

> à la longueur 7 pour les séquences ayant un 1 en tête est donné en: <http://complexitycalculator.com/D5.csv>

Comme on s'y attend, les séquences les plus simples se trouvent en tête, et plus on avance dans le classement, plus la complexité des séquences augmente, ainsi que leur longueur. Cela confirme expérimentalement deux choses: ce qui est simple est produit avec une fréquence plus grande que ce qui est complexe, et la méthode des petites machines de Turing donne une fine mesure de complexité où toute subjectivité est absente.

Selon la même méthode, on a déterminé une classification par complexité croissante des schémas en dimension 2 de quelques pixels (voir l'encadré 2).

LA SIMULATION DU HASARD

Ces mesures objectives ont permis de réaliser de nouvelles expériences de psychologie qui font comprendre comment l'esprit humain évalue le complexe et le hasard.

Depuis longtemps, la psychologie expérimentale cherche à mesurer les aptitudes

humaines pour les tâches consistant à produire du hasard ou à le reconnaître. Les psychologues ont ainsi mené des expériences où des sujets humains produisaient des suites de 0 et de 1 en imitant le mieux possible ce que donneraient des tirages équitables à pile ou face.

En moyenne, dans 60% des cas, derrière un 0, les sujets humains placent un 1, et derrière un 1, placent un 0, alors que des suites provenant de vrais tirages aléatoires ne présentent de tels basculements que dans 50% des cas. Pourquoi les humains sont-ils incapables de simuler correctement le hasard et favorisent-ils les alternances?

Une hypothèse subtile a été formulée: quand un humain produit des séquences de 0 et de 1 au hasard, il tente de confectionner les plus complexes. Cela lui est mentalement difficile, car il lui faudrait de meilleures capacités de calcul que celles dont il dispose et il ne réussit donc qu'à produire des séquences moyennement complexes: les trop simples sont consciemment évitées, les trop complexes sont inaccessibles ou ne sont pas bien distinguées des séquences moyennement

2

COMPLEXITÉ DES DESSINS

Il est plus facile et naturel pour un humain d'évaluer la complexité d'une image que celle d'une séquence finie de symboles alignés. Il se trouve que la notion de machine de Turing écrivant sur un ruban s'adapte sans difficulté à deux dimensions. La machine se déplace (vers le haut, le bas, la gauche ou la droite) sur un plan quadrillé et fait ses choix en fonction de son état interne et de ce qu'elle lit sur la case où elle est posée et qu'elle peut réécrire. La méthode utilisée pour mesurer indépendamment de toute subjectivité la complexité d'une séquence s'adapte ainsi et donne alors une mesure de complexité pour les petits dessins composés de quelques pixels.

Cette mesure est assez proche de celle qu'on obtient quand on produit des statistiques de la complexité des images découpées en tout petits morceaux tirées du monde réel. C'est un argument en faveur de l'idée que le monde réel est assimilable à un immense ensemble de micromachines à calculer qui produisent les structures qu'on y trouve (soit ramenées à des séquences de symboles, soit ramenées à des petits dessins, etc.).

On a reproduit ici le tableau des petits dessins d composés de pixels blancs et noirs les plus simples. Ils sont classés en fonction de leur probabilité $p(d)$ d'être produits par les machines de Turing fonctionnant sur un quadrillage, c'est-à-dire en fonction de leur complexité croissante telle que ces machines l'évaluent. Les dessins se déduisant par des symétries ou des rotations de ceux représentés ne sont pas mentionnés. Les nombres indiqués évaluent la complexité des dessins par la formule $K(d) = -\log_2(p(d))$ (pour plus de précisions, voir la quatrième référence citée en bibliographie).

15,8159 	16,5252 	17,0378 	17,1387 	17,2143 	17,3662
17,3962 	17,4226 	17,4449 	17,4606 	17,5578 	17,7094
17,7354 	17,7508 	17,7625 	17,9538 	17,9579 	17,9709
18,0757 	18,1145 	18,1361 	18,1773 	18,2346 	18,3558
18,3596 	18,6431 	18,7286 	18,7878 	19,0349 	19,1076
19,1206 	19,1456 	19,1540 	19,2406 	19,3673 	19,6605
19,6829 	19,7418 	19,9768 	20,0518 	20,1158 	20,1528
20,1732 	20,2261 	20,3931 	20,4098 	20,4760 	20,5012
20,5031 	20,5110 	20,5369 			

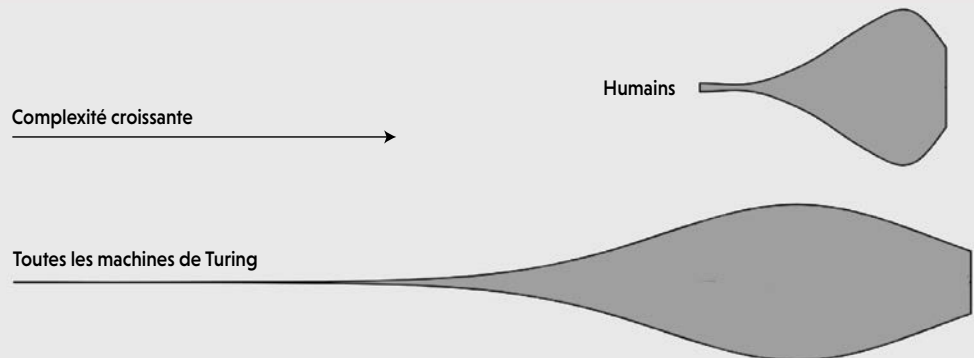
3

INACCESSIBLE COMPLEXITÉ

Quand les humains créent, en essayant d'imiter le hasard, des suites de longueur 10 composées de lettres prises parmi A, B, C et D, ils évitent de produire des suites très simples (comme AAAABBBBB), mais ne réussissent qu'imparfaitement à produire des suites complexes telles que la méthode des machines de Turing nous permet de les connaître.

Le schéma du haut montre la répartition des propositions faites par les humains en fonction de leur complexité : quand, pour une complexité

donnée (sur l'axe horizontal), le schéma est plus épais, c'est que davantage de suites ayant cette complexité ont été proposées. Le schéma du bas montre la répartition donnée par les machines de Turing quand toutes les suites sont prises en compte. On voit que les humains évitent de produire des séquences de très faible complexité (ce que l'on savait) et qu'ils ne réussissent pas à produire des séquences ayant les complexités les plus élevées (pour plus de détails, voir la troisième référence citée en bibliographie).



complexes. Or les séquences de complexité moyenne ont un biais d'alternance, alors que les séquences simples ou très complexes ont des répétitions de 0 et des répétitions de 1 plus nombreuses, et ont donc un taux d'alternance inférieur à 50%. Le biais d'alternance proviendrait donc de la difficulté à produire des séquences fortement complexes.

Une seconde expérience, avec cette fois des suites composées de 4 symboles, appuie l'hypothèse. Nicolas Gauvrit et son équipe ont demandé à des sujets d'âge compris entre 22 et 55 ans de produire des séquences de 10 lettres prises dans un alphabet de quatre lettres {A, B, C, D}, en veillant à ce qu'elles ressemblent le plus possible à des séquences tirées au hasard avec la même probabilité pour chaque lettre.

Les résultats font apparaître deux points intéressants (voir l'encadré 3) :

- La complexité des séquences proposées par les humains est en moyenne supérieure à la complexité moyenne mesurée par les productions de machines de Turing.

- Très peu de séquences proposées par les humains ont la complexité maximale.

C'est donc bien parce que l'esprit humain réussit mal à engendrer de la complexité forte que ses simulations de pile ou face exhibent un biais d'alternance.

COMPLEXITÉ LOCALE

D'autres expériences précisent ces limitations de l'esprit humain et en identifient d'autres caractéristiques. Cette fois, on demande à des sujets d'examiner des séquences

composées d'une vingtaine de symboles 0 ou 1. Les sujets doivent indiquer sur une échelle de 0 à 6 leurs évaluations de ces séquences, de la note 0 pour « certainement aléatoire » à la note 6 pour « certainement pas aléatoire ».

Le sujet humain élabore son jugement non pas en prenant en compte la séquence globale, mais en observant ce qui se passe dans une fenêtre glissant le long de la séquence. On s'interroge alors sur la taille de cette fenêtre quand le sujet examine de longues suites de symboles : examine-t-il trois symboles consécutifs, quatre symboles ou plus encore ? Grâce aux tables de complexité calculées par les machines de Turing, une réponse est proposée. Le protocole est le suivant :

- On demande à des sujets d'évaluer la complexité (entre 0 et 6) d'un certain nombre de séquences assez longues (trop longues pour être évaluées d'un seul coup d'œil).

- Pour chaque largeur L de fenêtre comprise entre 3 et 11, et pour chaque séquence utilisée avec les sujets, on évalue la complexité moyenne quand on déplace une fenêtre de largeur L sur la séquence et cela en se fondant sur le tableau déduit des calculs des machines de Turing pour évaluer la complexité des séquences de L symboles.

- On obtient donc, pour chaque séquence assez longue et pour chaque L , une note comprise entre 0 et 6, évaluant ce que donnerait un dispositif à machine de Turing ne voyant la séquence que par morceaux de largeur L .

- Pour chaque L , on a donc une série de notes (pour les différentes séquences assez >

4

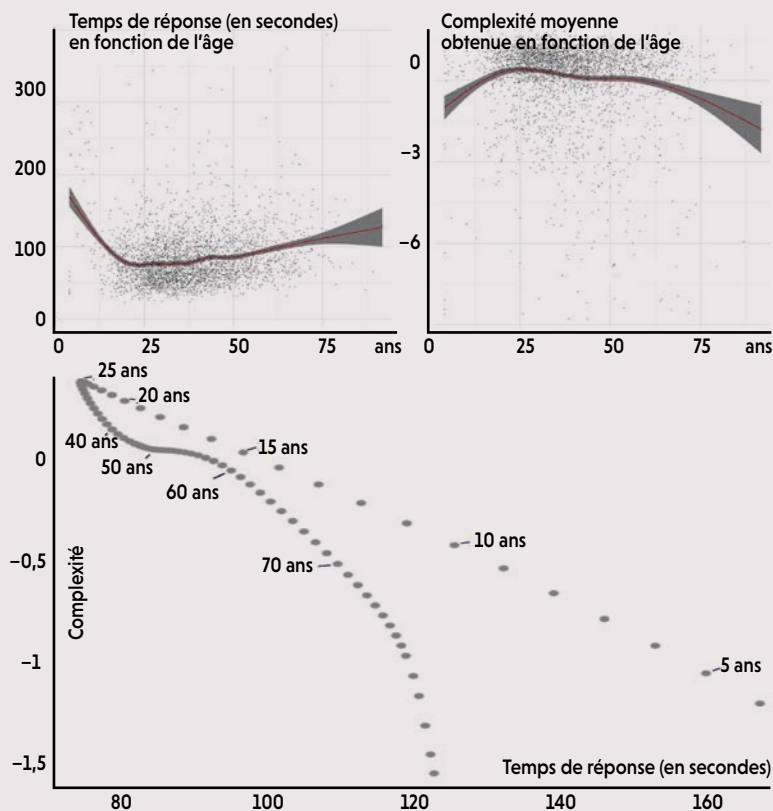
L'ÉVOLUTION DE NOS CAPACITÉS

La production de la complexité exige un effort intellectuel important et de bonnes capacités cognitives (attention soutenue, bonne mémoire immédiate, vivacité, etc.). La mesure de l'aptitude d'un sujet à produire cette complexité est donc un moyen d'évaluer le niveau d'éveil et de bonne santé intellectuelle d'un sujet. En opérant des moyennes sur un grand nombre de sujets d'âges variés, on en déduit l'âge de meilleure forme intellectuelle. Des tests impliquant plusieurs milliers de personnes et les tables de complexité déduites des calculs massifs de machines de Turing (voir le détail du protocole dans l'article), on conclut que l'âge optimal est de 25 ans.

Le graphique a représente le temps de réponse cumulé aux tests : il est minimal à 25 ans. Le graphique b représente la qualité des réponses

données : elle est maximale à 25 ans. L'épaisseur des courbes indique la zone de confiance à 95 %. La courbe c représente les deux variables simultanément : chacun de ses points représente un âge, avec en abscisse le temps de réponse, et en ordonnée une mesure de la qualité moyenne des réponses produites. Les données ont été lissées.

On voit qu'avant 25 ans, la vitesse de réponse comme la qualité augmentent linéairement : apprentissage et maturité produisent leurs effets. Ces bons résultats sont relativement stables pendant quelques années (accumulation des points en haut à gauche), puis se dégradent, et cela de plus en plus rapidement, surtout après 60 ans. Après 70 ans, la vitesse varie peu, mais l'aptitude à produire de la complexité chute sévèrement (voir la première référence bibliographique).



> longues testées) correspondant à une «vision de largeur L ».

– On compare les notations des sujets avec celles données par les machines de Turing pour chaque L , et on détermine le L qui rend compte le mieux de ce que les sujets proposent.

– On juge alors probable que cette largeur L est celle que les humains utilisent quand ils ont à évaluer la complexité d'une séquence assez longue qu'ils ne peuvent pas englober d'un seul regard.

Les largeurs de fenêtre qui donnent des réponses les mieux corrélées avec les évaluations (de 0 à 6) humaines sont 4 et 5. La corrélation devient ensuite de plus en plus mauvaise quand la taille de la fenêtre augmente. On en conclut qu'un humain détermine si une séquence longue est aléatoire en se fiant à sa complexité locale évaluée en examinant les sous-séquences d'environ 4 ou 5 éléments consécutifs. Ce résultat est conforme aux évaluations obtenues par ailleurs de la mémoire à court terme (voir l'encadré 3).

À QUEL ÂGE COMMENCE-T-ON À VIEILLIR ?

Puisque produire et reconnaître de la complexité nous est difficile, la mesure de nos capacités à le faire doit évoluer avec l'âge et pourrait, entre autres choses, servir à identifier des signes annonciateurs d'une maladie neurodégénérative.

Pour voir cette évolution, il fallait disposer de méthodes fiables et donnant des résultats stables et reproductibles. C'est pourquoi on a mis en place un protocole expérimental impliquant un grand nombre de personnes (3429) d'âges variés (entre 4 et 91 ans), et que l'on a collecté et analysé un ensemble de données suite à des tests passés par chacun. L'ensemble du travail est fondé sur les mesures de complexité obtenues par les calculs de milliards de machines de Turing. La cohérence des résultats, leur convergence, et les informations nouvelles qu'ils fournissent montrent l'intérêt des évaluations objectives de complexité en psychologie.

Nicolas Gauvrit, le principal organisateur de cette grande expérimentation, avance que la capacité à simuler le hasard est liée à nos capacités cognitives générales :

« Si l'on demande à quelqu'un de produire une suite aléatoire de chiffres, cette personne cherche à produire une suite complexe. Or, depuis longtemps, les psychologues pensent que la complexité permet de mesurer des fonctions cognitives de haut niveau, telles que l'attention, la mémoire à court terme ou encore l'inhibition. Par exemple, après un 1, nous avons tendance à vouloir mettre un 2 : il

BIBLIOGRAPHIE

N. Gauvrit et al., **Human behavioral complexity peaks at age 25**, *PLoS Comput. Biol.*, vol. 13(4), e1005408, 2017.

N. Gauvrit et al., **A preference for some types of complexity comment [...]**, *Acta Psychologica*, vol. 174, pp. 48-53, 2017.

N. Gauvrit et al., **Algorithmic complexity for psychology [...]**, *Behav. Res.*, vol. 48(1), pp. 314-329, 2016.

H. Zenil et al., **Two-dimensional Kolmogorov complexity [...]**, *PeerJ Comp. Sci.*, vol. 1, e23, 2015.

N. Gauvrit et al., **Algorithmic complexity for short binary strings applied to psychology [...]**, *Behav. Res.*, vol. 46(3), pp. 732-744, 2014.

F. Soler-Toscano et al., **Calculating Kolmogorov complexity [...]**, *PLoS One*, vol. 9(5), e96223, 2014.

J.-P. Delahaye et H. Zenil, **Numerical evaluation of the complexity of short strings [...]**, *Applied Math. and Comput.*, vol. 219, pp. 63-77, 2012.

H. Zenil et J.-P. Delahaye, **On the algorithmic nature of the world**, dans G. Dodig-Crnkovic & M. Burgin (eds.), *Information and Computation*, pp. 477-496, World Scientific, 2010.

faut inhiber ce réflexe pour créer de la complexité. Nous savons par ailleurs que ces fonctions cognitives de haut niveau reposent principalement sur le lobe frontal du cerveau. C'est pourquoi les neurologues demandent parfois au patient qui a subi un choc à la tête de produire une suite de chiffres la plus aléatoire possible. Si la région cérébrale est endommagée, le patient aura du mal à créer de la complexité et la suite présentera de nombreuses répétitions.»

Hector Zenil, l'un des organisateurs de la large étude, considère que la capacité à produire des suites complexes est aussi directement corrélée à la créativité, car celle-ci exige de produire de l'inattendu, ce qui est lié à la capacité d'accéder à chaque instant à un ensemble ouvert et large d'éventualités. Pour lui, cette étude doit aussi s'interpréter comme une évaluation de la créativité en fonction de l'âge.

Voyons quelques détails sur le protocole utilisé. Les sujets ont été invités à répondre à cinq exercices liés aux séquences aléatoires. On a évalué leurs réponses en considérant la complexité des séquences courtes qu'ils produisaient.

Exercice a. Les participants doivent créer une série de 12 «pile ou face» pour qu'elle semble aléatoire à quelqu'un d'autre. Ils choisissent sur un écran qui leur propose de manière répétée les deux options. Seul le dernier choix opéré reste visible quand ils effectuent un nouveau choix.

Exercice b. Dix fois de suite, les participants doivent sélectionner 5 cartes dans un jeu en tentant de faire comme si elles provenaient d'un jeu parfaitement mélangé.

Exercice c. Les participants doivent engendrer 10 nombres entiers compris entre 1 et 6, comme pourrait en produire un dé lancé 10 fois. Pour cette épreuve, les choix précédents restent visibles, mais ne peuvent pas être modifiés.

Exercice d. Les participants doivent pointer 10 fois de suite un petit cercle pris dans une série de 9 cercles présentée à l'écran, en choisissant au hasard.

Exercice e. Les participants doivent noircir des cases d'une grille 3×3 initialement blanche, de façon que la grille finale semble aléatoire. Ici, les participants sont autorisés à revenir en arrière sur leur choix jusqu'à en être satisfaits.

À chaque fois, on mesure à la fois la complexité des choix faits et le temps mis pour les faire. Comme le note l'article rapportant cette expérience, celle-ci est un «test de Turing inversé». On ne demande pas à un système informatique d'imiter les réponses que donnerait un humain pour savoir si on peut qualifier le système d'intelligent (c'est l'idée du célèbre test de Turing); on considère, à l'opposé, que c'est la capacité de l'humain à approcher ce que (collectivement) les machines de Turing nous disent de

la complexité qui mesure le niveau des aptitudes cognitives des personnes testées.

Malgré la nature variée des tâches demandées, la réussite des sujets évolue de la même façon en fonction de l'âge pour les cinq exercices: le niveau de réussite s'accroît jusqu'à 25 ans, puis décroît assez lentement jusqu'à 60 ans, âge à partir duquel le déclin s'accélère. L'analyse de la vitesse à laquelle les sujets répondent évolue de façon inverse: diminution des temps de réponse jusqu'à 25 ans, puis augmentation (voir l'encadré 4).

CE QUI INFLUE... ET CE QUI N'INFLUE PAS

Les données collectées sur les participants ont permis d'analyser d'autres facteurs que celui de l'âge. Le sexe des participants n'a aucun effet significatif sur les résultats de la simulation du hasard. En revanche, les participants de sexe masculin répondent plus rapidement que ceux de sexe féminin. Cette différence avait déjà été notée dans d'autres études: les hommes semblent plus pressés que les femmes!

L'analyse montre que les réussites aux quatre premiers exercices sont fortement corrélées: réussir assez bien à l'un indique le plus souvent qu'on réussira bien aux autres. Le cinquième exercice est moins nettement corrélé aux autres, ce qui est peut-être dû à la possibilité donnée aux participants à cet exercice de revenir en arrière sur leurs décisions jusqu'à en être globalement satisfaits.

Le rapport entre réussite aux exercices et croyance au paranormal a été examiné. À cette fin, on a demandé aux participants de noter sur une échelle de 0 à 6 leur degré d'approbation de la phrase suivante: «Certaines coïncidences dont j'ai été témoin ne s'expliquent bien qu'en envisageant la perception extrasensorielle ou des forces paranormales.» L'idée était que quelqu'un n'ayant pas l'aptitude à bien évaluer et produire du hasard chercherait à expliquer des situations parfaitement banales du point de vue statistique en invoquant des considérations fondées sur l'intervention de forces paranormales. Aucune corrélation n'a été trouvée pour les quatre premiers exercices, et une légère corrélation négative est apparue pour le cinquième: les participants créent un peu moins bien des grilles complexes quand ils croient au paranormal.

AUTRES APPLICATIONS

Terminons en mentionnant que, grâce à la mise à disposition des tables de complexité calculées par machines de Turing, plusieurs autres études ont pu être menées concernant la perception visuelle (voir l'encadré 2), la formulation des jugements esthétiques et la façon dont s'opère l'apprentissage d'un langage. ■

L'AUTEUR



LOÏC MANGIN
rédacteur en chef adjoint
à Pour la Science

UN HAVRE POUR LES ABEILLES

Au Havre, au cœur d'une exposition dédiée au génie des abeilles, des modèles façonnés en papier rendent compte de la diversité insoupçonnée de ces insectes.



La nomade jaune, *Nomada flava*,
l'anthilde à manchettes *Anthidium manicatum* et l'abeille orchidée
Euglossa hemichlora plus vraies
que nature, mais en papier.

En juin 2019, le chercheur Vincent Bretagnolle et ses collègues ont publié une étude montrant que l'agriculture biologique profite aux colonies d'abeilles mellifères, notamment pendant la période de disette à la fin du printemps. Au vu de l'engouement pour le bio et de l'expansion attendue des terres dévolues à ce type de produits, on peut se réjouir pour les abeilles. C'est certainement le cas d'Éric Tourneret, commissaire de l'exposition « Abeilles, une histoire naturelle », présentée au Muséum d'histoire

naturelle du Havre. On peut y voir de nombreuses photographies que ce spécialiste des abeilles a rapporté de dix années de reportages dans vingt pays. On découvre également les toutes dernières découvertes scientifiques sur ces insectes, aussi bien sur leur biologie (communication, prises de décision démocratique, performances cognitives, moyens de défense...) que sur les causes de leur déclin, à commencer par le rôle des pesticides.

Pour compléter l'exposition, l'artiste Hugo Boistelle a quant à lui façonné une série de vingt abeilles en papier illustrant la grande diversité des espèces (géantes ou naines, dotées d'un abdomen fin et long ou au contraire court et trapu...). De fait, les spécialistes en ont identifié plus de 20 000, dont plus de 1 000 en France ! Certes, on connaît l'abeille *Apis mellifera*, l'espèce dominante en apiculture : les ruches qui abritent ses colonies nous sont familières,

même au cœur des villes. Mais saviez-vous que plus de 1 000 espèces en France sont solitaires, c'est-à-dire qu'elles ne fondent pas de colonie sociale et hiérarchisée ? C'est notamment le cas (voir ci-dessus) de la nomade jaune *Nomada flava* et de l'anthilde à manchettes *Anthidium manicatum*. On peut aussi citer l'abeille orchidée *Euglossa hemichlora*, une solitaire qui vit quant à elle au Panamá.

On peut s'étonner que d'autres espèces soient sans dard : pourtant, dans le monde, environ 600 espèces sont dépourvues de cet aiguillon. Un héritage des caractères des abeilles ancestrales, celles qui sont apparues il y a 100 millions d'années, avant