

# Le paradigme informatique.

---

Extrait de "La machine univers"

Pierre Lévy (Points-Sciences, Paris, 1987), pages 131 à 141.

---

L'emploi croissant de concepts informatiques et cybernétiques dans diverses disciplines, l'usage de plus en plus fréquent de simulations numériques pour étudier les phénomènes, ou la tendance à la mise en algorithme des raisonnements scientifiques sont différentes manifestations de l'émergence d'un paradigme du calcul. Ce paradigme n'est pas un corps de doctrine systématique ayant ses partisans et ses adversaires déclarés. Il se dessine au contraire peu à peu au fil des métaphores, des procédures et des habitudes de la recherche. L'utilisation toujours plus massive des ordinateurs et des systèmes de documentation automatique favorise incontestablement son extension, bien qu'il ne la détermine pas automatiquement. Mais une multitude de petits changements, quoique dispersés et peut-être localement imperceptibles ne concourent pas moins à la combinaison d'un ensemble de critères, de schémas intellectuels, de modes de représentation et de principes méthodologiques d'une redoutable cohérence, dont la notion de calcul constitue justement le centre secret.

Les enjeux de l'opération en cours sont multiples. Sur le versant ontologique, les systèmes physiques, vivants, psychiques sont-ils des machines à traiter l'information ? Sur le versant méthodologique, y a-t-il une seule rationalité scientifique, codifiable et applicable en droit à tous les objets ? Enfin et surtout, le but de l'activité scientifique est-il de prévoir et de calculer toujours mieux ou de rendre intelligible et d'éclairer le monde qui nous entoure ?

## **CALCUL ET MÉTHODE**

Un nouveau rapport à l'expérience

La pratique de la simulation sur ordinateur se répand dans les champs les plus divers: mathématiques, physique nucléaire, astrophysique, cosmologie, météorologie, sismologie, chimie, avionique, gestion, économie, démographie, histoire, etc. Un des premiers effets de l'usage de la simulation numérique est de conférer un caractère expérimental à des disciplines qui ne le possédaient pas, comme la cosmologie ou la démographie. Le changement est essentiellement pratique. Rien n'interdisait, en effet, aux chercheurs en astrophysique ou en histoire économique de formaliser des ensembles de relations logiques explicites entre des variables, de faire varier systématiquement les différents paramètres pour étudier le comportement du modèle, puis de comparer les résultats au phénomène en grandeur vraie. Mais avant les ordinateurs, cette possibilité de principe se heurtait à une impraticabilité de fait, le moindre calcul un peu compliqué pouvant occuper plusieurs années d'une équipe de mathématiciens chevronnés.

Bien entendu, les simulations numériques ne sont pas de véritables expériences puisqu'elles ne portent pas sur les phénomènes mais sur des modèles de ceux-ci. Il faut néanmoins remarquer que l'expérience scientifique classique opère déjà une manière d'idéalisation, de purification et de construction artificielle de son objet. On peut donc considérer la simulation numérique comme un degré supplémentaire dans l'artifice visant le contrôle et la purification du phénomène.

La simulation numérique permet de tester des hypothèses beaucoup plus facilement que par la simple observation de phénomènes sur lesquels le chercheur scientifique est incapable d'agir. La pratique de la simulation met le modélisateur dans la nécessité d'explicitier et de justifier le choix de ses variables, ainsi que de formuler une description quantitative précise des relations supposées de cause à effet. La «testabilité» des hypothèses et la formalisation rigoureuse des modèles rapprochent fortement la psychologie,

l'économie, la sociologie ou l'histoire des sciences exactes, au moins sur le plan méthodologique.

Dans les disciplines, comme la physique, qui avaient traditionnellement accès à l'expérimentation, la simulation numérique ajoute un troisième terme à la théorie et aux résultats empiriques. Dans cette nouvelle configuration, l'activité théorique consiste moins à rendre compte de tous les résultats qu'à sélectionner les modèles qui pourraient avoir une signification physique. Le réel est ici pratiquement appréhendé comme un modèle parmi une prolifération de modèles possibles, alors que l'ancienne physique était censée partir uniquement de résultats réels pour dégager le modèle sous-jacent qui les rendrait cohérents. Certes, l'expérience de pensée et l'imagination de modèles possibles ont toujours fait partie de l'activité scientifique. Il reste que l'ordinateur transforme l'expérience de pensée, de bricolage artisanal qu'elle était, en entreprise systématique à grande échelle. Une des plus curieuses modifications liées à l'usage des simulations numériques est celle qui affecte aujourd'hui les mathématiques. Traditionnellement considérées comme le royaume de la déduction, elles prennent à leur tour un caractère expérimental. Des simulations numériques d'objets mathématiques peuvent infirmer, confirmer ou faire naître des conjectures. La conjecture ne devient cependant théorème qu'une fois démontrée. On rencontre ici les limites du calcul, car les démonstrateurs automatiques n'ont encore jamais démontré à ce jour de théorèmes significatifs qui n'aient déjà été prouvés par un mathématicien en chair et en os. Or, en mathématique, seule la démonstration éclaire et rend raison du théorème. Le calcul laisse prévoir, suscite des hypothèses, donne plus de poids à certaines conjectures, il n'explique pas. L'extraordinaire puissance de calcul statistique et d'analyse des données qui sont offertes par l'informatique se conjugue à la simulation pour renforcer le caractère empirique et inductif de la recherche informatisée. L'analyse des correspondances, l'analyse en composantes principales, l'analyse factorielle, les programmes de classification automatique ainsi que de nombreuses autres méthodes statistiques nées avec l'informatique autorisent des traitements extrêmement complexes sur des masses énormes de données qui seraient restées muettes sans les ordinateurs. Il faut ajouter à la puissance de calcul statistique proprement dite les efforts de mise en scène visuelle des résultats par image synthétique qui permettent au chercheur d'appréhender sur un mode sensible d'immenses tableaux de chiffres autrement illisibles.

## **Limites de la simulation numérique**

Comment un calcul, par essence discret, pourrait-il toujours simuler adéquatement un processus continu ? Dans son livre *le Calcul, l'Imprévu*, Ivar Ekeland donne l'exemple du carambolage de quelques boules sur une table de billard. Si nous entrons dans un ordinateur les positions et vitesses initiales de toutes les boules et que nous attendons en sortie les positions et les vitesses à un temps ultérieur, le résultat de la simulation risque fort de n'avoir rapidement plus aucune signification. A cela, deux raisons principales. Premièrement, l'ordinateur ne peut travailler qu'avec un nombre fini de décimales en négligeant les décimales supplémentaires. A chaque nouvelle opération, les erreurs d'arrondi s'amplifient jusqu'à dénaturer grossièrement le résultat final. Secondement, le système réel est soumis à une multitude de perturbations qui vont de la respiration de l'expérimentateur dans la pièce au mouvement d'un noyau d'hydrogène sur Alpha du Centaure. Or, ces perturbations deviennent vite significatives, si bien que le résultat du calcul, même exact, n'en sera pas moins très éloigné du résultat observé. On sait que les trajectoires des vaisseaux spatiaux naviguant vers la Lune doivent être corrigées à intervalles rapprochés en fonction de leur vitesse et de leur position réelle. Les ingénieurs de la NASA savent parfaitement que les ordinateurs, pourtant extrêmement puissants, dont ils disposent ne peuvent calculer le trajet exact des engins spatiaux uniquement à partir des conditions astronomiques et astronautiques initiales. Un autre

type de limite des simulations numériques tient au choix du langage de description. Dans le cas d'une simulation, toutes les décisions quant à ce qui pourrait être pertinent pour l'évolution du système étudié ont été prises d'un seul coup, au moment de la formalisation du modèle. L'expérience réelle, au contraire, peut toujours laisser apparaître l'importance d'un facteur auquel on n'avait pas pensé au moment de sa mise au point. Cela ne signifie pas qu'une simulation numérique est incapable de surprendre le chercheur, mais que les genres de surprises que peuvent offrir la simulation et l'expérience réelle ne sont pas du même ordre. Nous retrouvons ici la différence entre possible et virtuel. Le résultat d'une simulation, même surprenant, était logiquement contenu dans un ensemble de possibles précodés par l'algorithme et les données. Les surprises qui surviennent au cours d'une expérience réelle, en revanche, peuvent provoquer un changement dans la façon d'appréhender le phénomène et amener le chercheur à redéfinir l'objet même de son étude. La surprise «réelle» est capable d'ouvrir un nouveau champ de virtualités.

Il est des disciplines, comme l'histoire, où la légitimité de la simulation est problématique. Une école d'histoire économique américaine, la New Economic History, a mis à l'honneur l'utilisation de modèles formels précisément quantifiés simulés sur ordinateur pour étudier des questions du type: «L'esclavage dans le Sud avant la guerre de Sécession était-il réellement une entrave au développement économique ?», ou: «Peut-on vraiment considérer l'extension du chemin de fer comme un facteur fondamental de la croissance américaine au XIXe siècle ? » Les modèles mis au point essaient de prévoir ce qui serait arrivé si les planteurs avaient utilisé des salariés agricoles et si l'on avait creusé des canaux et construit des routes plutôt que d'étendre le réseau ferré. Pour certains historiens traditionnels, de telles simulations n'ont aucun sens car elles supposent que l'on puisse faire varier seulement les facteurs dont on veut peser l'importance historique, toutes choses étant égales par ailleurs. Or, disent les historiens traditionnels, le «toutes choses égales par ailleurs» n'a aucun sens en histoire. Il serait au contraire très raisonnable de penser que si le Sud n'avait pas été esclavagiste, ou si les chemins de fer ne s'étaient pas développés, énormément d'autres faits de mentalité, de société et d'économie auraient aussi été différents. Ce qui enlève toute valeur aux conclusions des simulateurs.

Les nouveaux historiens de l'économie répondent que presque tous les historiens utilisent implicitement le «conditionnel irréal» chaque fois qu'ils tentent de déterminer les causes d'un phénomène historique. (Si le nez de Cléopâtre avait été plus court, la face du monde en eût été changée.) Au lieu de le faire de façon vague et implicite, ils jugent beaucoup plus conforme à la démarche scientifique d'utiliser le «conditionnel irréal» de manière explicite et formalisée. C'est seulement à cette condition que les hypothèses sur le poids causal de tel ou tel facteur historique peuvent être testées.

Le débat sur l'usage de la simulation en histoire recouvre un enjeu plus profond. Y a-t-il, en droit, homogénéité des procédures méthodologiques et des modes d'explication légitimes d'une discipline à l'autre ? Doit-on reconnaître une seule méthode scientifique, quitte à l'adapter à des objets divers, ou bien, au contraire, faut-il admettre une hétérogénéité fondamentale entre les approches rationnelles d'objets aussi différents que l'univers physique et le monde historique? Si l'on se décidait pour une acception unique et rigoureuse du concept de causalité, par exemple, alors aucun historien, économiste ou psychologue ne pourrait parler de cause sans formaliser un modèle quantifié où la cause en question s'incarnerait dans la liaison fonctionnelle de quelques paramètres. On voit le risque de stérilité qu'une telle exigence pourrait faire courir à nombre de disciplines. Comment quantifier le rôle de l'imaginaire social, par exemple ? Cela ne signifie pas que la modélisation ne soit souvent utile et légitime par plaques locales, mais seulement qu'il serait néfaste d'en faire une condition de scientificité a priori. Les ordinateurs permettent d'utiliser des outils statistiques sophistiqués, donnent accès à la simulation de modèles numériques et contribuent largement à la formalisation des procédures de la recherche.

Ils peuvent ainsi conférer un caractère inédit de rigueur et d'exactitude aux sciences humaines. L'informatique serait donc plus qu'un instrument scientifique. Capable de faire passer une discipline d'un état préscientifique au statut de science exacte, elle serait un opérateur métascientifique. Dans le débat épistémologique en cours, l'ordinateur renforce le camp de la méthode scientifique unique parce qu'il lui donne pour la première fois les moyens de ses ambitions.

## **Expliquer, calculer**

Pour évaluer correctement le rôle de la simulation numérique, il importe de distinguer soigneusement la mise en algorithme et la modélisation numérique d'une part, la mathématisation et la théorie d'autre part. La simulation vise avant tout à prévoir ou à reproduire le comportement du système étudié. En intelligence artificielle, par exemple, le calcul sert souvent à obtenir une certaine performance cognitive, sans prétendre l'atteindre par les mêmes moyens que l'esprit humain. Le modèle sous-jacent au calcul est donc instrumental, ou opératoire, plutôt qu'explicatif.

La mathématisation d'un phénomène, en revanche, a pour objectif premier d'élucider sa structure fondamentale, de le rendre intelligible, de mettre en évidence les paramètres pertinents capables d'en rendre compte rationnellement. La capacité prédictive ne vient que par surcroît, pour confirmer la justesse de la théorie mathématique.

En un sens, la théorie physique d'Aristote était prédictive: le feu monte, la terre descend, les mouvements de la sphère céleste sont réguliers, etc. La théorie newtonienne, prédictive elle aussi, découvre d'abord une autre structure du réel: espace homogène, corps s'attirant en fonction de leurs masses et de leur distance. La formule mathématique de la loi de la gravitation exprime dans tous les cas la nature des interactions en jeu, elle permet de plus de prédire parfois le comportement de systèmes simples. Dès que trois corps sont en jeu, la théorie mathématique, bien qu'elle propose un schéma correct, ne permet plus de calculer précisément l'évolution du système à partir de ses conditions initiales. La différence entre calcul et mathématiques ressort encore de la comparaison classique entre les physiques de Newton et d'Einstein. Sur le plan du calcul, c'est-à-dire, au fond d'un point de vue opératoire, la théorie newtonienne peut être considérée comme un cas particulier de celle d'Einstein, moyennant une légère approximation. En revanche, sur le plan du modèle mathématique décrivant la structure de la réalité physique, les deux théories sont radicalement différentes. Disons qu'en général une théorie mathématique ne permet pas nécessairement de prédire et que les calculs, même utiles et corrects à des fins de prédiction, peuvent parfaitement s'effectuer sans théorie ou à l'aide d'une théorie fautive. Il existe en général un très grand nombre d'algorithmes capables d'arriver au même résultat. L'usage de l'ordinateur favorise massivement la simulation opérationnelle, la prédiction et le calcul. Ici encore, l'idéal scientifique est en jeu. Les théories mathématiques sont-elles des fictions utiles pour faciliter les calculs, mais dont on pourrait à la limite se passer au profit d'algorithmes performants? La science doit-elle au contraire viser à rendre intelligible, à expliquer la réalité, le calcul ne jouant à cet égard qu'un rôle instrumental? La fin ultime de la science est-elle l'opération ou l'interprétation?

## **La formalisation du raisonnement scientifique**

La mise en algorithme du raisonnement scientifique peut être envisagée sous deux points de vue, normatif et descriptif. L'approche normative est le fait de l'école logiciste en sciences humaines. L'abord empirique et descriptif est généralement soutenu par des informaticiens.

On rencontre sous la plume d'archéologues, d'historiens ou d'économistes des déclarations parfaitement explicites sur l'idéal du logicisme: toutes les opérations intellectuelles par lesquelles un scientifique passe d'une collection initiale de données à ses propositions finales doivent pouvoir être effectuées par un automate calculateur. Le raisonnement est ici entendu comme l'ensemble des techniques de traitement de l'information utilisées par le chercheur. Pour juger de la rigueur d'un raisonnement, sa programmabilité est un critère essentiel.

Les logicistes déplorent que, dans de nombreux travaux en sciences humaines, les propositions présentées par les chercheurs comme des conclusions valides ne résultent en réalité ni de déductions logiques dont on connaîtrait clairement les prémisses et les différentes étapes, ni de raisonnements inductifs probabilistes corrects à partir de séries définies de données d'expérience ou d'observation.

La programmation du raisonnement suppose l'explicitation de toutes ses étapes. Elle impose la transparence, la publicité des moyens employés par le chercheur. La possibilité de discussion rationnelle est augmentée, tandis que beaucoup d'erreurs sont évitées. Le critère de la programmabilité dissuade enfin de publier les auteurs qui comptent sur des artifices rhétoriques pour convaincre le lecteur plutôt que sur la cohérence de leurs raisonnements et la testabilité de leurs conclusions.

La définition des données initiales peut elle-même faire l'objet d'une procédure informatisée. En archéologie, comme on l'a vu, on se gardera de ranger immédiatement un artefact découvert sur un champ de fouille dans telle ou telle catégorie (hache, grattoir, etc.) en fonction d'une reconnaissance informelle immédiate. On préférera décomposer l'objet en caractères et confier son classement à un programme taxinomique, ou la détermination de son identité à un algorithme de reconnaissance de forme.

Un algorithme n'opère pas sur des données appréhendées dans leurs significations immédiates et informelles. Il ne traite que les symboles de la description des données, au sens typographique du mot «symbole». C'est pourquoi la programmation du raisonnement implique toujours en amont le codage uniforme de toutes les descriptions de faits. Comme elles doivent être testables, les conclusions seront elles-mêmes exprimées dans un langage congruent avec le code de description des faits. L'exigence d'un certain type de rigueur provoque donc une forte pression vers la standardisation du langage scientifique.

L'utilisation croissante des banques de données et des langages documentaires incline également à la normalisation des idiomes scientifiques, mais pour d'autres raisons que la rigueur logique. Une meilleure communication entre les équipes, une utilisation optimale des réseaux de télé. informatique, la portabilité des programmes, autant d'arguments en faveur de la standardisation des langages, dans la perspective d'une rationalisation du travail scientifique. La normalisation des langages peut entrer en opposition avec la stricte logique de la recherche. Supposons que la proposition obtenue au terme de la chaîne de calculs symbolique soit infirmée par les faits. Le chercheur doit alors soupçonner soit que la collection de faits qui fonde ses inductions est insuffisante, soit qu'une ou plusieurs des prémisses de ses déductions sont incorrectes, soit enfin que le langage de description utilisé ne reflète pas adéquatement les paramètres pertinents pour le problème à résoudre. Un langage implique une théorie. Plus encore que l'usage de la simulation numérique des phénomènes, l'exigence de mise en algorithme programmable du raisonnement scientifique accompagne une prise de parti épistémologique en faveur de l'unité de la méthode scientifique. Un des principaux objectifs du logicisme est la métamorphose des sciences humaines en « vraies » sciences. Les sciences humaines ont-elles tout à gagner en formalisant leurs opérations intellectuelles de la sorte ? Il est douteux que la dimension herméneutique ou interprétative de la recherche puisse se ramener entièrement à des calculs inductifs et déductifs ainsi qu'à des tests programmables.

La formalisation du raisonnement scientifique peut être envisagée sur un mode beaucoup plus descriptif et pragmatique. Cette deuxième voie est en général

adoptée par les informaticiens spécialistes de l'intelligence artificielle cherchant à constituer des systèmes experts reproduisant fidèlement les raisonnements des chercheurs. Dans ce cas, l'informaticien ne tente pas d'imposer une norme de scientificité à l'historien ou à l'archéologue, mais le conduit à expliciter le plus parfaitement possible un grand nombre de savoir-faire implicites et d'opérations intellectuelles sous-entendues.

Il est rare qu'on aboutisse par ce moyen à la construction de systèmes formels décidables et complets. Les structures ad hoc et les règles heuristiques auxquelles parviennent les programmeurs ne garantissent pas contre les erreurs, les contradictions ou les circularités logiques, mais elles permettent peut-être de les détecter plus facilement. Même si les systèmes obtenus ne sont pas totalement cohérents, la mise à plat des procédures intellectuelles autorise leur automatisation effective, leur multiplication et leur exportation. De plus, l'expérience de la formalisation de leur savoir-faire est souvent pour les chercheurs l'occasion d'une réflexion épistémologique féconde.

Dans l'état de sciences humaines telles que l'histoire et l'archéologie, les spécialistes en intelligence artificielle constatent que les raisonnements ne sont formalisables que par segments et non d'un bout à l'autre de la chaîne de la recherche. Il reste que la mise en calcul de plaques locales du travail intellectuel impose de proche en proche, en amont et en aval de ces zones calculées, un accroissement de rigueur formelle ainsi qu'une pression vers le codage uniforme des données.

L'intelligence artificielle rencontre dans la simulation du raisonnement scientifique le même type de problèmes théoriques que dans ses autres applications. Pour que les programmes soient autonomes, il faudrait qu'ils soient capables d'apprendre et qu'ils puissent mettre en reuvre une hiérarchie indéfinie de connaissances sur la connaissance (métaconnaissances), ce dont aucun programme n'est encore vraiment capable. D'autre part, les simulations automatisées de raisonnements scientifiques fonctionnent à l'intérieur d'une problématique fixe. On n'en connaît pas qui sache poser de nouvel/es questions, quoique cela semble être la principale vocation du chercheur.