

# Intelligence Artificielle

**Russell & Norvig**

3ème édition, 2010, éd. Pearson (1199 pages).

[...]

**Fondements Philosophiques**

[...]

Page 1090

## 26.3 Éthique et risques du développement de l'intelligence artificielle

Jusqu'à maintenant, nous nous sommes demandé s'il était possible de construire des intelligences artificielles. Nous devons également nous demander si c'est **souhaitable**. Si les effets de l'IA sont plus susceptibles d'être plus néfastes que bénéfiques, ceux qui travaillent dans ce domaine auraient la responsabilité morale de réorienter leurs recherches. Nombre de nouvelles technologies ont eu des effets de inattendues : la fission nucléaire a abouti à Tchernobyl et à la menace de la destruction totale, le moteur à combustion interne a conduit à la pollution atmosphérique et au réchauffement global, et transformé les petits chemins en autoroutes. En un sens, les automobiles sont des robots qui ont conquis le monde en se rendant indispensables.

Tous les scientifiques et les ingénieurs sont confrontés à des considérations éthiques: comment se comporter dans leurs travaux, quels projets entreprendre ou non et comment les gérer (voir l'ouvrage consacré à *l'éthique de l'informatique* par (Berleur et Brunnstein, 2001)). Toutefois, l'IA semble poser de nouveaux problèmes, qui dépassent celui de construire des ponts qui ne s'effondrent pas:

- Les gens pourraient perdre leurs emplois au profit de l'automatisation.
- Les gens pourraient avoir trop (ou pas assez) de temps libre.
- Les gens pourraient perdre le sens de leur unicité.
- Les systèmes d'IA pourraient être utilisés à des fins indésirables.
- L'utilisation de systèmes d'IA pourrait entraîner une perte de responsabilité.
- Le succès de l'IA pourrait signifier la fin de l'espèce humaine.

Nous allons examiner chacune de ces questions tour à tour.

**Les gens pourraient perdre leurs emplois au profit de l'automatisation.** L'économie industrielle moderne est devenue dépendante des ordinateurs en général et de certains programmes d'IA en particulier. Par exemple, une grande partie de l'économie, surtout aux États-Unis, dépend de l'existence du crédit à la consommation. Les applications de gestion de cartes de paiement, d'approbation de crédits et de détection des fraudes sont maintenant des programmes d'IA. On pourrait dire qu'ils ont remplacé des milliers de travailleurs, mais, en réalité, si l'on supprimait ces programmes, ces emplois n'existeraient pas parce que le travail humain ajouterait un coût inacceptable aux transactions. Jusqu'ici, l'automatisation, permise par les technologies de l'information en général et par l'IA en particulier, a créé plus d'emplois qu'elle n'en a éliminés, et qui sont plus intéressants et mieux rémunérés. Maintenant que le programme d'IA canonique est un «agent intelligent» conçu pour assister les humains, la perte d'emplois est une menace moins grave que lorsque l'IA se concentrait sur les « systèmes experts» destinés à les remplacer. Mais certains chercheurs pensent que c'est là le but que l'IA doit se fixer. Réfléchissant au vingt-cinquième anniversaire de l'AAAI, Nils Nilsson (2005) a posé comme défi la création d'une IA de niveau humain qui pourrait passer avec succès un test d'embauche au lieu du test de Turing: un robot qui pourrait apprendre à exécuter n'importe laquelle d'une gamme de tâches. Nous pourrions nous retrouver à l'avenir dans un monde où le taux de chômage serait élevé, mais où les chômeurs pourraient diriger des cohortes de robots.

**Les gens pourraient avoir trop (ou pas assez) de temps libre.** Dans le *Choc du futur*, Alvin Toffler (1970) écrivait que la durée de la semaine de travail avait diminué de 50 % depuis le tournant du siècle, et qu'il n'était pas absurde de prédire qu'elle serait de nouveau réduite de moitié avant l'an 2000. Arthur C. Clarke (1968b) écrivait que les hommes de 2001 seraient confrontés à un avenir d'ennui absolu, dans lequel le principal problème de l'existence serait de sélectionner une chaîne de télévision parmi plusieurs centaines disponibles. La seule de ces prédictions qui s'est approchée de la vérité concerne le nombre de chaînes de télévision. En revanche, ceux qui travaillent dans des industries grosses consommatrices de connaissances ont découvert qu'ils faisaient partie d'un système informatisé intégré fonctionnant vingt-quatre heures par jour; pour suivre le mouvement, ils ont été forcés de travailler plus longtemps. Dans une économie industrielle, les récompenses sont approximativement pro-

Page 1091

portionnelles au temps investi: travailler 10 % de plus tendrait à signifier une augmentation de revenus de 10 %. Dans une économie de l'information caractérisée par des communications à haut débit et une facilité de reproduction de la propriété intellectuelle (ce que Frank & Cook (1996) nomment la Winner-Take-All Society) une importante récompense est associée au fait d'être légèrement meilleur que la concurrence: travailler 10 % de plus pourrait signifier une augmentation de revenus de 100 %. Tous subissent donc une pression croissante pour travailler plus. En accélérant le rythme de l'innovation technologique, l'IA contribue à cette tendance générale tout en promettant de nous permettre de prendre un peu de temps libre et de laisser les agents automatisés gérer les choses un moment. Tim Ferriss (2007) recommande de recourir à l'automatisation et à la délocalisation pour parvenir à une semaine de travail de quatre heures.

**Les gens pourraient perdre le sens de leur unicité.** Dans *Computer Power and Human Reason*, Weizenbaum (1976), l'auteur du programme EUZA, nous alerte sur certaines des menaces potentielles que l'IA représente pour la société. L'un de ses principaux arguments est que la recherche en IA rend possible l'idée que les humains sont des automates - idée qui entraîne une perte d'autonomie, voire d'humanité. Notez que cette idée est beaucoup plus ancienne que l'IA et remonte au moins à *L'Homme-Machine* (La Mettrie, 1748). L'humanité a survécu à d'autres atteintes à son sens de l'unicité: De *Revolutionibus Orbium Coelestium* (Copernicus, 1543) délogeait la Terre du centre du système solaire, et *La Filiation de l'homme* (Darwin, 1871) plaçait l'*Homo sapiens* au même niveau que les autres espèces. Si elle se répand largement, l'IA peut être au moins aussi menaçante pour les postulats moraux de la société du XXI<sup>e</sup> siècle que la théorie évolutionniste l'a été pour ceux du XIX<sup>e</sup>.

**Les systèmes d'IA pourraient être utilisés à des fins indésirables.** Les puissants ont souvent utilisé les technologies avancées pour éliminer leurs rivaux. Comme l'écrivait le théoricien des nombres G. H. Hardy (Hardy, 1940), « Une science est dite utile si son développement tend à accentuer les inégalités existantes dans la répartition des richesses, ou promeut plus directement la destruction de la vie humaine. » Cela s'applique à toutes les sciences, et l'IA n'y fait pas exception. Les systèmes intelligents autonomes sont maintenant monnaie courante sur le champ de bataille, et l'armée des États-Unis a déployé plus de 5000 avions et 12 000 véhicules terrestres autonomes en Irak (Singer, 2009). Une théorie soutient que les robots militaires sont analogues aux armures médiévales poussées jusqu'au bout de leur logique: personne n'aurait d'objection morale à ce qu'un soldat veuille porter un casque pour se protéger des attaques d'ennemis furieux brandissant des haches, et un robot télécommandé est comparable à une forme d'armure très sûre. En revanche, les armes robotisées exposent à des risques supplémentaires. Si les humains venaient à être exclus de la boucle de décision, les robots pourraient se retrouver à faire des choix qui entraîneraient la mort de civils innocents.

À une plus grande échelle, la possession de robots puissants (comme celle de casques solides) pourrait donner à une nation un excès de confiance qui la conduirait à entrer en guerre avec plus de témérité que nécessaire. Dans la plupart des conflits, au moins l'une des parties présume trop de ses capacités militaires, sinon ils auraient été résolus pacifiquement.

Weizenbaum (1976) a également souligné que les technologies de reconnaissance de la parole pourraient entraîner la généralisation des écoutes clandestines et donc une perte de libertés civiques.

Il n'avait pas prévu un monde dans lequel les menaces terroristes modifieraient la dose de surveillance que les individus sont prêts à accepter, mais il avait bien reconnu que l'IA possédait un potentiel de production de surveillance de masse. Sa prédiction s'est en partie réalisée: le Royaume-Uni dispose maintenant d'un énorme réseau de caméras de

Page 1092

surveillance, et d'autres pays contrôlent systématiquement le trafic web et les appels téléphoniques. Certains acceptent que l'informatisation entraîne une perte de la vie privée -le P.-D.G. de Sun Microsystems, Scott McNealy, a dit: « Vous n'avez pas de vie privée, de toute façon. Il faut vous y faire.» David Brin (1998) soutient que cette perte est inévitable, et que la façon de combattre l'assymétrie entre le pouvoir de l'État et celui des individus était de rendre la surveillance accessible à tous les citoyens. Etzioni (2004) plaide en faveur d'un équilibre entre la vie privée et la sécurité, entre les droits individuels et l'intérêt de la communauté.

**L'emploi de systèmes d'IA pourrait entraîner une perte de responsabilité.** Dans l'atmosphère procédurière qui règne aux États-Unis, la responsabilité légale devient un problème important. Quand un médecin fait confiance au jugement d'un système expert pour émettre un diagnostic, qui est en faute si celui-ci est faux? Heureusement, en partie à cause de l'influence croissante des méthodes de la théorie de la décision en médecine, on accepte maintenant que la négligence ne peut être établie si le praticien exécute des procédures médicales dont l'utilité *escomptée* est élevée, même si le résultat *réel* est catastrophique pour le patient. La question devrait donc être: « Qui est en faute si le diagnostic est déraisonnable? » Les tribunaux ont considéré jusqu'ici que les systèmes experts jouaient le même rôle que les manuels et les ouvrages de référence: les médecins ont la responsabilité de comprendre le raisonnement sous-jacent à toute décision et d'exercer leur propre jugement pour déterminer s'ils doivent accepter les recommandations du système. Si l'on conçoit des systèmes experts médicaux comme des agents, on ne doit plus penser que leurs actions affectent directement le patient mais qu'elles exercent une influence sur le comportement du praticien. Si les systèmes parvenaient à émettre des diagnostics plus fiables et plus exacts que ceux des humains, les médecins pourraient devenir légalement responsables s'ils n'appliquaient *pas* leurs recommandations. Atul Gawande (2002) explore cette possibilité.

Des problèmes semblables commencent à se poser avec l'utilisation d'agents intelligents sur Internet. Des progrès ont été réalisés: on incorpore maintenant des contraintes aux agents intelligents afin qu'ils ne puissent pas, par exemple, endommager les fichiers d'un utilisateur (Weld et Etzioni, 1994). Le problème change d'ordre de grandeur quand l'argent change de mains. Si un agent intelligent effectue des transactions monétaires « au nom de quelqu'un », ce « quelqu'un » est-il responsable des dettes contractées? Serait-il possible qu'un agent intelligent possède lui-même des actifs et exécute des transactions en son propre nom? Jusqu'à maintenant, ces questions ne semblent pas très bien comprises. À notre connaissance, aucun programme n'a jamais bénéficié d'un statut juridique l'autorisant à des transactions financières, et, pour le moment, cette pratique semble déraisonnable. De même, les programmes ne sont pas considérés comme des «conducteurs» lorsqu'il s'agit de faire appliquer le code de la route sur des autoroutes réelles. Pour les lois californiennes, au moins, il ne semble exister aucune sanction légale susceptible d'empêcher un véhicule automatisé de commettre un excès de vitesse, alors que le concepteur du mécanisme de contrôle du véhicule serait responsable en cas d'accident. Comme pour les technologies de reproduction humaine, la loi n'a pas encore pris en compte ces nouveaux développements.

**Le succès de l'IA pourrait signifier la fin de l'espèce humaine.** Presque toutes les technologies sont potentiellement nuisibles lorsqu'elles se trouvent entre de mauvaises mains, mais l'IA et la robotique posent un nouveau problème: les mauvaises mains pourraient bien appartenir à la technologie elle-même. D'innombrables ouvrages de science-fiction nous ont mis en garde en nous racontant des histoires de robots ou de cyborgs mi-robots mi-humains pris de crises de folie meurtrière. Les premiers exemples remontent au livre de Mary Shelley,

Page 1093

*Frankenstein ou le Prométhée moderne* (1818) et à la pièce de Karel Capek, *R.U.R.* (1921), dans laquelle les robots conquièrent le monde. Au cinéma, *Terminator* (1984) combine le cliché de la conquête du monde par les robots et celui du voyage dans le temps, et *Matrix* (1999) associe ce même thème de la conquête à celui du **cerveau dans la cuve**.

Il semble que les robots sont les protagonistes de tant d'histoires de conquête du monde parce qu'ils représentent l'inconnu, tout comme les sorcières et les fantômes à des époques reculées, ou les Martiens de *La Guerre des mondes* (Wells, 1898). La question est de savoir si un système d'IA représente un risque plus important qu'un logiciel traditionnel. Nous allons examiner trois sources de risque.

Premièrement, l'estimation de l'état du système d'IA pourrait être incorrecte, ce qui le pousserait à commettre une erreur. Par exemple, une voiture autonome pourrait mal évaluer la position d'un véhicule roulant sur la voie adjacente, provoquant ainsi un accident qui pourrait tuer les passagers. Plus grave encore, un système de défense par missiles pourrait détecter à tort une attaque, déclencher une contre-attaque et causer la mort de millions de personnes. Ce risque n'est pas l'apanage des systèmes d'IA: dans les deux cas, la même erreur pourrait être commise par un humain ou par un ordinateur. Le meilleur moyen de l'atténuer est de concevoir un système doté d'un mécanisme d'équilibrage tel qu'une seule erreur d'estimation d'état passée inaperçue ne se propage pas dans tout le système.

Deuxièmement, la spécification de la bonne **fonction d'utilité** pour un système d'IA n'est pas si facile. Par exemple, on pourrait proposer une fonction destinée à *réduire la souffrance humaine*, exprimée sous la forme d'une fonction additive de récompense dans le temps, comme au chapitre 17. Toutefois, vu la nature humaine, nous trouverons toujours un moyen de souffrir, même au paradis, et la décision optimale pour le système d'IA est de mettre fin à l'espèce dès que possible: pas d'humains, pas de souffrance. Il convient donc d'être très prudents dans ce que nous demandons à de tels systèmes, alors que des humains n'auraient aucun mal à se rendre compte qu'il ne faut pas prendre la fonction d'utilité proposée au pied de la lettre. En revanche, les ordinateurs ne sont pas pollués par les comportements irrationnels décrits au chapitre 16. Les êtres humains utilisent parfois leur intelligence de façon agressive, parce qu'ils possèdent des tendances innées à l'agression dues à la sélection naturelle. Les machines que nous construisons n'ont pas besoin d'être naturellement agressives, sauf si nous décidons de les construire de cette manière (ou à moins qu'elles n'émergent comme le produit final de la conception d'un mécanisme qui encourage un comportement agressif). Heureusement, il existe des techniques, comme l'apprentissage par imitation, qui permettent de spécifier une fonction d'utilité par l'exemple. L'on peut espérer qu'un robot doté de suffisamment d'intelligence pour trouver le moyen de mettre fin à l'espèce humaine aura également celle de comprendre que ce n'était pas là la fonction d'utilité attendue.

Page 1094

Troisièmement, la fonction d'apprentissage du système d'IA peut l'amener à développer des comportements inattendus. Ce scénario, le plus sérieux, étant unique à ces systèmes, nous allons l'étudier plus en profondeur. I. J. Good a écrit (1965) :

« Définissons une **machine ultra-intelligente** comme une machine capable de surpasser toutes les activités intellectuelles de n'importe quel homme, aussi intelligent soit-il. Comme la conception de machines est l'une de ces activités intellectuelles, une machine ultra-intelligente pourrait concevoir des machines encore meilleures; il s'en suivrait incontestablement une « explosion d'intelligence » qui laisserait l'intelligence humaine loin derrière elle. En conséquence, la première machine ultra-intelligente est la *dernière* invention que l'homme aura jamais besoin de construire, pourvu que celle-ci soit suffisamment docile pour nous indiquer comment la garder sous contrôle. »

L'« explosion d'intelligence » a été également qualifiée de **singularité technologique** par le professeur de mathématiques et auteur de science-fiction Vernor Vinge, qui écrit (1993): « D'ici trente ans, nous aurons les moyens technologiques de créer une intelligence surhumaine. Peu après, l'époque de l'homme sera terminée.. Good et Vinge (et bien d'autres) notent avec raison que la courbe du progrès technologique est en train de croître (selon de nombreuses mesures) de façon exponentielle (voir la loi de Moore). De là à extrapoler que la courbe va continuer vers une

singularité de croissance quasi infinie, il y a très loin. Jusqu'à maintenant, toutes les technologies ont suivi une courbe en S, où la croissance exponentielle finit par diminuer progressivement. Il arrive que de nouvelles technologies apparaissent quand les anciennes ont atteint un palier, et l'on se heurte parfois à des limites infranchissables. Avec moins d'un siècle d'histoire des hautes technologies, il est difficile de prédire où nous en serons dans des centaines d'années.

Notez que le concept de machines ultra-intelligentes part du principe que l'intelligence est un attribut particulièrement important, et que, si l'on en possède suffisamment, on peut résoudre tous les problèmes. Mais nous savons qu'il y a des limites à la calculabilité et à la complexité calculatoire. Si le problème de la définition de machines ultra-intelligentes (ou même une approximation) appartient par exemple à la classe des problèmes NEXP-complets, et s'il n'existe pas de raccourcis heuristiques, même un progrès exponentiel de la technologie ne servira à rien: la vitesse de la lumière impose une limite supérieure stricte à la calculabilité, les problèmes qui la dépassent ne seront pas résolus, et nous ne savons toujours pas ce que sont ces limites supérieures.

Cette singularité à venir effraie Vinge, mais d'autres informaticiens et futurologues s'en délectent. Hans Moravec (2000) nous encourage à donner tous les avantages à nos « enfants spirituels », les robots que nous créons, qui peuvent nous dépasser en intelligence. Il existe même un nouveau mot - **transhumanisme** - pour désigner le mouvement social actif qui attend ce futur avec impatience. Inutile de dire que de telles questions représentent un défi pour tous les théoriciens de la morale, qui tiennent la préservation de la vie et de l'espèce humaine pour une bonne chose. Ray Kurzweil est actuellement l'avocat le plus visible de la singularité: dans *Humanité 2.0* (2005), il écrit:

« La singularité va nous permettre de transcender ces limitations physiques et cérébrales. Nous allons obtenir la maîtrise de notre destin. Nous allons prendre notre mortalité en main. Nous pourrions vivre aussi longtemps que nous le désirerons (ce qui ne signifie pas que nous pourrions vivre éternellement). Nous pourrions comprendre la pensée humaine et ses capacités. D'ici à la fin de ce siècle, la portion non biologique de notre intelligence sera des milliards et des milliards de fois plus puissante que l'intelligence humaine non assistée. »

Kurzweil note également les dangers potentiels: « Mais la singularité va aussi amplifier notre habileté à agir en suivant nos propres inclinaisons destructrices. L'histoire complète n'est donc pas encore écrite. »

Page 1095

Si les machines ultra-intelligentes sont une possibilité, nous ferions bien de veiller à concevoir leurs prédécesseurs de telle sorte qu'ils se conçoivent eux-mêmes pour nous traiter correctement. L'auteur de science-fiction Isaac Asimov (1942) a été le premier à aborder cette question, avec ses trois lois de la robotique:

1. Un robot ne peut porter atteinte à un être humain, ni, restant passif, permettre qu'un être humain soit exposé au danger.
2. Un robot doit obéir aux ordres que lui donne un être humain, sauf si de tels ordres entrent en conflit avec la première loi.
3. Un robot doit protéger son existence, tant que cette protection n'entre pas en conflit avec la première ou la deuxième.

Ces lois semblent raisonnables, au moins pour nous humains [8]. Mais leur mise en œuvre n'est pas chose aisée. Dans la nouvelle d'Asimov, *Cercle vicieux*, un robot est envoyé chercher du sélénium. Plus tard, on le retrouve en train de tourner en rond autour du gisement. Il détecte un danger chaque fois qu'il s'en approche, et la troisième loi le pousse à s'en écarter. Mais, chaque fois qu'il s'éloigne, le danger diminue et le pouvoir de la deuxième loi reprend le dessus, ce qui fait qu'il revient vers le sélénium. L'ensemble des points qui définit le point d'équilibre entre les deux lois décrit un cercle. Cela suggère que les lois ne sont pas des absolus logiques, mais qu'elles sont pondérées entre elles, avec un poids plus fort pour les premières lois. Asimov envisageait sans doute

une architecture fondée sur la théorie de la commande - peut-être une combinaison linéaire de facteurs - alors que, de nos jours, l'architecture la plus vraisemblable serait un agent probabiliste qui raisonnerait sur les distributions de probabilité des résultats et maximiserait l'utilité telle que définie par les trois lois. Mais l'on ne voudrait sûrement pas de robots qui empêcheraient un humain de traverser la rue sous prétexte d'une chance non nulle de danger. Cela signifie que l'utilité négative du tort causé à un humain doit être très supérieure à celle de la désobéissance, mais que chacune des utilités est finie, et non infinie.

Yudkowsky (2008) entre plus en détail dans la façon de concevoir une **IA amicale**, *friendly AI*. Il affirme que l'aspect amical (un désir de ne pas causer de tort aux humains) doit être incorporé dès le départ, mais que les concepteurs doivent reconnaître à la fois que leurs propres conceptions peuvent être imparfaites, et que le robot apprendra et évoluera au fil du temps. La difficulté est donc celle de la conception d'un mécanisme: définir celui-ci pour faire évoluer les programmes d'IA dans un système de «freins et contrepoids», et leur donner des fonctions d'utilité qui demeureront amicales face à de tels changements.

On ne peut pas se contenter de donner à un programme une fonction d'utilité statique, parce que les circonstances, et les réponses désirées à celles-ci, changent avec le temps. Par exemple, si la technologie avait permis de créer un agent intelligent superpuissant en 1800 et de lui inculquer la morale qui prévalait à l'époque, il se battrait aujourd'hui pour rétablir l'esclavage et abolir le droit de vote des femmes. En revanche, si l'on construit un tel agent aujourd'hui et qu'on lui dit de faire évoluer sa fonction d'utilité, comment être certain que son raisonnement ne sera pas: «Les humains pensent qu'il est moral de tuer les insectes qui les ennuient, en partie parce que le cerveau de ceux-ci est si primitif. Mais les cerveaux humains sont primitifs comparés à mes pouvoirs, et il doit donc être moral pour moi de tuer des humains.»

Page 1096

Omohundro (2008) émet l'hypothèse que même le plus innocent des programmes d'échecs pourrait représenter un risque pour la société. De même, Marvin Minsky a un jour suggéré qu'un programme d'IA conçu pour résoudre la conjecture de Riemann pourrait finir par accaparer toutes les ressources de la Terre pour construire des superordinateurs plus puissants qui l'aideraient à atteindre son but. Moralité: même si vous ne voulez qu'un programme qui joue aux échecs ou prouve des théorèmes, vous aurez besoin de garde-fous si vous lui donnez la capacité d'apprendre et de se modifier lui-même.

Omohundro conclut que «des structures sociales où les individus supporteraient le coût de leurs externalités négatives feraient beaucoup pour assurer un avenir stable et positif». Cette assertion semble une excellente idée pour la société en général, indépendamment de la possibilité de machines ultra-intelligentes.

Notons que la notion de garde-fous contre les changements de la fonction d'utilité n'est pas nouvelle. Dans *l'Odyssée*, Homère (autour de 700 avant J.-c.) décrivait la rencontre d'Ulysse avec les sirènes, dont le chant était si séduisant qu'il poussait les marins à se jeter à la mer. Sachant qu'il aurait cet effet sur lui, Ulysse ordonna à son équipage de l'attacher au mât, afin qu'il ne puisse pas accomplir cet acte d'autodestruction. Il est intéressant de réfléchir à la façon dont des garde-fous similaires pourraient être intégrés à des systèmes d'IA.

Enfin, considérons le point de vue du robot. Si les robots deviennent conscients, les traiter comme de simples «machines» (par exemple, les désassembler) pourrait être immoral. Les robots eux-mêmes devront agir conformément à la morale - et il faudrait les programmer avec une théorie du bien et du mal. Les auteurs de science-fiction ont abordé la question des droits et des responsabilités des robots, à commencer par Isaac Asimov (1942). Le célèbre film *A.I.* (Spielberg, 2001) s'inspirait d'une nouvelle de Brian Aldiss dont le protagoniste est un robot intelligent programmé pour croire qu'il est humain et qui ne parvient pas à comprendre que sa mère-propriétaire finisse par l'abandonner. L'histoire (et le film) plaide pour la nécessité d'un mouvement des droits civiques pour les robots.

[Fin du Chapitre]