

Group of Governmental Experts of the High Contracting Parties to the Convention on Prohibitions or Restrictions on the Use of Certain Conventional Weapons Which May Be Deemed to Be Excessively Injurious or to Have Indiscriminate Effects

7 November 2017

English only

Geneva, 13–17 November 2017

Item 6 of the revised provisional agenda

Examination of various dimensions of emerging technologies in the area of lethal autonomous weapons systems, in the context of the objectives and purposes of the Convention

Towards a definition of lethal autonomous weapons systems

Submitted by Belgium

1. Since 2013, States Parties to the Convention on Prohibitions or Restrictions on the Use of Certain Conventional Weapons Which May Be Deemed to Be Excessively Injurious or to Have Indiscriminate Effects (CCW) have informally tackled the issue of emerging technologies in the area of lethal autonomous weapons systems (LAWS).
2. Three meetings at experts' level have already taken place in the CCW framework. The discussions have allowed for a preliminary examination of the topic from various perspectives (legal, ethical, technical, etc.) without reaching a common characterization of LAWS. However, previous discussions have led to the unanimous acknowledgement that fully autonomous lethal weapons systems do not exist yet. Although excluding existing weapons systems (whether automated or partially autonomous) from the discussions, this recognition has not allowed for a clear conceptualization of LAWS.
3. In order to enhance the forthcoming discussions of the Group of Governmental Experts (GGE) related to emerging technologies in the area of Lethal Autonomous Weapons Systems that was set up by a decision of the 5th Review Conference of the CCW in December 2016, as well as to encourage debates based on shared and common concepts, Belgium wishes to summarize its view of the state of the question and to share its thoughts as follow : first, an introductory framework on the notions of autonomy, unpredictability, intentionality with lethal consequences ; then, as part of a conceptual approach, an enumeration of the concepts to be taken into consideration in order to set up the constitutive characteristics of LAWS allowing for a possible future definition; finally, elements to be taken into account as a result of legal, operational and technical frameworks that would apply to LAWS.

An introductory framework on the notions of autonomy, unpredictability, intentionality with lethal consequence

4. In Belgium's view, LAWS and non-autonomous weapons systems should be considered as being different in nature, and not in degree. As far as non-autonomous weapons systems, including automated weapons, are concerned, the ultimate decision to produce (even potential and/or delayed) lethal effects fully remains in the hands of a human being. There is no other authority involved in the decision and its effects. As for LAWS, on

GE.17-19491(E)



* 1 7 1 9 4 9 1 *

Please recycle 

the other hand, the human being would either not intervene in the ultimate decision to inflict a lethal effect, or he would intervene among other “artificial deliberative” authorities taking ultimately the decision whether to inflict or not the lethal effects. This latter observation has two consequences: (1) the human being would be sidelined in the decision-making process (he would only remain the originator of the lethal process) and (2) the ‘result’ (i.e. the implementation of lethal effects) would become potentially highly unpredictable for the human agent. In the case of LAWS, the level of unpredictability of the consequences of the human decision regarding targeting and collateral damages could be higher than in classical military actions.

5. In light of the above, the notion of ‘lethality’ goes hand in hand with the notion of intentionality. As a consequence, the decision whether to use LAWS with the specific objective to kill, i.e. the intended use with lethal consequence, or not, remains under human authority, be it political and/or military. The LAWS’ level of autonomy in the lethal decision-making presupposes a clear definition of the division of authority between the human authority (political and/or military) and the machine.

6. Existing automated weapons systems should be explicitly excluded from the debate on LAWS, since they do not differ in nature from non-autonomous weapons systems. Automated weapons systems, including mines, fall beyond the scope of this debate because such systems are programmed and their whole functioning is predictable and intended. Furthermore, partially autonomous weapons systems, such as anti-missiles systems, are also fully programmed and do not allow the emergence of an artificial decision-making power that goes beyond what is included in the human intention to use such systems.

7. A constant vigilance should also be given to the potential dual-use nature of robotic technology. A future characterization of LAWS should not hinder the development and use of robotic technology for civilian purposes.

Characteristics or constitutive elements regarding the notions of autonomy, intentionality with lethal consequence, control and unpredictability

8. In Belgium’s view, the following characteristics or constitutive elements regarding the notions of autonomy, intentionality with lethal consequence, control and unpredictability should be considered in the framework of a strictly conceptual exercise aiming at defining LAWS:

(a) Total autonomy in the lethal decision-making process, i.e. LAWS that would be able to switch to lethal mode – or to a mode in which they could inflict wounds to a human person – without any previous or marginal human decision;

(b) Full independence from human intervention, at any stage, in the ability to identify and select targets with the intent to maim or kill;

(c) An unclear or uncertain division of authority between the human agent and the machine in the intentionality with lethal consequence, as well as a division of authority that would not be subjected to a precise criteria based assessment;

(d) The impossibility to bring, at any time and upon human decision, LAWS working in autonomous mode back to remotely controlled mode, or to deactivate them;

(e) The openness (i.e. uncertain, unpredictable or unreliable character) or the limited knowledge of the entirety or only one of the potential behaviors of LAWS;

(f) LAWS’ ability to redefine by themselves the criteria according to which they would be able to operate in terms of environment, targeting or mission among others.

Elements of legal, operational and technical frameworks

9. Furthermore, Belgium wants to reaffirm that existing legal frameworks have been developed for human agents, not for autonomous non-human agents. Rights and obligations

apply to human agents, not to robotic agents. To presuppose that LAWS are immoral agents (thus not responsible) would obfuscate the fact that, ultimately, it's human agents who would decide to use LAWS. From this point of view, the possible introduction of LAWS as means of war could only be carried out within the existing normative framework, according to which the human agent bears the ultimate responsibility for the chosen means and methods of war. In this context, the following legal, operational and technical issues should be taken into consideration:

A. About the legal framework:

The impossibility, for LAWS operating in autonomous mode, to fulfill the requirements under international law and in particular, the law of armed conflict and the international human rights law, as the case may be.

B. About the operational framework:

- (i) The possibility for LAWS to move outside a clearly limited geographical perimeter, defined by military and/or political authority;
- (ii) The possibility, for LAWS that would be used according to a set of well-established behaviors, to operate within a zone where external persons to the conflict or to the assigned mission could be located/could enter;
- (iii) The possibility for LAWS to act in insufficiently defined, controlled and protected areas.

C. About the technical framework: securing of the system, legal assessment software, self-learning, reprogramming, hacking:

- (i) The absence of identification and of explicit consent by the human authority with regards to the laws and guiding principles of legal assessment software which would equip LAWS;
 - (ii) The ability of LAWS, after a self-learning process, to redefine their missions or objectives, with no evaluation nor human decision. And in the case where a self-learning mode would exist, the absence of human supervision and validation of what has been learnt;
 - (iii) The impossibility to deactivate the self-learning ability of LAWS during an operation;
 - (iv) The ability of LAWS to define their own evaluation criteria of actions;
 - (v) The absence of a high degree of protection against the hacking of LAWS.
-