

LES LEÇONS DE L'INTELLIGENCE ARTIFICIELLE :

Approche phénoménologique des savoirs informels

Raphaël Goubet

Je voudrais ici étudier, dans une démarche philosophique, la question de savoir si la pensée est fondée ou non dans des mécanismes formels. Parler, à l'heure actuelle, de mécanismes formels de la pensée, c'est irrésistiblement évoquer, comme point de comparaison, l'ordinateur. C'est pourquoi, pour illustrer mon propos, je me référerai à la recherche en Intelligence Artificielle (IA). Nous serons ainsi confrontés aux limites des modèles formels de la pensée, pour en venir à des démarches plus fructueuses, s'inscrivant dans le chemin de la phénoménologie.

Les principes fondamentaux de l'Intelligence Artificielle

Nulle part à l'époque moderne cette problématique des fondements des savoirs n'a été aussi vive que dans les sciences cognitives. Nées au cours des années 60, elles sont parties d'une conception particulièrement radicale de l'esprit suite à leurs bases conceptuelles, et appelée cognitivisme. Elles sont en effet le résultat d'une des plus

importantes inventions du XX^e siècle : l'ordinateur. Il est nécessaire de s'y attarder un bref instant.

L'ordinateur est une machine de manipulation logique de morceaux d'information, des bits. La chose est plutôt simple, mais elle va ouvrir à des perspectives inouïes grâce à deux thèses à leur propos :

(1) Alan Turing montre qu'on peut reproduire les inférences d'un système logique grâce à un type spécifique de machine, appelée machine de Turing, et dont l'ordinateur est la réalisation la plus aboutie. De plus, toute machine de Turing est simulable par une machine de Turing¹.

(2) Newell & Simon, pionniers de l'Intelligence Artificielle, voient dans l'ordinateur un modèle *littéral* de l'esprit humain. Ainsi, tout état interne d'un ordinateur peut être interprété comme un symbole représentant un quelconque fait du monde, et toute opération sur ces symboles constitue un modèle psychologique valable².

L'ordinateur lui-même n'est qu'une machine de manipulation, qu'on peut dire inférentielle dès lors que ses opérations sont mises en relation avec le monde. Pour effectuer ces manipulations, elle doit suivre un ensemble d'instructions appelé programme. Le programme est, par nature, fonctionnellement indépendant du substrat matériel qui le réalise : plutôt qu'un ordinateur électronique, ce pourrait aussi être un assemblage de poulies et de leviers. Cette thèse, qui dérive directement de celle de Turing, est connue sous le nom de

¹ A. TURING, "Computing Machinery and Intelligence", *Mind*, 59 (236), 1950, pp. 433-460 (trad. fr. : "Machines à calculer et intelligence", in A. PÉLISSIER et A. TÊTE (éds), *Science cognitives : textes fondateurs (1943-1950)*, Paris, Presses Universitaires de France, coll. "Psychologie et sciences de la pensée", 1995, pp. 247-295). Pour le dire autrement, la thèse de Turing "affirme que pour toute procédure, manipulation ou fonction d'entiers que l'esprit humain peut exécuter ou calculer concrètement, effectivement (à la manière dont les mathématiciens depuis l'aube des temps appliquent des algorithmes, et les cuisiniers des recettes), il existe une machine (de Turing) capable d'exécuter cette procédure ou de calculer cette fonction" (D. ANDLER, "Calcul et représentation : les sources", in D. ANDLER (éd.), *Introduction aux sciences cognitives*, Paris, Gallimard, coll. "Folio Essais", n° 179, 1992, p. 24).

² A. NEWELL & H.A. SIMON, "Computer Science as Empirical Inquiry : Symbols and Search", *Communications of the ACM*, 19(3), 1976, pp. 113-126.

*fonctionnalisme*¹. Elle revient à dire que nos états mentaux sont définis par une certaine organisation fonctionnelle abstraite, et non par la substance dans laquelle ils sont réalisés. Un "isomorphisme fonctionnel" entre un ordinateur fait de composants électroniques et un être humain de chair et d'os devient donc concevable.

Le tour était joué : un ordinateur peut réaliser presque instantanément les opérations les plus abstraites et complexes de l'esprit humain. Tout état interne d'un ordinateur peut alors être assimilé à un état mental. Il n'y avait qu'un pas qu'on franchit allègrement pour poser que l'esprit humain consiste en un programme spécifiant la manipulation de symboles abstraits dépourvus en eux-mêmes de sens, selon les règles de la logique formelle. Le savoir est donc l'objet d'une description purement syntaxique. De plus, pour réaliser une intelligence artificielle, il suffit de mettre au point un programme adéquat. Tout savoir est donc, par essence, purement formel.

La critique phénoménologique

Pourtant, très tôt, l'intelligence artificielle va buter sur des problèmes importants. Par exemple, dans la reconnaissance et la compréhension de texte (étudiée pour mettre au point notamment des traducteurs automatiques), on se rend compte très vite qu'il y a bien plus dans la langue que des règles syntaxiques à la Chomsky, fussent-elles universelles. Chaque terme n'a de sens que dans la situation concrète qui le fixe, question bien connue de la pragmatique. Or, il n'y a aucun moyen de décrire cette situation concrète et toujours particulière selon la forme abstraite et universelle requise par les ordinateurs. Ensuite, on s'est aperçu qu'un mot n'a de sens, dans une phrase, qu'en fonction de son insertion dans l'*ensemble* de la phrase. Cela signifie qu'on ne peut comprendre un mot avant d'avoir compris la phrase. Or, un ordinateur, manipulant des symboles discrets, ne peut analyser une phrase qu'à partir de ses parties selon une méthode strictement atomiste. La première leçon à en tirer est que l'esprit humain n'est pas atomiste, mais holiste.

¹ H. PUTNAM, "Minds and Machines", *Mind, Language and Reality : Philosophical Papers*, vol. 2, Cambridge, Cambridge University Press, 1975, pp. 362-385 ; J. FODOR, "Le corps et l'esprit", *Pour la Science*, n° 43, mai 1981, pp. 78-88.

Un autre exemple classique d'échec est le jeu du même nom. La difficulté, pour l'élaboration d'un programme assez intelligent pour jouer aux échecs au même niveau que les meilleurs joueurs humains est que, si l'ordinateur doit prendre en compte tous les coups possibles, toutes ses conséquences et toutes les réponses envisageables, la masse de calcul devient insupportable pour la machine. La solution a alors consisté à demander à des joueurs réels d'expliquer leurs performances, et d'implémenter ces petits trucs, qu'on appelle des heuristiques, dans la machine. L'une de ces heuristiques consiste à donner un poids à chaque pièce en fonction du nombre de pièces semblables disponibles et de leurs capacités dans le jeu.

Pourtant, la technique ne donna que des résultats forts moyens –il fallut attendre des ordinateurs beaucoup plus puissants, de type *Big Blue*. Quelles que soient les performances futures de la *machine*, le *programme*, lui, semblait très limité. C'est qu'il n'est pas très difficile de se rendre compte qu'un joueur d'échecs avancé ne suit pas des règles, mais au contraire fait usage de son intuition, de son expérience de milliers de situations concrètes et vécues.

Ces divers problèmes seront très vite remarqués par le philosophe Hubert Dreyfus, qui, inspiré par la phénoménologie continentale, mènera, dès les années 60, une polémique *critique de la raison artificielle*. Il apparaît ainsi que l'esprit humain ne fonctionne pas en décomposant le monde auquel il est confronté en des entités discrètes qu'il manipule ensuite selon des règles. Au contraire, il appréhende le monde dans lequel il se trouve comme un tout. Ce point avait été relevé déjà par la *gestalt theorie*. Il signifie que tous nos rapports intelligents au monde sont déterminés par notre *implication* dans ce monde, une contextualisation toujours vécue par un corps propre. Nous ne donnons un sens au monde que par notre attitude intentionnelle et téléologique à son égard : nous n'adoptons pas des comportements intelligents vis-à-vis de faits du monde en fonction du sens que nous leur donnons préalablement ; au contraire, ces faits ne prennent leur sens que par notre implication vis-à-vis d'eux, par les buts que nous poursuivons par notre être-au-monde¹. Le *sens* n'est pas logique, il est profondément *existentiel*. En aucun cas cette implication n'est-elle descriptible par des règles.

¹ Sur ce point de la philosophie de Dreyfus, voir D. ANDLER, "Context and Background : Dreyfus and Cognitive Science", in M. WRATHALL & J. MALPAS (eds), *Heidegger, Coping and Cognitive Science : Essays in Honor of Hubert L. Dreyfus, Volume 2*, Cambridge (MA), MIT Press, 2000, pp. 137-159.

Les objets de notre perception nous sont toujours donnés selon une structure figure/fond. Le monde ne nous apparaît donc pas comme entièrement déterminé, spécifié, formellement modélisable : notre esprit l'appréhende toujours sur un fond indéterminé, que Husserl appelait l'horizon extérieur. Ainsi, le joueur d'échec n'appréhende un problème donné au cours du jeu que sur le fond indéterminé que représente la partie, en considérant l'échiquier comme un tout déjà signifiant, et qu'il saisit par son intuition. Il en va de même pour tout un chacun lorsqu'il mène une conversation, où les mots sont pris ensemble. Comme le faisait remarquer Michael Polanyi, si l'on devait porter son attention sur chaque mot, un pianiste sur chaque note, nous serions incapables de donner un sens au monde¹.

Mais, outre cet horizon extérieur, les objets de notre perception sont dotés d'un horizon intérieur : nous ne percevons un objet que selon un profil donné, tous les autres restant indéterminés. Cependant, l'objet m'est immédiatement donné comme un tout, à travers le profil perçu, auquel il ne se réduit pas. Lorsque les autres profils nous sont présentés, ils ne remplacent pas la perception précédente : les différents profils sont perçus comme co-présents. Le sens du profil donné est déterminé par les profils qui me sont cachés : c'est le tout qui détermine les parties, alors qu'une modélisation formelle ne pourrait aller que des parties au tout. Au contraire, comme le dit Dreyfus, pour l'esprit humain, ce qui compte comme parties est défini dans les termes du tout : "une gestalt définit ce qui compte comme l'élément qu'elle organise ; un plan ou une règle organise simplement des éléments définis indépendamment"².

Nos savoirs, y compris les plus abstraits, sont fondés (notamment du point de vue de l'ontogenèse) dans nos compétences concrètes acquises au cours de l'expérience. Cette implication dans le monde par les actions que nous y menons est rendue possible par notre

¹ M. POLANYI, *Personal Knowledge*, Chicago, University of Chicago Press, 1962, pp. 56-57.

² H.L. DREYFUS, *What Computers Still Can't Do : A Critique of Artificial Reason* (3ème éd.), Cambridge, MIT Press, 1992, p. 245 (trad. fr. : *Intelligence artificielle : mythes et limites*, Paris, Flammarion, 1984). Il est intéressant de noter que c'est dans des cas pathologiques (*agnosie visuelle*) que l'esprit semble fonctionner comme un ordinateur, c'est-à-dire en identifiant systématiquement des traits indépendants pour reconnaître ensuite un objet ou un visage, en allant donc des parties au tout. Voir à ce sujet la célèbre histoire du Docteur P. racontée par Oliver Sacks dans le chapitre titulaire de *L'homme qui prenait sa femme pour un chapeau*, Paris, Éd. du Seuil, coll. "Points essais", n° 245, 1988.

incarnation, par notre inscription dans un corps situé. Or, lorsque je mène une action sur le monde à l'aide d'une compétence acquise par l'expérience, je ne le fais pas en suivant des règles, des formules, mais automatiquement, instinctivement, parce que j'ai déjà été confronté à des expériences semblables. Et si je les perçois comme semblables, c'est grâce à des dispositions corporelles préalables à l'action elle-même : mon corps retrouve une situation déjà connue, quoique différente, par exemple une disposition spatiale identifiable à un exemple prototypique déjà rencontré. Nos compétences sont ainsi mises en œuvre en faisant appel à une collection d'exemples concrets que mon corps reconnaît comme similaires, et non en suivant des règles abstraites et générales. Celles-ci ne sont impliquées qu'aux premiers moments où une compétence est acquise, mais ensuite disparaissent pour laisser la place à une action intuitive. Ainsi, lorsque j'apprends à conduire, c'est consciemment que je dois gérer les mouvements de mon pied sur la pédale d'embrayage, mais peu à peu, cette action analytique laisse la place à une action synthétique, réalisée non par un calcul abstrait, mais par les mouvements d'un corps vécu. Comme l'indique Merleau-Ponty :

(...) si mon corps peut être une "forme", et s'il peut y avoir devant lui des figures privilégiées sur des fonds indifférents, c'est en tant qu'il est polarisé par ses tâches, qu'il existe vers elles, qu'il se ramasse sur lui-même pour atteindre son but (...) ¹.

Pour Merleau-Ponty, apprendre à sentir la soie, ce n'est pas comparer des traits perçus avec une série de propriétés stockées en mémoire ; c'est apprendre à déplacer sa main vers le tissu, à confronter cette action à des attentes. Cela ne peut se faire que par un agir continu sur le monde, et ceci non pas guidé par un seul sens, mais par tous les sens. D'un objet que je vois, que je touche, dont j'entends le bruit lorsque je le frappe, je n'ai pas une image disparate d'informations diverses : au contraire, il m'apparaît comme un tout. Acquérir une compétence vis-à-vis d'une situation peut donc se faire d'une infinité de manières : si j'apprends à reconnaître de la soie par la vue, en voyant la manière dont elle reflète la lumière, je deviens tout aussi capable de l'identifier par le toucher. Il n'y a aucun moyen, pour un

¹ M. MERLEAU-PONTY, *Phénoménologie de la perception*, Paris, Gallimard, coll. "Tel", n° 4, 1945, p. 117.

ordinateur, de reproduire cette capacité synesthésique. Chaque sensation serait pour lui une source d'informations différentes et incommensurables.

Dreyfus, qui reste ici ma principale référence, montre donc que les actions compétentes menées par les humains n'impliquent pas nécessairement qu'ils suivent des règles : "parce que nous sommes incarnés, les règles nécessaires pour donner une analyse objective de notre compétence n'ont en aucun cas besoin d'être impliquées dans notre performance"¹. Pour décrire le processus par lequel on passe de l'action d'un débutant faisant consciemment appel à des règles à celle d'un expert ayant acquis une compétence intuitive, il détaille une évolution en cinq étapes².

Au long de ce parcours, l'élève part d'un stade où sa performance repose sur la reconnaissance de traits acontextuels, tout à fait généraux, à partir desquels il doit appliquer des règles pour déterminer l'action suivante adéquate. Peu à peu, sa compétence évolue de telle manière que l'usage de règles devient de plus en plus exceptionnel, remplacé par une appréhension plus intuitive de la situation et de ce qu'il convient de faire, grâce à une expérience d'un nombre croissant de situations concrètes différentes. Arrivé au dernier stade, celui de l'expert, plus aucun recours à des règles n'est nécessaire. Le sujet est capable d'appréhender la situation dans sa globalité, d'opérer des décisions immédiates par identification de la situation avec une autre déjà rencontrée. Il n'existe plus de règles : il n'y a plus qu'un long et complexe apprentissage au cours duquel un nombre énorme de situations ont été rencontrées, et une capacité à faire face intuitivement à n'importe quelle situation nouvelle. On estime ainsi que le grand maître d'échecs est capable de distinguer quelques 50 000 types de positions, auxquelles il peut répondre immédiatement, sans devoir les analyser.

Le dernier stade de ce parcours peut paraître exclusivement réservé à un nombre réduit d'individus particulièrement compétents. Pourtant, nous sommes tous des experts de la vie quotidienne, dans notre capacité à faire face à la contingence du réel, à un nombre incalculable de situations nouvelles et imprévisibles. Cette capacité,

¹ H. DREYFUS, *op. cit.*, p. 254.

² H.L. DREYFUS & St.E. DREYFUS, *Mind Over Machine : The Power of Human Expertise in the Era of the Computer*, Oxford, Basil Blackwell, 1986. Pour un texte en français, voir H.L. DREYFUS, "La portée philosophique du connexionnisme", *Introduction aux sciences cognitives, op. cit.*, pp. 352-373.

acquise au cours d'un plus ou moins long apprentissage qui a eu lieu tout au long de notre vie, est profondément liée à notre incarnation et notre contextualité, et ne peut donc se réduire à des savoirs purement formels, ce qui avait été le point de départ de l'intelligence artificielle classique. Pourtant, d'autres approches ont émergé. Les plus importantes dérivent de l'étude des réseaux neuronaux, mais leur complexité conceptuelle est telle que je ne pourrai ici les aborder¹. Pour les besoins de l'illustration, je préfère étudier une autre démarche radicale, que l'on doit à Rodney Brooks.

Brooks, au MIT, construit des robots, c'est-à-dire des agents incarnés et situés. Ceux-ci ne sont pas dotés d'un programme central, mais leur comportement émerge de la coordination de "couches" comportementales autonomes, chacune dédiée à un aspect particulier du comportement. Cette coordination passe par un échange minimum d'information entre les couches. Du fait de l'absence de programme central, ces robots ne possèdent aucune représentation du monde extérieur².

Un robot de Brooks n'est donc pas soumis à un programme central déterminant son comportement. Celui-ci émerge au cours de son interaction concrète avec son environnement, en fonction de ses besoins réels et des tâches particulières qui lui ont été fixées. Il est doué d'un véritable sens commun, sommaire, mais issu d'une véritable action délibérée et orientée vers un but. Il est capable de faire face à la contingence du réel parce que, comme nous, il ne porte attention qu'aux éléments de l'environnement qui sont significatifs pour lui à un moment donné, selon son objectif, le reste demeurant indéterminé.

Prenons l'exemple d'Herbert, un robot conçu pour récolter des canettes vides. Le comportement d'Herbert, rapide et souple, qui lui

¹ A ce sujet, voir la nouvelle introduction à H. DREYFUS, *What Computers Still Can't Do* et "La portée philosophique du connexionnisme", *op. cit.* ; H.L. DREYFUS & St.E. DREYFUS, "Making a Mind Versus Modeling the Brain : Artificial Intelligence Back at a Branchpoint", *Daedalus*, 117(1), 1988, pp. 15-43 ; sur le connexionnisme, voir D. ANDLER, "Connexionnisme et Cognition : à la recherche des bonnes questions", *Revue de Synthèse*, IV(1/2), 1990, pp. 95-127 ; A. CLARK, *Microcognition : Philosophy, Cognitive Science and Parallel Distributed Processing*, Cambridge (MA), MIT Press, Explorations in Cognitive Science, 1989.

² R. BROOKS, *Intelligence Without Reason*, Cambridge (MA), MIT, Artificial Intelligence Laboratory, A.I. Memo No. 1239, April 1991 ; "Intelligence Without Representation", *Artificial Intelligence*, 47(1-3), 1991, pp. 139-159 ; <<http://www.ai.mit.edu>>. Voir aussi A. CLARK, *Being There : Putting Brain, Body and World Together Again*, Cambridge (MA), MIT Press, 1997.

permet d'atteindre ce but, émerge grâce à l'interaction de diverses couches comportementales autonomes au sein de l'environnement. À aucun moment n'a-t-il de représentation globale du but qu'il poursuit ; il n'est pas, dès le départ, à la recherche de canettes, dont il aurait une représentation "en tête", parcourant son environnement pour les trouver. C'est l'environnement lui-même qui va le guider, en déclenchant des réponses comportementales appropriées. Pour reprendre les termes de Brooks : le monde est son propre modèle. Une première couche permet d'éviter les obstacles : Herbert s'arrête s'il se trouve près d'un obstacle immobile, et se réoriente dans une direction non-bloquée. Tant que cette couche est inactive, la couche locomotrice lui permet de se déplacer. Cependant, ces deux couches sont interrompues si une forme identifiable comme celle d'une table est rencontrée. A ce moment, une nouvelle routine est activée, comprenant la détection d'une éventuelle canette. Dès lors, un bras prend la relève, explore la table et saisit la canette. Jamais Herbert n'a-t-il suivi un plan global déterminant la recherche de canettes. Ce comportement émerge de diverses routines spécifiques à un but pratique : il sait se déplacer, repérer une canette, la saisir, mais il ne sait pas, lorsqu'on le met en fonction, que son but est de collecter des canettes.

Communiquer en action

Mais qu'en est-il alors de la communication ? Si les savoirs sont acquis et la pensée déterminée par l'expérience d'un sujet incarné dans un monde réel, comment peuvent-ils être transmis, communiqués ? La conception traditionnelle voit dans la forme abstraite la clé du mystère : en s'élevant au-dessus des particularités empiriques et contingentes de son expérience propre, en faisant usage d'un code abstrait, arbitraire, l'individu atteindrait l'universalité *a priori* nécessaire à toute communication. Mais n'est-ce pas poser comme condition préalable ce dont l'existence doit être élucidée ? Cette théorie semble plutôt faire seulement reculer le problème, puisqu'elle laisse inexpliquée la nature de cet *a priori* du langage (Habermas) ou de la communication (Apel).

Pourtant, chercher dans le formalisme le fondement de la communication, ainsi que l'avait fait la sémiotique structurale¹, c'est manquer la nature initialement pratique et existentielle de la communication. De cette implication du sujet dans le monde ne découle pas nécessairement son isolation : l'expérience n'en est pas moins sociale, partagée ; le contexte sur lequel j'ai insisté, avec Dreyfus, n'est pas qu'un contexte physique, il est aussi un contexte social, susceptible d'être vécu avec l'autre. Durant le parcours d'apprentissage décrit par Dreyfus, où l'élève acquiert une compétence sur base de son expérience propre, il ne le fait pas dans le repli de sa seule subjectivité. Il est, au contraire, accompagné par une personne plus avancée dans sa compétence, peut-être un expert, qui guidera l'élève pour que son expérience soit constructive, porte ses fruits, et ce, non à l'aide de formules abstraites qu'il faudrait assimiler (mais alors, sur base de quoi ?), mais grâce notamment à une série d'essais et d'erreurs, en coopération avec autrui.

Toute interaction entre sujets est ainsi fondée non pas tant dans le partage préalable de symboles arbitraires, que dans celui de buts, de desseins, de projets, d'expériences, une communication "en actes". La communication, écrit Dewey, est :

l'établissement de la coopération dans une activité dans laquelle il y a des partenaires, et dans laquelle l'activité de chacun est modifiée et régulée par le partenariat. Échouer à comprendre, c'est échouer à être en accord dans l'action ; mal comprendre, c'est faire reposer l'action sur des desseins hybrides².

Et la signification que nous accordons aux objets n'est pas une sorte d'entité universelle et immuable qui viendrait se poser sur eux aux termes d'un accord intersubjectif ; ce n'est pas plus une simple

¹ L'alternative pragmatiste à la sémiologie structurale, celle que l'on doit à Peirce, se distingue précisément en ce que, tout en étant une théorie des signes, elle n'en fait pas une entité purement formelle. On pourrait croire que l'approche pan-sémiotique de Peirce, pour qui toute pensée est signes, se rattache au computationnisme symboliste. Mais il n'en est rien : l'une des notions centrales de la théorie de Peirce, l'interprétant, ne se laisse jamais réduire à des symboles abstraits et discrets. Interprétant d'un signe, il est signe lui aussi et a son propre interprétant, d'où naît une sémiose infinie que seule boucle une habitude d'action pratique, c'est-à-dire qui se rapporte toujours à une situation donnée réelle : le sens a une nature pratique, et la pensée se résout dans l'action, et non dans la forme.

² J. DEWEY, *Experience and Nature*, La Salle, Open Court, 1929, pp. 148-149.

qualité de ces objets ; ce n'est toujours pas l'inclination du seul sujet vis-à-vis de cette chose :

(...) le sens est l'acquisition de signification par des choses dans leur statut à rendre possible et à remplir une coopération partagée, les choses signifiantes sont des choses en fait impliquées dans des situations de desseins et d'exécution partagés ou sociaux¹.

Dans cette interaction, l'autre possède un statut particulier, il n'est pas une simple chose sur laquelle je puis agir : il est un autre *avec lequel je peux interagir*. S'il n'est pas un objet, il partage cependant avec moi son implication dans un monde dans lequel il est aussi incarné. Tout comme la pensée est déterminée par l'incarnation et la contextualité du sujet, la communication l'est par la "co-incarnation" et la co-opération de deux individus au moins. Autrui pose vis-à-vis de moi des actions par son corps, qui se révèlent en accord ou en désaccord avec mes propres actions : nos projets se rencontrent, s'affrontent, s'évitent, se renforcent, se ressource l'un dans l'autre ou cherchent à s'éteindre, ils s'équilibrent ou se détruisent. Mais si cette communication est possible, c'est parce qu'à travers cette coopération et cette co-incarnation, autrui devient un prolongement de moi-même, son corps complète le mien, en quelque sorte, et je trouve (ou non) en lui une résonance particulière.

Quelqu'un se sert de mes objets familiers, écrit Merleau-Ponty. Mais qui ? Je dis que c'est un autre, un second moi-même, et je le sais d'abord parce que ce corps vivant a même structure que le mien. J'éprouve mon corps comme puissance de certaines conduites et d'un certain monde, je ne suis donné à moi-même que comme une certaine prise sur le monde ; or, c'est justement mon corps qui perçoit le corps d'autrui et il y trouve comme un prolongement miraculeux de ses propres intentions, une manière familière de traiter le monde ; désormais, comme les parties de mon corps forment ensemble un système, le corps d'autrui et le mien sont un seul tout, l'envers et l'endroit d'un seul phénomène et l'existence anonyme dont mon corps est à chaque moment la trace habite désormais ces deux corps à la fois².

¹ *Ibid.*, pp. 149 et 150.

² M. MERLEAU-PONTY, *op. cit.*, p. 406.

La seule forme *a priori*, pourrait-on dire, c'est donc celle du corps, par lequel passent toute notre perception et notre expérience du monde, sa structure nécessairement commune. Mais cette "forme" n'en est pas vraiment une : elle est une forme existentielle, vécue, et non essentielle et désincarnée. Elle est mouvante dans son expérience, mais reste inchangée dans sa référence à l'existence du sujet et de l'autre. Si je puis d'emblée imputer à autrui une conscience, et une expérience comparable à la mienne, ouvrant donc à la possibilité d'un partage, c'est parce qu'il se présente avec un corps, et donc avec une expérience qui prolonge la mienne. Toute signification n'est partageable et partagée que parce que je la communique à autrui en lui imputant une expérience similaire à la mienne, parce qu'il a un corps, tout comme moi, et le même corps que moi. Avant tout concept formel, il y a une présence de son propre corps et du corps de l'autre par laquelle toute communication est rendue possible¹. La communication émerge d'un sens du corps : sens de son propre corps, mais aussi sens du corps de l'autre, ce tout que forment mon corps et celui d'autrui.

Il est remarquable que cette analyse phénoménologique soit aujourd'hui soutenue empiriquement, notamment par la découverte des "neurones miroirs". Vittorio Gallese et ses collègues ont remarqué que ces neurones sont actifs lorsqu'une action volontaire est réalisée par un sujet, mais aussi quand ce dernier observe la même action chez l'autre :

Ce processus *implicite, automatique et inconscient* de simulation motrice permet à l'observateur de faire usage de ses propres ressources pour pénétrer le monde de l'autre sans avoir besoin de théoriser à ce sujet, sans avoir besoin d'utiliser nécessairement une attitude propositionnelle. (...) L'action est le principe "a priori" qui permet l'établissement initial de liens sociaux².

C'est ce qui amène Gallese à avancer que,

Loin d'être exclusivement dépendante d'aptitudes mentalistes/linguistiques, la capacité à comprendre les autres comme agents intentionnels est profondément enracinée dans

¹ A ce sujet, voir aussi M. JOHNSON, *The Body in the Mind : The Bodily Basis of Meaning, Imagination and Reason*, Chicago, University of Chicago Press, 1987.

² V. GALLESE, "The 'Shared Manifold Hypothesis' : from Mirror Neurons to Empathy", *Journal of Consciousness Studies*, 8(5-7), 2001, p. 41.

la nature relationnelle de l'action. (...) L'action intentionnelle [agency] constitue un point clé pour la compréhension de l'intersubjectivité et pour expliquer comment les individus peuvent interpréter leur monde social¹.

On voit donc que c'est à travers la perception du corps de l'autre et de ses actions qu'autrui m'est donné comme être doué d'intentions et de conscience. Il ne s'agit pas d'une simple analogie entre autrui et moi qui m'autoriserait à inférer formellement la présence d'une conscience étrangère, mais d'une prolongation de mon expérience en autrui : "quand mon regard croise un autre regard, écrit Merleau-Ponty, je ré-effectue l'existence étrangère dans une sorte de réflexion"². D'une manière presque littérale, je me mets dans sa peau, je me glisse en lui, je deviens comme le miroir de ses actions. Ce corps vécu est le siège d'une intersubjectivité *a priori*, préalable au langage et à la forme, et condition de toute rencontre effective avec l'autre³.

La communication est donc à la base le partage d'une expérience, et non celui d'un code. Les approches structuralistes avaient buté sur ce problème, et la recherche prit alors la direction de la pragmatique. Mais même elle, en cherchant à formaliser les usages, les actes de paroles, à les réduire à des concepts tels que la force illocutoire, n'a pas vu que le fondement de la communication est d'une nature proprement non-formalisable. La communication repose sur la coopération et la coïncarnation de sujets, dont elle ne peut donc être abstraite.

Est-ce à dire que la forme n'a finalement qu'un rôle secondaire dans la communication ? En aucun cas. Reconnaître un fondement informel à la communication, ce n'est pas l'y réduire. Il reste évident que la formalisation est une capacité humaine fondamentale, qui lui permet d'étendre les limites de son expérience. Confiné à sa seule expérience concrète, le sujet serait incapable de dépasser la contingence du réel, de placer celle-ci dans le cadre d'une expérience

¹ *Ibid.*, p. 34.

² M. MERLEAU-PONTY, *op. cit.*, p. 404.

³ Sur l'analyse phénoménologique de l'intersubjectivité, voir D. ZAHAVI, "Beyond Empathy : Phenomenological Approaches to Intersubjectivity", *Journal of Consciousness Studies*, 8(5-7), 2001, pp. 151-167. Pour d'autres exemples de rencontres entre la phénoménologie et les sciences empiriques, on consultera J. PETITOT, Fr.J. VARELA, B. PACHOUD & J.-M. ROY (eds.), *Naturalizing Phenomenology : Issues in Contemporary Phenomenology and Cognitive Science*, Stanford, Stanford University Press, coll. "Writing Science", 1999.

continue, d'en relier les composants et de leur donner une signification nouvelle. La formalisation de la pensée et des savoirs les fait accéder à une dimension nouvelle, elle permet la créativité et les merveilles de l'imagination. Mais précisément, cette créativité serait impossible si la pensée devait avoir son essence dans des formes immuables ; au contraire, elle émerge de situations concrètes souvent instables, voire chaotiques, dont le sujet parvient à extraire une signification plus déterminée, durable, mais qui le confrontent aussi à l'imprévu, l'inconnu, la nouveauté.

Ce passage de l'informel au formel, et ce retour permanent à l'informel, n'est pas le passage d'une frontière fixe. Si nos actions, nos pensées, nos paroles peuvent être décrites par des règles invariables, ce n'est qu'*a posteriori*. Mais cette accession au domaine du formel ne se produit que lorsque, d'une expérience, le sujet devient capable d'isoler une "essence" qu'il peut rapporter à une autre expérience, éventuellement nouvelle, pour en renouveler le sens. Cette essence, pourtant, n'est pas une Idée platonicienne, ce n'est pas une réalité ultime et immuable ; elle reste liée à la réalité dont elle émerge, elle reste conditionnée aux expériences partagées des sujets qui en reconnaissent l'existence, et c'est en cela qu'elle n'est pas immuable, qu'elle peut être recrée. La pensée n'est pas un processus imparfait tendant à la compréhension d'une réalité parfaite ; elle est une action mouvante suivant le flot et la richesse de l'expérience.

La condition, donc, pour qu'un symbole, par exemple, puisse prétendre être une forme intersubjectivement valide, ce n'est pas qu'il reflète exactement –et pour tous– la réalité dont il se veut le miroir. C'est plutôt que les sujets, dans l'usage qu'ils en font, reconnaissent sa valeur pratique, qu'ils s'accordent sur le fait que ce symbole correspond à la réalité dont ils font l'expérience. Si la réalité ou si l'expérience de celle-ci devait changer, le symbole lui-même, sa signification et sa valeur, changeraient, non pas tant pour correspondre mieux à la réalité à laquelle on le rapporte (plus qu'il ne s'y rapporte lui-même), que pour être recréé par la coopération et la coïncarnation des individus. Autrement dit, la communication est fondée dans des *pratiques* discursives, qui elles-mêmes ont leurs racines dans la coïncarnation des sujets.

Revenons un instant à l'intelligence artificielle. J'ai associé plus haut le nom d'Alan Turing à la notion de "machine de Turing", mais on lui doit aussi une autre idée très importante pour l'IA classique : le test de Turing. En substance, il s'agit de dire que, pour savoir si un

ordinateur pense, il doit être possible d'avoir avec lui une conversation de telle manière qu'on ne puisse le distinguer d'un interlocuteur humain. Donc, nous avons ici l'idée que l'intelligence et la pensée sont révélées par la communication, elle-même réduite à la manipulation linguistique de symboles formels¹. Pourtant, il est concevable qu'une machine puisse manifester un comportement intelligent, incluant des performances communicationnelles, et ce, sans même être capable de passer le test de Turing².

Retournons à l'exemple des étonnants robots de Brooks. Imaginons qu'Herbert soit secondé dans sa tâche ménagère par Simon, son alter ego — ainsi baptisé pour que le duo soit un hommage à l'un des pères de l'intelligence artificielle. En pratique, rien n'empêche Herbert et Simon de récolter les canettes chacun dans leur coin. Mais voici qu'arrive un cas particulier : tous deux ont repéré la même canette, et se la disputent, le bras de l'un l'arrachant à celui de l'autre. Faut-il, pour éviter ce cas de figure, insérer dans le programme d'Herbert et Simon, une théorie marxiste du partage du travail pour le bien-être de la collectivité ? Faut-il, plus sérieusement, leur apprendre un langage leur permettant de négocier l'accomplissement de leur tâche commune, dont, rappelons-le, ils n'ont aucune représentation ? Non : il suffit simplement d'ajouter une couche comportementale interrompant la saisie de la canette si un autre membre de l'espèce se l'est déjà appropriée. Sans qu'il soit nécessaire de faire appel à un principe formel représentant une essence de la coopération et de la communication (comme un langage), Herbert et

¹ A. TURING, *op. cit.* Tout ceci présume une série de présupposés assez stricts et guère maintenus aujourd'hui, du moins sous cette forme. Notamment, les animaux ne penseraient pas puisqu'ils ne parlent pas. On peut aussi se demander s'il faut considérer les aphasiques comme doués de pensée. Enfin, la validité du test de Turing présuppose que c'est le langage qui constitue la pensée, ce qui correspond bien au tournant linguistique de la philosophie analytique, alors dominante à l'époque de Turing.

² On doit à Searle la réfutation inverse du test de Turing, l'argument de la "chambre chinoise", selon lequel on peut manipuler des symboles avec succès (et donc passer le test) sans même en comprendre le sens (et il n'y a pas de pensée sans intentionnalité). Plus de vingt ans après, cet argument est toujours l'objet de vives polémiques. Voir J. SEARLE, "Minds, brains and programs", *Behavioral and Brain Sciences*, 3(3), 1980, pp. 417-457. En français, voir *De l'esprit au cerveau*, Paris, Hermann, 1985, chap. 2 et "L'esprit est-il un programme d'ordinateur ?", *Pour la Science*, 149, mars 1990, pp. 38-44.

Simon se révéleront capables de coopérer dans leur tâche, et donc, de communiquer¹.

On pourrait douter que cet exemple simpliste constitue un véritable cas de communication. Pourtant, un peu d'imagination devrait persuader de la véritable portée de l'argument. Si l'on voulait apprendre à Herbert et Simon à coopérer pour saisir des canettes isolément hors de leur portée, par exemple, en faisant la courte échelle, il ne serait toujours pas nécessaire de leur fournir une représentation globale de leur tâche, des formes que peut prendre une coopération, et moins encore un langage comme instrument de manipulation symbolique de ces représentations. De nouvelles couches comportementales leur permettraient de s'interpeller lorsque le besoin se fera sentir, et de s'engager dans un comportement coordonné et adéquat pour réaliser leur but. Et si l'on poursuit l'expérience de pensée, on s'apercevra que, s'ils devaient faire un jour usage d'un langage, celui-ci ne marquerait pas le passage brusque d'une limite vers la dimension symbolique et formelle, mais serait au contraire le fruit d'une évolution continue, dont la base serait encore l'implication pratique dans l'environnement ; il n'y a pas de frontière nette entre la coopération en action et la communication symbolique. Mais prétendre que les exemples plus élémentaires ne sont pas des exemples de communication car il n'y a pas d'échange de symboles revient à poser comme pétition de principe cela même qui doit être prouvé (et que je réfute) : que les symboles formels sont la condition nécessaire de la communication².

Conclusion

J'ai tenté ici de dessiner une image de la pensée et de la communication où celles-ci apparaissent fermement ancrées dans le monde auquel elles se rapportent. Ce rapport du sujet au monde est, on l'a vu, fondamentalement informel, ou plutôt, pré-formel, dans la

¹ Pour un exemple similaire, voir la construction d'une termitière évoquée par I. PRIGOGINE & I. STENGERS, *La nouvelle alliance : Métamorphose de la science*, Paris, Gallimard, 1986 : c'est à partir de comportements simples et coordonnés que les termites arrivent à bâtir une termitière complexe, sans qu'il soit nécessaire qu'ils en partagent une représentation.

² Une vue similaire est développée par E.S. REED, *Encountering the World : Toward an Ecological Psychology*, Oxford, Oxford University Press, 1996.

mesure où c'est dans ce rapport particulier au monde que le domaine du formel prend sa source. J'ai décrit le sens comme pratique et existentiel, et non formel et essentiel : c'est par notre attitude pratique, téléologique, que le monde tel que nous le vivons et dans lequel nous nous impliquons prend son sens. La formalisation de nos compétences n'est qu'une opération *a posteriori* de description de ce qui ne se vit pas par des règles. L'acquisition de compétences est avant tout la capacité à faire appel à une intuition qui, peu à peu, prend la place d'actions régulées. La communication elle-même ne se laisse pas réduire à la forme. Elle est acte, fondamentalement pratique, née de la rencontre de sujets coïncarnés dans le monde dans lequel ils sont impliqués. Le formel, encore une fois, est un *a posteriori* dont on ne doit pas oublier l'origine existentielle.

Pour tracer les contours de cette image, j'ai illustré mon propos par l'un des projets scientifiques les plus ambitieux jamais entrepris : la conception d'une intelligence artificielle. Des échecs de ce programme, on peut tirer de nombreuses leçons. L'esprit humain se découvre dans une complexité plus saisissante encore (en fait proprement vertigineuse¹). Mais nous sommes forcés aussi à l'humilité car nous ne sommes pas des machines surnaturelles. Au contraire, notre matérialité s'impose jusque dans ce que nous pensons être le plus immatériel en nous. Mais c'est précisément là... matière à penser.

¹ Qu'on songe au fait que, dans un seul cerveau humain, il y a des milliers de fois plus de connexions possibles qu'il y a d'atomes dans tout l'univers connu ! (cf. G. EDELMAN, *Biologie de la conscience*, Paris, Odile Jacob, 1992).