

des processus qui leur ont donné naissance. Nous donnons, pour illustrer ce propos, deux instantanés de motifs chimiques tels qu'apparus au cours d'états transitoires du système et qui, à long terme, auraient abouti à une structure en réseau hexagonal de taches. Un traitement approprié de la palette de couleurs met en évidence d'étonnantes similarités entre ces motifs chimiques et ceux portés par les peaux du guépard et du léopard (voir la figure 2).

Dans des conditions un peu particulières («loin du seuil d'instabilité», diraient les spécialistes des systèmes non linéaires), bien d'autres motifs surprenants peuvent apparaître. Dans certains cas, les motifs s'étendent par division des premières taches apparues, évoquant curieusement la division cellulaire au cours de la croissance de tissus biologiques. Enfin notre équipe a récemment observé des croissances en forme de doigts, donnant naissance à d'étranges et transitoires «fleurs» chimiques (voir la figure 6).

Analogies et universalité

Aussi extraordinaires que soient les ressemblances entre structures chimiques de réaction-diffusion et certains motifs observés dans la nature, de telles similarités ne permettent nullement de conclure à l'identité des processus qui président à leur développement. De plus, étant donné l'extraordinaire complexité du réseau d'interactions propre à tout système biologique, isoler chez un être vivant des auto-activateurs et des inhibiteurs afin d'en vérifier le comportement *in vitro* tient souvent de la gageure.

Tout au plus, dans certains organismes où apparaissent des structures périodiques, on a réussi à identifier des morphogènes (par exemple le calcium au cours du développement des ramifications de l'algue *Acetabularia*) ou des molécules participant au processus d'auto-activation (au cours du développement du membre supérieur de l'embryon de poulet). Des expériences de biologie confirment aussi que, lors de la formation des motifs pigmentés sur la peau des mammifères, il y a bien développement de pré-motifs; toutefois l'origine de ceux-ci reste mal établie, et l'on sait que ce processus met aussi en

œuvre des migrations de cellules et l'adhésion entre cellules.

Quoi qu'il en soit des difficultés inhérentes au milieu biologique, les systèmes biologiques, aussi bien que les systèmes chimiques non linéaires, sont capables d'évoluer d'un état quasi indifférencié vers un nouvel état plus structuré. Cette transition constitue ce que les physiciens appellent une «brisure de symétrie». Ces brisures de symétrie peuvent être décrites par les mêmes équations mathématiques, définissant ainsi des classes d'universalité. Certains systèmes biologiques, tout comme les systèmes chimiques de réaction-diffusion, appartiennent au moins en partie à la même classe d'universalité, et les connaissances que l'on acquiert sur les uns contribuent à élucider les comportements des autres. Si des biologistes de plus en plus nombreux pensent que les processus de réaction-diffusion jouent un rôle au cours de la morphogenèse, il est certain que, outre la réaction-diffusion, d'autres processus physiques contribuent à l'élaboration des formes lors des multiples modifications qui ont lieu au cours du développement embryonnaire. Ce domaine de recherches est en plein essor.

Sur *Archimède*, l'émission franco-allemande de vulgarisation scientifique d'ARTE, P. De KEPPEL présente «Les taches du léopard», le 13 mai 1997 à 20 heures.

P. De KEPPEL est directeur de recherche et É. DULOS est chargée de recherche au Centre de recherche Paul Pascal, laboratoire du CNRS à Bordeaux.

Oscillations and travelling waves in chemical systems, sous la direction R.J. Field et M. Burger, John Wiley & Sons, 1985.

Chemical waves and patterns, sous la direction de R. Kapral et K. Showalter, éditions Kluwer, 1995.

V. CASTETS, E. DULOS, J. BOISSONADE et P. De KEPPEL, *Experimental evidence of a sustained standing Turing-type non equilibrium chemical pattern*, in *Physical Review Letters*, n° 64, pp. 2953-2956, 1990.

P. De KEPPEL, J.J. PERRAUD, B. RUDOVICS et E. DULOS, *Experimental study of stationary Turing patterns...* in *International Journal of Bifurcation and Chaos*, n° 4, pp. 1215-1231, 1995.

INTERNET

Les promesses seront-elles tenues?

42 L'Internet :
du chaos à l'ordre?



44 La recherche
d'information
Clifford Lynch



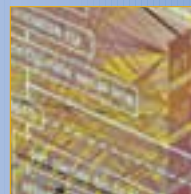
49 Les bibliothèques
numériques
Michael Lesk



52 Le filtrage
d'information
Paul Resnick



55 Le multilinguisme
Bruno Oudet



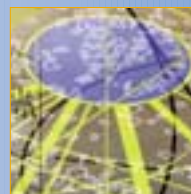
Les interfaces
de recherche
Marti Hearst

57



Les systèmes
sécurisés
Mark Stefik

62



Transferts rapides
Jacques Dupraz

66



Les archives
d'un réseau actif
Brewster Kahle

68

L'Internet : du chaos à l'ordre?

Le réseau Internet est victime de son succès : en quelques années, ce système, initialement confidentiel, de communications entre scientifiques ou ingénieurs est devenu un carrefour inévitable, emprunté par des personnes du monde entier, dans des secteurs aussi variés que l'analyse financière ou la haute couture. Pour remédier à la saturation du réseau, le gouvernement américain a même annoncé son intention de construire un réseau réservé aux scientifiques : l'Internet II.

Comment mettra-t-on le réseau existant au service du grand public? Le courrier électronique, voire les vidéoconférences, sont déjà répandus, mais ces applications sont bien en deçà de ce que le réseau mondial, le World Wide Web, peut offrir : la connexion, par le téléphone ou par des lignes spécialisées, des divers ordinateurs du monde a engendré la constitution d'une gigantesque mine d'informations et d'analyses, la base de données de toutes les bases de données. Les inquiétudes sur l'avenir d'Internet, qui portent habi-



tuellement sur les temps de chargement prohibitifs et sur les limitations d'accès dus à la surcharge du réseau, ne sont que des «douleurs de croissance», qui se résoudront. En revanche, on n'utilisera bien le réseau que si l'on parvient à structurer, à stocker et à classer l'information qui y circule.

Dans les pages qui suivent, des informaticiens expliquent comment l'organisation du savoir rendra l'Internet réellement efficace. À partir de points de vue variés, ils cherchent comment simplifier la recherche de l'information (oui, de nouveaux moteurs de recherche faciliteront l'exploration du réseau). Ils examinent les meilleurs moyens de structurer et d'afficher des données, de sorte que chacun (y compris les aveugles) accède facilement à l'information, de toutes les manières imaginables. Les solutions techniques qu'ils proposent ne sont peut-être pas celles qui seront finalement adoptées, mais leurs idées susciteront certainement des débats.

La structuration d'un milieu fondamentalement désordonné comme le réseau mondial contribuera peut-être à réaliser le rêve des encyclopédistes français du XVIII^e siècle : réunir en un seul lieu toutes les connaissances du monde. Deux siècles après eux, Vannevar Bush, qui dirigea le bureau américain de la recherche scientifique et technique durant la Seconde

Guerre mondiale, proposa le «memex», un poste de travail comprenant un lecteur de microfilms et quantité de films qui auraient été l'équivalent d'une bibliothèque de recherche complète ; le lecteur du memex aurait pu lier entre elles et annoter des informations provenant de films différents. Les idées de Bush influencèrent Ted Nelson, l'inventeur du système de l'hypertexte, qui fut finalement introduit sur le réseau. Le dynamisme intellectuel qui court, des encyclopédistes à T. Nelson, se retrouve dans les articles qui suivent.

Leurs auteurs y esquissent les grandes orientations techniques qui pourraient conduire à la constitution d'un fonds complet du savoir humain. Le réseau Internet deviendrait ainsi un lieu où les rêves d'Homère, de Shakespeare et de Lao-tseu seront à un clic de souris des menus des cantines scolaires ou de l'ordre du jour du prochain conseil municipal. Toutes les activités humaines seront ainsi représentées, des plus nobles aux plus communes.

Retrouvez les articles de ce dossier avec des informations complémentaires et des liens sur notre site Web <http://www.pourlascience.com>

Jeff Brice



La recherche d'informations

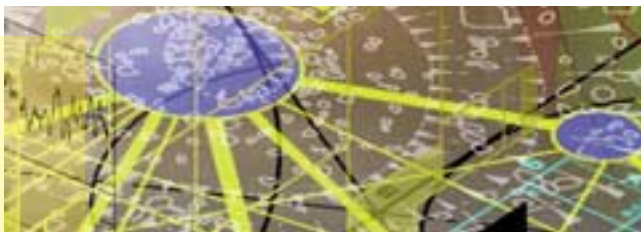
CLIFFORD LYNCH

Les bibliothécaires devront s'allier aux informaticiens pour que cesse l'anarchie qui règne sur l'Internet.

Certains prétendent que le réseau Internet est la bibliothèque mondiale de l'ère numérique. Cette idée ne résiste pas à l'examen. L'Internet, et en particulier l'ensemble de ressources multimédia qui constituent le World Wide Web, ou réseau mondial, n'ont pas été conçus pour servir de support à la publication et à la recherche de l'information. C'est plutôt une bibliothèque chaotique de l'ensemble des publications sortant des «presses» numériques de la planète. Cette mine d'informations contient non seulement des livres et des articles, mais aussi des données scientifiques, des menus, des comptes rendus de conférences, des publicités, des enregistrements audio et vidéo et des transcriptions de conversations interactives. L'anecdotique et l'éphémère sont jetés pêle mêle dans l'important et le durable.

En somme, l'Internet n'a rien d'une bibliothèque numérique. Il ne se développera et ne deviendra un nouveau moyen de communication universel que si l'on met au point des services analogues à ceux des bibliothèques classiques pour organiser l'information, la conserver et la retrouver. Même alors, l'Internet ne ressemblera pas à une bibliothèque classique, car son contenu restera dispersé. Les bibliothécaires, qui choisissent et classent les ouvrages, devront être épaulés par des informaticiens, qui mettront au point des techniques d'indexation et de stockage automatisés de l'information. Seule une synthèse de ces deux professions assurera la viabilité de ce nouveau média.

Aujourd'hui, la structuration de l'information sur l'Internet repose essentiellement sur l'informatique. En théorie, les programmes qui classent et indexent



automatiquement des données numériques peuvent gérer la surabondance d'information présente sur l'Internet. De surcroît, l'indexation automatique, peu onéreuse, est préférée à l'indexation humaine, lente et coûteuse.

Toutefois, ces programmes classent l'information différemment de nous. En un sens, le travail effectué par les divers outils automatiques d'indexation et de catalogage, les moteurs de recherche, est extrêmement démocratique : il donne un accès uniforme, non hiérarchisé, à toutes les données présentes sur le réseau. Dans la pratique, cet égalitarisme électronique a ses avantages et ses inconvénients. Les «surfeurs» du réseau qui cherchent une information se retrouvent souvent submergés par des milliers de réponses. De plus, les résultats d'une recherche renvoient fréquemment à des sites Web (un site Web est un ensemble de pages enregistrées dans la mémoire d'un ordinateur du réseau) sans rapport avec la demande et omettent d'autres sites contenant des informations pertinentes.

Arpenter le réseau

Les spécificités de l'indexation électronique apparaissent quand on examine comment les moteurs de recherche, tels *AltaVista*, de la Société *Digital Equipment*, ou *Lycos*, construisent leurs index et cherchent l'information réclamée par un utilisateur. Ces moteurs envoient périodiquement des programmes nommés «robots collec-

teurs» sur tous les sites qu'ils identifient sur le réseau. Ces robots récupèrent les pages, puis un logiciel en extrait une indexation qui assurera leur description. Selon le moteur, cette dernière procédure va de la simple localisation de

la plupart des mots présents dans les pages à l'analyse complexe, assortie d'une identification des mots et des groupes de mots importants. Ces données sont ensuite stockées dans la base de données du moteur de recherche, avec une adresse, une URL (*Uniform Resource Locator*, soit «localisateur unifié de ressources»), qui localise le document. Pour soumettre ses demandes à la base de données du moteur de recherche, le demandeur utilise un navigateur – tel le programme *Netscape* – qui lui donne une liste de ressources, les URL, sur lesquels il peut cliquer pour se connecter aux sites repérés par le moteur de recherche.

Aujourd'hui, les moteurs de recherche traitent quotidiennement des millions de demandes d'information. Cependant, ces moteurs sont loin d'être des outils idéaux pour s'y retrouver dans la masse d'informations, sans cesse croissante, disponible sur l'Internet. Contrairement aux indexeurs humains, les moteurs de recherche cernent mal les caractéristiques globales d'un document, telles que le sujet ou le genre – poème, pièce de théâtre, voire publicité.

En outre, le réseau mondial ne dispose pas encore de normes qui faciliteraient l'indexation automatique : les documents qu'il contient ne sont pas structurés, de sorte que les programmes ne peuvent en extraire l'information qu'un indexeur humain trouverait par un simple examen superficiel, comme l'auteur, la date de publication, la longueur et le sujet d'un texte (ces infor-