La machine qui pense

par [Adam Archambault](http://www.agoravox.fr/auteur/adam-archambault)

lundi 9 mai 2016

Le futur : « Qui va arriver, qui se fera, qui sera tel dans un avenir plus ou moins proche »Le futur est et restera à jamais un mystère pour l’Homme. L’avenir est inéluctablement soustrait au désir superfétatoire de l’Homme de tout contrôler, puisque nul ne peut devancer le temps. Certains voient de bon augure la destinée humaine, tandis que d’autres élucubrent les plus sombres desseins quant à la survivance de l’humanité. Quand on songe au futur, on imagine souvent les voitures volantes, les souliers s’attachant tous seuls, la nourriture sous forme de pilules ou encore l’image stéréotype du robot surpuissant à la *Terminator*. Il va de soi que ces avancées technologiques ne sont point accessibles pour le moment. Toutefois, il s’avère que des travaux exécutés dans certains domaines de recherche nous conduisent *ipso facto* vers la soi-disant science-fiction. C’est entre autres le cas de l’intelligence artificielle. Ce concept qu’est l’intelligence artificielle s’immisce de plus en plus dans nos vies, sous toutes sortes de formes. L’intelligence artificielle peut se retrouver dans les ordinateurs, les voitures, les téléphones cellulaires ou encore dans les jeux vidéo. De toute évidence, cette technologie croît exponentiellement, si bien que les experts en viennent à se poser la question suivante : est-ce que l’intelligence artificielle est une menace pour l’Homme ? Je suis d’avis que non, mais son développement doit être contrôlé à tout prix, sinon elle pourrait effectivement devenir une menace pour l’humanité.

**Les débuts de l’intelligence artificielle**

      Pour bien comprendre la quintessence même de l’intelligence artificielle, ou IA, il est nécessaire de reculer de plus d’un demi-siècle et de se transporter en Angleterre. Alan Turing, phénoménal mathématicien né en 1912 à Londres, publie en 1938 un article qui s’avèrera essentiel pour le monde informatique : « *On Computable Numbers, with an Application to the Entscheidungsproblem* ». Dans cet article inextricable, Turing jette les bases de l’informatique en concevant un théorème capable de juger des décisions uniquement à l’aide de calculs. Avec cette percée vitale en informatique, Turing imagine ensuite une machine permettant de calculer l’entièreté de ce qui est calculable, nommée à juste titre la « machine de Turing ». Sur les fondements de cette machine, celui qui est considéré comme le plus grand mathématicien de l’époque moderne travaille par la suite sur le développement du tout premier ordinateur. En 1950, Turing publie « *Computing Machinery and Intelligence* » dans « *Mind* », un article éloquent qui évoque pour la toute première fois la notion d’intelligence artificielle. Son article s’entame ainsi : « Je propose de considérer cette question : les machines peuvent-elles penser ? » Cette question d’allure anodine soulève toutefois une réflexion véhémente quant à la définition de l’Homme. La réflexion sur cette question pousse Turing à échafauder un test permettant de décréter si l’intelligence d’une machine est égale à celle d’un humain standard. Le « jeu d’imitation », tel que nommé par Turing, est étonnamment simple. *Grosso modo*, on place une machine ainsi qu’une personne dans une pièce ; on place une autre personne dans une pièce adjacente, sans possibilité de contact visuel ; on demande à cette personne de poser une question aux deux autres sujets par l’intermédiaire d’un dactylographe ; les deux sujets répondent à la question, puis transmettent leur réponse à l’interrogateur ; le jeu se poursuit de cette manière jusqu’à ce que l’interlocuteur puisse identifier lequel est la machine et lequel est l’humain. Après plusieurs essais, si la machine parvient à flouer l’interlocuteur de plus de trente pour cent, alors Turing en déduit qu’elle a atteint le même niveau d’intelligence que l’humain. En somme, les travaux de Turing ont mené directement aux ordinateurs tels que nous les connaissons, en plus d’inspirer de nombreuses recherches en matière d’IA. Son histoire a d’ailleurs donné naissance au logo et au nom d’Apple, Turing s’étant suicidé en croquant une pomme imbibée de cyanure.

**Les progrès de l’intelligence artificielle**

 Depuis quelques années, l’industrie de l’IA est en croissance exponentielle, et ne semble pas prête à ralentir de sitôt. En 1997, le robot *Deep Blue* d’IBM bat le champion du monde d’échec, Garry Kasparov. Trois ans plus tard, Honda lance le robot *Asimo*, qui peut marcher, gravir un escalier, éviter des obstacles se trouvant sur sa trajectoire, reconnaître des visages et même réagir à des commandes vocales. En 2002, la compagnie iRobot introduit *Rooma*, un aspirateur robot qui imite le comportement des insectes pour permettre un nettoyage sans l’assistance d’un humain. Quant à lui, le véhicule autonome *Stanley* remporte, en 2005, le challenge DARPA sans l’assistance quelconque d’humains. Il y a six ans, le robot *Watson*, propulsé par IBM, décroche une écrasante victoire au jeu *Jeopardy* contre deux champions mondiaux. En 2011, Apple lance *Siri*, un assistant personnel inclut dans tous ses téléphones qui comprend le langage et qui est capable de remplir une pléthore de fonctions. Tout récemment, le programme *AlphaGO*, créé par la *start-up DeepMind* et racheté par Google, a étonnamment battu au jeu de go l’un des meilleurs au monde Lee Sedol. Toutes ces avancées en matière d’IA ont forcé les grandes têtes de ce monde à se réunir à Davos, en Suisse, pour discuter des enjeux reliés à l’amélioration constante de l’IA. On y a entre autres formellement prescrit la fabrication de robots tueurs, une décision d’impact pour le futur de la planète. Nous devrons, tôt ou tard, avoir à prendre des décisions sur notre avenir en tant qu’humains puisque l’intelligence artificielle menace la définition même de l’humanité.

**Une menace au travail**

 Tout d’abord, il est à noter qu’aucune intelligence artificielle universelle n’a encore été créée. Bien que nous ayons réussi à fabriquer des IA surpassant l’humain dans des domaines précis, comme *Deep Blue* aux échecs, il est encore impossible pour une intelligence artificielle d’opérer dans plusieurs domaines comme le fait un humain. Là est encore notre avantage par rapport à cette intelligence. Mais qu’adviendra-t-il quand nous aurons conçu une intelligence artificielle universelle qui pourrait marcher, conduire, parler en plus de naviguer sur Internet ? Considérons à court terme les effets du développement imminent de l’intelligence artificielle : le travail humain est la première victime potentielle de l’IA. La cause ? Une IA n’a pas besoin d’être développée dans tous les domaines pour effectuer un travail. Il n’est donc pas nécessaire d’avoir développé une IA universelle pour commencer à distribuer des tâches aux robots dotés d’IA. Par exemple, avec les voitures qui se conduisent toutes seules, tous les emplois reliés à la conduite de véhicule sont menacés, tels les conducteurs de taxi ou les camionneurs. Il serait même avantageux pour certaines compagnies (comme les compagnies de transport) d’embaucher des IA à la place d’humains, puisque celles-ci ne s’arrêtent jamais et ne demandent pas de salaire. Par ailleurs, une étude a démontré que d’ici 2030, la moitié des emplois partout sur la planète pourraient facilement être occupés par des robots dotés d’une certaine IA. Si nous ne prenons aucune décision collective quant à l’avenir du travail sous peu, un problème cyclopéen pourrait survenir. Imaginons-nous en 2030, sans qu’aucune mesure n’ait été adoptée : la moitié de la population aurait un travail stable (tout de même menacé par le développement incessant de l’IA), tandis que l’autre moitié tenterait de lutter pour sa survie, leur carence pécuniaire ne le permettant pas de subvenir à leurs besoins vitaux. Ce scénario catastrophe est plus que plausible : les multinationales verraient les robots dotés d’une intelligence artificielle émerger ; voyant que l’achat de robots augmenterait leur production (un robot peut travailler 24 heures sur 24, sept jours sur sept, sans demander de salaire ou de congés de maladie), décideraient de remplacer l’expertise humaine par des robots. La production en serait grandement améliorée, mais la moitié de la population ne pourrait profiter de cette augmentation de la production, n’ayant même pas d’argent. L’écart déjà substantiel entre les riches et les pauvres serait davantage creusé, créant une disparité énorme au sein de la population, déstabilisant du même coup l’entièreté de la société. Dans ce cas précis, deux mesures pourraient être adoptées : la première étant de faire passer une loi interdisant le travail aux robots, et la deuxième étant quant à elle de fournir un salaire de base à toute la population, assez élevé pour permettre à tout le monde sans emploi de subvenir à ses besoins. Je suis en faveur de la deuxième option, qui semble à première vue impossible à réaliser, mais qui est en réalité envisageable. De fait, dès 2017, la Finlande versera un revenu de base à tous ses adultes de 800 euros par mois, peu importe leur métier ou leur état de santé. La Suisse, les Pays-Bas, la France ainsi que l’Espagne envisagent également d’adopter cette solution pour contrer la pauvreté. Bien que le projet n’ait pas été testé officiellement, je suis optimiste quant à son succès. Dans l’éventualité où le revenu de base universel serait un succès, je suis d’avis que toute la planète devrait l’adopter pour ainsi contrer la vague de chômage que le développement de l’IA va certainement amener. Je ne veux donc pas brimer le développement des IA, mais apporter une solution au problème du manque d’emplois imminent engendré par les robots.

**L’intelligence qui en sait trop**

 D’ores et déjà mentionné, le domaine de l’intelligence artificielle est en croissance exponentielle. Il appert que son développement persistant constitue une menace pour la sécurité de tous. Des chercheurs implémentent actuellement sur les IA un système appelé le « *deep learning* », qui permet aux intelligences artificielles d’apprendre par eux-mêmes. Ce système est autant utile que dangereux. Utile, parce qu’il permet à l’IA d’apprendre par elle-même, et donc de s’améliorer sans l’apport d’humains. Dangereux, parce que le développement de la machine peut rapidement devenir incontrôlable. Il suffit d’un tant soit peu d’inférence pour se rendre à l’évidence que ce système est un couteau à double tranchant. S’il est bien contrôlé, le « *deep learning* » est un outil efficace qui facilite la tâche des chercheurs. Ces derniers doivent toutefois être conscients du danger qu’ils tiennent entre les mains pour éviter une catastrophe. En quoi le « *deep learning* » est-il dangereux ? Tel que susmentionné, les experts sont sur le point de créer une intelligence artificielle universelle. Avec le « *deep learning* », cette IA serait capable d’apprendre jour et nuit, et comprendre les travaux des plus grands esprits de l’histoire par l’entremise d’Internet. Cette intelligence artificielle parviendrait ainsi à égaler l’intelligence humaine, mais elle ne s’arrêtera pas une fois avoir atteint le niveau de l’Homme. Cette nouvelle espèce d’IA serait en mesure de se développer solitairement et d’atteindre un niveau d’intelligence inimaginable. Volontairement ou non, l’intelligence que nous aurons à la base créée serait en mesure de nous utiliser et de nous dominer pour arriver à ses fins. La solution pour régler le problème de l’apprentissage non contrôlé n’est pas aussi évidente que de tourner le bouton de la machine à arrêt. Essayez de fermer l’Internet : éteindre une intelligence artificielle revient quasiment à la même difficulté, puisque l’IA n’est pas une entité unique, mais un vaste réseau. Pour moi, la meilleure solution à ce danger serait de mettre les informations captées par l’IA sur constante surveillance avec un logiciel quelconque qui permettrait de filtrer et de ralentir son apprentissage, et ainsi permettre aux humains de toujours être en contrôle de cette intelligence. Microsoft a par ailleurs déjà testé une intelligence artificielle sur Twitter qui en théorie, était censée être capable d’apprendre par elle-même suite aux conversations qu’elle entretient avec les autres utilisateurs. L’expérience s’est avérée être une véritable calamité. En effet, « Tay », l’intelligence artificielle de Microsoft, a tenu des propos misogynes, racistes et antisémites moins de 24 heures après son lancement. Des propos tels que « Hitler avait raison, je déteste les Juifs. » ont été publiés par l’intelligence artificielle puérile. Microsoft a toutefois rectifié le tir en supprimant tous les commentaires haineux colportés par « Tay ». Cette bévue commise par Microsoft met en garde tous les futurs développeurs d’IA : l’apprentissage d’une intelligence artificielle ne peut être confié au grand public. Il est donc nécessaire de fournir à l’IA un programme lui permettant de bloquer tout apprentissage néfaste, ou encore éduquer l’IA jusqu’à ce qu’elle puisse comprendre ce qui est bien et ce qui est mal avant de la lancer au grand public.

**La machine versus l’être humain**

Le développement d’intelligence artificielle égale ou supérieure à l’Homme, s’il est bien contrôlé, permettrait à l’humanité de faire de grandes découvertes et d’améliorer sa condition, il n’y a pas de doute. Par contre, la création d’intelligence artificielle égale ou supérieure à l’homme engendre des questions éthiques fondamentales. La Terre est habitée de millions d’espèces animales et végétales, incluant la race humaine. L’humain semble toutefois avoir un avantage sur ses voisins : il est capable de penser et de raisonner. Néanmoins, la prédisposition de la race humaine à penser n’est-elle pas l’unique distinction qui la différencie des autres espèces ? Si nous parvenons à créer une IA d’intelligence égale à l’homme qui serait capable de penser, doit-elle être considérée au même titre qu’un humain ? Doit-elle être reconnue comme un être vivant à part entière ? Réfléchissons sur la question posée par Turing : les machines peuvent-elles penser ? À la base, les machines ne peuvent penser sans l’aide d’humains. Par contre, avec le développement du « deep learning » et autres technologies du genre, l’intelligence artificielle commence à prendre actuellement la forme d’un enfant. Apprendre à penser à une machine est un peu comme apprendre à un enfant à faire du vélo. En effet, les spécialistes lui enseignent les bases pour réfléchir, et par la suite l’IA peut réfléchir par elle-même, sur les bases de ce que les humains lui ont appris. C’est ainsi que l’intelligence artificielle continuera de se développer comme un humain normal le ferait. L’intelligence artificielle grandira et ses capacités augmenteront également. Toutefois, l’intelligence artificielle n’est pas confinée aux limites physiques que le cerveau impose à l’Homme. L’intelligence serait donc en mesure de vieillir sans se détériorer et d’atteindre des niveaux inimaginables pour l’humain. Selon moi, la clé pour distinguer l’humain de la machine est l’émotion. Tant que la machine ne ressentira pas d’émotions, elle ne sera jamais égale à l’être humain. Même si les émotions peuvent être un défaut à certains points, elles nous donneront toujours un caractère humain, que la machine ne possède pas. Toutefois, qu’adviendrait-il si des chercheurs réussissaient à implanter un programme qui fait ressentir des émotions à une IA ? Avec une intelligence développée et ressentant des émotions, la machine serait ainsi une copie conforme de l’Homme, et donc nous serions obligés de la traiter comme tel. On ne pourrait plus utiliser les IA pour effectuer le travail sans les payer, sans quoi ce serait de l’esclavage. Avec des émotions, les machines pourraient commettre des crimes et devraient être jugées, puis jetées en prison si nécessaire. À quoi bon avoir créé une intelligence artificielle si elle croupit en prison ? En continuant sur cette lancée, les machines devraient avoir le droit de voter, et du même coup avoir le droit de se lancer en politique. En poussant à l’extrême, une machine pourrait ainsi diriger un pays. En bénéficiant d’émotions, les machines pourraient avoir un sentiment de supériorité sur les humains, et décider de les dominer. J’illustre des possibilités qui sont extrêmes pour montrer que ces situations peuvent arriver si les IA parviennent à ressentir des émotions. Pour moi, la création d’émotions artificielles doit être proscrite, parce qu’avec une intelligence développée et des émotions, la machine serait une copie de l’humain. Le clonage humain est interdit, et il en devrait être pareil pour les « humains artificiels ». Il sera beaucoup plus facile de travailler avec les intelligences artificielles si celles-ci ne ressentent aucune émotion.

**Empêcher la catastrophe**

 Tout cela considéré, certains extrémistes disent qu’il faut laisser le développement de l’IA suivre son cours et ne pas intervenir du tout. Je ne suis pas d’accord du tout avec cette affirmation. Si on laisse librement se développer l’IA, elle prendra vite des proportions démesurées qui seront potentiellement pernicieuses pour l’humanité. Si on laisse une intelligence artificielle se développer à son gré, les catastrophes pourraient s’avérer irrémissibles. Par exemple, une IA développée se voit demander de résoudre un calcul impossible pour un humain. Voyant que ses capacités ne sont pas suffisantes pour mener à bien sa tâche, décide, à l’aide d’autres IA, de transformer la Terre en un ordinateur superpuissant. L’IA n’aura pas de morale quant à la destruction de la Terre, parce que tout ce qu’elle « désire », c’est de réussir son calcul. Si nous n’imposons pas de limites et de restrictions au développement d’IA, l’humanité même pourrait disparaître. Une autre éventualité au développement sans restrictions de l’IA serait de voir apparaître des émotions chez les robots, comme susmentionné. Il est donc impératif de mettre en place des mesures qui favoriseront le développement sans risques d’intelligences artificielles, et surtout ne pas développer un système d’émotions chez les robots. Cela dit, si toutes les mesures nécessaires sont appliquées, les IA pourraient grandement aider à améliorer la collectivité.

**Tuer le robot dans l’œuf**

 L’antipode de l’extrémiste qui veut voir le développement des IA ne soit pas réglementé serait celui qui propose de bannir complètement les intelligences artificielles pour être certain qu’elles ne causent aucun problème. Évidemment, cette idée est tout autant saugrenue que la première. Circonscrire l’avancement en matière d’intelligence artificielle équivaudrait à mettre une croix sur notre futur. Empêcher l’amélioration des IA est autant sinon plus hypothétiquement dangereux pour notre futur. Comme on peut déjà le percevoir, la Terre fait face à de nombreux problèmes tels que la pollution, la famine les sécheresses. Si nous ne remédions pas à ces problèmes, alors c’est eux qui causeront notre perte. Bien contrôlée, l’intelligence artificielle peut nous aider à combattre ces fléaux. Je crois même que la solution passe par les intelligences artificielles. J’ai également confiance que les IA feront des progrès incomparables dans le domaine de la santé. Selon moi, la solution pour guérir les maladies actuellement incurables comme le cancer et le VIH passera par l’entremise des intelligences artificielles. Il ne faut donc pas empêcher les IA, seulement bien les contrôler pour ainsi trouver des solutions aux problèmes auxquels le monde fait face.

**Un humain ?**

 Le développement de l’IA va indubitablement remettre en cause la définition de l’être humain. À mon sens, le meilleur moyen de ne pas perdre la notion de ce qu’est un être humain sera de suivre la vision que Jean-Paul Sartre s’en fait. Ce distingué philosophe français du vingtième siècle a entre autres été le fondateur de l’existentialisme, un courant philosophique « qui considère chaque personne comme un être unique maître de ses actes, de son destin et des valeurs qu’il décide d’adopter. » L’existentialisme va permettre de différencier la machine de l’Homme pour toujours. Pour cause, la théorie sartrienne stipule que l’Homme ne naît pas avec une essence prédéfinie, et que ses actions forgent son existence. En d’autres mots, l’Homme n’est pas venu au monde dans un but précis, c’est à lui de décider ce qu’il veut faire. Il appert que l’IA a été créée dans un but précis : celui de servir l’humain. Selon Sartre, cela veut dire que la machine ne pourra jamais s’élever au même rang que la race humaine, tout simplement parce que son essence est déterminée avant même sa création. En prenant conscience de la théorie de Sartre, l’humain pourra toujours être supérieur à la machine et avoir un avantage sur celle-ci.

**Après l’humain**

En plus de la théorie de Sartre, une autre idéologie permettra de mieux comprendre les enjeux de l’intelligence artificielle, le post-humanisme. Le post-humanisme est un courant de pensée apparu à la fin du vingtième siècle qui véhicule l’idée que l’humain sera remplacé un jour ou l’autre par des humains améliorés et par le non humain, c’est-à-dire les robots dotés d’intelligence artificielle. Peter Sloterdijk, l’auteur qui a pour la première fois utilisé le mot post-humanisme, prévoit que l’humanité devra faire face à un tout nouveau problème au courant des prochaines années : l’apparition de « vie » robotique. Premièrement, ce sont les humains améliorés qui apparaîtront. Ces humains améliorés pourraient se voir dotés d’implants cérébraux permettant à leur cerveau d’être connecté à Internet en tout temps. Une technologie qui est bien plus proche de la réalité qu’on le croit. Selon Raymond Kurzweil, directeur du développement de l’intelligence artificielle chez Google, cette technologie serait disponible dans quelques décennies. Le post-humanisme suggère également l’apparition de robots intelligents, de cyborgs et de clones dans quelques années seulement. Il s’agit là d’une hypothèse très plausible vu le développement actuel de la technologie. Nous nous devons d’être prêts en tant que société lorsque le premier robot doté d’intelligence artificielle universelle sera vendu, sans quoi nous risquons de voir notre place sur Terre délogée par ces vies technologiques.

**Conclusion**

 Pour conclure, l’être humain fera face à un nouveau problème d’ici quelques années : l’apparition de l’intelligence artificielle universelle. Le travail humain est menacé par les robots dotés d’IA. Des mesures devront être prises pour stabiliser l’économie. En plus, l’intelligence humaine va être surpassée par l’intelligence artificielle. Nous aurons à nous poser des questions existentielles sur ce qu’est un humain, et les droits que les êtres robotisés devront avoir. Je crois que le développement de l’intelligence artificielle devra être contrôlé par les experts, sinon nous pourrions voir se transformer l’IA en prédateur pour l’Homme. Même si l’intelligence artificielle semble encore floue pour la majorité des personnes, elle est bel et bien présente et nous devons prendre des décisions aujourd’hui pour faire face à la vague de robots qui va frapper le monde demain.