De l’intelligence humaine à l’intelligence artificielle

David Bensoussan

De la matière au vivant

La science naît d'abord de la curiosité : celle de comprendre le monde qui nous entoure. Elle progresse en confrontant nos intuitions à l'observation et à l'expérimentation, puis en cherchant à formuler, grâce au langage des mathématiques, des lois capables de décrire et de prévoir les phénomènes naturels.

La mécanique classique, dont les fondements sont associés à Newton, a constitué une étape majeure de cette aventure. Elle a permis de décrire avec une remarquable précision le mouvement des corps, leur vitesse, leur accélération et les forces auxquelles ils sont soumis. L'étude de l'électricité et du magnétisme, puis leur unification dans la théorie de l'électromagnétisme, ont ouvert un nouveau chapitre de la physique et rendu possibles d'innombrables applications.

Notre civilisation bénéficie aujourd'hui de ces avancées dans presque tous les domaines : communications, moteurs, électronique, navigation, imagerie médicale, traitements thérapeutiques et exploration spatiale. La physique nous permet même d'étudier des objets situés à des milliards d'années-lumière et, grâce à la lumière qu'ils nous transmettent, de déterminer leur composition et de reconstituer une partie de l'histoire de l'Univers.

Mais en pénétrant dans le monde de l'infiniment petit, la physique a rencontré un domaine où notre intuition quotidienne cesse d'être fiable. À l'échelle des atomes et des particules, la mécanique classique doit laisser place à la mécanique quantique. Il ne s'agit pas simplement de lois différentes pour des objets plus petits : c'est une autre manière de décrire la réalité physique, fondée notamment sur la probabilité, la quantification de

certaines grandeurs et la dualité entre les aspects ondulatoires et corpusculaires de la matière.

Pour autant, la mécanique classique n'est pas devenue fausse. Elle demeure une excellente approximation dans le monde que nous connaissons. La difficulté fondamentale de la physique contemporaine se situe plutôt dans la recherche d'une description cohérente de phénomènes qui relèvent à la fois de la physique quantique et de la gravitation, dont la théorie générale de la relativité fournit la description actuelle à grande échelle.

Notre compréhension de l'atome a elle aussi connu une révolution. Le mot grec *atomos* signifiait « indivisible ». L'atome s'est pourtant révélé être une structure complexe. Il possède un noyau composé de protons et de neutrons, autour duquel se trouvent des électrons. Les électrons sont considérés comme des particules élémentaires, tandis que les protons et les neutrons sont eux-mêmes constitués de quarks, liés par l'interaction forte.

La matière apparaît ainsi comme une organisation hiérarchique de constituants et d'interactions. Les particules forment des atomes ; les atomes se combinent pour former des molécules ; les molécules peuvent à leur tour s'organiser en structures beaucoup plus complexes. De nouvelles propriétés apparaissent alors, qui ne sont pas nécessairement évidentes lorsqu'on considère séparément les constituants élémentaires.

La chimie constitue une étape essentielle de cette montée en complexité. Certaines molécules, notamment les grandes molécules organiques, peuvent atteindre une extraordinaire complexité. Parmi elles, les protéines jouent des rôles essentiels dans le fonctionnement des organismes, tandis que l'ADN permet de conserver et de transmettre l'information génétique.

À un niveau d'organisation supérieur apparaît la cellule, unité fondamentale du vivant. Une cellule n'est pas simplement un assemblage de molécules : elle possède une organisation interne, échange de la matière et de l'énergie avec son environnement, régule ses fonctions et, pour les organismes

capables de se reproduire, contribue à la transmission de l'information génétique.

C'est ici qu'apparaît l'une des grandes questions de la science : comment une matière constituée de particules, d'atomes et de molécules peut-elle donner naissance à des systèmes capables de se maintenir, de se reproduire et de transmettre de l'information ?

Il n'est pas nécessaire d'invoquer une « force vitale » particulière pour poser cette question. La biologie moderne cherche plutôt à comprendre comment des propriétés nouvelles peuvent émerger de l'organisation et de l'interaction d'un grand nombre de constituants physiques et chimiques. Le vivant ne viole pas les lois de la physique ; il manifeste des niveaux d'organisation et de complexité qui deviennent pertinents lorsque la matière s'organise de manière suffisamment élaborée.

L'évolution ajoute une nouvelle dimension à cette histoire. Des variations génétiques apparaissent au cours des générations, notamment par mutation et recombinaison. La sélection naturelle, ainsi que d'autres mécanismes évolutifs, modifient les fréquences de ces variations au sein des populations. Sur des temps extrêmement longs, ces processus ont conduit à une extraordinaire diversification du vivant.

À chaque étape, la matière ne change pas de nature fondamentale, mais son organisation devient plus complexe et fait apparaître de nouvelles propriétés. Cette idée d'émergence constitue l'une des grandes énigmes intellectuelles de la science : comprendre comment des comportements collectifs, parfois extrêmement sophistiqués, peuvent apparaître à partir d'un grand nombre de constituants soumis à des lois physiques relativement simples.

De l'infiniment petit à l'infiniment grand, la science nous conduit ainsi à une même interrogation : comment les lois fondamentales de la nature permettent-elles l'apparition de la complexité ? C'est cette question, à la frontière de la physique, de la chimie, de la biologie et de la cosmologie, qui

ouvre la porte à une compréhension plus profonde de notre Univers et de notre propre existence.

De l'évolution du vivant à l'évolution technologique

L'idée d'une montée en complexité ne s'arrête pas au monde vivant. Elle peut également nous aider à comprendre l'extraordinaire développement des technologies humaines. Il faut toutefois être prudent : l'évolution technologique n'est pas une évolution biologique. Elle n'obéit pas directement aux mécanismes de la mutation génétique et de la sélection naturelle. Elle résulte de l'intelligence humaine, de la transmission des connaissances, de l'expérimentation et de la sélection des solutions qui se révèlent les plus efficaces.

L'histoire des semi-conducteurs en fournit un exemple remarquable.

Un semi-conducteur possède des propriétés électriques qui se situent entre celles d'un conducteur et d'un isolant. Mais sa véritable importance tient à la possibilité de contrôler la circulation des charges électriques. Cette maîtrise a conduit à l'invention du transistor au milieu du XXe siècle.

Le transistor constitue une étape décisive. À lui seul, il est un composant relativement simple : il peut notamment fonctionner comme un interrupteur électronique ou comme un amplificateur. Mais des transistors peuvent être assemblés pour réaliser des fonctions logiques. Des millions, puis des milliards de ces composants peuvent ainsi être réunis sur une même puce électronique.

Une nouvelle étape est aujourd’hui explorée avec les ordinateurs quantiques, qui utilisent les propriétés de la mécanique quantique pour traiter certaines classes de problèmes selon une logique fondamentalement différente. Dans ces domaines, ils pourraient effectuer en quelques secondes des calculs qui demanderaient aux ordinateurs classiques un temps considérablement plus long, ouvrant ainsi des perspectives nouvelles en

cryptographie, en simulation moléculaire, en optimisation et dans certains domaines de l'intelligence artificielle.

L'évolution des ordinateurs a ensuite intégrée à la mise en réseau des machines. Un ordinateur isolé possède une capacité de calcul et de stockage limitée. Des ordinateurs reliés entre eux peuvent partager des informations et des ressources. Avec Internet, l'humanité a construit une infrastructure mondiale dans laquelle des milliards de dispositifs peuvent communiquer.

Cela n'est évidemment pas une chaîne biologique. Elle repose sur une autre forme d'hérédité : l'information et le savoir. Une innovation peut être copiée, améliorée, combinée avec d'autres innovations et transmise à la génération suivante d'ingénieurs. Le progrès technologique possède ainsi une forme de mémoire collective.

L'intelligence artificielle représente une étape nouvelle dans cette histoire. Les premiers ordinateurs étaient principalement conçus pour exécuter des instructions explicites. Avec l'apprentissage automatique, on cherche plutôt à construire des systèmes capables d'ajuster leurs paramètres à partir de données et d'exemples.

Dans les réseaux neuronaux artificiels, par exemple, un grand nombre d'unités de calcul sont reliées entre elles. Chaque unité est relativement simple, mais l'organisation collective du réseau permet de réaliser des opérations extrêmement complexes : reconnaître des images, traiter du langage, détecter des régularités ou produire des prédictions.

Une analogie avec le vivant devient alors particulièrement intéressante. Dans le cerveau biologique, les neurones sont reliés par des réseaux extraordinairement complexes et leur activité collective donne naissance à des capacités cognitives. Dans une intelligence artificielle, des éléments de calcul sont également organisés en réseaux et leurs paramètres sont ajustés au cours d'un processus d'apprentissage.

Il ne faut cependant pas en conclure qu'un réseau neuronal artificiel est un cerveau, ni que l'intelligence artificielle possède nécessairement les mêmes propriétés que l'intelligence biologique. Les architectures, les mécanismes physiques et les processus d'apprentissage sont très différents. L'analogie est surtout intéressante pour comprendre un même principe général : la complexité peut émerger de l'organisation d'un grand nombre d'éléments relativement simples.

Une autre différence est essentielle. Dans l'évolution biologique, la transmission de l'information génétique et la reproduction des organismes jouent un rôle central. Dans l'évolution technologique, ce sont les êtres humains qui conçoivent, sélectionnent, transmettent et combinent les innovations. Aujourd'hui, cependant, les systèmes informatiques participent eux-mêmes à la conception de nouveaux logiciels, de nouveaux circuits et de nouveaux modèles d'intelligence artificielle. La frontière entre l'outil et le concepteur commence ainsi à devenir plus subtile.

Dans le domaine de la technologie et du monde naturel, l'élément fondamental n'est peut-être pas seulement la matière, mais l'organisation de la matière et la circulation de l'information.

C'est ici que les deux histoires — celle du vivant et celle de la technologie — commencent à se rejoindre. La biologie a produit des systèmes capables de traiter et de transmettre de l'information. L'être humain, produit de cette évolution biologique, a ensuite construit des systèmes artificiels capables eux aussi de traiter, stocker et transformer l'information.

La question devient alors vertigineuse : sommes-nous simplement en train de prolonger l'évolution du vivant par des moyens technologiques, ou sommes-nous en train de faire émerger une nouvelle forme d'organisation de la matière et de l'information ?

Cette question ne possède pas encore de réponse définitive. Mais elle permet de regarder autrement l'histoire qui nous a conduits du transistor à l'intelligence artificielle.

L’intelligence artificielle : jusqu’où ira l’autonomie des machines ?

Cette progression de la matière vers l'information, puis de l'information vers des systèmes capables d'apprentissage, nous conduit naturellement à une dernière question : quels sont les risques associés à l'intelligence artificielle ?

De nombreuses voix tirent aujourd'hui la sonnette d'alarme. Ces inquiétudes ne portent pas seulement sur les erreurs ou les usages malveillants de l'IA, mais également sur la possibilité de voir apparaître des systèmes dont les capacités dépasseraient progressivement celles de leurs concepteurs dans certains domaines.

Il faut toutefois comprendre ce que l'on appelle ici « intelligence ». Une intelligence artificielle ne raisonne pas nécessairement comme un être humain. Selon les méthodes employées, elle peut explorer ou estimer un très grand nombre de possibilités, repérer des régularités dans d'immenses quantités de données et sélectionner les réponses qui apparaissent les plus pertinentes selon les critères qui lui ont été donnés ou qu'elle a appris. Dans certains cas, elle peut ainsi trouver une solution à un problème que l'être humain aurait eu beaucoup de difficulté à résoudre directement.

Cette capacité peut produire des résultats spectaculaires sans que la machine possède pour autant une compréhension du monde comparable à celle d'un être humain. Une IA peut être extrêmement performante dans une tâche et commettre, dans une autre situation, une erreur surprenante. Elle peut également produire une réponse plausible mais fausse, phénomène souvent désigné sous le terme d'« hallucination ». L'IA n'est donc pas infaillible.

Une difficulté plus subtile apparaît lorsque les systèmes deviennent capables d'apprendre et de réutiliser des informations dans de nouveaux contextes. Supposons, par exemple, qu'un système découvre ou identifie une démonstration mathématique nouvelle. Cette information peut être conservée dans son dossier de connaissances et être ultérieurement

mobilisée pour résoudre un problème dans un contexte différent. La capacité à accumuler, combiner et réutiliser des connaissances donne alors au système une forme de continuité que ne possédaient pas les programmes traditionnels.

Le véritable enjeu apparaît lorsque cette capacité est associée à l'autonomie. Un système capable de planifier, d'utiliser des outils, d'écrire et d'exécuter du code, d'interagir avec d'autres systèmes et de poursuivre un objectif sur une longue période pourrait accomplir des tâches dont il serait difficile pour un être humain de suivre chacune des étapes.

C'est dans ce contexte qu'est évoquée l'hypothèse d'une superintelligence artificielle : un système dont les capacités intellectuelles dépasseraient très largement celles des humains dans la plupart des domaines pertinents. Une telle situation demeure hypothétique, mais elle soulève une question essentielle : comment conserver un contrôle humain effectif sur un système dont les capacités de traitement, de mémorisation, d'apprentissage et d'action seraient supérieures aux nôtres ?

Le problème devient d'autant plus important que l'intelligence artificielle ne reste pas confinée aux laboratoires. Elle s'intègre progressivement à la finance, à l'industrie, aux communications, à la médecine, à la recherche, à l'éducation, à la sécurité et à l'administration. Elle devient ainsi une infrastructure susceptible d'influencer une part croissante des activités humaines.

Le risque ne réside donc pas uniquement dans une machine qui deviendrait soudainement « consciente » ou hostile. Il peut aussi résider dans une dépendance progressive : des décisions humaines pourraient être de plus en plus préparées, orientées ou exécutées par des systèmes dont le fonctionnement devient difficile à comprendre dans son ensemble.

L'histoire de la technologie nous enseigne pourtant que les dangers ne proviennent pas seulement des machines elles-mêmes, mais aussi de la manière dont les sociétés choisissent de les concevoir et de les utiliser. Le

défi de l'intelligence artificielle sera donc double : accroître ses capacités tout en maintenant la possibilité de les contrôler, de les vérifier et, lorsque cela est nécessaire, de les interrompre.

Nous sommes ainsi peut-être à une nouvelle étape de cette longue histoire de la complexité. La matière s'est organisée en structures vivantes ; le vivant a produit l'intelligence humaine ; l'intelligence humaine a créé des machines capables de traiter l'information et d'apprendre.

Les premières tentatives d'implantation de puces cérébrales pourraient permettre d'accroître certaines capacités intellectuelles. Mais rien ne permet encore de savoir comment ces technologies pourraient affecter la psyché, entendue comme l'ensemble des manifestations conscientes et inconscientes de l'activité cérébrale.

La question ultime n'est peut-être donc pas de savoir si la machine deviendra « intelligente » comme l'être humain, mais plutôt de savoir jusqu'où nous accepterons de déléguer à des systèmes artificiels notre capacité à comprendre, décider et agir.

C'est peut-être là le véritable paradoxe de notre époque : après avoir créé des machines pour nous assister, nous devons désormais apprendre à nous assurer qu'elles demeurent nos outils — et ne deviennent pas progressivement les arbitres de nos choix.