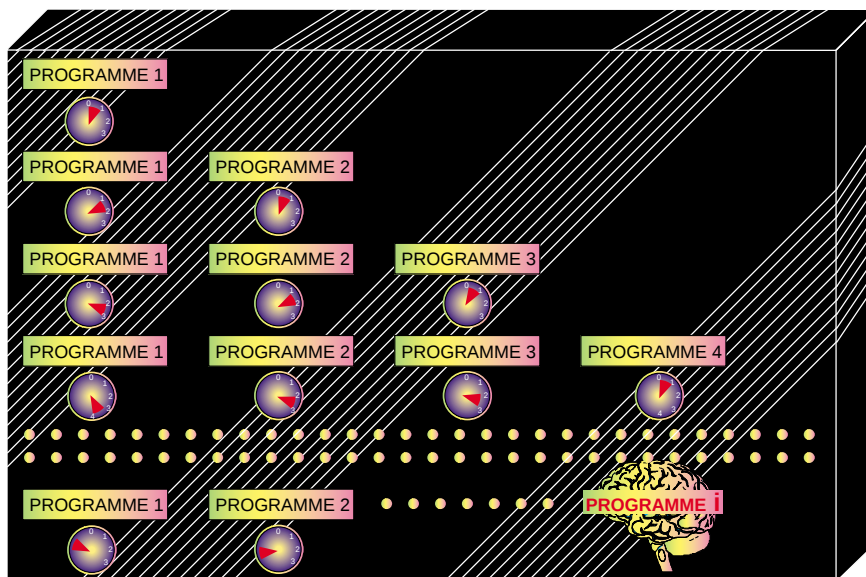




2. Le *déploieur universel* est un programme qui, une fois lancé, ne s'arrête jamais et fait fonctionner tous les programmes envisageables (en théorie de la calculabilité, on démontre que c'est possible). Il prend la liste de tous les programmes, *Programme 1*, *Programme 2*, *Programme 3*, ... et les 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 *déploieur universel* et qu'on le fait fonctionner indéfiniment, alors il finit par reproduire votre fonctionnement mental : un *déploieur 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 *déploieur universel*.

Un tel *déploieur 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 *déploieur 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 *déploieur 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 *déploieurs 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 *déploieur 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 *déploieur 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 *déploieur 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 *déploieur 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 *déploieur 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 :

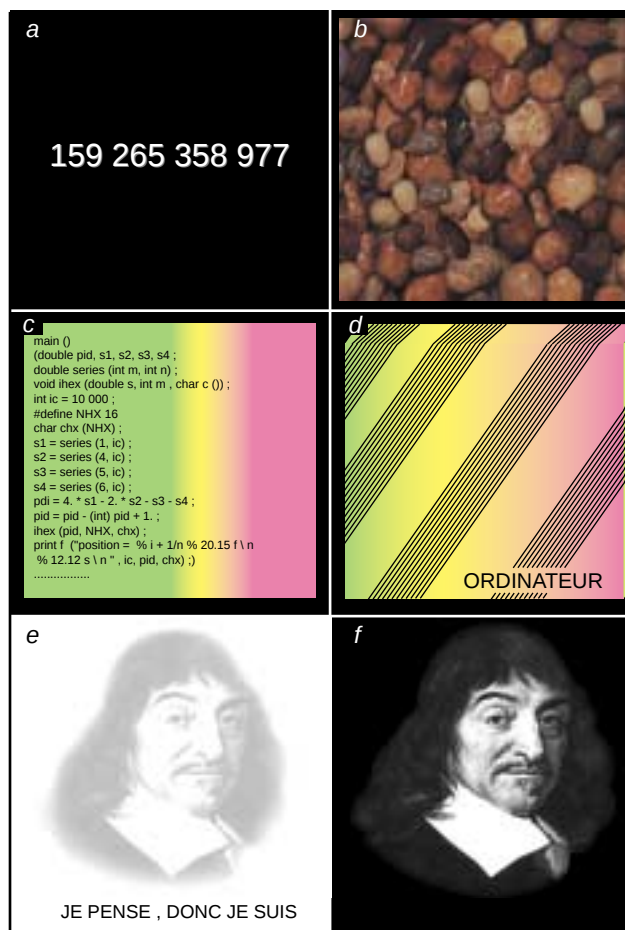
– si nous sommes des *machines numériques* (hypothèse 1) et,

– 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 s'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-