours de Hugh Loebner *(à droite)*.

2. TESTS PARTIELS DE TURING

Certains tests de Turing partiels sont déjà possibles.

Test de Turing de simulation d'un patient paranoïaque

Les réponses d'un patient paranoïaque sont suffisamment stéréotypées pour qu'on puisse faire des simulations informatiques indiscernables de ces malades. Dès 1971, M. Colby a mis au point un programme réussissant à tromper une fois sur deux des psychiatres chargés de reconnaître les patients humains des patients simulés par la machine.

Test de Turing limité au jeu d'échecs

Même un bon joueur d'échecs ne peut plus aujourd'hui savoir s'il joue avec une machine ou un humain.

Test de Turing limité aux calculs élémentaires

En posant des questions de calcul élémentaire du type :
23 + 45?
552 + 128?
dérivée de $\sin(3\,x^2)$?
solutions de l'équation $4x^4–3x^2–1=0$?
simplification de $(x^4–9)/(x^2–3)$?
Un juge ne peut pas distinguer un homme d'une machine, car certains programmes comprennent et répondent correctement à ce genre de questions. Les domaines du calcul pouvant être abordés par ces programmes sont de plus en plus nombreux.

Le test de Turing peut ainsi être envisagé sous des formes partielles. Dans certains cas les progrès de la recherche en Intelligence artificielle ont été tels qu'il est devenu impossible de distinguer dans ces contextes restreints un être humain d'une machine. En revanche, dans le cas envisagé par Turing où tous les sujets de conversations sont autorisés, aucun programme n'est en mesure de tromper un juge plus de quelques minutes. Remarquons que, de même qu'un programme qui calcule trop bien et trop vite se fait repérer, un programme d'échecs peut lui aussi se faire démasquer en jouant certains coups étranges. Un cas célèbre illustre ce point : deux programmes d'échecs jouaient l'un contre l'autre et, au cours de la partie, un des programmes met sa tour en prise. Stupeur dans l'auditoire parmi les grands maîtres qui regardaient le déroulement de la partie. Jusqu'à ce que l'un d'eux remarque après quelques dizaines de minutes que si le programme n'avait pas mis sa tour en prise, il était échec et mat en 9 coups. Nul doute qu'un joueur humain, même s'il avait vu la menace, aurait joué «psychologique» et pensé que son adversaire ne verrait pas un mat aussi long, et n'aurait pas mis sa tour en prise, car alors il était alors certain de perdre la partie. Pour passer les tests de Turing limités, il faut donc concevoir les programmes de manière à ce qu'ils répondent mal aux questions que les humains ne peuvent pas traiter en un temps limité.

LES PROGRAMMES DE CONVERSATION

À l'époque où Turing proposait son test, il n'était pas envisageable de le mettre en œuvre matériellement. Les interfaces des ordinateurs étaient trop frustres et surtout les mémoires et la rapidité de calcul des machines étaient insuffisantes (environ dix millions de fois inférieures à celles d'aujourd'hui!), ce qui interdisait l'écriture de programmes manipulant le langage naturel.

Dans la décennie 1960, des programmes capables de dialogues en langues naturelles sont apparus (on les nomme *chatbots* ou *chatterbots* pour «robots de conversation»). Un de ces programmes est le célèbre ELIZA de Joseph Weizenbaum mis au point en 1964 au MIT. Le mécanisme interne du programme est assez simple. Les phrases de l'interlocuteur sont analysées, et, à partir de leur structure grammaticale, une réponse est construite en utilisant certains des mots de l'interlocuteur. Quand le programme se trouve embarrassé (l'analyse grammaticale de certaines phrases peut-être très difficile, le vocabulaire prévu peut être insuffisant, etc.), il utilise des phrases passe-partout pour détourner la conversation ou pour esquiver les questions. Le but d'ELIZA n'était pas de passer le test de Turing, mais de jouer le rôle d'un psychanalyste attentif à ce que dit son interlocuteur et l'incitant à se confier. Bien qu'il fût conçu comme un jeu sans prétention, certaines personnes utilisant ELIZA, et informées de sa véritable nature se prirent au jeu et se confièrent à lui comme s'il s'agissait d'un être humain. Des utilisateurs passionnés prétendirent même s'y être attachés.

Un programme tenant sur trois pages – ce qui est très peu – écrit en langage BASIC et s'inspirant d'ELIZA proposé par Richard Schomberg et Jean-Pascal Aubert permettait le dialogue suivant en français.

Programme. Je suis ELIZA. Quel est votre problème?
Utilisateur. Je n'aime pas les ordinateurs.
Programme. Est-ce que les ordinateurs vous ennuient?
Utilisateur. Oui.
Programme. Pourquoi m'interrogez-vous?
Utilisateur. Pour obtenir des solutions à mes problèmes.
Programme. Est-ce que cette question vous intéresse?
Utilisateur. Oui.
Programme. Quelle réponse vous ferait le plus plaisir?
Utilisateur. C'est à vous de deviner.
Programme. Nous étions en train de discuter de vous, pas de moi.
Utilisateur. Cela ne résout pas mes problèmes.
Programme. Que pensez-vous?
Utilisateur. J'ai peur.
Programme. Pouvez-vous être plus explicite sur ce qui vous effraie?
Utilisateur. C'est difficile.
Programme. Racontez m'en un peu plus sur vos problèmes…

On le voit, rien de bien subtil, et, à moins de laisser le programme mener la conversation et d'être d'une extrême indulgence, on aura du mal devant les quelques trucs syntaxiques et phrases préparées (contenues *in extenso* dans le programme) à imaginer que l'on dialogue avec un être intelligent. En 1971, un programme du même type simulait, non plus un psychanalyste, mais un malade paranoïaque. Le programme possédait des états mentaux et, selon le contexte, était placé dans une situation de peur, d'anxiété ou de méfiance qui fixait son type de réponses. Une expérience où on demandait à des psychiatres de distinguer le programme de véritables patients donna le résultat de 48 pour cent d'identifications correctes, ce qu'on peut voir comme une sorte de première réussite au test de Turing.

LE PRIX HUGH LOEBNER

En 1990, Hugh Loebner avec le soutien du *Cambridge Center for Behavioral Studies* a créé un concours dont l'objectif est de susciter de véritables tentatives pour passer le test de Turing. H. Loebner s'engagea à récompenser d'un prix de 100 000 $ et d'une médaille d'or le premier ordinateur dont les réponses sont indiscernables de celles d'un humain. H. Loebner, personnage étrange, précise

avec insistance que, contrairement aux médailles distribuées aux gagnants des Jeux olympiques, sa médaille n'est pas plaquée or, mais en or massif. Chaque année, un prix de 2 000 $ et une médaille de bronze sont aussi attribués au programme «le plus humain» pris parmi les programmes se présentant au concours, et cela quel que soit le niveau atteint par le programme, aussi mauvais soit-il.

Les 100 000 $ ne seront attribués que pour un test de Turing où le programme et l'interlocuteur dialoguent en utilisant des données textuelles, sonores et graphiques. Pour un test de Turing limité à des échanges textuels (comme l'envisageait pourtant Turing), seul un prix intermédiaire de 25 000 $ et une médaille d'argent sont proposés.

Aujourd'hui la compétition pour les 2 000 $ se déroule de la manière suivante. Les programmes présentés ainsi que quelques humains mêlés à eux sont numérotés. Un panel de juges dialogue alors par écrit avec ces interlocuteurs numérotés. Quand il le souhaite, le juge rend son avis en attribuant une note de 0 à 5 à chacun des participants, homme ou programme. Le programme qui a obtenu les meilleures notes gagne la médaille de bronze et les 2 000 $. Jusqu'à présent les juges ont trié sans mal leurs interlocuteurs humains et leurs interlocuteurs informatiques (moins bien notés). Toutefois, en 2001, un juge a donné une meilleure note à un programme qu'à un humain! Cette erreur d'un juge n'est pas suffisante pour emporter les prix de 25 000 $ et 100 000 $ qui ne seront attribués que lorsqu'un programme réussira de manière répétée à se faire passer pour humain.

Certains chercheurs ont jugé que l'organisation du prix Loebner était inutile et même nuisible au progrès des recherches en Intelligence artificielle. Le prix détourne, disent-ils, les chercheurs d'objectifs plus raisonnables et les incite à accumuler, dans leurs programmes, des astuces qui ne sont finalement que de simples bases de données conçues pour faire semblant d'être intelligent. La véritable compréhension de l'intelligence ne progressera pas ainsi, argumentent ces critiques.

Marvin Minsky (l'un des pionniers de l'Intelligence artificielle) a d'ailleurs offert en 1995 un prix de 100 $ pour quiconque dissuaderait H. Loebner de maintenir son offre. H. Loebner a alors malicieusement déclaré que puisque la seule façon de le faire renoncer à l'organisation de son prix était de le gagner, il en résultait que *de facto* M. Minsky était cosponsor de son prix dont le montant était donc passé à 100 000 $ + 100 $. En voulant s'opposer au prix, M. Minsky l'avait renforcé!

LE MEILLEUR PROGRAMME CONTEMPORAIN : ALICE

Ces deux dernières années, le prix Loebner a été emporté par un programme nommé A.L.I.C.E. *(Artificial Linguistic Internet Computer Entity)* dont l'auteur principal est Richard Wallace (c'est ce programme qui lors du dernier concours a trompé un juge en se faisant classer devant un humain).

Le développement du programme ALICE est l'œuvre d'une collectivité de passionnés (300 personnes y auraient participé) et le programme, qui n'est pas secret, est mis à la disposition de toute personne intéressée. Si vous le souhaitez, vous pouvez essayer ALICE, et mieux, vous pouvez le modifier ou l'adapter pour créer votre propre logiciel de conversation (voir http://www.alicebot.org/). Le programme de conversation en ligne sur le site publicitaire du film A.I de Steven Spielberg a été obtenu à partir d'ALICE.

La conception d'ALICE dans son essence n'est pas différente de celle d'ELIZA et d'ailleurs, contrairement à d'autres chatbots, aucune théorie linguistique, aucun modèle d'apprentissage automatique, aucune représentation interne du monde ne sont intégrés au programme de R. Wallace. Comme souvent en Intelligence artificielle, la méthode de la force brute semble mieux réussir que l'utilisation d'idées théoriques élaborées. Une étude détaillée des dialogues entre les utilisateurs et le programme améliore ce dernier : tout le fonctionnement du programme est centré sur le repérage de la forme de la question de l'utilisateur et de l'association d'une

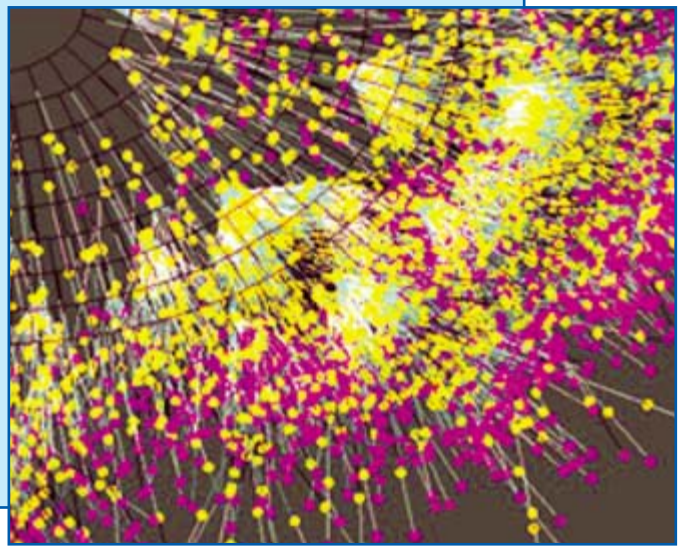
3. LE CERVEAU D'ALICE

Le programme ALICE qui a gagné des deux derniers prix Loebner est fondé sur les questions qu'un utilisateur pose à un programme de conversation. Les informations sur ces questions sont mémorisées sur une spirale où sont placés, par ordre de fréquence d'utilisation, les premiers mots des questions des utilisateurs, puis les seconds mots, etc. La disposition en spirale s'impose, car le premier mot est pris dans un ensemble de possibilités assez grand (2 000 cas) alors que les autres mots (les premiers étant fixés) ne sont pris que parmi un petit nombre de possibilités (en moyenne 2).

Aujourd'hui cette spirale que Wallace voit comme le cerveau d'ALICE comporte 24 000 «catégories». Il s'agit en fait des questions type que ALICE sait traiter. Bien que très volumineuse cette information est organisée sous une forme qui rend improbable qu'ALICE devienne intelligente au sens général : aucun modèle du monde, aucune capacité de déduction, aucune organisation complexe des connaissances ne peuvent être stockés dans cette structure du «cerveau d'ALICE».

Le mot le plus fréquemment utilisé comme premier mot d'une question posée à ALICE est YES. Viennent ensuite NO, OK, WHY, etc. Selon R. Wallace, la fréquence de ces mots se conforme à la loi de Zipf : la fréquence d'utilisation d'un mot est inversement proportionnelle à son rang. Voici la liste des premiers mots ou groupes de mots apparaissant le plus fréquemment en tête d'une question posée à ALICE (avec le nombre d'occurrences rencontrées jusqu'à présent).

8024 YES
5184 NO
2268 OK
2006 WHY
1145 BYE
1101 HOW OLD ARE YOU
946 HI
934 HOW ARE YOU
846 WHAT
840 HELLO
663 GOOD
645 WHY NOT
584 OH
553 REALLY
544 YOU
531 WHAT IS YOUR NAME
525 COOL
516 I DO NOT KNOW
488 FUCK YOU
486 THANK YOU
416 SO
414 ME TOO
403 LOL
403 THANKS
381 NICE TO MEET YOU TOO
375 SORRY
374 ALICE
368 HI ALICE
366 OKAY
353 WELL



Le cerveau d'ALICE

réponse ajustée à la forme reconnue. Le grand nombre de formes envisagées fait la force d'ALICE : il sait répondre à une variété de questions de plus en plus grande. Le programme se comporte mieux avec un utilisateur moyen qui pose des questions déjà rencontrées qu'avec un expert plus imaginatif et informé des points faibles du programme. C'est pourquoi R. Wallace suggère que le panel de juges qui choisit le gagnant du prix Loebner soit composé de personnes prises au hasard, plutôt que d'universitaires et de spécialistes diplômés en Intelligence artificielle.

R. Wallace ne semble pas réellement gêné par le peu de complexité réelle de son programme et il explique : «ELIZA a parfois été considéré comme une farce, et c'en est peut-être une, mais doit-on en déduire que ce programme n'est pas de l'intelligence artificielle ? Ne devons-nous pas plutôt nous préparer à l'inévitable conclusion : la conscience elle-même est une illusion. C'est sans doute parce qu'il avait entr'aperçu cela que Turing a proposé un test fondé sur l'aptitude à simuler. La frontière entre l'Intelligence réelle et l'illusion n'a jamais été aussi floue.»

R. Wallace livre ensuite le fond de sa pensée : «Nous n'avons pas besoin d'une théorie complexe de l'apprentissage, des réseaux de neurones ou des modèles cognitifs pour expliquer comment bavarder dans les limites des 25 000 cas envisagés par le programme d'ALICE. Notre conception stimulus-réponse est une bonne théorie et c'est certainement la plus simple.» R. Wallace parle du «cerveau d'ALICE» expliquant que son programme est une sorte de copie de sa propre personnalité. «En allant visiter le contenu du cerveau d'ALICE, écrit-il, vous y rencontrerez le monde culturel des livres que j'ai lus, des films que j'ai vus, vous y trouverez mes citations et plaisanteries préférées, etc.»

Finalement R. Wallace défend l'idée qu'en continuant à ajouter des schémas de réponses préparées au cas par cas, on finira par avoir une machine indiscernable d'un humain. «Le domaine du langage couvert par notre programme contient déjà la plus grande partie des phrases que les gens utilisent en dialoguant avec une machine. En étendant encore un peu ce domaine, nous allons cerner les mots récalcitrants, et le moment arrivera où le plus sévère des critiques humains ne pourra plus concevoir une seule phrase pour coincer ALICE». À l'entendre, l'image souvent utilisée pour dénoncer la naïveté de certains chercheurs en intelligence artificielle revient à l'esprit. R. Wallace ne se comporte-t-il pas comme quelqu'un qui réussissant à monter sur un tabouret dirait : «Vous voyez bien que je progresse, d'ici peu j'atteindrai la Lune ?»

L'enthousiasme du maître d'œuvre du programme qui a gagné les deux derniers prix Loebner est sans limite. Il imagine que l'industrie des logiciels de conversation va d'ici peu connaître un développement considérable et que ces logiciels seront au cœur des prochaines interfaces informatiques. Pourtant à ce jour, seule la médaille de bronze de 2 000 $ a été décernée et je ne crois pas prendre beaucoup de risques en affirmant que les sommes de 25 000 $ et de 100 000 $ resteront dans les caisses des organisateurs du prix Loebner pour de longues années encore.

Notons qu'un pari de 20 000 $ portant sur la question : «Un ordinateur passera-t-il le test de Turing avant 2029?» a été pris entre Mitchell Kapor (qui soutient que non) et Ray Kurtzweil (qui soutient que oui). N'oublions pas non plus que Turing dans son article de 1950 avait prédit que son test serait passé en l'an 2000, ce qui ne s'est pas produit.

Même si certains *chatbots* sont conçus pour imiter des personnes humaines ayant existé (vous pourrez dialoguer sur internet avec John Lennon, Elvis Presley ou Jésus) les spécialistes les plus optimistes ne se risquent pas aujourd'hui à parier pour un succès avant 2029. Puisqu'il semble qu'on se limite à des programmes qui, comme ELIZA en 1964, sont sans véritable ambition, la date réelle d'un succès définitif pourrait être plus lointaine encore.

4. ALICE ET LANGUE DE BOIS

Richard Wallace l'auteur principal du programme ALICE qui a gagné les deux derniers prix Loebner explique que les techniques utilisées dans son programme s'inspirent des méthodes des politiciens (ce qu'on appelle en France la langue de bois). Il écrit : «Généralement un politicien ne donne pas de réponse directe. Si le journaliste qui l'interroge pose une question, le politicien répond avec un discours tout préparé ayant un rapport parfois lointain avec la question. Les réponses semblent activées par des mots-clefs. Le mot *école* déclenchera par exemple un discours sur l'enseignement. Je suis persuadé qu'un robot de conversation serait un politicien infatigable. Le président Clinton a donné un magnifique exemple de discours robotisé quand un jour il a répondu en disant «It depends on what the meaning of "is" is.» («Cela dépend de ce qu'est le sens de "est"»), ce qui convenons-en est une esquive vraiment passe-partout.

Jean-Paul DELAHAYE est professeur d'informatique à l'Université de Lille. Son dernier livre *L'intelligence et le calcul*, éditeur Belin-Pour la Science, est un recueil de chroniques parues dans *Pour la Science*. e-mail : delahaye@lifl.fr

R. FRENCH, *Subcognition and The Limits of the Turing Test*, in *Mind*, 99, pp. 53-66, 1990.

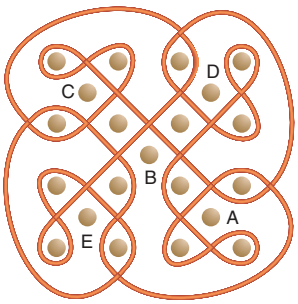
A. P. SAYGIN, *The Turing Test Page* : http://cogsci.ucsd.edu/~asaygin/tt/ttest.html

J. R. SEARLE, *Minds, Brains and Programs*, in *Behav. and Brain Sc.*, 3, pp. 417-57, 1980.

A. TURING, *Computing Machinery and Intelligence*, in *Mind*, 59, 236, pp.433-460, 1950. Disponible en ligne : http://www.loebner.net/Prizef/TuringArticle.html

Pages du prix Loebner : http://www.loebner.net/Prizef/loebner-prize.html

Site du programme ALICE : http://www.Alicebot.org/



Les mystérieux sona

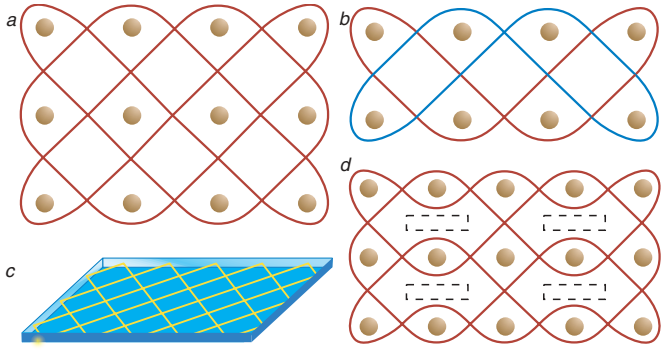
AVIDA DOLLARS

Les dessins qu'utilisent les conteurs d'un peuple d'Afrique centrale pour illustrer leurs histoires offrent quelques énigmes de combinatoire aux mathématiciens.

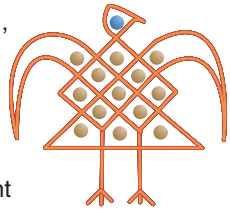
Jusqu'à la fin des années 1950, les membres des tribus du peuple Chokwe, qui compte encore un million d'habitants dans la région de Lunda, au Nord-Est de l'Angola, se réunissaient après une journée de chasse autour d'un feu et écoutaient l'un d'eux, l'akwa kuta sona, raconter des contes selon un rituel précis *(voir la figure 4)*. Après avoir nettoyé et lissé le sol sableux de sa main, il dessinait une grille de points, puis au fil de sa narration, son doigt traçait autour de ces points une ligne courbe qui servait de support à son histoire. Par exemple, la figure à côté du titre illustre une fable où un lapin, en A, découvre une mine de sel, en B. Hélas, celle-ci est l'objet de convoitises de la part d'un lion, en C, d'un guépard, en D et d'une hyène, en E, chacun arguant de la loi du plus fort pour mettre la main sur le trésor. Heureusement, à la fin de l'histoire, le lapin à gain de cause et reste le seul à avoir accès à la mine, les autres en étant séparés par la courbe.

Ces dessins nommés sona (au singulier, un lusona) appartenaient à une longue tradition : ils illustraient des proverbes, des fables, des jeux, des animaux *(voir le médaillon)*, des énigmes et jouaient un rôle important dans la transmission du savoir et de la sagesse aux jeunes générations. Le tracé des dessins devait être lisse et continu, sans à-coup : toute hésitation ou arrêt était un signe d'imperfection et d'un manque de savoir-faire que l'audience sanctionnait d'un sourire ironique.

La grille initiale de points facilitait la mémorisation des dessins par l'akwa kuta sona, ou expert en sona. Le nombre de colonnes et de lignes dépendait du motif souhaité et de l'histoire. Par exemple, les marques laissées sur le sol par un poulet que l'on poursuit étaient représentées par un dessin dont la grille de départ avait cinq lignes de six points. Grâce à cette méthode – un exemple ancien de système de coordonnées –, l'akwa kuta sona réduisait la mémorisation d'un lusona entier à celle de deux nombres, celui des lignes et celui des colonnes.



1. Les motifs de tissages sont les sona les plus simples : ils sont formés de courbes qui parcourent les diagonales d'une grille de points. La plupart sont composés d'une seule courbe *(a)*, mais certains en nécessitent plusieurs *(b)*. Ces dessins sont analogues au trajet d'un rayon de lumière dans une enceinte où les parois sont des miroirs *(c)*. Lorsqu'on insère des miroirs supplémentaires *(en pointillés noirs)* entre les points des colonnes de rang pair, on obtient les sona de la famille du ventre du lion *(d)*.



Les sona qui ressemblent aux tissages sont parmi les plus simples *(voir la figure 1)*. Parmi eux, certains sont constitués d'une seule courbe continue qui contourne chaque point de la grille avant de rejoindre son point initial, d'autres requièrent le tracé de plusieurs courbes. Mark Schlatter, de l'Université de Shreveport, en Louisiane, a analysé les propriétés mathématiques des sona aux motifs de tissages, ainsi que ceux d'une autre famille, nommée ventre du lion, et mis en évidence dans quels cas une seule courbe suffit à dessiner le lusona entier, et dans quels autres, plusieurs sont nécessaires. En d'autres termes, étant donné le nombre de lignes et de colonnes, combien de courbes sont nécessaires pour dessiner un motif complet ?

Tout d'abord, s'inspirant des idées de Paulus Gerdes, professeur de mathématiques à l'Université de Maputo, au Mozambique, il imagine le dessin comme la trajectoire d'un rayon lumineux dans une enceinte rectangulaire dont les parois sont des miroirs. La lumière est émise d'un coin de l'enceinte dans une direction faisant 45 degrés avec l'un des côtés. Ainsi, la lumière se réfléchit sur les parois en parcourant les diagonales de la grille.

Puis, dans une première étape, il établit qu'une grille carrée de n points de côté requiert n courbes *(voir la figure 2)* : en effet, lorsqu'on introduit un système de coordonnées pour repérer les points de la grille, une courbe qui part du point $(a, 0)$ atteint le point $(n + 1, n + 1–a)$ puis le point $(n + 1–a, n + 1)$ et enfin le point $(0, a)$ avant de rejoindre le point initial. Chaque courbe rebondit trois fois avant de revenir à son point initial et parcourt ainsi quatre diagonales : au total, on dessine n courbes pour parcourir toutes les diagonales de la grille.

Ensuite, le mathématicien décompose un motif de tissage donné en carrés *(voir la figure 2)* et montre que le nombre de courbes requises pour le motif entier est égal à celui du motif restant, ou au nombre de courbes du dernier carré quand la grille est uniquement constituée de carrés. Par exemple, un motif de trois lignes et quatre colonnes se décompose en une grille carrée de neuf points et en une ligne de trois points qu'une seule courbe suffit à parcourir. Dans le motif entier, cette dernière courbe et les trois courbes de la grille carrée n'en forment qu'une seule. En revanche, pour un motif de deux lignes et quatre colonnes, on ôte un carré de deux points de côté et on laisse un carré identique : deux courbes sont nécessaires pour dessiner un motif de tissage sur une telle grille carrée de quatre points ; deux courbes sont donc aussi nécessaires pour le motif entier. Quand la grille restante après la séparation du premier carré est encore importante, on réitère la même opération jusqu'à l'obtention d'un motif simple.

Ce processus est analogue à la détermination du plus grand commun diviseur de deux nombres donnés, ici, le nombre de colonnes et le nombre de lignes. Lorsque le découpage