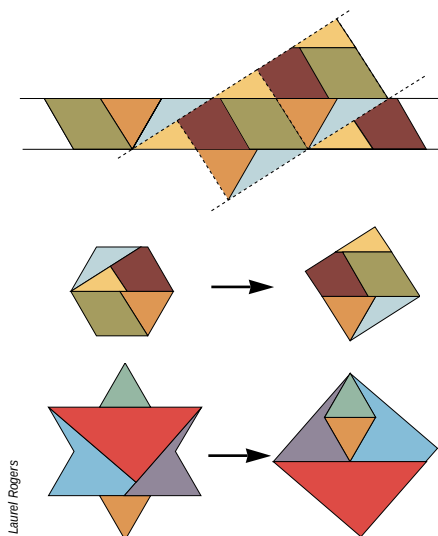




4. Des bandes infinies pavées par des formes différentes nous inspirent de possibles dissections ; en haut, un hexagone en un carré. Une étoile de David est, similairement, dissectée en carré (en bas).

nuages de poussières sphériques infiniment compliqués qu'à des formes juxtaposées.

Il n'y a, faut-il le dire, aucune possibilité pratique de réaliser cette dissection sur une sphère réelle – les trafiquants en métaux précieux peuvent soupirer de soulagement –, mais la dissection démontre combien le concept de volume est subtil. Tarski a démontré en 1935 qu'il n'existe pas de dissections « paradoxales » modifiant les aires des surfaces, quelle que soit la complexité des parties découpées – mais il en existe, cependant, à la surface d'une sphère.

Lorsque les parties issues du découpage de l'objet ont des aires, ou des volumes, bien définis, il n'y a pas de construction contraire à l'intuition. En 1833, un lieutenant de l'armée prussienne, P. Gerwein, résolut le problème essentiel des dissections posé par le mathématicien hongrois Farkas Wolfgang Bolyai. Gerwein prouva que, étant donné deux polygones plans d'aires

égales, on peut découper l'un en un nombre fini de pièces polygonales qui peuvent être assemblées pour former l'autre. Ce résultat est dénommé théorème de Bolyai-Gerwein, bien qu'un certain William Wallace l'ait démontré auparavant, en 1807.

Le théorème de Bolyai-Gerwein ne se généralise pas à la dimension trois. En 1900, le mathématicien allemand Hilbert demanda si deux polyèdres d'égal volume étaient toujours « équivalents par dissection », c'est-à-dire pouvaient être la réunion du même ensemble de parties polyédrales. Une année après, le mathématicien Max Dehn, spécialiste de topologie, prouva un surprenant théorème : un cube et un tétraèdre régulier ne sont pas équivalents par dissection polyédrale, répondant par la négative à Hilbert.

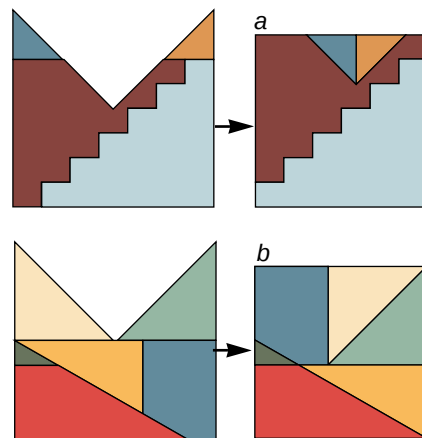
PAVAGES, BANDES ET DISSECTIONS

Des techniques mathématiques apparemment sans relation avec les dissections vous aident dans les recherches. Vous pouvez progresser dans ce domaine à l'aide de tâtonnements inspirés, mais seulement si vous avez une excellente imagination spatiale. Une des vertus de *Dissections* est qu'il explique nombre de techniques générales utilisées dans leurs recherches.

Une de ces techniques consiste à couper une forme suivant un « escalier », qui peut alors être décalé d'une marche pour créer une forme différente. David Colli-son, un fervent des dissections, conçut des dissections élaborées fondées sur ce principe. Sa dissection du pentagone (voir la figure 5) démontre l'égalité pythagoricienne : $5^2 + 12^2 = 13^2$.

Une autre méthode générale est celle de la mosaïque. Nombre de formes intéressantes peuvent être découpées et réassemblées en formes qui pavent le plan. Si deux mosaïques distinctes, chacune formée de pavés de même aire, sont superposées, il est souvent possible d'y « lire en clair » une dissection d'une des formes dans l'autre. Par exemple, une croix grecque peut de la sorte être « dissectée » en carré.

Un usage plus élaboré de la même idée de base – la dissection d'un dodécagone en une croix latine – est dû à



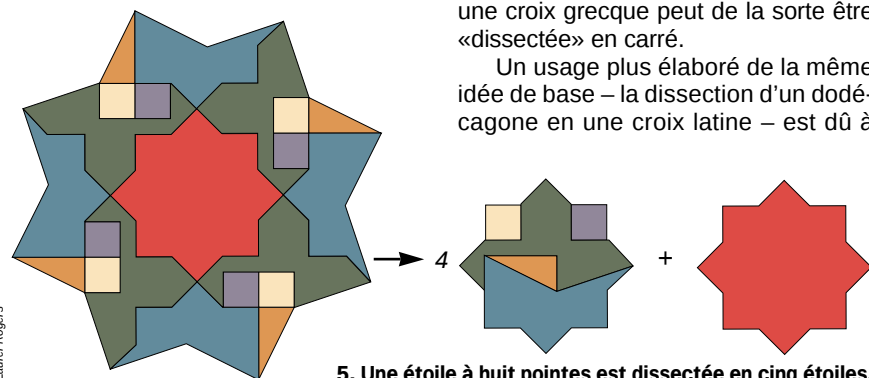
6. Dissection trompeuse : Sam Loyd pensait pouvoir disséquer une mitre en un carré. Malheureusement le carré (a) était un rectangle ! La bonne dissection est représentée en (b).

Harry Lindgren, l'auteur de *Geometric Dissections* (Van Nostrand, 1964). Le premier pas, et le plus difficile, est le découpage du dodécagone et son réarrangement en une forme pavant le plan. La croix latine (de même aire que le dodécagone) peut aussi se réarranger pour paver le plan. La superposition des deux pavages conduit à la dissection désirée.

Une troisième technique est la méthode des bandes. Chaque forme est découpée et réarrangée en pavé qui reproduit une bande infinie. Un recouvrement adéquat des bandes peut alors déterminer une dissection. Celle d'un hexagone en carrée, obtenue par Paul Busschop, provient de cette technique des bandes (Busschop était belge ; solitaire, il écrivit un livre dont son frère assura la publication posthume en 1879). Harry Bradley, un ingénieur américain, découvrit en 1897, alors qu'il était assistant à l'Université, comment dissectionner une étoile de David en carré par la même méthode.

Les problèmes de dissection ne se limitent pas à changer une seule forme en une autre. Souvent tout un ensemble de formes doivent être découpées et réassemblées, ou vice versa.

Les dissections vous hypnotisent, et, à l'occasion, l'œil peut vous induire en erreur. En 1901, Loyd commit une bourde mémorable : il affirma pouvoir dissectionner un carré privé d'un de ses quarts, une mitre, en un carré. Malheureusement, le « carré » apparent était, en définitive, un rectangle aux côtés proportionnels à 48 et 49. Son grand rival, Dudeney, fut ravi de signaler l'erreur en 1911, et indiqua une dissection correcte, illustrée sur la figure 6. Aussi, si vous voulez découvrir de nouvelles dissections, suivez le conseil prodigué par nombre de parents à leur progéniture : « Amusez-vous bien, mais faites attention. »



5. Une étoile à huit pointes est dissectée en cinq étoiles.



Le monde des machines

JEAN-PAUL DELAHAYE

Envisageons sérieusement l'hypothèse selon laquelle nous sommes des ordinateurs.

Nous sommes en l'année 2098, Monsieur J. Martin souffre d'une grave maladie qui atteint progressivement chacun de ses organes. Il possède déjà un rein, un poumon, un cœur, des jambes et des bras artificiels, une peau de remplacement... seul son cerveau reste biologique. À chaque opération, son médecin l'a assuré que l'organe substitué était fonctionnellement identique à celui enlevé, et J. Martin, lorsqu'il s'est réveillé après les greffes, n'a jamais rien perçu d'anormal.

Aujourd'hui, les infos sur les neurones indiquent que son cerveau est touché, mais son médecin-informaticien connaît les techniques pour mesurer l'état instantané du cerveau et pour «reprogrammer» son équivalent, neurone par neurone, dans un ordinateur, lequel se loge facilement dans la boîte crânienne en plastique de Monsieur Martin.

Notre médecin décide d'opérer la substitution, et un nouveau cerveau-ordinateur convenablement programmé est placé dans le corps de Monsieur Martin. Celui-ci se réveille deux heures plus tard et, comme lors des précédentes greffes, ne ressent aucun changement majeur. Le lendemain, il rentre chez lui où son épouse et ses enfants l'accueillent avec chaleur et soulagement.

Tout serait parfait si, à l'hôpital, le cerveau de Monsieur Martin n'avait été conservé. Après observation, l'organe se révèle totalement sain, contrairement aux analyses erronées d'un laborantin inexpérimenté. Le même médecin remplace le cerveau dans un corps artificiel en tout point identique à celui de Monsieur Martin (les caractéristiques précises de son corps étaient décrites dans son dossier médical). Monsieur Martin retourne à son domicile. Il est, bien sûr, quelque peu surpris d'y découvrir son double. Mais qui est le double de l'autre ? Cette fable montre que la croyance que nous sommes des machines, croyance baptisée *mécanisme*, conduit à des situations inattendues et choquantes.

Reste que la question de la philosophie mécaniste est bel et bien posée et qu'à défaut de prévoir ce que l'avenir de la médecine autorisera, on peut quand même s'interroger sur sa validité.

Nous voudrions vous faire découvrir une théorie proposée à titre d'hypothèse et explorée avec le plus grand soin depuis une dizaine d'années par Bruno Marchal, qui travaille à l'Institut de recherche interdisciplinaire et de développement en intelligence artificielle de Bruxelles.

Les difficultés que rencontre B. Marchal à faire reconnaître son travail montrent qu'il est plus aisé de faire accepter une recherche concernant un sujet précis, même inintéressant (comme «L'influence de la direction du vent sur les pratiques sexuelles des fourmis unijambistes de l'île de Pâques»), que d'explorer des hypothèses que certains jugent déplaisantes pour des raisons extra-scientifiques. Le travail de B. Marchal explore les conséquences de l'hypothèse mécaniste, sans préjuger toutefois de sa validité, et son travail se fonde sur la théorie de la calculabilité, la logique mathématique, et sur certains formalismes de la mécanique quantique, domaines où il fait preuve de minutie et d'érudition. Le monde universitaire est cloisonné en disciplines qui s'ignorent, quand elles ne se méprisent ou ne se combattent pas. Nulle place ne semble réservée aux travaux fondés sur une culture multidisciplinaire impliquant de délicats mélanges de mathématiques, de physique et de philosophie.

B. Marchal étudie le *mécanisme numérique*, théorie selon laquelle, si l'on décrit un être humain avec une précision suffisamment grande, on saisit complètement son identité mentale. À ce niveau de précision, le fonctionnement de l'esprit d'un individu est mécaniquement décrit. Autrement dit, B. Marchal s'interroge sur ce qu'on peut déduire de l'idée que «je ne suis peut-être après tout qu'une machine». Il précise son hypothèse en adoptant le principe que les

machines envisagées, que nous appellerons machines numériques, sont du type de nos ordinateurs actuels : tous les calculs s'y ramènent à des manipulations réglées de «0» et de «1».

Le premier résultat que B. Marchal fait apparaître – et cela seul est déjà très intéressant – est l'indéterminisme que son hypothèse entraîne. Comme nous allons le voir, cette nouvelle forme d'indéterminisme est différente de celles rencontrées jusqu'à présent en physique.

UN INDÉTERMINISME RÉCALCITRANT

Procédons à quelques expériences de pensée, comme on le fait fréquemment en physique et comme Albert Einstein l'a fait avec délectation et génie pour élaborer et justifier sa théorie de la relativité générale.

Si l'hypothèse du mécanisme numérique est vraie, on imagine que l'on saura construire un jour des machines qui analyseront un corps humain tout entier, en enverront la description codée à une machine de reconstitution, par exemple sur la Lune, machine qui alors le recréera ou recréera son équivalent mécanique. La science-fiction a depuis longtemps envisagé ce moyen de transport à la vitesse de la lumière, qu'elle désigne par *téléportation*. Celui qui monte dans une machine de téléportation a le sentiment d'un simple déplacement. De son point de vue, avant que la machine soit mise en marche et sauf en cas de pannes, qu'on espère rares, il est certain de se retrouver sur la Lune.

Imaginons maintenant que le signal avec la description de votre corps soit envoyé à la fois sur la Lune et sur Mars, où deux machines réceptrices différentes vous reconstituent chacune de leur côté. Avant de monter dans la machine, vous savez qu'elle va vous envoyer dans deux directions ; si vous vous posez la question : «Vais-je me retrouver sur la Lune ou sur Mars ?», vous n'avez aucune rai-

son de donner plus d'importance à la machine sur la Lune qu'à celle sur Mars et vous répondez : «Il y a 50 pour cent de chances que j'aille sur Mars, et 50 pour cent que j'aille sur la Lune.»

En recommençant l'expérience un très grand nombre de fois, et en interrogeant les gens qui sortent des machines sur Mars et sur la Lune, on trouverait que la moitié exactement ont fait le voyage Terre-Lune et l'autre moitié Terre-Mars. La prédiction «Je vais arriver sur la Lune» est donc juste une fois sur deux. Nous avons affaire à une indétermination absolue. Même si nous croyons que toutes les lois de la physique sont parfaitement déterministes, cet indéterminisme persiste. C'est bien sûr un indéterminisme psy-

chologique – intime, pourrions-nous dire – qui n'en est pas moins total. Contrairement à l'indéterminisme quantique, il ne peut être évité par aucune théorie spéciale comme celles «à variables cachées et non locales» qu'on a proposées pour rétablir le déterminisme physique.

IMPOSSIBLES PROBABILITÉS

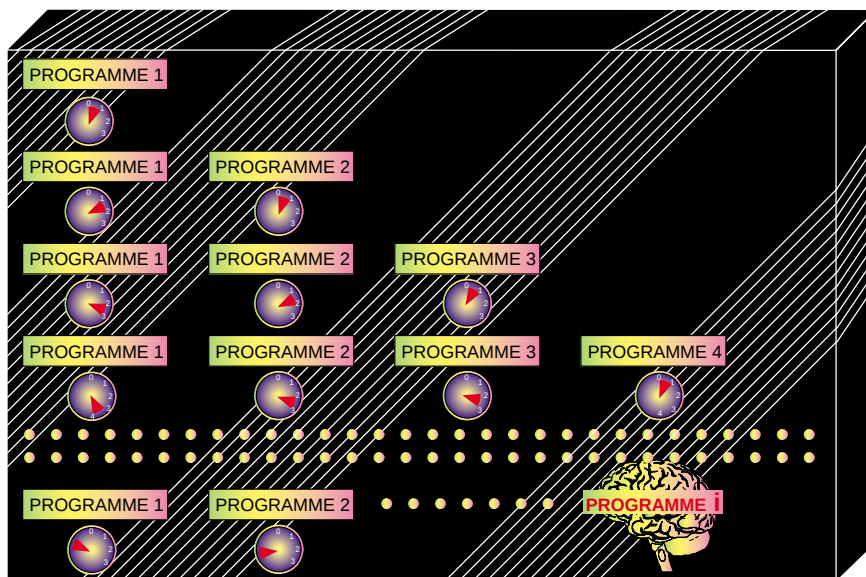
Tout serait simple si l'on ne pouvait envisager qu'une troisième machine ayant capté secrètement le message radio ne reconstitue aussi le voyageur, maintenant sur Saturne. La probabilité est-elle donc de 33,33 pour cent pour chaque destination plutôt que 50 pour cent?

Allons encore un peu plus loin et imaginons que le message codant le voyageur a été capté par une autre machine, qui conserve en mémoire le message reçu et qui, toutes les heures régulièrement et sans fin, produit une nouvelle copie du voyageur.

Quelle est dans ce cas la probabilité pour le voyageur (s'interrogeant avant le départ) d'être la version numéro 1 432 ou la version numéro 323 846 264 338 ? Il est impossible de répondre, car une infinité de versions du voyageur vont être créées. Aucune théorie des probabilités ne semble s'adapter à un tel contexte. Il y a bien indéterminisme, mais il est tellement grave qu'on ne peut même plus parler de probabilité. Troublant, non ?



1. Si nous sommes des machines, alors il sera possible de réaliser des appareils de téléportation : après avoir enregistré (soigneusement !) sa description, on annihile la personne qui doit voyager. On fait voyager la description par onde électromagnétique. On reconstitue la personne à l'arrivée. On peut aussi la reconstituer deux fois (par exemple, sur la Lune et sur Mars). Celui qui entre dans la machine vit une expérience où ce qui va se passer est pour lui absolument indéterminable à l'avance : il n'a aucun moyen de savoir où il va arriver. Cet indéterminisme «intime», sans lien avec l'indéterminisme physique (quantique ou autre) est troublant. L'illustration est tirée de la bande dessinée *Clarke et Kubrick T.1 agents très spatiaux* d'Alfonso Font, Éditions Glénat.



2. Le *dépoyeur universel* est un programme qui, une fois lancé, ne s'arrête jamais et fait fonctionner tous les programmes envisageables (en théorie de la calculabilité, on démontre que c'est possible). Il prend la liste de tous les programmes, *Programme 1*, *Programme 2*, *Programme 3*, ... et les fait fonctionner à tour de rôle selon le schéma indiqué. Si votre esprit est équivalent à un programme, et si l'on réalise concrètement un *dépoyeur universel* et qu'on le fait fonctionner indéfiniment, alors il finit par reproduire votre fonctionnement mental : un *dépoyeur universel* simule tous les esprits possibles.

LE PROGRAMME DE SIMULATION DE TOUT

La chose peut être rendue encore plus inquiétante en imaginant un «programme de simulation de tout», que B. Marchal dénomme un *dépoyeur universel*.

Un tel *dépoyeur universel* fonctionne de la manière suivante : il écrit dans une zone de mémoire le plus simple des programmes possibles (un langage de programmation a été fixé, et l'on mesure la simplicité des programmes par leur longueur ; si plusieurs ont la même longueur, on les classe par ordre alphabétique). Le *dépoyeur universel* le fait fonctionner durant une seconde. Il écrit ensuite le second programme dans l'ordre de simplicité et fait fonctionner le premier programme et le second une seconde chacun. Le *dépoyeur universel* écrit ensuite le troisième programme et fait fonctionner le premier, le second et le troisième une seconde chacun, etc. (Le schéma d'appel des programmes est donc 1 12 123 1234 12345, etc.)

Un tel méta-programme engendrant et faisant fonctionner tous les programmes est possible, car l'ensemble des programmes peut être ainsi énuméré et qu'il n'y a aucun obstacle à les faire fonctionner par tranches de une seconde, selon le schéma proposé.

La mise au point de tels programmes ne présente pas de difficultés particulières : B. Marchal en a écrit. Bien sûr, les *dépoyeurs universels* n'ont guère de fonctions pratiques, car on est obligé

de les arrêter avant que quelque chose d'intéressant ne se produise.

Cependant, si nous ne sommes que des machines numériques (c'est l'hypothèse que nous explorons), alors dans le cours d'un fonctionnement ininterrompu d'un *dépoyeur universel* vient nécessairement un moment où le programme envisagé qui fonctionne pendant une seconde correspond à celui de mon cerveau à l'instant présent. Pendant une seconde, tout se passera donc dans l'ordinateur qui exécute le *dépoyeur universel* comme dans mon cerveau maintenant : ma conscience de l'instant aura été simulée. Mon esprit, pendant un court intervalle de temps, existera dans la machine.

Remarquons qu'il n'est pas nécessaire de supposer que le programme simule l'environnement extérieur, car, pour que j'aie une sensation identique à celle que j'ai actuellement, il suffit que les influx nerveux provenant de mes divers organes sensoriels soient correctement simulés.

Un peu plus tard, le même programme sera de nouveau exécuté une seconde de plus, et la simulation de mon esprit se poursuivra donc. Ce que je vis en ce moment sera en fait parfaitement reproduit par tranches d'une seconde, sans que bien sûr j'aie conscience des interruptions.

Toujours sous réserve que le *mécanisme numérique* soit vrai, le *dépoyeur universel*, en simulant tous les programmes, simulera en réalité tous les esprits possibles. Je suis donc dedans,

comme vous l'êtes, comme toute conscience passée, présente ou future est dedans. En fait, chacune y est une infinité de fois.

Cela rappelle l'histoire de *La bibliothèque de Babel* de Borgès, qui contient tous les livres possibles (voir *Pour La Science*, juillet 1996), mais ici ce n'est pas la description extérieure de ce que nous vivons qui est reproduite, c'est l'esprit et le sentiment intime que nous avons de nous-même : la conscience de soi.

La perspective de tout cela donne le vertige, et il y a quelque difficulté à y croire sérieusement. L'hypothèse du mécanisme numérique, même si elle est loin d'être prouvée, ne paraît pas aussi absurde que les conséquences que nous venons d'en tirer. Aurions-nous adopté une hypothèse non mentionnée en plus de celle du *mécanisme numérique*?

Oui, et cette hypothèse que nous allons expliquer s'appelle le *fonctionnalisme*. Il a donné lieu à de nombreuses discussions en philosophie de l'esprit depuis une vingtaine d'années.

LE FONCTIONNALISME

Nous avons supposé dans la présentation du *dépoyeur universel* que le fonctionnement d'un esprit, l'essence de ce qu'est une personnalité ou une conscience, est indépendant de la base matérielle qui en constitue le support.

Lorsque l'on change un bras de Monsieur Machin, le fait que ce soit un bras en plastique qui se substitue au bras de chair n'a pas d'importance, car on suppose que la science des greffes a tellement progressé que la fonction du bras est préservée dans le moindre détail. De même, quand il s'agit d'un ordinateur, si je remplace un composant par un autre fonctionnellement équivalent, tout sera identique, et les mêmes programmes fonctionneront à l'identique.

De même encore, si je dispose d'un programme qui calcule les nombres premiers, je peux le faire tourner sur toutes sortes de machines ; à chaque fois, le détail de ce qui se passe est assez différent, mais, à un niveau d'abstraction suffisant, le calcul est le même et, en définitive, ce qui est important lors du fonctionnement de mon programme ne dépend pas de la machine (le langage Java, si à la mode aujourd'hui, a justement cette propriété : les programmes écrits avec lui peuvent fonctionner dans une grande variété d'environnements différents, PC, Macintosh ou UNIX).

L'hypothèse fonctionnaliste qu'on a utilisée sans le dire à propos du *dépoyeur universel* est l'idée que, pour l'esprit, il en est de même : ce qui est important pour

caractériser mon esprit et son fonctionnement, ce ne sont pas les neurones biologiques particuliers qui se trouvent dans ma boîte crânienne, mais l'organisation abstraite de leur réseau et des flux électriques et chimiques qui les parcourent.

Le *fonctionnalisme* en philosophie de l'esprit est considéré comme acceptable par nombre de gens : ce serait revenir au vitalisme des siècles antérieurs que de soutenir que la pensée et l'esprit présents dans nos cerveaux le sont à cause du type particulier de chimie ou à cause de propriétés biologiques uniques que seules possèdent les cellules humaines.

Il est beaucoup plus naturel de croire que l'esprit, la pensée, la conscience sont le résultat abstrait de processus abstraits réalisés d'une certaine façon dans les cerveaux des humains, mais qui pourraient être réalisés matériellement d'autres façons, comme le programme de calcul des nombres premiers peut exister et fonctionner de mille façons différentes dans mille machines différentes.

Lorsque nous avons envisagé le *dépoyeur universel* et que nous avons dit que chacun de nous était dedans, nous avons, sans le dire, utilisé l'hypothèse *fonctionnaliste*. Maintenant que nous en avons pris conscience, la conclusion de l'expérience de pensée du *dépoyeur universel* s'exprime ainsi :

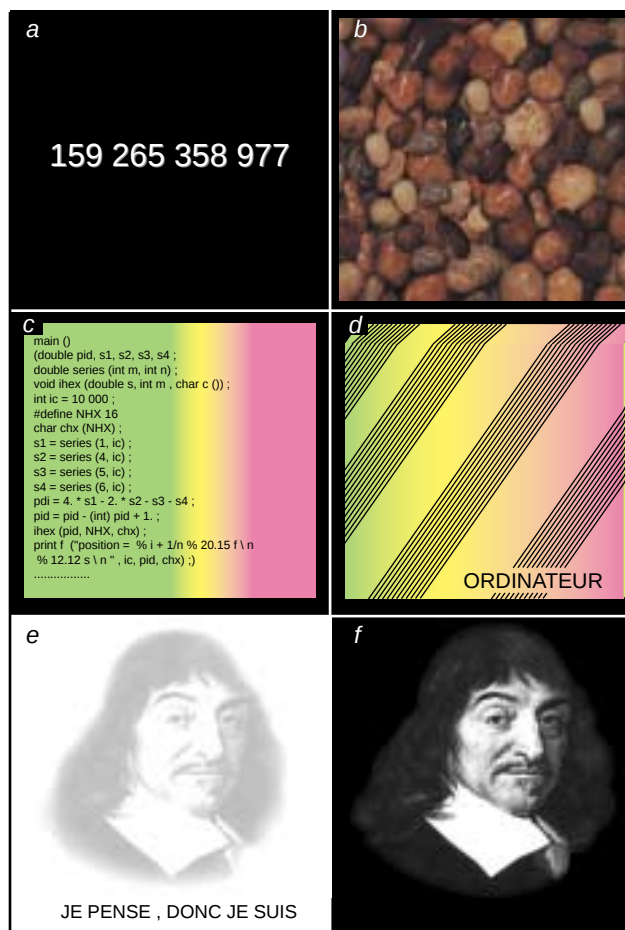
– si nous sommes des *machines numériques* (hypothèse 1) et,

– si le *fonctionnalisme* est exact (hypothèse 2), alors le *dépoyeur universel* fait fonctionner tous les esprits possibles ayant existé, existants ou qui existeront.

LA SUPERVÉNIENCE PHYSIQUE

Pour franchir l'étape suivante, nous devons nous interroger sur l'importance des réalisations matérielles nécessaires à l'existence d'un esprit.

Il semble naturel de soutenir que, si aucune réalisation



3. LA SUPERVÉNIENCE PHYSIQUE. Un objet mathématique (ainsi l'entier naturel 159 265 358 979) peut être physiquement réalisé : on l'écrit (a) ou on rassemble un nombre de cailloux correspondant (b). Avant que cet entier n'ait été conçu, il existait déjà dans le monde abstrait des objets mathématiques. Il en va de même des programmes informatiques (c) : avant d'avoir été pensé, un programme existe comme un nombre entier. Donc, si je suis un programme (hypothèse mécaniste), mon existence n'exige pas vraiment qu'un monde physique la supporte (e). Cette idée qui dispense d'appuyer la conscience sur la matière est refusée par les philosophes qui soutiennent la thèse de la supervénience physique : pour qu'une conscience existe, il faut qu'un support physique lui soit associé. Cette thèse exige que nous adoptions vis-à-vis des programmes une attitude différente de celle que nous avons vis-à-vis des entiers.

LE CORPS ET L'ESPRIT	
DUALISME (deux types d'entités fondamentales)	MONISME (un seul type d'entité fondamentale)
Le principe du rasoir d'Occam (il faut être le plus économe possible dans les entités que l'on postule) et la difficulté à concevoir une théorie de l'interaction entre la matière et l'esprit font pencher vers le monisme.	
La double évidence que le monde physique existe et que j'ai un esprit relativement autonome fait pencher vers le dualisme.	
MONISME MATÉRIALISTE (tout résulte du monde matériel)	MONISME IDÉALISTE (tout provient de l'esprit)
Le fait que la matière a préexisté à l'esprit, en fait d'apparition récente, fait pencher vers le monisme matérialiste.	
Les difficultés de la théorie de la supervénience physique font pencher vers le monisme idéaliste.	

4. Les oppositions de base entre le corps et l'esprit.

matérielle ne résulte d'un processus de calcul, non seulement celui-ci ne sert à rien, *stricto sensu*, mais n'existe pas. Cette idée en termes techniques est que l'esprit «supervient à la matière» (qui lui est nécessaire). Autrement dit encore, l'esprit et la conscience naissent de la matière et ne peuvent se passer de ce support.

B. Marchal explore cette hypothèse avec soin et, au terme d'une longue analyse menée avec minutie, il retrouve la conclusion (aussi énoncée par le philosophe T. Maudlin) qu'à partir du moment où l'on accepte le *fonctionnalisme* il faut abandonner la *supervénience physique*. Autrement dit, il n'est pas raisonnable de n'accorder d'existence qu'aux pensées dont une réalisation matérielle existe quelque part dans l'univers physique.

On comprend directement la position de B. Marchal en revenant au cas du programme de calcul des nombres premiers : pour qu'il existe, il n'est pas nécessaire d'admettre que je l'ai vraiment écrit. Il existe en tant que structure mathématique, même si personne n'a jamais vu ce programme, ne l'a jamais écrit ni même jamais envisagé. La conclusion de B. Marchal n'est que la généralisation de cette idée à l'esprit et à la conscience. Une telle conclusion repose sur une option particulière en philosophie des mathématiques appelée le *réalisme* : les objets mathématiques, en particulier les nombres et les programmes (qu'on peut, grâce à des méthodes introduites par Kurt Gödel, ramener à des nombres) existent indépendamment de nous.

Le *réalisme*, quand il ne porte que sur des nombres et des programmes, est difficile à récuser (quoique ce soit l'option retenue par les philosophes *intuitionnistes*), donc la nouvelle hypothèse que propose B. Marchal est plausible. Ainsi, à partir d'hypothèses parfaitement identifiées, B. Marchal défend l'idée qu'un esprit existe en tant qu'être arithmétique abstrait, même si aucune réalisation du programme correspondant n'a été physiquement réalisée ni mise en fonctionnement.

Toutefois, si la conscience que nous avons de nous-même n'a plus besoin de support maté-

1. MÉCANISME NUMÉRIQUE
(il existe un niveau de détail où nous sommes des machines numériques, c'est-à-dire des programmes)

2. FONCTIONNALISME
(en changeant des composants d'une machine par d'autres fonctionnellement équivalents, sa nature n'est pas modifiée ; donc mon esprit n'est pas vraiment dépendant de sa réalisation).

3. RÉALISME ARITHMÉTIQUE
(entiers et programmes existent indépendamment du monde physique)

→ La solution du problème du corps et de l'esprit est l'idéalisme, qui sera vérifié par notre réussite à déduire la physique de la théorie de la calculabilité.

5. Si l'on admet les trois prémisses de Bruno Marchal, on obtient une nouvelle théorie philosophique et une conception révolutionnaire de la physique. Si l'on pense que la conclusion est invraisemblable,

alors il faut renoncer à l'une des hypothèses au moins. Nombreux sont ceux qui renoncent au mécanisme, mais peut-être que le réalisme arithmétique est l'hypothèse qui doit être abandonnée.

riel, la physique n'est-elle pas une illusion? Dit autrement, la physique pourrait-elle se déduire de l'existence des machines numériques? La physique peut-elle être élaborée purement logiquement à partir de l'étude de la théorie de la calculabilité?

DUALISME, MATÉRIALISME ET IDÉALISME

B. Marchal défend que cette idée doit être examinée et se fixe le projet de retrouver les lois de la physique par l'étude de la théorie de la calculabilité. Il a déjà obtenu quelques résultats. Avant de les évoquer, je vais situer la conception globale qu'il envisage en la comparant aux théories plus classiques portant sur ce qu'on appelle le problème du corps et de l'esprit.

Très schématiquement, on distingue le *dualisme*, le *matérialisme*, et l'*idéisme*. Le *dualisme* défend que l'esprit et la matière existent et entretiennent des relations l'un avec l'autre. La communication entre les deux mondes est une énigme délicate qui a conduit Descartes (qui soutenait un point de vue dualiste) à proposer qu'elle se faisait grâce à la glande pinéale. Des versions modernes du dualisme ont récemment été défendues par le prix Nobel de médecine 1963 John Eccles, qui propose une théorie fondée sur la possibilité de certains «mécanismes quantiques dans les vésicules du réseau présynaptique des cellules pyramidales du cortex».

Pour des raisons de simplicité et pour éviter le délicat problème de l'interaction entre la matière et l'esprit, le *matérialisme* propose de se débarrasser de l'esprit en décrétant que seule existe vraiment la matière. Dans les versions modernes, le mot *matière* est remplacé par les concepts de particules ou de champs. L'esprit, pour un matérialiste, n'est qu'une illusion ou se ramène à son support physique. Cette conception est celle majoritairement défendue aujourd'hui, car elle semble plus économique en nombre d'entités de base que le dualisme et donc satisfait mieux au principe

du rasoir d'Occam, qui veut qu'en cas de doute, on doive toujours donner la préférence à l'option la plus simple.

Une autre conception satisfait au principe du rasoir d'Occam l'*idéisme* : seul l'esprit existe, la matière n'étant qu'une illusion. À partir du mécanisme, B. Marchal tombe sur un *idéisme*. Son *idéisme* est toutefois très spécial, puisqu'il admet comme élément de base, non pas une substance spirituelle vague et toute-puissante, mais l'univers des *machines numériques abstraites*, qui ne sont qu'un type d'objets mathématiques. Pour cet idéisme, ce qui existe avant tout, c'est une catégorie complexe d'objets mathématiques, et tout se construit à partir de là.

Bien évidemment, la solution de B. Marchal est inattendue, choquante, et ceux qui ont fait l'effort de l'écouter ou de le lire ont rarement réussi à le suivre, tant les deux conceptions dominantes d'aujourd'hui que sont le *dualisme* et le *matérialisme* occupent le terrain et masquent les autres alternatives.

Heureusement sa conception est susceptible d'être testée ou attestée (je n'ose dire prouvée!) : si l'on réussit à déduire la physique de la théorie des machines numériques, la conception de B. Marchal devra certainement être prise au sérieux, quel que soit l'étonnement qu'elle provoque au premier abord. Un premier pas a peut-être été franchi dans cette direction. Pour B. Marchal, la théorie de la calculabilité doit servir de base à une reconstruction de la physique. C'est une théorie non triviale (elle se construit sur l'arithmétique, qui est loin de l'être), dont certaines parties développées pour approfondir la compréhension des théorèmes d'incomplétude de Gödel explicitent ce qu'une machine peut connaître d'elle-même sans se tromper, c'est-à-dire en restant non contradictoire. Cette *logique de la prouvabilité* développée pour comprendre mieux le sens des résultats de Gödel peut être vue comme une *logique des êtres conscients*. B. Marchal en a élaboré une version nouvelle qui semble s'adapter au point de vue qu'il

défend et qui constitue l'apport technique et mathématique de son travail, indépendamment de son interprétation philosophique.

Un point remarquable est que la logique qu'il a développée ressemble à une autre logique qui avait été proposée par G. Birkhoff et J. von Neumann pour rendre compte des observables de la mécanique quantique.

Cette coïncidence doit-elle être interprétée comme le début de la réalisation du programme de reconstruction de la physique par la théorie de la calculabilité? Il est trop tôt pour le dire. Cependant, si c'est le cas, alors cela signifie que la conception philosophique générale qu'envisage B. Marchal doit être examinée avec attention. En explorant avec détermination les conséquences de l'hypothèse du *mécanisme numérique*, B. Marchal a ouvert un champ de recherche nouveau qui, si l'on réussit à y progresser, remettra en cause nos conceptions philosophiques du corps et de l'esprit.

Jean-Paul DELAHAYE est directeur adjoint du Laboratoire d'informatique fondamentale de Lille du CNRS.

e-mail : delahaye@lilfl.fr

B. MARCHAL, *Informatique théorique et philosophie de l'esprit*, Actes du Colloque de l'Association sur la recherche en cognisciences, Toulouse, 1988.

B. MARCHAL, *Mechanism and Personal Identity*, in *Proceedings of WOCFAI, World Conference on the Fundamental of Artificial Intelligence*, sous la direction de M. De Glas et D. Gabbay, Paris, 1991.

B. MARCHAL, *Des fondements théoriques pour l'intelligence artificielle et la philosophie de l'esprit*, *Revue internationale de philosophie*, 172, 1990.

B. MARCHAL, *Calcul, physique et cognition*, Rapport de recherche, LIFL, 1997.

T. MAUDLIN, *Computation and Consciousness*, in *The Journal of Philosophy*, pp. 407-432, 1989.

Écrans à cristaux liquides

Des écrans à cristaux liquides de bonne qualité équipent aujourd'hui tous les ordinateurs portables. Les cristaux liquides sont utilisés comme de petits interrupteurs qui allument ou éteignent des unités élémentaires de l'image nommées pixels. Ces matériaux possèdent deux propriétés utiles à cette fin : ils sont transparents, mais ils peuvent modifier la polarisation de la lumière (l'orientation du champ électrique de l'onde) qui les traverse ; en outre, l'alignement des molécules de cristaux liquides et leurs propriétés polarisantes sont modifiables par application d'un champ électrique.

Un écran est formé de plusieurs couches, emprisonnées entre deux plaques de verre : un film mince, transparent, qui porte de minuscules électrodes, est en contact avec une couche de cristaux liquides. La taille des électrodes définit celle des pixels. Les faces externes des plaques de verre portent elles-mêmes des filtres polarisants, orientés à 90 degrés l'un par rapport à l'autre. Les écrans couleur contiennent en outre des filtres colorés dans les trois couleurs fondamentales (rouge, vert et bleu), découpés en sous-pixels : la combinaison de trois sous-pixels forme un pixel.

La lumière émise à l'arrière de l'écran, par fluorescence, est filtrée une première fois : seules les ondes dont le plan de polarisation est vertical pénètrent dans l'écran (a). La lumière qui traverse un sous-pixel «éteint» reste polarisée verticalement (b) : elle est absorbée par le filtre polariseur de la face avant, et le pixel reste sombre sur l'écran. En revanche, la polarisation de la lumière qui traverse un sous-pixel «allumé», où une orientation est imposée aux molécules de cristaux liquides, est modifiée : cette lumière traverse la face avant de l'écran (c).

L'allumage des sous-pixels du réseau est réalisé ligne par ligne. En même temps qu'un potentiel est appliqué à une ligne, un logiciel détermine quels potentiels doivent être appliqués aux colonnes où les sous-pixels doivent être allumés. À l'intersection d'une ligne et d'une colonne activées, un champ électrique qui commande l'orientation des molécules de cristaux liquides est créé par l'intermédiaire d'un transistor. L'activation successive des 1 280 lignes d'un écran requiert entre 16 et 33 millisecondes. On améliorera la qualité des images en augmentant la densité des pixels et en accélérant l'activation des lignes.

Samuel MUSA, Université du Michigan

