

Tangence



L'axiomatique et *Les robots* d'Isaac Asimov Axiomatics and Isaac Asimov's Robot Trilogy

Fabrice Gagnon

Numéro 68, hiver 2002

Littérature et mathématiques

URI : <https://id.erudit.org/iderudit/008250ar>

DOI : <https://doi.org/10.7202/008250ar>

[Aller au sommaire du numéro](#)

Résumé de l'article

Établissant un certain nombre de rapports entre la méthode axiomatique, telle qu'elle est employée en mathématiques, et l'extension de son champ d'application à la littérature, telle qu'elle peut être pressentie à la lecture d'une oeuvre comme *Les robots*, d'Isaac Asimov, cet article se veut une tentative de rapprochement entre ces deux domaines. Ce rapprochement éclaire non seulement l'oeuvre étudiée, mais aussi les transformations que subit l'axiomatique lorsqu'elle est transposée en principe narratif.

Éditeur(s)

Presses de l'Université du Québec

ISSN

0226-9554 (imprimé)

1710-0305 (numérique)

[Découvrir la revue](#)

Citer cet article

Gagnon, F. (2002). L'axiomatique et *Les robots* d'Isaac Asimov. *Tangence*, (68), 89–100. <https://doi.org/10.7202/008250ar>

Tous droits réservés © Tangence, 1992

Ce document est protégé par la loi sur le droit d'auteur. L'utilisation des services d'Érudit (y compris la reproduction) est assujettie à sa politique d'utilisation que vous pouvez consulter en ligne.

<https://apropos.erudit.org/fr/usagers/politique-dutilisation/>

Érudit

Cet article est diffusé et préservé par Érudit.

Érudit est un consortium interuniversitaire sans but lucratif composé de l'Université de Montréal, l'Université Laval et l'Université du Québec à Montréal. Il a pour mission la promotion et la valorisation de la recherche.

<https://www.erudit.org/fr/>

L'axiomatique et *Les robots* d'Isaac Asimov

Fabrice Gagnon, Université Laval

Établissant un certain nombre de rapports entre la méthode axiomatique, telle qu'elle est employée en mathématiques, et l'extension de son champ d'application à la littérature, telle qu'elle peut être pressentie à la lecture d'une œuvre comme *Les robots*, d'Isaac Asimov, cet article se veut une tentative de rapprochement entre ces deux domaines. Ce rapprochement éclaire non seulement l'œuvre étudiée, mais aussi les transformations que subit l'axiomatique lorsqu'elle est transposée en principe narratif.

Première Loi: Un robot ne peut porter atteinte à un être humain ni, restant passif, laisser cet être humain exposé au danger.

Deuxième Loi: Un robot doit obéir aux ordres donnés par les êtres humains, sauf si de tels ordres sont en contradiction avec la Première Loi.

Troisième Loi: Un robot doit protéger son existence dans la mesure où cette protection n'est pas en contradiction avec la Première ou la Deuxième Loi.

Manuel de la robotique, 58^e édition (2058 apr. J.-C.)¹

Alan Sokal lui-même, dans son livre qui polémique pourtant contre la manière qu'ont eue certains intellectuels (philosophes, critiques littéraires, théoriciens des sciences sociales) d'« utiliser »

1. Isaac Asimov, *Les robots* [1950], Paris, Éditions J'ai lu, 2001, p. 5.

dans leurs écrits la science et les mathématiques, admet toutefois n'avoir en soi « rien contre la tentative d'établir des analogies entre divers domaines de la pensée humaine² ». C'est rasséréner par cet aveu que nous voulons tenter de dégager un certain nombre de rapports pouvant exister entre la « méthode axiomatique », telle qu'elle est employée en mathématiques, et l'éventuelle « extension » de son domaine d'application à la littérature, telle que nous croyons la pressentir à la lecture d'une œuvre comme *Les robots* d'Isaac Asimov.

Nous pouvons en effet lire, dans un ouvrage de l'épistémologue Robert Blanché (ayant précisément pour titre — et pour objet — *L'axiomatique*), que la « méthode axiomatique n'est pas seulement un procédé technique des mathématiciens ; on peut y trouver une illustration [...] de la manière dont procède la pensée dans la connaissance³ », processus ou « double mouvement » de la pensée qui est explicité plus loin : « Les génies sont divers, [...] tel excelle à lire dans le concret l'abstrait, tel autre à interpréter l'abstrait par le concret » (A, p. 107). Or Asimov ne s'est-il pas livré, en formulant d'abord ses Trois Lois de la Robotique pour en « dériver » ensuite quelques dizaines de nouvelles littéraires, au travail de l'axiomaticien ? Ce n'est sans doute pas sans raison qu'on a justement pu dire qu'un passage obligé dans l'élaboration de ces nouvelles avait dû être la simplification du « problème en le réduisant à ses données principales pour en chercher d'abord une solution approximative, cette situation idéale [étant] ensuite progressivement compliquée, de façon à se rapprocher de la réalité⁴ ».

Si l'on veut bien concéder qu'à l'origine des nouvelles de robots d'Asimov on retrouve des préoccupations proches de celles des axiomaticiens « scientifiques », on peut se demander dans quelle mesure les notions propres aux axiomatiques mathématiques ont pu être « importées » dans l'œuvre de cet auteur de

2. Alan Sokal et Jean Bricmont, *Impostures intellectuelles*, Paris, Éditions Odile Jacob, 1997, p. 19.

3. Robert Blanché, *L'axiomatique* [1955], Paris, Presses universitaires de France, coll. « Quadrige », 1990, p. 104. Désormais, les références à cet ouvrage seront indiquées par le sigle A, suivi de la page, et placées entre parenthèses dans le corps du texte.

4. Demètre Ioakimidis, « Introduction » à Isaac Asimov, *Le livre des robots* [1950], traduit par Pierre Billon, Paris, Éditions Opta, 1967, p. XI. Désormais, les références à cet ouvrage d'Asimov seront indiquées par le sigle LR, suivi entre guillemets du titre de la nouvelle citée, suivi de la page, et placées entre parenthèses dans le corps du texte.

science-fiction — cas particulier d'une question plus générale qu'on pourrait formuler ainsi : sur quels points les axiomatiques mathématiques et « littéraires » (si tant est que ces dernières existent, bien entendu) convergent-elles ou divergent-elles ? Pour étayer nos remarques à ce sujet, nous considérerons *i*) *axiomatique* dans son sens moderne, en tant que les mathématiques se sont constituées en systèmes hypothético-déductifs ; *ii*) que de telles axiomatiques ou systèmes tentent, dans la mesure du possible, d'être à leur base entièrement « explicites » ; *iii*) qu'en regard d'un tel système, ce n'est plus un quelconque recoupement exact avec une vérité « externe » qui compte véritablement, mais bien la cohérence interne dudit système ; *iv*) que pour atteindre son état de plus parfait achèvement, un tel système devra passer par des étapes de symbolisation et de formalisation.

Les systèmes hypothético-déductifs

Dans son ouvrage, Blanché rappelle l'évolution des axiomatiques qui, des constructions apodictiques qu'elles étaient, sont devenues des systèmes hypothético-déductifs dont l'essentiel du « fonctionnement » pourrait se résumer ainsi : « Si l'on pose, arbitrairement, tel ensemble de principes, voici les conséquences qui, formellement, en résultent » (A, p. 15). N'est-ce pas précisément ce que fait Asimov en posant « arbitrairement » les Trois Lois de la Robotique, puis en décrivant « objectivement » les implications encore cachées de ces axiomes, en « déroulant » ni plus ni moins que la multitude de situations pouvant être engendrées *par* — on serait tenté d'écrire plutôt *malgré* — ces lois⁵ ?

Mais revenons un instant sur le pouvoir apparemment discrétionnaire dont jouit l'axiomaticien : celui de doter le système qu'il entend construire des axiomes de son choix. « Arbitraire à certains égards, le choix des postulats qu'on met à la base d'une axiomatique n'est pas pour cela laissé au hasard : il demeure assujéti à des *exigences internes* diverses, plus ou moins impérieuses [nous soulignons] » (A, p. 48), affirme Blanché. Asimov en tant qu'axiomaticien n'échappe pas plus qu'un autre à cette règle. La nature des « exigences internes » qui le poussent à poser comme axiomes ses Trois Lois est-elle cependant la même que celle des exigences qui

5. On remarquera que ce « malgré » introduit une dimension proprement argumentative qui est étrangère à l'axiomatique.

président d'ordinaire à la sélection des assises d'une axiomatique mathématique ?

En un sens, oui : dans un cas comme dans l'autre, ce sont des exigences « logiques » qui peuvent expliquer le choix qui sera fait. Dans le cas des axiomatiques mathématiques, on choisira des axiomes qui permettent de construire un système non contradictoire. Pour ce qui est d'Asimov, il a choisi des axiomes qui lui permettaient de bâtir un « système » qui ne soit pas, tout au moins, en contradiction avec son projet narratif. Car, en écrivant ses nouvelles, il souhaitait avant tout prendre le contre-pied d'une tradition d'histoires de robots qui avaient toujours toutes la même conclusion : le créateur (l'homme) est détruit par sa créature (le robot). Scénario insensé aux yeux d'Asimov qui refusait de croire les hommes suffisamment illogiques pour ne pas doter chacune de leurs inventions « technologiques », potentiellement dangereuses, d'un dispositif de sécurité. Ses robots seraient munis d'un tel dispositif, n'agissant que « selon les règles logiques implantées dans leurs “cerveaux” au moment de leur construction⁶ ». On voit que le choix des axiomes posés par Asimov en tête de son système, loin d'être laissé au hasard ou à des considérations purement formelles, répond en fait à certaines exigences, du moins à celle de refléter dans ses écrits sa foi (positiviste et humaniste) dans la science.

Caractère « explicite » des axiomatiques modernes

La « nécessité » la plus impérieuse qui sous-tend la postulation des Trois Lois de la Robotique selon la *formulation* qu'on leur connaît est peut-être cependant d'ordre purement « littéraire ». Nous savons en effet que,

d'une manière générale, on appellera antérieures à un système axiomatique toutes les connaissances auxquelles ce système fait [...] appel. On remarquera que, si une axiomatique se présente comme un système purement formel, les connaissances dont elle a besoin pour se constituer sont, elles, *des notions entendues dans la plénitude de leur sens* et des thèses prises dans leur vérité matérielle. Ce recours à des connaissances antérieures, surtout s'il n'est pas déclaré, répugne à l'esprit de l'axiomatique, qui s'impose comme loi de tout expliciter, sans rien présupposer [nous soulignons] (A, p. 32).

6. Isaac Asimov, « Préface à la première partie par l'auteur », dans Isaac Asimov, *Les robots*, ouvr. cité, p. 12.

Or il est bien évident que l'« antériorité » présupposée par le système d'Asimov est extrêmement vaste, puisque ses axiomes sont formulés en langue naturelle, c'est-à-dire avec des mots dont n'est pas explicité le sens — parmi ceux, potentiellement nombreux, auxquels chacun renvoie — qui sera privilégié dans les « développements » construits à partir de ces termes premiers.

Peut-être Asimov aurait-il pu faire montre de plus de précision, par exemple en énumérant dès le départ, à la façon des axiomaticiens scientifiques, les significations qu'il entendait attribuer à chacun des termes entrant dans la composition de ses axiomes. Mais la création littéraire n'aurait-elle pas été entravée ? Asimov avoue, en effet, avoir dû s'arranger pour qu'

il subsiste juste assez d'ambiguïté dans les Trois Lois de la Robotique pour susciter les conflits et les *incertitudes nécessaires* à l'élaboration de nouvelles trames de récits. En fait, à [s]on grand soulagement, il [lui] était toujours possible de prendre sous un nouvel angle les cinq douzaines de mots composant les Trois Lois et leur *implication* [nous soulignons]⁷.

Une différence fondamentale entre les axiomatiques mathématiques et cette axiomatique « littéraire » serait donc mise en évidence par cette recherche pour les unes de la plus grande explicitation possible, la condition d'existence *sine qua non* de l'autre étant bien au contraire de contenir une part plus ou moins importante d'implicite. L'exigence de précision n'est donc pas confiée aux « axiomes » (aux Lois) mais aux « théorèmes » (les nouvelles) qui en découlent. Un bon exemple de cette nécessité de la potentialité polysémique des Trois Lois nous est donné par la nouvelle « Menteur », dans laquelle la résolution de l'intrigue passe précisément par une attention particulière accordée au(x) sens à prêter aux termes de la Première Loi :

— Un robot ne peut porter atteinte à un être humain ni, restant passif, laisser cet être humain exposé au danger.

— Merveilleusement exprimé, ironisa Calvin. Mais quel genre de danger ? Quel genre d'atteinte ?

— Mais... tous les genres.

— Exactement ! Tous les genres d'atteintes ! Mais pour ce qui est de blesser les sentiments, d'amoindrir l'idée que l'on se fait

7, Isaac Asimov, « Préface à la deuxième partie par l'auteur », dans Isaac Asimov, *Les robots*, ouvr. cité, p. 264.

de sa propre personne, de réduire en poudre les plus chers espoirs, sont-ce là des choses sans importance ou au contraire... ? (LR, «Menteur», p. 133)

La nouvelle se conclut une fois cet implicite explicité; mais s'il l'avait été dès le départ, si les termes des Trois Lois n'avaient pas laissé place à interprétation, la nouvelle n'aurait tout simplement pas eu raison d'être.

Cohérence interne versus vérité externe

Rappelons maintenant qu'une fois posés les axiomes d'une axiomatique « moderne » (donc d'un système hypothético-déductif), ce qui compte, ce n'est plus le rapport qu'entretiennent avec une vérité « externe » les résultats des raisonnements qu'on peut faire à partir d'eux, mais bien la cohérence interne du système. Les nouvelles d'Asimov semblent concorder avec une telle conception. « Raison » en présente même une illustration exemplaire. Cette nouvelle met en scène le robot Cutie qui adhère, en quelque sorte, aux principes de ces « anciennes » axiomatiques apodictico-intuitionnistes qui ne répugnaient pas à recourir au « sentiment de l'évidence » et dont le « mot d'ordre » était, *grosso modo*, celui-ci : tels axiomes « étant vrais absolument, telle proposition, que j'en déduis, est donc vraie aussi » (A, p. 18). C'est ni plus ni moins ce qu'affirme le robot en question quand il réplique à ses créateurs humains qui le somment de s'expliquer sur son refus de s'imaginer engendré par eux :

— Appelez cela de l'*intuition*. Je ne vois pas plus loin pour l'instant. Mais j'entends édifier une explication rationnelle. *Une suite de déductions logiques ne peut aboutir qu'à la détermination de la vérité*, et je n'en démordrai pas avant d'y être parvenu. [...] Tels sont les faits qui, avec le postulat *évident* qu'aucun être ne peut créer un autre être supérieur à lui-même, réduisent à néant votre stupide hypothèse [nous soulignons] (LR, « Raison », p. 68 et 73).

Son mode de raisonnement l'amène toutefois à considérer comme vraies un grand nombre de propositions que l'on sait par ailleurs fausses sur le monde qui l'entoure (il croit avoir été créé par le convertisseur d'énergie du vaisseau spatial qu'il commande, il pense être un prophète de ce dieu-convertisseur, etc.). Mais cet échafaudage chimérique n'est que d'une gravité toute relative, peut-on comprendre en conclusion de la nouvelle. En effet, cette « ancienne » conception de l'axiomatique se trouve finalement

réfutée tout simplement parce que prime sur elle, dans les faits, la conception hypothético-déductive selon laquelle seule la cohérence interne d'un système donné importe : Cutie peut bien « croire » ce qu'il veut, il n'en demeure pas moins qu'il affiche par ses actes un total respect des Trois Lois de la Robotique. La cohérence du système asimovien est maintenue, le robot accomplit la tâche qui lui a été assignée : c'est tout ce qui compte véritablement.

En fait, la cohérence interne du système des Trois Lois de la Robotique est « prouvée » dans chacune des nouvelles du *Livre des robots* — ce qui appelle, croyons-nous, d'autres commentaires. Une axiomatique mathématique est dite cohérente si elle est non contradictoire. Les mathématiciens disposent de trois méthodes pour démontrer la non-contradiction d'une axiomatique : dans le cas des axiomatiques symbolisées et formalisées, ils peuvent en faire une preuve purement mathématique ; dans les autres cas, reste la *réduction* d'une axiomatique à une théorie antérieure — la non-contradiction de celle-ci étant, elle, établie —, ou la construction d'un *modèle* correspondant à l'axiomatique étudiée. Ce dernier procédé

consiste à donner, de la théorie en question, une *réalisation* dans le monde des choses. Au lieu de ramener la théorie à une théorie antérieure dont la consistance soit mieux assurée, on descend au contraire vers le concret, on en construit un modèle physique. Comme tout ce qui est réel est *a fortiori* possible, l'existence de ce modèle garantit la consistance de l'axiomatique qui lui correspond (A, p. 50).

Sans vouloir jouer sur les mots, ne peut-on pas justement voir chaque nouvelle d'Asimov — introduisant chacune un nouveau « modèle » de robot — comme autant de modèles conférés aux Trois Lois de la Robotique ? Au cœur de l'intrigue de chacune de ces nouvelles se trouve une situation problématique où semble compromise, en première analyse, la cohérence de l'axiomatique asimovienne — par exemple dans « Cycle fermé », où un robot affectueusement surnommé Speedy (simple dérivation des lettres d'appel propres à tous les robots de modèle SPD) n'obéit pas particulièrement « promptement » à l'ordre qui lui a été donné par ses patrons humains d'extraire le sélénium d'un gisement situé à proximité de leur base mercurienne. La Deuxième Loi serait-elle transgressée ? Que se passe-t-il donc ? C'est qu'il

existe un danger quelconque dont le centre se situe dans le filon de sélénium. Il s'accroît à [...] l'approche [du robot], et à une certaine distance le potentiel de la Loi trois [...] équilibre exac-

tement le potentiel de la Loi deux [...]. Si bien qu'il [le robot] tourne en rond autour du filon de sélénium, et se tient sur le lieu des points de l'équilibre potentiel (LR, « Cycle fermé », p. 59).

Speedy tournera perpétuellement en rond, précisément parce qu'il se conforme « trop » bien aux deux dernières des Trois Lois. Les deux ingénieurs, héros de la nouvelle, qui le feront sortir de ce cercle vicieux en s'exposant volontairement à un danger — le forçant à les sauver, comme l'exige la Première Loi —, vont du même coup prouver la consistance du système des Trois Lois par la « fonctionnalité » du modèle SPD — et il en est ainsi pour les autres « modèles » présentés dans les autres nouvelles.

Notons enfin l'*isomorphie* de la quasi-totalité de ces nouvelles : comme le rappelle Blanché,

lorsque des modèles ne se distinguent [...] entre eux que par la diversité des interprétations concrètes qu'on donne à leurs termes, et coïncident exactement quand on fait abstraction de celles-ci pour s'installer sur le plan de l'axiomatique formelle, on dit qu'ils sont *isomorphes* : ils ont [...] la même structure logique (A, p. 46).

Les nouvelles d'Asimov sont elles-mêmes toutes « soutenues » par la même structure « axiomatique » (les Trois Lois de la Robotique) — sauf une : en effet, dans « Le petit robot perdu », un des « axiomes » a subi une légère modification. Les « cerveaux » des robots du modèle NS-2 n'ont pas été imprégnés de la Première Loi en entier : ils « furent construits de telle sorte qu'ils ne conserveraient que les aspects positifs de la Loi, dont les termes devenaient simplement : "Un robot ne peut porter atteinte à un être humain." Et c'est tout » (LR, « Le petit robot perdu », p. 145). Les axiomes de cette nouvelle n'étant pas *exactement* les mêmes que ceux des autres nouvelles, on ne peut dire qu'elle leur est isomorphe ; on dira simplement que, tout en leur étant très apparentée, elle constitue un système *voisin*⁸.

Symbolisation et formalisation

Le but ultime de tout axiomaticien scientifique, pourrait-on dire, est d'amener son système, dans la mesure du possible, jusqu'au stade où celui-ci se trouve formalisé. L'étape préalable à toute

8. Voir Robert Blanché, *L'axiomatique*, ouvr. cité, p. 47.

formalisation d'un système est sa symbolisation. Mais une tentative de symboliser un système tel que celui d'Asimov déboucherait-elle sur des résultats concluants? Par exemple, en employant les symboles de la logique déontique de Georg Henrik von Wright — puisqu'il « paraîtrait bien paradoxal, au moment même où l'on crée un symbolisme pour une théorie qui n'en possède pas encore, de négliger d'utiliser en même temps celui des théories qui en possèdent déjà un » (A, p. 56) et qui lui siérait bien —, on pourrait en arriver à des propositions exprimées sous la forme suivante :

- p = les robots ne portent pas atteinte aux êtres humains et ne les laissent pas exposés au danger ;
 q = les robots obéissent aux ordres des humains ;
 r = les robots protègent leur existence.

Ces propositions permettraient de reformuler les Trois Lois de la Robotique ainsi :

- Première Loi = Op
 Deuxième Loi = $Oq \rightarrow (q \rightarrow p)$
 Troisième Loi = $Or \rightarrow Op \wedge Oq$

Ces lois, on pourrait respectivement les lire à peu près comme suit : *i*) il est obligatoire que les robots ne portent pas atteinte aux êtres humains et ne les laissent pas exposés au danger ; *ii*) s'il est obligatoire que les robots obéissent aux ordres des humains, alors c'est que l'obéissance à ces ordres implique que les robots ne portent pas atteinte aux êtres humains et ne les laissent pas exposés au danger ; *iii*) l'obligation pour les robots de protéger leur existence implique l'obligation qu'ils ont (toutefois) de ne pas porter atteinte aux êtres humains et de ne pas les laisser exposés au danger *et* l'obligation d'obéir aux ordres des humains (qui elle-même implique que cette obéissance aux ordres implique que les robots ne portent pas atteinte aux êtres humains et ne les laissent pas exposés au danger).

Mais le chemin vers une éventuelle formalisation du système d'Asimov s'arrête là : il n'est pas possible de déduire « formellement » un théorème à partir de ces axiomes. Du fait qu'il s'agit d'une axiomatique « littéraire » qui table précisément sur la virtualité des significations de ses termes premiers, il apparaît en effet difficile de détacher tout à fait ses axiomes symbolisés de leur contenu interprétatif et de les laisser se « réorganiser » en un jeu purement formel. D'ailleurs, malgré les constants appels à la logique lancés par Asimov et qui sont le véritable moteur du déroulement de la trame narrative — ils prennent en gros la forme suivante :

puisque nous sommes confrontés à une situation problématique à résoudre, « quelle est la déduction logique que l'on peut tirer de [nos] observations ? » (*LR*, « Attrapez-moi ce lapin », p. 105) —, on est en droit de se demander quelles règles strictes d'inférence pourraient bien s'appliquer au système des Trois Lois.

Peut-être faut-il tout simplement constater que l'axiomatique d'Asimov constitue *une* interprétation « figée » d'une partie de ce système plus vaste qu'est la logique déontique de Wright. Cette « logique » s'attache à l'étude formelle des énoncés qui comportent des expressions comme « il est permis que », « il est interdit que », « il est obligatoire que » et « il est indifférent que⁹ »... Notons au passage que les situations problématiques que ne manque pas de présenter chacune des nouvelles du *Livre des robots* trouvent toutes leur origine dans cette mince frange d'actions qui, bien qu'« à peine » permises par les Trois Lois, sont en bout de ligne indifférentes au système puisqu'elles n'en remettent pas en cause les termes « forts » (c'est-à-dire les obligations et les interdictions).

Un regard jeté sur les autres pans de la théorie de Wright (de prime abord non directement impliqués dans la version restreinte qu'en donnent les Trois Lois de la Robotique) peut cependant permettre d'entrevoir sous un jour nouveau ce système *affaibli* qu'est, par rapport à celle-ci, l'« axiomatique » d'Asimov — comme quoi il serait vérifié une fois de plus que, pour ce qui est de notre perception des choses, « un élargissement purement formel entraîne quelquefois l'apparition de solutions significatives [pouvant] prétend[re] avoir droit de cité dans l'interprétation d'une expérience réelle non encore aperçue¹⁰ ». Ainsi, l'un des théorèmes premiers de la logique wrightienne — qui stipule ceci : $O(\sim p \rightarrow p) \rightarrow Op$ (lu : s'il est obligatoire, pour que *non-p*, que *p*, alors il est obligatoire que *p*) — semble nous inviter à une tierce lecture de la conclusion de la nouvelle « Évidence ».

Rappelons que celle-ci raconte la joute électorale opposant Quinn, présenté comme un politicien véreux, et Byerley, être d'une droiture remarquable. Quinn, pour se faire élire, entend faire la preuve que Byerley n'est pas un humain, mais bien un

9. Pour une introduction à la logique déontique, voir Georges Kalinowski, *La logique des normes*, Paris, Presses universitaires de France, coll. « Sup/Le philosophie », 1972, en particulier les p. 79-100.

10. Marie-Antoinette Tonnelat, « Limites d'extension du concept de doctrine informelle », dans Georges Canguilhem (sous la dir. de), *La mathématisation des doctrines informelles*, Paris, Hermann, 1972, p. 113.

robot androïde parfaitement ressemblant. Byerley, mis en demeure de s'expliquer lors d'un discours qu'il prononce en public, finit par frapper un auditeur. Première conclusion possible de la nouvelle : puisque Byerley a frappé un homme, il ne peut qu'être un humain, un robot ne pouvant enfreindre la Première Loi de la Robotique qui lui interdirait de poser un tel geste. La perspicace robopsychologue Susan Calvin entrevoit cependant une autre possibilité : rien n'empêche que Byerley soit bel et bien un androïde ayant frappé un *autre* androïde, aucune loi de la robotique n'interdisant aux robots de s'agresser entre eux.

Mais, à la lumière de la logique déontique, une troisième interprétation nous paraît encore valable. En effet, en apportant de légères modifications à nos définitions de départ — de telle sorte que : *i*) p devienne « robot(s) porte(nt) atteinte à n humain(s) » ; *ii*) n reçoive les valeurs des quantificateurs logiques ; *iii*) « ne pas porter atteinte à » soit entendu dans sa connotation très positive d'« apporter le bien à » —, on peut en venir à lire le théorème wrightien mentionné plus haut ainsi : s'il est obligatoire, pour que \exists robot « ne porte pas atteinte » [ou plutôt « apporte le bien »] à presque- \forall humains, qu'il porte atteinte à \exists humain, alors il est obligatoire que « ce » robot porte atteinte à « cet » humain. Puisque l'élection de Byerley, même s'il s'avérait être un robot, était plus à même d'apporter le bien à la société que celle de Quinn, il se devait — conformément à la Première Loi de la Robotique ainsi qu'au théorème de la logique déontique — de frapper l'auditeur, même si celui-ci était un humain !

Notons que la nouvelle ouvre d'ailleurs la porte à une telle interprétation. Y est en effet évoquée l'hypothèse suivante :

— Qu'arriverait-il si un robot surprenait un fou en train de mettre le feu à une maison pleine d'habitants ? Il réduirait le fou à l'impuissance, n'est-ce pas ?

— Naturellement.

— Et si la seule façon d'y parvenir était de le tuer ?

[...]

— Je pense, quant à moi, qu'il [pourrait] [...] enfreindre la Première Loi pour obéir à cette même Première Loi, *mais sur un plan plus élevé* [nous soulignons] (*LR*, « Évidence », p. 215)¹¹.

11. Près de quarante ans après la publication d'« Évidence », Asimov confirmera l'existence d'une « Loi Zéro de la Robotique » dans *Prélude à Fondation* : « Il

Ainsi donc, seul le bien-être du plus grand nombre compterait véritablement. Ceci n'est pas sans rappeler l'utilitarisme moral du philosophe anglais Jeremy Bentham et son principe du « plus grand bonheur possible pour le plus grand nombre possible de personnes¹² ». Bentham parlait d'ailleurs, à propos de cette recherche calculée des grands et petits bonheurs qu'il supposait inéluctable, d'une « arithmétique des plaisirs » — comme quoi, qu'il s'agisse de science, de littérature ou de philosophie, on n'échappe jamais tout à fait à la sphère d'influence des mathématiques !

*

Nous croyons avoir montré, par le biais de comparaisons entre ce qu'est l'axiomatique mathématique et ce que pourrait être l'axiomatique « littéraire » des *Robots* d'Asimov, la possibilité de rapprocher deux domaines de la pensée humaine apparemment aussi éloignés l'un de l'autre que le sont les mathématiques et la littérature. Ne sachant que trop bien que comparaison n'est pas toujours raison, nous espérons néanmoins ne pas nous être livré dans notre étude à ces « abus de langage » tant décriés par Sokal — bref, de n'avoir pas succombé, pour reprendre les mots de Mallarmé, au « démon de l'analogie »...

[...] semblait qu'il aurait dû y avoir une loi encore plus générale que les trois précédentes. [...] Elle s'énonce ainsi : « Zéro : un robot ne peut porter atteinte à l'humanité ni, en restant passif, laisser l'humanité exposée au danger ». La Première Loi est donc modifiée comme suit : « Un : un robot ne peut porter atteinte à un être humain ni, en restant passif, laisser un être humain exposé au danger, sauf si cette démarche est en contradiction avec la Loi Zéro ». Et les autres lois doivent être modifiées de manière analogue », dans Isaac Asimov, *Prélude à Fondation* [1988], Paris, Presses de la Cité, 1989, p. 438. Cela nous laisse croire que l'interprétation de la nouvelle « Évidence » que nous déduisons à partir d'un théorème de la logique déontique n'est pas abusive.

12. Peter Kunzmann (sous la dir. de), *Atlas de la philosophie*, Paris, Le Livre de Poche, coll. « Encyclopédies d'aujourd'hui », 1993, p. 165.