Elements Supporting the Prohibition of Lethal Autonomous Weapons Systems

***

An autonomous weapon system is a weapon system capable of identifying, selecting and triggering action on a target without human supervision. Among these types of systems, we can distinguish different levels of autonomy. For example, it is conceivable that weapons meet the definition that we have given, but which are not likely to re-program themselves (according to some learning, self-learning, or adaptation to a certain environment). Moreover, one can think of systems capable of learning and reprogramming (system based on genetic algorithm). This could allow an armed autonomous robot, for example, to redefine itself and the objectives that it must achieve, without the mediation of a human control.

When considering autonomous systems, it is important to specify whether one contemplates its use in controlled environments (geographical areas well identified wherein we know all the parameters: objects and persons therein and events that can occur there) or its use in open environments (i.e., not defined). It is also important to know if all the possible behaviors are unpredictable or predictable (in the case of reprogrammable machines, it is certainly not the case, by definition).

If an unarmed autonomous robot works in a defined and known environment (we know exactly what this environment contains), and if all of its possible behavior is predictable (even if his particular behavior at some point and at some time, is meanwhile not predictable), it poses no particular problem (if such behavior is in accordance with the law and ethical principles that protect people and property).

In the contrary, if a robot, without human supervision, operates in an open environment, i.e., it is not controlled, and if all of its behavior is not predictable, it may be dangerous to be used even if it is not armed. Think for example of an autonomous robot that has entered a nuclear power plant (well marked perimeter) to repair some facilities, but whose behavior is not predictable because of its ability to reprogram itself. One can easily imagine what this would mean in terms of risks. In rescue and repair operations, it is obvious that we should leave to the machine some degree of freedom, however it is important to ensure that it will not perform actions that we condemn or dread.

The definition of autonomy of an armed robot involves the following:

1) the absence of human supervision (this absence can be permanent or temporary. In the last case, there is a way to prevent the autonomous mode. An operator can perform from outside this deactivation or an internal system to the machine can redirect the control to a human operator).

Differences between the supervised machines can appear when one specifies:

2) all the possible behaviors of the robot (behavior determined by a given algorithm or algorithms enabling learning and reprogramming).

3) the environment in which the robot operates: as specified from the point of view
- (a) its geographical boundaries
- (b) its contents (of persons and goods)
- and (c) all likely events.

We can characterize armed autonomous robots using these three points: (1) the degree and duration of supervision, (2) the predictability of the behavior of the robot, (3) and the characteristics of the environment in which it operates.

The "human control" of a robot with a certain degree of autonomy, in the sense that its behavior in an environment are not prescribed by an operator, can be in fact "significant" if there is a human supervision (cf. (1) allowing in any time to disable the autonomous mode of operation), if the environment in which the robot is perfectly circumscribed and known (see (3a, b, c)), and finally, if all behaviors are well known (see (2)).

If the robot is armed with a lethal or a non-lethal weapon and that, moreover, it is not supervised or that it operates in an open environment in which events and everything in it are not fully known, it is likely to pose serious problems from the standpoint of international humanitarian law or ethics.

What are the risks of using weapons systems with a large degree of autonomy?

Besides the fact that it delegates to a machine the decision of life or death of a human being, one of the dangers is that these weapons could lead to strategies diluting or concealing true responsibilities in case of collateral damages. If armed machines provoke such damages it is easy and tempting, on the part of those who use them, to invoke technical malfunctions rather than face their responsibility. This could then lead some authorities to be less conservative, concerning precisely these collateral damages, knowing that there will be a way to protect themselves, on the pretext of failing technological mediations.

The autonomous robotic weapon necessarily induces a form of lack of accountability. It puts a distance, in a new way, between the human being and the gravity of the battle and its material and human consequences. We can already see this, for example, in the field of social networks. Violence can be unleashed in the virtual space while the actors, who are at the origin, do not realize the damage they can produce at the people who are attacked or ridiculed. The screen of technological mediations carries with it a significant risk of the occultation of our responsibilities and the concrete effect of our actions.

Another danger is the risk of hacking or the use of these systems by terrorist groups who could have recovered them. We know that complex systems are vulnerable to hacker attacks and consequently we cannot use machines with such destructive capabilities with the possible risk of misuse by malicious groups.

The development of armed autonomous robots will induce and stimulate an arms race, with its attendant costs and risks of reinforcing oppositions between nations. The speed of development in the field of computer science and robotics is such that the balance of forces will quickly be broken. Each side will try to take advantage of

very rapid technical innovations in the field to gain superiority. A frenetic race may ensue. The more a military technique has the ability to advance, the higher the risk of seeing an arms race with permanent instability. This is the case of this type of technology (see the "Moore's Law" in computer science).

We should not underestimate the psychological impact of these weapons on people (already remotely operated drones pose such problems). In addition to the anxiety of the risks of nuclear, chemical or biological war, we will add the anxiety of seeing machines deciding about our life and death. As already Pope Paul VI said in his speech to the United Nations (4 October 1965): "The weapons, especially the terrible weapons that modern science has given you, even before causing victims and ruins, cause bad dreams, feeding bad feelings, create nightmares, distrust, dark resolutions; they demand enormous expenditures; they stop the projects of solidarity and useful work; they distort the psychology of peoples (No. 5)."

Far from contributing to the defense of peace, the accumulation of robotic arms could turn into a progressive incitement to war. This accumulation is all the easier as the technology of robots is relatively accessible.

One of the most common hazards associated with the use of robots is the fascination they hold for their performance. This can lead to overestimated value them and to even to the concealment of certain technological limitations. One could, in fact, believe that these machines by using their artificial intelligence capabilities, could replace the human person and meet the requirements of international humanitarian law. However, this is not the case. Why? Even assuming that the machines can correctly apply the principle of distinction between combatant and non-combatant (which does not happen by itself), it could not properly implement the principle of proportionality which requires careful judgment involving unquantifiable concepts. The reference to the meaning of the military value of a target is also problematic. Prudence and respect for the law sometimes require going beyond the "letter" of the law and so as to interpret it according to the context to preserve the "spirit". This ability of interpretation and going beyond the rules is not programmable. Even if a machine had reprogramming capabilities, it would still follow the rules. But humans have this ability to be able at any time to innovate out of the shackles of rigid rules, taking into account the unexpected, the unusual event. Human creativity is essential in certain circumstances, to get out of situations that threaten the human being, his dignity and the stability of societies. And it is precisely this creativity that is lacking in machines even autonomous ones. To be convinced, just think about the problem of whether it is legitimate or even possible to replace, in a court, a judge with a machine (reflecting on the fact that the mode of the exercise of justice is not reduced to the mechanical and algorithmic application of rules, even if they are programmed in a suitable logic adapted to the legal world, the deontic logic, for example).

If we want peace, we must not only avoid accumulation of arms, but we must also convert minds. "Peace must be born of mutual trust between nations, instead of being imposed on nations by the terror of weapons," said Vatican II (Gaudium et Spes 81). This confidence is based on an "ethics of brotherhood" between nations. But the accumulation of lethal autonomous weapons could undermine that trust. Indeed, can a nation really have trust in another if it has the capacity without incurring any risk, constantly to observe and hit the first with robotic arms? One could also point out that the disappearance of the human fighter will induce the disappearance of what

the relationship of a person to a person and the discovery of the face of the other could provoke. A machine cannot have real empathy (this requires the experience to feel in one's body what the other feels in his body. The machine has no real corporeality). A machine is not open to the unexpected forgiveness and to a real possibility of reconciliation or pacification.

We must reflect on the use of technologically advanced weapons in the context of the new dimensions of the war. The battlefield has become global and extends both in the physical and virtual dimensions (cyberspace). The history of conflicts since the fifties of the twentieth century showed that this use has never resolved the conflict against determined specific groups or nations and knowing well enough the land on which they operate. It is clear that the huge spending on sophisticated weapons fails to restore peace. Quite the contrary! It appears that this sophisticated technology does not protect us against attacks and terrorism of all kinds perpetrated by people using rudimentary methods, but ready to sacrifice their lives. We should perhaps ask whether the race for technological sophistication has not reached its limits. The balance of nuclear terror has shown its limits, and with the research and development on lethal autonomous weapons, we are still in a logic that bears no fruit. We must face the facts, the race for sophisticated weaponry, including Lethal Autonomous Weapon Systems, is not able to fight against the current scourge of terrorism. The real fight is the one which will restore justice, respect for human rights, respect for minorities' rights, political participation, integral development, etc. This fight will not be won with technologically powerful weapons (robotized or not). The use of LAWS will only lead to false security and to instability. In any case, it will not establish the conditions for peace.

One could say that LAWS should not be prohibited simply because they are not able to meet the principles of international humanitarian law and the requirements of a prudent and responsible decision, but also because they lead us in a way that showed its limits. The fight against the forms of the most widespread wars now, namely terrorism, urban warfare, asymmetrical conflicts, cannot be won by increasing the firepower and sophistication of weapons. It has to happen in the field of dialogue and the common search for peaceful solutions to the real problems of our shared planet.

The refusal of LAWS must be based on the fact that even from an operational point of view, they cannot provide a lasting solution to the conflicts of today. We should draw conclusions from recent history and not be locked in a vision that is quite outdated and unrealistic. We must also be very careful about the economic logic that could push the development of these robotic technologies simply for reasons of financial gain and not for genuine needs of security and defense. For all the reasons mentioned before, we must be cautious about the research and development of LAWS. Now is the time to prevent LAWS from becoming the reality of tomorrow's warfare. The CCW should make a courageous decision of prohibiting lethal autonomous weapons like it did in the past concerning other types of weapons.