

Sciences
En questions

Jean-Paul Delahaye

**Complexité
aléatoire
et complexité
organisée**

éditions
Quæ

Complexité aléatoire
et complexité organisée

La collection « Sciences en questions » accueille des textes traitant de questions d'ordre philosophique, épistémologique, anthropologique, sociologique ou éthique, relatives aux sciences et à l'activité scientifique. Elle est ouverte aux chercheurs de l'Inra mais aussi à des auteurs extérieurs.

Raphaël Larrère, Françoise Lescourret,
directeurs de collection

Le groupe de travail « Sciences en questions » souhaite favoriser la réflexion critique des acteurs de la recherche sur l'activité scientifique et ses implications. Son ambition première est d'enrichir la réflexion interne à l'Inra en l'alimentant de contributions propres à éclairer, sous une forme accessible et attrayante, les débats contemporains sur la science et la recherche.

Texte rédigé par l'auteur à la suite de la conférence, avec la collaboration de François Rodolphe et d'Elena Rivkine.

© Quæ, Versailles, 2009 ISSN : 1269-8490 ISBN : 978-2-7592-0321-5

Le code de la propriété intellectuelle du 1^{er} juillet 1992 interdit la photocopie à usage collectif sans autorisation des ayants droit. Le non-respect de cette proposition met en danger l'édition, notamment scientifique. Toute reproduction, partielle ou totale, du présent ouvrage est interdite sans autorisation de l'éditeur ou du Centre français d'exploitation du droit de copie (CFC), 20 rue des Grands-Augustins, 75006 Paris, France.

Jean-Paul Delahaye

Complexité aléatoire et complexité organisée

*Une conférence-débat organisée
par le groupe Sciences en questions
Inra, Centre de Paris, 19 octobre 2006*

Éditions Quæ, c/o Inra, RD 10, 78026 Versailles cedex

Préface

Jean-Paul Delahaye, je vous remercie au nom du groupe Sciences en Questions d'avoir accepté de parler de complexité aléatoire et complexité organisée. Il nous a semblé en effet intéressant d'aborder un tel sujet dans le cycle de nos conférences, car la « complexité » est un mot omniprésent et actuellement à la mode, dans le langage courant comme dans les milieux professionnels et scientifiques, sans toutefois que son sens soit clairement défini.

La complexité est attachée à de nombreuses autres notions qui forment une liste assez hétéroclite et le titre de votre conférence indique déjà clairement une séparation majeure. « Complexe » renvoie à « élaboré, organisé », donc porteur de beaucoup d'information et difficile à comprendre. Mais, de façon assez contradictoire, « complexe » renvoie aussi à « déstructuré, désordonné, aléatoire », donc non résumable et non explicable. Le point commun serait la difficulté de compréhension, possiblement due à son impossibilité !

Ces notions d'intérêt très général sont présentes depuis toujours dans les sciences et, en particulier, dans les sciences de la vie. On aimerait bien savoir caractériser le degré d'organisation de systèmes perçus comme complexes, par exemple un organisme ou un écosystème, et disposer de descripteurs et de critères objectifs pour le mesurer. L'autre versant de la complexité, celui de l'aléa, du désordre, mène à la théorie des probabilités et à la statistique. La théorie des probabilités permet de modéliser les phénomènes aléatoires, mais elle ne dit rien sur la nature du hasard. Pourtant éclairer cette nature est intéressant comme le montrent la variété des domaines d'intervention de la statistique (analyse de données expérimentales ou d'échantillonnage, ou analyse de l'information) et celle de ses interprétations, ainsi que les controverses qui ont eu cours dans cette discipline. Les mathématiques sont ainsi concernées par la complexité. Elles apportent aussi un éclairage très intéressant et fondamental sur cette question.

Votre parcours professionnel vous a fait côtoyer ces questions par plusieurs points d'entrée. Vous avez d'abord suivi un cursus classique de mathématiques pures orienté vers l'enseignement et la recherche, et vous avez été professeur agrégé de mathématiques dans l'enseignement secondaire durant quatre années. À l'issue de cette période, vous avez soutenu une thèse d'État en mathématiques en 1982. Vous êtes ensuite passé dans l'enseignement supérieur à l'université de Lille I à Villeneuve d'Ascq où vous êtes professeur depuis 1988.

Votre thèse portait sur la notion d'algorithme manipulant des nombres réels et présentait des résultats sur les limitations logiques de ces algorithmes. Elle se situait bien dans le courant des recherches mathématiques qui s'est développé à la frontière de la logique mathématique et de l'informatique théorique à partir de travaux de personnages comme Kurt Gödel et Alan Turing. À titre anecdotique, cette thèse fut classée en analyse numérique : l'informatique théorique n'avait pas encore émergé comme discipline indépendante.

Après ces travaux sur l'algorithmique à nombres réels, vos travaux de recherche ont porté sur des thèmes d'intelligence artificielle : programmation logique, moteurs d'inférences, théorie computationnelle des jeux. Depuis 1988, vous travaillez également sur des sujets liés à la théorie de la complexité de Kolmogorov.

Vous avez été successivement membre du Laboratoire d'analyse numérique et optimisation de l'Université de Lille I, puis membre du Laboratoire d'informatique fondamentale de Lille (LIFL, UMR CNRS 8022). Vous serez successivement directeur de l'équipe Méthodes et outils théoriques pour la programmation en logique, puis directeur de l'équipe de bio-informatique. Aujourd'hui vous avez rejoint l'équipe d'intelligence artificielle, simulation multi-agents et comportements du LIFL.

Parallèlement, vous avez participé activement au développement de l'enseignement de l'intelligence artificielle et du génie logiciel. Vous avez aussi animé de nombreuses structures, réseaux ou équipes, notamment, en présidant le Conseil scientifique du GDRIA

(groupe de recherche « Intelligence artificielle » coordonné par le CNRS et fédérant 350 chercheurs français dans le domaine de l'intelligence artificielle) en 1995.

Votre intérêt pour les applications de l'informatique, en tant que science et non pas seulement comme moyen de calcul ou de manipulation de données, vous a conduit à prendre part à des recherches orientées vers d'autres disciplines, la biologie surtout, mais aussi l'économie et la finance, et actuellement la psychologie, autour de la perception du hasard (avec Nicolas Gauvrit). C'est ainsi que vous avez été, en 2000 et 2001, directeur de la plate-forme de bio-informatique de la Génopole de Lille, membre de son comité de pilotage scientifique et responsable de la coordination de ses projets transverses technologiques en bio-informatique.

Vous avez aussi été l'un des créateurs du DESS de bio-informatique ouvert en 2001 et responsable avec Max Dauchet de l'équipe « Complexité algorithmique appliquée à la biologie, à l'automatique et à l'apprentissage » de Lille du LIFL depuis 1994. Cette équipe a pris le nom d'équipe de Bioinformatique en janvier 2000.

Cependant c'est surtout par votre contribution diversifiée et prolifique à la vulgarisation et à la promotion de la culture scientifique que beaucoup parmi nous vous connaissent. La vulgarisation des mathématiques est un art difficile et je me permets de dire que vous y excellez. Vous êtes auteur d'un article mensuel dans la revue *Pour la Science* pour la rubrique « Logique et calcul », conseiller auprès de l'*Encyclopædia Universalis* (et auteur de neuf articles liés à l'informatique et à la logique) et auteur de conférences. Vous êtes également impliqué dans de très nombreuses structures comme auteur, animateur ou organisateur dans ce domaine : groupe de travail de la Fête de la science, du Conseil de la culture et de l'information scientifiques et techniques et des Musées en 2001 et 2002. Vous êtes également membre du jury du Prix d'Alembert de la Société mathématique de France en 2000 et 2002, du comité de rédaction de la revue *Les nouvelles d'Archimède*, de l'équipe d'animation culturelle de l'université de Lille I depuis 1998 (chargé de la rubrique « Paradoxes ») et

du comité éditorial de *Interstices* (Site internet d'information sur les thèmes des sciences et technologies de l'information et de la communication, créé à l'initiative de l'Inria).

Vous avez écrit une quinzaine de livres dont plusieurs ont été traduits et vos travaux ont été honorés à plusieurs reprises : par le prix d'Alembert en 1998 (ce prix est décerné tous les deux ans par la Société mathématique de France pour les actions de promotion de la culture scientifique et mathématique) et par le premier Prix auteur en 1999 de la culture scientifique et technique du ministère de l'Éducation nationale, de la Recherche et de la Technologie.

François Rodolphe
Unité Mathématique, Informatique et Génome,
INRA (Jouy-en-Josas)

Complexité aléatoire et complexité organisée

Introduction

On entend souvent dire aujourd'hui que la complexité est devenue un problème fondamental de la science contemporaine, car toutes les disciplines y sont confrontées, et cela sous une multitude de formes. Avant de savoir si une telle affirmation est vraie, une série de questions se posent :

- Qu'est-ce que la complexité ?
- Y a-t-il un concept précis derrière le mot ?
- Que nous disent les mathématiques sur de possibles définitions formelles ?
- Quelles sont les conséquences et les applications du travail de définition théorique ?

C'est à ces questions que nous allons nous consacrer ici.

Nous verrons que certaines réponses sont disponibles depuis que la *théorie de la calculabilité* s'y est intéressée et qu'en prendre connaissance est certainement utile dans bien des domaines. Nous soutiendrons aussi que, loin de fournir une vision dogmatique de la complexité, l'approche proposée par la théorie du calcul laisse ouvertes de nombreuses pistes et permet la prise en compte d'intuitions qu'il faut seulement traduire en tentant de les rendre rigoureuses.

La théorie de la calculabilité est née entre 1930 et 1936 des travaux de Kurt Gödel, Alonso Church et Alan Turing. Elle définit des modèles de calcul – dont la fameuse *machine de Turing* – et précise le concept de *fonction calculable*¹. Son développement s'est

¹ Une machine de Turing est un modèle abstrait d'ordinateur composé d'un mécanisme de calcul et d'un ruban sur lequel la machine écrit et qu'elle peut effacer en déplaçant une tête de lecture-écriture. À chaque machine de Turing est associé un programme qui détermine avec une précision absolue les opérations qu'elle effectue. Les fonctions qu'on peut calculer avec une machine .../...

poursuivi dans les décennies suivantes mettant à la disposition des mathématiciens et des logiciens une famille d'outils remarquables (*machine universelle*, *degrés de calculabilité*, classes des *fonctions primitives récurrentes et fonctions récurrentes*, etc.). Le tout constitue un socle sur lequel l'informatique théorique s'est développée depuis.

En 1965, ces outils conduisirent le mathématicien russe Andreï Kolmogorov à formuler une définition de *la complexité d'un objet fini*. Cette définition, à la fois générale et puissante, est à l'origine d'autres concepts importants utilisés aujourd'hui par les physiciens, les épistémologues, les psychologues et les biologistes. Des applications inattendues de cette définition de la complexité ont été proposées dans les domaines de la classification des données, de la comparaison des images et même du repérage du plagiat et de l'élimination des *pourriels* (les fameux « *spam* » qui encombrant nos boîtes à lettres électroniques).

Cette première définition de la complexité – *la complexité de Kolmogorov* – a été complétée un peu plus tard par une définition surprenante – *la profondeur logique de Bennett* – qui permet de comprendre et de formuler en termes précis l'opposition naturelle entre *complexité aléatoire* et *complexité organisée*. Nous pensons que ces définitions sont fondamentales aussi bien sur le plan scientifique que philosophique. Nous présenterons ces idées théoriques en menant ce qu'on pourrait appeler une *analyse conceptuelle* ; celle-ci réalise le passage de l'idée intuitive et vague, à une série de définitions formelles et puissantes possédant un champ étendu d'applications.

Notre ouvrage vise un public non spécialisé. Nous ferons pour lui un effort particulier pour limiter le recours aux formalismes

.../... de Turing ou avec un algorithme – par définition c'est la même chose – sont les fonctions *calculables* (on dit aussi *récurrentes*). Les fonctions usuelles ($n \rightarrow n^2$, $n \rightarrow n$ -ième nombre premier, etc.) sont calculables. On démontre toutefois qu'il existe des fonctions non calculables : la première d'entre elles est la fonction qui, pour tout programme donné P, indique (a) si le programme finit par s'arrêter ou (b) s'il continue indéfiniment de calculer (indécidabilité de l'arrêt).