Lecture 7 - Depth of interactions

Government, autonomity, responsibility

- So, if we pose the fundamental question of what results we would want from our increasing interactions with technology and dependence on such technology, we could pose many measures

Government, autonomity, responsibility: answers?

- Happiness:

Advantage:

- Certainly seems appropriate;

Disadvantage:

- A difficult and amorphous one, notoriously dependent on individual tastes

Government, autonomity, responsibility: answers?

- Freedom and dignity:

Advantages:

- Appeals to traditional American sensibilities,
- Not so formidable to measure as happiness;

Disadvantages:

- Not universally recognized as advantages;
- indeed, there is the very (in)famous book “Beyond Freedom and Dignity” by B.F. Skinner

Government, autonomity, responsibility: answers?

- Social justice, also

Advantage:

- Popular subject over the last few decades

Disadvantage:

- Formidably difficult to formulate, much less measure

The IEEE and AI-enabled technology

- Editorial from IEEE Intelligent Systems, Volume 30, Issue 3

The IEEE's work on ethical principles in systems design

- http://standards.ieee.org/develop/project/7000.html

The IEEE's work on formulating principles for autonomous interactions

- With specific thought to autonomity: The IEEE's "Ethically Aligned Design" document or here, website

The IEEE's work on formulating principles for autonomous interactions

- Three general principles have been proposed in the EAD document:
  - Embody the highest ideals of human rights
  - Prioritize the maximum benefit to humanity and the natural environment
  - Mitigate risks and negative impacts as AI/AS

Other issues

- Safety and beneficence of artificial general intelligence (AGI) and and artificial superintelligence (ASI)
  - Unanticipated capabilities and behavior may be dangerous
  - The Humpty-Dumpty problem: retrofitting safety into a system may not be possible
  - Cascading ethical and technical issues from the very existence of increasingly autonomous and increasingly capable systems: experiencing science fiction firsthand?

Other issues

- Autonomous weapons systems and the challenges of autonomous systems designed to harm
  - Should professional organizations hold increasingly autonomous products to the same sort of standards that their creators are held to?
  - Autonomous weapon systems have the dangerous characteristic of covert, non-attributable, and repudiable use.

Other issues

- Autonomous weapons systems and the challenges of autonomous systems designed to harm
  - Any learning system might not learn well; such a characteristic is quite literally dangerous with a weapons system.

Other issues

- Autonomous weapons systems and the challenges of autonomous systems designed to harm
  - The battlefield is already a savage enough place *with* humans; will it become even more so under autonomous weapons systems since the existing Geneva Conventions clearly don't apply to those. Cross-reference Tallinn Manual, website, on Amazon

Politics clearly are involved

- War is a political activity with political ends, as Clausewitz observed long ago:

"But however powerfully this may react on political views in particular cases, still it must always be regarded as only a modification of them; for the political view is the object, War is the means, and the means must always include the object in our conception." On War, at Gutenberg

Politics and laws clearly are involved

- Politicians create laws – even laws of war, as discussed here, the famous Hague Conventions, the previously referenced Tallinn Manual, even the environment, the impuissant Kellogg-Briand Treaty, humanitarian concerns

Politics and laws clearly are involved

- And enforcement of laws means trials: International Military Tribunal for the Far East, Nuremberg trials

The ethical implications of increasing depth of interactions

- As we have been discussing, autonomity brings on a new measure of concern; being able to explain the activities of an autonomous war system is of paramount concern

The ethical implications of increasing depth of interactions

- Currently deployed state-of-art is probably the Phalanx defense and its derivatives like C-RAM
- Scale this to offense capabilities and future improvements

Or swarming effects

- CICADA

What do we want from the study of ethics?

- Clearly ethics is more about questions than answers; we frame ethics not as a study of specific actions but of questions of why and what is important; what principles should we apply, and what principles do we not apply
- What framework do we want to apply in any ethical question