**Chair’s introduction of Agenda Item 6 (a), 9 April 2018**

This afternoon we begin consideration of the first substantive item on our agenda by looking at approaches to characterization of the systems under consideration in order to promote a common understanding on concepts and characteristics relevant to the objectives and purposes of the CCW. Broadly speaking, we could take four approaches to the characterization issue.

I. What I call the ‘separative’ approach. You would recall that during the discussion on legal/ethical aspects last November we spoke of the Via Negativa/Via Positiva.

*Via Negativa* would mean:

What are the characteristics and concepts not relevant to the objectives and purposes of the CCW? E.g. tele-operated drones, which are a platform that is part of a capability (weapon system) that cannot function without humans.

*Via Positiva* would mean on the other hand:

Characteristics and concepts that are definitely relevant to the objectives and purposes of the CCW? A lethal weapon system that programmes itself without any human involvement would definitely be of interest to the CCW community.

Thus by a “separative” process of exclusion and inclusion, concepts and characteristics of relevance to the GGE and its mandate will become clearer.

II. A second approach would be a cumulative approach.

This could also be called an “aggregative” approach in contrast to the previous “separative” approach. This is the approach we have willy-nilly followed in our discussions in Geneva. Categories of characteristics could be added to a master-list of concepts and characteristics based on what High Contracting Parties bring to the discussion and then evaluated.
against certain criteria – technical, legal-humanitarian or political-security - to decide on their relevance.

Examples of such categories could include physical performance (mobility, speed and survivability/energy autonomy), targeting performance and other technical characteristics.

Or they could include a human-machine interface related category such as system type, which divides systems into Automated, Autonomous, Hybrid or others, a human-machine relationship category such as human in the loop, on the loop or out of the loop - or secondary characteristics such as reliability, predictability, command & control etc.

Thus, within the cumulative approach we could take a technology-centric approach or a human-control centric approach.

III. A third approach would be to avoid characterizing autonomous systems altogether either by using various levels of autonomy or other technical characteristics or by categories related to loss of human control. This is because in the view of some machines cannot be autonomous by definition or because an abstract approach to human-control implies a lot of assumptions, which need to be made explicit.

In this third approach, let us call it an accountability-oriented approach, we could consider a set of characteristics related to removal of functions i.e. the type of decisions and actions being handed over to machines. One example would be ATR or automatic/autonomous target recognition. This third approach would naturally be more context dependent and involve a combination of technical and human-interface evaluations but centered around what use the system is being put to. This approach recalls broadly SIPRI's ‘functional approach’. It is also not unrelated to ICRC’s critical functions approach.

IV. A fourth approach could be a Purpose-oriented Approach (or as the Chair mentioned later, an effects-oriented approach), in which you iterate backwards from the desired purpose to characteristics of the systems under consideration. For example, in the general debate earlier today, Mexico mentioned their preference for a law-centric approach rather than a lethality-based or autonomy-based approach.
These remarks, in addition to the compilation of ideas e-mailed to all of you, are meant to stimulate your thinking. This is just an aid. Nothing more and nothing less.

In this exercise, there are perhaps some caveats we could keep in mind. Let me suggest some based strictly on what you have yourselves suggested in your Working Papers.

First, intelligibility. We should make sense to each other and to the wider world outside. This is one reason I circulated the draft glossary of terms that IEEE is working out. I particularly like the three-broad sources they are tapping into for each term – the technology domain, the social science domain and the policy domain.

Second, let us try not to prejudge the regulatory response with our characterization. I doubt if anyone of us can sneak anything past the others in this room. Let us play it as we say in cricket with a ‘straight bat’.

Third, let us try not to stigmatise technology as such.

There could be other guiding principles that you may wish to share today or tomorrow as we deepen this discussion.

Let us start.