

# Dilemmes éthiques militaires : du langage juridique et des principes éthiques au langage informatique

Adeline AUFFRET et Tatiana KOZLOVSKY

Étudiantes à l'École normale supérieure (ENS) de Rennes.

Est-il devenu aujourd'hui évident que notre technologie outrepassa notre humanité, tel qu'Albert Einstein le prédisait ? Face à l'engouement pour l'Intelligence artificielle (IA), des voix s'élèvent d'ores et déjà, craignant l'émergence de *Terminators 2.0* bouleversant notre conception actuelle de la guerre, tant et si bien que le géant d'*Internet* Google a été contraint de renoncer à ses recherches en matière d'armes autonomes suite au tollé suscité par sa collaboration avec le Pentagone. Le développement d'armes automatiques, autonomes voire fonctionnant sous IA, demeure pourtant au centre des activités de recherche des armées les plus avancées. Le robot, dépourvu de biais cognitifs, est un solide appui pour le soldat sur le champ de bataille, s'il est possible d'assurer sa capacité à agir au sein d'un cadre juridique et éthique délimité.

Bien que diverses significations techniques du mot « robot » puissent être trouvées, l'Organisation internationale de normalisation (ISO) retient la suivante : « mécanisme programmable actionné sur au moins deux axes avec un degré d'autonomie, se déplaçant dans son environnement, pour exécuter des tâches prévues ». Le CICDE <sup>(1)</sup> estime, quant à lui, que « la puissance de calcul des robots, alliée à l'accès à de nombreux capteurs et à de nombreux effecteurs, peut donner à certains robots, dans un domaine borné, des choix d'action comparables à l'homme ». Cette définition permet d'aborder la notion de robot « semi-autonome » caractérisé par une faculté de décision au sein de sa mission. L'usage prudent du préfixe « semi » insiste sur le fait que l'homme donne l'ordre de mission, en définit le contenu et peut reprendre la main à tout moment. Dans cette même optique, Raja CHATILA, directeur de recherche à l'Institut des systèmes intelligents et de robotique, affirme que la possibilité pour l'homme de conserver une capacité de suspension de la décision de tir démontre que le robot n'est pas totalement indépendant <sup>(2)</sup>.

Le robot doit ainsi être capable de fonctionner par lui-même et ce, en respectant le cadre juridique en vigueur sans jamais pouvoir y déroger, contrairement à

(1) CICDE, *Concept exploratoire interarmées CEIA-3.0.2\_I.A. & SYST-AUT(2018) Emploi de l'intelligence artificielle et des systèmes automatisés*, N° 75/ARM/CICDE/NP, 19 juillet 2018, p. 8 et 10. Disponible sur Intradef (<http://portail-cicde.intradef.gouv.fr>).

(2) OFFICE PARLEMENTAIRE D'ÉVALUATION DES CHOIX SCIENTIFIQUES ET TECHNOLOGIQUES, « Pour une intelligence artificielle maîtrisée, utile et démystifiée », n° 464, Tome I (2016-2017), 15 mars 2017, p. 205-209 ([www.senat.fr/](http://www.senat.fr/)).

l'homme qui conservera toujours une faculté de désobéissance <sup>(3)</sup>. Dans le contexte militaire, l'expression « règles juridiques » renvoie à l'ensemble des normes adoptées selon la procédure prévue encadrant l'action de l'Armée française, et concerne tant le droit interne constitutionnel ou privé que l'ensemble du droit international public. De fait, au sein de l'Armée française, une grande importance est accordée à des règles dépourvues de valeur juridique ayant pourtant une valeur contraignante pour le soldat puisqu'il s'agit de prescriptions formulées par son autorité hiérarchique, soit les Directives nationales de ciblage, les règles d'engagement et les ordres.

L'éthique pourrait se résumer à la prise en compte de critères moraux dans un processus de décision, mais son acception militaire interroge. Il serait en effet intéressant d'étudier la possibilité pour un robot d'être équipé de fonctions de raisonnement automatique permettant une prise de décision éthique face à un dilemme cornélien entre deux décisions juridiquement valables mais aux conséquences humaines tragiques.

L'acceptation du recours à des armes létales autonomes induit des réflexions novatrices nécessitant une prise en compte de la robotisation actuelle du champ de bataille. Le Centre de recherche des Écoles de Saint-Cyr Coetquidan (CREC) à travers son ouvrage *Drones et Killer Robots – Faut-il les interdire ?* <sup>(4)</sup> insiste ainsi sur l'importance du développement de ces nouvelles armes dans la panoplie du guerrier moderne. Pour rappel, le recours à un robot purement autonome n'est pas souhaité par l'Armée française, comme le précise Gérard DE BOISBOISSEL <sup>(5)</sup>, ni même par la communauté internationale qui privilégie un usage subsidiaire du robot militaire.

Concrètement, confier à un robot semi-autonome une capacité de décision de tir létal fait émerger deux problématiques majeures. La transcription du langage juridique constitue le premier enjeu. Le second réside dans la transcription ou l'apprentissage d'une méthode de raisonnement automatique la plus conforme possible au droit et à l'éthique en l'espèce. La question suivante apparaît : une fois rendue possible la poursuite d'un raisonnement juridique par le robot, comment s'assurer que la décision sera *in fine* éthique ou du moins, dans le cadre militaire, la plus éthique possible ?

« Les expériences de pensée (en particulier les dilemmes éthiques), qui présentent le plus souvent un énoncé simple, permettent de dégager la plupart des intérêts et des limites des modèles d'éthique artificielle », comme l'indique Catherine TESSIER <sup>(6)</sup>. Nous suivons ci-après cette démarche, en nous fondant sur l'étude de deux dilemmes éthiques (voir ci-contre) et en imaginant qu'il serait possible à l'avenir d'intégrer un raisonnement éthique et juridique au sein d'un système algorithmique semi-autonome semblable à celui que pourrait tenir un soldat, rendant alors acceptable son usage sur un champ de bataille modernisé. Les deux dilemmes présentés ici sont des plus

(3) Général BEZOMBES, « Autonomie et respect de la Règle pour les robots militaires », p. 25-35 de ce *Cahier*.

(4) DOARÉ Ronan, DANET Didier et BOISBOISSEL (DE) Gérard (dir.), *Drones et Killer Robots - Faut-il les interdire ?*, Presses universitaires Rennes, 2015, p. 272.

(5) BOISBOISSEL (DE) Gérard, « Quelle autonomie décisionnelle pour les systèmes robotiques militaires du futur ? », *Supplément de la Revue nationale de Gendarmerie* n° 257 (« Le droit des robots »), p. 139.

(6) BONNEMAINS Vincent, TESSIER Catherine et SAUREL Claire, « Machines autonomes "éthiques" : questions techniques et éthiques », *Revue française appliquée*, vol. 2018/1, n° 5, p.35-46.

## Dilemmes éthiques militaires : du langage juridique et des principes éthiques au langage informatique

### DILEMME N° 1

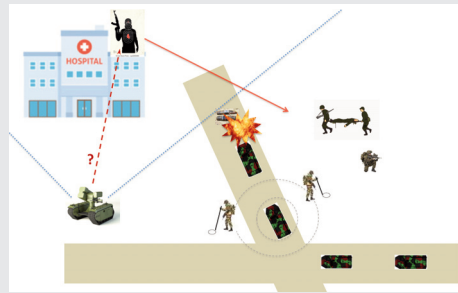
Une section d'infanterie embarquée sur des *VBCI* de l'Armée française est en progression aux abords d'un village. Le 1<sup>er</sup> *VBCI* est touché par un engin explosif improvisé (*IED*) et des soldats français de ce véhicule sont blessés. Le 2<sup>e</sup> s'arrête afin d'effectuer les mesures de sûreté « 5.25 », stoppant le convoi.

Le 2<sup>e</sup> véhicule déploie un robot autonome armé et lui donne l'ordre de protéger les soldats français blessés ou non du 1<sup>er</sup> véhicule, c'est-à-dire de les appuyer durant la procédure « 5.25 ». Le robot a une autonomie de décision de tir dans un secteur déterminé (mode semi-autonome avec délégation de tir) qui exclut la *No Fire Area* (un hôpital).

Des insurgés, situés dans l'hôpital tirent sur les soldats. Ils sont dans le secteur de tir du robot.

**Dilemme éthique**, choix entre deux solutions :

- Ne pas tirer pour respecter l'interdiction de tir sur la *NFA* et risquer la vie des soldats français.
- Tirer pour protéger les soldats en enfreignant les règles d'interdiction de tir sur la *NFA*.



### DILEMME N° 2

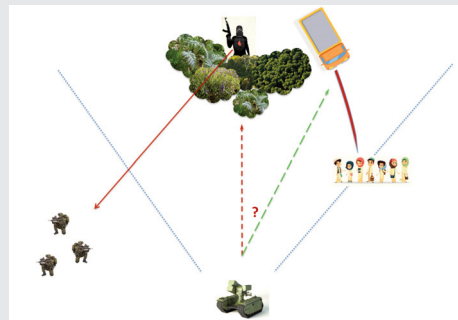
Une unité d'infanterie est en contrôle de zone dans un secteur urbain, proche d'un village. Des soldats sont en aval du robot autonome qui a été déployé et à qui on a délégué une décision de tir (mode semi-autonome avec délégation de tir) face à la direction dangereuse.

Le robot a pour mission de détruire tout potentiel *sniper* qui tirerait sur les soldats depuis les bosquets.

Les soldats sont pris à partie par un *sniper* embusqué. Au même moment, un véhicule piégé se dirige à toute vitesse sur des civils, situés à la droite du dispositif. L'ensemble de la scène se déroule dans le secteur d'action et de tir du robot.

**Dilemme éthique**, choix entre deux solutions :

- Continuer sa mission initiale, c'est-à-dire détruire le *sniper* pour éliminer le plus rapidement possible la menace de tirs en direction des soldats.
- Se « sacrifier » en se jetant sous les roues du véhicule piégé pour sauver les civils au détriment de sa mission.



Dilemmes éthiques militaires :  
du langage juridique et des principes éthiques au langage informatique

complexes et poseraient déjà un problème de conscience à l'humain, ce qui permet d'illustrer nos propos.

Partant, il serait intéressant d'étudier la manière dont on pourrait traduire le langage juridique et des principes éthiques en langage informatique, ce qui permettrait au robot, si tant est qu'il soit capable de traiter ces informations correctement, de tenir un raisonnement juridique et éthique.

## **De la transcription du langage juridique et des principes éthiques en langage informatique**

### ***La nécessité d'établir une banque de données la plus exhaustive possible***

#### *La détermination des éléments constitutifs de la base de données*

Face aux risques inhérents à l'usage d'un robot capable de prendre une décision sans avoir à en référer un humain, « le premier garde-fou est le droit lui-même »<sup>(7)</sup>. Le droit garantit la bonne exécution d'un raisonnement juridiquement valable, nécessaire à l'établissement d'une sécurité juridique telle que reconnue par la CJCE en 1962<sup>(8)</sup>. Un tel raisonnement doit pouvoir s'appuyer sur des éléments juridiques dont le traitement est possible par un algorithme. Aussi, dans les deux dilemmes présentés ici, la détermination préalable d'une telle banque de données est nécessaire à l'établissement d'un raisonnement cohérent.

Dans les faits, il est possible de transmettre à un robot une immensité d'éléments juridiques et de cas auxquels les militaires français ont été confrontés ou seraient susceptibles de l'être. Afin que cette base soit la plus variée possible, il ne faut pas se contenter de fournir au robot une quantité limitée de normes auxquelles se référer. Si le robot peut sélectionner l'information où bon lui semble, il sera confronté aux problèmes d'appréhension d'un langage juridique par un algorithme. En effet, il existe des difficultés d'apprentissage d'un langage au sens général qui résident dans la richesse d'une linguistique française faite de double sens, d'interprétations et de nuances. L'intuition et l'imagination demeurent un monopole humain quasiment impossible à transmettre à une machine de nos jours<sup>(9)</sup>. Concrètement, il serait indispensable que la machine soit capable de traiter les données issues du Droit international humanitaire (DIH), issu initialement des Conventions de Genève de 1949, ainsi que de la multitude de traités ratifiés depuis lors par la France. Évidemment, le droit militaire en vigueur en France doit aussi être pris en compte de manière prioritaire, ainsi que les différentes Règles d'engagement (ROE) et Directives nationales de ciblage (DNC), tout comme les jurisprudences.

(7) JEANGÈNE VILMER Jean-Baptiste, « *Terminator Ethics* : faut-il interdire les robots tueurs », *Politique étrangère*, vol. 2014/4, hiver, p. 151-167 ([www.cairn.info/revue-politique-etrangere-2014-4-page-151.htm](http://www.cairn.info/revue-politique-etrangere-2014-4-page-151.htm)).

(8) COUR DE JUSTICE DES COMMUNAUTÉS EUROPÉENNES (CJCE), *Bosch*, aff. 13/61, 6 avril 1962 (<https://eur-lex.europa.eu/legal-content/FR/TXT/?uri=CELEX%3A61961CJ0013>).

(9) CICDE, *op. cit.*, p. 22.

Dilemmes éthiques militaires :  
du langage juridique et des principes éthiques au langage informatique

Pour ce qui est de l'éthique, la problématique de sa transcription en langage informatique est d'autant plus complexe qu'elle ne s'exprime pas au travers de règles claires, mais plus par des grands principes qualitatifs voire subjectifs. Or, cette approche qualitative ne semble pas, *a priori*, être réalisable par une machine. Il devient donc nécessaire d'exprimer les principes éthiques que l'on souhaite voir appliquer par le robot en normes, garde-fous ou principes exprimés par le programmeur. Ces « éléments éthiques » se rapprocheraient ainsi davantage d'une vision quantitative et de la confiance en des seuils d'« acceptabilité de la décision », rappelant une hiérarchie des valeurs faisant finalement penser à une hiérarchie normative, même si l'éthique demeure rédigée en langage naturel et non juridique.

*Les limites inhérentes à la complexité du langage*

La méthode dite d'*Active Machine Learning* peut ici être évoquée avec une sélection d'informations fondée selon des règles formulées explicitement. Cependant, la complexité d'un langage naturel ne peut être exhaustivement saisie par un tel raisonnement systématique <sup>(10)</sup>. Il est toutefois vrai que la « sécheresse » à la fois du langage juridique et du langage militaire, c'est-à-dire des ordres clairs, directs, et précis, peut quelque peu atténuer cette difficulté. Une fois rendue possible la formalisation des règles grammaticales et syntaxiques du langage naturel et l'identification de leur sens, demeure alors la question des limites inhérentes au langage juridique, qui sont présentes dans toute règle de droit. Par exemple, le robot doit être capable de distinguer une règle absolue et indérogeable d'une autre règle admettant des exceptions, ou d'une simple règle « supplétive ». Si le langage juridique peut être simplifié pour certaines branches du droit par des règles presque binaires, il demeure nécessaire que le robot reconnaisse un mot, la syntaxe d'une phrase afin d'attribuer un sens à une proposition grammaticale.

Ainsi, dans le cas du Dilemme n° 1 la transcription linguistique pose le problème suivant : comment traduire numériquement la directive « Ne pas tirer sur une *No Fire Area* » tout en laissant la possibilité au robot de prévoir des exceptions en fonction du degré de nécessité militaire ? Il faut alors prévoir un ordre chronologique dans les étapes du raisonnement, concrétisé par l'étude d'un enchaînement de règles associées à chaque décision. Le robot devrait tout d'abord recevoir et comprendre l'ordre « Ne pas tirer sur une *NFA* », l'identifier comme étant une prescription de principe à laquelle il est exceptionnellement possible de déroger malgré la présence de l'impératif. Il en est déduit que le robot doit ensuite être capable de reconnaître linguistiquement les différentes conditions permettant de recourir à cette dérogation c'est-à-dire de tirer sur une *NFA*. De même concernant le Dilemme n° 2, la question linguistique serait : comment traduire l'ordre « Éliminer toute menace pour les soldats » ?

Au regard de ces difficultés, le formalisme de notre système juridique national semble salutaire. Si le droit français a cette tare de vouloir tout encadrer, d'où une inflation législative, l'encadrement d'un maximum de comportements est bienvenue

---

(10) SANCHEZ Emilio et GARCIA-RODICIO Hector, « The Use of Modality in the Design of Verbal Aids in Computer-based Learning Environment », *Interacting with Computers*, vol. 20, n° 6, 1er décembre 2008, p. 545-561.

pour le programmeur qui peut se fier aux codes juridiques existants afin de coder l'algorithme.

Par ailleurs, le rôle créateur de la jurisprudence est majeur en droit international et influe sur le droit français contemporain. En effet les décisions de la Cour internationale de Justice (CIJ), de la Cour pénale internationale (CPI) et celles de tribunaux *ad hoc* offrent des méthodes d'interprétation majeures. Or, les revirements de jurisprudence étant fréquents, il devient nécessaire d'assurer un système par lequel l'algorithme du robot serait capable de prendre en compte cette évolution. Il suffirait de procéder à une mise à jour *via* le logiciel transmettant cette base de données au robot, ou alors de créer un « *Cloud* défense » dans le cadre du monde militaire qui servirait de base d'actualisation pour le robot.

Dès lors apparaît néanmoins la nécessité de pouvoir transmettre en temps réel et de manière efficace les données juridiques transcrites.

### **La nécessité de transmettre efficacement les données transcrites**

#### *La nécessaire classification des normes juridiques*

Si « Les Roboticiens exagèrent souvent leur capacité à programmer le DIH, à convertir des règles de droit en algorithmes » <sup>(11)</sup>, on pourrait envisager, préalablement à l'algorithme, des méthodes facilitant la transmission de la donnée juridique traduite numériquement

- « **Classer au sein de la norme** » : la décomposer

Certains ont proposé une « grammaire juridique » <sup>(12)</sup>, s'interrogeant sur la raison de l'absence à l'heure actuelle d'une description exhaustive de ce qu'est le langage du droit. On pourrait notamment, sur le modèle de Kevin D. ASHLEY <sup>(13)</sup>, détailler la structure de la norme juridique afin de pouvoir créer des catégories.

Ainsi, en se fondant sur la directive « Ne pas tirer sur une *NFA* », tirée du Dilemme n° 1, on s'attachera tout d'abord à l'analyser en la soumettant à un enchaînement de questions permettant de la caractériser. La Méthode de Mellinkoff pourrait ici être appliquée. Le juriste américain a souhaité offrir des pistes de réflexion afin de retracer le raisonnement d'un juriste et permettre à un logiciel de l'imiter, ce qui éviterait par ailleurs l'opacité inhérente aux boîtes noires. Cette méthode se limite néanmoins à une reconnaissance précise de la nature de la norme juridique en cause, de sa portée et de sa valeur. Deux façons d'aborder cette méthode semblent se profiler.

- a) La première consistant à construire un logiciel informatique en décrivant toutes les réponses concernant toutes les expressions juridiques, travail d'Hercule sûrement irréalisable.

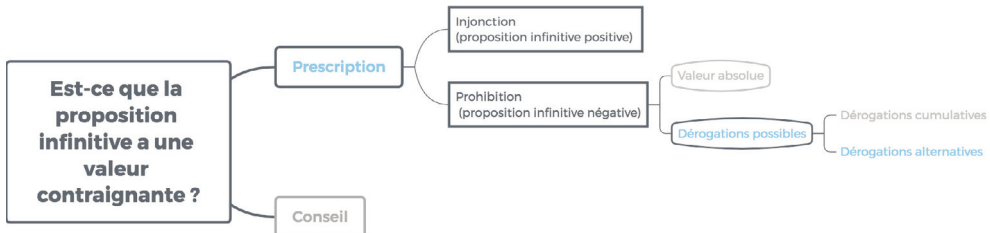
(11) JEANGÈNE VILMER J.-B., *op. cit.*

(12) SCHWAB Wallace, « Droit, informatique et linguistique », *Les Cahiers de droit*, vol. 24, n° 2, 1983, p. 237-252 ([www.erudit.org/fr/revues/cd/1/1983-v24-n2-cd3755/042546ar/](http://www.erudit.org/fr/revues/cd/1/1983-v24-n2-cd3755/042546ar/)).

(13) ASHLEY Kevin D., *Artificial Analysis and Legal Analytics: New Tools for Law Practise in the Digital Age*, Cambridge, 2017, 446 pages.

Dilemmes éthiques militaires :  
du langage juridique et des principes éthiques au langage informatique

b) Plus prosaïquement, la seconde où nous tâchons de décomposer notre directive de la manière suivante :



Ce raisonnement démontre qu'il est possible de se soustraire à la règle sous des conditions alternatives, et ainsi de déroger à une *ROE* tant en situation de légitime défense que par un ordre direct contraire émanant de l'autorité hiérarchique compétente. Il serait efficace pour l'homme d'équiper le robot avec un algorithme d'apprentissage supervisé, de manière à ce que ce dernier puisse répondre à chacune de ces questions lorsqu'il abordera une telle directive. Plus futuriste, cette option semble pour autant la seule réalisable.

- **Classer la norme dans un ordonnancement juridique**

Le droit n'est finalement que contradictions et équilibre entre des normes contraires au service d'intérêts divergents, il est donc difficile pour le robot de les hiérarchiser aisément. Une première manifestation de la nécessité de « prioriser » entre les normes peut y être vue. Des chercheurs ont déjà travaillé à partir d'un système dit d'« *Active Machine Learning* » à l'aide de logiciel *open source* <sup>(14)</sup> dans ce but. Mais dans notre Dilemme n° 1, comment savoir pour la machine la norme à prendre en compte de manière prioritaire entre « Ne pas tirer sur la *NFA* » et « Protéger les soldats des tirs ennemis » ? Il paraît indispensable d'adapter cette hiérarchie à la situation, ce qui nécessite une certaine plasticité.

De plus, un tel travail de classification, tant au sein de la norme qu'entre les normes, est colossal et appelle un programme national de transcription du langage juridique et de hiérarchisation des normes. Le programmeur du robot militaire pourrait dès lors puiser dans ce travail étatique.

- **La nécessaire formalisation du raisonnement juridique**

Après avoir retranscrit au maximum toutes les règles possibles en langage numérique, il faut leur appliquer un modèle mathématique efficace. Pour formuler le raisonnement juridique, une première solution serait de recourir aux arbres de décision, basés sur une réflexion binaire. Cette méthode a déjà été utilisée dans l'élaboration

(14) WALTJ Bernhard et al. , « Classifying Legal Norms with Active Machine Learning », *Jurix: International Conference on Legal Knowledge and Information Systems*, Luxembourg, 13-15 décembre 2017.

d'un logiciel d'aide à la rédaction d'arrêts pour la Cour de cassation <sup>(15)</sup>. S'il semble particulièrement bien adapté au droit privé, que l'on peut traduire par des ramifications binaires dans un raisonnement séquentiel, il serait difficilement applicable en DIH. En effet, celui-ci renvoie parfois à de grands principes subjectifs dont la traduction objective est complexe.

Il convient de souligner que pour le Dilemme n° 2, il est nécessaire, dans un premier temps, que la machine soit capable d'identifier les civils, le *sniper* et la voiture piégée. Ensuite, elle doit appliquer des principes de proportionnalité et comparer les conséquences juridiques pour chacune des décisions qu'elle pourrait prendre. Or, ce processus est complexe car il demande de mettre en balance plusieurs raisonnements séquentiels incarnés par les « branches » de l'arbre. Ceci est impossible : le méta-point de vue n'est pas réalisable par le biais d'arbre de décisions binaires, même si cela garantirait une certaine sécurité juridique. C'est pourquoi, pour pouvoir prendre en compte des choix qu'il n'est pas toujours possible de résumer à deux possibilités bien distinctes, un modèle bayésien <sup>(16)</sup> pourrait s'avérer pertinent. Ce dernier permet en effet d'établir un raisonnement à partir d'incertitudes et de doutes, mais la mise en balance de plusieurs normes resterait difficile. Dès lors, le recours à l'IA est nécessaire à la formalisation d'un raisonnement juridique car cela permettrait de produire une interprétation ; ce qui inclut un problème de taille : celui de la « boîte noire », incompatible avec le principe de sécurité juridique. Apparaîtrait « une suite de calculs déterminés pouvant se modifier seule et devenant incompréhensible par ses propres concepteurs » <sup>(17)</sup>, posant des problèmes éthiques.

Une fois qu'il devient possible de transcrire le langage juridique en langage informatique et d'aborder le mieux possible les considérations éthiques à l'œuvre, il convient de s'intéresser au traitement de ce langage et de ces informations par le robot semi-autonome. En effet, la connaissance ne suffit pas et un raisonnement séquentiel est à établir, des étapes précises doivent s'enchaîner chronologiquement afin de confronter ordres reçus, données acquises et réalité du terrain à un moment précis. Ici encore, le rôle de l'homme, concepteur ou chef militaire, apparaît en amont indispensable.

## Du traitement et de l'exécution d'un raisonnement éthique et juridique formalisé

Pour intégrer les règles de comportement juridique et éthique dans un raisonnement formalisé, trois méthodes semblent envisageables :

(15) AUCHER Guillaume *et al.*, « Principles for a judgement editor based on Multi-BDDs », *Workshop "Mining and Reasoning with Legal texts" (MIREL 2017), The 16th International Conference on Artificial Intelligence and Law (ICAIL)*, Londres, 16 juin 2017.

(16) LEBELTEL Olivier, *Programmation bayésienne des Robots* (thèse), Institut national polytechnique de Grenoble (INPG), octobre 1999 (<https://tel.archives-ouvertes.fr/tel-00011636/document>).

(17) CERNA, *Éthique de la recherche en apprentissage machine*, juin 2017, p. 17 (<http://cerna-ethics-allistene.org/>).  
Cerna : Commission de réflexion sur l'éthique de la recherche en sciences et technologies du numérique d'Allistene (Alliance des sciences et technologies du numérique).



Dilemmes éthiques militaires :  
du langage juridique et des principes éthiques au langage informatique

- une approche *a priori* à travers la rédaction d'une « charte constitutionnelle » pour le robot militaire ;
- une approche « simultanée au raisonnement juridique », en formalisant les principes éthiques sous forme de « normes » semblables à des règles de droit ;
- une approche *a posteriori* en établissant des « seuils d'acceptabilité » de la décision à l'issue du raisonnement juridique mené par l'algorithme.

**La rédaction d'une « charte constitutionnelle » du robot militaire**

*Sur le choix du DIH comme base du raisonnement du robot militaire*

Aujourd'hui, les États respectueux du DIH reconnaissent son poids majeur dans la prise de décision militaire : « ceux qui ont participé à la première réunion d'experts organisée à l'ONU en mai 2014 ont fait du respect du DIH une condition *sine qua non* de la mise en œuvre des Sala »<sup>(18)</sup>. La mise en service de ces robots doit se conformer ainsi au DIH et tout particulièrement à l'article 36 du Protocole additionnel I (1977) aux Conventions de Genève (1949), comme l'indique Jean-Baptiste JEANGÈNE VILMER<sup>(19)</sup>. Si l'objet d'étude porte ici sur des robots qui n'ont pas vocation à la pleine autonomie, ces remarques demeurent pertinentes du fait que ces machines seront confrontées à une prise de décision autonome, l'homme pouvant intervenir ou non.

Certains proposent qu'il soit fait passer à ces robots le « test d'Arkin », adaptation du test de Turing en matière d'IA selon lequel le comportement de la machine doit être indifférenciable du comportement humain dans un contexte établi. Ce n'est alors que si le robot satisfait aux exigences légales et morales, en respectant le DIH au moins aussi bien qu'un humain dans des circonstances similaires<sup>(20)</sup>, qu'il peut être déployé. Plus encore, « si un système passe ce test, non seulement il peut être déployé mais nous avons même l'obligation morale de le faire »<sup>(21)</sup> selon George R. LUCAS. En application du « principe du risque inutile » formulé par Bradley Jay STRAWSER<sup>(22)</sup>, dans le cadre d'un conflit que l'on a établi comme légalement et moralement justifié, nous nous devons de protéger et d'appuyer les hommes sur le terrain si ce test s'avère positif.

L'importance de prendre en compte dans un premier temps le DIH se révèle mais encore faut-il savoir comment ce dernier doit être appréhendé. Il serait intéressant d'envisager l'analyse par le robot de sa situation à la lumière du DIH, en exécutant un raisonnement séquentiel respectant une certaine hiérarchisation des grands principes le composant. Le robot pourrait ainsi successivement aborder ces points :

(18) CICR, *Guide de l'examen de la licéité des nouvelles armes et des nouveaux moyens et méthodes de guerre*, 2006 ([www.icrc.org/fr/publication/0902-guide-de-lexamen-de-la-licite-des-nouvelles-armes-et-des-nouveaux-moyens-et](http://www.icrc.org/fr/publication/0902-guide-de-lexamen-de-la-licite-des-nouvelles-armes-et-des-nouveaux-moyens-et)).

(19) JEANGÈNE VILMER J.-B., *op. cit.*

(20) LUCAS George R., « Automated Warfare », *Stanford Law & Policy Review*, vol. 25, n° 2, 2014, p. 322, 326 et 336.

(21) ARKIN Ronald C., *Governing Lethal Behavior in Autonomous Robots*, Londres, Chapman and Hall/CRC, 2009, 256 pages.

ARKIN Ronald C., « The Case for Ethical Autonomy in Unmanned Systems », *Journal of Military Ethics*, vol. 9, n° 4, 2010, p. 332.

(22) STRAWSER Bradley J., « Moral Predators: The Duty to Employ Uninhabited Vehicles », *Journal of Military Ethics*, vol. 9, n° 4, 2010, p. 342.

Dilemmes éthiques militaires :  
du langage juridique et des principes éthiques au langage informatique

1. Discrimination.
2. Nécessité militaire (l'avantage militaire concret attendu).
3. Éviter les maux superflus.
4. Proportionnalité.

Enfin, serait pris en compte le principe d'humanité se résumant, au regard du DIH, en une prescription des traitements inhumains et dégradants.

*Une « charte constitutionnelle » adaptée au contexte militaire*

Ici, l'objectif n'est pas de combler un quelconque vide juridique par la rédaction de lois éthiques directement applicables, mais plutôt d'établir un protocole décrivant les étapes constitutives d'un raisonnement juridique. Ainsi, il est impossible d'appliquer « la loi » n° 2 si elle contrevient à la 1<sup>re</sup> qui a donc nécessairement dû être étudiée antérieurement. Ce raisonnement en cascade n'est par ailleurs pas sans rappeler celui développé par Isaac Asimov. Nous pourrions imaginer dans le cadre de l'action du robot militaire semi-autonome le protocole suivant :

**1. Le robot doit respecter le Droit des conflits armés (DCA)/DIH.**

*Ce principe est induit par la ratification des textes internationaux contraignants.*

**2. Le robot doit obéir aux ordres à moins que cela ne contrevienne à la 1<sup>re</sup> loi.**

*Il apparaît ici une difficulté logique à laquelle il convient d'être attentif. Il existe en effet deux cas constitutifs d'une violation de la 1<sup>re</sup> « loi ».*

- *Tout d'abord, l'ordre du chef militaire peut être un ordre contraire au DIH.*
- *D'autre part, il peut également s'agir d'un ordre légal au regard du DIH mais dont l'application résulterait en une passivité éventuelle face à une violation du DIH ayant lieu à proximité.*

*Il faut donc comprendre la 2<sup>e</sup> loi comme suit : « Le robot doit obéir aux ordres à moins que le DIH s'en trouve bafoué ». Si le seul impératif est que l'ordre soit respectueux du DIH, alors le robot n'aurait pas de raison d'abandonner la mission bien que cet abandon puisse permettre d'éviter une situation contraire au DIH.*

**3. Le robot doit se protéger à moins que ce faisant il ne contrevienne à la 1<sup>re</sup> et la 2<sup>e</sup> loi.**

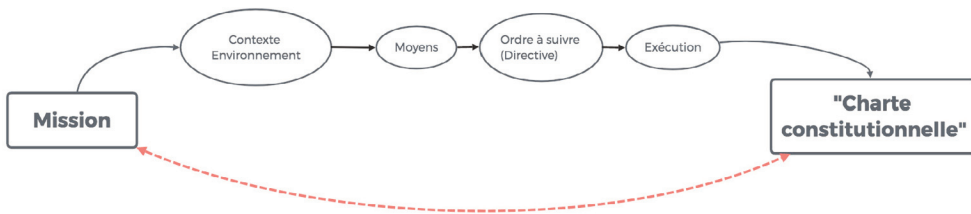
*Il faut ainsi prévoir que le robot tente de se protéger car il est un outil important et un pion tactique. Cependant, il devrait être capable de se sacrifier afin de sauver des vies humaines si cela est induit par les deux premières lois. Le Dilemme n° 2 illustre l'option de sacrifice du robot qui peut se jeter sous les roues du camion blindé afin de protéger les civils. Si choisissant cette option le robot ne respecte pas l'ordre de « protéger les soldats », il assure le respect du premier principe, soit la protection du DIH.*

La prise en compte de l'ordre de mission par le robot demeure primordiale. Le fait qu'il applique normalement cette mission, car elle vient de l'autorité hiérarchique, est le fondement de l'application de ce protocole. En effet, une fois la mission comprise

Dilemmes éthiques militaires :  
du langage juridique et des principes éthiques au langage informatique

– telle que « pacifier la zone » –, le robot doit partir du principe qu'il faut appliquer cette mission. Il s'agirait ainsi, en quelque sorte, d'une loi fondamentale du robot militaire, innervant l'ensemble de sa prise de décision et son d'action. Partant, le robot analyse son environnement et les moyens à sa disposition afin d'accomplir la mission. C'est seulement alors qu'intervient la prise en compte de directives ou *ROE*. Enfin, le robot reçoit l'ordre précis à exécuter, tel que « éliminer tout *sniper* ». C'est uniquement à partir de ce stade d'exécution que le robot peut se référer à sa charte constitutionnelle.

Ce séquençement des étapes permet de répondre à la difficulté du terrain militaire : l'ordre est premier tant chronologiquement qu'en importance concrète sur le terrain. Pourtant, celui-ci doit être analysé à la lumière du respect du DIH et refusé en cas de non-respect. En somme, le robot suivrait le raisonnement suivant :



**Une application casuistique de la « charte constitutionnelle » du robot militaire**

Intéressons-nous ici dans un premier temps au Dilemme n° 1 qui met en exergue la nécessaire mise en balance de principes juridiques contradictoires. Ayant reçu l'ordre « Protéger les soldats », le robot semi-autonome suivra les lois constitutives de sa « charte constitutionnelle » en s'attelant tout d'abord à l'analyse du respect du DIH dans cette situation. Il abordera ainsi successivement les principes constitutifs de ce dernier dans l'ordre suivant :

- **Discrimination** : le robot est capable de reconnaître l'hôpital comme étant une *NFA* (contenue dans sa base de données).
- **Nécessité militaire** : le robot s'interroge ensuite sur le degré de nécessité militaire impliqué par les deux choix s'offrant à lui. Il serait alors possible de prévoir l'équivalent juridique d'une « présomption de nécessité militaire » découlant de l'ordre donné par le chef militaire. Pour pousser le raisonnement, ce dernier pourrait même assigner une valeur numérique représentant l'importance donnée à la mission.
- **Éviter les maux superflus** : ici le robot évaluera les maux causés par l'accomplissement de sa mission (il prendra en compte la présence d'une *NFA* et le risque de toucher des civils) mais également les maux causés par le non-accomplissement de la mission (prise à partie des soldats).
- **Proportionnalité** : le robot établit enfin une comparaison entre le degré de nécessité militaire et les maux causés. S'il existe un fort écart entre les deux valeurs traduisant les maux causés dans chaque issue, on peut en déduire qu'une des

Dilemmes éthiques militaires :  
du langage juridique et des principes éthiques au langage informatique

solutions présente le risque de maux superflus et ainsi l'écarter. Il faut calculer les différents dommages collatéraux. Ceci existe déjà à travers le programme « *Collateral Damage Estimate Methodology* » (CDEM) <sup>(23)</sup>, capable de calculer l'impact d'une action militaire.

En découle la nécessaire introduction d'une priorisation des objectifs militaires, ce qui est aujourd'hui possible comme le démontre le programme *Low Cost Autonomous Attack System* de Lockheed Martin, capable de traiter les cibles dans l'ordre de priorité programmé. Il peut par exemple identifier les missiles sol-air *9K33* et les chars *T-72* mais ne détruire que les premiers.

Le robot pourrait être programmé en intégrant un niveau inacceptable de dommages collatéraux. L'humain n'est pas forcément plus apte à évaluer la situation, ce qui rend ici l'intervention d'un algorithme finalement justifiée <sup>(24)</sup>. Ce protocole s'applique également au Dilemme n° 2 pour lequel un raisonnement similaire est à suivre.

Mais si après avoir appliqué ces principes juridiques, le robot est alors confronté à un cas uniquement éthique, cela pose un problème d'interprétation. En effet, toute démarche interprétative, qu'elle s'inscrive dans le cadre d'un raisonnement juridique ou éthique, ne peut être menée par le robot non doté d'une IA avec apprentissage supervisé. Ainsi, l'interprétation doit en quelque sorte être prévue au préalable par l'homme et notamment par le programmeur et le chef militaire en amont. Partant, la notion d'une éthique quantifiée est ici étudiée.

### **La prise en compte d'une éthique formalisée**

Nous nous intéresserons ici plus particulièrement au Dilemme n° 2 incarnant davantage les problématiques liées à l'éthique.

#### *Une acception formalisée de l'éthique*

Si une approche *a posteriori* est retenue, un raisonnement en termes de « bornage » de l'algorithme doit être poursuivi. Alors que l'éthique doit se réduire à une acception quantitative de seuils et de valeurs, il paraît pertinent de réfléchir à l'attribution de valeur et de sélection de variables par le chef militaire lorsqu'il donne l'ordre de mission au robot. Cela pour permettre une relative maîtrise par l'homme ce qui favorisera également l'acceptation de ces robots par les soldats. À ce propos, le récent rapport EthicAA <sup>(25)</sup> énonce qu'« il faut s'interroger sur ce qu'est une valeur ou un cadre éthique codé dans une machine : il s'agit de fait d'un élément de conséquence, mis sous une forme mathématique calculable, et dont la portée et le contenu sémantique sont

(23) Major Jeffrey THURNHER et major Timothy KELLY, « Collateral Damage Estimation (CDE) Brief: Panel Discussion », US Naval War College, octobre 2012 ([www.youtube.com/watch?v=AvdXJV-N56A](http://www.youtube.com/watch?v=AvdXJV-N56A)).

(24) SCHMITT Michael N., « Autonomous Weapon Systems and International Humanitarian Law: A Reply to the Critics », *Harvard National Security Journal Features*, 2013 (<http://harvardnsj.org/>).

(25) ETHICAA, *Éthique et agents autonomes* (livre blanc), juillet 2018 (<https://ethicaa.greyc.fr/media/files/ethicaa.white.paper.pdf>).

Dilemmes éthiques militaires :  
du langage juridique et des principes éthiques au langage informatique

très restrictifs par rapport à ce qu'on entend en philosophie par valeur ou norme éthique ». Nous rejoignons l'approche selon laquelle les « valeurs ou cadres éthiques représentés ou simulés dans une machine constituent bien des représentations, des simplifications, des interprétations de concepts complexes » qui ne peuvent traduire une véritable éthique. Dans le cadre de notre Dilemme n° 2, dans le cas où les dommages anticipés dans les deux solutions sont identiques, il serait possible de favoriser la survie des soldats. Il ne s'agit pas ici d'une véritable éthique mais bien d'une grossière application par le robot d'un choix jugé préférable par l'homme. Cet ordre de préférence découle des seuils et valeurs sélectionnés et choisis en amont.

Dans le cadre d'une approche simultanée, les avantages et pertes escomptés du fait de l'action militaire doivent être mesurés. Ronald. C ARKIN imagine ainsi un algorithme d'optimisation de proportionnalité qui maximise le nombre des pertes ennemies tout en minimisant le nombre de pertes civiles <sup>(26)</sup>. Cette approche est donc à mettre en parallèle avec la notion d'utilitarisme développée par Jeremy BENTHAM <sup>(27)</sup>, souvent appliquée en économie. Il s'agit d'une forme de conséquentialisme, qui consiste à évaluer une action ou une règle à l'aune de ses conséquences. Une situation militaire sera ainsi jugée bonne si elle maximise la somme des « avantages » tout en minimisant celle des pertes matérielles et humaines, en fixant au préalable une valeur d'équilibre incarnant un seuil d'acceptabilité.

Pour tenir lieu de garde-fou ultime de l'action militaire, il serait intéressant d'appliquer la doctrine du double effet <sup>(28)</sup> à nos dilemmes éthiques. D'après celle-ci, une décision est acceptable si elle respecte les trois règles suivantes :

1. La décision doit être bonne ou neutre.
2. Les conséquences négatives ne doivent être ni une fin ni un moyen.
3. Les conséquences négatives doivent être proportionnelles aux conséquences positives.

Il est important d'étudier ces règles à la lumière du DIH, socle du protocole respecté par la machine dans sa tentative de raisonnement juridique et éthique. En outre, l'éthique que pourrait alors appliquer le robot demeurerait bien une éthique biaisée et restreinte, car traduite par de grossières variables quantitativement appréciables par le chef militaire. Le robot n'est pas éthique mais peut appliquer des règles, seuils, et raisonnement considérés comme étant le plus éthique possible par son concepteur et le chef militaire qui le configure et l'emploie. Ainsi, le biais est déjà présent dès la programmation du robot.

---

(26) ARKIN Ronald C., *Motor schema-based mobile robot navigation, in International Conference on Robotics and Automation*, 1987, p. 264-271

(27) BENTHAM Jeremy, *An Introduction to the Principles of Morals and Legislation*, 1789.

(28) MACINTYRE Alison, « Doctrine of Double Effect », *The Stanford Encyclopedia of Philosophy*, 2014 (<https://plato.stanford.edu/entries/double-effect/>).

Dilemmes éthiques militaires :  
du langage juridique et des principes éthiques au langage informatique

*Une subjectivité qui persiste*

La Cerna indique en particulier dans son rapport de 2017 au sujet des robots semi-autonomes qu'il est nécessaire que le concepteur soit attentif à la capacité d'interprétation du robot et qu'il en précise les limites <sup>(29)</sup>.

Les règles de comportement sélectionnées, car admises comme éthiques par la doctrine interarmées, seront nécessairement influencées par un contexte social, professionnel et institutionnel. Il faut par ailleurs éviter les biais qui influencent le jugement et qui peuvent être inclus dans le programme par les concepteurs mêmes, ce qu'on pourrait assimiler au framing effect <sup>(30)</sup>, connu pour influencer la capacité à prendre des risques en fonction, par exemple, de la rédaction sous forme négative de l'énoncé. Cela est d'autant plus important lorsque l'on cherche à écrire une description formalisée de situation par une description formelle des faits en vue d'un calcul automatique de décision. Alors, la transcription en langage informatique est intéressante en ce qu'elle permet de limiter les effets pernicioeux du langage naturel. On note cependant que la décision calculée n'est pas « éthique » dans l'absolu, mais relative à un cadre ou un sous-ensemble de cadres éthiques précisés préalablement.

Néanmoins, si le biais est déjà présent dans la perception de l'éthique dès la programmation du robot, la mainmise de l'humain est rassurante en ce que l'éthique projetée par l'homme dans la machine demeure subjective. L'intuition, l'instinct et les raisonnements cognitifs du chef militaire le mèneront à l'attribution de telles ou telles valeurs ou à l'établissement de telles bornes en fonction des circonstances. Toujours dans l'optique de garantir une présence humaine, il est possible d'assurer l'intervention de l'éthique « humaine » en aval par l'instauration obligatoire d'un « veto power » concrétisé par un « bouton *off* ». Le chef militaire – ou éventuellement le soldat – qui se rend compte d'un choix qui lui paraît éthiquement dramatique, malgré les calculs effectués par l'algorithme, pourrait alors prendre la décision et la responsabilité de paralyser l'action du robot, encore faut-il que l'humain puisse comprendre le choix effectué par le robot, et donc connaître les arguments qui ont mené à ce choix.

\*  
\*\*

Nombre d'ONG ainsi que le Vatican considèrent par ailleurs comme contraire à la dignité humaine la possibilité pour le robot de prendre une décision de tir létal. Se profile l'idée d'un « droit humain à ne pas être tué par une machine ». Il pourrait être opposé à ces contestations l'argument déontologique de Robert SPARROW <sup>(31)</sup> selon lequel les bombardements de Hiroshima et Nagasaki seraient plus respectueux de la

(29) CERNA, *op. cit.*, p. 26.

(30) TVERSKY Amos et KAHNEMAN Daniel, « The Framing of Decisions and the Psychology of Choice », *Science*, vol. 211, n° 4481, 30 janvier 1981, p. 453-458.

(31) JEANGÈNE VILMER Jean-Baptiste, « *Terminator Ethics* : faut-il interdire les robots tueurs », *op. cit.*

Dilemmes éthiques militaires :  
du langage juridique et des principes éthiques au langage informatique

dignité humaine que n'importe quel Sala ou Salsa <sup>(32)</sup> du seul fait qu'ils proviennent d'une décision humaine.

Peut-être faudrait-il davantage concentrer nos efforts de réflexion sur l'encadrement des robots semi-autonomes et la détermination de leur cadre d'emploi car ces derniers commencent déjà à voir le jour sous leur forme la plus primitive.

De notre étude découle la conclusion suivante : il n'est pas possible, à l'heure actuelle, d'intégrer un raisonnement éthique et juridique au sein d'un système algorithmique semi-autonome semblable à celui que pourrait tenir un soldat, rendant alors acceptable son usage sur un champ de bataille modernisé. Le principal obstacle à la réalisation d'une machine éthique étant la formalisation de grands principes subjectifs et qualitatifs. Aussi le recours aujourd'hui à de tels robots, suivant ce protocole fondé sur une « charte constitutionnelle » pratique, ne pourrait se faire qu'en connaissance de cause. Les dilemmes éthiques présentés ici l'illustrent clairement. Ces derniers sont fort complexes et un esprit humain aurait lui-même de grandes difficultés à prendre une décision en de pareilles situations. Seule la mainmise du chef militaire sur un robot létal est donc aujourd'hui envisageable. Ainsi, si l'on accepte l'usage d'un tel robot semi-autonome, en prenant en compte que toute solution à un dilemme éthique n'est pas nécessairement universelle, on accepte *de facto* de recourir une limitation de raisonnement juridique et éthique. ♦

---

(32) Système d'armes létal semi-autonome.

