INSTITUT DE FRANCE
Académie des sciences

Séance solennelle de l'Académie des sciences / 17 juin 2008
Réception sous la coupole de l'Institut de France des Membres élus en 2007

Le codage correcteur ou l'art de la redondance

Claude Berrou

Lorsque j'étais sur les bancs de la petite école, on m'enseignait que la redondance, de la même famille que les redites et les pléonasmes, était inutile et disgracieuse. Aujourd'hui, je sais, tout comme vous, que la redondance est indispensable au développement et à la survie des organismes vivants ainsi qu'au bon fonctionnement des machines. La redondance, lorsqu'elle est bien construite - c'est un problème essentiellement mathématique - et judicieusement exploitée – c'est alors affaire de mathématique et de physique - est aussi au cœur des systèmes d'information. Les télécommunications modernes ne peuvent se passer de redondance pour fonctionner dans des conditions toujours plus difficiles de rapport signal à bruit et d'interférences. Dans ce monde des télécommunications, la « science de la redondance » s'appelle le codage correcteur d'erreurs ou codage de canal, dont les principes relèvent d'une théorie plus générale qui est la théorie de l'information.

Les fondements de la théorie de l'information et des communications numériques modernes ont été posés par Claude Shannon il y a exactement soixante ans. En particulier, Shannon a établi une importante limite théorique concernant la qualité envisageable d'une transmission numérique, par le moyen d'un code correcteur d'erreurs, lequel restait à découvrir.

Durant cinquante ans, ce résultat théorique a constitué pour des milliers de chercheurs et d'ingénieurs un défi scientifique majeur car l'enjeu économique était important. Améliorer le pouvoir de correction d'un code, c'est à même qualité d'information reçue (par exemple en téléphonie numérique, pas plus d'une information binaire fausse sur 10.000 reçues), permettre au système de transmission de fonctionner dans des conditions plus sévères. Il est alors possible de réduire la taille des antennes, le niveau de puissance à l'émission ou le poids des batteries d'alimentation. Dans les systèmes spatiaux (satellites, sondes, ...), l'économie peut être considérable, car le poids des équipements et la puissance du lanceur s'en trouvent notablement réduits. Dans les systèmes cellulaires de téléphonie mobile, améliorer le code, c'est aussi permettre à l'opérateur d'augmenter le nombre d'utilisateurs potentiels dans la cellule ou d'accroître l'autonomie en énergie du portable.

En 1990, l'état de l'art était fixé par le code correcteur aujourd'hui utilisé dans la télévision numérique terrestre. Il s'agit d'un code concaténé ou code gigogne car deux codes emboîtés protègent mieux qu'un seul le message à transmettre comme le font deux enveloppes, au lieu d'une seule, autour d'une feuille de papier. Mais la limite calculée par Shannon n'était toujours pas

atteinte, d'un facteur deux à trois sur le rapport signal à bruit, et accroître le nombre de codes concaténés n'apporte apparemment aucun gain supplémentaire. On commençait donc à s'habituer à l'idée que la limite théorique était inaccessible.

C'est sur les conseils d'un collègue de TELECOM Bretagne, Alain Glavieux, éminent spécialiste des communications numériques bien trop tôt disparu, que j'ai alors orienté l'activité du laboratoire circuits intégrés que je venais de mettre en place vers les modulations et le codage. Je m'intéressais de près aux travaux d'un autre professeur de l'Institut TELECOM, Gérard Battail, l'un des pionniers de la théorie de l'information en France. Les résultats que j'ai obtenus par la suite doivent beaucoup à l'expertise et à la gentillesse de ces deux collègues.

L'algorithme de Viterbi à entrée et sortie pondérées que Gérard Battail a développé est un excellent sujet d'études pour ce que l'on appelle l'interaction algorithme-silicium. Admettant des probabilités à son entrée et capable d'en produire à sa sortie, cet algorithme de décodage peut être utilisé avec profit dans le traitement en série de fonctions diverses. Cependant, trop complexe dans sa version originale, il doit être simplifié pour une mise en œuvre sur silicium, sans perte de performance si possible. Me voilà donc passant des journées entières dans les treillis de codes convolutifs, en compagnie de logarithmes de probabilités et sous le double commandement d'inférences bayésiennes et de contraintes de complexité. C'était en 1990 et ce travail fut certainement le plus intense de tous ceux que j'ai entrepris. J'étais loin d'imaginer ce qui allait en découler et qui marqua pour toujours mon parcours de chercheur. François Jacob avait bien raison de dire « L'imprévisible est dans la nature même de la science ».

Grâce à l'algorithme de Viterbi modifié que j'avais entre les mains, j'étais capable de décoder deux codes convolutifs concaténés, voire plus, ce qui était mon objectif initial. Me vient alors une idée saugrenue : si l'on admet qu'un décodeur de Viterbi à entrée et sortie pondérées est assimilable à un amplificateur de rapport signal à bruit parce qu'il amplifie les logarithmes de rapport de vraisemblance, pourquoi donc n'utilise-t-on pas la technique de contre-réaction dans une cascade de décodeurs ? La contre-réaction est un principe largement exploité en électronique. "Inutile" répondent les experts et, sous un certain angle, les mathématiques : lorsqu'on enchaîne des traitement localement optimaux, le résultat global est également optimal.

Je persiste malgré tout et imagine coup sur coup concaténation parallèle, codes récursifs, permutation pseudo-aléatoire et information extrinsèque, des concepts aujourd'hui bien ancrés dans la théorie de l'information et qui sont les ingrédients de la contre-réaction probabiliste. Avec le concours d'Alain Glavieux et le soutien sans failles de France Télécom, j'obtiens enfin des résultats pratiquement conformes aux prévisions de Shannon. Je donne alors à ce nouveau procédé le nom de turbo-décodage par analogie avec le moteur turbocompressé dans lequel les gaz d'échappement – ce qui est normalement perdu – sont réutilisés. Le moins que l'on puisse dire est que notre première communication en 1993 fut reçue avec beaucoup d'étonnement et même de scepticisme, notamment de la part de l'école américaine de codage jusqu'à ce que, quelques mois plus tard, la NASA fût en mesure de confirmer nos résultats.

Pour reprendre l'image toute relative des deux enveloppes successives autour de la feuille de papier, il était donc démontré que l'enveloppe interne est également capable de protéger l'enveloppe externe et qu'il n'est pas nécessaire de multiplier le nombre d'enveloppes. Plus concrètement, deux composantes de redondance issues de deux codes différents peuvent se protéger l'une l'autre quel que soit leur ordre de construction si l'on autorise les deux décodeurs élémentaires à s'échanger des informations dans les deux sens.

Le principe turbo fut plus tard reconnu comme une instance particulière du principe très général de propagation de croyance qui gouverne le traitement probabiliste et distribué de l'information. Les turbocodes qui en furent la première application ont été adoptés dans de nombreux standards de télécommunications, notamment la norme mondiale de troisième génération de téléphonie mobile. J'ai également été, avec plusieurs collègues de TELECOM Bretagne, à l'origine de l'extension des applications du principe turbo à des fonctions de traitement de l'information autres que le décodage correcteur d'erreurs, telles que la détection, la démodulation ou l'égalisation. Aujourd'hui, qu'on l'appelle principe turbo ou propagation de croyance, le concept de communication distribuée probabiliste continue d'ouvrir des perspectives en dedans comme en dehors du champ des télécommunications. Il ne serait pas surprenant par exemple qu'il contribue, dans les années à venir, au progrès des sciences cognitives.

Je suis heureux d'être parmi vous aujourd'hui pour différentes raisons, particulièrement pour y représenter cette dynamique région de Bretagne qui a tant apporté aux sciences de l'information et aux télécommunications.