Hartley définit la quantité d'information contenue dans une suite de N symboles élémentaires comme étant égale à $N \log S$, tout en soulignant que ces considérations sont indépendantes de la nature ou de la signification de l'information véhiculée.

C'est dans le contexte de la Seconde Guerre mondiale que les idées de Nyquist, Hartley et d'autres ingénieurs ou chercheurs convergent vers une « théorie de l'information », essentiellement grâce aux travaux de Shannon, également des Laboratoires *Bell*.

Ce mathématicien-ingénieur quelque peu excentrique rédige en 1945 *Une théorie mathématique de la cryptographie*, rapport classé secret jusqu'en 1957. Dans ce document, il a l'intuition de comparer les perturbations du signal engendrées par la clef de cryptage au « bruit » qui parasite les lignes de télécommunication et que les ingénieurs essayent de minimiser. Shannon y définit en outre la quantité d'information, dans le cas discret, comme la moyenne logarithmique des probabilités de sélection des éléments constitutifs du message *(voir l'encadré ci-dessous)*.

Parallèlement à ses recherches cryptographiques, Shannon a travaillé à l'automatisation de la défense antiaérienne. De la Première Guerre mondiale à la Seconde, les avions volaient deux fois plus haut et deux fois plus vite. Les opérateurs des radars n'avaient alors pas le temps de communiquer efficacement avec les équipes d'artillerie. Pour transmettre les données automatiquement, il fallait là encore, pour l'équipe de Shannon, trouver un codage adéquat et définir l'information à transmettre.

Une autre équipe des Laboratoires *Bell*, dirigée par Norbert Wiener, se préoccupait des perturbations dues aux dents des engrenages, qui affectaient la stabilité du bras déterminant l'axe de tir. Assisté de l'ingénieur Julian Bigelow, Wiener eut l'idée de consulter son ami cardiologue, Arturo Rosenblueth, qui avait déjà analysé des mouvements perturbés caractéristiques de certaines formes d'ataxie (difficulté de coordination) dans le tabès, une forme tardive de la syphilis.

De l'information dans les boucles de rétroaction

Ces échanges transdisciplinaires, liés à l'effort de guerre des États-Unis, ont conduit à l'élaboration d'une théorie générale de la régulation, que Wiener nomma cybernétique dès 1947. Cette théorie portant sur des systèmes quelconques, que ce soit des machines ou des organismes vivants, était

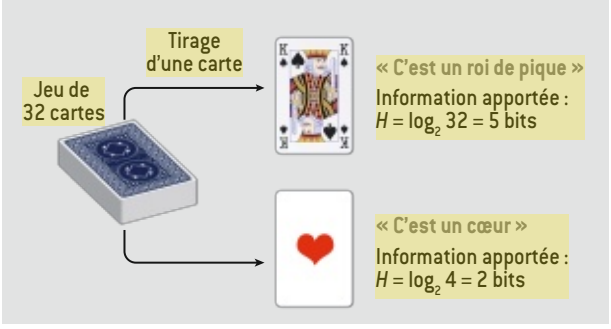
QUANTITÉ D'INFORMATION, INCERTITUDE, ENTROPIE

Pour définir la quantité d'information contenue dans un message d'une certaine longueur et formé à partir d'un certain alphabet de lettres ou symboles, Claude Shannon a cherché à quantifier l'incertitude ou le choix – ou encore le manque d'information – associé à l'ensemble des messages possibles ayant la même longueur et utilisant le même alphabet.

Supposons que l'on ait N messages possibles et que la probabilité de choisir le i-ème message soit p_i, avec $p_1 + p_2 + \ldots + p_N = 1$. Shannon a cherché une fonction $H(p_1, \ldots, p_N)$ qui quantifie l'incertitude associée à ces possibilités. En exigeant qu'elle satisfasse à certaines propriétés naturelles, il a prouvé que la fonction H, nommée entropie informationnelle ou entropie statistique, devait être de la forme : $H = -K(p_1 \log p_1 + p_2 \log p_2 + \ldots + p_N \log p_N)$, où K est une constante positive arbitraire. Quand un seul choix est possible (tous les p_i sont nuls, sauf un), l'incertitude est nulle ; l'information apportée en spécifiant ce choix est donc elle aussi nulle. Dans le cas où les probabilités sont toutes égales (à p), on a $H = K \log (1/p)$; cela correspond à une entropie ou une incertitude maximale, et l'information apportée en spécifiant le choix effectué est maximale.

Un exemple très simple est celui d'un message constitué d'un seul symbole binaire, généralement noté 0 ou 1. Il y a deux choix possibles ; s'ils ont même probabilité ($p_1 = p_2 = 1/2$), l'entropie de Shannon devient $H = K \log 2$. Par commodité, on prend $K = 1$ et des logarithmes de base 2 ; l'entropie de Shannon d'un symbole binaire devient alors $H = 1$, ce qui définit un bit d'information (le terme « bit », contraction de *binary digit*, a été proposé par le statisticien américain John Tukey). De la même façon, l'affirmation « Mon rendez-vous est le matin » apporte une information de un bit, alternative élémentaire, s'il est équiprobable que ce rendez-vous ait lieu le matin ou l'après-midi.

Autre exemple : l'annonce de la carte que l'on a choisie dans un jeu de 32 cartes. Les choix possibles ayant la même probabilité, 1/32, l'entropie de Shannon est égale à $\log_2 32$, c'est-à-dire cinq bits (puisque $32 = 2^5$). La quantité d'information apportée par cette annonce correspond au nombre de questions à réponse binaire (« oui » ou « non ») qu'il faut poser pour trouver la carte : « Est-elle de couleur rouge ? », « Est-elle égale ou supérieure au valet ? », etc.



Shannon a aussi montré qu'on pouvait calculer l'« entropie » d'une langue à partir d'un calcul sur sa redondance. Bien des phrases restent compréhensibles lorsque des lettres sont enlevées ; c'est le principe des techniques de sténographie et à titre d'exemple, sans vouloir flatter les lecteurs, Shannon écrivait : « Is lctrs ntl-lgnts cmprnnnt fclmnt n txt sns vlls. »

En même temps, Shannon a toujours tenu à préciser que la dimension sémantique de l'information était absente de ses travaux. Que ce soit à travers la définition statistique ou par l'utilisation de la redondance, l'ingénieur en télécommunications aura la même application à transmettre une déclaration de guerre, des cours de la Bourse ou un message dépourvu de sens pour lui.

La formule de l'entropie au sens de Shannon, $H = -K \sum p_i \log p_i$, ressemble beaucoup à la formule qui définit l'entropie en mécanique statistique, telle qu'elle a été établie par les physiciens Ludwig Boltzmann et Josiah Gibbs au XIX^e siècle. Il suffit d'adopter les logarithmes népériens et de prendre pour K la constante de Boltzmann ($k_B = 1{,}38 \times 10^{-23}$ joule/kelvin) pour obtenir une expression formellement identique, p_i désignant alors la probabilité du i-ème état microscopique accessible au système physique considéré. Toutefois, l'identification entre l'entropie/information de Shannon et l'entropie de Boltzmann-Gibbs a été contestée par de nombreux scientifiques ; le physicien et bio-informaticien américain Hubert Yockey, notamment, a souligné que les espaces de probabilités dans les deux situations ne sont pas isomorphes (leur correspondance n'est pas rigoureuse).

centrée sur la notion de « boucle de rétroaction » (*feedback* en anglais). Le thermostat est l'archétype d'un dispositif réalisant une telle régulation : il mesure la température et, en fonction du résultat, modifie son fonctionnement dans un sens ou dans un autre afin d'atteindre ou maintenir la température souhaitée.

Après la guerre, de 1946 à 1953, une dizaine de conférences interdisciplinaires, organisées avec le soutien de la Fondation Macy, ont permis à Wiener et d'autres scientifiques de classer ces phénomènes de rétroaction et d'y souligner le rôle clef de l'information – celle-ci étant définie par Wiener de façon similaire à l'information de Shannon, avec la même analogie par rapport à l'entropie statistique.

Les boucles de rétroaction permettent de déclencher automatiquement un système de régulation par action de mécanismes agissant dans le sens opposé à celui ayant causé un écart avec l'objectif désiré. Ces phénomènes étaient connus depuis longtemps (par exemple en mécanique pour les machines à vapeur ou en biologie pour la régulation de la température ou de la pression sanguine). Mais en considérant que c'était de l'information qui circulait dans ces boucles, les cybernéticiens ont pu construire une théorie abstraite des systèmes de régulation.

Parmi les scientifiques impliqués dans ces réflexions théoriques, voire philosophiques, figuraient bien entendu Shannon, mais aussi d'autres célébrités – des physiciens et mathématiciens comme John von Neumann, des biologistes comme Max Delbrück (l'un des pères de la biologie moléculaire), ou encore des anthropologues comme Margaret Mead ou Gregory Bateson.

Les fondements mêmes de plusieurs disciplines scientifiques furent réinterprétés à l'aune de la théorie de l'information et de la cybernétique. Dans le domaine des techniques, l'application de la théorie de l'information à la Shannon a eu de nombreuses retombées. Il suffit de penser aux codes correcteurs d'erreurs, utilisés aujourd'hui universellement pour protéger l'information numérique contre sa dégradation, aux techniques de compression de données, etc. Dans l'exemple de la compression de données, les travaux de Shannon et de ses continuateurs ont permis de conceptualiser la redondance : ainsi, pour coder numériquement une image montrant une partie de ciel bleu uniforme, au lieu de coder chaque pixel (cordonnées du pixel, valeur de la couleur), un codage compressé de type JPEG utilisera une syntaxe permettant de traduire en langage informatique « de tel pixel à tel autre, afficher telle valeur de bleu ».

Au-delà des services rendus aux technologies de la communication, ce qui était le cadre originel de la théorie de l'information édifiée par Shannon et son école, beaucoup de scientifiques ont été tentés de voir dans l'information une entité fondamentale, intervenant dans tous les domaines du savoir.

Matière, espace, temps et information : les nouveaux éléments ?

Il y a une vingtaine d'années déjà, Rolf Landauer (1927-1999), physicien chez IBM aux États-Unis, affirmait que « l'information est physique » et qu'elle doit avoir la même place, dans l'analyse du monde qui nous entoure, que la matière, l'énergie, l'espace et le temps. De même, en France, le physicien des hautes énergies Gilles Cohen-Tannoudji proposait d'introduire une quatrième constante fondamentale, une constante « informationnelle » (combinant la constante de Boltzmann et celle de Planck) ayant le même statut que la vitesse de la lumière c, la constante de gravitation G et la constante de Planck h pour rendre compte de l'importance de la notion d'information en physique théorique.

Une vision encore plus radicale est prônée par le physicien d'origine serbe Vlatko Vedral, qui travaille aux Universités d'Oxford et de Singapour et dont les recherches portent sur l'information quantique (un domaine actif de la physique moderne, où l'on cherche à exploiter les phénomènes quantiques pour mieux stocker, traiter, coder ou crypter l'information). Pour ce physicien, tout est information, et c'est cette notion, et non l'énergie, la matière, l'espace ou le temps, qui devrait apporter la clef de la compréhension de l'Univers. Dans cette conception, l'Univers équivaut à une sorte de gigantesque ordinateur.

Du côté de la biologie, la théorie de l'information a également marqué les esprits. Dans les années 1950, elle avait apporté un langage commun aux biologistes dans leurs recherches pour établir le « code » génétique, après la découverte

L'AUTEUR

Jérôme SEGAL est maître de conférences à l'Université Paris IV Sorbonne, actuellement en disponibilité à Vienne où il coordonne le Collège doctoral d'histoire et philosophie des sciences de l'Université de Vienne.

BIBLIOGRAPHIE

F. Conway et J. Siegelman, **Héros pathétique de l'âge de l'information - En quête de Norbert Wiener, père de la cybernétique**, Hermann, 2012.

J. Segal, **Le zéro et le un**, Matériologiques (édition électronique, www.materiologiques.com), octobre 2011.

P. M. Binder et A. Danchin, **Life's demons : information and order in biology**, *EMBO Reports*, vol. 12(6), pp. 495-499, 2011.

V. Vedral, **Decoding Reality – The Universe as Quantum Information**, Oxford University Press, 2010.

M. Triclot, **Le moment cybernétique – La constitution de la notion d'information**, Champ-Vallon, 2008.

G. Cohen-Tannoudji, **Les constantes universelles**, Hachette, 1996.

H. Yockey, **Information Theory and Molecular Biology**, Cambridge University Press, 1992.