

HOMO SAPIENS INFORMATICUS chronique de Gilles Dowek

Pince-mi et pince-moi à l'aune de Kolmogorov

Les notions de complexité et de compression de l'information nous dévoilent une partie des mécanismes de l'humour.



Une devinette lors d'un dîner : « Justine et Juliette sont dans le bateau du marquis de Sade. Justine tombe à l'eau. Qui est-ce qui reste ? Pince-moi bien sûr. » Tout le monde rit, sauf l'un des convives – un Japonais. Notre hôte entreprend de lui expliquer l'innocente farce des cours de récréation, le sulfureux marquis, ses séjours à la Bastille et à l'asile de Charenton, les héroïnes de ses romans : Justine et Juliette. Tout un pan de notre culture. Peine perdue. Le Japonais écoute poliment, opine du chef, mais, à la fin de la laborieuse explication, ne rit toujours pas : pour une raison mystérieuse, une histoire cesse d'être drôle quand on l'explique. La théorie algorithmique de l'information, cependant, éclaire peut-être ce mystère d'une lumière nouvelle.

S'il n'est pas évident de définir la quantité d'information qui se cache derrière l'énoncé de la devinette, il est plus simple de le faire avec une suite de caractères. En effet, quand un programme produit une telle suite, nous pouvons comparer la taille de la suite produite à celle du programme lui-même. Le programme `for i=1 to 500 000 print "ab"`, par exemple, produit une suite de un million de caractères : `ababab...ab`, bien qu'il n'en comporte lui-même qu'une trentaine. Ce programme est beaucoup plus court que la suite qu'il produit, parce que cette suite est très régulière : elle se laisse compresser en un très petit programme.

En théorie de l'information, la « complexité de Kolmogorov » d'une suite de caractères est la taille du plus petit programme qui la produit. Elle mesure la quantité d'informa-

tion de la suite bien mieux que la taille de la suite elle-même.

Cette notion de complexité sert à définir la notion d'aléa : selon Andreï Kolmogorov, une suite est « aléatoire » quand elle est totalement irrégulière, impossible à compresser, c'est-à-dire quand sa complexité est égale à sa taille. Mais cette notion est également intéressante en soi, car l'art de dire beaucoup avec peu de mots se retrouve partout : en grammaire, où



UNE BLAGUE, UN ALGORITHME
ou un geste comme celui peint dans ce portrait de Gabrielle d'Estrées (vers 1594) véhiculent des informations compressées.

l'usage de pronoms permet de remplacer un mot par un autre plus bref, qui signifie provisoirement la même chose ; dans la caricature, qui croque en quelques traits une personne et ses défauts ; dans le roman, qui fait tenir la vie d'un personnage en quelques centaines de pages ; dans la poésie, qui évoque davantage qu'elle n'explique ; dans la science, qui fait tenir tout l'Univers dans une seule équation ; etc.

Parler, écrire, écouter, lire, comprendre, etc. nous demandent un effort. C'est pour-

quoi nous préférons dire « ketch », plutôt que « voilier à deux mâts, dont le mât avant est plus grand que le mât arrière ». Ainsi, dans une énonciation, la taille de l'énoncé, « ketch », est un coût, tandis que celle de la suite engendrée dans la tête de l'interlocuteur, « voilier à deux mâts, dont le mât avant... », est un bénéfice.

Parfois, le rapport bénéfice/coût est extraordinaire : la simple expression « Pince-moi » convoque dans notre tête l'innocente farce des cours de récréation, le sulfureux marquis, Justine et Juliette, etc., y provoque la collision de la candeur et de la cruauté, de deux âges, de deux époques : un véritable feu d'artifice. Tant d'information véhiculée à si peu de frais nous ravit. Et ce ravissement provoque la contraction des muscles de notre visage, l'ouverture de notre bouche, les mouvements de notre appareil respiratoire et l'enchaînement de petites expirations saccadées : nous rions.

Mais quand l'histoire nous est expliquée, le coût de l'énonciation croît et le rapport entre son bénéfice et son coût décroît en conséquence. En deçà d'un certain seuil, les muscles de notre visage ne se contractent plus, notre bouche reste fermée et notre appareil respiratoire au repos.

Cela ne nous empêche pas d'être polis : « Merci. Cette histoire de Pince-mi et Pince-moi était très intéressante. »

Gilles DOWEK est chercheur à l'Inria et membre du conseil scientifique de la Société informatique de France. Il vient de publier avec Serge Abiteboul le livre Le temps des algorithmes (Le Pommier, 2017).