



HAL
open science

Faut-il conférer la personnalité juridique aux (robots) androïdes ?

Pierre-Emmanuel Audit

► **To cite this version:**

Pierre-Emmanuel Audit. Faut-il conférer la personnalité juridique aux (robots) androïdes ?. 2024. hal-04541223

HAL Id: hal-04541223

<https://hal.science/hal-04541223v1>

Submitted on 10 Apr 2024

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution - NonCommercial - NoDerivatives 4.0 International License

Faut-il conférer la personnalité juridique aux (robots) androïdes ?

Pierre-Emmanuel AUDIT

Maître de conférences

Université Paris-Panthéon-Assas, LDC, Paris, France

Avertissement

Cet article constitue le développement d'une communication réalisée dans le cadre du cycle de séminaires organisé en 2018 par le laboratoire de sociologie juridique de l'Université Paris 2 Panthéon-Assas sur la thématique de « La personne » et sous la direction scientifique de Dominique Fenouillet. Sa rédaction a été achevée décembre 2020.

1. La perspective d'une machine humanoïde – Il y a plus de 2300 ans, un génial inventeur de l'Antiquité nommé Philon de Byzance créait ce que l'on peut considérer comme le premier robot : un automate aux traits humanoïdes – une servante – dont la fonction était de verser du vin aux convives¹. Toutes les fois que l'un d'entre eux plaçait sa coupe dans la main gauche de la servante, celle-ci versait, depuis la cruche qu'elle tenait dans sa main droite, d'abord du vin, puis de l'eau, conformément à l'usage de l'époque qui voulait que le vin soit servi « allongé ». Pour doser son breuvage, le convive n'avait qu'à retirer sa coupe de la main de l'automate au moment opportun. Nul doute que les contemporains de cette invention furent fortement impressionnés par cette prouesse technique qui, parce qu'elle prenait des traits humains, n'en était que plus spectaculaire.

Près de 23 siècles plus tard, la société *Alderaban* a élaboré le prototype d'une version moderne de la servante de Philon de Byzance. Son nom est *Roméo* ; il s'agit d'un robot de 1,40 mètres constituant une plateforme de recherche permettant d'explorer les solutions envisageables pour l'accomplissement de tâches ménagères au profit des personnes âgées perdant leur autonomie. On pourra donc demander à Roméo de verser de l'eau ou du vin, mais également de vider le lave-vaisselle.

Après une très longue parenthèse historique, jonché certes de quelques réalisations, le fantasme du robot humanoïde – un androïde – commence donc

¹ PHILON DE BYZANCE, *Le livre des appareils pneumatiques et des machines hydrauliques*, trad. LE BARON CARRA DE VAUX, Imprimerie nationale, Lib. C. Klincksieck, Paris, 1902, § 30.

à se réaliser. En 2020, le phénomène reste anecdotique ; mais il ne faut pas s’y tromper : l’irruption des robots humanoïdes dans notre société est en cours et semble inéluctable. Pour s’en convaincre, il faut avoir vu, par exemple, ces ouvriers d’une usine japonaise faisant le matin des exercices physiques à l’unisson avec les robots humanoïdes de la chaîne de montage² ; le but est que les ouvriers aient l’impression que les robots sont des collègues³. *Pepper*, un autre petit robot humanoïde, est réceptionniste dans un hôpital belge⁴ et peut conduire les personnes qu’il rencontre en différents endroits ; il est capable de percevoir certaines émotions humaines ainsi que certains aspects du langage corporel et d’agir en conséquence. Les media se délectent également de l’avènement de robots sexuels ; sur ce point, toutefois, les réalisations restent encore primitives⁵. Et que dire des robots militaires, par exemple le très inquiétant *Atlas* de la société Boston Dynamics, capable de faire des sauts particulièrement impressionnants⁶ et laissant augurer d’un futur Terminator digne du film éponyme ?

2. La fascination exercée chez certains par la perspective de robots humanoïdes – les robots de chaîne de montage n’ont jamais suscité la même passion – est appuyée par les développements récents en matière d’intelligence artificielle. Ces développements permettent d’envisager des robots doués d’une certaine autonomie, entendons par là pour l’instant une possibilité limitée pour ces machines de prendre des décisions à partir de leur perception et de leur état interne, plutôt qu’en exécution de commandes prédéterminées⁷. Cette autonomie limitée, de type décisionnelle, permet l’avènement d’une autonomie spatiale, le robot acquérant la possibilité de se mouvoir dans des lieux qui ne figurent pas dans sa programmation initiale en décidant lui-même de la meilleure route à suivre⁸. Dans le sillage des algorithmes permettant ces prouesses, on trouve des algorithmes dotant la machine d’une certaine faculté

² V. le documentaire d’ARTE intitulé *Au Cœur des robots*, Juin 2017.

³ Le débat sur raréfaction des emplois causé la robotisation grandissante de tâches ouvrières permet de se convaincre de l’intérêt psychologique d’une telle initiative...

⁴ *AZ Damiaan hospital*, Ostende Belgique.

⁵ V. par ex. le reportage de la BBC « The future of sex ? | Sex robots and us », disponible en ligne. Sur ce thème, v. J. DIRINGER, « Sexualité et robotique. Réflexions sur les enjeux d’un antisépisme robotique », in M. TOUZEIL-DIVINA et M. SWEENEY, *Droit(s) au(x) sexe(s)*, éd. L’Épilogue Lextenso, vol XIX, 2017, p. 77. Il est en revanche indéniable que les poupées gonflables des années 1980 ont fait de sérieux progrès et laissent augurer de ce à quoi pourra ressembler un robot sexuel du futur : v. par ex. les produits de la société *Realdoll*.

⁶ V. sur Youtube la vidéo « What’s New, Atlas », le robot exécutant parfaitement un salto arrière.

⁷ Approche de l’autonomie empruntée à M. SCHEUTZ et C. CROWEL, « The Burden of Embodied Autonomy: Some Reflections on the Social and Ethical Implications of Autonomous Robots », disponible [en ligne](#).

⁸ Les rovers martiens sont capables depuis plusieurs années d’atteindre un objectif désigné sans intervention humaine ; v. par ex. M. BAJRACHARYA, M. W. MAIMONE, and D. HELMICK, « Autonomy for Mars Rovers: Past, Present, and Future » (2008), disponible [en ligne](#).

d'apprentissage, certes pour l'instant limitée, mais dont il n'est pas douteux qu'elle s'améliorera à terme de manière substantielle.

La perspective est donc celle d'une machine humanoïde – un androïde – douée d'une certaine autonomie spatiale mais aussi décisionnelle et de surcroît capable de s'éduquer au contact de son environnement : l'Homme du XXI^e siècle tente ainsi de créer une machine à son image, pour le meilleur ou pour le pire.

3. Les propositions de personnification juridique de l'androïde – Il est remarquable que cette tentative fasse l'objet d'un prolongement sur le plan juridique : certains proposent en effet que l'on confère aux robots la personnalité juridique, ce qui les extrairait de la catégorie des choses pour les hisser sur un plan comparable sinon identique au nôtre. La formulation même de cette proposition est de nature à susciter instinctivement, chez la plupart des juristes, une réaction de réticence ; une lettre ouverte la critiquant a d'ailleurs été adressée à la commission européenne⁹. Tout paraît en effet opposer les concepts de chose et de personne, et il semblera aberrant de vouloir ériger ce qui est appartenant indubitablement à la première catégorie au rang des entités appartenant à la seconde.

Le concept de personne est à la source de cette polarisation. Instinctivement, il renvoie en effet à l'idée de personne *humaine*, ce qui conduit l'esprit à effectuer une séparation très nette entre personne et chose¹⁰. Ce mouvement instinctif doit cependant être réfréné. La personnalité juridique est avant tout une construction intellectuelle et juridique, une technique susceptible d'être affectée à la réalisation d'une certaine fin¹¹. Il suffit, pour s'en convaincre, de se remémorer l'existence des personnes morales, dont le caractère utilitaire est particulièrement apparent. L'Histoire, de même que le droit comparé, offre également de nombreux exemples d'entités non-humaines, autres que des sociétés, auxquelles l'on a jugé utile de reconnaître la personnalité juridique. Dans l'Antiquité, par exemple, les Dieux se voyaient reconnaître la qualité de sujets de droit¹². Au Moyen-Âge, l'on a semble-t-il pas hésité à faire des procès aux animaux¹³, ce qui revenait à leur prêter la

⁹ <http://www.robotics-openletter.eu>

¹⁰ Séparation dont la netteté n'est finalement qu'assez récente puisque – faut-il le rappeler – l'esclave a été, du moins dans un premier temps à Rome, considéré purement et simplement comme une chose, quoique la religion et les mœurs apportaient certains correctifs (E. CUQ, *Manuel des institutions juridiques des romains*, éd. Plon et LGDJ, 2nde éd., Paris, 1928, p. 79).

¹¹ V. à cet égard R. DEMOGUE, « La notion de sujet de droit », *RTD civ.* 1909.611-655, disponible également en tiré à part sur [Gallica.fr](http://gallica.fr).

¹² R. DEMOGUE, *La notion de sujet de droit*, tiré à part cité note préc., p. 13 et art. préc. not. préc., p. 623, citant notamment F. SAVIGNY, *Traité de droit romain*, t. II, trad. Ch. GHENOUX, éd. Firmin Didot Frères, 1841, p. 246 (en réalité, p. 245).

¹³ Ainsi M. JOUSSE, *Traité de la justice criminelle*, t. IV, éd. Debure Père, 1771, fait état p. 122 et s. des procès pour bestialité, aux termes desquels on brûla vif non seulement l'humain mais aussi l'animal qui a été sa victime. On trouvera quelques éléments de bibliographie

qualité de sujet de Droit. En Inde, certaines idoles indoues ont une personnalité juridique¹⁴. En Nouvelle-Zélande, une loi du 15 mars 2017¹⁵ a doté le fleuve *Whanganui* d'une telle personnalité. La personnalité juridique est donc un instrument dont il ne faut pas oublier qu'il a un champ d'application bien plus large que les êtres humains.

On ne saurait toutefois ravalier entièrement le concept de personnalité juridique au rang de simple construction technique. Reconnaître une entité juridique sur le plan du Droit n'est en effet pas neutre : il y a au moins un semblant de jugement de valeur dans une telle consécration¹⁶. Reconnaître la personnalité juridique, c'est en effet reconnaître et consacrer l'existence singulière d'une entité ainsi que le fait qu'elle a des intérêts *propres* qu'elle doit pouvoir défendre ou, du moins, faire défendre. Lorsque l'on reconnaît par exemple la personnalité juridique à un fleuve ou à une idole, ce n'est pas à titre principal parce que cela est juridiquement commode : c'est parce que l'on souhaite consacrer juridiquement les aspirations et croyances ontologiques de certains peuples ou de certaines sociétés. La personnalité juridique est ainsi non seulement une construction technique mais également un *instrument politique*, une manière d'affirmer la prééminence et le respect qui doit être témoigné à une chose à ou à ce qu'elle représente¹⁷. Or cette considération prend un relief tout particulier lorsque la chose à l'égard de laquelle on s'interroge est une chose d'apparence humaine et à laquelle on cherche à enseigner le comportement de l'humain : on s'explique sans peine que la proposition de conférer une personnalité juridique au robot soit étroitement liée à la figure du robot humanoïde. L'intérêt technique que l'on pourrait trouver dans l'octroi de la personnalité juridique au robot doit donc être pondéré avec la charge symbolique substantielle qui s'attacherait irrémédiablement à une telle reconnaissance.

Le constat que le juriste n'est pas le seul à s'intéresser et à prendre parti sur cette question révèle d'ailleurs sans ambiguïté l'existence de cette charge symbolique. Outre le fait que le grand public ne semble pouvoir prendre en compte que cette dimension, d'autres disciplines scientifiques en étudient les ramifications dans leur domaine. Y aurait-il par exemple un intérêt *psychologique* à conférer au robot la personnalité juridique ? L'insertion de ces machines à apparence humaine dans nos sociétés humaines ne serait-elle pas facilitée par cette reconnaissance ? A force d'autonomie, le robot ne finirait-il

chez R. GARRAUD, *Traité théorique et pratique du droit pénal français*, Sirey, 3^e éd., 1913, t. 1, p. 534.

¹⁴ D. ANNOUSSAMY, « La personnalité juridique de l'idole hindoue », *Revue historique de droit français et étranger*, Quatrième série, Vol. 57 (1979), pp. 611-621.

¹⁵ Te Awa Tupua (Whanganui River Claims Settlement) Act 2017; le §14 (1) déclare expressément : « Te Awa Tupua is a legal person and has all the rights, powers, duties, and liabilities of a legal person ».

¹⁶ En ce sens, E. GAILLARD, *Le pouvoir en droit privé*, Economica, 1985, n° 286 et s.

¹⁷ L'ensemble du débat relatif à la reconnaissance de la personnalité juridique aux animaux l'illustre parfaitement.

pas par accéder à une forme d'existence dont l'importance qualitative mériterait notre reconnaissance ? Conférer au robot une personnalité juridique revêt ainsi une dimension sociologique au sens large que le juriste ne saurait ignorer lorsqu'il envisage cette question.

Cette dimension symbolique doit d'autant plus être prise en compte qu'il faut prendre garde aux implications potentielles d'une telle reconnaissance : octroyer à une machine d'apparence humaine la qualité de personne juridique est susceptible de provoquer, sur le long terme, une irrésistible attraction d'éléments de la personnalité juridique « humaine » vers celle de la personnalité juridique « robot »¹⁸ ; quand bien même la communauté des juristes s'y refuserait, la société pourrait finir par l'y contraindre¹⁹.

4. Plan – La proposition de conférer aux robots une personnalité juridique doit donc être éprouvée non seulement sur un plan juridique, mais également étudiée de manière plus générale à l'aune de l'ensemble des arguments qui ont pu être avancés en faveur de cette idée. De cet examen, il ressort qu'une telle proposition, à l'heure actuelle et même dans un futur proche, ne paraît pas justifiée (I). Il semble toutefois qu'il faille se garder de l'exclure définitivement : dans un futur lointain, les conditions pourraient se trouver réunies pour que nous y soyons fortement incités (II).

¹⁸ Pour reprendre l'expression de A. BENSOUSSAN (v. par ex. « La personne robot », *D.* 2017.2044).

¹⁹ Il faut rappeler ici la fascination qu'exerce déjà les robots humanoïdes sur une partie de la population humaine : les personnes revendiquant déjà le droit de se marier avec un robot ne symbolisent-elles pas parfaitement les prémises d'une telle dérive ?

I.

UNE PROPOSITION DIFFICILEMENT JUSTIFIABLE AUJOURD'HUI

5. Plan – La proposition de conférer la personnalité juridique au robot étonne : pourquoi faire ? Quel intérêt ? On ne peut s'empêcher, en première approche, de penser qu'elle est le produit d'une fascination aussi étrange qu'excessive pour un objet auquel on a conféré une apparence externe humaine. La lecture de la littérature sur le sujet permet toutefois de se faire une idée un peu plus précise ; elle révèle un certain nombre d'arguments assez disparates au premier abord, mais qu'un effort de classification permet de ramener à trois catégories. Certains justifient l'octroi de la personnalité juridique au robot sous un angle utilitariste : la personnalité juridique serait une technique juridique qui pourrait se révéler utile pour appréhender et traiter un certain nombre de problèmes en rapport avec le robot (A). D'autres justifient l'octroi de la personnalité juridique au robot en se fondant sur la liberté décisionnelle que ceux-ci sont en train d'acquérir (B). D'autres, enfin, énoncent une série d'arguments qui, en fin de compte, peuvent se ramener à l'idée que conférer au robot une telle personnalité sert l'intérêt de l'Homme (C).

A. L'argument utilitariste

6. Champ – L'argument utilitariste trouve sa source dans l'idée que le concept de personnalité juridique est avant tout un instrument technique qui pourrait se révéler utile et même indispensable pour appréhender certains problèmes juridiques en relation avec le robot. La réflexion a principalement porté, à ce jour, sur l'intérêt que pourrait présenter cette technique pour appréhender la responsabilité civile extracontractuelle dérivant de l'action du robot ; seule²⁰ cette dernière nous retiendra ici.

²⁰ Dans le champ contractuel, l'hypothèse dans laquelle un robot causerait un dommage au créancier d'une prestation pour l'exécution de laquelle le robot est employé s'appréhende facilement : le prestataire devra contractuellement répondre du dommage (la dernière proposition de loi portant réforme la responsabilité civile datée du 29 juillet 2020 proposant d'ouvrir au créancier une option entre le domaine contractuel et délictuel pour la réparation de son dommage corporel, v. le projet d'art. 1233 civ.) ; conférer la personnalité juridique au robot ne modifierait en rien la responsabilité du prestataire, pas plus que la personnalité juridique d'un employé n'exonère par principe son employeur de devoir rendre compte d'une inexécution contractuelle.

Personne ne semble par ailleurs suggérer que le robot, personne juridique, doive répondre *pénalement* de son comportement : l'aspect punitif inhérent à la sanction pénale paraît en effet dénué de sens à l'égard d'un robot, constat que l'on peut rattacher, juridiquement, à l'absence de discernement de la machine (en ce sens que le discernement fait défaut à l'imputation d'une responsabilité pénale aux robots, G. COURTOIS, « Robot et

7. Difficulté : l'autonomie du robot – Le robot appartenant pour l'instant indiscutablement au royaume des choses, la responsabilité dérivant des dommages qui lui sont imputables a naturellement vocation à être appréhendée en fonction des régimes juridiques applicables à cette catégorie juridique²¹, lesquelles conduiront à imputer ces dommages à une personne physique ou morale. L'idée que ces régimes juridiques puissent poser difficulté et qu'il faille même les évincer n'a germé qu'à partir du moment où l'on a commencé à prendre en considération la question des robots « autonomes ». Mais que faut-il, au juste, entendre par cette expression ? Il est possible, de manière grossière, d'en proposer quatre acceptions, lesquelles renvoient en fin de compte à des degrés d'autonomie différents.

8. Première acception possible – Dans une première acception, la plus faible, il est possible de définir l'autonomie d'un robot comme l'aptitude de celui-ci « à exécuter des tâches prévues à partir de son état actuel et de sa capacité de détection, sans intervention humaine »²². Cette définition de l'autonomie est tellement large qu'elle semble pouvoir même s'appliquer à un robot d'usine assemblant des pièces d'une voiture, c'est-à-dire réalisant de manière automatique une série de tâches prédéfinies. Elle n'est donc pas particulièrement utile ici.

9. Deuxième acception possible – Dans une deuxième acception, légèrement plus restrictive, on pourra dire d'un robot qu'il est doué d'autonomie dès lors que son programme a été conçu pour lui permettre de

responsabilité », in *Les robots, objets scientifiques, objets de droits*, éd. Mare & Martin, pp. 129-156, n° 2.1) ; pour un parallèle entre responsabilité pénale des personnes morales et responsabilité pénale du robot, H. DE VAUPLANE, « La personnalité juridique des robots », *Mélanges Jean-Jacques Daigre*, Joly Éditions, Lextenso, 2017, pp. 79-84., spéc. p. 83.

Enfin, à moins de prêter au robot une volonté et une liberté de l'accomplir qui devrait être respectée en elle-même, l'idée de permettre par ailleurs au robot de *conclure en son nom* des actes juridiques ne semble avoir d'intérêt que si cette possibilité sert ou facilite l'intérêt d'une autre personne située à l'arrière-plan du robot, en premier lieu son propriétaire. Il a ainsi pu être suggéré qu'il serait utile de conférer au robot une personnalité juridique afin d'en faire un mandataire. L'idée est étrange : a-t-on par exemple éprouvé le besoin de conférer à un distributeur automatique la personnalité juridique pour en faire le « mandataire » de l'entreprise qui vend par cet intermédiaire boissons et nourriture ? Il y aurait au demeurant paradoxe à vouloir conférer à quelque chose la personnalité juridique pour *ensuite* lui appliquer une technique juridique dont l'objet est précisément de rendre la personne du mandataire transparente au profit de la personnalité du mandant au nom et pour le compte de qui l'acte est conclu ! Tout au plus pourrait-on élaborer des règles spécifiques régissant certaines interactions de l'homme avec le robot lorsqu'elles ont une finalité juridique, un peu comme l'on réglemente aujourd'hui les contrats conclus par voie électronique et conclus en traitant avec un système de traitement automatisé des données (sur « l'intermédiation contractuelle d'un robot », v. G. LOISEAU et M. BOURGEOIS, « Du robot en droit à un droit des robots », *JCP G* 2014, doct. 1231, spéc. n° 14 à 18).

²¹ Il restera toujours possible d'engager la responsabilité pour faute d'une personne physique ou morale à raison de la négligence ou de la faute intentionnelle que cette personne aura commise et qui a conduit à ce que le robot soit une source de dommages. Cette responsabilité ne semble pas devoir appeler de développements particuliers.

²² Définition de la norme ISO 13482:2014 « Robots et composants robotiques — Exigences de sécurité pour les robots de soins personnels ».

gérer et de surmonter un éventail de situations et de difficultés susceptibles de se poser à lui dans l'accomplissement d'une tâche prédéfinie. L'exemple type est celui d'un rover martien, qui, une fois qu'on lui a indiqué une destination, est capable de gérer seul les obstacles qui se trouvent sur sa route, en décidant de rouler au-dessus ou de les contourner. Ces types de programmes, destinés en fin de compte à assurer l'exécution d'une tâche en apparence simple, sont déjà très complexes ; mais la manière dont ils vont gérer les difficultés rencontrées n'en reste pas moins réductible à l'exécution d'un programme soigneusement conçu pour envisager toutes les hypothèses que les concepteurs du programme ont pu imaginer.

10. Troisième acception possible – Dans une troisième acception, l'autonomie de la machine s'entendra de surcroît de sa possibilité de *tirer les conséquences* de ses constatations ou erreurs afin de ne plus reproduire telle stratégie de résolution d'un problème s'avérant non fonctionnelle et ainsi d'améliorer de manière itérative sa manière d'exécuter la tâche correspondante ou la vitesse de son accomplissement. Par exemple, le rover martien, après s'être engagé dans une voie, pourrait constater que *ce type de voie*, qui ne figurait pas préalablement dans sa base de données, est en réalité impraticable et *retenir cette information* pour le futur afin de privilégier d'autres routes dont l'expérience lui aura montré qu'elles sont sans surprises ou facilement surmontables. Ce mode de fonctionnement, par lequel la machine apprend par elle-même comment aboutir de manière plus efficace à un certain résultat, renvoie aux processus qualifiés de manière générique par l'expression « machine learning » (apprentissage automatique).

11. Quatrième acception possible – Dans une dernière acception, forte et encore lointaine, on dira d'un robot qu'il est doué d'autonomie s'il est doté d'une capacité, postérieurement à sa mise en service, de se doter *par lui-même* nouvelles compétences, de nouvelles stratégies de résolution d'un problème ou de nouveaux savoirs. Le robot devient alors capable d'apprendre et de développer de nouveaux comportements. Imaginons par exemple un robot spécialisé dans le service des personnes âgées. Un soir, alors que la personne âgée et affaiblie est couchée, le robot constate avec ses capteurs qu'elle présente un taux de déshydratation alarmant. Le robot se met alors en quête d'eau conformément à son programme mais constate qu'aucune eau ne sort du robinet (coupure d'eau) et qu'il n'y a pas d'eau dans le réfrigérateur. Jugeant le besoin d'eau critique, le robot prend l'initiative de se rendre dans le magasin où la personne âgée lui demande ordinairement de faire des courses afin de se procurer de l'eau en urgence ; constatant que les portes sont fermées, il accède alors sur Internet aux horaires des magasins avoisinants et se rend dans le plus proche. Dans cet exemple, le robot aura développé seul une stratégie de résolution d'un problème pour lequel il n'a pas été spécifiquement programmé (surmonter une coupure d'eau dans une hypothèse jugée critique).

Relèvera également de cette acception forte de l'autonomie d'un robot la capacité de ce dernier d'apprendre de nouvelles compétences par apprentissage direct auprès d'un humain, compétence qu'il pourra ensuite « cultiver » et

mettre en œuvre seul ; dans ce genre d'hypothèses, le robot acquiert en effet également des capacités qui n'ont pas été prévues par son concepteur et pouvant étendre son champ d'action. On pourrait ici faire l'analogie avec le dressage d'un animal.

12. Hypothèses problématiques – C'est la faculté d'apprentissage de la troisième et surtout de la quatrième acception de l'autonomie, au sens de la typologie proposée précédemment, qui est fondamentalement à l'origine de la difficulté que pose l'appréhension de la responsabilité dérivant de l'action d'un robot « intelligent »²³. Comme l'expose un auteur, « l'apprentissage réduit les possibilités de traçabilité, ce qui peut poser des problèmes pour déterminer l'origine de la responsabilité dans les prises de décision du système »²⁴. Cette difficulté atteint son paroxysme avec les *black box algorithms*, qui tirent leur nom de ce qu'il est impossible de rendre compte de la logique qui a conduit la machine à tel ou tel résultat ou action.

13. Cette possibilité pour un robot de déterminer dans une certaine mesure *lui-même* sa conduite invite à réexaminer tant la possibilité²⁵ que la légitimité²⁶ de faire peser la responsabilité dérivant de ses actions aux personnes à qui l'on s'adresse traditionnellement pour répondre des dommages causés par une chose : en l'espèce l'« utilisateur » du robot, son propriétaire ou encore son producteur, cette dernière catégorie englobant le concepteur de l'algorithme qui a abouti au comportement dommageable. Si l'on accepte que l'autonomie dont sera douée le robot interdit de rechercher la responsabilité de ces personnes, il n'y a alors plus qu'un pas à faire pour aboutir à l'idée qu'il faut imputer cette responsabilité *directement au robot*, ce qui suppose d'en faire un sujet de droit et donc de lui conférer une personnalité juridique.

14. Il incombe donc de commencer par examiner si l'état actuel des textes applicables en France conduit ou non à l'impossibilité de rechercher la responsabilité d'une personne humaine – ou morale – à raison des dommages que pourront causer le robot autonome (1°). Au terme de cet examen, dont on verra qu'il ne conduit pas à une impasse mais soulève quelques difficultés, on essaiera de montrer que la personnalité juridique ne paraît pas être un remède technique utile ou satisfaisant pour les résoudre (2°), avant de conclure que la

²³ Cette difficulté est notamment citée dans le rapport de la Commission européenne du 19 février 2020 intitulé *Rapport sur les conséquences de l'intelligence artificielle, de l'internet des objets et de la robotique sur la sécurité et la responsabilité*, COM/2020/64 final, spéc. p. 7 et s.

²⁴ R. CHATILA, « Intelligence artificielle et robotique : un état des lieux en perspective avec le droit », *Dalloz IP/IT* 2016.284.

²⁵ Certains évoquent tout simplement une faille dans le système de responsabilité : v. A. MATTHIAS, « The responsibility gap : Ascribing responsibility for the actions of learning automata », *Ethics and Information Technology*, Vol. 6, Issue 3, 2004, pp. 175-183.

²⁶ C'est alors la question du *fondement* la responsabilité des différentes personnes impliquées qui est posée.

meilleure solution serait sans doute l'adoption d'un régime spécifique de responsabilité (3°).

1°) Le recours aux régimes juridiques existants

15. Plan – Comment appréhender, en l'état du droit positif, la responsabilité issue des dommages attribuables à l'action d'un robot autonome et « intelligent » ? On fera ici abstraction des actions qui dériveraient de l'éventuelle existence d'un contrat²⁷ pour se concentrer tout d'abord sur l'examen du régime des produits défectueux (a), qui transcende les deux ordres de responsabilité et intéresse la responsabilité des « producteurs », avant d'étudier le régime de la responsabilité du fait des choses (b), qui aboutit à déclarer responsable les personnes réputées « gardiennes » du robot.

a) La responsabilité du fait des produits défectueux

16. Origine – Le régime des produits défectueux, que l'on trouve aux actuels articles 1245 et suivants du Code civil issus de la transposition tardive et imparfaite par la loi du 19 mai 1998²⁸ de la directive européenne du 25 juillet 1985 sur les produits défectueux²⁹, semble constituer un terrain naturel pour appréhender certaines des conséquences dommageables attribuables au fait d'un robot. Son application aux robots est cependant de nature à susciter quelques difficultés, comme l'a relevé la Commission européenne dans son rapport sur les implications de l'intelligence artificielle, de « l'Internet des choses » et de la robotique en matière de sécurité et de responsabilité publié en février 2020³⁰.

17. Responsables – On rappellera qu'aux termes de l'article 1245 du Code civil, « Le producteur est responsable du dommage causé par un défaut de son produit, qu'il soit ou non lié par un contrat avec la victime »³¹.

²⁷ *Supra*, note 20.

²⁸ L. n° 98-389 du 19 mai 1998, remanié par L. n° 2004-1343 du 9 décembre 2004 pour tenir compte de certains griefs européens. A l'origine figurant aux articles 1386-1 et suivants du Code civil, le dispositif se trouve, depuis la renumérotation opérée par la réforme du droit des contrats de 2016, aux articles 1245 et s. du Code civil.

²⁹ Dir. 85/374/CEE du Conseil des Communautés européennes du 25 juillet 1985 relative au rapprochement des dispositions législatives, réglementaires et administratives des États membres en matière de responsabilité du fait des produits défectueux.

³⁰ [Rapport sur les conséquences de l'intelligence artificielle, de l'internet des objets et de la robotique sur la sécurité et la responsabilité](#), 19 février 2020, accompagnant le [Livre blanc sur l'intelligence artificielle](#), publié à la même date, tous deux disponibles en ligne.

³¹ Doit également répondre des défauts du produit celui qui se présente comme producteur en apposant son nom, sa marque ou un autre signe distinctif ainsi que l'importateur du produit dans la communauté européenne ; l'article 1245-6 civ. prévoit au demeurant, à défaut pour le producteur de pouvoir être identifié, la responsabilité subsidiaire de celui qui a vendu ou fourni le produit à titre professionnel.

18. Fabricant de la structure physique du robot – La *structure physique* qui constitue le robot étant incontestablement un produit au sens de ce régime de responsabilité, son fabricant répondra, sauf cause exonératoire applicable, des dommages que le robot causera dès lors qu’il peut être démontré que sa chose « n’offre pas la sécurité à laquelle on peut légitimement s’attendre » (art. 1245-3 civ.), indépendamment du fait que le dommage soit imputable ou non au logiciel qui l’anime.

19. Concepteur de la partie logicielle du robot ? – Mais peut-on dire que le concepteur de la composante logicielle, à le supposer différent du producteur de la structure physique du robot, est « fabricant d’une partie composante » au sens de l’article 1245-5 alinéa 1^{er} du Code civil³² et qu’à ce titre il devra répondre de la défectuosité de la chose qu’il a conçue aux côtés du fabricant du produit fini³³ ? La réponse à apporter est malaisée du fait qu’elle est à beaucoup d’égards dans la dépendance d’une autre question, qui a toujours embarrassé et que la Commission elle-même a récemment qualifié de « question ouverte »³⁴ : celle de savoir si un logiciel *seul et en lui-même*³⁵ peut relever du régime des produits défectueux³⁶.

20. Arguments *contra* – Pour justifier leur exclusion du champ de la directive, on peut avancer que le dysfonctionnement d’un logiciel révèle moins le dysfonctionnement d’un produit que celui d’une prestation de service³⁷, l’algorithme susceptible d’opérer en son sein devant recevoir un traitement identique³⁸. Mais la source essentielle de la gêne éprouvée tient surtout à l’incorporalité d’un logiciel, l’application de la directive aux biens incorporels n’allant pas d’elle-même³⁹. Celle-ci emploie en effet des concepts ne

³² « *Est producteur, lorsqu’il agit à titre professionnel, le fabricant d’un produit fini, le producteur d’une matière première, le fabricant d’une partie composante* » (nous soulignons).

³³ Le producteur qui a réalisé l’incorporation est en effet solidairement responsable des dommages causés par la partie composante (art. 1245-7 civ.). Le producteur de la partie composante ne sera toutefois pas responsable s’il « établit que le défaut est imputable à la conception du produit dans lequel cette partie a été incorporée ou aux instructions données par le producteur de ce produit » (art. 1242-10 al. 7 civ.).

³⁴ *Livre blanc sur l’intelligence artificielle* (dans sa version anglaise), préc. note 30, p. 14.

³⁵ La version anglaise du livre blanc, qui utilise le terme « stand-alone » (p. 14), est beaucoup plus claire que la version française qui, employant le terme « autonome » (p. 16), prête à confusion dans le contexte de l’intelligence artificielle. L’expression *seul et en lui-même* paraît, faute d’être parfaitement claire, la plus proche du sens entendu.

³⁶ Pour un exposé assez complet des opinions sur la question, v. E. MONTERO, *La responsabilité civile du fait des bases de données*, Travaux de la Faculté de Droit de Namur, n° 21, Presses Universitaires de Namur, 2002, spéc. n° 95 et s.

³⁷ O. SABARD, « Les produits, Rapport français », in *La responsabilité du fait des produits défectueux*, Recueil des travaux du Groupe de Recherche Européen sur la Responsabilité Civile et l’Assurance (GRERCA), IRJS Éditions, 2013, pp. 93-97 spéc. p. 95.

³⁸ L. GODEFROY, « Les algorithmes : quel statut juridique pour quelles responsabilités ? », *Communication Commerce électronique*, n° 11, nov. 2017, Études 18, spéc. p. 13.

³⁹ Sur cette idée, v. not. A. LUCAS, « La responsabilité du fait des choses immatérielles », *Mélanges Pierre Catala*, pp. 817-826.

convenant guère à l'immatériel ; de surcroît, une telle application paraît n'avoir jamais été discutée lors de son élaboration⁴⁰ et conduirait dans l'absolu à des conséquences que personne n'est prêt à admettre⁴¹.

21. Arguments pro – Pour défendre en revanche l'application de la directive aux logiciels, on peut souligner qu'est qualifiable de produit au sens de la loi « tout bien meuble » (art. 1245-2 civ.), aucune exigence de corporalité n'étant formulée, de sorte qu'il ne faut pas distinguer là où la loi ne distingue pas⁴² ; on peut également tirer les logiciels vers la corporalité⁴³ en observant qu'ils reposent nécessairement sur un *support corporel* et ne pourraient provoquer des dommages matériels s'ils étaient purement incorporels⁴⁴.

22. Que choisir ? – En vérité, il est peu probable que le législateur européen ait eu les logiciels à l'esprit lorsqu'il a imaginé la directive : non seulement son champ d'application naturel est à l'évidence celui des objets de grande consommation susceptibles de causer des dommages physiques, mais encore la perspective qu'un logiciel puisse causer de tels dommages devait sembler bien lointaine en 1985. Cela n'a pas empêché la Commission, lorsqu'elle est a été interrogée trois ans plus tard sur cette question, de se prononcer pour l'applicabilité de la directive au logiciel en relevant que celle-ci s'appliquait à « tout meuble »⁴⁵ ; elle soutint toutefois par la suite une position non compatible avec cette prise de position initiale en énonçant dans son livre vert sur les produits défectueux que les produits doivent être définis comme des « biens corporels meubles »⁴⁶. La Garde des sceaux en fonction durant l'année 1998 devait quant à elle, interrogée par une question parlementaire formulée à ce sujet à propos de la loi transposant la directive sur

⁴⁰ V. J.-S. BORGHETTI, *La responsabilité du fait des produits, étude de droit comparé*, LGDJ, 2004, n° 494 ; sur la problématique spécifique du logiciel, v. n° 495.

⁴¹ A. LUCAS, art. préc. note 39, spéc. n° 24, citant de nombreuses prestations de services qui devraient alors être considérées comme des produits.

⁴² F.-X. TESTU et J.-H. MOITRY, « La responsabilité du fait des produits défectueux. Commentaire de la loi 98-389 du 19 mai 1998 », *D. Aff. supplément* au n° 125, 16 juillet 1998, pp. 3-18, spéc. n° 12.

⁴³ Bien que l'électricité ne soit pas à proprement parler un bien incorporel, il n'est pas anodin que la directive l'ait expressément désigné comme un produit (art. 2) : la volonté d'évacuer tout débat à ce sujet fournit un indice sur la volonté du législateur européen de ne pas exclure les biens qui tendent vers l'incorporalité.

⁴⁴ V. J.-S. BORGHETTI, art. préc. citant J.-P. TRAILLE et G. CATTANEO dans un sens voisin.

⁴⁵ Question écrite n° 706/88 du 5 juillet 1988, réponse du 15 novembre 1988, JOCE n° C 114 du 8 mai 1989, p. 12 (en ce sens également, F.-X. TESTU et J.-H. MOITRY, « La responsabilité du fait des produits défectueux. Commentaire de la loi 98-389 du 19 mai 1998 », art. préc. note 42, spéc. n° 12) La Commission semble toutefois avoir changé d'avis par la suite puisqu'elle n'évoque que les meubles corporels dans son livre vert de 1999 sur les produits défectueux.

⁴⁶ Livre vert *La responsabilité civile du fait des produits défectueux*, Commission des Communautés européennes, 28 juillet 1999, COM(1999) 396 final, p. 28.

les produits défectueux, apporter une réponse⁴⁷ conforme⁴⁸ à la prise de position initiale de la Commission, mais en relevant que les hypothèses concrètes dans lesquelles un logiciel serait « à l'origine directe d'une atteinte à la sécurité des personnes et des biens » seraient « pour le moins résiduelles ».

23. Bien que la portée de cette dernière affirmation dépende dans l'absolu de ce que l'on peut qualifier d'« origine directe », il est facile d'observer qu'aujourd'hui le rôle et place des logiciels se sont considérablement accrues dans la société et que les hypothèses dans lesquelles leur défaillance ou « défaut de sécurité » (au sens informatique, c'est-à-dire leur vulnérabilité à des attaques ou à des virus) est susceptible de porter atteinte à la sécurité des personnes se sont considérablement accrues. On peut à cet égard formuler un argument de politique juridique : l'importance et l'ubiquité des logiciels dans le monde d'aujourd'hui font qu'indépendamment des prévisions des rédacteurs de la directive à cet égard, il pourrait simplement être *opportun* de les soumettre à celle-ci. La directive fournirait en effet, dans l'attente d'un éventuel instrument rénové ou dédié, un cadre juridique préétabli au problème de l'encadrement de la responsabilité à raison des dommages causés par un défaut de sécurité attribuable à un logiciel et *a fortiori* une intelligence artificielle défectueuse. A tout le moins pourrait-on – et cela couvrirait un grand nombre d'hypothèses pratiques dont celle que nous examinons à titre principal ici – attirer le concepteur du logiciel dans le champ du régime de la directive en le considérant comme « fabricant d'une partie composante » toutes les fois qu'il a conçu un logiciel destiné à être intégré à un produit corporel, à en régir le fonctionnement et qu'un défaut du logiciel conduit ce produit à causer un dommage. C'est peu ou prou l'opinion timide émise par la Commission européenne dans son récent rapport sur les implications de l'intelligence artificielle en matière de sécurité et de responsabilité ; on y lit qu'un « logiciel qui commande le fonctionnement d'un produit concret *pourrait* être considéré comme une partie ou un composant de ce produit »⁴⁹ mais qu'il conviendrait dans le futur de clarifier le champ du concept de produit pour mieux refléter les technologies émergentes et s'assurer qu'une indemnisation est toujours possible pour les dommages causés par des produits qui sont défectueux à cause de leur logiciel ou d'autres caractéristiques numériques.

Il serait donc opportun nous semble-t-il d'attirer le concepteur de la couche logicielle d'un robot dans le champ de la directive en le qualifiant de fabricant d'une partie composante ; il ne faut toutefois pas perdre de vue qu'indépendamment de la position qu'il convient d'adopter sur cette

⁴⁷ Question écrite n° 15677, réponse publiée au JOAN Q, 24 août 1998, p. 4728 et *D. Aff.* 17 décembre 1998, n° 142 p. 1999.

⁴⁸ La Garde des sceaux jugea que cette loi avait vocation à « englober l'intégralité de la catégorie juridique des meubles, à laquelle appartiennent les logiciels ».

⁴⁹ *Rapport sur les conséquences de l'intelligence artificielle, de l'internet des objets et de la robotique sur la sécurité et la responsabilité*, préc. note 23, spéc. p. 16.

question⁵⁰, le fabricant du produit fini et pleinement matériel que constitue le robot reste lui indubitablement responsable en première ligne vis-à-vis de la victime des dommages qui peuvent être imputés à un défaut de sécurité de celui-ci.

24. Autonomie et défaut de sécurité – Hors celle des responsables, l'autre difficulté essentielle que pose l'application de la directive aux robots autonomes tient dans l'appréciation de la condition de défectuosité. Les algorithmes conférant au robot un certain degré d'autonomie rendent en effet délicate l'appréciation de la sécurité qui peut être légitimement attendue de leur mise en œuvre. Il faudra sans doute ici distinguer selon la couche logicielle impliquée.

25. Couche logicielle « primaire » – Au niveau le plus élémentaire du fonctionnement d'un robot, on trouvera l'ensemble des instructions et routines⁵¹ régissant son fonctionnement et ses comportements de base. Par exemple, les instructions primitives régissant le mouvement de sa main droite, de ses doigts (s'il en est pourvu) ou de sa tête ; à un degré de complexité supérieure, des routines impliquant l'exécution, dans une séquence complexe et ordonnée, de ces instructions primitives et lui conférant, par exemple, la capacité de se saisir d'un objet ou de se mouvoir d'un endroit à un autre. Pour ces routines simples, assimilables à un programme traditionnel⁵², la caractérisation d'un défaut de sécurité ne sera pas trop difficile : si l'on peut montrer, par exemple, que le robot bute systématiquement face à telle catégorie d'obstacles, ou écrase la main de son interlocuteur chaque fois qu'il la serre, on pourra en conclure qu'il ne présente pas la sécurité à laquelle peut légitimement s'attendre : on peut en effet légitimement s'attendre à ce qu'un robot susceptible de déplacement et jugé suffisamment avancé pour être commercialisé soit capable de contourner les obstacles qui se situent sur sa route et n'écrase pas les mains des humains qu'il serre. Ce raisonnement peut, dans une certaine mesure, être étendu à l'exécution de tâches plus complexes⁵³ dès lors que la manière d'exécuter ces tâches a été explicitement programmée dans le robot au moment de sa commercialisation.

26. Couches logicielles avancées – Les hésitations vont naître à l'égard des couches logicielles plus avancées et plus particulièrement de celles conférant au robot des capacités d'apprentissage au sens de la quatrième acception proposée précédemment⁵⁴, qu'il s'agisse d'*auto-apprentissage* ou d'apprentissage *sous la direction d'humains* lui enseignant de nouvelles

⁵⁰ Une autre question se pose : que décidera-t-on si le logiciel est *open source* ? (v. G. COURTOIS, « Robot et responsabilité », art. préc. note 18, spéc. p. 140).

⁵¹ En informatique, une routine peut être décrite comme un ensemble d'instructions dédiées à l'accomplissement d'une tâche déterminée et formant donc une unité fonctionnelle.

⁵² On entend ici un programme entièrement écrit par un humain, dont on peut lire et examiner le code source afin d'en comprendre le fonctionnement et en déterminer les failles.

⁵³ Par exemple : l'exécution d'une recette de cuisine ; la composition d'un bouquet de fleur.

⁵⁴ *Supra*, n° 11.

compétences⁵⁵. Pourra-t-on imputer au producteur le dommage ayant pour origine un comportement résultant de tels apprentissages ?

27. Défaut inexistant à la mise en circulation – L’objection immédiate que feront valoir les producteurs à toute tentative d’engagement de leur responsabilité pour ce type de dommage est qu’à supposer qu’il puisse être considéré que le comportement dommageable de la machine soit constitutif d’un défaut de sécurité au sens de la directive, son article 7.b, transposé à l’article 1242-10 2° du Code civil, prévoit que la responsabilité de plein droit du producteur cesse s’il est en mesure de démontrer que, « compte tenu des circonstances, il y a lieu d’estimer que le défaut ayant causé le dommage *n’existait pas au moment où le produit a été mis en circulation par lui ou que ce défaut est né postérieurement* »⁵⁶. L’argument paraît imparable, puisque l’hypothèse est celle d’un comportement appris postérieurement à la mise en circulation de la machine ; et la conséquence juridique inéluctable... à moins de considérer que la *possibilité même* pour le robot d’apprendre des comportements susceptibles de causer des dommages aux personnes ou des biens *soit en elle-même constitutive d’un défaut de sécurité* existant au jour de la mise en circulation. Cette proposition ambitieuse et intellectuellement soutenable neutraliserait cette cause d’exonération ; mais est-elle réaliste ?

28. Faculté d’apprentissage d’un comportement dommageable = défaut de sécurité ? — Plusieurs arguments conduisent à douter de la validité d’une telle idée. Rappelons déjà à titre liminaire que le simple fait qu’un produit soit susceptible de causer un dommage, et le cause effectivement, n’est aucunement constitutif, à lui seul, d’un défaut quant à la sécurité à laquelle on peut légitimement s’attendre⁵⁷. Pour juger du défaut de sécurité, « il doit être tenu compte de toutes les circonstances et notamment (...) de l’usage qui peut en être raisonnablement attendu et du moment de sa mise en circulation » (art. 1245-3 civ.), l’information de l’utilisateur étant par ailleurs déterminante⁵⁸. Il faut donc se garder de tout lien d’automaticité entre dommage causé par un robot autonome et responsabilité de son producteur.

29. Ce rappel effectué, l’objection essentielle que l’on peut faire à la thèse selon laquelle serait constitutif d’un défaut de sécurité la possibilité même pour un robot autonome d’apprendre des comportements susceptibles de causer des

⁵⁵ Cette distinction va certainement se réduire à mesure que les algorithmes d’auto-apprentissage progresseront, l’algorithme intégrant l’enseignement des humains comme une simple composante de son apprentissage général.

⁵⁶ Nous soulignons.

⁵⁷ V. par ex. G. VINEY, P. JOURDAIN, S. CARVAL, *Traité de droit civil, Les régimes spéciaux et l’assurance de responsabilité*, LGDJ, 4^e éd., 2017, n° 21 ; Y. BUFFELAN-LANORE et V. LARRIBAU-TERNEYRE, *Droit civil, Les obligations*, 15^e éd. par V. LARRIBAU-TERNEYRE, Sirey, 2017, n° 3076.

⁵⁸ Comp., par ex., Cass. civ. 1^{ère}, 7 novembre 2006, n° [05-11.604](#) : responsabilité du fait des produits défectueux encourue à raison l’insuffisance de l’information fournie à l’utilisateur quant à l’ensemble des précautions qu’il fallait prendre pour éviter les brûlures causés par le contact prolongé du béton frais avec la peau.

dommages est de politique juridique : elle aboutit concrètement à condamner l'avènement de telles machines, le risque semblant *inhérent* au processus utilisé pour les doter des qualités d'adaptabilité et d'évolutivité qui feront leur intérêt. Concevoir un algorithme permettant à un robot d'apprendre par lui-même, c'est en effet nécessairement laisser à ce programme une latitude pour créer des règles que *l'on ne peut prévoir à l'avance*. Le fait que le robot puisse, à l'issue de ce processus, avoir appris ou inventé des comportements présentant un risque du point de vue de la sécurité et imprévisibles pour le concepteur de l'algorithme apparaît ainsi la rançon des qualités dont on veut le doter⁵⁹. Y voir un défaut de sécurité reviendrait donc à limiter sérieusement le développement et la commercialisation de ces machines, ce qui paraît au rebours de l'intention des pouvoirs publics.

30. Le fait que la possibilité d'apprentissage par le robot autonome d'un comportement susceptible de causer des dommages paraisse inhérente au procédé utilisé pour le doter d'une telle capacité conduit à la conclusion que les producteurs devraient échapper à responsabilité non parce que le défaut à l'origine du dommage n'existait pas au jour de la mise en circulation mais parce que plus fondamentalement l'on ne saurait, pour en revenir aux critères de l'article 1245-3 du Code civil, « légitimement attendre » d'un robot doué de facultés d'apprentissage avancées la même sécurité qu'un robot qui n'en est pas doué, cela au regard des usages qui en sont « raisonnablement attendus ».

31. Limite à l'absence de responsabilité – Il ne s'agit pas toutefois de suggérer que les producteurs ne devraient jamais être responsable du dommage résultant d'un apprentissage autonome du robot : le critère de *l'attente légitime* permet de circonscrire le champ de leur responsabilité. A cet égard, les producteurs devraient voir leur responsabilité engagée toutes les fois qu'une victime parviendrait à démontrer que, si l'auto-apprentissage qu'a fait le robot était par nature imprévisible, *le potentiel dommageable d'une action particulière* que celui-ci a adopté en exécution de cet apprentissage pouvait être prévu, de sorte qu'il y a un défaut de sécurité imputable à ses concepteurs à ne pas le lui avoir interdit. L'exemple plus évident est celui d'un mouvement ou d'une action du robot dont l'effet direct serait de causer une lésion corporelle à un humain ou un animal. On peut en effet légitimement attendre des concepteurs des logiciels du robot qu'ils aient interdit à la machine toute action dont le résultat direct serait de léser un tissu humain ou animal et que cette directive gouverne tant les actions initialement prévues par le programme que celles qui résulteraient de son auto-apprentissage ou même de l'apprentissage « inculqué » au robot postérieurement à sa mise en service.

⁵⁹ Cette considération devient dirimante s'agissant des algorithmes d'apprentissage pour lesquels les erreurs sont une *condition même* de l'apprentissage de la machine : ce sont les algorithmes d'apprentissage dits « par renforcement ». V. A. MATTHIAS, art. préc. note 25, spéc. p. 179 et 182. L'algorithme apprend en effet directement dans le milieu dans lequel il opère, à l'aide d'un système d'essais et d'erreurs à partir duquel il modifie par tentatives successives son approche d'un problème. On ne peut alors plus soutenir que l'erreur de la machine est attribuable à une erreur du programme et donc du programmeur.

32. Cas des enseignements dispensés au robot – Le cas du comportement dommageable résultant d'un enseignement dispensé au robot postérieurement à sa mise en circulation devrait obéir aux mêmes principes : il ne saurait être imputé aux concepteurs du robot – car on ne saurait légitimement s'attendre à une telle sécurité – sauf à démontrer que le résultat auquel a conduit cette apprentissage pouvait être d'emblée identifié comme constituant un risque de sécurité évident.

33. Participation de la victime au dommage – Si l'on devait néanmoins considérer que la possibilité même pour un robot de faire l'apprentissage d'un comportement dommageable soit constitutif d'un défaut de sécurité, les producteurs pourraient au moins, lorsque cet apprentissage a été dispensé par la victime, compter de l'article 1245-12 du Code civil qui prévoit « La responsabilité du producteur peut être réduite ou supprimée, compte tenu de toutes les circonstances, lorsque le dommage est causé conjointement par un défaut du produit et par la faute de la victime ou d'une personne dont la victime est responsable ». Il en ira différemment lorsque l'enseignement dommageable aura été dispensé par un autre que la victime, du fait de la règle de l'article 1245-13 du Code civil prévoyant que « La responsabilité du producteur envers la victime n'est pas réduite par le fait d'un tiers ayant concouru à la réalisation du dommage »⁶⁰. Cette règle ne semble pas opportune en matière de robotique autonome puisqu'il pourrait paraître plus juste que le producteur voit sa responsabilité d'emblée réduite lorsque le dommage a pour origine des enseignements dommageables fournis par un tiers à sa machine ; mais son application suppose déjà de considérer que la possibilité d'apprentissage d'un comportement aux effets potentiellement dommageables constituait un défaut de sécurité ce qui, on l'a vu, ne paraît pas opportun en premier lieu.

34. Risque de développement – Les raisonnements qui précèdent ne sont-ils pas réductibles à la fameuse exception de risque de développement, consistant pour le producteur à démontrer que « l'état des connaissances scientifiques et techniques, au moment où il a mis le produit en circulation, n'a pas permis de déceler l'existence du défaut » (art. 1245-10 4° civ.) ? Bien que ce moyen de défense soit souvent cité en matière de robotique autonome, il ne semble pas constituer l'échappatoire que beaucoup sont tentés d'y voir. En effet, l'état des connaissances scientifiques et techniques *d'ores et déjà existant* permet de se convaincre qu'en permettant à l'algorithme d'un robot d'apprendre par lui-même de son environnement et d'inventer des réactions supposées adaptées à des situations pour lesquelles il n'a jamais été programmé, tout comme en acquérant de nouvelles compétences par le biais de l'enseignement d'un humain, il existe une possibilité que l'algorithme apprenne ou élabore un comportement dommageable. Le risque de sécurité qui en résultera ne sera donc aucunement attribuable aux lacunes de nos connaissances scientifiques ou techniques actuelles, mais à la *nature même du programme* que l'on a mis à la base de son fonctionnement : un programme

⁶⁰ Art. 1245-13 civ.

d'apprentissage dont les infinies possibilités de résultat ne pourront jamais être intégralement ni *testées*, ni *prédites*⁶¹ par son concepteur avant la commercialisation du robot.

35. Responsabilité du fait des produits défectueux : bilan – La responsabilité du fait des produits défectueux permettra très certainement la condamnation des producteurs du robot toutes les fois qu'il s'agira d'obtenir réparation des dommages causés par l'exécution de routines préprogrammées dans le robot et dont on peut montrer qu'elles comportent une erreur de programmation aboutissant de manière systématique ou prévisible à mettre en jeu la sécurité des personnes et des biens. Il semble *a priori* bien plus délicat – et à certains égards inopportuns – d'engager la responsabilité du producteur lorsque le dommage aura été causé par un comportement appris par le robot postérieurement à sa commercialisation, cela non pas tant du fait de l'inexistence du défaut au jour de la circulation que parce que le risque d'apprentissage d'un tel comportement paraît consubstantiel au procédé permettant de doter le robot de facultés d'apprentissage ; cela, pour autant que ne puisse être démontré que le comportement dommageable du robot pouvait facilement être pré-identifié et interdit à celui-ci dans sa couche logicielle primaire⁶². C'est en fin de compte le critère de la sécurité que l'on peut « légitimement attendre » qui doit gouverner la responsabilité des producteurs de ces machines et l'on ne saurait, au regard de l'usage qui en est « raisonnablement attendu », légitimement attendre la même sécurité de robots serviles que de robots dotés de facultés d'apprentissage avancées.

36. Le recours à la responsabilité des produits défectueux présente pour la victime d'un robot autonome certaines limites : il ne permet que d'atteindre les personnes *réputées productrices* du robot au sens de cette législation⁶³ ; il suppose que le dommage puisse être attribuable à un défaut de sécurité de la machine ; il ne permet d'être indemnisé d'une atteinte aux biens que si le préjudice est supérieur à un certain montant⁶⁴. La vénérable responsabilité du fait des choses est-elle de nature à ouvrir d'autres perspectives pour les potentiels victimes de ces robots intelligents ?

b) La responsabilité du fait des choses

37. La responsabilité du fait des choses, dont on rappellera avec ironie qu'elle a précisément été inventée pour appréhender les dommages causés par des machines, permet d'atteindre une autre classe de personnes en relation avec les robots : celles qui seront réputées être leurs *gardiens*. La difficulté

⁶¹ Sur cette non-prédictibilité, A. MATTHIAS, art. préc. note 25.

⁶² La solution la plus englobante étant d'interdire au robot non pas tant certains comportements que d'aboutir, *in fine*, à certains résultats.

⁶³ Et à défaut, son fournisseur (art. 1245-6 civ.).

⁶⁴ Art. 1245-1 al. 2 civ. (500 euros).

essentielle, ici, est de savoir qui l'on va réputer être gardien des robots doués d'autonomie⁶⁵.

38. Principes gouvernant la garde — Rappelons succinctement que la Cour de cassation, dans son célèbre arrêt Franck, a consacré la conception matérielle de la garde sur une toile de fond de garde juridique : si l'arrêt énonce qu'est gardien celui qui a l'usage, le contrôle et la direction de la chose, le propriétaire de celle-ci n'en reste pas moins le gardien de principe et il lui incombera de démontrer, s'il entend échapper à cette responsabilité, qu'il a transféré les pouvoirs caractérisant la garde matérielle à un tiers, ce transfert pouvant résulter d'un acte purement matériel ou d'un titre juridique. Comment cette grille d'analyse s'applique-t-elle à la responsabilité dérivant de l'utilisation d'un robot autonome ?

39. Garde juridique – Des principes gouvernant la garde, on peut déjà tirer une certitude : par principe, la victime des faits dommageables d'un robot doué d'autonomie pourra toujours s'adresser à son *propriétaire* pour demander réparation des dommages que celui-ci aura causés. C'est en effet tout l'intérêt de la garde juridique posée à titre de principe que de désigner un responsable par défaut. Les difficultés que pose l'application de la responsabilité du fait des choses aux robots autonomes résident donc dans la manière dont il convient d'apprécier à leur égard les critères traditionnels du transfert de la garde, différentes personnes pouvant, selon la conception que l'on retiendra, se retrouver désignées gardiennes.

40. Garde purement matérielle du robot autonome ? — La conception *purement* matérielle de la garde – avoir *physiquement* l'usage, le contrôle et la direction de la chose – ne paraît pouvoir que très exceptionnellement recevoir application dans le cas d'un robot susceptible de se mouvoir et d'exécuter des tâches complexes de manière autonome. Il sera en effet rare que le dommage se produise à l'occasion d'une manipulation matérielle du robot par une personne humaine⁶⁶ ; c'est *sous son propre mouvement* que le robot causera le dommage.

41. La jurisprudence ne s'en est toutefois jamais tenue à une conception aussi stricte du concept de garde. La maîtrise matérielle de la chose n'a pas à être contemporaine du dommage et il suffira que la personne ait à un moment acquis, sans les avoir perdus au moment du dommage, les *pouvoirs* d'usage, de contrôle et de direction sur la chose pour en être réputée gardienne⁶⁷. La question devient donc celle de savoir ce qui pourrait caractériser l'acquisition et l'exercice de tels pouvoirs sur un robot autonome.

⁶⁵ Le robot étant doué de mouvement, c'est à cette occasion que le dommage se produira très fréquemment de sorte que son rôle actif sera présumé.

⁶⁶ Auquel cas, sa responsabilité sera très aisément engagée.

⁶⁷ Rapp., par ex., le cas d'un caddy abandonné par un client dans un port et ayant endommagé l'hélice d'un navire, Cass. civ. 2^e, 13 janvier 2012, n^o [11-11.047](#) (inédit).

42. La garde ramenée à l'existence d'un ordre – La réponse la plus simple que l'on peut imaginer est de faire dépendre la garde de l'existence d'un ordre donné par un humain⁶⁸ au robot⁶⁹. On pourra alors dire de cet humain qu'il a bien « l'usage » du robot puisqu'au moment du dommage, le robot était occupé à exécuter une tâche sur l'impulsion et pour le compte de cet humain ; il est également possible de défendre l'idée qu'à cet instant, il en a également le « contrôle », puisque le robot est *a priori* asservi exclusivement à ses ordres, du moins pendant leur exécution. Quant à la caractérisation de la « direction », la jurisprudence pourra soit l'assimiler à la simple existence de l'*ordre* donné au robot (idée d'« acte directif »), soit la caractériser de manière plus matérielle dans le constat que la direction dans laquelle le robot se mouvait lui a été intimée. Si, par exemple, une personne âgée ordonne à son robot domestique d'aller lui chercher une bouteille d'eau dans la cuisine, l'indication de la cuisine remplirait le critère de la direction⁷⁰.

La jurisprudence pourrait donc, nous semble-t-il, trouver dans l'*ordre* donné au robot les critères de la garde matérielle sans s'attirer trop de critiques. Il faut toutefois noter la part de fictivité que ce raisonnement comporte. En effet, entre l'ordre donné au robot et l'accomplissement effectif de la tâche, il y a la réalisation matérielle de celle-ci : et cette partie capitale *sera entièrement effectuée par l'algorithme* gouvernant les actions du robot. Ne doit-on pas, dès lors, considérer que c'est l'algorithme gouvernant le robot et donc son concepteur qui a la maîtrise de celui-ci et donc sa garde à l'instant auquel le dommage se produit ? A tout le moins semble-t-il en avoir le « contrôle » en cet instant⁷¹. Et que dire lorsque le robot aura pris lui-même l'initiative de son déplacement, ce que l'on pourra attendre d'un robot autonome ? Par exemple, lorsqu'il « décidera tout seul » d'aller chercher une bouteille d'eau pour la personne âgée au service de laquelle il est affecté parce qu'il a détecté sur sa peau des signes de déshydratation et que la bouteille qui est à côté de son lit se trouve être vide ?

Imputer la garde matérielle du robot à celui qui lui donne un ordre ne va donc pas nécessairement de soi, car cette personne n'a pas la maîtrise concrète de la manière dont le robot s'exécute et donc la latitude nécessaire pour pouvoir

⁶⁸ On n'ose imaginer ce qu'il faudra décider lorsque l'ordre aura été donné par un autre robot....

⁶⁹ En ce sens, évoquant le contrôle « intellectuel » du robot, G. COURTOIS, « Robot et Responsabilité », art. préc. note 18, p. 132 et 134 et s.

⁷⁰ Ce raisonnement pourrait fonctionner également si aucune indication géographique n'est donnée mais que le robot connaît le lieu dans lequel il doit s'acquitter de sa tâche ; mais on se rapproche alors du critère tenant à simplement à l'existence d'un ordre donné au robot.

⁷¹ *Contra* : G. LOISEAU, « La personnalité juridique des robots : une monstruosité juridique », *JCP G* 2018.597 : « si des prises de décision automatique non-maîtrisées par l'utilisateur peuvent survenir dans l'exécution des tâches, elles n'excluent pas le contrôle au besoin largement entendu comme la capacité d'arrêter le fonctionnement de la machine ».

prévenir le dommage⁷², ces deux considérations allant de pair ; et les critères d'usage, de contrôle et de direction paraissent encore moins applicables lorsque le robot aura causé le dommage sous l'empire d'une directive qu'il s'est lui-même donnée ou dont il a pris l'initiative.

43. Garde de la structure et du comportement ? – Pour appréhender le fait qu'un algorithme intervienne dans la production du dommage, il a pu être suggéré⁷³ de remettre au goût du jour la fameuse distinction de Goldman entre garde de la structure et garde du comportement. Cette distinction, on le sait, a été proposée⁷⁴ pour appréhender la responsabilité dérivant des objets présentant un vice interne ou doués d'un « dynamisme propre », telle la bouteille d'oxygène liquide qui a donné lieu à sa consécration jurisprudentielle⁷⁵. Le dommage causé par une mauvaise manipulation de la bouteille a vocation à être imputé à celui qui avait l'usage, le contrôle et la direction du *contenant* ; tandis que celui qui a pour source la maîtrise inadéquate de son *contenu* sera imputé à celui qui a procédé à son conditionnement, et donc le fabricant⁷⁶. La légitimité de cette distinction, fort ingénieuse, tient à l'idée sous-jacente au critère de la garde matérielle selon laquelle la responsabilité doit être imputée à celui qui avait le pouvoir de prévenir le dommage. On voit bien, dès lors, quelle est la transposition suggérée : si la responsabilité est imputable au logiciel gouvernant les actions du robot, c'est au concepteur de ce logiciel, gardien de l'aspect « structurel » du robot, que la responsabilité doit être imputée et non à son utilisateur.

44. Critique de la transposition – Pour tentante qu'elle soit, la transposition au robot autonome de la distinction entre garde de la structure et garde du comportement ne s'opère que malaisément et sa mise en œuvre se heurterait partiellement à un obstacle jurisprudentiel.

45. Confusion entre structure et comportement – La première difficulté que pose cette transposition tient à ce que distinguer entre *structure*

⁷² V. par exemple, faisant dépendre la garde de cuve de la fourniture d'une information adéquate au regard de la dangerosité de l'installation des moyens de prévenir le dommage, Civ. 2, 12 octobre 2000, n° [99-10.734](#). Plus encore, un auteur observe que vouloir imputer la garde du robot à un gardien motif pris du contrôle qu'il exerce sur la machine est incohérent dans la mesure où cela « contredit la fonction substantielle à tout robot intelligent : celle de servir l'homme en prétendant justement le soulager d'un contrôle qui lui reviendrait sinon » (C. COULON, « Du robot en droit de la responsabilité civile : à propos des dommages causés par les choses intelligentes », *RCA* avril 2016 Étude n° 6, spéc. n° 8).

⁷³ Par ex. G. LOISEAU et M. BOURGEOIS, « Du robot en droit à un droit des robots », *JCP* 2014 doct. 1231, spéc. n° 13.

⁷⁴ B. GOLDMAN, *De la détermination du gardien responsable du fait des choses inanimées*, préf. P. Roubier, Sirey, Paris, 1947.

⁷⁵ Cass. civ. 2^{ème}, 5 janvier 1956, n° [56-02.126](#) et [56-02.138](#), *Bull. civ.* 2 n° 2, *GAJC*, n° 205, *Dalloz*, 13^e éd., 2015.

⁷⁶ Dans la conception de Goldman, la garde de la structure devait peser sur le propriétaire. V. G. VINEY, P. JOURDAIN et S. CARVAL, *Traité de droit civil, Les conditions de la responsabilité*, LGDJ, 4^e éd., n° 691.

et *comportement* est ici assez délicat dans la mesure où, dans l'hypothèse d'un robot doué d'une capacité de mouvement autonome, *c'est la structure qui commande le comportement concret de la chose* : c'est bien sous l'effet de son programme que le robot se meut dans l'espace, décide d'éviter des obstacles et règle les menus détails qui permettent l'accomplissement concret de ses tâches⁷⁷. Chercher le gardien du comportement concret du robot conduit donc au gardien de la structure, ce qui semble nier l'intérêt de la distinction.

46. Garde du comportement et ordre donné au robot – Il est évidemment possible d'adopter une conception plus large de la « garde du comportement » en réputant cette garde résulter de l'*ordre* donné par l'utilisateur au robot⁷⁸. Mais une autre difficulté surgit alors ; comment fera-t-on la part entre ce qui est imputable à l'*ordre* qui a été donné au robot et ce qui est imputable au *logiciel* de la machine ? Si par exemple la machine blesse une personne en se rendant d'un point A à un point B, dira-t-on que cela est imputable au logiciel sous l'action duquel la machine se meut ou au fait que la machine ne s'est trouvée là que parce que l'utilisateur lui avait intimé de réaliser ce déplacement ? De surcroît, si l'on considère que le dommage est imputable au programme qui avait la maîtrise concrète du déplacement du robot, et donc à la garde de la structure, on peut alors se demander quand est-ce qu'un dommage sera concrètement jugé attribuable à la garde du comportement, en dehors de l'hypothèse d'un ordre malveillant de l'utilisateur. Si, à l'inverse, on décide que le dommage relève de la garde du comportement en ce que c'est sous l'impulsion d'un ordre que le robot se trouvait en action au bénéfice de son utilisateur, on sera porté symétriquement à se demander, lorsqu'un ordre aura été émis, quand cette responsabilité devra être jugée par exception être imputable au concepteur de l'algorithme gouvernant les actions du robot⁷⁹. Cette difficulté fait implicitement apparaître une question de politique juridique : la responsabilité devrait-elle plutôt peser sur celui dont on peut présumer qu'il avait les meilleurs moyens de prévenir le dommage lié au mouvement du robot (concepteur du logiciel) ou sur celui pour le bénéfice duquel la chose opérait lorsque le dommage s'est produit (utilisateur) ?

⁷⁷ Il est également possible que l'inverse se produise : lorsqu'un algorithme d'apprentissage interne à la machine a déterminé son comportement, « c'est le comportement de l'entité matérielle ou immatérielle, au cours de l'apprentissage, qui influe sur la structure de la programmation » (L. GODEFROY, « Les algorithmes : quel statut juridique pour quelles responsabilités ? », art. préc. note 38, *Communication Commerce Électronique*, Lexis-Nexis, 2017, Étude n° 18, spéc. n° 24 note 64.

⁷⁸ C'est la conception de MM. LOISEAU et BOURGEOIS, « Du robot en droit à un droit des robots », art. préc. note 20, spéc. n° 13 : l'utilisateur resterait responsable au titre de la garde du comportement toutes les fois « que le comportement dommageable du robot aurait son origine dans les instructions de son utilisateur (...) ».

⁷⁹ G. COURTOIS, art. préc. note 18, sans soulever ce problème, suggère de laisser au concepteur du robot la charge des dommages survenant « suite à un comportement aberrant sur lequel l'utilisateur n'a eu aucun contrôle » (p. 136). Mais outre qu'il est alors probable que l'on tombe dans le champ du défaut de sécurité et donc de la responsabilité du fait des produits défectueux (*infra*, n° 16 et s.), quand dira-t-on que l'utilisateur avait « un certain contrôle » sur la machine ?

47. Absence possible du fondement de la distinction – Une troisième difficulté posée par cette transposition est que, dans l'esprit de Goldman, la distinction de la structure et du comportement n'avait de légitimité qu'en ce qu'elle permettait d'imputer le dommage à un « acte humain imparfait » ou un comportement humain défectueux tel qu'un défaut d'entretien de la chose⁸⁰. En conséquence, lorsque des « imperfections, bien que connues, [sont] absolument inévitables en l'état de la technique au jour où ces imperfections causent un dommage », on ne saurait présumer le gardien responsable « si l'on entend tenir ferme sur la position d'une responsabilité digne de ce nom, c'est-à-dire d'une responsabilité ayant pour fondement un comportement humain dans lequel on puisse reconnaître la cause du dommage » de sorte que l'« on sera amené à donner au gardien la faculté de s'exonérer de cette présomption lorsqu'il peut établir que le dommage a eu pour cause un vice de la chose échappant par sa nature, à tout contrôle. A vrai dire, dans cette hypothèse, personne ne sera gardien responsable »⁸¹. La description de Goldman semblant pouvoir s'appliquer sans peine à l'hypothèse d'un dommage qui aurait été causé par le comportement qu'un robot aurait appris par lui-même ou spontanément adopté, on en déduira que l'éminent auteur se prononcerait sans doute, dans une telle hypothèse, pour l'inapplicabilité de sa distinction ainsi que pour une absence de responsabilité.

48. Responsabilité du fait des choses et incorporalité – Une quatrième difficulté que pose cette transposition, rarement soulevée, est qu'imputer au concepteur du logiciel au titre de la garde de la structure la responsabilité d'un dommage causé par le robot reviendrait à appliquer la responsabilité du fait des choses à une entité incorporelle ; or l'on sait que la doctrine n'est pas favorable à cette idée⁸² et que la jurisprudence ne semble jamais lui avoir donné suite⁸³. Les difficultés et débats que suscite l'application de ce régime de responsabilité aux choses incorporelles sont très comparables à ceux portant sur l'application de la directive sur les produits défectueux aux logiciels, ce qui n'est pas étonnant car la difficulté initiale – le caractère incorporel de la chose envisagée – est identique. On trouvera un argument supplémentaire pour s'opposer à l'application de la responsabilité du fait des choses à l'immatériel dans le constat que l'actuel proposition de réforme de la responsabilité civile limite expressément son application aux choses corporelles⁸⁴. Il ne serait

⁸⁰ V. par ex. B. GOLDMAN, thèse préc. note 74, n° 113 spéc. p. 188 et n° 116 p. 191, cet acte humain imparfait étant pour l'auteur en toute probabilité la cause du dommage puisqu'« une chose inanimée, au mouvement comme au repos, est toujours en réalité, inerte – entendons qu'elle est dépourvue de toute initiative propre » (p. 188).

⁸¹ B. GOLDMAN, thèse préc. note 74, n° 117 pp. 192-193.

⁸² V. not. A. LUCAS, art. préc. note 39, spéc. I.

⁸³ La seule décision qui a pu sembler donner corps à l'idée est une décision du TGI de Paris du 27 février 1991 (*JCP G 1992.II.21809* note Ph. LE TOURNEAU) toujours citée à cet égard et qui a décidé qu'une image de télévision pouvait constituer une chose au sens de l'ancien art. 1384 civ.

⁸⁴ Proposition de loi préc. note 20, spéc. article 1242 al. 1^{er} : « Chacun est responsable de plein droit du dommage causé par le fait des choses corporelles qu'il a sous sa garde ».

toutefois pas impossible à la jurisprudence, aujourd'hui comme demain, de neutraliser la difficulté en « corporalisant » le logiciel sur le fondement du constat que celui-ci produit des effets concrets à travers l'action du robot ; il s'agit là, on se souvient, de l'une des pistes également suggérée pour l'application de la directive aux logiciels.

49. Conflit avec la directive sur les produits défectueux – S'agissant enfin de l'obstacle jurisprudentiel au recours à la distinction entre garde de la structure et du comportement, il tient à ce que la CJCE⁸⁵, que la Cour de cassation n'a eu d'autre choix que de suivre⁸⁶, a décidé que la possibilité laissée à la victime par l'article 13 de la directive sur les produits défectueux de se prévaloir « du droit de la responsabilité contractuelle ou extracontractuelle ou au titre d'un régime spécial de responsabilité »⁸⁷ ne permet *pas* à la victime d'invoquer un régime national de responsabilité qui reviendrait à mettre en cause le défaut de sécurité d'un produit⁸⁸. Or, rechercher la responsabilité du concepteur du logiciel du robot au titre de la garde de sa structure paraîtra souvent difficilement dissociable d'une démarche visant à imputer à son logiciel un défaut de sécurité : c'est bien en effet la sécurité des produits qui était surtout en cause lorsqu'on invoquait la garde de la structure de la chose⁸⁹. Ce raisonnement, qui aboutit à évincer la responsabilité du fait des choses dès lors que serait invoquée la garde de la structure (concrètement soit par le propriétaire utilisateur victime, soit par celui-ci en défense à une action en responsabilité intentée contre lui), ne vaut toutefois pleinement que si le concepteur du logiciel est aussi le fabricant du robot. Si en effet ce concepteur est un tiers, le raisonnement ne vaut alors qu'autant que l'on accepte de considérer qu'il peut être réputé *fabricant d'une partie composante* au sens de la directive⁹⁰ et ainsi soumis à son régime juridique, lequel entraîne l'éviction

⁸⁵ CJCE, 25 avr. 2002, aff. C-183/00, *González Sánchez c/ Medicina Asturiana SA*, D. 2002.2462 note C. LARROUMET et 2002.2458 obs. J. CALAIS-AULOY (« Menace européenne sur la jurisprudence française concernant l'obligation de sécurité du vendeur professionnel ») ; v. spéc. le considérant n° 30 : la directive ne peut être interprétée comme laissant la possibilité « de maintenir un régime général de responsabilité du fait des produits défectueux différent de celui prévu par la directive ».

⁸⁶ Com. 26 mai 2010, n° [08-18.545](#), *Bull. civ.* IV n° 99, *JCP* 2010 n° 1015 note Ph. STOFFEL-MUNCK ; Civ. 1^{ère}, 10 décembre 2014, n° [13-14.314](#), *CCC* 2015 com. 58 obs. L. LEVENEUR.

⁸⁷ Article dont le contenu se trouve transposée l'article 1245-17 c. civ., lequel ne reprend toutefois pas la précision selon laquelle le régime spécial doit exister au moment de la notification de la directive.

⁸⁸ Sur cette limitation appliquée au robot, v. par ex. N. NEVEJANS, *Traité de droit et d'éthique de la robotique civile*, LEH éd., 2017, spéc. n° 716 et s.

⁸⁹ Ce qui explique pourquoi laquelle la distinction entre structure et comportement est tombée en quasi désuétude depuis l'avènement des dispositions européennes sur les produits défectueux tout comme l'obligation de sécurité que la jurisprudence avait introduite comme équivalent à la directive postérieurement à son entrée en vigueur et antérieurement à sa transposition législative. V. toutefois Civ. 1^{ère}, 30 septembre 2009, n° [08-12.625](#), dont une interprétation *a contrario* implique la survivance du concept de garde de la structure ainsi que Civ. 2^{ème}, 28 février 2013, n° [11-27.699](#) utilisant expressément le concept.

⁹⁰ *Supra*, n° 19 et s.

du droit commun. A défaut, on peut concevoir que le concepteur du logiciel resterait responsable au titre d'une forme de garde de la structure vis-à-vis de la victime sur le fondement de la responsabilité du fait des choses ; la victime pourrait également se fonder sur la faute contractuelle du concepteur du logiciel vis-à-vis de l'intégrateur du robot pour l'assigner sur le terrain de la responsabilité délictuelle pour faute⁹¹. Ce cas de figure illustre les méfaits qui résulteraient du refus de réputer fabricant d'une partie composante au sens de la directive le concepteur du logiciel du robot. Non seulement l'action de la victime s'en trouverait compliquée, mais encore le créateur du logiciel échapperait au régime de la directive alors qu'il est pleinement impliqué dans le défaut de sécurité que l'on reprochera à l'intégrateur du robot ; et l'on ne voit pas pourquoi le concepteur de la partie logicielle d'un robot devrait ou non se trouver soumis à la directive selon qu'il a fabriqué ou non la partie matérielle de celle-ci.

50. Quand bien même il serait décidé que le concepteur du logiciel relève du champ de la directive, on pourrait concevoir que la responsabilité du fait des choses – si tant est que l'on accepte de l'appliquer à une chose immatérielle – lui reste applicable au titre d'une forme de garde de la structure lorsque le dommage sera imputable à son programme sans pour autant pouvoir être qualifié de défaut de sécurité au sens de la directive. Les hypothèses concrètes paraissent de prime abord difficiles à imaginer, mais s'avèreraient moins rares si l'on adoptait la conception du défaut de sécurité suggérée plus haut s'agissant des robots doués de facultés d'apprentissage : ne serait pas constitutif d'un défaut de sécurité le fait pour la machine d'avoir adopté, en application de ses facultés d'apprentissage, un comportement dommageable qui ne pouvait être raisonnablement pré-identifié et interdit à la machine en amont. Cela reviendrait alors à faire assumer au concepteur du logiciel, sur le fondement de la responsabilité du fait des choses, les conséquences de la dangerosité imprévisible et pour ainsi dire « normale » du robot autonome, non attribuables à un défaut de conception constitutif d'un défaut de sécurité.

51. Cette dernière conception, qui laisserait entière la question de savoir la part de responsabilité qu'il conviendrait de réserver à l'utilisateur et attribuable à une forme de garde du « comportement » du robot, se heurte à une objection d'ordre pratique et de politique juridique : serait-il véritablement légitime de faire peser sur les fabricants et concepteurs de logiciels de robots autonomes une responsabilité objective à raison des dommages que peuvent causer les algorithmes gouvernant ces robots et que l'on ne pourrait qualifier de défaut de sécurité au sens de la directive ? On se rappelle, certes, que l'un des acquis de l'arrêt *Jand'heur* a été que le vice de la chose n'est aucunement une condition de la mise en œuvre de la responsabilité du fait des choses. Mais cela ne revient-il pas à faire reposer sur les épaules des producteurs les

⁹¹ Selon le fameux principe réaffirmé par la Cour de cassation dans l'arrêt Ass. Plén., 13 janvier 2020, n° [17-19.963](#), *Sucrierie de Bois rouge* : « le tiers à un contrat peut invoquer, sur le fondement de la responsabilité délictuelle, un manquement contractuel dès lors que ce manquement lui a causé un dommage » (n° 24).

conséquences dommageables liées à une utilisation « normale » de leur produit ? Les risques qu'engendre cette utilisation normale du robot, non attribuable à une défectuosité du logiciel, ne devrait-il pas être imputé à celui *qui en profite ordinairement*, c'est-à-dire son utilisateur ? C'est en effet lui qui, par l'utilisation d'un robot doué d'autonomie et susceptible d'adopter par apprentissage des comportements dommageables, a créé la situation dommageable. On peut en tous cas imaginer que les fabricants de robots combattront énergiquement le moment venu une construction juridique qui aboutirait à un tel résultat.

Une manière de justifier tout à l'inverse une responsabilité extensive de *l'utilisateur* de robot serait de raisonner par analogie avec la responsabilité du fait des animaux.

52. Analogie avec le fait des animaux ? – Pour envisager la responsabilité du fait des robots, il a pu être suggéré d'étendre⁹² ou du moins de raisonner comme en matière de fait des animaux et donc au regard de l'actuel article 1243 du Code civil disposant que « Le propriétaire d'un animal, ou celui qui s'en sert, pendant qu'il est à son usage, est responsable du dommage que l'animal a causé, soit que l'animal fût sous sa garde, soit qu'il fût égaré ou échappé ». Tout comme le gardien d'un animal doit répondre des dommages que celui-ci cause sous l'effet de son propre dynamisme, l'utilisateur d'un robot devrait assumer les conséquences que ce dynamisme cause aux tiers. L'analogie est fructueuse car le robot autonome est susceptible de se mouvoir par lui-même et d'adopter, s'il est doué de fonctions avancées d'autonomie, des comportements non prévisibles pour son utilisateur, voire d'échapper à son contrôle. L'emprunt de ce régime juridique permettrait au demeurant de justifier plus facilement que l'utilisateur réponde du robot même s'agissant des actions que celui-ci a entreprise sous l'effet d'une règle qu'il a lui-même élaboré.

Si l'analogie est pertinente, elle ne paraît pas présenter d'intérêt d'un point de vue technique. Outre qu'il paraît peu probable que la jurisprudence décide elle-même d'étendre l'application de l'article 1243 aux robots autonomes, le régime juridique qu'il consacre est pour ainsi dire identique à celui que la jurisprudence a créé sur le fondement de l'article 1242⁹³, pour autant qu'il y ait mouvement et contact avec le siège du dommage. Il serait donc infiniment plus simple pour le juge de raisonner classiquement sur le fondement de l'article 1242 du Code civil. Raisonner sur le fondement de l'article 1243 présenterait au demeurant l'inconvénient technique d'évacuer complètement l'éventuelle responsabilité du concepteur du logiciel dans la

⁹² Par ex. N. NEVEJANS, *Traité de droit et d'éthique de la robotique civile*, op. cit. note 88, n° 742.

⁹³ Ce qui n'est guère étonnant puisque le régime du premier a servi de modèle pour créer le régime du second (M. BACACHE-GIBEILI, *Les obligations, La responsabilité civile extracontractuelle*, *Traité de droit civil*, dir. C. LARROUMET, Economica, 3^e éd., 2016, n° 239).

survenance du dommage : la Nature ne répond en effet aucunement en dernier ressort de la configuration neuronale des êtres vivants qui naissent en son sein.

53. Bilan sur la responsabilité du fait des choses – L'étude de l'application de la responsabilité du fait des choses aux robots « autonomes » permet de dégager au moins un élément de certitude : la victime trouvera toujours un défendeur de principe à son action en responsabilité en la personne du *propriétaire* du robot.

Hormis l'existence d'un titre juridique organisant ou jugé emporter un transfert de la garde du robot, le propriétaire pourra espérer s'exonérer de cette responsabilité en désignant *l'utilisateur* de celui-ci si le dommage a été causé alors que le robot se livrait à une tâche impulsée sous l'ordre de l'utilisateur et que la jurisprudence accepte d'assimiler cet ordre à l'acquisition de l'usage, du contrôle et de la direction du robot. Dans une conception plus ambitieuse, la jurisprudence pourrait décider qu'acquiert la garde du robot celui *au bénéfice duquel* le robot exécutait l'opération qui causé le dommage. Cette dernière conception permettrait d'imputer au surplus à la personne qui devait profiter des actions du robot certaines initiatives prises par celui-ci en exécution de son programme d'apprentissage et donc sans ordre formel préalable en ce sens ; mais l'on serait plus éloigné encore de la justification traditionnelle de la garde matérielle qui réside en la maîtrise de la chose et le pouvoir corrélatif d'empêcher le dommage.

Il sera *a priori* beaucoup plus difficile au propriétaire – et même à l'utilisateur le cas échéant – d'alléguer, en l'état des conceptions actuelles, que la garde du robot autonome doit être imputée au *concepteur de son logiciel*. Outre que cela supposerait au préalable accepter d'appliquer la responsabilité du fait des choses à une entité immatérielle, la distinction entre ce qui devrait être imputé au concepteur du logiciel et ce qui devrait être imputé à l'utilisateur du robot – et dont la classique distinction entre garde de la structure et garde du comportement ne rend pas bien compte – sera souvent très délicate. Cette action ne pourrait en outre prospérer toutes les fois que le dommage apparaîtra imputable à un défaut de sécurité du logiciel dès lors que l'on accepte de qualifier son concepteur de fabricant d'une partie composante au sens de la directive sur les produits défectueux. La responsabilité résiduelle qu'il pourrait conserver au titre des dommages non imputables à un défaut de sécurité l'obligerait alors à répondre à de la dangerosité que l'on pourrait qualifier de « normale » d'un robot autonome.

On voit bien que les difficultés posées par l'application de la responsabilité du fait des choses au robot en général révèlent, en fin de compte, un problème sous-jacent de politique juridique. Sur qui *veut-on faire peser* le poids de la responsabilité qui dérivera du comportement journalier et dommageable du robot, non attribuable à un défaut de sécurité ? Le risque doit-il peser sur celui qui *profite* du robot au jour le jour, en lui donnant des ordres ? Sur son propriétaire, parce que ce dernier est le mieux placé pour souscrire à une *assurance* ? Ou sur son fabricant, parce qu'il s'est enrichi en

commercialisant un robot doué de fonctions dont l'exécution échappe dans une certaine mesure à leurs utilisateurs ?

On vient de voir que les deux principaux régimes susceptibles de s'appliquer pour appréhender la responsabilité dérivant des actions du robot autonome pose chacun quelques difficultés ; le recours à la technique de la personnalité juridique constitue-t-il une réponse adéquate à celles-ci ?

2°) Le recours à la technique de la personnalité juridique

54. L'accélérateur européen – L'idée de recourir à la technique de la personnalité juridique afin d'appréhender la responsabilité civile dérivant des actions du robot, si elle n'est pas née de cette initiative, a pris un relief particulier lorsqu'elle a été reprise par le Parlement européen dans une résolution du 16 février 2017 contenant des recommandations à la Commission concernant des règles de droit civil sur la robotique. Cette résolution demandait à la Commission d'examiner, parmi les solutions juridiques envisageables, « la création, à terme, d'une personnalité juridique spécifique aux robots, pour qu'au moins les robots autonomes les plus sophistiqués puissent être considérés comme des personnes électroniques responsables, tenues de réparer tout dommage causé à un tiers ; il serait envisageable de conférer la personnalité électronique à tout robot qui prend des décisions autonomes ou qui interagit de manière indépendante avec des tiers »⁹⁴. La proposition étant dénuée de réserves, on pourrait penser que le Parlement ne s'était pas interdit de considérer une personnalité juridique des plus complètes. Le rapporteur du texte, la députée européenne Mady Delvaux, s'est toutefois employée à restreindre la portée de cette proposition à plusieurs reprises. C'est ainsi que, dans une interview ultérieure, après avoir souligné que la personnalité juridique du robot « n'est pas la première proposition » envisagée par le texte, la députée a expliqué que cette e-personnalité ne voulait « pas dire que le robot deviendrait une personne mais qu'en cas de dommages causé par lui ce statut permettrait de gérer la responsabilité entre les différents acteurs concernés. Cela se rapprocherait de la personnalité juridique des personnes morales »⁹⁵. De prime abord, l'affirmation est contradictoire : vouloir se rapprocher de la personnalité juridique des personnes morales, *c'est bien vouloir faire des robots des personnes* au sens du Droit. On comprend toutefois que la députée a sans doute voulu signifier à son interlocuteur que la résolution ne visait aucunement à suggérer qu'il faudrait traiter les robots *à l'instar de personnes physiques* et donc *humaines*. Ces précisions ont le mérite de clarifier que, dans l'intention des auteurs du texte, la personnalité juridique envisagée ne serait qu'une personnalité « technique » ou utilitaire. Mais alors, que gagne-t-on à

⁹⁴ n° 59 f de la résolution.

⁹⁵ M. DELVAUX, « Pour un droit européen des robots » (interview), *Expertises* mars 2017, pp. 91-95, spéc. p. 84.

utiliser cette technique ? Faute d'explications précises sur cet intérêt⁹⁶, on en est réduit à examiner les conséquences qui résulteraient de la mise en œuvre d'une telle proposition.

55. Inopportunité de la suggestion – A s'en tenir à la conception classique de cette institution particulière, assigner une personnalité juridique une chose emporte deux conséquences. La première est de faire émerger une nouvelle « entité », un nouveau *sujet* sur le plan du Droit ; la seconde est de lui assigner un patrimoine. Ces deux aspects permettent-ils de vaincre les difficultés que l'on a rencontrées en étudiant la problématique de l'application des textes actuels à la responsabilité des robots autonomes ? Non seulement il ne le semble pas, mais encore, ils semblent compliquer l'appréhension ainsi que la mise en œuvre cette responsabilité, tout en ayant pour effet de limiter la réparation à laquelle la victime pourra prétendre.

56. La première conséquence de l'octroi au robot d'une personnalité juridique sera d'en faire le responsable principal des dommages qu'il cause. La vertu⁹⁷ essentielle de ce schéma, si l'on raisonne par rapport aux difficultés que pose l'application des régimes de responsabilité traditionnels, est de résoudre les difficultés d'imputabilité en tirant les conséquences de l'autonomie décisionnelle du robot.

57. Il n'est pas sûr, cependant, que ce gain compense les inconvénients qui vont découler de l'utilisation de cette technique juridique ; et les moyens qui permettront de remédier à ces inconvénients paraissent réduire en néant l'intérêt de recourir en premier lieu à celle-ci.

58. Traiter avec un robot ? – Sur un plan très pratique, tout d'abord, la victime éprouvera le plus grand mal à traiter avec le responsable, c'est-à-dire le robot, des conséquences de sa responsabilité⁹⁸ : le jour où un robot sera capable d'entretenir une conversation élaborée avec la victime ou son avocat dans la perspective d'une transaction ou alors d'aller défendre ses droits en justice est encore très lointain. La victime devra donc, en réalité, s'adresser à la *personne chargée de représenter le robot*. Sur ce point, on se trouvera

⁹⁶ M. Delvaux suggère que « c'est pour régler la question de la réparation et de l'assurance que nous avons suggéré cette piste de solution » (M. DELVAUX, interview préc. note préc., p. 94).

⁹⁷ Sur les difficultés que poserait la reconnaissance d'une responsabilité du fait personnel du robot, v. C. COULON, art. préc. note 72, spéc. n° 7 (en substance, isoler la faute du robot de celle de ses concepteurs ; distinguer cette faute de celle imputable à son utilisateur).

En tant qu'entité juridique reconnue, le robot devrait également pouvoir être jugé *responsable du fait des choses* dont on le réputerait avoir la garde, conséquence à laquelle il est d'autant plus facile à parvenir que le discernement n'est plus une condition de la responsabilité civile. On peut même concevoir qu'il puisse devenir *responsable du fait d'autrui*. Si l'on a du mal à imaginer, par exemple, un robot commettant d'une personne physique, on pourrait imaginer des robots commettants d'autres robots ; et viendra très certainement le jour où l'on confiera à un robot la responsabilité d'organiser à défilé de majorettes robots...

⁹⁸ Quand on songe à la frustration qu'engendre aujourd'hui l'expérience de devoir traiter d'une quelconque question avec un serveur vocal, la perspective est cauchemardesque.

concrètement ramené à la situation qui aurait été celle dans laquelle le robot n'est pas doué de personnalité : il faudra s'adresser à un humain et traiter avec lui. Qu'aura-t-on gagné au change ?

59. Le patrimoine, facteur limitant la responsabilité – Mais la déconvenue la plus substantielle pour la victime résultant de l'instauration d'une personnalité juridique pour le robot tiendra à la seconde conséquence attachée à la personnification : celui-ci se voyant doté d'un *patrimoine*, l'étendue de la réparation à laquelle la victime pourra prétendre sera concrètement limitée par l'étendue du patrimoine du responsable, c'est-à-dire du robot⁹⁹. Faux obstacle, dirons certains : il suffit, par exemple, d'avoir un responsable en deuxième ligne – un responsable *du fait du robot* – qui serait constitué sur l'un des modèles des responsabilités du fait d'autrui que connaît le droit civil¹⁰⁰ ; ou encore, de conserver un responsable « d'arrière ligne » en la personne du constructeur, à raison des défauts de sa machine ; ou encore, d'assurer la mise en place d'un système assurantiel couvrant la portion de la créance de dommages-intérêts qui dépasserait les capacités patrimoniales du robot. S'engager dans une telle voie serait cependant parfaitement incohérent tant avec la raison qui paraît justifier en premier le recours à cette technique juridique qu'avec les implications logiques et, si l'on peut dire, « naturelles » qui s'attachent à son utilisation.

60. Incohérence des propositions pour remédier à cette difficulté – Quant à la raison qui paraît justifier le recours à cette technique, tout d'abord, on a vu qu'il s'agissait à titre principal d'un problème *d'imputabilité* : c'est parce que l'autonomie du robot rendrait illégitime ou impossible¹⁰¹ l'imputation de la responsabilité de ses actions à d'autres que lui qu'il paraît indiqué de les imputer directement au robot. Mais n'est-il pas dès lors incohérent, pour remédier aux difficultés que cette situation va créer, de se tourner vers ces mêmes personnes comme s'il était normal qu'elles répondent

⁹⁹ Conséquence qui, notons-le, se trouve en contradiction avec la point n° 52 de la résolution du Parlement européen du 16 février 2017, lequel expose « quelle que soit la solution juridique choisie dans le futur instrument législatif [...] ledit instrument ne devrait en aucune manière limiter [...] l'étendue des dommages qui peuvent faire l'objet d'un dédommagement ».

¹⁰⁰ La logique n'indique-t-elle pas ici de désigner comme responsable son représentant légal, ce qui, évidemment, invite le parallèle avec la responsabilité des parents du fait de leurs enfants ? Reproduire ce régime juridique pour appréhender la responsabilité du robot ne manquerait pas de sel. En l'état actuel de la jurisprudence, en effet, les parents voient leur responsabilité engagée pour n'importe quel fait dommageable du mineur. Quel pourrait être alors l'intérêt d'avoir conféré au robot une personnalité juridique ? Et si l'on devait se rabattre sur le modèle traditionnel – peut-être ravivé demain – d'une responsabilité du fait d'autrui supposant en toute hypothèse une faute de « l'autrui », comment caractériser une faute du robot ? Devrait-on appliquer tout simplement une conception objective de la faute ? L'importation de ce concept, destiné à régir les comportements humains, ferait-il sens pour appréhender le comportement d'un robot ?

¹⁰¹ Par ex. L. GODEFROY, art. préc. note 38, n° 24 : si le comportement dommageable est lié à des « incidences négatives de l'apprentissage et de l'autonomie décisionnelle du robot [alors] « le dommage ne semble pas réparable par les régimes de responsabilité en vigueur » (l'auteur ne prônant par aucunement la personnalisation du robot, v. n° 26).

du fait du robot ? Soutenir une telle solution, c'est avouer que l'autonomie n'est, en fin de compte, pas un obstacle à l'imputation « traditionnelle » de la responsabilité qui découle des actions du robot à des personnes physiques ou morales.

61. A supposer que l'on passe outre cette première remarque et que l'on accepte donc que puisse subsister concurremment d'autres responsables, il faut maintenant souligner que vouloir *responsabiliser* le robot – ou n'importe quel entité d'ailleurs – en lui conférant la personnalité juridique est indissociable d'une démarche ou inspiration visant à *déresponsabiliser* ceux qui devraient, ou auraient dû, sans l'existence de cette construction juridique, assumer une responsabilité au titre de l'action du robot¹⁰². Attribuer au robot une personnalité juridique, c'est en effet intrinsèquement vouloir déplacer sur ses épaules le poids de la responsabilité qui découle de ses actions. N'est-ce pas d'ailleurs précisément ce qu'implique le parallèle qui est fait, pour cette personnalité « électronique », avec une personne morale ?¹⁰³ La technique de la personnalité morale – c'est encore plus vrai de nos jours du fait de la banalisation de la figure de la société unipersonnelle – est effet généralement utilisée comme une technique de *cantonnement* de la responsabilité. L'analogie vient d'ailleurs parfaitement conforter une observation pratique que l'on peut faire relativement à l'attribution d'un patrimoine au robot : *qui abondera ce patrimoine ?* Son constructeur ? Son propriétaire ? Quelle que soit la ou les personnes que le législateur désignera comme ayant à charge de constituer ce capital, celles-ci, une fois ce capital constitué, s'attendent légitimement à ce que la loi les préserve, en contrepartie, de toute forme de responsabilité supplémentaire à raison des dommages causés par les actions du robot. N'est-ce pas au demeurant la conséquence mécanique découlant de l'affirmation qu'en premier lieu il ne serait pas légitime de leur imputer le dommage ? Assigner un patrimoine au robot soutient en tous cas parfaitement l'analogie avec la constitution d'un capital social à une société de capitaux : c'est ce capital et non les associés – ici, les personnes impliquées dans « l'avènement » du robot – qui a vocation normalement exclusive à répondre des dettes de cette dernière. Quel intérêt pourrait d'ailleurs avoir cette constitution du patrimoine du robot si elle ne préservait pas la responsabilité de la ou des personnes qui devront assumer la charge de cette constitution ?¹⁰⁴ On ne peut de ce point de vue qu'être surpris lorsqu'on lit¹⁰⁵, chez certains partisans de la personnalité juridique du robot, la proposition de schémas de responsabilité en cascade impliquant les personnes qui, sans cet artifice technique, auraient vocation à répondre des conséquences dommageables liées à son action.

¹⁰² En ce sens, G. LOISEAU, « La personnalité juridique des robots : une monstruosité juridique », art. préc. note 71.

¹⁰³ V. les obs. de M. DELVAUX, préc. note 95.

¹⁰⁴ On pourrait répondre que l'on gagne une sécurité tenant à ce que la somme d'argent est immédiatement disponible. Mais alors, il y a une autre solution plus simple pour arriver concrètement à un résultat presque voisin : l'assurance.

¹⁰⁵ A. et J. BENSOUSSAN, *Droit des robots*, Larcier, 2016, n° 212.

62. La volonté d'assigner au robot une personnalité juridique paraît donc indissociable, tant conceptuellement que d'un point de vue politique, d'une démarche visant à *isoler les conséquences juridiques* dérivant du dommage causé par le robot *au robot lui-même* et donc à préserver l'intérêt des personnes qui, sans l'existence de cette personnalité, aurait dû assumer cette responsabilité. Vouloir préserver une action en responsabilité contre ces dernières personnes, afin de contrer l'effet limitatif de responsabilité qu'induit l'existence du patrimoine robot, c'est nier en premier lieu le sens et l'intérêt du recours à cette technique juridique.

63. **Faux remède de l'assurance « par » le robot** – Reste alors le dernier argument : cette limite pourrait être facilement surmontée en « obligeant » le robot à disposer d'une assurance pour couvrir la part de responsabilité que son patrimoine ne permettrait pas d'absorber. A cette idée, deux objections peuvent être faites. Tout d'abord, qui souscrira et paiera les primes de cette assurance ? Le robot, avec son capital ? Qu'advient-il alors lorsque ce capital, qui était déjà supposé servir à désintéresser les éventuelles victimes, sera épuisé¹⁰⁶ ? Il n'y aura plus d'assurance et, comme celle-ci a épuisé le capital, la victime ne pourra pas être désintéressée. On voit à nouveau que la technique de personnification du robot rime avec limitation de responsabilité. Pour éviter cette conséquence, il faudrait que le paiement de l'assurance incombe à une personne physique (ou une société) en lien avec le robot : son utilisateur ou, pour plus de sécurité, son propriétaire¹⁰⁷. Mais on en arrive alors à la seconde objection : procéder ainsi, n'est-ce pas s'avouer *qu'attribuer une personnalité juridique au robot n'aura en fin de compte servi à pas grand-chose* ? En effet, on en est finalement réduit à se tourner, pour contracter et abonder cette assurance, vers le ou les personnes qui auraient dû, sans l'existence de cette personnalité juridique, assumer les conséquences juridiques des dommages causés par le robot. N'est-il alors pas plus simple d'adopter un schéma juridique permettant à la victime de s'adresser *directement* à ces personnes pour obtenir réparation des dommages causés par le robot et d'obliger ces dernières à souscrire une assurance à cet effet ?

64. Le recours à la personnalité juridique, en tant que technique juridique, ne paraît donc présenter aucun intérêt pour appréhender la responsabilité civile dérivant de l'action du robot¹⁰⁸ ; elle semble surtout créer des difficultés dont la résolution aboutit à nier l'utilité de cette personnification du robot. Quelles

¹⁰⁶ S'il n'est pas déjà épuisé par un recours antérieur en responsabilité.

¹⁰⁷ La sécurité vient de ce que la propriété du robot est une situation plus stable que celle de simple utilisateur, ce qui rend facile la systématisation de l'obligation d'assurance. Quant à rendre le producteur du robot responsable de cette assurance, elle ne paraît pas réaliste ; sauf à imaginer qu'ils abondent un fonds de garantie venant pallier l'absence de toute indemnisation de la victime.

¹⁰⁸ Pour des opinions selon lesquelles doter le robot d'une personnalité juridique est inutile pour traiter des questions de responsabilité en cas de dommage causé par le robot, v. N. NEVEJANS, *op. cit.* note 88, n° 60.

solutions peut-on suggérer pour appréhender la responsabilité qui dérivera des actions dommageables d'un robot autonome ?

3°) Propositions de solution

65. Méthode – On a vu que les régimes de responsabilité civile existants ne s'appliquent que malaisément à la question du dommage causé par un robot exerçant ses facultés d'autonomie et que les règles qui pourront être amenées à régir cette question varieront selon l'interprétation que l'on fera des concepts et principes qu'ils contiennent. Étant donné qu'il serait erroné de penser qu'il existe une interprétation « exacte » des textes et principes en la matière, il est plus rationnel de commencer par se demander quel régime de responsabilité *paraîtrait opportun ou souhaitable* pour appréhender ce nouveau problème puis (a), dans un second temps, s'il est possible de mettre en place ce régime juridique en adoptant simplement l'interprétation qui permet d'aboutir à ce résultat, quitte à suggérer quelques règles complémentaires pour aboutir à un système pleinement satisfaisant (b).

a) Que veut-on ?

66. Critère de solution – La difficulté essentielle que pose l'appréhension de la responsabilité civile dérivant des actions d'un robot autonome et à laquelle la proposition de doter le robot d'une personnalité juridique se voulait une réponse est celle de l'imputabilité des dommages causés par le robot autonome en exécution d'un comportement qu'il aura auto-appris ou dont il aura pris l'initiative en vertu de son algorithme d'autoapprentissage. Or cette difficulté ne paraît pas si difficile à résoudre si l'on en revient aux fondements qui peuvent être adoptés pour un régime de responsabilité.

67. Quel fondement ? — On se trouve, s'agissant des robots « autonomes », face à une question qui n'est en vérité guère différente de celle qui s'est posée à l'aube de l'arrêt *Teffaine* : veut-on un système de responsabilité qui reflète au moins implicitement la conception selon laquelle la responsabilité d'un dommage ne doit peser que sur celui ou celle à qui, d'une manière ou d'une autre, on peut imputer une forme *de faute* en relation avec le dommage ? Ou souhaite-t-on un système qui, libéré de cette préoccupation, privilégie les intérêts de la victime et impute la responsabilité des dommages non clairement attribuables à une faute humaine à un responsable pré-identifié – *a priori* celui qui profite du risque créé par la chose – et dont le rôle essentiel sera de contracter et donc de fournir une assurance ? Vu sous cet angle, il semble évident que c'est le second système qui doit être privilégié puisque le postulat de départ de l'hypothèse que nous examinons ici est précisément que le comportement dommageable résultant des actions autonomes du robot ne sera souvent pas attribuable à une faute mais à processus algorithmique dont la nature même implique un risque de dommage.

68. Privilégier un système de responsabilité pour risque n'implique aucunement, au demeurant, de faire disparaître le classique système de

responsabilité pour faute qui pourra être mis en œuvre si la situation le justifie : faute du constructeur, faute de l'utilisateur qui donnerait au robot des ordres dans un dessein dolosif ou encore le placerait inutilement dans une situation de nature à susciter de sa part des réactions susceptibles de provoquer un dommage, faute d'un tiers qui piraterait la machine afin de lui faire accomplir des actes dommageables, etc. On peut cependant penser que relier certains comportements dommageables des robots autonomes au comportement fautif ou défectueux d'une personne va devenir, à mesure de l'évolution technologique, de plus en plus difficile : l'autonomie des robots les conduira à agir de plus en plus sous l'empire de règles qu'ils auront eux-mêmes élaborées et, on l'a vu, les doter d'une personnalité juridique pour leur imputer ces dommages au titre d'une « faute » de leur part serait inopportun ; ce à quoi l'on peut ajouter ici que si l'on suivait cette dernière voie, les fonctions préventives et punitives de la responsabilité se trouveraient parfaitement inopérantes¹⁰⁹ ce qui réduit d'autant plus sa pertinence.

69. Quel responsable ? – Qui devrait maintenant être le responsable de principe des dommages attribuables à des comportements résultant des fonctions d'apprentissage du robot ? Cela devient à se demander sur qui devrait peser à titre de principe ce risque de dommage.

70. Concepteurs du robot ? – Selon nous, les conséquences dommageables de ce risque ne devraient pas être imputées aux *concepteurs* du robot. Leur imposer cette responsabilité – en dehors de toute faute de leur part – reviendrait en effet à leur imposer la charge de répondre de dommages apparaissant comme une conséquence inéluctable du mode de fonctionnement de ces machines¹¹⁰. Cette opinion ne manquera pas de soulever des objections. L'une d'entre elles est qu'absoudre les producteurs de l'obligation de répondre des dommages attribuables aux apprentissages défectueux des algorithmes dont ils dotent leurs robots revient à leur donner une forme de licence ou d'immunité qui ne va pas dans le sens de l'intérêt de la société, concrètement des consommateurs et des potentiels victimes. A cela, on répondra tout d'abord qu'il existe une arme presque aussi dissuasive qu'une action en dommages-intérêts pour responsabiliser et motiver ces acteurs à produire une machine aussi sûre que possible du point de vue de ses algorithmes d'apprentissage : la réputation du constructeur. Aucun concepteur de robot n'a intérêt à produire ou à laisser sur le marché des robots à l'algorithme trop imprévisible et dangereux, sous peine de péricliter et disparaître. On peut donc attendre du simple fonctionnement du marché un effet disciplinaire sur les producteurs de ces machines. Ensuite, que les producteurs n'aient pas à répondre de ces dommages n'implique aucunement qu'on ne puisse les mettre à contribution

¹⁰⁹ Encore que l'on puisse imaginer que l'algorithme du robot intègre la condamnation comme une réponse non gratifiante l'éduquant sur l'inadéquation de son comportement...

¹¹⁰ On peut comprendre, de ce point de vue, qu'ils puissent voir de manière favorable le recours à la technique juridique de la personnalité, dont on a vu qu'elle aurait pour conséquence normale de les mettre à l'abri d'une telle responsabilité.

autrement en considération des risques qu'ils contribuent à créer : par exemple, en leur imposant d'abonder un fonds de garantie.

71. Utilisateur du robot ? – Ce risque de dommage devrait-il alors être imputé à *l'utilisateur* ? Cette solution pourrait sembler plus évidente car c'est lui qui, par définition, profite de l'action du robot au moment où la machine cause le dommage. Il nous semble cependant qu'elle ne devrait pas être retenue pour deux raisons. La première est qu'elle aurait pour inconvénient rendre possiblement incertaine, au gré des espèces, l'identité concrète du débiteur : pourra-t-on en effet facilement déterminer, au moment précis où le robot a causé le dommage, la personne au profit de laquelle le robot accomplissait une activité ? Dira-t-on que l'ensemble des personnes qui bénéficiait potentiellement de cette activité devront répondre de ses conséquences dommageables ? La seconde raison de ne pas imputer ces dommages à l'utilisateur tient à ce que cela introduirait une incertitude sur la capacité de la personne finalement désignée comme responsable à faire face à l'obligation de réparation qui sera mise sa charge.

72. Propriétaire du robot – C'est donc sur le *propriétaire* du robot qu'il nous paraît en fin de compte le plus correct de faire penser le risque de dommage attaché au fonctionnement des algorithmes d'apprentissage et d'autonomie du robot. Cette solution présente deux avantages : le propriétaire peut être rendu facilement identifiable et la stabilité de son identité facilite considérablement l'instauration d'une obligation d'assurance couvrant ce type de dommages. Comment, maintenant, instaurer un tel régime ?

b) Comment y parvenir ?

73. Relative simplicité de la mise en œuvre – En l'état actuel des textes, la mise en place d'un tel système, à supposer qu'on le juge opportun, est à *portée de main* : il suffit d'adopter une interprétation des textes et principe existants qui soit concordante, quitte à parfaire le système par quelques règles législatives complémentaires.

74. Mise en œuvre par interprétation des dispositifs actuels – Si l'on suit l'interprétation proposée précédemment des règles relatives aux produits défectueux¹¹¹, le producteur – on inclut ici le concepteur du logiciel qui devrait selon nous répondre de son fait selon le régime des produits défectueux – reste responsable des défauts de sécurité attribuables à une *erreur de programmation*¹¹² ainsi que des défauts de sécurité qui, s'ils résultent de comportements auto-appris par la machine, constitueraient des comportements *que l'on pouvait raisonnablement prévoir et préprogrammer le robot à ne pas adopter*. La responsabilité du producteur se verrait par ailleurs réduite, comme

¹¹¹ *Supra*, n° 24.

¹¹² Ce qui suppose un dommage causé par le robot en exécution de l'une de ses routines « primaires » s'avérant structurellement défectueuse.

le prévoit ce régime, en cas de combinaison entre le défaut du logiciel et une faute de la victime ou d'une personne dont elle est responsable¹¹³.

75. S'agissant des comportements dommageables résultant de comportement *auto-appris postérieurement à la mise en circulation de la machine* et que ses concepteurs *ne pouvaient raisonnablement prévoir*, les producteurs n'encourraient aucune responsabilité du fait que l'on ne saurait « légitimement attendre » une telle sécurité de la part d'un robot doué de facultés d'apprentissage autonome. Les options ouvertes à la victime dépendront alors de l'interprétation que les tribunaux accepteront de retenir de la règle posée par l'actuelle CJUE sur l'exclusion du droit commun lorsqu'est allégué un défaut de sécurité d'un produit¹¹⁴.

Si les tribunaux adoptent une interprétation stricte de cette règle en jugeant que l'absence d'un défaut de sécurité au sens de la directive ferme la porte de toute action en justice en rapport avec la sécurité du produit, alors l'indemnisation de la victime dépendra de l'adoption de certaines mesures législative suggérées plus bas¹¹⁵.

Si les tribunaux devaient en revanche accepter de considérer que le droit commun peut être sollicité car l'existence d'un défaut de sécurité a été techniquement écarté, la victime pourrait assigner le propriétaire du robot sur le fondement de la responsabilité du fait des choses¹¹⁶. Le système que nous suggérons ici nécessiterait alors que la Cour de cassation instaure la règle prétorienne selon laquelle le propriétaire d'un robot autonome, hormis l'existence d'un défaut de sécurité relevant du régime spécial des produits défectueux, et sous réserve d'une faute de la victime ou du fait d'un tiers, est responsable de plein droit des dommages qu'elle cause, *sans pouvoir démontrer qu'il en a transféré la garde matérielle* afin de couper court comme proposé à toute difficulté d'imputation¹¹⁷. Il pourrait paraître tentant de réserver l'hypothèse d'un transfert de la garde par titre juridique, mais ce serait en fin de compte réintroduire une partie des difficultés que la désignation irrévocable du propriétaire a vocation à résoudre¹¹⁸. Réputer le propriétaire du

¹¹³ Art. 1245-12 civ.

¹¹⁴ *Supra*, note 84.

¹¹⁵ *Infra*, n° 78 et 79.

¹¹⁶ Ici, l'idée que nous avons suggérée plus haut selon laquelle le comportement auto-appris d'un robot à l'origine du dommage *n'est pas un défaut* au sens de la directive s'avèrera particulièrement opportune puisqu'elle justifierait que le recours au droit commun reste possible (dans l'interprétation jurisprudentielle ici examinée). Admettre au contraire qu'il y a un défaut du robot mais que l'on ne saurait l'imputer à son concepteur car le défaut *est apparu postérieurement* à la mise en circulation de la machine fermerait en revanche tout recours.

¹¹⁷ Il faudrait donc décider que l'utilisateur d'un robot n'en acquiert jamais à proprement parler *l'usage, le contrôle et la direction* ; il conviendrait également dans le sillage de cette règle de ne pas ressusciter le concept de la garde de la structure.

¹¹⁸ Soit un robot de service de maison loué par une personne : admettre que le contrat de prêt opère un transfert de la garde reviendrait en réalité à réintroduire implicitement le critère comme la logique de la garde matérielle (c'est le preneur qui a l'usage de la chose) ; au

robot devoir irrévocablement répondre de ces dommages aurait également pour conséquence, du fait de la prévisibilité de la règle, de provoquer spontanément le développement d'une assurance venant couvrir ce risque de responsabilité. La responsabilité pour faute resterait toutefois disponible en arrière-plan afin d'atteindre toute personne dont le comportement défectueux a contribué en relation avec le comportement du robot à causer le dommage, que celui résulte d'ailleurs de l'algorithme d'apprentissage ou de tout autre cause : ainsi placer le robot dans une situation clairement inadéquate ou lui apprendre à dessein des comportements destinés indirectement à provoquer des dommages.

Il est probable que la jurisprudence se rangera à la première des deux interprétations suggérées plus haut, d'où la nécessité de prévoir certaines mesures législatives visant à assurer une réparation des dommages causés par les fonctions autonomes d'un robot et ne pouvant être imputées à un défaut de sécurité de celles-ci. Dans l'hypothèse où la jurisprudence retenait la seconde interprétation, ces mesures n'en seraient pas moins utiles, mais apparaîtraient comme des compléments venant parfaire la construction jurisprudentielle suggérée.

76. Mesures législatives – Trois mesures législatives seraient utiles pour appréhender les dommages causés par les robots autonomes et non attribuables à un défaut de sécurité au sens strict du terme.

77. La première mesure serait l'instauration d'une obligation d'immatriculation des robots autonomes¹¹⁹ laquelle aurait pour conséquence la création d'un registre dédié. Les robots, à l'instar des navires, auraient donc une *fiche matricule* qui permettrait d'identifier de manière certaine leur propriétaire, mais aussi les transferts de propriété dont ils ont pu faire l'objet et, pourquoï pas, les sûretés qui les grèvent.

78. La seconde mesure, dans le sillage de la précédente, serait l'instauration d'une *assurance obligatoire* à la charge des propriétaires de robot. Cette solution est sans doute la plus efficace pour répondre des dommages causés par les robots autonomes, quelle que soit l'attitude qu'adoptera la jurisprudence sur la possibilité de la victime de recourir au droit commun.

79. Enfin, il conviendrait d'instaurer un *fond d'indemnisation* venant couvrir les hypothèses dans lesquelles la victime, malgré le dispositif proposé, se retrouverait, hors faute de sa part, privée d'indemnisation ou dans l'impossibilité d'obtenir la réparation pleine et entière de son préjudice. Ce fonds serait alimenté à titre principal par une contribution des producteurs de

demeurant, on pourrait craindre les clauses contractuelles de style par lesquelles les propriétaires imposeraient à leurs cocontractants d'assumer les conséquences juridiques de ces dommages, ce qui compromettrait également leur prise en charge assurantielle faute pour les personnes désignées responsables par le contrat de s'être concrètement assurées.

¹¹⁹ Cette mesure est d'ailleurs proposée dans la Résolution du Parlement européen du 16 février 2017, mais à pour objet de pouvoir identifier le fonds dont dépendrait le robot.

robot, lesquels s'acquitteraient ainsi de la part de risque qu'ils génèrent en concevant ces machines douées d'autonomie.

80. Il serait enfin opportun que ces mesures soient adoptées à l'échelle européenne, non seulement à des fins d'harmonisation mais également parce que cela donnerait au législateur européen l'occasion de clarifier le texte de la directive sur les produits défectueux en imposant l'interprétation permettant au dispositif d'être efficace. Il conviendrait plus précisément de réputer expressément le concepteur du logiciel d'un robot autonome fabricant d'une partie composante et de réputer non constitutifs d'un défaut de sécurité la possibilité d'un robot autonome d'adopter un comportement dommageable en exécution de son algorithme d'apprentissage dès lors que celui-ci ne pouvait être raisonnablement prévu et interdit à la machine dans sa couche logicielle initiale.

81. Responsabilité civile du robot : conclusion – L'étude des difficultés que suscite l'appréhension des dommages causés par les robots autonomes permet deux conclusions. La première est que les régimes de responsabilité actuels, s'ils posent quelques difficultés, peuvent néanmoins être interprétés de manière constructive pour englober ces hypothèses, à supposer que l'on s'entende sur les objectifs de politique juridique qu'il convient de poursuivre. La seconde, qui nous intéresse le plus immédiatement dans le cadre de cette étude, est que la personnalisation juridique du robot ne constitue en aucune manière une solution adéquate aux difficultés qui doivent être résolues ; bien plus, recourir à cette figure juridique tout en souhaitant assurer une réparation adéquate des victimes nécessiterait d'adopter des mesures incohérentes avec le principe même de la personnalisation du robot.

L'argument de la liberté décisionnelle du robot permet-il de mieux fonder la revendication de l'octroi d'une personnalité juridique de ces machines ?

B. L'argument de la liberté décisionnelle du robot

82. Exposé de l'argument – On a vu lors des développements précédents que l'une des raisons qui pouvaient être avancées pour conférer une personnalité juridique au robot doué d'une certaine autonomie tient à ce que cette autonomie rend délicate, voire illégitime, l'imputation de son comportement dommageable à des humains. L'autonomie était alors considérée comme un critère impliquant et justifiant *l'exclusion* de la responsabilité humaine (ou d'une personne morale).

83. L'argument tiré de la liberté décisionnelle est en quelque sorte le versant positif de l'argument tiré de l'autonomie. L'idée ici est qu'il faut conférer au robot une personnalité juridique afin de tirer les conséquences ontologiques de cette liberté décisionnelle. Cette liberté du robot serait en effet

le « gage de son “existence”¹²⁰. Il s’agit « d’offrir à l’entité robot (certains diront espèce artificielle), douée d’intelligence et dotée de liberté décisionnelle, une catégorie juridique alignée sur ses capacités »¹²¹. Le *principe* même d’existence de cette liberté, qui fait du robot en partie « son propre législateur », suffirait à provoquer un changement de paradigme¹²². Quant au *degré* de liberté du robot, il servirait de critère de répartition entre les règles de droit pertinentes : plus la liberté d’action du robot s’avèrera importante, et plus il conviendra de lui faire quitter le régime du droit des biens afin de lui appliquer d’autres règles juridiques « plus proches de celle des humains et imprégnées de références éthiques et culturelles »¹²³. La personnalité robot serait une étape entre différentes possibilités qui s’échelonnent entre d’un côté du spectre la soumission du pur régime du droit des biens et, de l’autre, l’assimilation sur le plan juridique du robot à la personne humaine¹²⁴. En somme, plus la liberté décisionnelle du robot s’avèrera importante, plus il conviendra de l’appréhender juridiquement sous l’angle d’une personne juridique, apte à acquérir et à exercer des droits¹²⁵.

84. Analyse et critique – La proposition d’accorder une personnalité juridique (et en conséquence des droits) au robot en considération de la « liberté décisionnelle » dont il jouirait semble reposer sur une erreur de perspective fondamentale : celle selon laquelle cette liberté constituerait un critère valable ou légitime d’attribution de la personnalité à cette machine.

85. La liberté décisionnelle d’un robot ne paraît être, ni plus ni moins, que la capacité de celui-ci à s’engager dans une action d’une manière qui ne résulte pas *mécaniquement* de sa programmation initiale et qui, à ce titre, ne peut être prévue avec certitude par son concepteur¹²⁶. Il s’agit d’une latitude d’action de la machine face à une situation donnée. Cette latitude pourra, à mesure que les robots se perfectionnent, devenir de plus en plus étendue. Mais cette liberté du robot dans la mise en œuvre d’une *tâche pour laquelle il a été conçue*, hypothèse qui restera pour encore assez longtemps la seule « liberté » que l’on pourra lui prêter, ne mérite *aucune considération ontologique* devant se

¹²⁰ A. et J. BENSOUSSAN, *op. cit.* note 105, n° 178.

¹²¹ *Ibidem*, n° 177. Ces auteurs ne défendent toutefois pas l’attribution au robot une personnalité juridique équivalente à celle des humains.

¹²² *Ibid.*, n° 179.

¹²³ *Ibid.*, n° 187 et s., spéc. n° 191.

¹²⁴ *Ibid.*, n° 192. Cette dernière hypothèse n’ayant toutefois, d’après ces auteurs, pas lieu d’être envisagée « en l’état des technologies existantes et prévisibles (2015-2020) » (n° 193).

¹²⁵ *Ibid.*, n° 197.

¹²⁶ C’est *mutatis mutandis* la définition qu’en donne les partisans précités de cette proposition, ceux-ci considérant qu’« en creux de la notion de liberté se trouve celle d’intelligence », qu’ils définissent chez le robot comme la capacité de formuler des réponses qui ne « relèvent pas d’un automatisme ni d’une recherche lexicographique en base de données » et qui, couplé avec la fonctionnalité d’apprentissage associée, permet au robot « d’agir de manière autonome par rapport aux concepteur, fabricant, utilisateur et éducateur » (A. et J. BENSOUSSAN, *op. cit.* note 105, n° 187 et 188).

traduire sur le plan du Droit. La « liberté décisionnelle » que l'on cherche à conférer aux robots n'est qu'utilitaire ; comme l'a exprimé le Parlement européen dans sa résolution de 2017, « l'autonomie d'un robot (...) est de nature purement technique et le degré d'autonomie dépend du degré de complexité des interactions avec l'environnement prévu par le programme du robot »¹²⁷.

86. Il y a quelque chose de préoccupant à constater qu'il puisse être proposé que, même en supposant que la latitude d'action autonome des robots devienne de plus en plus substantielle, il faille en tirer comme conséquence la nécessité de conférer au robot une essence et des droits sur le plan juridique. Cette démarche n'est pas sans rappeler cette expérience, réalisée pour un documentaire animalier, dans laquelle on plaça dans l'enclos d'un groupe de singes une réplique robotique d'un de leurs congénères afin de pouvoir filmer plus discrètement la vie de la communauté. Les singes adoptèrent alors le robot comme l'un des leurs ; mais plus encore, lorsque l'un d'entre eux fit accidentellement tomber le robot d'une certaine hauteur, ce qui eut pour conséquence que le robot cessa de bouger comme s'il venait de mourir, tout le groupe se réunit autour de lui exactement comme s'il s'agissait d'un véritable jeune singe qui venait de décéder, avec des réactions d'empathie et de tristesse¹²⁸. La scène, attendrissante, laissera également le spectateur songeur quant à la capacité que nous humains avons eu à berner ce groupe de singes sur l'appartenance d'un automate à leur famille. N'est-ce pas nous laisser berner à l'identique de ces singes que de vouloir reconnaître une existence juridique et des droits corrélatifs à un robot capable de singer – expression particulièrement à propos – avec une latitude plus ou moins grande une autonomie dans l'adoption de certains de ses comportements ? Un automate singeant la capacité de décision d'un agent rationnel et libre ne saurait être assimilé à un tel agent. Car imiter la liberté, ce n'est certainement pas être libre !

87. Liberté ? — Il faut à cet égard souligner le caractère insidieux du raisonnement tirant argument de la « liberté » d'un robot pour en conclure qu'il faudrait lui reconnaître une existence juridique et des droits le rapprochant des humains¹²⁹ ; cela revient, *in fine*, à reconnaître que lui est dû à ce titre une forme de *respect* comparable à celui qui est dû aux humains.

La « liberté » du robot dont il est question ici n'a pourtant rien à voir avec la liberté dont bénéficie l'humain. La capacité d'apprentissage, d'adaptation et de décision subséquente d'un robot ne saurait être confondue avec la possibilité

¹²⁷ Point AA de la [résolution du 16 février 2017 contenant des recommandations à la Commission concernant des règles de droit civil sur la robotique](#).

¹²⁸ La scène, tournée lors de la réalisation du documentaire de la BBC « Spy in the wild » est devenue virale sur Internet et mérite d'être visionnée.

¹²⁹ C'est bien le contenu de la proposition étudiée, qui propose d'appliquer au robot des règles juridiques « plus proche des humains » aux robots doués de liberté décisionnelle.

de formuler ou d'avoir des « intentions »¹³⁰, des « désirs » ou des « projets », autant de notions participant du champ lexical du concept de liberté¹³¹ et qui traduisent des jugements de valeurs, des sentiments, ressentiments ou frustrations dont la machine dotée d'une « liberté décisionnelle » est dépourvue.

88. La liberté intellectuelle et décisionnelle humaine a de surcroît ceci de fondamental qu'elle existe par elle-même et pour elle-même dans toute sa plénitude ; elle surgit pour ainsi dire du néant, sans avoir été préconçue, contrairement à celle dont on dote le robot et dont l'objet est de faciliter ou d'optimiser l'accomplissement par lui d'une certaine catégorie de tâches. La « liberté » d'un robot est pleinement fonctionnalisée, asservie à un but que nous spécifions : elle ne vise pas à lui permettre de concevoir des « projets » sans aucun rapport avec la résolution d'une tâche pour laquelle il a été conçu¹³². Il faut donc prendre garde au fait qu'employer le terme de « liberté » pour désigner la latitude dont on dote une machine pour qu'elle s'acquitte d'une tâche revient d'une certaine manière à importer insidieusement à son profit – si l'on ose dire – et de manière erronée tout un ensemble de présupposés fondant le paradigme de la liberté humaine.

89. La liberté dans l'exécution d'une action se reconnaît également à la capacité de son auteur de rendre compte des raisons qui ont conduit à cette action¹³³. Sur ce point, l'idée d'une « liberté décisionnelle » conférée au robot faillit également dans la mesure où il semble que les algorithmes d'apprentissage avancés ouvrant cette capacité rendent pour beaucoup impossibles, du fait même de leur processus de fonctionnement, la détermination des facteurs qui ont conduit à l'adoption d'une règle¹³⁴ et conséquemment d'un comportement par la machine. Cette dernière est également dépourvue des moyens de comprendre les processus qui l'ont amené à ce comportement et en « subit » le résultat plus qu'elle ne l'adopte.

90. Le concept de liberté a également une dimension politique et la dérive à laquelle conduit le parallèle entre « liberté décisionnelle » du robot et liberté humaine est ici patente. Qu'est-ce en effet que militer pour la reconnaissance au robot, sous prétexte de sa liberté décisionnelle, d'une personnalité juridique et de droits, si ce n'est porter un projet politique visant à imposer aux humains

¹³⁰ « L'intention » d'un robot paraît devoir se résumer, *in fine*, au déclenchement d'une action qui se rapporte au champ des actions pour lesquelles il a été programmé (par ex. aller chercher la bouteille d'eau pour une personne âgée lorsque les conditions sont remplies) ou alors à mettre en œuvre une règle comportementale qu'il a auto-apprise par observation de son milieu.

¹³¹ V. P. RICOEUR, « Liberté », *Encyclopædia Universalis* [en ligne], consulté le 2 mai 2020. URL : <http://www.universalis.fr/encyclopedie/liberte/>.

¹³² Et l'on ajoutera qu'il serait sans doute sage de ne jamais s'orienter dans cette voie, mais les recherches dans cette direction paraissent inéluctables (v. *infra*).

¹³³ P. RICOEUR, « Liberté », art. préc. note 131.

¹³⁴ Ce sont les fameux « black box algorithms ».

la reconnaissance et le respect, sanctionné par des obligations contraignantes, d'une nouvelle population d'êtres ? Ici encore, les fondements d'une telle lutte politique, parfaitement identifiés et légitimes à l'égard d'un groupe humain, font défaut¹³⁵ s'agissant des robots. Outre le fait que ces derniers, on l'a vu, ne saurait être considérés comme ayant une « liberté » de même nature que celle des hommes, ils *ne souffrent aucunement* ni d'une contrariété qui serait apportée à leur prétendue liberté, ni d'une absence de reconnaissance politique et juridique de leur « essence ». Militer pour habiller d'existence juridique et de droits un robot parce qu'il manifeste une « liberté décisionnelle », n'est-ce pas d'une certaine manière se comporter tel un enfant qui s'aviserait d'habiller sa poupée nue parce qu'elle a froid ?

91. Contenu des droits reconnus aux robots ? — On peut enfin s'interroger sur les « droits » qu'il conviendrait de reconnaître aux robots en conséquence de leur autonomie décisionnelle. Les partisans de cette approche proposent bien une « charte des droits des robots », mais cette charte, pourtant forte de 10 articles, ne contient en réalité à proprement parler qu'un seul droit pour les robots, au demeurant fort étrange : « *Un robot a le droit au respect de sa dignité limitée aux données à caractère personnel conservées* »¹³⁶. Le constat est surprenant car il n'est tiré concrètement aucune conséquence de la liberté décisionnelle du robot, laquelle était pourtant au cœur de l'argumentation défendant qu'on leur attribue la personnalité juridique. Ce constat renforce l'impression que cette volonté semble avant tout procéder d'une certaine fascination pour ces machines et leur apparente « liberté » que d'un besoin impérieux.

92. Conclusion – Le simple fait pour un robot de disposer de ce qui s'apparente à une « liberté décisionnelle », c'est-à-dire une certaine latitude de sa part dans l'engagement ou la réalisation d'une action, ne semble donc en soi *aucunement* constituer un critère d'acceptabilité d'attribution d'une personnalité juridique à cette machine.

C. Les arguments sociétaux

93. Une troisième justification ayant pu être avancée pour justifier l'octroi au robot d'une personnalité juridique consiste en réalité en une famille d'arguments reposant sur l'idée que cette attribution présenterait *un intérêt pour la société*.

94. Prévisibilité, apaisement, commercialisation – Certains font valoir, tout d'abord, que donner la personnalité juridique aux robots permettrait « *de répondre à des impératifs pratiques de prévisibilité tendant à sécuriser*

¹³⁵ Pour l'instant du moins (*infra*, Partie II).

¹³⁶ A. et J. BENSOUSSAN, *op. cit.* note 105, n° 219. Pour une réflexion plus aboutie sur les droits fondamentaux qu'il conviendrait de conférer aux robots, mais sur le fondement d'un autre critère, S. SLAMA, « Les robots-androïdes, de quels droits fondamentaux ? », *RDLF* 2019 chron. n° 50 (disponible en ligne).

l'insertion des robots dans le tissu social »¹³⁷, d'*apaiser* cette insertion¹³⁸ ou encore s'avérerait être un « *moyen discret au service de la commercialisation de robots autonomes* »¹³⁹. Ces arguments peinent à convaincre. Que le régime juridique qui sera applicable au robot soit rendu *prévisible*, c'est un souhait que tout le monde partagera ; absolument rien, en revanche, n'implique que cette prévisibilité soit mieux atteinte par l'utilisation de la technique de la personnalité. Qu'on répute par exemple tous les propriétaires de robot objectivement responsables des dommages que ces derniers causent, et la prévisibilité souhaitée sera atteinte. L'on ne voit également pas bien en quoi l'utilisation de la technique de la personnalité juridique permettrait une insertion *apaisée* des robots dans le tissu social ; il nous semble, au contraire, que cette perspective pourrait effrayer la population¹⁴⁰, du moins la population occidentale¹⁴¹. Quant à l'argument tiré de la *commercialisation* facilitée du robot, si l'on écarte, comme on vient de le faire, l'argument tiré de la prétendue prévisibilité juridique que cette personnalité apporterait, on ne voit plus trop ce qui subsiste. On imagine mal, au surplus, que les fabricants de robots puissent tirer un argument commercial et marketing du fait que leur robot a la personnalité juridique : quel enfant, par exemple, demanderait avec instance à ses parents de lui acheter un robot parce que celui-ci est doué de cet attribut ?

95. Éducation des humains – Un autre argument de nature « sociétal » avancé pour défendre l'attribution d'une personnalité juridique aux robots est que leur attribuer certains droits¹⁴² contribuerait à promouvoir certaines de nos valeurs sociales et donc à éduquer la société¹⁴³. Le champ de ces droits serait limité à cet objectif : il s'agirait de droits contre la maltraitance ou des pratiques jugées répulsives. Comment a-t-on pu en venir à une telle proposition ? L'idée est simple : accorder aux humains une totale liberté d'action vis-à-vis des

¹³⁷ A. et J. BENSOUSSAN, *op. cit.* note 105, n° 202.

¹³⁸ *Ibid.*, n° 177.

¹³⁹ Argument rapporté par T. DAUPS, « Le robot, bien ou personne ? Un enjeu de civilisation ? », *LPA* 11/05/2017 p. 7 et s.

¹⁴⁰ Le non-juriste à qui l'on dit que le robot est une personne ne va-t-il pas en conclure rapidement qu'il pourrait avoir à lui rendre des comptes ?

¹⁴¹ On rappellera que la population asiatique et particulièrement japonaise semble parfaitement à l'aise avec la personnalisation des choses. Cela tient à ce que « dans la tradition shintoïste et bouddhiste, il n'y a pas de ligne de démarcation nette entre le vivant et le non vivant. Il n'y a pas de conception binaire des êtres et des choses » (S. SLAMA, « Les robots androïdes, quels droits fondamentaux », *RDLF* 2019 chron. n° 50, spéc. I. A., disponible en ligne).

¹⁴² La rhétorique de l'attribution de « droits » au robot, dans le débat public, est le plus souvent déconnectée de la question de l'attribution de la personnalité juridique, cela du fait de la méconnaissance des non-juristes du lien de nécessité entre ces deux notions ainsi que de la rhétorique répandue selon laquelle les animaux « ont des droits ». On ne peut ici toutefois que traiter cet argument en conformité avec la logique juridique et donc imputer aux partisans des droits des robots la volonté de lui reconnaître une personnalité juridique.

¹⁴³ V. K. DARLING, « Extending Legal Rights to Social Robots », *We Robot Conference*, University of Miami, April 2012, *Preliminary draft*, disponible sur Internet (l'auteur s'en tenant à l'idée de conférer des droits).

robots nous ressemblant pourrait conduire à diffuser des modèles de comportements néfastes qui seraient *in fine* préjudiciable aux humains. On a par exemple émis l'idée que défendre à un jeune enfant de maltraiter son robot animal de compagnie aurait pour vertu de lui enseigner à ne pas maltraiter de véritables animaux¹⁴⁴ ; l'impératif serait ici d'autant plus fort qu'un très jeune enfant a du mal à distinguer le robot d'un véritable animal. L'observation vaut également pour les robots humanoïdes, quand bien même un adulte ne saurait confondre un être artificiel avec un humain. La question mérite en effet d'être posée : doit-on laisser la liberté aux humains, par exemple, de maltraiter des robots humanoïdes ou de singer des pratiques de viol et de torture sur un robot sexuel ? Peut-on admettre que des humains se livrent à des pratiques sexuelles sur des robots ayant l'apparence d'un enfant ?¹⁴⁵ Ce dernier exemple suffit à souligner l'acuité du problème. La réponse adéquate ne paraît toutefois pas être de personnifier juridiquement le robot en lui accordant des droits. La création d'interdits objectifs, sous la forme d'incriminations pénales, paraît suffire autrement, cela d'autant plus que l'on pourrait, autrement, contribuer à alimenter la tendance suivante.

96. Demande sociale de protection des robots – L'argument précédent se fondait sur l'idée que donner des droits aux robots contre la maltraitance permettrait, *in fine*, de protéger les humains contre le développement de comportements dangereux ou antisociaux à l'égard des humains. Mais la volonté de protéger les robots contre la maltraitance pourrait naître directement du sentiment de malaise qui naît chez les humains à la vue d'un robot *maltraité*. Sur ce point, il n'est pas nécessaire de faire de la science-fiction. On peut déjà observer que, chez les internautes, des vidéos montrant des robots se faisant maltraiter par un humain suscitent des réactions d'indignation. On renverra le lecteur intéressé, par exemple, à la vidéo du petit dinosaure robot *Pleo* soumis à diverses tortures¹⁴⁶ ou à celle du robot humanoïde *Atlas*, malmené par ses concepteurs pour prouver ses facultés d'adaptation et d'équilibre¹⁴⁷. Ces réactions d'empathie¹⁴⁸ sont-elles absurdes ou déplacées ? Elles ne relèvent en tous cas apparemment pas de l'artifice : une étude scientifique japonaise a par exemple mis en évidence que les humains ont une réponse neuronale empathique similaire à la vue d'un stimuli de douleur appliqué à une main robotique et une main humaine¹⁴⁹. Ce que l'on constate

¹⁴⁴ K. DARLING, « Extending Legal Rights to Social Robots », art. préc. note préc., spéc. pp. 13-14.

¹⁴⁵ *Ibidem*, pp. 14-15.

¹⁴⁶ Voir la vidéo « Pleo R.I.P. » sur Youtube, vidéo d'autant plus spectaculaire que le petit robot pousse des plaintes...

¹⁴⁷ Voir la vidéo « Atlas, The Next Generation » sur Youtube.

¹⁴⁸ Sur l'utilisation de l'empathie pour favoriser l'acceptation des robots, N. NEVEJANS, *op. cit.* note 88, n° 1266 et s.

¹⁴⁹ Y. SUZUKI, *et alii*, « Measuring empathy for human and robot hand pain using electroencephalography » *Sci. Rep.* 5, 15924, (2015) (doi: 10.1038/srep15924, disponible en ligne).

ici est une forme d'anthropomorphisation de la machine animée, effet d'autant plus fort que les robots sociaux¹⁵⁰ – qu'il s'agisse de faux animaux ou de robots humanoïdes – sont expressément conçus pour susciter des projections anthropomorphiques¹⁵¹, en imitant des signaux que nous associons automatiquement à certains états d'esprit ou sentiments¹⁵². Ce sont ainsi nos propres sentiments d'empathie qui pourraient conduire à développer des règles de protection des robots contre des actes de maltraitance. Mais ces sentiments, à défaut de pouvoir être supprimées, méritent-ils d'être en quelque sorte encouragés en leur donnant le prolongement juridique substantiel que constituerait la personnalisation de ces robots ? Formuler la question, c'est y répondre : ce serait concrètement encourager la population à prendre des vessies pour des lanternes. L'argument de l'empathie et de la compassion peut amener de surcroît à faire un parallèle avec la lutte menée par certains pour l'attribution de droits aux animaux.

97. Parallèle avec les animaux – La lutte, ou du moins la proposition d'attribuer des droits et une personnalité juridique aux animaux, est plus ancienne que celle qui concerne les robots : Demogue évoquait déjà cette solution en 1909¹⁵³. Cette lutte peut se prévaloir aujourd'hui de quelques victoires discrètes. En 2014, par exemple, la chambre fédérale de cassation pénale d'Argentine a reconnu à un orang-outan prénommé Sandra la qualité de « personne non humaine » et conclu que la privation de liberté dont elle faisait l'objet était illégale¹⁵⁴. La reconnaissance par la loi du 16 février 2015 à l'article 515-14 du Code civil français de ce que les animaux sont des « êtres vivants doués de sensibilité » pourrait être également interprétée comme une marche discrète dans cette direction, quoique la seconde phrase de cet alinéa, qui réaffirme que les animaux sont, sous réserve des lois qui les protègent, soumis au droit des biens, semble précisément vouloir condamner une telle

¹⁵⁰ Il faut noter que cet effet se constate même à l'égard d'un robot non social, tel un aspirateur automatisé. Sur ce thème, v. not. M. SCHEUTZ, «The Inherent Dangers of Unidirectional Emotional Bonds between Humans and Social Robots », in *Robo-Ethics, The ethical and social implications of robotics*, dir. P. LIN, K. ABNEY et G. A. BEKEY, The MIT Press, 2012, pp. 205-222 ; J.-Y. SUNG, L. GUO, R. E. GRINTER, H. I. CHRISTENSEN, « My Roomba Is Rambo”: Intimate Home Appliances », 9th International Conference on Ubiquitous Computing, Innsbruck (Sept 2007) (disponible en ligne).

¹⁵¹ K. DARLING, « Extending Legal Rights to Social Robots », art. préc. note 143, spéc. p. 2.

¹⁵² *Ibidem.*, spéc. p. 5.

¹⁵³ R. DEMOGUE, « La notion de sujet de droit », *RTD civ.* 1909.611-655, spéc. p. 637 et s. Sur ce sujet, J.-P. MARGUENAUD, « Actualité et actualisation des propositions de René Demogue sur la personnalité juridique des animaux », *Revue juridique de l'environnement*, 2015/1 Vol. 40, pp. 73-83.. Sur la personnalisation de l'animal, v. l'étude fouillée de J.-P. MARGUENAUD, *L'animal en droit privé*, PUF, 1992, spéc. pp. 361-430 et « La personnalité juridique des animaux », *D.* 1998 chr. p. 205.

¹⁵⁴ Cámara federal de casación penal, Sala II, Orangutana Sandra s/ Recuso de Casación s/ Habeas Corpus, CCC 68831/2014/CFC001, 18 décembre 2014. Hasard du calendrier, la troisième division d'appel de la Cour suprême de New York avait dénié quelques jours plus tôt la personnalité juridique à des chimpanzés au motif qu'ils sont incapables d'assumer les responsabilités corrélatives à la titularité de droits (*People ex rel. Nonhuman Rights Project, Inc. v Lavery*, 124 AD3d 148 [3d Dept 2014]).

extension. A moins que cette personnification de l'animal ne soit pas déjà réalisée dans les faits, si l'on suit la proposition du Professeur Marguénaud de transposer aux animaux la théorie de la réalité technique des personnes morales¹⁵⁵.

98. Les positions parmi les partisans d'une personnalité juridique animale sont variées. Certains y voient un simple instrument technique utile tirant les conséquences de la protection actuelle dont bénéficient certains animaux et permettant de leur conférer à certains droits limités additionnels¹⁵⁶ ; d'autres militent pour que l'on reconnaisse aux grands singes des « droits-dignité » dans le cadre d'une personnalité que certains observateurs qualifient de personnification anthropomorphique¹⁵⁷. D'autres encore, rattachables à cette dernière tendance, ont une vision encore plus large en proposant¹⁵⁸ de reconnaître à toutes les espèces sensibles certains droits fondamentaux comme celui de ne pas être tué¹⁵⁹, torturé, mutilé ou enfermé. On ne cherchera évidemment pas ici à prendre parti sur ce débat ; ce qui est importé ici est ce qu'il peut apporter à la réflexion sur la personnalité juridique des robots.

99. Sur ce point, on observera qu'au-delà de la simple prétendue commodité technique que pourrait apporter le recours à cette technique, la proposition de doter les animaux d'une personnalité juridique part toujours, fut-ce implicitement, du constat suivant : les animaux sont doués d'une sensibilité et, à cet égard, il convient de les protéger. Comme l'expose le Professeur Marguénaud, hostile à la personnification anthropomorphique¹⁶⁰ mais fervent défenseur de la personnification purement technique, « la "personnalité animale" [a] vocation [à] la défense de l'intérêt physiologiquement et psychologiquement ressenti par la bête en tant qu'être de chair et de sang. [Elle est] une technique juridique mise au service d'êtres corporels qui ne peuvent que souffrir ; c'est, plus précisément, une technique

¹⁵⁵ J.-P. MARGUENAUD, « La personnalité juridique des animaux », art. préc. note 153, spéc. II B, résumant les propos que l'on trouve dans son ouvrage *L'animal en droit privé, op. cit.* note 153, p. 396 et s.

¹⁵⁶ C'est la proposition de J.-P. MARGUENAUD.

¹⁵⁷ V. D. CHAUVET, « Quelle personnalité juridique est digne des animaux ? », *Droits* 2015/2 (n° 62), pp. 217-234 (DOI 10.3917/droit.062.0217), analysant les travaux de Steven WIZE à la lumière de la distinction proposée par J.-P. MARGUENAUD.

¹⁵⁸ V. par ex. l'interview de la philosophe F. BURGAT par J. HARAU dans le quotidien *Libération* du 30 décembre 2014 sous l'intitulé « Il faut réviser notre façon de traiter les animaux ».

¹⁵⁹ Perspective dont on peine à croire qu'elle puisse être un jour atteinte puisque cela signifierait notamment la fin de la consommation de viande animale.

¹⁶⁰ J.-P. MARGUENAUD, « La personnalité juridique des animaux », art. préc. note 153, spéc. p. 207 : « Conférer aux animaux une personnalité juridique à taille humaine (...) serait une promotion inadaptée parce qu'elle leur conférerait des droits inutiles et les exposerait à des obligations grotesques ».

de protection individuelle des êtres non humains qui souffrent »¹⁶¹. A cet égard, le parallèle avec la situation des robots autonomes ayant vocation à se démocratiser dans notre quotidien ces prochaines années n'est simplement pas soutenable. Les animaux sont des êtres biologiques mus tout autant que nous par le « désir » d'éviter des souffrances naturelles comme la faim, la soif, le froid, etc. ; il apparaît par ailleurs de plus en plus clair qu'un grand nombre d'entre eux peuvent être sujets à des troubles psychologiques tout à fait similaires aux nôtres. Et, après tout, nous ne sommes en fin de compte que des animaux très évolués. Il y a, pour genre de raisons, une forme de *dignité* qui s'attache aux animaux et dont on peut comprendre que certains souhaitent la traduire ou la consacrer par une personnalité juridique¹⁶². Mais telle n'est pas la situation des robots doués des fonctions d'autonomie. Faire le parallèle avec les animaux c'est, comme le font certains internautes s'indignant du mauvais traitement du robot Pleo, leur prêter dans un élan quasi-anthropomorphe des réactions de souffrance ou de détresse qu'ils n'ont évidemment pas. A cela, on ajoutera que, contrairement à la personnalité animale qui paraît nécessairement devoir rester cantonnée à certaines limites, la personnalité juridique des robots porte en germe un abus caractérisé d'anthropomorphisme du fait que la chose a été créée pour nous ressembler.

100. Bilan de la situation et des perspectives actuelles – Après examen, les différents arguments avancés pour justifier l'attribution d'une personnalité juridique aux robots humanoïdes peinent à convaincre. D'un point de vue technique, l'utilisation de cette construction juridique paraît non seulement inutile mais aussi contreproductive. On ne saurait par ailleurs fonder cette attribution sur la « liberté décisionnelle » du robot : la latitude fonctionnelle laissée à une machine pour accomplir une tâche ne saurait être comparée à la liberté humaine et ne mérite aucune forme de protection qui puisse s'approcher de la dignité attachée à cette dernière. Enfin, les arguments sociétaux étonnent : la personnalité juridique ne présente aucun intérêt pour la commercialisation des robots et le reste des préoccupations invoquées ne nécessitent aucunement le recours à cette figure juridique.

En définitive, l'ensemble de ces arguments semble être au service d'une justification implicite : c'est en fin de compte parce que l'on crée une chose qui nous ressemble extérieurement que germe l'idée de la personnifier sur le plan juridique¹⁶³. Il ne semble s'agir, ni plus ni moins, que d'animer un

¹⁶¹ J.-P. MARGUENAUD, *L'animal en droit privé*, op. cit. note 153, spéc. p. 405, l'auteur (qui est à l'origine des différents soulignements) formulant une opposition avec la personnalité morale ; v. aussi K. DARLING, art. préc. note 143, spéc. p. 11 et 13.

¹⁶² Sur cette question, D. CHAUVET, « Quelle personnalité juridique est digne des animaux ? », art. préc. note 157.

¹⁶³ En ce sens que l'anthropomorphisme est la base d'une telle revendication : X. BIOY, « Vers un statut juridique des androïdes ? », *Journal International de bioéthique* 2013/4 vol. 24 pp. 85-98, spéc. p. 88 et s. ; rapp. G. LOISEAU, art. préc. note 71, conclusion du I., en ce sens que la personnalité n'est qu'un « alibi » qui sert une proposition de nature idéologique convoitant la valeur symbolique du statut de personne ; l'auteur avance également que cela conduira à vouloir par égalitarisme leur conférer des droits égaux aux personnes humaines.

fantasme nourri par la science-fiction selon lequel cette chose qui nous ressemble, capable de mouvements similaires aux nôtres et pouvant donner l'illusion de soutenir une conversation mériterait un respect du même niveau que celui qu'il convient d'accorder à un véritable être humain. Or la perspective de machines qui puissent être réellement considérées comme nos « égaux artificiels » paraît bien utopique... à moins qu'elle ne soit simplement que très lointaine ?



II. UNE PROPOSITION EVENTUELLEMENT JUSTIFIABLE DEMAIN ?

101. Perspectives de très long terme – Nous venons de voir qu'en l'état actuel et immédiatement prévisible de la robotique, la proposition de conférer aux robots une personnalité juridique apparaît concrètement dénuée de justification sérieuse. Si cette opinion se défend assez aisément au premier quart du XXI^e siècle, sera-t-elle aussi facilement défendable au début du siècle prochain ?

Si nous sommes à l'heure actuelle encore bien loin de la machine qui, par son apparence, son comportement et surtout ses réflexions pourra nous sembler être dirigée par une intelligence quasi-humaine au point de nous faire douter de son statut ontologique, cette perspective ne saurait pour autant être considérée comme entièrement utopique. Qu'on le veuille ou non – et que cela soit souhaitable ou non – la recherche dans le domaine de la robotique est irrémédiablement orientée vers le développement de techniques et d'algorithmes avancés dont l'objet est la simulation partielle du fonctionnement du cerveau ou de certains traits de l'intelligence humaine, avec des perspectives d'en réaliser une simulation totale. (A). Il semble ainsi inéluctable que, tôt ou tard, le logiciel que l'on mettra aux commandes d'un robot humanoïde finira par donner tous les apparences, par ses initiatives et ses réactions, d'une *pensée* humaine. Ce n'est qu'arrivé à ce stade que les conditions seront réunies pour que l'on puisse se poser la seule question dont il est rationnel de faire dépendre l'attribution à ces machines d'une personnalité juridique: la technique que l'on aura créée pour parvenir à ce résultat pourra-t-elle être jugée avoir créé un *état de conscience* chez la machine (B) ?

A. Les perspectives technologiques

102. État de l'art – Si l'on ne peut préjuger des technologies qui seront à l'œuvre dans une cinquantaine ou une centaine d'années, les recherches en cours permettent d'ores et déjà de prendre conscience des directions poursuivies en vue de réussir de doter les machines de certains aspects de l'intelligence humaine, voire même de simuler le fonctionnement du cerveau humain avec la perspective implicite ou explicite de reproduire, *in fine*, la pensée humaine. Le seul moyen de donner un peu de crédit à cette perspective étant de prendre conscience de la nature des travaux qui pourrait y mener, on tentera ci-dessous de donner un aperçu de l'état actuel de certaines réflexions en la matière à travers trois exemples.

Certaines recherches ont d'ores et déjà abouti à des techniques d'intelligence artificielle nées de l'observation du fonctionnement des neurones (1°). Un projet de recherche ambitieux par ailleurs de simuler et reproduire fidèlement le fonctionnement du cerveau humain (2°). Enfin, un auteur a proposé un ensemble de bases théoriques qui pourraient être utilisées pour créer une machine pensante (3°).

1°) Les neurones artificiels

103. Origine et intérêt – La technique des neurones artificiels, théorisée dans les années 1940¹⁶⁴, d'ores et déjà connue du grand public du fait qu'elle est à la base d'un processus auquel il est abondamment référé lorsqu'il est question d'intelligence artificielle, et qui se trouve au demeurant d'ores et déjà utilisé dans ces objets de grande consommation que sont nos téléphones portables : le *deep learning* ou apprentissage profond. C'est en effet sur la base de cette technique que nos téléphones sont par exemple capables de reconnaître, dans nos photos, des individus, des animaux ou des objets.

Si les apports théoriques qui ont permis le développement de ces réseaux de neurones artificiels ont été réalisés dans la seconde partie du XX^e siècle, ce n'est qu'à partir des années 2010 que l'apprentissage profond a pris son essor grâce à la conjugaison de deux facteurs : l'importante augmentation des jeux de données numériques nécessaires pour entraîner un réseau de neurones artificiels et l'atteinte d'une puissance de calcul suffisante pour réaliser ces entraînements.

104. Principes de fonctionnement – D'un point de vue concret, ces « neurones » sont des fonctions mathématiques reliées entre elles. Cette dénomination provient de ce que leur schéma de fonctionnement, comme leur nom l'indique de manière transparente, a été modelé sur l'observation du fonctionnement du neurone biologique. Tout comme ce dernier émet un signal électrique si les signaux envoyés par les neurones qui le précèdent atteignent une certaine intensité, la fonction mathématique symbolisée par l'idée de neurone artificielle retournera une valeur de sortie si les valeurs d'entrées pondérées qui lui sont soumises atteignent, lorsqu'elles sont combinées, une certaine valeur. La fonction d'apprentissage qui résulte du fonctionnement de l'ensemble est permise par la logique de renforcement des liens entre neurones artificiels « excités » conjointement, toujours selon le modèle du renforcement du lien synaptique pouvant exister entre deux neurones biologiques.

105. Applications et avenir – A l'heure actuelle, les réseaux de neurones artificiels sont d'ores et déjà utilisés dans de nombreux domaines. Les cas d'application les plus apparents pour le grand public sont sans doute la reconnaissance d'images (ainsi, la reconnaissance des individus présents sur

¹⁶⁴ Ces neurones artificiels ont été théorisés par W. MCCULLOCH et W. PITTS, « A logical calculus of the ideas immanent in nervous activity », *Bulletin of Mathematical Biophysics* 5, 115–133 (1943).

une photo ou la reconnaissance de l'écriture manuscrite) ainsi que la reconnaissance vocale (ainsi, la fonction de dictée d'un téléphone), deux compétences que l'on aurait classifiées jusqu'à récemment comme éminemment humaines et qui maintenant peuvent être accomplies par des machines avec un haut degré de réussite. Mais cette technique est également mise en œuvre dans des cas d'application aussi divers que la prédiction de la météo, de molécules chimiques en développement ou des marchés financiers.

106. La technique de neurones artificiels constituera-t-elle le fondement de l'élaboration d'une machine qui, à terme, sera capable de fonctions avancées de réflexion ? Si le néophyte aurait très certainement tendance à le penser, telle n'est cependant pas l'ambition des chercheurs en ce domaine. Ces réseaux de neurones, bien qu'ayant été conceptualisés sur le modèle de leurs cousins biologique, ne prétendent en effet pas ni n'ambitionnent reproduire le fonctionnement du cerveau humain : ils constituent « simplement » un moyen d'entraîner un programme à reconnaître des patrons – *patterns* – de manière très performante. La simulation d'un acte de réflexion ou de pensée nécessiterait des bases théoriques bien plus considérables. Il reste qu'en empruntant conceptuellement la logique qui gouverne le fonctionnement primaire des neurones biologiques afin de produire des résultats qui, pour l'accomplissement de tâches spécifiques telles la reconnaissance d'une tumeur cancéreuse, dépassent aujourd'hui semble-t-il les performances d'un cerveau humain, les spécialistes en intelligence artificielle ont réussi à simuler un pan spécifique de l'intelligence humaine en prenant le cerveau humain pour modèle. La qualité et la rapidité de ces algorithmes d'apprentissage ne pouvant que s'améliorer avec la recherche ainsi que le développement de la puissance de calcul à disposition de l'espèce humaine, on peut conjecturer aisément que, dans un futur lointain, les descendants de ces algorithmes constitueront l'un des pans d'un système plus général d'intelligence artificielle se rapprochant des capacités d'apprentissage et de raisonnement d'un cerveau humain.

Si les réseaux de neurones artificiels ne prétendent pas reproduire le fonctionnement du cerveau humain, il existe un projet qui s'est donné pareille ambition : le projet *Whole Brain Emulation*.

2°) Le projet *Whole Brain Emulation*

107. But – L'objectif du projet *Whole Brain Emulation*¹⁶⁵ (ci-après *WBE*), dont les buts et jalons ont été explicités dans un livre blanc daté de 2008 et intitulé *Whole Brain Emulation, A Roadmap*¹⁶⁶ consiste à scanner¹⁶⁷ un

¹⁶⁵ On peut mentionner un autre projet scientifique dédié à l'étude du cerveau humain, que l'on ne fera que mentionner ici : le *Human Brain Project* <<https://www.humanbrainproject.eu>>.

¹⁶⁶ A. SANDBERG & N. BOSTROM, *Whole Brain Emulation, A Roadmap*, Technical Report #2008-3, Future of Humanity Institute, Oxford University.

¹⁶⁷ Le verbe « scanner » tout comme le substantif « scannage » fleure bon l'anglicisme, toutefois la langue française n'a pas su se doter d'un équivalent acceptable, le terme de

cerveau humain en détail puis à construire, sur le fondement des données recueillies, un logiciel qui, s'exécutant sur un matériel approprié, se comporterait en substance de la même manière que le cerveau original¹⁶⁸. Le logiciel en question entend *émuler* le cerveau humain, c'est-à-dire reproduire l'ensemble des propriétés du cerveau modèle¹⁶⁹.

La simplicité de l'idée ne doit pas masquer la grande complexité que présente sa réalisation et l'un des grands mérites du livre blanc précité est d'explorer méthodiquement les technologies envisageables pour réaliser ce projet, les difficultés prévisibles ainsi que les différentes voies qui s'ouvriront en fonction des résultats partiels de sa mise en œuvre, tout en réservant l'apport des développements parallèles qui se produiront dans le domaine des sciences cognitives.

108. L'acquisition des données – La première étape – la numérisation à haute résolution d'un cerveau humain – se heurte à deux contraintes essentielles. La première tient à la technologie qui sera utilisée pour acquérir les données. L'étude suggère¹⁷⁰ en effet qu'une résolution de 5×5×50 nanomètres sera très certainement nécessaire pour obtenir le degré de détail requis pour émuler les processus primitifs dont il s'agit de reproduire les fonctionnalités. Or, s'il existe des technologies susceptibles d'atteindre de telles résolutions, la plupart sont destructives en ce sens qu'elle requiert la découpe du cerveau source en lamelles et ne peuvent donc être pratiquées sur un organe vivant, ce qui est susceptible d'empêcher de saisir le fonctionnement de certains processus qu'il s'agit d'émuler¹⁷¹. Une seconde contrainte tient au volume des données qui seront engendrées ainsi qu'à leur traitement ; les derniers développements de la technologie informatique laissent penser que ce second point sera surmonté sans trop de difficulté.

109. L'émulation logicielle – Une fois les données convoitées acquises avec le degré de précision et de fidélité envisagé, la seconde étape – la plus longue et la plus difficile – consistera à créer un logiciel capable de reproduire les propriétés résultant des structures anatomiques numérisées. Pour ce faire, les auteurs de ce livre blanc estiment qu'il n'est pas nécessaire de comprendre tout le système d'origine, mais « simplement » de disposer d'une base de données contenant toutes les informations de bas niveau à propos du cerveau ainsi que de la connaissance des mises jour locales modifiant les états du cerveau d'un moment à un autre¹⁷². Ce logiciel nécessitera un modèle

« balayage » étant assez réducteur. Quant au substantif *numérisation*, il désigne plutôt le processus par son résultat.

¹⁶⁸ A. SANDBERG & N. BOSTROM, *Whole Brain Emulation, A Roadmap*, *op. cit.* note 166, p. 7.

¹⁶⁹ *Ibidem*, p. 7.

¹⁷⁰ *Ibid.*, p. 14.

¹⁷¹ Sur ces technologies et leur difficultés, *ibid.* p. 40 et s.

¹⁷² *Ibid.*, p. 8.

mathématique de l'activité neuronale ainsi que les moyens de mettre en œuvre efficacement ce modèle sur des ordinateurs¹⁷³. La construction de ce modèle logiciel relève du domaine des « neurosciences computationnelles », discipline s'attachant à construire des algorithmes susceptibles de refléter le traitement de l'information par notre système nerveux et aboutissant à la mise en œuvre de nos fonctions cognitives. Le WBE requiert plus spécifiquement des modèles décrivant les processus chimiques internes aux neurones ainsi que des modèles de l'expression génétique et des mécanismes épigénétiques susceptibles de rendre compte de la plasticité neuronale¹⁷⁴.

On décèle ici une importante différence avec les recherches concernant les neurones artificiels : s'agissant de ces derniers, toute invention mathématique susceptible d'en optimiser le fonctionnement et les performances sera adoptée si elle est efficace pour accomplir le résultat voulu. Le projet WBE ambitionne au contraire de refléter le mieux possible, d'un point de vue logiciel, le fonctionnement des neurones humains, de même que ses interactions avec le reste du corps humain, plus précisément des signaux neuronaux impliqués par son existence et l'environnement dans lequel il est plongé¹⁷⁵.

110. Matériel utilisé – La question du matériel qui sera utilisé pour faire tourner cette simulation reste très ouverte du fait de sa dépendance aux évolutions technologiques rapides dans le domaine des processeurs. Les auteurs envisagent la possibilité d'utiliser du matériel dédié pour obtenir des gains de performance appréciables¹⁷⁶. Concrètement, il s'agirait de « coder en dur »¹⁷⁷ certaines fonctions logicielles de base dans le but d'optimiser grandement les performances de calcul, une technique d'ores et déjà usuelle¹⁷⁸ mais dont l'opportunité sera dépendante de la capacité de l'industrie informatique à trouver à ces optimisations matérielles spécifiquement orientées vers la simulation de l'activité neuronale biologique¹⁷⁹ une utilité économique.

111. Jalons – De manière réaliste, les auteurs du libre blanc envisagent en premier lieu des objectifs modestes, dont l'atteinte conditionnera la réalisation d'objectifs plus ambitieux. Quelques grands jalons sont la

¹⁷³ *Ibid.*, p. 16.

¹⁷⁴ *Ibid.*, p. 64 et 65.

¹⁷⁵ *Ibid.*, p. 16 et 74.

¹⁷⁶ *Ibid.*, p. 27.

¹⁷⁷ L'expression désigne le fait d'incorporer directement dans les unités de calcul d'un processeur des opérations logiques précises dont l'exécution était auparavant réalisée par l'intermédiaire d'un logiciel.

¹⁷⁸ Cette technique est actuellement utilisée dans les processeurs pour accélérer le décodage de vidéos ou encore les processus de *deep learning*, v. par exemple le « neural engine » des processeurs produits par la société Apple.

¹⁷⁹ Une discipline que l'on appelle l'ingénierie neuromorphique (*Neuromorphic Engineering*).

simulation d'une cellule, l'émulation partielle d'un système nerveux comme la rétine, puis l'émulation complète d'un organisme simple comme un vers, l'émulation d'un organisme invertébré comme un escargot ou un insecte, puis l'émulation d'un petit mammifère (souris, rat), d'un grand mammifère et enfin l'émulation d'un cerveau humain.

112. Simulation de l'intelligence humaine ? – Le projet WBE se donnant pour ambition d'émuler méticuleusement le fonctionnement du cerveau humain dans son moindre détail, il semblerait naturel d'en conclure qu'il ambitionne à terme, d'émuler une intelligence artificielle qui serait l'égale de celle de l'humain. Telle n'est pourtant pas l'ambition explicite ni immédiate de ses partisans. Ces derniers exposent à ce sujet que le WBE « ne présume pas de la nécessité d'une compréhension de haut niveau du cerveau ou de l'esprit. En fait, si une telle compréhension devait être atteinte, il est probable qu'elle pourrait être utilisée pour produire une intelligence artificielle (IA). L'IA de niveau humain (ou IA super-intelligente) n'empêcherait pas nécessairement le WBE, mais certaines des raisons scientifiques et économiques disparaîtraient, ce qui pourrait rendre le domaine moins pertinent. D'un autre côté, une IA puissante pourrait accélérer considérablement les progrès des neurosciences et peut-être aider à développer le WBE à d'autres fins. Inversement, le succès de certaines parties du projet WBE pourrait aider l'IA, par exemple si les microcircuits corticaux et les règles d'apprentissage pouvaient être simulés efficacement en tant que système général d'apprentissage/comportement ». Ainsi le projet WBE n'est-il pas conçu dans le but de créer une IA forte, mais pourra très certainement y contribuer de manière substantielle.

113. Un autre effort intellectuel, particulièrement intéressant pour tenter d'appréhender la nature de la situation en considération à laquelle les juristes de demain devront se prononcer, s'est au contraire tout entier tourné vers l'objectif de créer une machine susceptible de produire une pensée de type humaine.

3°) Les travaux d'Alain Cardon

114. Ambition – Les neurones artificiels ont été inspirés par le fonctionnement des neurones biologiques, mais ne visent pas à les reproduire ; le WBE tente au contraire d'émuler fidèlement le fonctionnement neuronal du vivant, mais sans se donner spécifiquement pour but l'émulation de l'intelligence ou de la pensée humaine. M. Alain Cardon, dans son ouvrage *Modéliser et concevoir une machine pensante*¹⁸⁰, se fixe quant à lui d'emblée pour objectif de déterminer les éléments nécessaires à la création d'une machine pensante. L'ambition n'est pas toutefois de fournir un modèle précis

¹⁸⁰ A. CARDON, *Modéliser et concevoir une machine pensante, approche de la conscience artificielle*, Vuibert, 2004.

de la pensée humaine telle qu'elle se construit dans le cerveau humain¹⁸¹ : il ne s'agit pas de simuler le fonctionnement des neurones, mais de recourir à des agents logiciels constituant des entités symboliques complexes susceptibles de produire un résultat du même genre¹⁸². L'idée de M. Cardon, *in fine*, est de disposer d'une machine qui puisse élaborer ses propres idées et agir en fonction de ces idées et non de manière purement automatique. Le système doit prendre conscience de ce qu'il a formulé une idée¹⁸³ et à cet égard il aura une propriété « d'auto-observation de l'organisation des états du système, et ce pour pouvoir définir la conscience de soi »¹⁸⁴.

115. Qu'est-ce penser ? Ambitionner de créer une machine pensante suppose déjà de s'interroger sur ce qu'est la pensée. Sur ce point, M. Cardon observe que s'est imposée l'idée que la mainmise de l'Homme sur sa planète s'expliquerait par le fait qu'il est l'unique possesseur du pouvoir de la pensée ; les animaux, eux, en sont jugés dépourvus et uniquement dominés par leur instinct. L'auteur s'inscrit en faux contre cette conception : « les animaux dits supérieurs ressentent, pensent, évaluent des situations parfois complexes sur leurs intérêts et les possibilités de sensation et d'action de leurs corps »¹⁸⁵. La discipline de l'intelligence artificielle, qui repose sur cette conception de l'Homme doué d'une raison supérieure, a par ailleurs réussi à désincarner l'intelligence : « penser est, selon cette discipline, une pratique abstraite, langagière, très sophistiquée et surtout non sensible. Le raisonnement artificiel est symbolique, fondé sur des logiques complexes et même très complexes. *Mais personne, nulle part dans le vivant, ne pense comme ça.* On pense parce qu'on ne peut pas faire autrement et l'on pense en général à des choses très banales, liées à des situations très courantes, en éprouvant des émotions et en se trompant la plupart du temps »¹⁸⁶. On ne saurait donc prétendre simuler la pensée si l'on se contente de vouloir construire une capacité à raisonner¹⁸⁷.

S'interrogeant ensuite sur les rapports entre la pensée – non matérielle – et le cerveau, l'auteur observe qu'il est usuel de croire qu'il y a une différence de genre radicale et absolue entre les deux ; qu'il n'y a pas de rapport direct entre un ensemble de neurones et ce à quoi il est pensé en recourant au cerveau¹⁸⁸. Pour M. Cardon, la nature réelle de ce rapport est la suivante : « le réseau de neurones est un support *permettant de représenter des mouvements*, des activités à la fois locales et relationnelles, des flux de mouvements localisés dans des zones précises, qui sont fortement communicantes. Et la pensée est

¹⁸¹ *Ibidem*, p. 137.

¹⁸² *Ibid.*, pp. 69 et 166.

¹⁸³ *Ibid.*, p. 136.

¹⁸⁴ *Ibid.*, p. 148.

¹⁸⁵ *Ibid.*, p. 56.

¹⁸⁶ *Ibid.* (nous soulignons).

¹⁸⁷ *Ibid.*, p. 57, p. 77, p. 79 et s.

¹⁸⁸ *Ibid.*, p. 68.

là, elle n'est que là, dans ces mouvements très complexes »¹⁸⁹. La pensée – qu'il définit comme représentation intentionnelle et perçue de quelque chose¹⁹⁰ – est ainsi analysée comme un *mouvement* survenant dans des ensembles spécialisés que sont les aires neuronales, ensembles pouvant évoluer, se reconfigurer ou changer de fonction¹⁹¹.

116. L'hypothèse de la pensée calculable – Partant de là, l'idée maîtresse de l'ouvrage de M. Cardon est que la pensée est un processus susceptible d'être produit dans le domaine du calculable¹⁹². « [L']informatique peut générer de la pensée si elle peut représenter des mouvements sur ses supports habituels, si elle peut représenter des calculs distribués, communicants, concurrents, sur le support de systèmes très plastiques comme les systèmes multiagents massifs, s'exécutant sur des grappes de machines »¹⁹³ ; cet effort devant être nourri par les apports de la neurobiologie, la psychologie, et les sciences cognitives¹⁹⁴.

Envisageant les processus informatiques susceptibles de produire des pensées, M. Cardon refuse l'approche behavioriste, qui est déterministe et selon laquelle il serait possible d'élaborer une fonction qui, pour un état initial du sujet – la situation de l'individu à un certain moment – produirait l'état suivant – la réaction de l'individu en pensant, face à cette situation et déterminant son éventuelle action dans le futur¹⁹⁵. L'auteur estime au contraire que l'acte de penser se caractérise par une complète liberté de génération et postule qu'il « doit y avoir en fait un processus de construction de la fonction génératrice de pensées, génératrice à chaque instant, et un processus qui englobe l'état initial dans sa définition même. La génération d'une pensée peut être appréhendée d'une façon calculable si l'on admet que la fonction qui générerait un état de pensée artificielle peut se déterminer elle-même, pour chaque pensée, à chaque fois que le système produit de la pensée »¹⁹⁶. On se risquera à résumer l'idée de la manière suivante : il s'agit de créer une fonction complexe ayant la propriété de s'auto-redéfinir à mesure qu'elle produit des résultats, ce qui rend ces résultats à la fois non-prédictibles mais pourtant en la maîtrise du système qui les produit¹⁹⁷. Il s'agit de « concevoir les mouvements

¹⁸⁹ *Ibid.* (nous soulignons).

¹⁹⁰ *Ibid.*, p. 74.

¹⁹¹ « La pensée est mouvement dans et de ces ensembles spécialisés, mouvements de communication et d'activation des amas de neurones. Elle est donc aussi influx, transfert de charges représentant le mouvement effectif. C'est essentiellement dans le mouvement de l'influx, supporté par le substrat des neurones regroupés en aires, que la pensée se fera » (*ibid.*, p. 73).

¹⁹² *Ibid.*, p. 2.

¹⁹³ *Ibid.*, p. 69.

¹⁹⁴ *Ibid.*, p. 135.

¹⁹⁵ *Ibid.*, p. 2.

¹⁹⁶ *Ibid.*, p. 138.

¹⁹⁷ « Il s'agit de concevoir, et de construire un système qui agit pour son compte, qui construit des représentations aux caractères très changeants, des représentations qui se déroulent

d'une organisation à la recherche délicate de sa stabilité », le système étant caractérisé par « un fonctionnement non stable de relations entre une structure complexe très plastique et un extérieur très différent interprété pour produire des références »¹⁹⁸. L'auteur définit ainsi la pensée artificielle comme « un état émergent produit par un ensemble de programmes fortement communicants, en fait comme la forme de cette organisation dynamique de programmes communicants très nombreux en exécution et en auto-observation »¹⁹⁹.

117. L'organisation logicielle simulant la pensée – D'un point de vue technique, M. Cardon envisage la production d'un acte de pensée artificielle par le recours à la coactivité de deux systèmes. Le premier regroupe des composants qualifiés d'*agents aspectuels*, lesquels sont des « formes dynamiques » décrivant des processus primaires et constituant les entités de base permettant de construire le système²⁰⁰. Ces agents communiquent entre eux et sont regroupés en ensembles spécifiques aux fonctionnalités précises et eux-mêmes communicants²⁰¹. Un second système, regroupant des *agents morphologiques*, a lui pour objectif d'exprimer simultanément « la conformation, la morphologie, l'état organisationnel de ces groupes d'agents aspectuels »²⁰² ; ce dernier système influe sur la dynamique du système qu'il observe : « la pensée est une co-activation autoréférente dans un double système dynamique »²⁰³. Globalement, « les deux systèmes par leurs co-activations incessantes, se stabilisent sur un état de concordance où le premier système est représenté dans l'autre par sa morphologie, et qui est l'état de pensée artificielle. Il y a une état de pensée, qui est saturant un instant pour l'organisme »²⁰⁴. Le système étant conçu comme en altération continue, il se déploie sans cesse par une succession d'états particuliers, qui sont autant d'états de pensée. Il est au demeurant conçu comme ayant la propriété de

comme des scènes, sur de multiples niveaux, qui sont des activités locales et générales, qui se renvoient leurs activités à elles-mêmes et qui forment un tout » (*ibid.*, p. 164).

¹⁹⁸ *Ibid.*, p. 145.

¹⁹⁹ *Ibid.*, p. 148.

²⁰⁰ Un agent aspectuel est défini plus précisément comme un « élément minimal [qui] doit représenter à la fois un trait de caractère élémentaire, une partie de connaissance à propos des émotions ou des pensées. Il doit être actif, autonome, proactif et fortement coopératif, il doit avoir des caractères organisationnels précis et définis essentiellement par son activité, il a des aspects de caractère perceptif » ; « c'est un agent logiciel (...) léger, qui décrit un aspect élémentaire d'une propriété comportementale d'un organe de l'organisme ou un trait sémantique relatif à un concept. L'activation de cet agent reviendra à prendre effectivement en considérations, à un moment, un aspect de la propriété qu'il représente » (A. CARDON, *op. cit.* note 180, p. 166).

²⁰¹ *Ibid.*, p. 5.

²⁰² *Ibid.*

²⁰³ *Ibid.*, p. 6.

²⁰⁴ *Ibid.* ; v. aussi p. 75, à propos de la pensée humaine : « Une pensée ressentie est une forme claire, distinctive au-dessus d'un mouvement informe de groupes de neurones. C'est une émergence organisationnelle ».

s'auto-observer, ce qui permet de réaliser la notion de réaction, d'intention à agir et à penser, à appréhender ces actions²⁰⁵.

118. Un corps support de la pensée – Un autre hypothèse posée par M. Cardon pour la réussite de la conception d'une machine pensante est qu'« il n'y a pas de pensée artificielle sans corps matériel »²⁰⁶. « L'animal ainsi que l'enfant nouveau-né, pensent strictement selon leur corps, avant que de penser à l'être et à des abstractions »²⁰⁷. Il est au demeurant nécessaire à l'Homme de ressentir certaines choses du réel pour concevoir certains concepts²⁰⁸. Une machine dont on prétend qu'elle aurait la capacité de penser se doit donc d'être pourvue d'un corps qui l'informe à chaque instant sur de multiples aspects du monde sensible qui l'entoure, tout comme notre propre corps nous renseigne continuellement sur notre environnement et influe sur le cours ou la nature de notre pensée. Il faut ainsi concevoir un robot pourvu de motricité, d'organes de préhensions et de nombreux capteurs lui permettant de saisir les informations visuelles, sonores, tactiles de son environnement et de mesurer les efforts exercés par ses différents organes dans leur mouvement²⁰⁹. L'ensemble de ces informations auront vocation participer à l'ensemble de la mécanique logiciel débouchant sur l'acte de penser. Simplement, contrairement aux organismes vivants, ce corps pourra ne pas être unique mais « distribué » sur multiples robots, possiblement très éloignés, l'intégration se faisant au niveau logiciel ; c'est l'hypothèse du « multi-robot »²¹⁰.

119. La simulation des émotions – Dans le sillage du corps, il y a la production d'émotions que M. Cardon définit comme « un certain état de réaction automatique de l'organisme devant un choc ou une situation inattendue et qui n'est pas senti explicitement par l'organisme, qui n'est pas conscient », au contraire des sensations ou sentiments²¹¹. Ces émotions doivent impérativement être modélisées²¹² pour parvenir à reproduire un acte de pensée, car la pensée repose en partie sur l'influence des émotions : « l'organisme de tous les mammifères, et notamment celui de l'homme, a une appréciation des choses du monde qui est d'abord émotionnelle »²¹³. La

²⁰⁵ *Ibid.*, p. 149.

²⁰⁶ *Ibid.*, p. 5, p. 79 et p. 146 et s.

²⁰⁷ *Ibid.*, p. 146.

²⁰⁸ *Ibid.*, p. 151.

²⁰⁹ *Ibid.*, p. 7.

²¹⁰ *Ibid.*, p. 153, la définition étant empruntée à M. Antonio Damasio.

²¹¹ *Ibid.*, p. 156. L'auteur décrit également ailleurs les émotions comme étant « l'effet visible de l'altération passagère de tout organisme vivant de par l'existence de sa corporéité sensible » ; une « réponse comportementale automatique et donc subie permettant de mettre l'organisme en situation d'adaptation à ce qui se déroule dans son environnement » (*ibid.* p. 79).

²¹² *Ibid.*, p. 2 ; sur les émotions, v. p. 80 et s. ; sur les aspects techniques de la création d'états émotionnels artificiels, p. 200 et s.

²¹³ *Ibid.*, p. 80. Le modèle proposé par A. Cardon, parce qu'il modélise les émotions à partir de stimulus externes, répond finalement à l'objection selon laquelle il serait aberrant de concevoir un cerveau artificiel comme une reproduction du cerveau humain « comme si

simulation de ces émotions reposera sur l'action d'« éléments minimaux », qui sont les agents aspectuels précédemment cités. Ces éléments minimaux ne représentent pas en eux-mêmes des émotions ou des symboles : par analogie avec les neurones, ils représentent des traits particuliers (par ex. une couleur, une forme) qui, par association, pourront s'activer, étayer, renforcer ou inhiber des caractères plus généraux au nombre desquels figurent des émotions.

120. La conscience noyau – L'étape suivant la survenance d'une émotion chez le sujet pensant sera la perception par ce dernier de l'état nouveau qu'elle vient de provoquer sur son organisme : c'est éprouver *un sentiment*, c'est-à-dire « la conscience immédiate d'une transformation qui vient d'arriver et qui se continue et peut être s'amplifie »²¹⁴. Ce nouvel état, qui consiste à éprouver sans pour autant analyser, contribue à produire ce que M. Cardon appelle, à la suite de M. Antonio Damasio, la « conscience noyau » ou « proto-soi ». Celle-ci « est la perception de l'état de soi-même en état d'appréhension immédiate des choses du monde, en état de sentir observateur pouvant devenir acteur, par son corps et dans son monde (...) la notion de conscience noyau se réfère à une situation d'immédiateté de la mise en action de l'organisme et perçue instantanément par celui-ci comme telle. C'est le fait d'être dans l'instant, simplement. (...) L'organisme se perçoit et perçoit sa mise en situation dans le monde »²¹⁵. Il ne s'agit pas d'une observation de soi, qui supposerait une mise à distance, mais de la perception que l'on est en train d'exister.

121. La conscience étendue – Enfin, l'état le plus complexe du système sera la création de la conscience étendue ou « Soi autobiographique »²¹⁶. C'est « la mise en situation de l'individu dans la temporalité »²¹⁷ : il va faire appel à des souvenirs, des faits ou des situations vécues ce qui lui permettra de désigner, juger, évaluer, prévoir. Cette conscience étendue reposera sur un système générateur de pensée distinct mais comparable à celui générateur d'émotions²¹⁸. Ce système sera fondé sur l'existence de « grains de significations », qui sont des entités minimales capables de s'activer, dotées de caractères sémantiques et regroupées dans des structures de signification²¹⁹. Le

celui-ci pouvait être considéré seul, indépendant du corps. Le cerveau fait partie intégrante du corps avec lequel il est hyper-connecté et avec son environnement au travers des sens » (J.-C. HEUDIN, « Intelligence artificielle et robot: entre utopie et dystopie », in *Les robots, objets scientifiques, objets de droits*, éd. Mare & Martin, pp. 45-55, spéc. p. 52). *Adde.* p. 53 : « En outre, le mode de prise de décision dans les deux cas est très différent : une intelligence artificielle commence par une analyse des données, systématique ou avec des heuristiques, puis en déduit logiquement une décision alors qu'un humain prend généralement une décision émotionnelle puis justifie cette décision par un raisonnement rationnel ».

²¹⁴ A. CARDON, *op. cit.* note 180, p. 82.

²¹⁵ *Ibid.*, p. 83.

²¹⁶ Sur la pensée artificielle et le soi, *ibid.* p. 227 et s.

²¹⁷ *Ibid.*, p. 84.

²¹⁸ *Ibid.*, p. 231.

²¹⁹ *Ibid.*, p. 246.

système agrègera ces structures de signification de manière cohérente dans un processus d'émergence, processus qui sera bouclé pour obtenir « un point fixe ayant une certaine durée » qui sera la représentation courante de la pensée de l'organisme artificiel²²⁰. La machine se saura en état de penser parce que ses deux système générateurs – l'un générant les émotions, l'autre les idées artificielles – lui offrent « un choix inévitable d'être dans son monde d'une façon émotionnelle et immédiate, ou bien d'être située d'une façon conceptuelle et évaluée. En levant ce choix (...) l'organisme artificiel se place en posture de se savoir pensant, et ceci à chaque instant, à tout moment de son fonctionnement. Ce choix qui est à faire et qui est fait est la façon, pour cet organisme, de se savoir être, dans le temps. Il est fait intentionnellement »²²¹. Cette intentionnalité²²² sera le fruit de l'opération d'entités appelées par l'auteur « ensemble d'anticipation », lequel appréhende les caractères des ensembles morphologiques des structures de signification et a la propriété d'altérer leur état. L'ensemble d'anticipation réalisera « une action réflexive sur les ensembles morphologiques de description »²²³.

122. La lecture de l'ouvrage de M. Cardon, dont le résumé qui précède mutile quelque peu complexité des raisonnements qui y sont déployés, est impressionnante pour le néophyte, car elle lui permet d'entrevoir sérieusement la perspective d'une machine qui serait douée d'une faculté de penser qui puisse s'approcher de celle dont nous sommes dotés. Et ce n'est que face à une machine aussi prodigieuse et complexe que celle imaginée par M. Cardon que l'on pourra se poser sérieusement la question d'attribuer ou non à celle-ci la personnalité juridique, avec comme critère le seul qui mérite véritablement de commander une telle solution.

B. Le critère ultime : la conscience du robot

123. La question fondamentale – L'étude des arguments actuellement avancés pour accorder la personnalité juridique au robot humanoïde nous a amené conclure qu'en réalité, c'est parce l'on a créé des machines qui nous ressemblent extérieurement qu'a germé l'idée de les personnifier sur le plan juridique²²⁴. Si cette considération ne saurait en aucune manière justifier l'octroi à une machine, sur le plan juridique, d'une individualité qui soit comparable à celle qui caractérisent les humains, il en va différemment si l'on

²²⁰ *Ibid.*, p. 247.

²²¹ *Ibid.* p. 266. *Adde.* p. 274 : «Le système ressent le fait qu'il produit une pensée lorsque se produit un état de stabilité de ses deux systèmes générateurs qui fixe les états de ses structures sur la représentation de quelque chose de distingué et qui est désigné : la chose précisément pensée dont il avait l'intention de la penser. Il éprouve la particularité de cet état intérieur et tente alors, par nécessité, de réduire cette tension qui inhibe l'action externe de son système émotionnel et du système permettant l'intentionnalité, pour satisfaire à une intention et produire une autre pensée ».

²²² Sur l'intentionnalité du système, v. *ibid.*, p. 280 et s.

²²³ *Ibid.*, p. 287.

²²⁴ *Supra*, n° 100.

considère la création d'une machine dont on pourrait alléguer qu'elle nous ressemble *intérieurement*, c'est-à-dire sur le plan de notre vécu personnel et de notre psychisme. Les travaux du type de ceux qu'ont mené M. Alain Cardon nous amène à devoir considérer une telle hypothèse : tôt ou tard, des ingénieurs créeront une machine qui s'efforcera non pas de faire acte « d'intelligence » artificielle, mais bien faire un acte de *pensée* entendu comme un acte *réflexif* d'une grande complexité, construit sur des bases les plus similaires possibles à celles dont nous aurons compris qu'elles gouvernent la pensée humaine, et mené par une entité dont il sera allégué qu'elle sait désormais qu'elle pense. C'est alors que l'on devra se poser la question fondamentale devant commander l'attribution ou non à cette machine d'une personnalité juridique : la machine *a-t-elle pris conscience d'elle-même* ?²²⁵

124. Si nos descendants venaient à s'accorder sur le fait que c'est effectivement le cas, *alors il serait difficile de ne pas reconnaître à cette machine la personnalité juridique*²²⁶ ; cela pour des raisons ontologiques. Ne pas le faire serait en effet maintenir une entité dont on aurait établi qu'elle est consciente, intelligente et sensible dans le giron juridique des choses et donc susceptible d'être appropriée, abîmée et détruite au mépris de sa capacité à sentir, souffrir, concevoir, communiquer et espérer. Ce serait maintenir des entités jugées conscientes, ayant des capacités intellectuelles et sensibles égales ou supérieures aux nôtres dans un état d'esclavage.

Mais pour en arriver à une telle conclusion, encore faudra-t-il, pour nos descendants, s'entendre ou avoir découvert ce qu'est réellement la conscience : une machine peut-elle réellement prendre conscience d'elle-même ou cet état est-il irrémédiablement lié à notre condition biologique ? La réponse à cette question sera fortement dépendante des avancées de la recherche dans différentes disciplines, au premier rang desquels la biologie.

125. La place de la biologie dans la conscience – Si l'on ne peut présumer des découvertes qui seront faites dans les siècles à venir sur la question de la conscience, on peut du moins tracer un cadre des futurs possibles ; l'on suggèrera même que ce cadre peut se résumer à l'alternative suivante.

126. Première hypothèse – Dans un premier futur possible, les recherches en biologie ne feront que préciser le fonctionnement du cerveau humain sur le modèle de ce que nous avons déjà compris à l'heure actuelle. En d'autres termes, nous passerons de la connaissance de l'existence d'aires

²²⁵ En ce sens, X. BIOY, art. préc. note 163, spéc. p. 92 et s. ; N. NEVEJANS, « Le robot qui voulait devenir un homme... ou le statut juridique de l'androïde » in *Le droit saisi par la science-fiction*, éd. Mare & Martin, 2016, pp. 137-164, spéc. n° 36 : « En dehors de la forme humanoïde du robot, de l'existence d'une communication et d'une interaction, ou de son autonomie, la démonstration de son altérité pourrait reposer sur sa conscience, son âme, sur une forme propre d'intériorité, ou encore sur ses sentiments ».

²²⁶ L'étendue de celle-ci restant à discuter ; sur cette question, v. not. S. SLAMA, art. préc. note 136, spéc. II. A.

spécialisées et interconnectées du cerveau à une cartographie détaillée de l'intégralité d'un cerveau humain, neurone par neurone, avec une compréhension exacte du rôle de chacun et de l'ensemble des interactions électriques et chimiques qui les lient²²⁷. Le résultat de ces recherches aura été mis à profit pour créer une machine semblable à celle décrite par M. Cardon, pouvant reproduire les mouvements qui, dans l'esprit de cet auteur, sont la source de la pensée et douée par définition d'une capacité réflexive. En l'état de la science de ce premier futur possible, face à une telle machine dont il serait allégué que ses processus fonctionnent selon des canons semblables au cerveau, et dont chacun pourrait éprouver qu'elle se comporte en tous points comme un humain intelligent, voire comme une entité plus intelligente que lui, il serait difficile à l'Humanité de renier qu'elle a bien réussi à créer une machine douée d'une conscience similaire à celle des êtres humains. Partant, sur un plan politique, il lui sera délicat – bien que toujours possible – de refuser de reconnaître à ces individualités nouvelles la consécration juridique de leur existence alors qu'elles sont douées de cette capacité qui commande cette attribution pour notre espèce. Il ne serait guère cohérent en effet d'avoir construit une machine susceptible selon l'état de la science le plus avancé de répliquer les ressorts les plus intimes du cerveau humain et de constater que cette machine se comporte effectivement comme les humains tout en lui reniant une reconnaissance et une protection²²⁸ semblable à celle dont jouissent les humains ; cela d'autant plus que ces machines seront capables de revendiquer elles-mêmes cette reconnaissance politique et ontologique. Mais il y a un autre futur possible.

127. Seconde hypothèse – Dans un second futur possible, des recherches similaires à celles entreprises dans la première hypothèse conduiront à la découverte qu'il existe, en sus de toute la « machinerie » propre au cerveau, un facteur agissant de manière décisive sur celui-ci – on pense à la prise de décision – mais dont l'essence ne pourrait être jugée comprise en lui. Pour que cette hypothèse fasse sens, il faut nécessairement postuler que les scientifiques se seront entendus sur le fait que nous vivons dans un Univers contenant plus de dimensions que celles accessibles à nos sens – une hypothèse d'ores et déjà sérieusement considérée – car cet autre facteur, à supposer son existence démontrée²²⁹, devra nécessairement agir de « quelque part ». Cette seconde

²²⁷ Il s'agirait, somme toute, d'avoir accompli toutes les étapes du projet WBE (*supra*, n° 107) ou d'un projet similaire.

²²⁸ Sur la question des droits fondamentaux qu'il conviendrait de reconnaître à des robots conscients, S. SLAMA, art. préc. note 136, spéc. II.

²²⁹ La logique voudrait que l'on commence par détecter l'existence indirecte de cet autre élément par la mise en évidence de son implication causale dans le fonctionnement du cerveau : concrètement, des manifestations identifiables de cet élément s'avèreraient corrélables avec des réactions précises mais *postérieures* du cerveau ou du corps humain par rapport à ces manifestations, qui apparaîtraient alors comme à l'*origine* de ces réactions. Un nouveau champ de la science s'ouvrirait alors afin de comprendre ce qui peut être à l'origine de ce phénomène, recherches qui impliqueront non seulement la Biologie mais aussi la Physique et, dans son sillage, la Cosmologie.

hypothèse, qui sans doute fera sourire à l'heure actuelle, est très ambitieuse en l'état de nos connaissances actuelles mais doit résolument être prise en considération²³⁰. Elle permettrait au demeurant de trancher de manière décisive une interrogation que certains scientifiques se posent déjà depuis un certain temps : dans le cerveau, où est l'observateur ? C'est la question dite de « l'aspect phénoménologique » de la conscience, c'est-à-dire de la localisation de l'aspect vécu de celle-ci. Dans la première version du futur possible exposée précédemment, cette conscience est simplement la résultante de l'ensemble des interactions internes du cerveau ; ce serait en somme le « mouvement » dont parle M. Cardon. Dans cette autre version du futur de la science ici suggérée, le cœur de ce qui fait notre expérience de la conscience – l'observateur – aura été démontré comme étant externe à notre cerveau, lequel resterait néanmoins une pièce indispensable de la réalisation de notre expérience vécue²³¹.

Il n'est pas nécessaire de pousser plus loin ce dernier effort de prospection scientifique pour en tirer l'enseignement utile à ces développements : si cette seconde branche de l'alternative des futurs possibles se réalise, alors il ne sera ni utile ni indiqué de conférer aux machines prétendument pensantes que l'on aura créées une personnalité juridique. Dans ce second futur, en effet, il sera démontré que l'expérience que nous faisons tous de la conscience dépend ultimement d'un ressort biologique et physique qu'il ne sera pas possible de reproduire car dépendant d'une propriété de la Nature dont on entr'apercevra l'existence et qui, le jour où l'on en percera le secret, ne sera sans doute pas reproductible puisque reposant sur un aspect de la structure physique de l'Univers. Les machines prétendument pensantes que nous aurons créées ne devront alors pas être mieux considérées que des pantins extrêmement élaborés, c'est-à-dire des machines capables de simuler de la manière la plus admirable toutes les réactions possibles d'un cerveau humain, mais sans pouvoir un seul instant prétendre bénéficier de la même reconnaissance ontologique, ces machines étant dépourvues de la propriété même qui conditionnerait cette reconnaissance²³².

²³⁰ Des recherches actuelles suggèrent par exemple un lien entre la conscience et la structure de l'espace-temps ; v. W. BROWN, « Unified Physics and the Entanglement Nexus of Awareness » *Journal of Neuroquantology*, Vol. 17, n° 7 (2019) ; pour une vulgarisation de cet article par son auteur et en français : <https://fr.resonancescience.org/blog/le-role-de-la-conscience-dans-l-univers>.

²³¹ Cette alternative n'est pas sans évoquer l'opposition entre J.-P Changeux et P. Ricoeur sur la nature de la conscience : alors que pour le premier, la conscience résulte uniquement des interactions que nous pouvons modéliser entre les zones du cerveau, pour P. Ricoeur, le connu et le vécu sont sur deux plans qui ne se rencontrent pas (ce dernier ne faisant aucunement l'hypothèse que nous formulons ici sur la cause possible de cette distinction). Sur cette opposition, X. BIOY, art. préc. note 163, spéc. p. 92 et 93).

²³² A. CARDON, *op. cit.* note 180, spéc. p. 419, conclut son ouvrage en exposant une alternative compatible avec celle que nous suggérons ici, mais en remontant un cran dans l'échelle causaliste pour ce qui relève de la seconde hypothèse que nous avons suggérée. La question ultime que se pose l'auteur est de savoir si l'organisation des systèmes vivants est dirigée de l'extérieur vers une efficacité qui cause leur évolution ou est le résultat de mouvements de co-évolutions conduits par le seul hasard. Alors que dans le second cas, il

s'avèrera parfaitement possible de « construire notre double sous le paradigme du calculable », dans le premier, « nous ne pourrons pas aller au-delà de l'artefact. Nous ne nous atteindrons jamais et il manquera toujours l'essentiel de l'être que nous sommes, dans les constructions que nous serons capables de faire (...). Il nous restera à nous amuser à construire des avatars, des simulacres, de produits ludiques, qui nous étourdiront et nous hallucineront peut-être, mais nous serons définitivement prisonniers de notre destin, venu d'ailleurs ».