Fielding and securing autonomous combat units

JM Pelletier
Agenda

» What is happening?
» Why is it happening?
» What does it mean for the future?
» What can/should we do about it?
What is happening?

» Historical context:

1940 - TT-26 teletank (RUS)

1944 - Goliath (DEU)
What is happening?

» Historical context:


Strong AI funding and optimism

AI “winters” and downturns

Turing Letter

Dartmouth Workshop

IPTO formed

Lighthill Report

Fifth Generation

Internet Bubble

Major HCI labs form

HCI accepted in many CS departments

First Autonomous Car Fatality (2017)

Watson wins Jeopardy (2011)
What is happening?

Demystifying Artificial Intelligence

Goals of Artificial Intelligence
- Natural Language Processing
- Robotics
- Computer Vision
- Pattern Recognition
- Machine Learning
- Deep Learning
- Data Mining
- Reasoning

Stages of Artificial Intelligence

Artificial intelligence
- 1950
- 1960
- 1970
- 1980
- 1990
- 2000
- 2010

Machine learning
- Cognitive computing

Machine Consciousness
- Machine Intelligence
- Machine Learning

Watson Lives Here
What is happening?

» $$$ poured into AI systems

2017-2018: Mohammed bin Salman (KSA) and Masayoshi Son (JPN) announce commitment of over $45bn to fund/build superintelligent machines

2017: Pentagon spent $7.4bn (in 2017 alone) as part of overall strategy to increase use of AI in combat, combat support & combat service support

Project Maven
FEDOR (RUS)

TALON (USA)

Guardium (ISR)

MQ-1 Predator (USA)

Hermes 450 (ISR)

Shahed 129 (IRN)
Why is it happening?

“Artificial intelligence is the future, not only for Russia but for all humankind. Whoever becomes the leader in this sphere will become the ruler of the world.”

-Vladimir Putin, 9/1/2017
What does it mean for the future?

Update September 4th, 5:40AM ET: Elon Musk offers his cheery perspective on Twitter:

Elon Musk @elonmusk 4 Sep
It begins ... twitter.com/verge/status/9...

Elon Musk @elonmusk
China, Russia, soon all countries w strong computer science. Competition for AI superiority at national level most likely cause of WW3 imo.
5:33 AM - Sep 4, 2017
47K 22.9K people are talking about this
» United States ~ 2025
» China ~ 2030
» Russia ~ 2035

» Asymmetric option
“Any actor with the financial resources to buy, or steal, an AI APT (advanced persistent threat) system could gain access to tremendous offensive cyber-capability, even if that actor is relatively ignorant of internet security technology.

- Ministry of Defense (GBR), 1/2018
Figure 10. Schematic of Blitzkrieg Encirclement.
Demystifying Artificial Intelligence

Goals of Artificial Intelligence
- Natural Language Processing
- Robotics
- Computer Vision
- Pattern Recognition
- Machine Learning
- Deep Learning
- Data Mining
- Reasoning

Stages of Artificial Intelligence
- Machine Learning
- Machine Intelligence
- Machine Consciousness

Watson Lives Here

Artificial intelligence timeline:
- 1950
- 1960
- 1970
- 1980
- 1990
- 2000
- 2010

Cognitive computing
What can/should we do about it?

» Embed protections in AI to prevent harm to humans
  » DARPA Assured Autonomy Program
  » IEEE Global Initiative for Ethical Considerations in Artificial Intelligence and Autonomous Systems
  » Asilomar AI Principles
Proposed experimental design: From generative adversarial nets toward a Confidence-Emitting Diode for embedded goal security in autonomous combat machines

All inputs (goal hypotheses) | Adversarially-Screened outputs (accepted goals)

Deception-detector (Discriminator) | Lie-generator (Generator)

Discrimination agent screens the input; outputs a confidence level for the trusted dataset

Lying agent generates output designed to trick the trusting agent

"as iron sharpens iron, so one learning agent sharpens another"