



Les découvertes oubliées d'Alan Turing

JACK COPELAND • DIANE PROUDFOOT

Connu pour la machine qui porte son nom, il fut aussi le père des réseaux de neurones.

Tous les ordinateurs actuels sont des descendants de la «machine de Turing» conçue en 1935 par Alan Mathison Turing. Ce mathématicien britannique fut également un pionnier de l'intelligence artificielle : il proposa un test – fameux, mais controversé – permettant de déterminer si un ordinateur «pense». Au cours de la Seconde Guerre mondiale, il contribua à décrypter le code allemand *Enigma* dans le cadre d'une opération britannique secrète qui, d'après les historiens, abrégua de deux ans la guerre en Europe. Au moment de sa mort à 41 ans, Turing travaillait sur ce que l'on nomme aujourd'hui la «vie artificielle», simulant la chimie de la croissance – de la morphogénèse.

Tout au long de sa carrière, Turing s'est peu préoccupé de faire connaître ses idées, de sorte que certains aspects de ses travaux sont passés inaperçus ou ont été oubliés. Rares sont ceux, même parmi les spécialistes de l'informatique, qui connaissent les travaux de Turing sur l'informatique «neuronale». De surcroît, ses conceptions novatrices des «hyperordinateurs» n'ont pas été prises en compte ; des spécialistes pensent que les hyperordinateurs de Turing pourront un jour résoudre des problèmes jugés insolubles aujourd'hui.

Les ordinateurs dévorent les nombres. En quelques secondes, ils calculent la trajectoire des fusées ou les résultats comptables des grandes sociétés. En revanche, la programmation des actions apparemment simples, que nous accomplissons sans même y penser, telles que reconnaître un visage ou lire un livre, est autrement difficile. Les réseaux de neurones du cerveau humain ont des capacités dont sont encore dépourvus les ordinateurs ; aussi, les informaticiens tentent-ils de reproduire, par l'informatique, le fonctionnement du cerveau humain.

Le connexionnisme est cette nouvelle branche de l'informatique où les calculs sont effectués par des réseaux de neurones artificiels. Les informaticiens simulent les neurones et leurs

connexions sur des ordinateurs (tout comme les ingénieurs créent des modèles virtuels d'ailerons d'avion ou de gratte-ciel). Un algorithme d'apprentissage ajuste les connexions entre neurones, pour que la machine, dédiée à une tâche précise, l'accomplisse de mieux en mieux.

Les connexionnistes modernes considèrent Frank Rosenblatt, qui publia le premier article sur le sujet en 1957, comme le fondateur de la discipline ; or, dès 1948, Turing avait étudié les réseaux de neurones dans un article intitulé *Intelligent Machinery*.

LA MACHINE DE TYPE B

Turing écrit cet article alors qu'il travaille au Laboratoire de physique de Londres. Considérant que l'article ne vaut guère mieux qu'une «dissertation d'écolier», son directeur, Charles Darwin – le petit-fils du naturaliste – le rejette. En réalité, cet article visionnaire était le premier de tous ceux qui allaient être publiés sur l'intelligence artificielle. Dans ce manuscrit, qui ne fut publié qu'en 1968, 14 ans après sa mort, le mathématicien britannique formulait les principes du connexionnisme, et introduisait plusieurs concepts fondateurs de l'intelligence artificielle, que d'autres redécouvrirent ultérieurement.

Dans cet article, Turing invente une sorte de réseau de neurones, qu'il dénomme «machine non organisée de type B», constituée de neurones artificiels et de dispositifs qui modifient leurs connexions. Les machines de type B incluent autant de neurones que l'on veut, connectés comme on veut, mais chaque connexion entre deux neurones passe par un dispositif modificateur.

Tous les modificateurs de connexions ont deux fibres d'apprentissage. Une impulsion sur l'une d'entre elles met le modificateur en «mode passage», dans lequel une entrée, soit 0 soit 1, traverse sans être modifiée et devient une sortie. Une impulsion sur l'autre fibre met

le modificateur en «mode interruption», où la sortie est toujours égale à 1, quelle que soit l'entrée. Dans cette configuration, le modificateur détruit toute information qui traverse la connexion à laquelle il est lié.

Une fois initialisé le modificateur conserve sa fonction («passage» ou «interruption»), à moins qu'il ne reçoive une impulsion sur l'autre fibre d'apprentissage. Ces modificateurs de connexions, dont le principe est très astucieux, permettent à une machine non organisée de type B d'apprendre au moyen de ce que Turing appelait «des interférences appropriées», qui imitent les mécanismes éducatifs. Selon Turing, «le cortex d'un jeune enfant est une machine non organisée, qui se structure grâce à un entraînement adapté».

Chaque neurone du modèle de Turing a deux fibres d'entrée, et la sortie est une fonction logique simple de ses deux entrées. Chaque neurone du réseau effectue la même opération logique suivante : si l'une des deux entrées est 0, la sortie est 1 ; si les deux entrées sont 1, la sortie est 0. Ces neurones sont de type NAND (de l'anglais *not and*)

Turing avait choisi ce type de porte parce que toutes les autres opérations logiques (ou booléennes) sont exécutables par des ensembles de neurones de ce type. De plus, il a montré que les modificateurs de connexions eux-mêmes peuvent être créés à partir de neurones NAND. Ainsi, Turing construisait un réseau constitué uniquement de neurones NAND et de leurs fibres de connexion, un modèle de cortex parmi les plus simples que l'on puisse imaginer.

En 1958, Rosenblatt définit le connexionnisme : «L'information est mémorisée sous forme de nouvelles connexions, ou de canaux de transmission dans le système nerveux (ou par la création de conditions équivalentes à de nouvelles connexions)». La destruction de connexions existantes peut être fonctionnellement équivalente à la création de nouvelles connexions ; aussi, pour

construire un réseau destiné à accomplir une tâche spécifique, les informaticiens peuvent partir d'un réseau qui aurait trop de connexions et en détruire certaines. Les machines de type B de Turing apprennent par création et destruction de connexions.

Initialement, les machines de type B contiennent des connexions interneuronales aléatoires dont les modificateurs ont été mis au hasard sur «passage» ou «interruption». Au cours de l'apprentissage, les connexions inutiles sont détruites par basculement sur le mode interruption des modificateurs qui leurs sont connectés. Inversement, en faisant

passer un modificateur du mode interruption au mode passage, on crée une nouvelle connexion. Ces inactivations et activations sélectives des connexions façonnent le réseau aléatoire initial en un réseau organisé pour une tâche donnée.

Turing étudia d'autres sortes de machines initialement non organisées ; il rêvait de simuler un réseau neuronal et les mécanismes d'apprentissage à l'aide d'un ordinateur. Il voulait «laisser tourner le système pendant un bon moment et l'examiner ensuite, comme un inspecteur d'école, pour constater les progrès réalisés». Hélas, ses travaux venaient trop tôt : ce n'est qu'en 1954,

l'année de la mort de Turing, que Belmont Farley et Wesley Clark réussirent la première simulation d'un petit réseau de neurones sur un ordinateur.

Un réseau spécialisé pouvait-il devenir universel? Turing muni d'un crayon et de papier a montré qu'un réseau de neurones de type B suffisamment grand pouvait être configuré (grâce à des modificateurs de connexions) et devenir un ordinateur généraliste. Cette découverte éclairait l'un des principaux aspects de l'apprentissage humain.

En effet, les mécanismes cognitifs sont séquentiels et complexes, faisant intervenir le langage ou d'autres formes

Turing et le connexionnisme

Dans un article qui ne fut publié que 14 ans après sa mort (ci-contre), Alan Turing décrivait un réseau de neurones artificiels connectés de façon aléatoire. Cherchant à reproduire les réseaux de neurones du cerveau (a), Turing a conçu une «machine inorganisée de type B» (b) : chaque connexion traverse un modificateur qui est réglé de façon à laisser passer la donnée inchangée (fiche verte), ou à détruire l'information transmise (fiche rouge). Le basculement des modificateurs d'un mode à un autre assure l'apprentissage du réseau. Chaque neurone a deux entrées (voir le cartouche) et exécute l'opération logique simple suivante : si les deux entrées sont 1, la sortie est 0 ; sinon la sortie est 1.

A typical example of an unorganised machine would be as follows. The machine is made up from a rather large number N of similar units. Each unit has two input terminals, and has an output terminal which can be connected to the input terminals of other units. We may imagine that for each integer r , $1 \leq r \leq N$ two numbers, $i(r)$ and $j(r)$ are chosen at random.

Dans le réseau de Turing les neurones s'interconnectent librement. En revanche, dans les réseaux modernes (c) le flux d'informations ne circule que d'une couche à l'autre.

