

Like robots ?

Aimer les robots ou souhaiter être comme des robots ? Tout l'enjeu se situe dans cette alternative. Les robots se situent dans un *no man's land* et dans un *no god's land* ; la façon dont on les appréhende leur confère un statut très particulier : ce ne sont pas des hommes au sens strictement biologique, mais pas non plus des dieux ou des divinités. Nous avons inventé les robots pour en faire un sosie artificiel du modèle naturel de l'humanité. La peur de voir les machines échapper à la maîtrise humaine est un motif récurrent de la culture populaire (cinéma, vidéo, série...). Même si ce motif connaît des variations selon les périodes, on observe qu'au XX^e siècle, toute l'attention se porte sur ce qui est considéré comme caractérisant le mieux l'humanité. L'enjeu est double : d'une part reproduire l'intelligence humaine (le fonctionnement du cerveau humain) ; de l'autre reproduire l'apparence humaine. Si l'on a tendance à se focaliser sur la première (enjeu existentiel et déontologique), il ne faut pas minorer l'importance du second enjeu (qui semble davantage s'inscrire dans le champ esthétique). Selon les versions et les périodes, le robot peut perdre son anthropomorphisme (taille, corpulence, voix...), parfois au contraire, l'apparence humaine du robot est renforcée, notamment la beauté (à l'instar des récits avec des femmes artificielles : Galatée, Olympia, L'Ève future...). On peut y voir un désir masculin de maîtriser la capacité de reproduction des femmes. Le robot participe à cette volonté de contourner les modalités biologiques de la fabrication de l'humain.

Mais pas seulement. Il faut également prendre en compte la dimension technique intrinsèque à la conception des robots, et plus précisément la notion d'innovation et son rôle moteur. Pour Abraham Moles, le mythe du robot fait partie des « mythes dynamiques », c'est une façon de rapatrier un motif religieux dans le domaine de l'histoire des sciences et des techniques. Il existe selon lui un rapport étroit entre le mythe et la technique, les fables robotiques sont donc une façon pour l'être humain de réaliser ses rêves en asservissant la nature à la loi. Dans cette perspective, on peut postuler que le robot est une représentation paradoxale de l'humain. Le robot est l'une des façons concrètes (techniques) dont cette représentation se forme et se diffuse dans la société. C'est pour cela que le robot est révélateur de la façon dont *nous nous percevons*. En cela, il permet une compréhension de nos sociétés (du rapport de celles-ci à une technique qui nous dépasse). Parce que le robot n'est pas seulement occidental (il existe tout un paradigme robotique en Asie avec ses enjeux propres et ses spécificités ontologiques et culturelles), il renvoie aussi à une disposition centrale de la culture humaine (comme trope, comme forme, et pas seulement comme contenu). Philippe Breton considère que le robot occupe même une position stratégique dans la genèse et le maintien des civilisations (Breton 1995). On peut interpréter cette occupation du champ civilisationnel de deux manières : 1) soit comme une volonté de montrer la différence du robot avec l'être humain, donc une volonté de montrer l'étrangeté du robot ; 2) soit comme une façon

de faire du robot l’image de l’homme. De la sorte, les représentations robotiques renvoient donc plutôt à la question : qu’est-ce qu’un être humain ? Dans les deux cas, les questions de la connaissance de l’homme et de savoir s’il est possible d’en construire une réplique robotique se posent. Ou plus exactement, ce qui est interrogé, ce sont le rôle et l’évolution de nos représentations de l’être humain.

Le robot conjugue deux critères : un critère d’ingénierie (parce qu’il présuppose tout un mécanisme automatique autonome) ; il est potentiellement capable de s’adapter à certains changements de son environnement et de transformer son comportement en conséquence ; le second critère est sociofonctionnel : les robots ont été conçus pour travailler à notre place, c’est tout l’enjeu de l’invention du mot par Kapek dans sa pièce *RUR* (et que reprend fidèlement Fritz Lang dans *Metropolis*). Or ce seul critère pose en soi un problème du point de vue social et culturel, car il est lié à une notion historiquement changeante : le travail humain. C’est ce qui est au cœur de la réflexion du dernier ouvrage d’Antonio Casili, *En attendant les robots*, où celui-ci insiste plutôt sur les enjeux économiques, commerciaux des robots qui sont là pour accomplir des tâches à la place des êtres humains, qui, donc, potentiellement peuvent les remplacer, voire les rendre inutiles. Ce qui est faux dans la réalité actuelle (on a besoin d’êtres humains pour faire marcher les robots). Antonio Casili a montré à quel point nous sommes fascinés par l’illusion de l’automatisation intelligente (*digital labour*), par l’apprentissage automatique (*deep learning*) et par la robotique mobile, alors même que nous sommes pris au contraire dans les pièges du « travail du consommateur » (ce sont les usagers, les consommateurs qui « font fonctionner les machines » (Casili 2019 : 45). Il s’agit donc moins en réalité d’un « remplacement du travail » que de « son déplacement, à savoir une délégation d’un nombre croissant de tâches productives à des non-travailleurs (ou à des travailleurs non rémunérés et non reconnus comme tels »). Dès lors, notre approche du robot s’infléchit légèrement : ce n’est pas un hasard si certaines caractéristiques leur sont refusées : pas d’insubordination, pas d’indépendance, pas de faiblesse, pas d’émotion, pas de peur non plus. Les robots n’ont donc pas été conçus comme des doubles de nous-mêmes. Apparaît même une contradiction intrinsèque aux robots : on veut à la fois qu’ils soient autonomes et qu’ils ne le soient pas. On souhaiterait qu’ils soient à la fois comme nous tout en restant différents. Ce n’est pas un hasard si l’un des motifs récurrents des œuvres robotiques dans la culture moderne et contemporaine va porter sur leur révolte. Les robots nous fascinent mais ils continuent à nous faire peur.

En raison des progrès des neurosciences et sciences cognitives, on assiste actuellement à une forte remise en cause de l’analogie entre cerveau et machine, notamment de l’idée que la pensée puisse être totalement indépendante de son support biologique se trouve de plus en plus contestée. C’est même l’inverse : on s’intéresse de plus en plus à la neurobiologie de la rationalité – avec cette idée que la rationalité ne peut jamais être totale dans les comportements humains ; elle est limitée par différents facteurs comme l’équilibre entre exploration et exploitation, l’imprédictibilité (parce qu’il perdure toujours une part d’aléatoire dans les décisions humaines). On se demande actuellement si cela ne résulterait pas de contraintes évolutives de l’être humain. On prend ainsi en compte la théorie de l’évolution concernant l’adaptation de l’espèce, et notamment l’adaptation qui se produit en fonction du

couple : bénéfice/coût énergétique. C'est ce qu'on appelle une phylogenèse de la décision : celle-ci se définit par combinaison de deux facteurs : un processus qui décide (acteur), un système d'évaluation (critique) qui conserve la mémoire des conséquences des choix successifs (ce qu'on appelle un système d'apprentissage par renforcement). Or parfois le système critique envoie des informations erronées au système acteur. Selon Peirce, quand nous avons le choix entre deux décisions, on ne prend pas toujours la plus optimale. Il s'agit des limites mêmes de la rationalité qui seraient intrinsèques aux propriétés du système nerveux. En outre, le corps joue un rôle dans nos réflexions et dans nos prises de décision (ce qui est évacué avec les robots), notamment parce que la mémoire naît de la relation entre le corps et l'image du corps (donc du cerveau). Or justement l'un des problèmes que posent les androïdes (dans les représentations filmiques et dans les séries) est qu'ils n'ont pas de mémoire (ou tout au moins pas de mémoire sensible), tout l'enjeu va consister à pallier ce manque : soit en leur implantant de faux souvenirs soit en rendant caduque et non avenue la mémoire elle-même.

Jacques Paillard a précisément comparé les modes de fonctionnement de l'ordinateur et du cerveau (Paillard 1999 : 26-37). Il ne s'agit pas de vouloir comprendre le cerveau pour mieux l'imiter dans une machine mais d'observer comment ces nouvelles machines fonctionnent pour mieux comprendre le cerveau. Pour lui, le contraste entre l'ordinateur et le cerveau est réel, alors même qu'on est enclin à penser qu'il y aurait une analogie de fonctionnement entre le cerveau (les rhizomes que composent ses neurones par exemple) et un ordinateur. Il recense quatre différences : par leur mode de fabrication (le cerveau est biologique et génétique) ; en raison d'une différence d'échelle dimensionnelle : le cerveau et l'ordinateur ne relèvent pas de la même échelle (même si l'on prend en compte les nanotechnologies) ; le neurone fonctionne comme un ultra-microprocesseur, il est autonome, c'est un système vivant, il est doté d'une véritable reconnaissance de formes (ce qui pose des problèmes à la programmation par ordinateur).

En outre, que ce soient les synapses (interfaces entre les neurones) et les circuits entre neurones, la circulation des informations entre les neurones et les synapses est électrique, puis chimique, puis électrique. Enfin, la quatrième différence est d'ordre quantitatif : le nombre des synapses et des neurones fait que le cerveau dispose d'une capacité de traitement incommensurable à celle des ordinateurs les plus complexes. Si les ordinateurs peuvent dépasser les performances du cerveau, ils sont néanmoins encore inaptes à fournir un modèle biologiquement plausible des mécanismes qui sont à la base du fonctionnement cérébral (notamment sa souplesse adaptative et sa surprenante fiabilité). Mais plus encore, le cerveau se construit lui-même et il ne cesse de se modifier, c'est ce qu'on appelle sa plasticité intrinsèque et que François Jacob a appelé « la logique du vivant » (Jacob 1970). Or on ne sait pas totalement reproduire sa logique opératoire et l'efficacité de ses performances, il existe notamment dans le cerveau une dimension historique individualisée et une mémoire des contextes émotionnels. On distingue dans le cerveau des circuits câblés (les circuits câblés ont la capacité de catégoriser, d'évaluer, de raisonner, de décider) et des circuits humides (les circuits humides ou flottants intègrent des données en fonction du contexte émotionnel ou même postural). Dès lors, c'est surtout la

perspective inverse qui semble opérante : non pas le cerveau comme objet de reproduction robotique, mais c’est l’étude des ordinateurs (que ce soient l’étude des options connexionnistes des ordinateurs ou l’étude de ses options computationnelles) qui nous livre des éléments sur la façon dont se font les connexions dans le cerveau. Celles-ci seraient régies selon deux options. Dans l’option computationnaliste, on regarde dans l’ordinateur ses capacités combinatoires et ses fonctions calculables (elles se rapprochent de l’intelligence artificielle), et les options connexionnistes (plus récentes) : on regarde le fonctionnement de réseaux de l’ordinateur (notamment comment fonctionnent les réseaux de nœuds interconnectés), et l’on se rend compte qu’ils sont sources d’information pour la neurobiologie.

Revenons sur l’un des problèmes que pose le robot dans ses différentes représentations actuelles : son anthropomorphisation. La plupart des questionnements et des angoisses qu’il suscite tiennent à sa forme humanoïde (forme que l’être humain a précisément choisi de lui donner). Il est un fait qu’on cherche à reproduire des apparences humaines (surtout en Asie), ce qu’on appelle les robots androïdes. Ce principe d’imitation n’est pas sans multiples effets : d’abord, on a tendance à projeter une intentionnalité aux robots, alors que ce qu’ils accomplissent correspond à des boucles d’asservissement (des tâches et des objectifs prédéterminés). Ensuite, le robot ne comprend pas le sens de ce qu’il est, ni la portée de ses actions (tout montre en revanche que la conception d’un robot comme Pepper a été faite pour produire cette confusion). Pepper est un *robot humanoïde* qui a été conçu par Aldebaran (une entreprise fabriquant des robots) pour donner l’illusion de se comporter comme un être humain, ce qui est tout particulièrement évident lorsqu’il utilise une formule comme « Je suis heureux de partager avec toi cette musique que j’apprécie ». Dire « je suis heureux », c’est donner à croire que Pepper est capable de sentiment ; de même lorsqu’il dit une « musique que j’apprécie » : cela donne à croire qu’il est capable d’émotion, et même qu’il posséderait un goût esthétique. Si l’on examine encore d’un peu plus près Pepper, c’est un robot 1m20 (donc la hauteur d’un enfant) qui parle d’elle au féminin. Les informations perceptives dont elle dispose lui proviennent de sa caméra 3D, « elle » voit la forme du corps et en tire des informations, elle adapte son comportement à ce qu’il perçoit (ce qui apparaît pour l’usager comme le signe, sinon la preuve, de son intelligence). Mais surtout, si on lui pose des questions, on voit qu’elle dit qu’elle a des émotions positives. Quand on lui demande si elle a déjà pleuré, elle répond oui mais elle est incapable de dire quand. Il est intéressant pour nous d’examiner notre réaction d’êtres humains par rapport à ces robots. L’enjeu n’est donc pas de savoir si elle ressent des émotions (ce qui est pourtant très exactement en question dans de nombreux films), mais de voir ce que produit sur nous le fait qu’elle a recours à toute une phraséologie émotionnelle et sentimentale (« Je suis contente », « ça me fait plaisir », « je suis chatouilleuse »), toutes ces phrases sont là pour susciter une relation empathique (voire sympathique, avec toute l’ambivalence que cela présuppose puisque la sympathie, comme le rappelle Michel Foucault dans *Les mots et les choses*, produit du semblable, elle reconduit le sujet au même, donc cela nous rapproche de l’autre).

Un autre enjeu important pour nous porte sur la relation homme-robot. L’enjeu n’est pas seulement existentiel (théorique) ou même esthétique, il fait partie

actuellement des questionnements sociaux forts : à savoir dans quelle mesure le robot compagnon peut constituer un assistant à l'autonomie des personnes (il y a une forte demande en gérontologie). Le robot est non seulement une présence (il possède une fonctionnalité conversationnelle, il est capable d'aborder plusieurs sujets de conversations ; il possède également une fonction divertissante : il est capable de jouer aux cartes, de se promener dans l'appartement sous la forme d'applications que l'on peut télécharger), mais il peut aussi rappeler aux personnes âgées perdant la tête qu'elles doivent boire, prendre leurs médicaments. Un autre exemple fameux tient dans ce robot qui a battu un joueur au jeu de go (le robot avait été conçu et formé par quatre joueurs professionnels aguerris selon le principe de l'apprentissage par renforcement (*deep learning*)), il se caractérise également par un principe de récompense positive.

Actuellement, un certain nombre de questions restent encore en suspens. On se demande ainsi si un robot pourrait se représenter son environnement comme une entité cohérente ? Se distinguer de son environnement ? Avoir une intentionnalité, une volonté, prendre des initiatives ? Être capable d'autoévaluation ? Comprendre les conséquences de ses actions ? Être capable de raisonner sur lui-même ? À cet égard, rappelons une donnée importante : un robot qui ouvre la porte ne sait pas qu'il a ouvert cette porte. Toute la difficulté tient à la compréhension de l'environnement : l'interaction avec l'autre exige une capacité de raisonner sur l'autre, d'avoir un modèle de l'autre. Or on acquiert cela en agissant et *en regardant* les autres agir. Une autre difficulté épistémologique/éthique tient à la question de l'autonomie, notamment les véhicules autonomes ou les armes autonomes. Toute la question consiste à évaluer leurs capacités d'anticipation, et par là leurs responsabilités. Le robot pose des questions éthiques sur l'identité humaine. Certains savants considèrent par exemple que le robot est une personne et donc qu'on peut parler de personne-robot, ils parlent aussi de « robot-humanité », ce qui nous interroge sur la nature des liens affectifs et du statut social que l'on est prêt à accorder aux robots. En effet, certains scientifiques proposent de créer un droit de la personnalité de l'Intelligence Artificielle, ils postulent ainsi qu'il existe une « personnalité robot ». Cette réflexion prolonge le débat autour de ce qu'on a appelé « l'homme augmenté » (Besnier 2009), au premier sens du terme augmenté par des prothèses (on implante des prothèses aux humains). Le robot est la représentation d'un des fantasmes actuels : non pas seulement recréer l'humain mais en mieux. Il représente notre fantasme d'un *être humain augmenté*, mais il répond aussi à un fantasme démiurgique (se mettre à la place d'un dieu créateur). Il est une mise en abyme de notre désir de perfection. Par rapport à l'être humain qui est imparfait et perfectible (il ne se maîtrise pas totalement, il n'a pas tout pouvoir sur les autres et le monde, il est la proie de son inconscient), le robot constitue au contraire une (re)création parfaite. Mais il est également le reflet d'angoisses contemporaines (et notamment de nos angoisses présentes concernant notre futur). Le robot interroge aussi l'adaptabilité de nos valeurs aux réalités créées par nos technologies : c'est peut-être cela qui est en jeu avec l'homme augmenté (plus que notre disparition éventuelle à cause des robots).

Parler du robot comme d'une personne pose un problème par rapport à la notion d'espèce. On se souvient que Linné a classé le vivant. Or le robot ne l'est pas

(vivant), donc ce n’est pas une nouvelle espèce au sens strict du terme, à l’instar de l’étymologie d’*homo sapiens* (*homo* : le genre et *sapiens* : l’espèce). Un autre problème qu’on a vu émerger tient à celui des limites que l’être humain doit se fixer dans son comportement vis-à-vis du robot. On se souvient du robot-chat mis au point par Microsoft, robot qui a dû être retiré rapidement (car en une journée, il tenait des propos racistes, antisémites, puérils). En effet, les chercheurs avaient programmé la personnalité d’une adolescente. Ce robot-chat avait effectivement une personnalité, et il était évolutif, mais ce qui a posé problème est qu’il était un robot non-supervisé avec un modèle évolutionniste. Ce qui lui est arrivé est un exemple de ce que deviendrait un robot seul (non encadré, non délimité) sur Twitter par exemple. D’où la nécessité de mettre en place un cadre juridique.

Si le mythe d’une humanité dépassée par sa propre technique existe depuis longtemps, il véhicule actuellement d’autres idéologies, et notamment une fascination pour un futur dominé par les machines. Ce n’est pas un hasard s’il existe tout un mouvement critique appelé la cyborgologie (ou anthropologie cyborg). Ce nouveau champ intellectuel possède quelques années déjà, il est apparu aux États-Unis après la publication en 1985 de l’œuvre fondatrice de Donna Haraway : *A Manifesto for Cyborgs*. La cyborgologie se fonde sur le principe que le corps postmoderne est un amalgame de processus biologiques et technologiques. Il serait vain de diaboliser la technologie (ce serait même anachronique). En outre, les nouvelles technologies corporelles (prothèses, manipulation génétique, stéroïdes, chirurgie plastique...) remettent en cause l’opposition tranchée entre naturel et artificiel. Pour Haraway, le cyborg est un corps reconstruit de manière technologique, comme une possibilité prometteuse permettant de multiples incarnations polymorphes. Le cyborg transcende les frontières, la *perte* d’humanité n’est pas perçue comme un problème ou un dilemme mais un projet épistémologique. Cette fascination est liée à l’explosion technique que l’on connaît actuellement (dans les domaines : sciences de la vie, informatique, génie génétique, biotechnologie, imagerie médicale, les nanosciences...). L’essor du robot tient au complexe biologie-informatique, le robot en est devenu l’utopie concrète, de sorte que la technologie est en train de prendre la place des grandes idéologies qui sont actuellement en faillite, comme si les récits anticipateurs de science-fiction se trouvaient concrétisés, comme si l’avenir ne s’imaginait plus mais se fabriquait.

Tout ceci fait terriblement écho aux tendances de notre époque : le construit plutôt que le vivant ; le robot et le clone plutôt que l’humain ; le virtuel plutôt que le réel. Tout se passe comme si le désir d’être machine s’était emparé de la culture occidentale. On notera également un double mouvement symétrique : d’un côté, les machines convergent vers l’humain (on les anthropomorphise) ; de l’autre l’homme veut ressembler au modèle de la machine, ce qu’on repère tout particulièrement dans le fait que l’homme a tendance à se réduire à la notion d’instrument, comme si la fascination pour la technique avait pour effet que le but (idéal) de l’homme soit d’être un instrument. On peut se demander si les angoisses et les inquiétudes du monde actuel (conflits, attentats, réchauffement climatique...) n’ont pas pour effet de rendre le robot encore plus désirable pour l’être humain (en tant qu’il est rassurant). Le robot offre l’illusion d’une fausse sécurité (en tant qu’image d’une perfection et d’une maîtrise de la technique et de contrôle). Perfection, maîtrise, contrôle,

technique : on tient là les différentes caractéristiques qui rendent le robot désirable et rassurant, ce qui présuppose que les sociétés modernes sont trop complexes pour être gouvernées par les hommes et que seules les machines vont être capables de les diriger. On peut se demander si ce n'est pas une nouvelle forme de *dictature machinique* : l'homme asservi à ses propres dispositifs techniques et technologiques ?

Dès lors, au-delà de l'acceptabilité sociétale, quels sont les enjeux éthiques que pose l'insertion des robots dans notre vie ? Il existe en effet des enjeux éthiques de la robotique sociale, on parle ainsi d'éthique robotique. L'éthique renvoie à la manière dont les êtres interagissent avec leur milieu, ce qui présuppose que le robot est à même de jouer un rôle dans notre quotidien autant qu'il a un impact sur l'image que nous nous faisons de nous-mêmes. De nos jours, le non-humain revendique une dignité ontologique que l'anthropocentrisme lui refusait. L'éthique reflète cette revendication : le bien-vivre des humains est lié à une pacification avec le non-humain (machines, animaux), tout ce qui nous organise et stabilise notre relation avec les autres. Cela fait écho au concept de « cosmopolitique » défini entre autres par Isabelle Stengers comme « une institution publiquement réglée des liens entre humains et non-humains » (Stengers 2003 : 45), autrement dit une écologie politique comme relation multidimensionnelle et complexe. C'est ici que le rôle des films ou des fictions robotiques prend toute son importance, des films comme *AI*, *Blade runner*, *HER*, ou encore *I, robot* sont des films qui essaient de prévoir l'impact comportemental et éthique des robots, de l'intelligence artificielle et des systèmes d'exploitation dans nos vies humaines. C'est même tout l'enjeu des sciences humaines : poser ces questions, exposer les dilemmes, envisager tous les possibles dysfonctionnements. Ce qui montre bien que le robot nous apprend autant des choses au sujet de qui nous sommes que de notre façon de vivre ensemble (et même plus précisément encore des formes que va prendre ce vivre-ensemble). C'est même devenu un enjeu actuel essentiel des sciences humaines : aborder les problèmes éthiques que pose l'intégration des robots, l'IA, la cobotique.

Ce n'est pas un hasard si l'on cite toujours les lois d'Asimov (comme fondateur éthique du comportement du robot). Isaac Asimov est un savant d'origine russe (1920-1992) émigré aux États-Unis, il a fait un doctorat de biochimie puis il s'est mis à écrire. Il possède une approche militante des robots, il ne se contente pas d'inventer des fictions. Après sa rencontre avec John Campbell, il se lance dans deux cycles qui vont occuper une place centrale dans son œuvre : *Le Cycle des robots* et *Fondation*. Ces lois sont citées pour la première fois dans une nouvelle d'Asimov intitulé *Runaround* et publiée entre 1940 et 1950 : 1) première loi : un robot ne peut nuire à un être humain ni laisser sans assistance un être humain en danger ; 2) deuxième loi : un robot doit obéir aux ordres qui lui sont donnés par les humains sauf quand ces ordres sont incompatibles avec la première loi ; 3) troisième loi : un robot doit protéger sa propre existence tant que cette protection n'est pas incompatible avec les deux autres lois.

Ces lois sont gravées dans les cerveaux des robots, mais elles ont assez d'ambiguïtés pour être contournées (comme le montrent *Les Robots de l'aube*). En 1985, Isaac Asimov ajoute une loi zéro qui a pour objectif de placer l'intérêt de l'humanité au-dessus de celle de l'individu. Or on considère qu'il est l'inventeur de 3

concepts : la psycho-histoire, la robotique et les cerveaux positroniques (ce qui donne aux robots accès à l’intelligence et à la conscience) dont le comportement est contrôlé par les trois lois de la robotique définies par Asimov. Si les robots ne sont pas des monstres systématiquement en révolte contre leurs créateurs pour Asimov, celui-ci a néanmoins examiné dans ses différents récits les rapports conflictuels entre l’homme et la machine (*I, Robot* est directement inspiré des nouvelles d’Asimov).

L’enjeu technologique actuel consiste à limiter de l’intérieur la marge de manœuvre des robots en introduisant un module éthique au sein de leur système de commande, ce qui est une façon de limiter leur possibilité d’action. Cela est loin d’être anodin puisque cela révèle l’ambiguïté fondamentale de l’éthique robotique (et qui tient notamment à l’emploi du terme autonome). Dès lors, l’autonomie est une notion qui mérite une réflexion approfondie. D’un côté, on doit se demander quel comportement humain est opérant face à l’automatisation des robots et par rapport à ce qu’on appelle délégation de conduite ? Il convient de réfléchir au caractère (nécessité) réellement autonome du robot et de se demander s’il y a des hiérarchies entre ces valeurs (parfois il peut y avoir des conflits de valeurs, donc il faut des arbitrages, ce qui relève de l’humain, non du robot). En un mot, ce que nous appelons autonome ne doit pas l’être totalement. Cela révèle une certaine confusion quant au rôle de l’autonomie en philosophie morale : à la fois on suggère que l’action des robots pourrait poser des problèmes moraux, et d’un autre côté on présuppose qu’ils ne sont pas de véritables agents moraux (puisque la définition d’un agent moral est un agent qui agit comme il le doit, donc qui par définition pourrait agir *autrement*).

Or l’éthique robotique refuse précisément aux robots cette autonomie-là (ce qui montre aussi qu’il ne s’agit pas réellement d’éthique mais plutôt de codes de lois) où tout l’enjeu consiste à remplacer la capacité de choisir par l’impossibilité de faire autrement. Précisons bien que ce ne sont pas les robots qui ont des problèmes éthiques, mais bien nous qui en avons vis-à-vis d’eux. L’un des enjeux consiste à enseigner aux robots la différence entre le bien et le mal, à élaborer donc des règles morales pour des agents artificiels par le couplage d’une dimension technique avec une dimension morale. Cela signifie qu’il faut développer des robots programmés pour pouvoir intégrer des règles morales, autrement dit leur faire intégrer des règles morales qui limitent leurs possibilités d’action. Il ne s’agit nullement d’inventer des règles éthiques propres aux robots, mais plutôt d’être sûrs qu’ils vont agir conformément à la morale (et non pas agir moralement). Ces questionnements trouvent un écho amplifié dans le domaine de la guerre et des conflits, c’est sans doute là que les robots sont les plus fascinants/terrifiants.

Ce n’est pas un hasard si l’armée est tentée par l’intégration et l’utilisation de robots : parce qu’ils sont obéissants et disciplinés, parce qu’ils appliquent les ordres qu’on leur donne, parce qu’ils ne font pas d’erreur d’inattention, ils n’ont pas peur, ils sont plus résistants. Autrement dit, ce sont des super-soldats, ou pour le dire autrement des soldats idéaux (car ils poussent la discipline jusqu’à la perfection). L’utilisation des robots (qu’on appelle soldats robotiques autonomes) permet à l’armée de consolider sa stratégie militaire selon un double principe : d’une part, limiter la capacité de décider et de prendre des décisions à quelques-uns ; de l’autre s’assurer que les décisions seront exécutées sans états d’âme ni atermoiements. Le

robot associe différents facteurs : 1) un facteur technique (les robots sont plus précis, plus rapides) ; 2) un facteur économique (un robot coûte moins cher qu'un soldat ou un pilote) ; 3) un facteur humain et politique (la mort d'un être humain est problématique, alors qu'un robot ne pose pas cette question) ; 4) un facteur militaire (actuellement dans la course aux armements, on confère une place importante aux robots. La présence du robot dans un contexte de conflit guerrier se pose de manière accrue parce qu'elle n'est rien moins qu'une façon de déléguer aux robots la responsabilité de tuer. Or un autre effet en découle : la prise de décision (qui revient à l'être humain) se trouve concentrée sur lui. Ce qui produit un effet de concentration des décisions qui revient à quelques agents.

Le cinéma s'est emparé à sa façon de ces problématiques. C'est ce que montre *Robocop* de Paul Verhoeven et notamment à la scène liminaire des robots-cops (policiers) qui tuent un enfant (qui était armé). Les robots appliquent les ordres de manière automatique/systematique. Le rôle du héros Murphy s'inscrit dans ce contexte, il s'agit d'un humain reconstruit par des greffes d'organes artificiels, il est donc ce qu'on appelle un cyborg, ce qui est une façon de montrer que la frontière entre homo sapiens et création artificielle est ténue. En effet, d'un côté Murphy a gardé des caractéristiques humaines (la mémoire, les souvenirs, ses émotions personnelles), de l'autre il est supervisé, donc contrôlé en permanence. Tout l'enjeu va consister à réfléchir sur la possibilité (ou non) de conserver un statut proprement humain en étant un robot. *Terminator* déploie ce questionnement sur une ligne un peu différente, c'est un film qui dépeint un futur où les robots sont en rébellion (l'action se passe en 2029) contre les humains. L'un d'eux, Terminator, va remonter le temps jusqu'en 1984, son objectif consiste à empêcher la naissance de John, fils de Sarah Connor, qui va mener la bataille contre les robots jusqu'à la victoire.

L'un des enjeux du film est de montrer qu'une machine n'aura jamais d'autre morale que celle qui lui est programmée. Le succès est immédiat et mondial, montrant l'intérêt mais aussi l'inquiétude que suscite la présence des robots. Un autre des films paradigmatiques des films de robots est sans nul doute *2001 Space Odyssey* (2001, *L'odyssée de l'espace*) et notamment la scène où HAL est débranché. Scène culte en ce qu'elle est filmée par Kubrick comme si c'était un humain qu'on débranchait, tout en montrant que HAL est un robot malfaisant et qui peut prendre des initiatives. HAL est un ordinateur chargé d'une mission dont il est le seul à connaître les enjeux, cet ordinateur va considérer que les humains sont en conflit avec sa mission, il décide de les éliminer. L'un des membres de l'équipage réussit à déconnecter petit à petit les circuits vitaux de l'ordinateur, c'est une des plus belles scènes du genre (au point de vue émotionnel), mais elle est également emblématique de l'une des croyances de l'IA : on voit l'ordinateur régresser et retourner en enfance vers les phases les plus significatives de son apprentissage, au fur et à mesure où ses circuits sont débranchés. Dans le dernier plan de cette séquence, la machine est désormais inoffensive, elle chante une comptine qui était enfouie dans sa conscience artificielle.

Dès lors, on se rend compte qu'il convient d'inverser la question sans cesse posée au cinéma (les robots seront-ils un jour à même d'éprouver des sentiments ?). La question des sentiments renvoie à un questionnement beaucoup plus profond qu'il n'y

paraît. Elle correspond à un questionnement de la société sur ce qu’on appelle la robotique sociale. Il ne s’agit donc pas de savoir quels seraient les sentiments des robots que de nous interroger sur les sentiments que nous pourrions éprouver vis-à-vis d’eux. Ce qui est en question, c’est bien notre capacité (propension/désir) à nous attacher à eux. Le robot est le reflet et même le symptôme de notre socialité actuelle (de la façon dont nous la concevons, et de la façon dont nous sommes en train de la faire évoluer). Les robots nous interrogent sur nos relations sociales, ils nous permettent de nous poser des questions sur la *nature même de la socialité humaine*, ou plus exactement sur les mutations de cette socialité. Toute réflexion sociale sur le robot oscille toujours entre deux pôles : d’un côté, la représentation des robots ne fait pas l’économie d’une réflexion sur l’humain (ce qui fait notre part d’humanité, ce qui nous définit comme être humain). Le robot est une représentation technologique que qui nous permet de sentir ce qui en nous *ne l’est pas* (spécifiquement humain). Ce qui est en jeu dans la représentation des robots, c’est donc bien la nature de relation qu’ils peuvent entretenir avec les humains : une relation conflictuelle ? une relation destructrice ? une relation pacifique ? une relation affective ?

Il convient d’ajouter encore deux remarques. Tout d’abord, il y a des différences culturelles évidentes et même parfois massives dans les représentations du robot, notamment entre la culture occidentale et la culture asiatique. La représentation du robot à l’Ouest est très américanisée et par là très conflictualisée, alors que la représentation du robot en Asie est beaucoup plus pacifiée, elle est davantage représentée dans une certaine normalité dans la culture asiatique (manga, drama par exemple). Le robot fait partie de l’univers de pensée en Asie, en ce qu’il est une image du futur mais qu’il s’articule étroitement avec le passé (il n’entre pas en conflit avec lui). Une deuxième remarque porte sur la question de l’attachement/affection vis-à-vis des robots est souvent latente dans les représentations filmiques et littéraires. Ce qui est interrogé : la nature de cet attachement (de quel sentiment s’agit-il au juste (à l’instar d’*AI* de Spielberg). Le robot propose par là même une réflexion sur l’ambivalence de certains sentiments comme l’antipathie, le rejet, le dégoût, l’angoisse, ou au contraire l’empathie, l’identification. Quelle est la légitimité, la normalité/conventionnalité de cet attachement ? Et dans le cas où cette relation ne serait pas représentée comme normale : de quelle pathologie notre relation au robot est-elle donc le symptôme ?

Dès lors que le robot ressemble à un être humain (d’autant plus quand il est joué par un acteur), cela augmente la propension du spectateur à s’attacher à lui. Il convient d’ajouter une autre donnée à ce raisonnement dans la mesure où le robot est aussi un emblème du progrès technologique, plus précisément sa représentation dans des films révèle la place qu’on projette de lui donner dans l’avenir. Les différents films-robots montrent qu’il est tentant de lui donner une place presque équivalente à celle d’un humain. D’autre part, ils mettent en évidence que ce qui se joue actuellement, c’est est la technologisation des corps : à quel point on a envie d’intégrer en nous-mêmes des éléments robotiques ou mécaniques (se faire greffer des éléments inorganiques). Nous nous situons à une période où on cherche à tester à quel point l’être humain peut se robotiser (effectuer des choses mécaniques sans y penser). La réflexion sur le robot va de pair avec les progrès sur le clonage et l’uniformisation :

il y a un principe d'uniformité et de reproduction à l'identique dans le robot). Le fantasme sous-jacent que révèle la représentation robotique fonctionne comme un slogan mimétique : tous pareils !

Le robot révèle les questionnements existentiels actuels : qu'est-ce qui nous définit finalement comme réellement humain : notre intelligence ? notre corporéité ? notre culture ? notre relation aux autres ? notre individualité ? Mais les films-robots nous permettent également de réfléchir à la spécificité humaine non pas d'une manière abstraite mais dans ses aspects les plus concrets, parce qu'ils mettent concrètement en jeu notre peur de souffrir, notre peur de mourir, la nécessité (ou non) d'avoir des liens sociaux, d'avoir des sentiments pour ses proches, de désirer fonder une famille, de vivre en collectivité, de posséder des souvenirs/nécessité de la mémoire, d'avoir conscience du bonheur et du malheur, d'avoir conscience du bien et du mal, d'être capable de mettre fin à ses jours. Alors même que ces caractéristiques rappellent en définitive certaines différences fondamentales entre les robots et les êtres humains, on notera qu'elles ne nous empêchent pas de nous poser la question suivante : à quel point peut-on (doit-on) considérer les robots comme nous ? Faut-il ou non leur appliquer nos principes et nos règles de vie (on ne peut pas les maltraiter, les tuer...) ou peut-on leur faire faire tout ce qu'on veut (c'est la question que pose l'un des épisodes de *Black Mirror*) ? Ce qui est dérangeant dans le robot, c'est qu'il se positionne en concurrence directe avec l'être humain (et non pas avec la machine). Une dernière question mérite d'être posée : plusieurs des films cités nous montrent des robots *qui veulent devenir humains*. De quoi est-ce l'expression ou le symptôme ? D'une part, c'est une réflexion sur la notion de devenir (on ne naît pas complet/délimité, on devient quelqu'un en se dotant de caractéristiques qui nous singularisent). D'autre part, il existe des spécificités humaines qui sont enviables (et non imitables ou reproductibles par les robots) : or le robot ne peut pas se reproduire et il ne connaît pas le plaisir, il ne connaît pas l'affection, l'attachement, le souvenir, l'émotion. Tout cela nous interroge en filigrane sur des notions comme le lien du sang, la question de la filiation, le sentiment maternel ou paternel...

Le robot est une image paradoxale de la diversité, c'est même une allégorie de l'altérité. C'est le sens même du titre de cet article qui joue sur les deux sens de « like » en anglais : comme et aimer. En tant que spectateur, cela me permet de voir à quel point je suis capable d'accepter, de supporter (voire d'aimer) quelqu'un ou quelque chose de différent de moi, il suscite donc une réflexion sur l'altérité en tant qu'elle n'est pas perçue comme normalité, mais au contraire une réflexion sur l'anormalité (et donc une réflexion sur des différences qui dérangent et non pas des différences minimales). Le robot en ce sens reste un défi artistique. Il libère l'imagination et plus particulièrement une imagination scientifique. Le robot est une stimulation pour l'imaginaire, mais c'est aussi un défi qui articule esthétique et technique (les images de synthèse et les effets spéciaux sont au service des mutations des représentations de l'espèce humaine).

UNIVERSITÉ DE LIMOGES

professeure

frederique.toudoire-surlapierre@unilim.fr

BIBLIOGRAPHIE

BESNIER, Jean-Michel (2009). *Demain les posthumains. Le futur a-t-il encore besoin de nous ?*, Paris : Hachette, « Pluriel ».

BRETON, Philippe (1995). *À l’image de l’homme. Du Golem aux créatures virtuelles*, Paris : Le Seuil, « Science ouverte ».

CASILI, Antonio (2019). *En attendant les robots. Enquête sur le travail du clic*, Paris : Le Seuil, « La couleur des idées ».

HARAWAY, Donna (1985). *A Manifesto for Cyborgs : Science, Technology ; and Socialist Feminism in the 1980s*, [En ligne] <https://sites.evergreen.edu/politicalshakespeare/wp-content/uploads/sites/226/2015/12/Haraway-Cyborg-Manifesto-2.pdf>. Consulté le 12 mars 2021.

JACOB, François (1970). *La logique du vivant. Une histoire de l’hérédité*, Paris : Gallimard, « Tel ».

MALABOU, Catherine (2004), *Que faire de notre cerveau ?*, Paris : Bayard.

MALABOU, Catherine (2017). *Métamorphoses de l’intelligence. Que faire de leur cerveau bleu ?*, Paris : PUF.

PAILLARD, Jacques (1999). « L’ordinateur et le cerveau : le contraste demeure saisissant », Vincent Bloch (dir), *Cerveaux et machines*, Paris : Hermes Science Éducation, 26-37.

STENGERS, Isabelle (2003). *Cosmopolitiques I*, Paris : La Découverte, « Poche ».