

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@lifl.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.