

Intervention de Dominique Lambert

Nous voudrions partir de l'éthique : qui est le champ philosophique dont le but est d'étudier les jugements de valeurs. D'une autre manière on peut dire que l'éthique est la science qui étudie les références, les fondements par rapport auxquels les comportements sont jugés comme acceptables ou non.

Evaluer éthiquement les LAWS par exemple exige que l'on donne et que l'on clarifie un système général de valeurs qui vont servir de références pour juger si tel ou tel système d'armes va être ou non acceptable

Ceci n'est pas simple, car dans un contexte multiculturel et multinational on pourrait avoir tendance à nier l'existence d'un tel système de valeurs de référence. Cette difficulté est réelle, mais on pourrait penser qu'il existe une solution qui pourrait engendrer une communauté de valeurs autour d'une conception assez générale de l'humain, de ce qui est ou non humain.

Nous sommes encouragés à cela par le fait que bon nombre de corpus juridiques ou de clauses importantes comme celles de Martens sont fondées en fait sur une notion assez large mais importante de conscience publique et d'humanité. Le contenu d'une telle clause joue le rôle d'une sorte d'horizon éthique qui même si son contenu est parfois difficile à cerner, indique que l'on ne peut se passer et que l'on se réfère ultimement à ce qui constitue la personne humaine.

Mais alors resurgit la question quels sont les traits de l'humain, à quoi pourrait-on reconnaître l'humanité ? Pour employer un vocabulaire philosophique : sur base de quoi peut-on construire une anthropologie, une représentation de l'être humain dans ce qu'il a d'essentiel ?

Cette question ne s'épuise pas en quelques minutes, mais il nous suffit ici de nommer quelques éléments qui pourront nous servir pour analyser notre question et qui structurent une conception, une définition de l'humain :

1 La personne humaine est un sujet responsable. Doué de conscience et de liberté, il peut répondre (et c'est cela étymologiquement le sens du mot responsable) de ses actes. Comme le dit Kant dans sa *Doctrine du Droit* : « Une personne est ce sujet dont les actions sont susceptibles d'imputation. La chose est ce qui n'est susceptible d'aucune imputation. »

2 La personne humaine se révèle et se constitue en partie dans une relation interpersonnelle, dans un face-à-face. C'est Levinas qui nous a appris que la rencontre avec le visage d'un autre, de l'autre fait entrer dans une dimension véritablement éthique et humaine.

3 Ce qui est humain c'est aussi la capacité de créer, d'inventer, d'innover, de transgresser les règles, de les suspendre, pour en inventer d'autres. Cette créativité, scientifique, artistique, éthique est le propre d'une être qui ne se laisse enfermer dans aucun système, dans aucune clôture de théories qui sont toujours trop courtes pour dire l'humain. Comme le dit encore Pascal, « L'homme

« passe infiniment l'homme », l'homme est caractérisé par le fait qu'il est toujours en-avant de lui-même, un être ouvert à l'infini.

4 L'humain se reconnaît aussi dans sa capacité à prendre en compte les être vulnérables. C'est Darwin qui dans la conclusion de son livre de 1871 : *The Descent of Man, La filiation humaine* avoue que la partie la plus noble de l'homme, ce qui fait l'homme en définitive c'est le fait de faire des lois pour les pauvres, de construire des hôpitaux, etc. Bref de se préoccuper de ce que les règles utilitaristes viendraient considérer comme des réalités sans grande valeur. L'humain se voit, dès la préhistoire, dans une sorte de comportement d'empathie, de compassion, en même temps que son intelligence peut l'amener à perpétrer les pires exactions. Mais il a chez lui un trait spécifique qui lui fait respecter les humains vulnérables ; les écraser relèverait au contraire de la barbarie, de l'inhumanité ! ; un trait qui lui fait aussi respecter les êtres fragiles de son environnement.

Responsabilité, capacité relationnelle, créativité, compassion, voici quelques traits qui, dans le fond, sont associés à l'humanité et qui pourrait être intéressants pour fonder une sorte de référence éthique que l'on pourrait appeler la nature ou la dignité humaine.

Nous pourrions énoncer comme principe régulateur d'une éthique basée sur cette référence, l'impératif suivant : « agis de telle manière que tu restes toujours cohérent avec ce qui définit l'humain en profondeur », l'humain envisagé, entre autres, au travers des quelques traits que je viens de mentionner. Ceci n'est somme toute qu'un **principe d'auto-cohérence anthropologique** : qui demande d'agir de telle manière à préserver l'humanité de la femme et de l'homme d'aujourd'hui et de demain.

Appliquons ce principe de cohérence à notre question des robots.

Les robots font des choses mieux et plus vite que ce que nous pouvons faire c'est assez clair et les développer fait partie de notre nature d'*homo faber*, mais il faut reconnaître en vertu du principe d'auto-cohérence anthropologique :

- Que l'on ne voudrait pas d'un robot qui se retourne contre l'humain, ou qui ne réalise pas ce que le décideur humain politique ou militaire lui a prescrit. Le refus de l'imprédictibilité est une conséquence de cette cohérence. L'homme aidé de la robotique ne doit pas devenir un esclave de la robotique.
- Que l'on ne voudrait pas non plus un système dans lequel l'humain perdrait ou diluerait voire cacherait sa responsabilité qui fait de son acte un acte vraiment assumé. Une machine est un objet, un bel objet, mais jamais un sujet responsable.

Les deux premiers points signifient que l'on ne voudrait pas de robots qui sortiraient d'une chaîne de commandement et d'autorité, légitime et responsable.

- Que l'on ne voudrait pas d'un robot ou d'une société, envahie par des robots, qui induirait la disparition de toute relation entre les humains et l'abandon des humains à une foule de machines qui imiteraient peut-être bien les émotions, les pensées des humains mais qui ne seraient pas des humains avec leur visage en chair et en os, qui ne seraient pas les êtres sensibles, capables d'inventions intellectuelles ou affectives, de compassion ou de pardon. Humains capables de nous tirer de nos calculs pour nous faire entrer dans une dimension fraternelle.

Je pense que l'on ne peut peut-être pas définir avec précision ce que sont les LAWS. Mais on peut cependant exhiber et définir, au minimum et de manière préventive, **les LAWS que l'on ne voudrait pas** pour des raisons de cohérence humaine. Ici ce ne sont pas les définitions qui vont précéder les questions et cadres éthique, c'est plutôt l'éthique qui va fonder les définitions négatives des machines que l'on ne voudrait certainement pas au nom d'une conception minimale et commune de ce que constitue l'humanité. C'est d'ailleurs peut-être cela qui pourrait donner un sens à ce que nous appelons un contrôle humain significatif, lorsqu'on aborde les fonctions critiques des LAWS.

Remarquons que l'on pourrait admettre des robots autonomes qui servent l'humain, sa vie, sa survie dans un contexte complexe, l'autonomie voulant dire ici une absence d'un pilotage humain et aussi d'une supervision du pilotage, s'il est entendu que les robots se situent dans un espace délimité et n'exécutent que des actions qui sont connues ou en permanente cohérence avec ce que les autorités responsables et l'éthique prescrivent.

Ce que l'on ne voudrait donc pas c'est en fonction des références éthiques de cohérence avec notre humanité dans ce qu'elle de plus profond et de plus commune à nous toutes et tous, c'est des robots autonomes tellement innovants du point de vue de l'auto-programmation et de l'auto-apprentissage qu'ils nous privent de notre responsabilité, de nos projets, de nos décisions, de nos intentions, mais aussi de notre capacité de compassion, de pardon et aussi de tendresse.

Une éthique de la commune humanité partagée, nous permet donc peut-être, à défaut d'une définition de robots armés qui n'existent pas, de préciser une définition minimale des LAWS que nous ne voudrions pas ! Le pas serait minime, mais il serait fondé, ce qui ne serait déjà pas si mal.