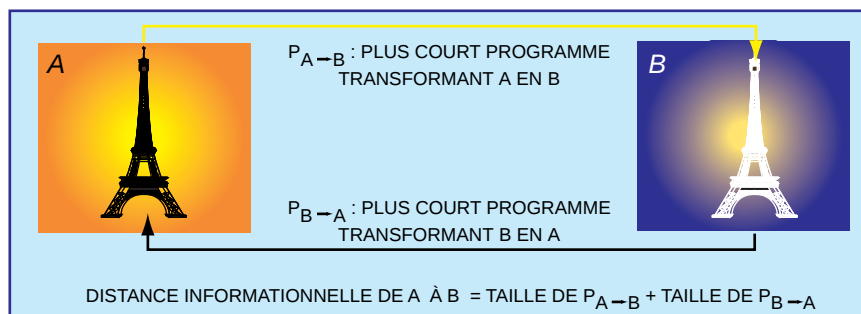




6. La distance informationnelle entre deux objets A et B s'obtient en additionnant la taille des plus courts programmes permettant de passer de A à B et de B à A. Elle possède des propriétés de minimalité qui lui confère une importance particulière en théorie algorithmique de l'information. De plus les rapports qu'elle a avec la thermodynamique du calcul lui donnent un sens physique que les autres distances ne possèdent pas.

en effet de très longues séquences d'événements assimilables à des opérations élémentaires de calculs dont les produits sont le langage, les sciences et plus généralement nos facultés d'analyse. Ces très longs calculs, sans que nous sachions précisément comment, ont produit dans nos cerveaux une notion d'analogie très fine et très efficace, essentielle à notre survie. Cette cristallisation est incomplète (il n'y a pas de raisons sérieuses de croire que notre intelligence échappe aux limitations formulées par la logique), mais elle est bien meilleure que celle que nous réussissons à insérer dans nos programmes : aujourd'hui nous ne savons pas écrire des programmes qui égalent notre capacité à percevoir des analogies.

LA DISTANCE ENTRE SÉQUENCES GÉNÉTIQUES

L'indécidabilité rend-elle la distance informationnelle stérile ? Heureusement non, car les majorations que l'on peut faire la rendent utilisable dans des cas sans doute assez nombreux.

Un exemple d'application de la distance informationnelle est la distance définie dans l'équipe de bio-informatique de Lille, composée de Max Dauchet, Éric Rivals, Jean-Stéphane Varré et moi, pour évaluer la distance entre séquences génétiques. Depuis longtemps les généticiens, et tout particulièrement ceux qui souhaitent faire de la reconstitution d'arbres phylogénétiques (ce sont les arbres qui indiquent les parentés entre espèces animales et végétales), utilisent des mesures de ressemblance entre séquences génétiques. L'idée à la base des distances utilisées jusqu'à présent était le décompte des mutations, délétions ou insertions de nucléotides (les lettres de l'alphabet génétique) : plus le nombre des changements ponctuels nécessaires pour passer d'une séquence à l'autre

est grand, plus les séquences sont considérées comme éloignées. Récemment nous avons proposé une nouvelle distance qui mesure l'éloignement entre séquences en considérant d'autres événements possibles comme les déplacements de morceaux de séquences ou leur duplication. Cette distance, que nous avons conçue comme une approximation calculable de la distance informationnelle, est plus précise que les distances utilisées classiquement (qu'elle généralise). On peut donc espérer qu'elle fournira des arbres phylogénétiques plus exacts.

Notons encore que les algorithmes modernes de compression de données pour les images de films sont fondés sur l'idée que deux images successives A et B d'une séquence sont proches et que, pour gagner de l'espace, il suffit de ne coder que leur différence : ce qui change dans A pour obtenir B. Cette quantité d'informations permettant de passer de A à B est en fait la distance informationnelle entre A et B (ou presque, puisque l'on ne s'occupe pas du passage de B à A).

INTERPRÉTATION THERMODYNAMIQUE

L'interprétation thermodynamique de la distance informationnelle est une autre confirmation de son intérêt. Les cinq chercheurs physiciens, mathématiciens et informaticiens ont en effet démontré un résultat remarquable concernant la distance informationnelle : elle est liée aux coûts thermodynamiques minimaux de la transformation de A en B et de B en A.

On sait depuis quelques années, à la suite des travaux de Rolf Landauer et de Charles Bennett, qu'il est possible, en théorie, de réaliser des calculs sans dépense d'énergie et n'entraînant donc aucun accroissement de l'entropie physique. Lors d'un calcul, les seules dépenses d'énergie inévitables proviennent de l'utilisation d'opérations irréver-

sibles comme l'effacement d'une mémoire dont on peut se passer si l'on accepte de garder des informations inutiles en fin de calcul.

Il est alors intéressant de considérer le flux minimum d'informations (ajout d'informations à A au départ, effacement d'informations à la fin du calcul une fois B obtenu) pour transformer A en B de manière réversible. Cette distance déduite de considérations thermodynamiques est équivalente à la distance informationnelle : elle en diffère d'un facteur deux au plus. D'autres propriétés de minimalité confèrent, parmi toutes les distances envisageables, un statut privilégié à la distance informationnelle.

La distance informationnelle entre A et B est donc une quantité ayant un sens en thermodynamique, ce qu'aucune autre distance ou notion mathématique de similitude ne possédait jusqu'à présent. L'adéquation entre ce que donne la nouvelle distance et l'intuition, le lien avec le monde physique, ainsi que d'autres résultats sur la distance informationnelle trouvés par le groupe des cinq chercheurs autour de Charles Bennett montrent clairement qu'une notion profonde a été identifiée.

Jean-Paul DELAHAYE est directeur adjoint du Laboratoire d'informatique fondamentale de Lille du CNRS.

e-mail : delahaye@lil.fr

C. BENNETT, et al. *Thermodynamics of Computation and Information Distance*. Proc. 25th ACM Symp. Theory of Computation, pp. 21-30, 1993.

C. BERGE, *Espaces topologiques et fonctions multivoques*, Éditions Dunod, Paris, 1966 (voir le chapitre 6).

J.-P. DELAHAYE, *Information, complexité et hasard*, éditions Hermès, Paris, 1994.

M. LI et P. M. B. VITANYI, *An Introduction to Kolmogorov Complexity and Its Applications*, Springer-Verlag, 1997 (voir le chapitre 8).

J. SERRA, *Image Analysis and Mathematical Morphology*, Academic Press, New York, 1982.

J.-S. VARRÉ, *Phylogénie et compression de données*, Publication du Laboratoire d'informatique fondamentale de Lille, URA CNRS 369, juillet 1996.

J.-S. VARRÉ, E. RIVALS, M. DAUCHET et J.-P. DELAHAYE, *Les distances transformationnelles et applications à la phylogénie*. Journées Analyse des séquences génomiques, École Polytechnique (Palaiseau), 20 et 21 juin 1996.

Articles de P. Vitanyi sur Internet :

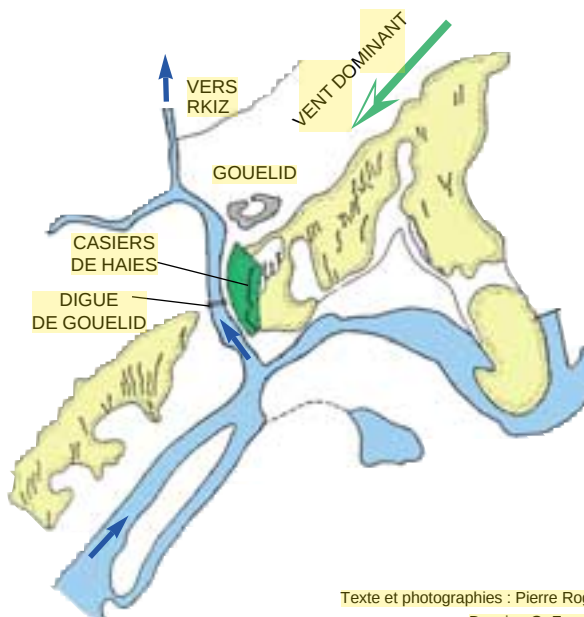
<http://www.cwi.nl/~paulv/publications.html>

L'avancée des dunes

Presque dépourvue de terres irriguées, la Mauritanie fonde de grands espoirs sur l'aménagement des 12 000 hectares de la cuvette de Rkiz, au Sud-Ouest du pays. Cette dépression est actuellement inondée chaque année par un marigot qui y déverse le surplus de la crue du Sénégal en période de hautes eaux. La construction du barrage de Diama, en 1985, permet progressivement de remplacer les cultures de décrue, peu rentables, par des cultures irriguées en permanence.

Toutefois, le projet ne tenait pas compte de la progression de dunes vives qui, poussées par l'harmattan, vent de Nord-Est, menacent de combler le marigot qui amène l'eau à Rkiz. En effet, la présence de ces dunes dans une région sahélienne couverte d'acacias est inhabituelle. Elle s'explique par l'extrême dégradation de la végétation due au surpâturage et à la surexploitation du bois, dans un milieu fragilisé par les sécheresses persistantes qu'a connues cette région pendant plus de 20 ans.

Pour s'opposer à l'ensablement, des haies de branches finement tressées ont été disposées en casiers, perpendiculairement à la direction de l'harmattan (flèches sur les photographies). Ce dispositif traditionnel n'arrête toutefois que des apports modérés de sable. Aujourd'hui, ces dunes déjà formées s'avancent en vagues successives (voir la photographie 1, face au vent) et submergent rapidement le dispositif de haies pourtant très dense, installé sur la dernière colline avant le marigot (voir la photographie 2). Ayant comblé les casiers (voir la photographie 3), elles redescendent déjà sur le flanc de la colline vers le marigot que l'on aperçoit à l'arrière-plan de la photographie 4 (dos au vent). Il faudrait mettre en œuvre de nouvelles techniques de lutte contre ces dunes mobiles si on ne veut pas voir disparaître les espoirs d'aménagement de la cuvette de Rkiz.



Texte et photographies : Pierre Rognon
Dessin : G. Forgiarini





Biologie

Non, l'avenir n'est pas héréditaire

Hervé Ponchelet. Éditions Belin, 1997.

Le livre qu'Hervé Ponchelet, journaliste à l'hebdomadaire *Le Point*, consacre à certaines dérives actuelles de la génétique est le cri de protestation, parfois indigné, parfois effaré, d'un observateur extérieur au monde scientifique, mais attentif. Toute la réflexion d'H. Ponchelet se fonde sur les travaux et les commentaires de généticiens et de spécialistes de la procréation eux-mêmes engagés dans le débat ; d'ailleurs, son ouvrage comporte de nombreuses citations. De l'une à l'autre, l'auteur dévide le fil de sa propre analyse et, parfois, manifeste sa révolte. Cette construction très journalistique a l'avantage de fournir un beau reflet des débats actuels sur les dérives de la génétique et de la «procréatique». En revanche, elle amène parfois l'auteur à adhérer, successivement, à des analyses dont les principes sont notoirement différents. Peut-être cela rend-il encore plus authentiques les sentiments exposés par un observateur confronté à différentes prises de positions, parfois irréductibles les unes aux autres, mais véhiculant toutes au moins des «vérités partielles».

Cette hésitation se ressent particulièrement dans la première partie, qui est consacrée aux liens du sang et à la filiation. L'auteur distingue bien les deux types de filiations caractéristiques du monde de l'homme : la filiation du sang et la filiation de l'esprit. Adversaire évident de la tyrannie du «droit du sang», opposé aux droits qui reposent sur des liens intellectuels et affectifs, H. Ponchelet n'en n'est pas moins, de manière parfois un peu singulière, très sensible à la revendication virulente du «droit à la connaissance de ses origines biologiques», qu'il s'agisse de la fécondation par sperme de donneur, de l'adoption, voire de l'accouchement sous X.

Cette revendication a de nombreux supporters, notamment parmi les psychanalystes, auxquelles l'auteur donne abondamment la parole. Les arguments présentés sont de trois ordres : droit à la vérité, importance de «l'enracinement» dans le développement harmonieux d'une personnalité et intérêt biologique de l'enfant. Ces trois arguments sont de qualité, mais ignorent habituellement le fait, bien connu de tous les généticiens,

qu'environ dix pour cent des enfants venus au monde «naturellement» et élevés par le couple géniteur théorique... ne sont en fait pas les enfants biologiques de leur père.

Collectivement, le nombre de ces enfants est bien supérieur à celui de ceux issus de l'adoption ou de l'assistance médicale à la procréation. Au nom des principes évoqués précédemment, qui de bonne foi prétendrait qu'il est essentiel de leur indiquer toujours, le moment venu, l'identité du géniteur ? En réalité, l'intransigeance et le systématisme de cette revendication d'une connaissance obligatoire de ses origines biologiques procède, qu'on le veuille ou non, de l'analyse que la filiation du sang, la loi des gènes, s'imposent de façon incontournable, dans l'espèce humaine comme dans le monde animal. Ce sentiment correspond à un fort courant de nos sociétés modernes où la conjonction du retour à un certain intégrisme religieux et d'un matérialisme de fait amène à considérer que, après tout, les hommes et les femmes n'ont peut-être pas grand-chose d'autre à transmettre à leurs enfants... que leurs gènes.

Dans sa deuxième partie, H. Ponchelet s'inquiète très justement d'une certaine évolution de la génétique des comportements dont les objectifs et les résultats, mal présentés et mal compris, sont «pain béni» pour les vieilles idéologies déterministes et racistes. Le caractère idéologique des assertions plus que centenaires sur l'inégalité des aptitudes intellectuelles des différentes races est bien souligné, de même que l'absurdité d'un certain réductionnisme ramenant un comportement humain à un déterminisme génétique rigoureux et simple. L'auteur aurait également pu noter l'incroyable erreur de perspective de beaucoup des généticiens du comportement et des neurobiologistes, qui franchissent allégrement le pas entre la découverte qu'un gène joue un rôle essentiel dans un comportement adaptatif inné animal, et la conclusion que son influence chez l'homme est similaire. Ce que néglige totalement une telle démarche, c'est le caractère proprement humain de la réappropriation de certains comportements innés. Ainsi, un même trait comportemental sélectionné par l'évolution depuis des centaines de millions d'années s'exprime-t-il de manière bien différente chez des mammifères humains et non humains : le désir sexuel, par exemple, ne se manifeste chez les chevaux que par la saillie de l'étalon, mais, au-delà de l'acte sexuel lui-même, par près de la moitié de la production poétique et artistique des sociétés humaines !

Là est vraiment le monde de l'homme, et non pas dans le fait évident que de mêmes gènes gouvernent la différenciation mâle ou femelle dans l'espèce humaine et dans les autres espèces !

Dans la troisième partie, l'auteur aborde le grave sujet d'actualité de la «tare génétique». Il illustre la vieille notion de l'atavisme par la saga des Rougon-Macquart de Zola. Il aurait, en fait, pu remonter beaucoup plus loin, par exemple à l'image du destin dans la tragédie grecque. Fardeau génétique, stigmatisation du lignage, certitude d'un destin contraire, la tare génétique est au centre des angoisses et des fantasmes humains. Vouloir y échapper, de toutes ses forces, par tous les moyens, fait en effet planer le spectre de la réminiscence des aspirations eugéniques qui, sous une forme ou sous une autre, sont présentes dans les sociétés humaines depuis des millénaires. Avec l'évolution de la génétique moderne, c'est à la fois le «gène prison» et le «gène libérateur» qui s'annoncent, et cela H. Ponchelet le présente avec beaucoup de sensibilité et de lucidité. Le gène impliqué dans la survenue d'une maladie, qu'il ait le statut du coupable principal ou de facteur de prédisposition, c'est en effet à la fois l'espoir d'un traitement de nature à atténuer le fardeau pathologique de l'humanité, mais c'est aussi la menace de transformer les droits de l'homme en les droits laissés à un homme particulier en fonction de ses gènes ! L'ampleur de ces enjeux et l'importance de ces débats sont telles que leur exposition claire par un journaliste attentif et humaniste est naturellement la bienvenue.

Axel KAHN

Astronomie

Imago Mundi

Francesco Bertola (traduit de l'italien par A. Hayli et É. Schelstraete). La Renaissance du Livre, 1996.

Ce luxueux ouvrage sur l'imagerie cosmique nous rappelle que la cosmologie, science de l'Univers considéré dans son ensemble, est aussi un projet esthétique. L'étymologie même du mot l'indique : en grec, le *kosmos* désignait d'abord la parure des femmes, les ornements, le bel aspect physique ou moral. La notion de cosmos appliquée à l'organisation de l'Univers apparut vers le