

Le robot, étranger de l'éthique.

Depuis l'apparition du robot autonome capable de prendre des « décisions », certains (C Allen et W. Wallach) sont tentés d'élaborer une éthique qui lui serait propre et de le considérer comme un agent moral. Certains (Arkin) pensent même qu'il pourrait agir de façon plus « éthique » que l'humain. En conséquence, ce dernier propose de créer des robots autonomes éthiques à vocation militaire en espérant que cela moralisera davantage les guerres. Dès lors que la position éthique que l'on concéderait au robot touche des questions de vie ou de mort, il devient nécessaire de s'assurer au préalable de ce dont nous parlons. Car avant qu'on se demande à quel type d'éthique un robot ferait référence, avant qu'on se demande si ce type d'éthique serait implémentable dans un robot, et avant qu'on se demande à quel type de responsabilité on aurait alors affaire, il est nécessaire de rappeler ce qu'est l'éthique et ce en quoi consiste un robot autonome si l'on ne veut pas manquer l'essentiel du débat : Un robot « autonome » peut-il être éthique ? L'éthique est une discipline qui fait appel à certaines capacités d'un agent en vue d'un but. Nous allons montrer que le but de l'éthique n'est pas atteignable par un robot, même « autonome », ce qui implique qu'il n'est pas un agent moral, qu'il ne saurait l'être puisqu'il est dépourvu des capacités qui le lui permettraient, que par conséquent le champ de l'éthique lui est étranger et pourquoi il serait inutile que ce dernier soit modifié pour l'intégrer. Le robot ne nous soulagera pas de la responsabilité que nous portons d'agir éthiquement.

L'éthique aristotélicienne :

Pour le comprendre, retournons d'abord à l'une des sources de l'histoire de la pensée éthique : l'éthique aristotélicienne. Nous évoquerons ensuite la position kantienne. Selon Aristote, le but de l'éthique est la vie bonne, sensée donner le bonheur. La vie bonne est considérée comme telle en fonction de la poursuite d'un but, un *télos*, qui consiste pour l'homme à être vertueux. La *phronèsis*, la prudence, la sagesse pratique, est ce qui nous permet de juger et d'agir selon un juste milieu et selon les circonstances. C'est cette disposition qui nous permet d'être vertueux. Cette définition donne un cadre pratique à l'éthique, il s'agit d'une éthique appliquée vécue au quotidien. L'éthique est donc la capacité de se comporter de la meilleure façon selon les circonstances, pas de délibérer théoriquement sur ce qui serait le Bien absolu.

En nous basant sur cette compréhension historique de l'éthique, il semble douteux qu'un robot puisse s'y frotter. En effet, si le but de l'éthique est le bonheur, autrement dit un état durable de bien-être, de satisfaction, quel serait le bonheur atteignable par un robot ? Ronald Arkin mentionne dans *Governing Lethal Behavior: Embedding Ethics in a Hybrid Deliberative/Reactive Robot Architecture* : « they can be designed without emotions that could their judgment or result in angers and frustration with ongoing battlefield events. (...) Autonomous agents need not suffer similarly ¹ ». Si donc le robot est affranchi des sentiments humains, il est affranchi du bien-être et du mal-être, le bonheur est un mot absurde pour lui. Il s'en suit que le but qui motivait le comportement éthique selon Aristote est étranger au robot.

¹ R. ARKIN, *Governing Lethal Behavior: Embedding Ethics in a Hybrid Deliberative/Reactive Robot Architecture*, Chapman & Hall/CRC, 2009, p 6.

S'il est exact que le robot a une certaine attention aux circonstances puisqu'il mesure un ensemble de variables, on peut cependant se demander s'il opère selon un jugement, autrement que par un glissement du sens de ce mot. En effet, le robot se borne à calculer arithmétiquement en fonction de mesures qu'il effectue via ses capteurs et à l'aide d'algorithmes qu'on lui a implantés. Son action se borne à adopter la réponse prévue dans son programme selon un ensemble de paramètres prédéfinis. Peut-on parler de jugement lorsque ne sont à l'œuvre que des réponses automatiques ? Un jugement suppose plus qu'une sélection machinale de moyens prédéfinis à mettre en œuvre. Le jugement nécessite une délibération réfléchie, qui évalue les pour et les contre, et un peu d'imagination pour trouver des solutions originales dans des circonstances imprévues ou inédites. La sélection machinale, elle, s'opère suivant un logiciel, autrement dit, un schéma issu des travaux des roboticiens et des scientifiques qui l'ont élaboré. En conséquence on ne saurait dire, sauf à projeter sur le robot nos conceptions anthropomorphes, qu'il est l'auteur de l'action, encore moins du jugement qui y a conduit. Le robot ressemble au mercenaire auquel ferait appel le commanditaire d'un meurtre. Cette comparaison même n'est pas suffisante, le mercenaire choisi encore lui-même son plan d'action, tandis que le robot n'est qu'un moyen, un instrument, le prolongement de l'action et de la décision humaine. On ne saurait dire qu'il est l'agent responsable d'une délibération dont le schéma lui a été imposé par des humains, si tant est d'ailleurs qu'appliquer un schéma fixe puisse s'apparenter à une délibération et à un jugement.

C'est un fait, les paramètres auxquels un système autonome est sensible dépendent des capteurs dont il est doté. Nous pouvons considérer qu'il possède un certain degré d'analyse des conditions extérieures. Mais, si l'on se base sur l'observation que le comportement moral est plus aisé pour ceux qui ont une forte capacité d'empathie, l'une des capacités essentielles pour agir moralement ferait alors cruellement défaut au robot. L'empathie est la capacité à se mettre à la place d'un autre pour comprendre ce qu'il doit ressentir en fonction des réactions que nous aurions eues dans cette même situation. Nous avons une « pré-réponse » instinctive, issue de notre capacité de mimétisme, qui nous guide vers le comportement que l'autre souhaiterait nous voir adopter, ou celui qui le soulagerait. Dès lors, une gamme de situations semble inaccessible au robot. Par exemple, comment doter un robot d'un capteur qui lui ferait comprendre pourquoi un bébé pleure ? Il se limiterait à voir s'il a mangé, est propre, n'a aucune douleur physique, ne montre pas des signes de fatigue... Comment pourrait-il comprendre qu'il a eu peur de quelque chose, ou qu'il se demande soudain où sont ses parents, ou encore qu'il ne retrouve pas son doudou, etc. ? S'il ne peut pas savoir pourquoi, il saura difficilement comment le reconforter efficacement. Quand bien même il déterminerait que le bébé doit être changé et qu'il le fasse lui-même, il ne saurait pas lui apporter l'affection qui doit accompagner ces soins pour que l'enfant s'apaise. On répondra sans doute que cet exemple ne concerne pas l'éthique. Il s'agit pourtant d'un exemple de relation d'aide. Parce qu'il est limité dans sa capacité à analyser les éléments pertinents d'une situation, limité dans sa capacité à imaginer des solutions face à l'imprévu, limité dans sa capacité à éprouver de l'empathie, mais aussi de la compassion, il semble bien incapable de poursuivre et d'acquiescer la « vertu » aristotélicienne, à savoir la capacité de bien juger selon les circonstances, qui est le propre de l'homme et de la vie bonne. Par ailleurs, « une vie bonne », n'a pas de sens pour un robot qui n'a ni sentiment ni sensation personnelle. Il n'a pas de vie, ni bonne ni mauvaise, mais seulement une durée d'exploitation.

La position kantienne d'universalité de la règle :

Certains espèrent voir le robot agir plus éthiquement que l'humain parce qu'il ne serait que « pure application » de la règle. La volonté (et la nécessité pour la programmation) de donner au robot des lois générales, automatiques, valables partout et en tout lieu, conduit à lui

donner une éthique « automatique ». On peut le rapprocher du projet kantien d'universalisation des règles de comportement. Cette position permet d'esquiver nos arguments précédents, à savoir que le robot n'est pas « éthique » parce que le bonheur lui est inconnu et parce qu'il est incapable d'obtenir la vertu, la prudence aristotélicienne. Mais cette éthique déontologique a d'autres inconvénients : l'inattention aux cas particuliers, aux circonstances, qui peuvent conduire à des actions immorales mais justifiées par le respect de la règle (comme l'interdit de mentir même pour protéger un réfugié). Si faire prévaloir la règle au dessus de toute autre considération peut conduire à des effets pervers au point de vue moral, on peut se demander s'il est raisonnable de prendre ce risque, surtout quand il est question de robots militaires.

Les promoteurs du robot éthique comme Arkin ne seront peut-être pas sensibles à cet argument puisqu'ils admettent que « an ethically-infallible machine ought not to be the goal now² », leur projet se limitant à « to design a machine that performs better than humans do on the battlefield, particularly with respect to reducing unlawful behavior or war crimes³ ». Cependant, les humains eux sont responsables de leurs actes, in fine. Ils ont l'obligation morale de désobéir à des ordres s'ils sont inadaptés aux circonstances et si leurs conséquences peuvent être néfastes. On peut se demander si, *structurellement*, des robots programmés avec une éthique générale ne produiront pas de façon certaine un certain nombre d'actions immorales, dictées par le respect de la règle.

Thomas M. Powers, dans son article « Prospect for a Kantian Machine », propose de partir des règles déontologiques kantienne pour créer un robot éthique. Penser néanmoins qu'il suffirait de se reporter à Kant pour démontrer qu'un robot pourrait agir éthiquement serait altérer sa pensée. Powers se réfère, pour sa « kantian machine » à la première formulation de l'impératif catégorique : « Act only according to that maxim whereby you can at the same time will that it should become a universal law without contradiction⁴ ». Il semble donc bien faire abstraction de la deuxième formulation : « Act in such a way that you treat humanity, whether in your own person or in the person of any other, never merely as a means to an end, but always at the same time as an end⁵ ». Manifestement, il semble difficile à un robot de traiter l'humanité dans sa propre personne comme une fin, et donc de satisfaire à l'impératif catégorique, avant même qu'on se demande de quelle manière ce comportement pourrait être traduit en langage mathématique. Considérer que le robot serait éthique au sens kantien parce qu'il satisferait à la première version de l'impératif catégorique n'est donc possible qu'au prix d'une réduction de la morale kantienne. Mais ce n'est pas la seule.

Car c'est oublier que la valeur morale de l'impératif catégorique nécessite la rectitude de l'intention (la volonté d'agir par devoir, une bonne volonté). Or, on ne peut attribuer aucune intention, ni bonne ni mauvaise, à un robot puisqu'il n'a pas de volonté propre, car il ne peut avoir de volonté puisqu'il n'a pas de liberté, nous reviendrons sur ce point. Ajoutons enfin que Kant ne justifiait le devoir moral qu'en vertu de l'existence de notre liberté, de l'immortalité de notre âme et de l'existence de Dieu. Si le robot n'a pas de liberté, qu'il n'a certainement pas d'âme, et que par conséquent il ne risque pas, si Dieu existe, d'encourir le jugement dernier, il n'a aucune raison de respecter la règle. Dès lors, même s'il avait eu une volonté propre et même s'il se conformait à la règle, il ne pourrait agir par respect de la loi morale, par bonne volonté, donc son action ne serait pas morale au sens kantien, mais ne

² P. LIN, G. BEKEY, K. ABNEY, « Autonomous Military Robotics : Risk, Ethics, and Design », rapport à destination du *US Department of Navy, Office of Naval Research*, préparé par « *Ethics + Emerging Sciences Group* » de la *California Polytechnic State University, San Luis Obispo*, 2008, p 2.

³ Idem.

⁴ KANT, *Grounding for the Metaphysics of Morals 3rd ed*, 1785, trad. by James W. Ellington, Hackett 1993, pp. 30.

⁵ Idem.

saurait être immorale non plus. D'après la philosophie kantienne, puisque le devoir moral ne saurait s'appliquer au robot, il s'en suit que le champ de l'éthique ne le concerne pas.

Minimum éthique et objections :

Après notre analyse des objections issues de deux courants opposés dans la tradition de éthique, à savoir l'éthique téléologique et l'éthique déontologique, on pourrait arguer qu'il suffirait de créer une nouvelle éthique pour qu'elle puisse s'adapter au robot. Autrement dit, de modifier la définition de l'éthique et de l'agent moral pour qu'elle s'applique à un robot. Ce serait oublier la confusion et les abus de langages concernant l'étendue des capacités du robot. Si revenir aux racines de deux grands courants de pensée ne semblait pas encore suffisant pour convaincre qu'un robot ne saurait être qualifié d'éthique, rappelons alors sans quel minima on ne saurait parler d'éthique.

Si nous tournons la question différemment en nous demandant comment l'action d'un robot pourrait être morale, qu'on se rappelle que la première condition pour déterminer si une action est morale ou non, est la responsabilité (ou non) de l'agent qui l'a posé. Or, pour qu'il puisse y avoir responsabilité, il faut que l'agent ait été libre de son choix.

Qu'est-ce que la liberté de choix ? Il s'agit d'une capacité plus vaste que de choisir comme les enfants, entre ce qui se cache dans la main droite, et ce qui se cache dans la main gauche. Il est rare que les choix que nous avons à opérer soient aussi simple, souvent le plus difficile consiste à élaborer et répertorier les différentes options qui s'offrent réellement à nous, ou les conditions que nous devons mettre en place pour pouvoir choisir certaines voies (étudier ses examens pour pouvoir choisir les bonnes filières après par exemple). Dans le domaine éthique, ces choix s'avèrent souvent beaucoup plus compliqués, parfois même il semble n'y avoir aucune solution pleinement satisfaisante moralement. Ce sont les dilemmes moraux, où l'individu, confronté à un conflit de valeurs, est forcé de choisir. Il sera alors responsable de son choix. Comment un robot, confronté à ce type de dilemme, pourrait le trancher autrement qu'en laissant faire le hasard ? Une telle solution ne serait pas un choix moral.

Le robot est-il libre ? La liberté consiste en autre chose qu'avoir plusieurs possibilités de réponses imposées (ou apprises) sachant que même le « choix » entre différentes réponses est déterminé par une règle qu'il n'a pas choisie (par exemple faire un « choix » qui maximise tel type de données). Avant que nous ne propositions une réponse concernant la liberté du robot, examinons les arguments contraires : On affirmera que l'homme, lui, est entièrement déterminé par sa nature, son ADN, son instinct, son environnement psycho-social... qu'il n'a qu'une illusion de liberté, ou bien, parce qu'il se soumet à un ensemble de lois et de règlements, il n'est pas complètement libre non plus. Au vu de toutes ces contraintes et ce déterminisme qui pèsent sur l'homme, on estimera peut-être que la programmation et la construction du robot sont leurs équivalents. Si malgré tout on estime que l'homme est libre, alors le robot le serait tout autant ? Un raisonnement par l'absurde montrera que non.

L'homme peut s'arracher dans une certaine mesure à son déterminisme naturel : il n'a pas d'aile mais n'a qu'à prendre l'avion pour voler et une fusée s'il veut découvrir l'apesanteur. L'homme est contraint par la loi, mais il ne s'y soumet pas toujours, voilà pourquoi il existe des prisons. Le robot, lui, peut-il échapper à sa programmation ? Peut-il choisir de désobéir ? Et s'il le peut, gardera-t-on un tel robot en service ? Paradoxalement, si un robot pouvait choisir d'échapper à sa programmation, il pourrait devenir un agent moral mais au comportement immoral. Il faudrait alors s'en séparer. Imaginons qu'il soit libre et ait décidé tout de même de respecter son programme, nous n'aurions pas pu distinguer sa liberté et lui attribuer le statut d'agent moral, puisqu'il n'aurait fait qu'appliquer son programme. Si le robot ne saurait être libre, ni reconnu comme tel, il ne peut être considéré comme un agent

moral. Quand bien même il aurait pu être libre, et reconnu comme tel, nous avons vu qu'il aurait été immoral.

L'une des raisons pour lesquelles on suppose que le robot pourrait être un agent moral, est sa capacité d'autonomie. Il ne s'agit pas de considérer que n'importe quel robot pourrait être un agent moral. Cette réflexion ne concerne que les robots appelés « autonomes ». Certains considéreront qu'un robot est autonome s'il n'est pas dépendant d'un fil électrique. Ce n'est pas ce genre d'autonomie qui nous intéressera non plus. Il s'agit de robots dotés de sensor-processor-actuator. Ils peuvent prendre des mesures, calculer la réponse à donner, agir sur l'environnement en retour, selon les capacités d'action dont ils sont dotés. Un robot est autonome lorsqu'il n'est pas dépendant d'un homme pour fonctionner une fois qu'il est mis en marche. Par un glissement de sens, on en vient à confondre l'autonomie que possède le robot et celle que possède un agent moral. Autrement dit, on confond l'autonomie technique du robot avec la liberté comme autonomie telle qu'on la trouve chez Kant. Mais le robot n'est pas « autonome » au sens étymologique, puisqu'il ne se donne pas à lui-même sa propre loi. Comme nous l'avons exposé plus haut, il est dépendant de sa programmation, il ne la choisit pas, il ne peut être libre. Son autonomie se résume à être capable de faire tourner un programme informatique sans intervention d'un utilisateur.

Certains voudraient voir le robot adopter une morale d'esclave. Cette idée mérite d'être analysée. Ils entendent sans doute par les mots « morale d'esclave » le fait de se conformer à une loi morale sans motivation, sans espoir de bonheur ni de récompense. Ce serait obéir sans même vouloir agir par devoir au sens kantien, car pour les raisons que nous avons exposées, cette obligation de respect de la loi morale ne s'impose pas au robot. La morale d'esclave du robot serait celle de l'obéissance à une loi morale dictée par d'autres. Dans ce cas, le robot est perçu comme l'esclave de sa propre programmation éthique. Mais peut-on dire qu'il s'agirait là d'un agent moral ? Une fois de plus, puisqu'il n'est pas libre d'y adhérer ou non, il ne porte pas la responsabilité de ce qu'il provoque comme événement, bon ou mauvais. S'il n'est pas responsable de son acte, ce n'est pas un acte moral, même s'il peut en être la cause matérielle.

On ne peut pas même considérer que le robot se soumet à sa programmation comme l'esclave antique était soumis à la loi de son maître. Dans l'antiquité, certains hommes, nés libres, pouvaient décider de se faire esclaves pour payer leurs dettes. Quant aux autres, ils pouvaient racheter leur liberté ou être affranchis. La perte ou l'absence de liberté pouvait donc être temporaire, ce ne saurait être le cas d'un robot. Lorsqu'un esclave désobéissait, il pouvait être puni pour lui enlever l'envie de recommencer. Nous connaissons tous l'histoire de Spartacus, son statut d'esclave ne l'a pas empêché de mener une rébellion. Ceci montre que l'esclave était capable de se rebeller contre la loi, ce qu'un robot ne peut pas non plus faire contre sa programmation. Si l'esclave antique ne disposait pas de sa liberté civile, il disposait au moins de sa liberté intérieure, celle d'obéir ou non, de prendre le risque d'encourir la punition. En ce sens, bien qu'il soit esclave, il était responsable de ses actes, ce que ne saurait être le robot.

Précisons ce point. Le robot n'a pas la possibilité d'avoir une analyse autoréflexive de sa programmation, entendue comme l'ensemble des ordres auxquels il obéit, ni celle de se rebeller contre celle-ci. Un humain, lorsqu'il est placé dans une position d'obéir à des ordres, reste toujours responsable in fine des conséquences des actes qu'il a posés, car sa conscience aurait dû le pousser à désobéir à un ordre immoral. Un robot n'a pas la capacité de déterminer si les actions qu'il va accomplir en suivant sa programmation sont morales ou non. S'il y a une erreur dans sa programmation, le robot n'est pas responsable des actions immorales qui en ont découlé. Mais il n'est pas non plus responsable des actes moralement acceptables qu'il ferait. Dans les deux cas, il ne fait qu'appliquer sa programmation. Nous devons constater que le robot est complètement soumis à sa programmation, en dehors des bugs et de

l'impossibilité de prévoir complètement son comportement en raison de sa complexité. En conséquence, considérer le robot comme un agent moral va à l'encontre de l'idée de la responsabilité. La responsabilité est la capacité de répondre de ses actes et de ses prises de positions, d'en endosser la paternité, de reconnaître la participation prise dans une action. Dire que quelque chose peut être de la responsabilité d'un robot autonome, surtout quand les conséquences sont néfastes, revient à donner une excuse similaire à celle du collégien dont le chien a mangé le devoir. La seule responsabilité possible pour un robot, nous l'avons déjà dit, est d'être cause matérielle.

Lorsque certains comme Arkin affirment que le robot peut être éthique, voire plus éthique que l'humain, ils l'analysent selon son respect des lois. Lin, Bekey et Abbey écrivent ceci à propos des robots militaires: « the relationship of morality to legality –a minefield for ethics- is likewise largely avoided; the Laws of War and Rules of Engagement make clear what actions are legal and illegal for robots, and for military situations, that can serve as a reasonable approximation to the moral-immoral distinction. »⁶. Ils reconnaissent donc que le caractère éthique du robot militaire se limiterait à son respect du droit. Dans ce cas, il est incorrect de l'appeler « robot éthique », car il est seulement un « robot légal ». Rappelons-le, le droit et l'éthique ne sont pas identiques, appeler « éthique » ces robots est inapproprié. Obéir au droit n'est pas suffisant pour être dit « éthique », car les lois ne peuvent jamais tout prescrire du champ moral. Rappelons aussi que les robots militaires sont une exception, car dans le champ de la moralité, il est normalement immoral de tuer des êtres humains. On pourrait s'appuyer sur cet argument pour dire que ce type de robot reste immoral comparé à la morale commune.

Cette position contient également un problème intrinsèque à la technique, surtout si l'on souhaitait l'élargir aux robots non-militaires. Elle nécessite que le droit soit un objet assez précis, complet, cohérent, autrement dit un objet logique non-contradictoire. C'est l'une des conditions pour qu'il puisse être implémenté dans un programme informatique. Mais l'existence des avocats et des procès démontre s'il était besoin, que le droit n'est jamais ni parfaitement complet, ni parfaitement cohérent. Les avocats remportent des procès en démontrant que tel ou tel article de loi n'était pas clair, ou que tel article était incompatible avec tel autre dans ce cas-là, ce qui nécessite le jugement d'une cour de justice. Le droit ne peut pas tout prescrire. De plus, en accumulant les précisions législatives, puisque le droit évolue, un article de loi risque d'être en contradiction avec un autre, égaré dans la masse des règlements de toute sorte. On retrouve sans doute là une application possible du théorème d'incomplétude de Gödel.

Lorsque nous avons évoqué la nécessité de l'empathie pour se comporter moralement, on arguera peut-être qu'il existe des recherches tentant de faire ressentir de la douleur ou de l'empathie aux robots. Mais dans ce cas, si tant est que ces recherches aboutissent, cela reviendrait à donner des sentiments au robot, et donc à perdre ce qui était présenté comme son atout majeur : un agir purement rationnel. Si en conséquence il n'était plus rationnel, ce serait prendre le risque de le voir reproduire les débordements immoraux des humains emportés par leurs sentiments. Quel serait alors la contribution d'un robot éthique s'il peut lui aussi devenir fou de douleur ?

Enfin, espérer programmer l'éthique est peut-être voué à l'échec compte tenu de la diversité des situations dans le champ éthique et de l'imprévisibilité de celles-ci. Sur ce point, l'homme est peut-être plus adapté que le robot. Pour n'en prendre qu'un exemple, prenons le célèbre jugement de Salomon. Deux femmes affirment chacune être la mère d'un même bébé. Salomon propose donc une répartition « équitable » : couper l'enfant en deux pour le partager.

⁶ P. LIN, G. BEKEY, K. ABNEY, « Autonomous Military Robotics : Risk, Ethics, and Design », rapport à destination du *US Department of Navy, Office of Naval Research*, préparé par « *Ethics + Emerging Sciences Group* » de la *California Polytechnic State University, San Luis Obispo*, 2008, p 42.

L'une des deux femmes renonce à sa part pour sauver la vie de l'enfant. Salomon le lui rend alors. Ce qui a permis à Salomon d'appliquer le comportement le plus juste, le plus éthique, est sa créativité et son intelligence affective. Les situations complexes le sont souvent parce qu'elles sont inédites. Il ne suffit pas toujours de traiter efficacement les données du problème pour trouver une réponse équitable. Dans cet exemple, l'équité seule aurait donné un assassinat. Il fallait le concours de l'imagination, de l'instinct, de l'intelligence affective pour dénouer la situation. Si l'on oppose à cet argument que l'on tente de doter les robots de capacités pour déduire les intentions, les désirs et les buts des autres, il reste d'une part à analyser les résultats de ces recherches, et d'autre part, même s'ils accéderaient à une certaine compréhension des sentiments humains, il leur manquerait encore la capacité d'y répondre de façon adéquate, et avec psychologie. Puisqu'un robot ne peut qu'appliquer les comportements prévus dans son programme, nous pouvons douter qu'un robot dispose un jour d'une sagesse digne de Salomon.

Conclusion de la partie :

La tentation de croire qu'un robot autonome peut être un agent moral est le produit de réductions de la réalité et de glissements de langages successifs. Ils ne sont pas en eux-mêmes condamnables si on ne s'y laisse pas tromper. Le robot possède une intelligence artificielle qu'on oublie parfois n'être qu'artificielle. Sa capacité d'autonomie se résume à faire tourner seul son programme. Son raisonnement n'est que calcul, sa décision n'est qu'une sélection parmi des réponses prévues dans son programme. Le robot n'est pas libre, il n'a pas d'autonomie morale puisqu'il n'a pas la possibilité d'agir in-éthiquement, ni de prendre du recul vis-à-vis des ordres auxquels il obéit. Il est complètement soumis à sa programmation, c'est d'ailleurs ce qu'on lui demande. Si malgré tout il pouvait s'en arracher, il serait alors immoral et bientôt mis hors service. Si l'on considère qu'il n'a qu'une moralité d'esclave, ce serait insatisfaisant au point de vue éthique, car même l'esclave antique restait responsable de ses actes. La responsabilité du robot ne saurait être autre que matérielle, en cas d'indécidabilité il ne saurait faire jouer que le hasard pour trancher, ce qui ne serait pas moral.

Au vu de tous ces éléments, il semble impossible de créer une définition de l'éthique qui intégrerait le robot au sein des agents moraux et dont les exigences seraient suffisantes pour être appelée éthique. Quand aux traditions philosophiques existantes, nous avons vu qu'il semblait difficile de faire du robot un agent moral si l'on suit l'ensemble de la pensée de leurs auteurs. Il n'est pas un agent éthique au sens aristotélicien car il ne peut pas atteindre le bonheur ni être vertueux. On ne peut le croire éthique au sens kantien qu'à condition de limiter l'étendue de la pensée éthique de Kant. Sa soumission à son programme ne s'apparente pas à l'action accomplie par respect de la loi morale. L'obligation de respect de la loi morale ne s'impose pas au robot. Enfin, on le considère comme éthique lorsqu'on réduit l'éthique au droit, alors qu'il ne sera au mieux qu'un robot légal.

Quel avenir reste-t-il pour le « robot autonome éthique » ? Il peut se borner à suggérer des réponses dans le domaine de l'éthique, il peut être une aide à la prise de décision éthique par des humains sans être un agent lui-même. Malheureusement, ce ne sera probablement possible que dans un cadre limité et simple où l'humain aurait fait tout aussi bien de lui-même. Les situations les plus complexes, autrement dit celles où il aurait été utile que les capacités humaines soient secondées, nécessitent souvent plus de créativité et d'imagination éthique pour résoudre les situations cornéliennes que de simples réflexions calculatoires.