

Herschel ne comprenait pas vraiment les causes de ces phénomènes, et il avait une capacité limitée à formuler mathématiquement ses découvertes. Ses atouts étaient la connaissance pratique de l'optique, combinée à son habileté en matière de fabrication d'instruments et ses talents d'observateur.

Son dernier article, présenté le 6 novembre 1800, contient le premier graphique représentant les distributions spectrales de la lumière visible et du rayonnement infrarouge (voir la figure 5). Il nomma ces courbes le « spectre d'éclairement » et le « spectre de chaleur ». Sur l'axe vertical, Herschel reporta la température mesurée et la luminosité perçue. Il leur fixa des maximums égaux afin de comparer l'étendue relative et la forme des distributions ainsi que la localisation de leur maximum. Pour l'axe horizontal, la notion de longueur d'onde étant encore inconnue, il utilisa la distance par rapport à l'endroit où les couleurs visibles éclairaient la table. Cet axe est inversé par rapport à la convention actuelle, qui consiste à représenter les longueurs d'onde croissantes de gauche à droite.

Un graphique trompeur

Le graphique de Herschel est le produit de son imagination, de son intuition et de mois de travail laborieux, mais il était trompeur. Son allure générale est sans doute le facteur décisif qui le conduisit à conclure que la lumière et la chaleur radiante sont, après tout, fondamentalement différentes.

La représentation des données peut influencer sur leur interprétation. Même aujourd'hui, il est difficile d'examiner le graphique de Herschel sans avoir l'impression que la lumière et la chaleur radiante sont deux entités distinctes. Les courbes de Herschel sont toutes deux exactes, mais elles correspondent à des quantités différentes et presque sans rapport, et ne devraient pas être représentées ensemble. Son erreur ne vient pas de ses données de base, mais du fait qu'il supposait les courbes comparables.

L'un des éléments manquants pour l'interprétation des courbes, et qui était un concept inconnu en 1800, est l'effet de la réponse du détecteur. La mesure réalisée

par tout instrument ou détecteur, y compris l'œil, résulte du produit de deux courbes : la distribution de la puissance reçue, et la courbe de réponse de l'instrument. Herschel supposait que le spectre mesuré représentait directement la distribution du rayonnement, sans imaginer que l'œil déformait ce signal par sa réponse.

Tout spectre créé à partir de ce que l'œil enregistre s'annule en dehors des limites de la vision. À l'intérieur de ces limites, la puissance reçue à n'importe quelle longueur d'onde est pondérée par la sensibilité de l'œil à cette couleur. Le pic de sensibilité, comme Herschel l'a correctement déterminé, est le jaune-vert à une longueur d'onde de 0,555 micromètre.

Par conséquent, le spectre d'éclairement de Herschel est davantage une carte de la façon dont l'œil répond aux couleurs que de la distribution de la lumière. Il voyait dans la forme de sa courbe une preuve que la lumière n'est pas également répartie sur le spectre. Son hypothèse était correcte, mais ses données ne le montraient pas. Il aurait obtenu une courbe pratiquement identique même si la distribution avait été uniforme.

En novembre 1800, Herschel est arrivé au bout de ce qu'il était possible de réaliser avec les connaissances et les techniques de son époque. Il retourna sans hésitation à son domaine de prédilection : l'astronomie.

L'apparence trompeuse de son graphique et les différentes sensations que sont la vision de la lumière et la sensation de chaleur l'ont finalement emporté sur ses données soigneusement recueillies, l'amenant à conclure que les deux énergies ne sont pas une seule et même grandeur physique.

Néanmoins, notre compréhension actuelle du rayonnement électromagnétique s'est développée à partir des mesures simples de température que Herschel a réalisées dans la lumière solaire. Elle s'est poursuivie avec l'unification du spectre par James Clerk Maxwell en 1861 et la formulation de la théorie quantique au XX^e siècle. Herschel ne pouvait pas prouver de façon concluante que la lumière et la chaleur radiante sont de même nature, mais ses expériences ont apporté des indices que d'autres allaient compléter pour résoudre l'énigme. ■

Le télescope spatial Herschel

L'Agence spatiale européenne a lancé en 2009 un télescope spatial nommé Herschel en l'honneur de l'astronome. Cet instrument observe le rayonnement infrarouge dans l'Univers. Il permet d'étudier la formation des nouvelles étoiles et l'évolution des galaxies. L'image ci-dessous est un nuage moléculaire (en fausses couleurs) au sein duquel se forment des étoiles.



NASA/Herschel

■ BIBLIOGRAPHIE

J. Heilbron, *Weighing imponderables and other quantitative science around 1800*, *Historical Studies in the Physical Sciences*, vol. 24, pp. 127-131, 1993.

D. J. Lovell, *Herschel's dilemma in the interpretation of thermal radiation*, *Isis*, vol. 59(1), pp. 46-60, 1968.

W. Herschel, *Experiments on the solar, and on the terrestrial rays that occasion heat*, *Philosophical Transactions of the Royal Society*, vol. 90, pp. 293-326 et 437-538, 1800.

W. Herschel, *Experiments on the refrangibility of the invisible rays of the Sun*, *Philosophical Transactions of the Royal Society*, vol. 90, pp. 284-292, 1800.

LOGIQUE & CALCUL

L'homme, meilleur joueur que la machine

Grâce à de nouvelles méthodes de coordination, l'intelligence humaine produit parfois des résultats supérieurs à ceux des programmes informatiques.

Jean-Paul DELAHAYE

En mai 1997, le lendemain de la victoire de l'ordinateur *Deep Blue* sur Gary Kasparov, le quotidien français *Libération* titrait « L'homme est-il foutu ? » Aux échecs, l'ordinateur bat aujourd'hui le meilleur humain et cela semble irréversible. Cette supériorité de la machine concerne de nombreux autres jeux dits de stratégie. En fait partie le jeu de dames anglaises, pour lesquelles l'ordinateur sait calculer des stratégies optimales, c'est-à-dire définitivement imbattables, ce qu'il ne sait pas faire dans le cas des échecs où les programmes jouent bien, mais d'une façon perfectible (voir l'article de cette rubrique de janvier 2008).

Cependant, le champion humain reste meilleur que la machine au jeu de go. Le niveau « sixième dan » atteint en 2012 par le programme *Zen19D* a inquiété des professionnels, mais il ne fait pas de l'ordinateur un égal de l'homme à ce jeu.

Moins attendu, l'ordinateur Watson d'IBM a battu l'être humain en 2011 au jeu *Jeopardy* (du type *Questions pour un champion*), dont on pensait que la nature non combinatoire et fondée sur des énigmes linguistiques exigeait des connaissances générales et une intelligence hors de portée des machines. L'ordinateur a montré que l'intelligence artificielle progresse même là où certains pensaient qu'elle resterait bloquée. Bien sûr, la base de données de 15 téraoctets de Watson, qui lui donne accès instantanément à une quantité colossale d'informations, a contribué au

succès de l'ordinateur et s'il a gagné, c'est en utilisant des algorithmes de traitement du langage naturel très différents des procédures du cerveau humain. Comme aux échecs, la machine gagne, mais elle n'imité pas.

L'ordinateur progresse aussi dans la résolution d'autres tâches où sa logique binaire semblait le handicaper. Il y a la résolution automatique des mots-croisés, où le programme *Dr Fill* approche les meilleurs joueurs humains. Pour la conduite automobile, les progrès sont tels que dans certains États américains, des autorisations ont été accordées pour que des systèmes automatisés de conduite circulent sur les routes au milieu des autres véhicules... à la condition qu'ils soient surveillés par un humain pouvant reprendre le contrôle à tout instant.

Or de nouvelles méthodes d'utilisation et de coordination de l'intelligence humaine produisent des résultats étonnants, et même des renversements inespérés ! Plusieurs problèmes dont la nature combinatoire semblait déterminer une implacable victoire de la machine sur l'humain voient celui-ci produire de meilleures solutions que les algorithmes les plus perfectionnés fonctionnant sur les plus puissants des systèmes.

À la condition de transcrire les problèmes sous de bons formats (ce que l'on fait avec l'ordinateur), les cerveaux humains sont dans ces cas de meilleurs calculateurs que les ordinateurs. L'indispensable présentation sous la forme de jeu, et les interfaces informatiques graphiques complexes pour

que l'intelligence ludique des humains produise ces prouesses, suggèrent que c'est l'association ordinateur-humain qui résout les problèmes posés. Une telle analyse serait injuste, car l'ordinateur n'est qu'un serviteur qui présente les données, exécute les commandes de manipulation formulées par le joueur humain, et évalue s'il progresse vers la solution ou l'atteint : toutes les décisions sont prises par le joueur.

Plier des protéines

La connaissance de la suite linéaire des nucléotides d'une séquence génétique codant une protéine, une suite de lettres prises parmi A, C, T et G, détermine la séquence des acides aminés qui composent la protéine et son repliement en trois dimensions. Cette forme dans l'espace détermine les parties actives de la molécule et comment la protéine interagit avec les autres molécules. Connaître ce repliement est donc essentiel au biologiste. En première approximation, c'est un problème de minimisation d'énergie, mais les degrés de liberté sont si nombreux (plusieurs dizaines, parfois beaucoup plus) que la détermination du bon repliement est extrêmement difficile ; les programmes informatiques spécialisés n'y réussissent aujourd'hui qu'assez médiocrement.

Un programme interactif en ligne dénommé *Foldit* (« Replie-le ») a été développé conjointement par le *Center for Game Science* de l'Université de Washington et le

Département de biochimie de la même université. Ce programme, que l'on télécharge facilement (<http://fold.it/portal/>), propose des énigmes de repliement. Une molécule dessinée en trois dimensions est affichée à l'écran, et le joueur peut alors la faire tourner, en examiner les parties de près et en manipuler les diverses composantes. Il est donc en mesure de la replier en essayant de faire au mieux. Le joueur est guidé par des informations diverses, dont bien sûr le niveau d'énergie qu'il doit minimiser. Il peut revenir en arrière, agiter aléatoirement la molécule, en figer certaines parties, etc. Il n'est pas nécessaire de connaître la biochimie pour tenter sa chance, car ce qui est

important sera appris par des exercices d'introduction et de découverte. Le but d'un joueur est d'améliorer l'énergie du meilleur repliement connu.

Le programme *Foldit*, qui fonctionne depuis 2008, a aujourd'hui été utilisé par 240 000 joueurs enregistrés. Les meilleures structures produites par les joueurs sont mémorisées, examinées et testées par les experts du domaine. Pour obtenir de bons résultats, les joueurs sont incités à se grouper en équipes. Des tableaux de meilleurs scores sont maintenus à jour, et diverses formes de compétitions entre joueurs et équipes accroissent les motivations des participants.

En 2011, une équipe de joueurs dénommée *Contenders* a donné la forme exacte de la protéase d'un rétrovirus responsable d'une des formes du sida chez le macaque rhésus. Il s'agissait d'un problème non résolu depuis une quinzaine d'années. La solution de l'équipe *Contenders* a été publiée dans la revue *Nature Structural and Molecular Biology*. Le nom de l'équipe apparaît comme auteur de la publication parmi les autres chercheurs ayant travaillé sur le projet *Foldit*. Le repliement découvert aidera peut-être à la mise au point de nouveaux antiviraux s'appliquant au sida humain.

Le jeu a aussi été utilisé pour modifier des protéines et améliorer leur efficacité

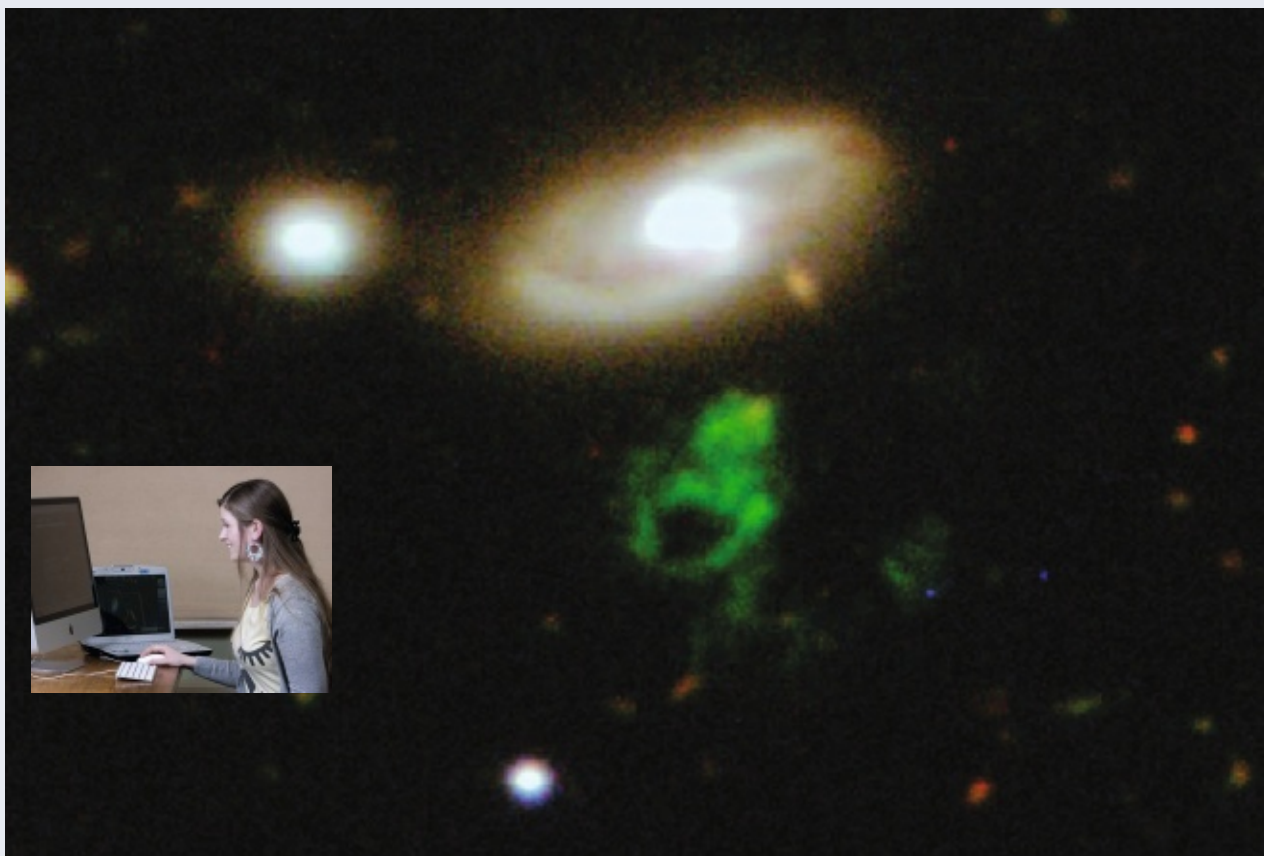
1. La grenouille de *Galaxy Zoo*

Les données recueillies par le télescope spatial *Hubble* sont trop nombreuses pour être traitées par les astronomes et trop complexes pour que la classification des galaxies soit

faite par un algorithme. Seul le recours à l'intelligence des bénévoles permet de mener le travail de classification nécessaire et rend disponibles aux chercheurs des données

scientifiques importantes (projet *Galaxy Zoo*). Durant ce travail de fouille des données portant sur des millions de galaxies, un étrange objet a été repéré par une internaute néerlandaise

prénommée Hanny (à gauche). L'objet, ici représenté en fausse couleur verte, a été nommé *Hanny's Voorwerp*. Il suscite conjectures et spéculations chez les astronomes.



Nasa et Université d'Arizona, *Galaxy Zoo Team*

2. Le jeu *Foldit*



Un joueur dispose à chaque instant d'une information sur le score (c'est-à-dire l'énergie qu'il faut minimiser) de la configuration moléculaire affichée. Pour améliorer ce score, le joueur a le choix entre plusieurs types d'actions sur la molécule à replier : saisir un élément de la molécule et le déplacer ; secouer (le squelette, les chaînes, ou tout) pour voir si cela améliore le score ; disposer quelques élastiques ou en enlever, ce qui aide à mieux contrôler la structure, etc. Un cahier de « recettes » (à gauche sur le côté) est aussi à sa disposition : chaque recette est une succession d'actions prédéfinies dont l'exécution a été reconnue efficace.

Étonnamment, les joueurs et les équipes de joueurs réussissent souvent à mieux replier les molécules que les meilleurs algorithmes spécialisés.

Foldit wiki

dans une fonction donnée. En janvier 2012, une enzyme qui catalyse la réaction dite de Diels-Alder a ainsi été perfectionnée, augmentant sensiblement son efficacité. Selon Seth Cooper, coordinateur de ce travail, « les gens possèdent des facultés naturelles de raisonnement spatial que les ordinateurs n'ont pas encore ».

Découvrir aussi des algorithmes

Récemment, les mécanismes de jeu du programme *Foldit* ont été complétés pour que les joueurs ou groupes de joueurs puissent définir des méthodes de repliement générales, autrement dit des algorithmes, et puissent les échanger, les copier et les faire évoluer. Ces « recettes » évolutives et en compétition ont amené de fécondes interactions et conduit à 5 400 propositions résultant de combinaisons de recettes élémentaires et de modifications variées des meilleures recettes trouvées précédemment. Deux recettes eurent un succès particulier auprès des joueurs. On remarqua alors des similitudes entre les deux méthodes et un algorithme récemment mis au point par des spécialistes de la programmation du repliement des protéines. Cet algorithme dépasse les meilleurs algorithmes publiés.

Les auteurs de l'article où est décrite cette découverte concluent que « les jeux scientifiques en ligne, non seulement résolvent des problèmes difficiles, mais

découvrent et formalisent des algorithmes nouveaux et efficaces ».

Ainsi, des joueurs utilisant leurs facultés naturelles de manipulations d'objets en trois dimensions et leur bon sens battent les meilleurs algorithmes et en conçoivent de nouveaux d'une complexité et d'une efficacité équivalentes à ceux que les spécialistes imaginent. Certes, les joueurs qui ont participé à ces expériences sont nombreux et, pris dans l'excitation de la compétition, certains ont consacré beaucoup de temps, d'énergie et d'ingéniosité pour arriver à ces exploits.

On ne peut s'empêcher de penser que les milliards d'heures perdues dans le monde à jouer à toutes sortes de jeux informatiques deviendraient utiles si plus de jeux scientifiques de cette nouvelle catégorie étaient pratiqués... Zoran Popovic, le directeur du *Center for Game Science*, énonce un objectif audacieux : « Notre but ultime est que toute personne puisse jouer et soit donc en mesure de gagner un prix Nobel. »

Un autre problème difficile de biologie computationnelle est celui de l'alignement multiple des séquences de molécules biologiques de type ADN, ARN ou protéines. On cherche à disposer ces séquences les unes sous les autres, de manière à ce qu'apparaissent les régions similaires et que ces zones de concordance soient, autant que possible, les unes sous les autres. On utilise des programmes qui maximisent le nombre de coïncidences entre les séquences. L'introduction de trous (c'est-à-dire de ruptures

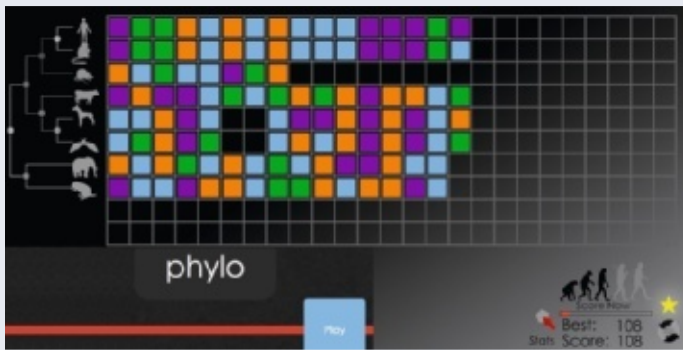
dans les séquences) est permise. Le résultat final est représenté par un tableau noté de la manière suivante : [a] deux séquences ayant un élément en commun au même endroit confèrent un point ; [b] si, en un endroit donné, deux séquences diffèrent, cela retire un point à la note du tableau ; [c] l'introduction d'un trou coûte cinq points et l'agrandissement d'un trou déjà existant en fait perdre un.

Les règles étant fixées, le calcul du meilleur alignement ne serait qu'une routine : en essayant tous les alignements possibles (en nombre fini) et en les notant tous, on déterminerait celui obtenant la meilleure note. Malheureusement, ce problème simple à énoncer est « complexe » à résoudre et aucun moyen algorithmique efficace n'est connu, car il appartient à la classe des problèmes NP-complets, classe de problèmes que personne ne sait traiter en temps raisonnable quand la taille des données augmente.

L'alignement des séquences est important pour déterminer la fonction des protéines et la phylogénie. L'alignement sert à identifier les zones fonctionnelles des protéines qui sont bien conservées lors de l'évolution et correspondent donc aux parties bien alignées des tableaux. Lorsqu'on détecte l'homologie d'une protéine A avec une protéine B de fonction connue, l'alignement suggère les fonctions de la protéine A.

Une seconde utilité des alignements multiples est liée à la recherche des parentés

3. Le jeu *Phylo*



Le programme *Phylo* transforme en jeu le problème de l'alignement de séquences génétiques. Le joueur déplace des cases colorées pour obtenir le plus grand nombre de couleurs identiques en colonnes. Des grands alignements sont examinés par le programme qui y repère des zones susceptibles d'être mieux alignées. Ces zones sont traduites en puzzles colorés.

On présente au participant l'un de ces problèmes en réserve. S'il trouve un meilleur alignement, celui-ci est inséré dans le grand alignement dont a été extraite l'énigme. De nombreux alignements ont ainsi été perfectionnés grâce aux joueurs qui, comme dans le cas de *Foldit*, obtiennent par leurs capacités visuelles et leur intelligence des résultats supérieurs à ceux des algorithmes.

entre espèces : quelle espèce a un ancêtre commun proche avec telle autre, et comment les espèces ancestrales se disposent-elles en arbre [équivalent aux arbres généalogiques] ? Un alignement multiple entre éléments d'une famille de protéines aide à calculer l'arbre phylogénétique le plus probable entre les protéines, qui est aussi un arbre reliant les organismes porteurs des protéines alignées. Les phylogénies ainsi reconstituées sont d'autant plus exactes que les alignements sont précis et optimaux. Toutes sortes d'algorithmes sont utilisés, complétés, perfectionnés et mis en commun par les chercheurs pour calculer des alignements multiples et en déduire des arbres phylogénétiques.

Alexander Kawrykow et ses collègues du Centre de bio-informatique de l'Université McGill à Montréal, inspirés par *Foldit*, ont mis au point un programme interactif nommé *Phylo* (<http://phylo.cs.mcgill.ca>) qui transforme les problèmes d'alignements multiples en jeux et propose à toute personne, même si elle n'a aucune connaissance de biologie, de rechercher de bonnes configurations. Des tableaux colorés sont affichés à l'écran et les joueurs les manipulent pour rechercher des alignements. Le programme indique à chaque instant la note du tableau affiché qu'il s'agit de faire progresser en déplaçant de petits carrés, dont il est inutile de savoir ce qu'ils représentent.

Encore plus facile à utiliser que *Foldit*, le programme *Phylo* s'est révélé d'une grande utilité. Les problèmes d'alignements soumis

aux utilisateurs concernent des séquences de taille limitée de façon à ce que l'œil humain s'y repère bien. Si un problème soumis à un joueur est résolu et obtient une note meilleure que celle de l'alignement connu au préalable (qui a été calculé par un algorithme rapide, produisant des alignements de qualité moyenne), alors l'alignement du joueur remplace celui qui était jusqu'alors enregistré dans la base de données.

Les organisateurs du jeu affirment que 70 pour cent des alignements soumis aux joueurs ont été améliorés et ont donc été remplacés par les résultats des joueurs humains. Après six mois de mise en place, le jeu avait attiré plus de 12 000 internautes qui avaient produit plus de 300 000 alignements variés. Comme le signalent les chercheurs de l'Université McGill, cela prouve à nouveau que le talent visuel des joueurs en ligne dépasse celui des algorithmes connus.

Une méthode d'avenir ?

Le programme *Phylo* et son usage sont gratuits. Un biologiste qui souhaite en bénéficier afin d'améliorer un alignement important pour ses travaux l'utilisera très facilement. Il suffit pour cela qu'il envoie l'alignement à améliorer à une adresse électronique donnée en respectant un format convenu ; le tableau sera transformé en jeu graphique et si les joueurs améliorent l'alignement initial, la solution graphique trouvée sera retraduite dans les notations

des biologistes, puis insérée là où il faut dans le grand alignement soumis, en même temps qu'il sera envoyé au chercheur à l'origine de la demande (voir la figure 3). Les méthodes de *Foldit* et de *Phylo* n'ont-elles d'intérêt que temporairement, ou permettront-elles longtemps à l'homme de battre la machine ?

Il paraît raisonnable de soutenir que dès qu'on aura mieux compris les problèmes, ou analysé comment travaille le cerveau humain et qu'on l'aura copié, ou encore dès qu'on disposera d'une puissance de calcul supérieure, alors le talent humain cessera de faire le poids face à la machine électronique, comme cela s'est produit pour le jeu d'échecs. *Foldit* et *Phylo* seront alors sans intérêt.

À l'inverse, on peut argumenter que les circuits neuronaux présents dans les aires visuelles du cortex sont le résultat d'une optimisation très fine après des dizaines ou centaines de millions d'années d'évolution ; en conséquence, ces circuits fonctionnent d'une manière extraordinairement subtile qui restera longtemps inaccessible à notre compréhension. Autrement dit, nous ne saurons pas programmer leurs mécanismes dans le silicium avant des lustres. En allant plus loin, on peut même imaginer qu'il sera impossible de comprendre dans le détail ce que font notre cerveau ou certaines de ses parties car, pour les comprendre, il faudrait y pénétrer et donc les détruire... Il est ainsi impossible de conclure aujourd'hui sur l'avenir de *Foldit* et de *Phylo*.

Transformer en jeu un problème scientifique réel et difficile est une belle idée mise en œuvre depuis peu, mais une idée proche l'a anticipée. Un groupe de chercheurs peut demander de l'aide à l'extérieur pour réaliser un travail qu'il n'arrive pas à effectuer seul, car cela exige trop de temps ou d'énergie. C'est ce qui s'est passé avec le projet *Galaxy Zoo*.

Des millions de galaxies

Pour comprendre la formation et l'évolution des galaxies, il est nécessaire de les classer selon leur forme et de disposer de nombreux exemples. Une classification sommaire distingue les galaxies spirales et les galaxies elliptiques. Un spécialiste disposant d'une image de galaxie indiquera assez rapidement son type et en classera quelques centaines sans se donner beaucoup de mal. La difficulté provient du nombre de galaxies que les chercheurs veulent prendre en compte (plusieurs millions), car il est important pour eux de

disposer de classifications aussi exhaustives que possible. Une première version du projet *Galaxy Zoo* (<http://www.galaxyzoo.org>) a permis, grâce aux internautes volontaires, de classer un million de galaxies. Les images provenaient d'une caméra numérique installée à l'Observatoire d'Apache-Point, au Nouveau-Mexique (États-Unis). Sans la mobilisation des 80 000 bénévoles, la caractérisation des 60 millions de galaxies repérées par le télescope spatial *Hubble* aurait été impossible.

Par ailleurs, lors de ces travaux d'observation et de classification, un objet étrange a été identifié par une internaute, Hanny van Arkel. L'objet, nommé *Hanny's Voorwerp*, est mystérieux et a suscité discussions et conjectures chez les spécialistes, prouvant à nouveau le bien-fondé d'un examen aussi exhaustif que possible, par des humains, des données astronomiques aujourd'hui disponibles en quantité... astronomique.

Si l'intelligence humaine réussit souvent mieux à calculer ou à classer que la machine, c'est que le problème a été bien traduit et mis en images pour que la recherche de sa solution devienne amusante ou très simple (dans le cas de *Galaxy Zoo*) grâce à un logiciel interactif mis en ligne et conçu dans ce but. Cela est possible depuis que les performances graphiques des ordinateurs de bureaux ont atteint les extraordinaires niveaux actuels.

Une autre raison du succès de ces méthodes est l'existence des réseaux informatiques et de la communication instantanée qu'ils permettent entre les organisateurs du travail et une multitude de joueurs potentiels dispersés dans le monde. Grâce à ces progrès technologiques, une masse importante de joueurs tente sa chance et fournit un travail irremplaçable. L'idée est de confier à une « foule » une tâche de calcul qu'on ne sait pas obtenir des machines et qu'un humain seul a peu de chance de mener à bien. L'intelligence humaine bat l'ordinateur dans ces exemples parce qu'elle s'est massivement mobilisée et coordonnée.

Cette idée de s'adresser au grand nombre pour obtenir ce qu'on ne réussit pas seul ou avec ses machines n'est pas

4. Karpov et Kasparov battent la foule

L'idée de confier à la foule une tâche trop difficile pour un être humain, et dont les exemples présentés dans l'article montrent qu'elle fonctionne souvent à merveille, a pu donner l'illusion que le *crowdsourcing* était apte à tout faire.

Il n'en est rien : dans un match sur Internet en 1996, Anatoli Karpov écrasa son adversaire, « Le reste du monde ». Toutefois, les coups étaient décidés à la majorité et certains joueurs

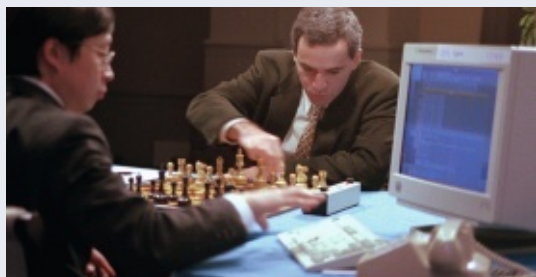
faibles votaient pour les mauvais coups.

Le match qui a été organisé en 1999 entre Gary Kasparov et « 50 000 bons joueurs » était plus difficile et a prouvé à nouveau que la foule, dans un tel cas, se fait battre par un individu seul. Alors que 50 000 personnes provenant de 75 pays différents et conseillées par quatre joueurs professionnels, dont Étienne Bacrot, collaboraient pour jouer et battre G. Kasparov en une fantastique

partie d'échecs, ce fut lui qui, en définitive, l'emporta après des échanges étendus sur 62 coups (voir www.chessgames.com/perl/chessgame?gid=1252350). Puisque l'ordinateur *Deep Blue* abattu G. Kasparov aux échecs, et que G. Kasparov a battu la foule, doit-on en conclure que, nécessairement, *Deep Blue* ou l'un de ses successeurs battrait lui aussi la foule ?

A. Karpov rejoua en juillet 2012 une partie majoritaire et gagna en 37 coups (www.europe-echecs.com/art/parte-majoritaire-contre-karpov-fin-de-la-partie-majoritaire-de-karpov-4421.html).

Ni les échecs ni la science ne sont démocratiques : ce n'est pas la majorité qui gagne. En science, le novateur a raison contre les habitudes de pensée de tous.



akora.com

5. Le service reCAPTCHA

Pour certaines opérations, on vous demande de retranscrire quelques symboles mal écrits formant un mot. Le but est d'empêcher un programme des substituer à vous et, par exemple, de créer massivement des comptes de courrier électronique. On suppose que seul un être humain peut traduire le dessin présenté en symboles. Comme ce sont des tests pour distinguer homme

et machine, tel celui qu'Alan Turing proposa en 1950, on les nomme CAPTCHA, de l'anglais *Completely Automated Public Turing Test to Tell Computers and Humans Apart*.

Une utilisation des CAPTCHA pour mettre en place un « crowd-computing » massif est pratiquée depuis plusieurs années. Le principe, nommé reCAPTCHA, consiste à demander à la personne qui s'au-

thentifie non pas une seule lecture de lettres brouillées, mais deux. Le premier mot sert à tester si l'on a bien affaire à un humain ; la seconde est une image tirée d'un document que les gestionnaires du reCAPTCHA veulent numériser et que les algorithmes de reconnaissance automatique de caractères ont repéré comme délicat ou totalement illisible. La seconde réponse de l'utilisateur permet de numériser le morceau incompréhensible, travail contrôlable par un autre utilisateur pour que la sûreté soit accrue. Ainsi sont traités des documents trop difficiles pour les numérisations automatiques. Le système est utilisé pour les archives du *New York Times*.



nouvelle. En 2006, Jeff Howe, rédacteur au magazine *Wired*, l'a nommée *crowdsourcing* (« externalisation vers la foule » ou « distribution du calcul à grande échelle »). Des exemples en dehors des sciences ont été mis en œuvre, dont un programme particulièrement étonnant, nommé reCAPTCHA, où les internautes participent à la validation des numérisations d'ouvrages et travaillent sans même en avoir conscience.

Quand la foule calcule

En réfléchissant à ce *crowdsourcing*, certains ont noté des utilisations anciennes du principe datant d'une époque où les ordinateurs n'existaient pas. Ainsi, le *Oxford English Dictionary* a été élaboré au milieu du XIX^e siècle avec l'aide de 800 volontaires qui copiaient des passages de livres, ce qui mit à la disposition des éditeurs un grand nombre de citations précises et une information fiable sur l'usage des mots et leur sens en langue anglaise. *Wikipedia* est bien sûr une forme de *crowdsourcing* et, bien qu'écrite par ceux qui le souhaitent, sa fiabilité n'est pas loin d'égaler celle des autres encyclopédies, tout en proposant une quantité d'articles bien supérieure à

la plus grosse des encyclopédies fondées sur le travail d'experts.

Notons toutefois que le fait de confier une tâche à un très grand nombre de personnes ne produit pas nécessairement des résultats meilleurs que ceux qu'on tire d'un individu isolé et compétent. L'exemple le plus spectaculaire d'un échec du *crowdsourcing* est donné par le match organisé en 1999 entre G. Kasparov et le reste du monde, constitué de 50 000 personnes provenant de 75 pays différents. Durant toute une partie qui s'étendit sur 62 coups, à raison d'un coup tous les deux jours, G. Kasparov affronta la foule des joueurs d'échecs du monde entier qui tentait de le battre. Malgré une organisation permettant aux bons joueurs de peser suffisamment dans les décisions collectives des coups à jouer, G. Kasparov gagna. La partie a été jugée l'une des plus importantes et des plus riches de l'histoire du jeu d'échecs, et un commentaire de 200 pages a été publié pour en analyser les nombreuses nouveautés et subtilités. Reste que la foule a perdu ! N'oublions pas non plus que la « sagesse des masses » a conduit aux plus grandes bulles spéculatives, à de multiples catastrophes économiques, et à l'avènement des pires dictateurs. ■

L'AUTEUR



Jean-Paul DELAHAYE est professeur à l'Université de Lille et chercheur au Laboratoire d'informatique fondamentale de Lille (LIFL).

BIBLIOGRAPHIE

A. Kawrykow et al., *Phylo : A citizen science approach for improving multiple sequence alignment*, *PLoS ONE*, vol. 7(3), e31362, 2012.

Kasparov versus the world, *Wikipedia*, 2012 : http://en.wikipedia.org/wiki/Kasparov_versus_the_World

S. Thomas, *9 examples of crowdsourcing, before crowdsourcing existed*, 2012 : <http://memeburn.com/2011/09/9-examples-of-crowdsourcing-before-crowdsourcing-existed/print/>

F. Khatib et al., *Algorithm discovery by protein folding game players*, *PNAS*, vol. 108(47), pp. 18949-18953, 2011.

F. Khatib et al., *Crystal structure of monomeric retroviral protease solved by protein folding game players*, *Nat. Struct. Mol. Biol.*, vol. 18, pp. 1175-1177, 2011.

C. Lintott et al., *Galaxy Zoo : Morphologies derived from visual inspection of galaxies from the Sloan Digital Sky Survey*, *Monthly Notices of the Royal Astronomical Society*, vol. 389(3), pp. 1179-1189, 2008.

J. Howe, *The rise of crowdsourcing*, *Wired Magazine*, juin 2006.

J. Surowiecki, *The Wisdom of Crowds*, Anchor Books, 2004.

SCIENCE & FICTION

Alien : enquête sur un monstre culte de la SF

Alien concentre les traits les plus repoussants du règne animal.

Une bave abondante parfait le tableau.

Jean-Sébastien STEYER et Roland LEHOUCQ

Le monstre du film *Alien*, réalisé par Ridley Scott en 1979, incarne l'effroi. Aujourd'hui encore, il fait la renommée du film, de ses suites et de ses nombreux « dérivés » (jeux vidéo, etc.). *Alien* a pour origine une créature imaginaire d'aspect phallique, créée par l'artiste suisse Hans Rudi Giger en 1976. Le style de la bête, vaguement humanoïde, osseux et tubulaire, s'inscrit dans un univers graphique sombre et anxiogène, mélangeant design métallique et lignes organiques. Arrêtons-nous sur ce monstre, probablement le plus terrifiant des extra-terrestres : que peut-on en dire en le voyant évoluer dans les films ? Quelles sont ses caractéristiques ? À quelles espèces terrestres ressemble-t-il ?

Dès les premières phases de son développement, *Alien* est repoussant. Tout commence par un œuf oblong et sombre, évoquant celui d'un dinosaure théropode (par exemple le *Tyrannosaurus rex*), sauf qu'il est recouvert d'une sorte de cuir gélatineux. L'espèce est ovipare, avec une larve capable de rester longtemps dans un œuf avant de se développer dans un organisme hôte : quand elle le repère, l'extrémité de l'œuf s'ouvre comme une fleur, et une énorme ventouse à huit pattes en jaillit. Elle saute au visage de l'hôte, en général un humanoïde qui tombe rapidement dans le coma. Ce *facehugger* (littéralement « qui étreint le visage »), disque muni d'appendices articulés et d'une queue élastique, présente une morphologie souvent comparée à celle des limules, bien que ces arthropodes marins présentent une queue rigide

et dix pattes. Les quatre paires d'appendices évoquent plutôt les arachnides, tels les scorpions, les araignées ou les acariens. En revanche, pour le reste du corps, le *facehugger* se rapproche plutôt d'un groupe méconnu d'échinodermes uniquement fossiles, les stylophores. Accroché à la face, le *facehugger* introduit une trompe dans la bouche de son hôte et pond un mini-alien dans la trachée.

Une croissance accélérée

Après quelques heures, ce modèle réduit de l'adulte sort de l'hôte en faisant exploser sa poitrine, d'où son sobriquet anglophone de *chestbuster*. Le monstre atteint sa taille adulte de grand prédateur en quelques heures. Ce développement rapide et complexe rappelle celui des insectes : la plupart des parasitoïdes répertoriés sur Terre sont des hyménoptères, telles les abeilles, les guêpes ou les fourmis. Soulignons toutefois que la succession de formes fixes et libres est également une caractéristique du cycle des cnidaires (le groupe comprenant les méduses) et que la ventouse du *facehugger* évoque aussi la sangsue ou la lamproie.

L'*Alien* adulte concentre les caractéristiques les plus effrayantes des formes vivantes. Incontestablement humanoïde, il présente de nombreuses pointes, épines et tubes dorsaux, paraît couvert d'un revêtement peu classique dans le monde animal : ni peau de mammifères, ni écailles de reptiles, ni chitine d'invertébrés. La tête est originale. Avec cette longue excroissance



1. ALIEN, Le 8^e passager, le premier film de la saga. L'affiche précisait : « Dans l'espace, personne ne vous entendra crier. » Ce qui, étant donné l'absence totale d'air susceptible de propager les sons, est tout à fait juste !