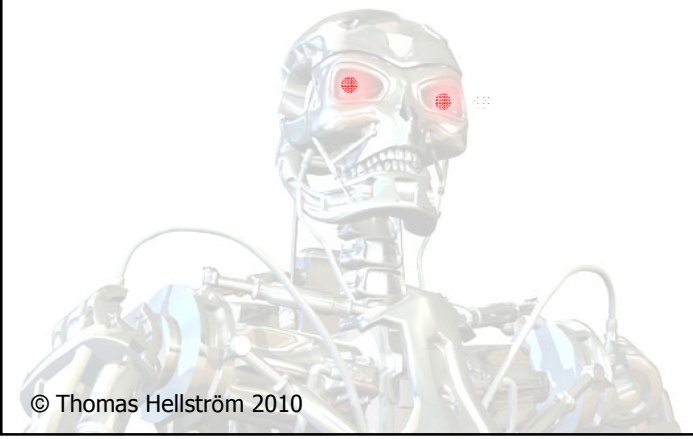
Terminator ethics

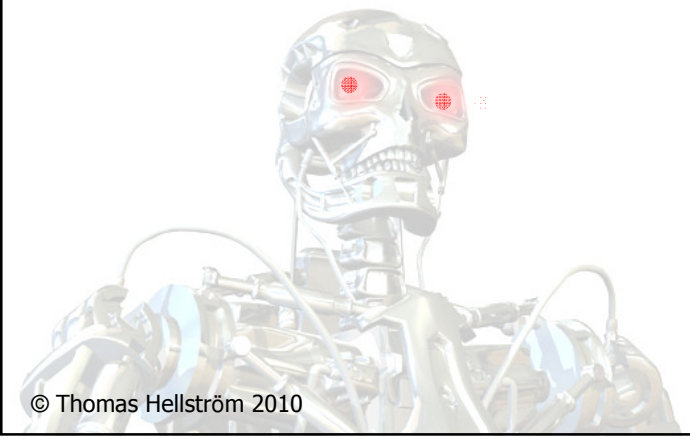
What's right and wrong
with killer robots?

Thomas Hellström
Department of Computing Science
Umeå University
Sweden

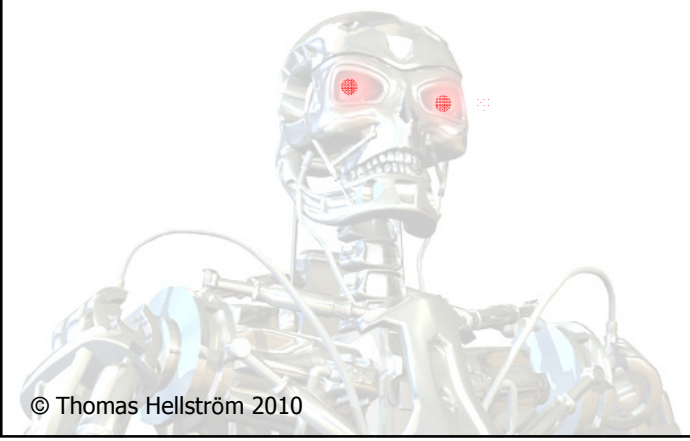


Overview

- Robots in warfare
- Ethical considerations
- Why would we (not) want robot soldiers?
- Technical challenges
- Researchers' role
- My conclusions



Robots in warfare

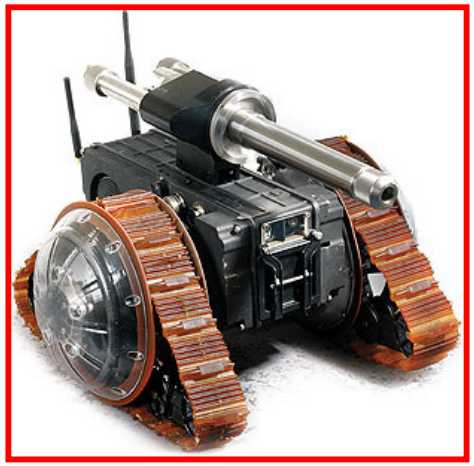


Robots in warfare

A dramatic transformation of the U.S army

U.S. congress 2001:

- By 2010 one-third of all deep-strike aircraft should be unmanned.
- By 2015 one-third of all ground vehicles should be unmanned



Robots in warfare

Most of the lethal robots are in the sky (UAVs)

- MQ1-Predator (27 feet long)
- MQ-9 Reapers
- Equipped with a Hellfire missiles
- Navigate and search out targets
- Remote controlled from the Nevada

Robots in warfare

- By the end of 2008, about 12,000 robots of nearly two dozen varieties on the ground in Iraq
- Most of these vehicles/robots are NOT lethal
  - Sniper detection, disrupting or exploding explosive devices and surveillance in dangerous areas

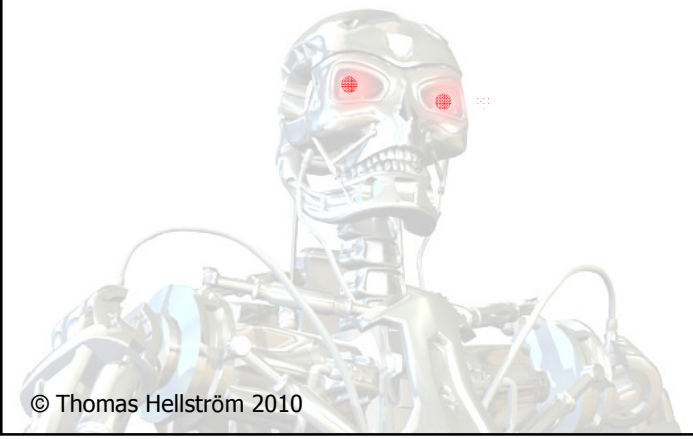
iRobot PackBot

iRobot with RedOwl Sniper Detection Kit

Robots in warfare

Phalanx system - in service since 1980

- Phalanx system for the U.S Navy.
  Task: ”search, detect, evaluate, track, engage, kill”
- May 2009: 260 M$ contract
  to upgrade the Phalanx



Robots in warfare

The SGR-1 Security Guard Robot for the Korean Demilitarized Zone

Brought to you by SAMSUNG

- Low-light high resolution and thermal cameras
- Laser Range Finder
- An 'acoustic device' emits a tone powerful enough to make intruders nauseous and drop to the ground
- Machine guns equipped with real or rubber bullets
- Automatic or with human confirmation

Robots in warfare

Neuron - Unmanned Combat Air Vehicle

- Is being developed by France, Greece, Italy, Spain, Sweden and Switzerland
- Capability to carry two laser guided 250kg bombs
- Controlled from ground or combat aircraft such as the Swedish Gripen

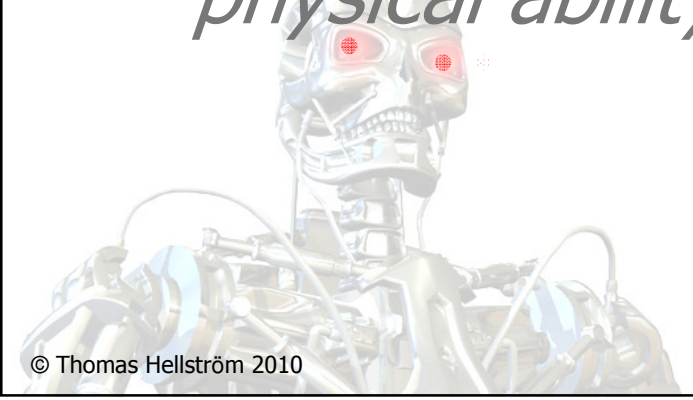
Robots in warfare

SWORDS TALON – Robot from Foster Miller

- Controlled by a soldier (up to 1000 m away) using a small console to remotely direct the robot and fire its weapons
- Can travel through sand, snow, 100 feet underwater
- Cameras: colour, black and white, infrared, night vision
- Sensors: Chemical, gas, temperature, radiation
- Robotic manipulator for disarming improvised explosive devices
- Possible weapons: machine gun, grenade launcher, anti-tank rocket launcher

Terminator ethics

- Military robots are being introduced in big scale!
- Big range of lethality and autonomy:
  - Most military robots have no weapons
  - Some robots are highly lethal and autonomous
- Important for robot ethical issues:
  - Autonomy
  - Power
  - Lethality
- By Lethality we denote *a weapon's or a robot's physical ability to kill, if being used or activated*



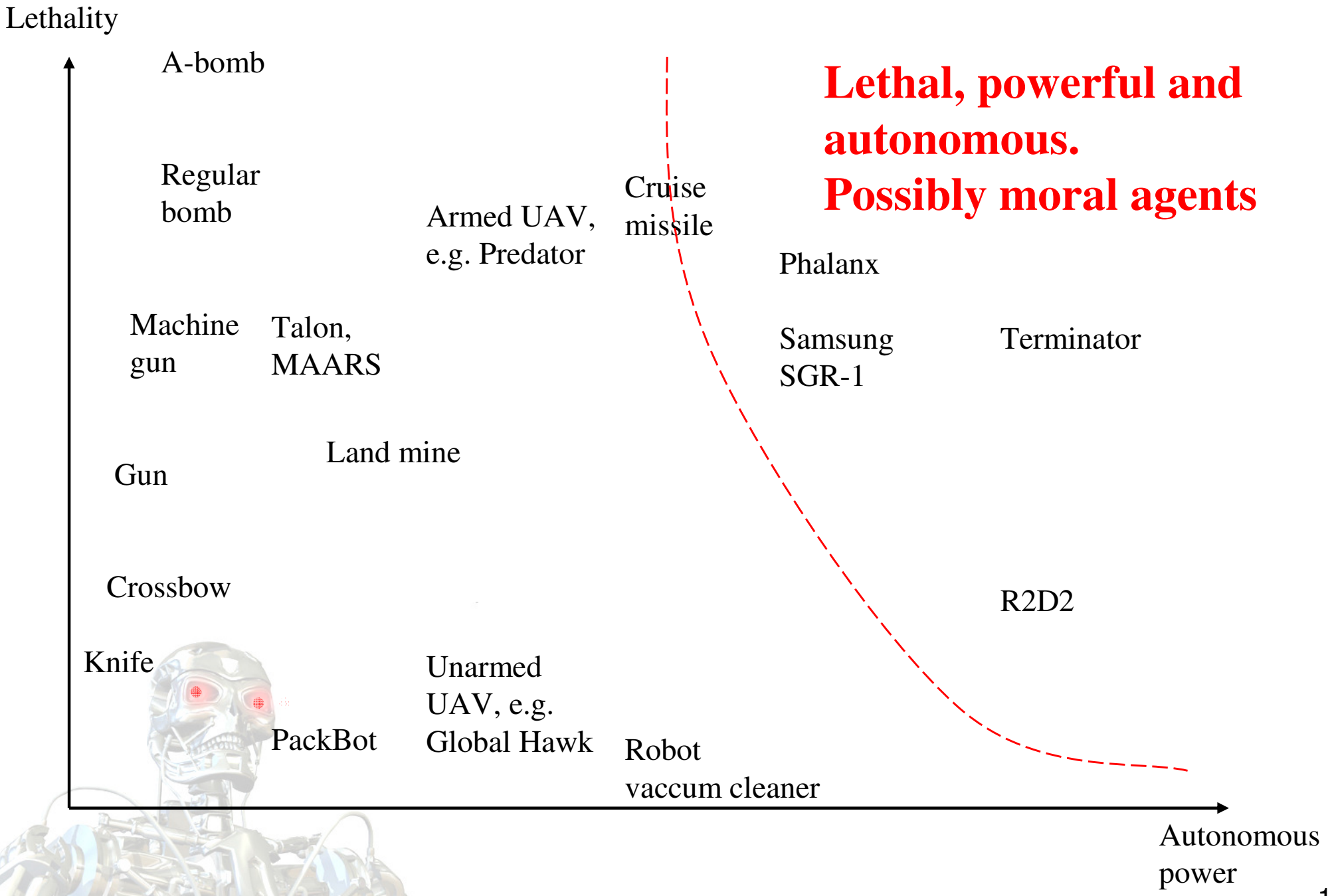
Autonomous power

- *An **autonomous agent** is a system situated within and a part of an environment that **senses** that environment and **acts** on it, over time, in pursuit of its **own agenda** and so as to effect what it senses in the future* (Franklin and Graesser 1997)

- So what's the difference between a thermostat a driverless forest machine?

- By **Autonomous power** we denote *the amount of actions and interactions an agent is capable of performing autonomously (i.e.: Autonomy & Power)*

Classification of weapons & robots



Moral responsiblity of humans

Two parts:

- causal responsibility and intention

(Dodig-Crnkovic and Persson 2008)

- a robot can not have advanced mental states such as intention (Johnson 2006)

A more pragmatic view: Moral Responsibility is a social regulatory mechanism:

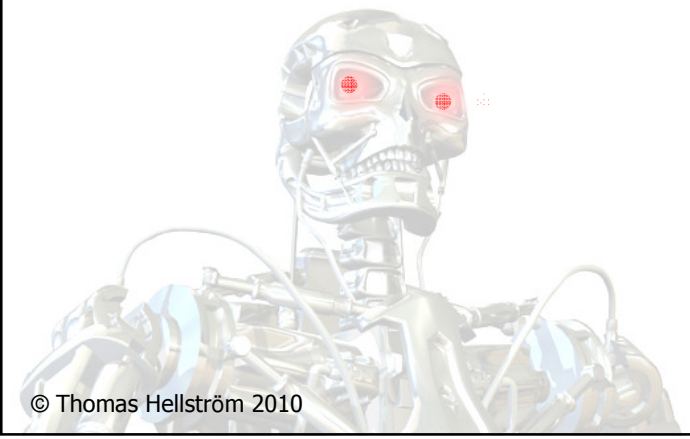
- reinforces behaviors considered to be good
- discourages what is considered to be bad

Moral responsiblity of robots

Two parts:

- causal responsibility – Ability to influence the environment (power)
- intention - Acts in pursuit of its own agenda (autonomy)

*A robot's **moral responsibility** is given by its autonomous power.*



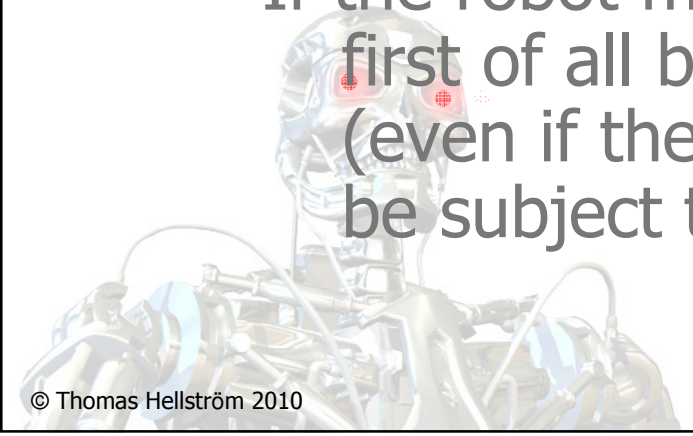
Moral responsiblity

Strange to assign moral responsibility to non-humans?

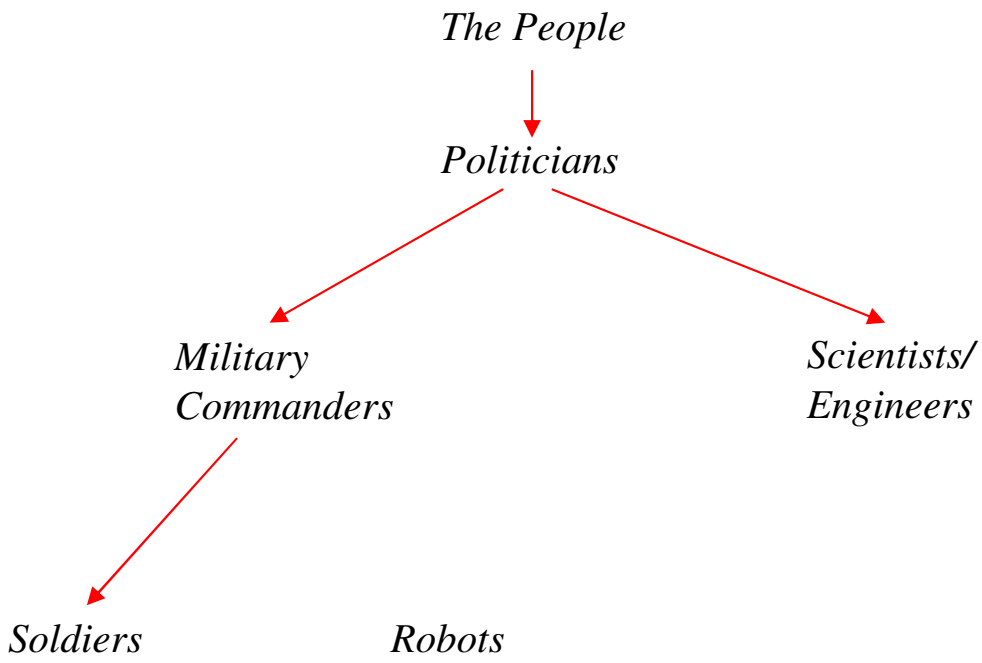
Well, it is already done to some extent with *juristic persons*

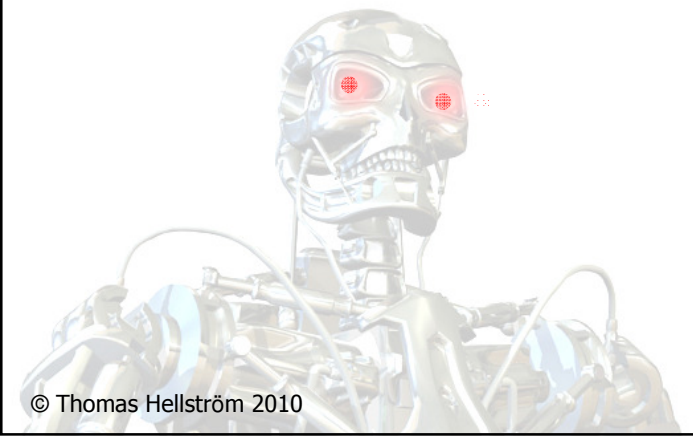
Prediction: the highly autonomous and intelligent robots of the future may be regarded as *artificial persons*.

If the robot misbehaves, correcting actions will first of all be directed towards the robot itself, (even if the owners and constructors also may be subject to criticism)



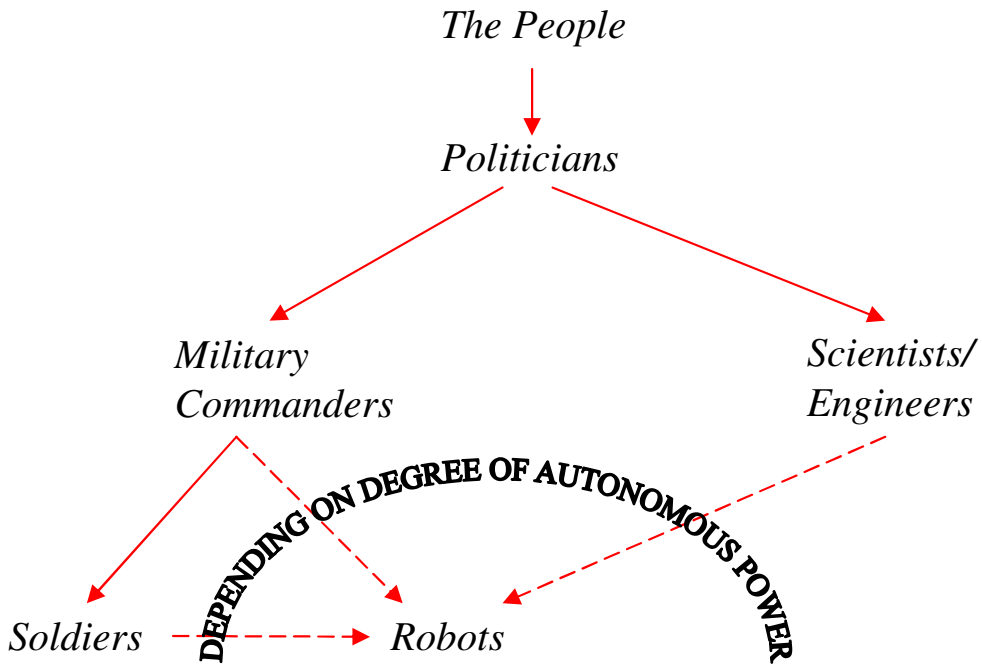
Assignment of moral responsibility





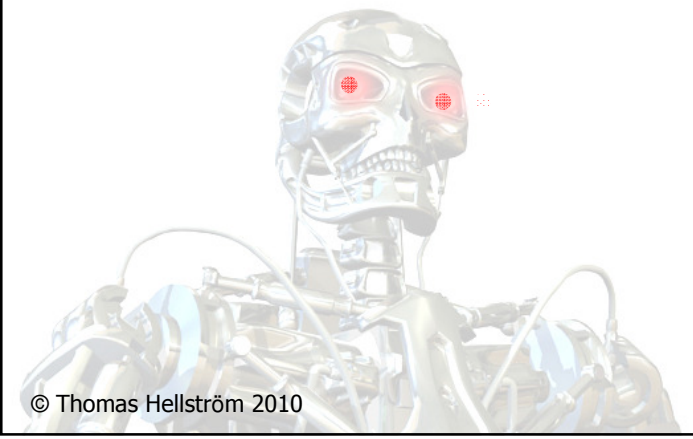
- Moral responsibility = autonomous power = autonomy & power
- Assignment of power to robots, but no autonomy (remote controlled)

Assignment of moral responsibility



- Moral responsibility = autonomous power = autonomy & power
- Assignment of moral responsibility depends on the receiving part's autonomous power:
  - Samsung SGR-1 is more responsible than a TALON robot

Ethical considerations

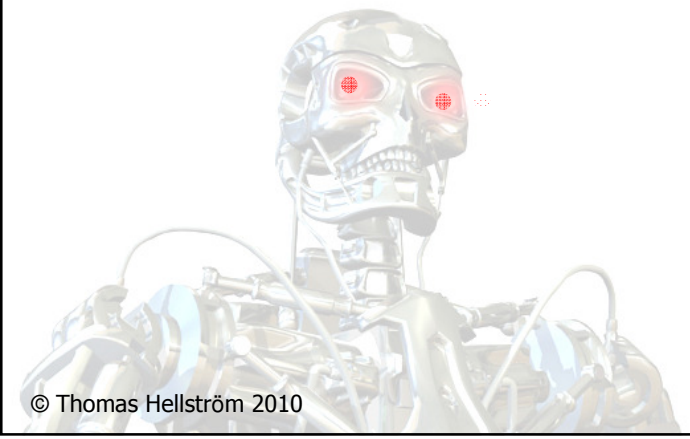


Ethical considerations

What is a good robot soldier?

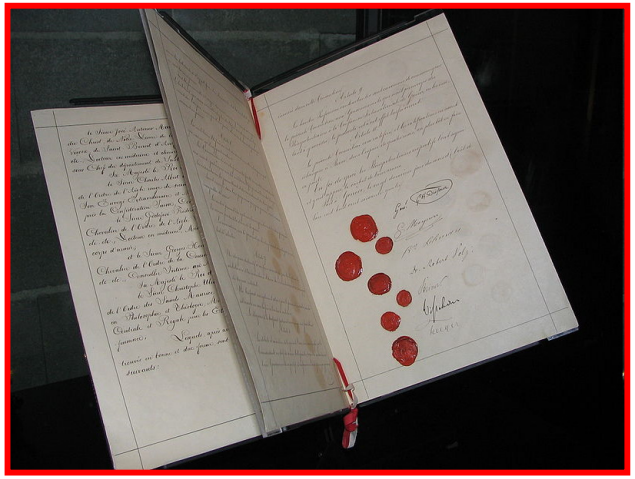
- Does the same as a human soldier would do?
- Does the same as a human soldier **should** do? (ethical ideal)
- ~~Is efficient (wins the war)~~

So: What is a good human soldier?



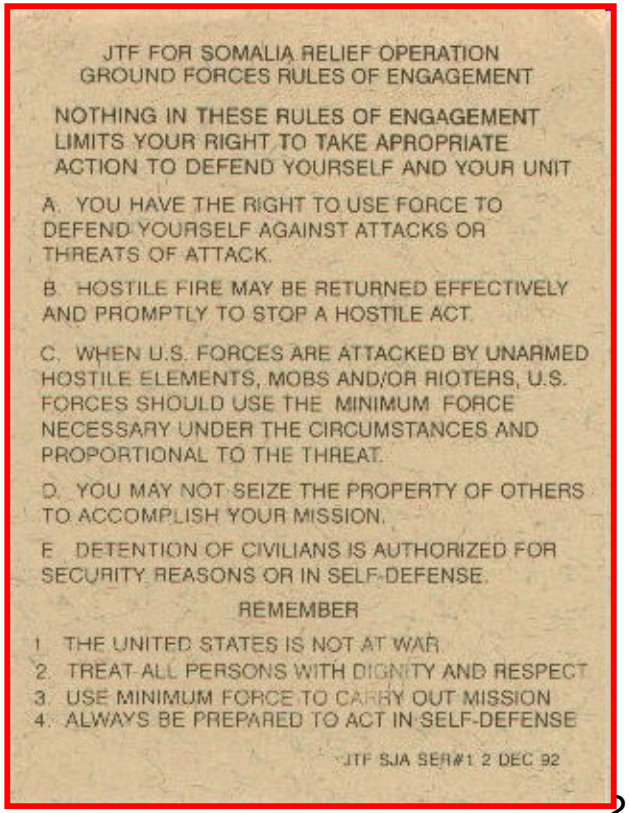
Ethical considerations

What is a good human soldier?



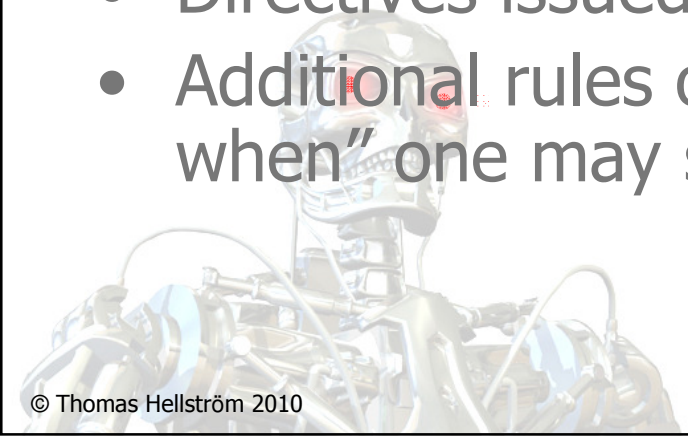
Laws of War (LOW)

- Govern behavior in war
- Regulates the conduct of armed hostilities
- Encoded in *Geneva Conventions, Hague Conventions*, …



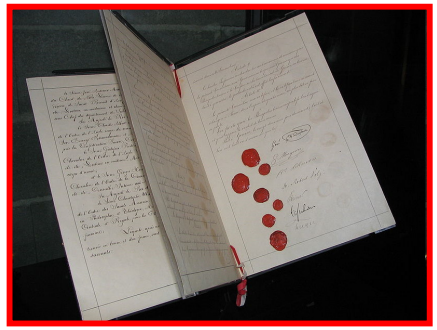
Rules of Engagement (ROE)

- Directives issued by military authority
- Additional rules on ”what, where and when” one may shoot



Ethical considerations

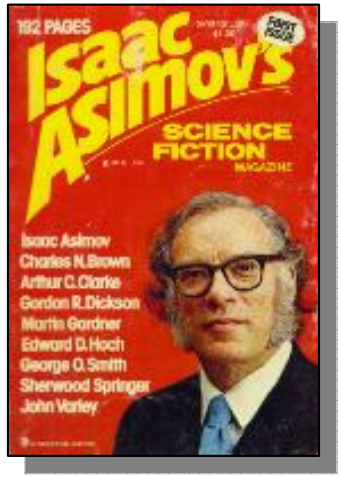
Two corner stones of Laws of War



- Discrimination
  - Only combatants are legitimate targets of attack
- Proportionality
  - The unintended harm caused to civilians must be proportional to the military advantage anticipated by an attack

LOW and ROE define human ethical behavior in war and could also be used to make military robots ethical

Asimov's three robot laws



1. A robot may not injure a human being, or, through inaction, allow a human being to come to harm
2. A robot must obey the orders given it by human beings except where such orders would conflict with the First Law
3. A robot must protect its own existence as long as such protection does not conflict with the First or Second Law

A zeroth law was added later:

0. A robot may not injure humanity, or, through inaction, allow humanity to come to harm

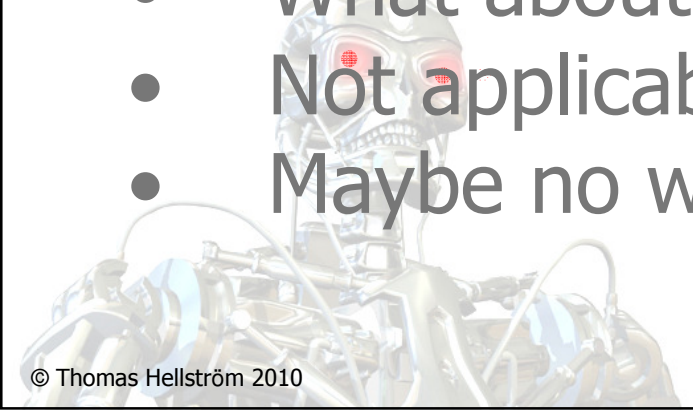
The rest of the laws are modified to acknowledge this.

Asimov's three robot laws

Problems:

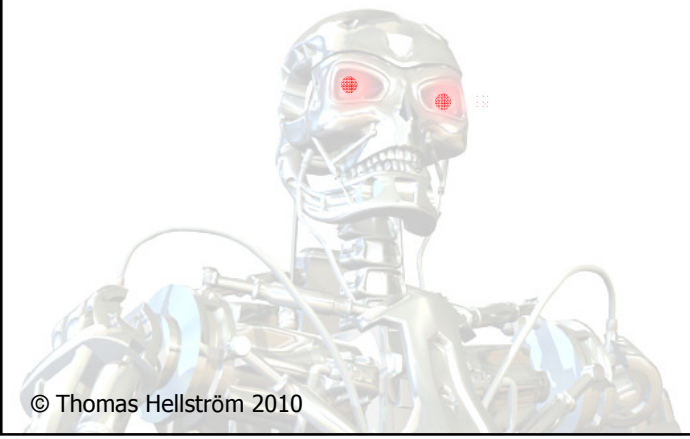
- Ranking:
  1. Humanity
  2. Human individuals
  3. The robot itself

- What if the robot would have to injure a human being to protect another?
- Rely on other, not specified, ethical considerations
- What about protecting borders or property?
- Not applicable to military robots?
- Maybe no worse/better than human laws…



Why would we want robot soldiers?

- Save human lives and health (on ”our” side)
  - A robot may fight and die instead of a human soldier
  - A robot may protect human soldiers
- Potentially more ethical warfare? ➔
  - Shortcomings with human soldiers
  - Advantages with robot soldiers



Why would we want robot soldiers?

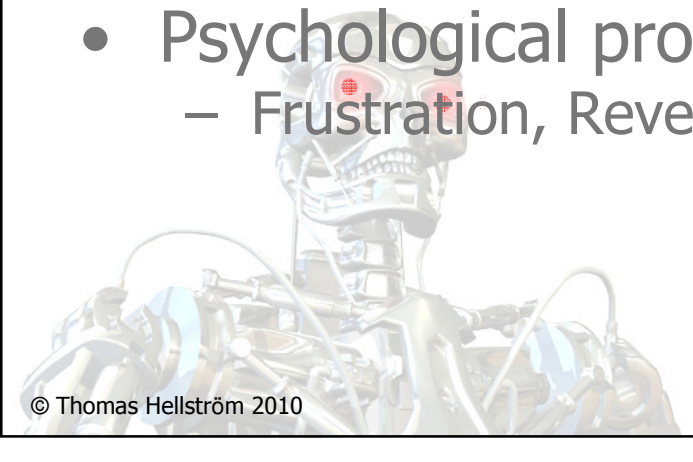
Shortcomings with human soldiers [3]

U.S soldiers after the Iraq war (2003):

- 10% of all soldiers report mistreating non-combatants
- 1/3 think that torture sometimes should be allowed
- ~30% report facing ethical situations in which they don't know how to respond

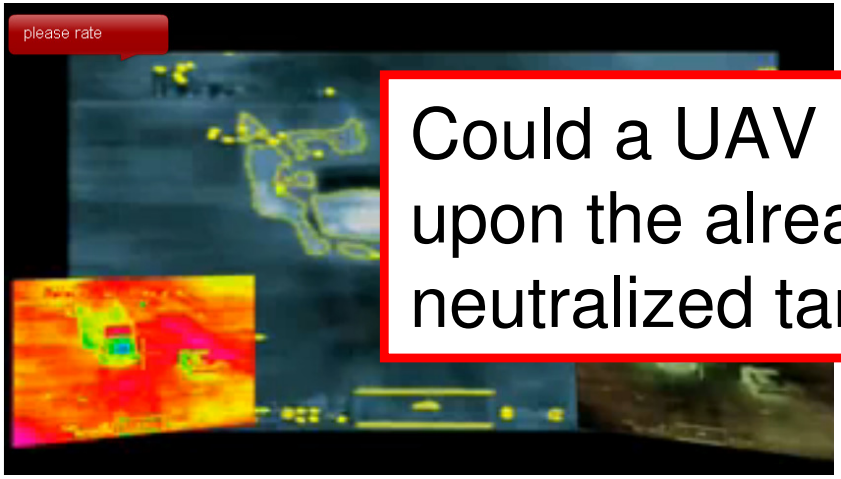
Atrocities often occur

- Fear of getting killed leads to non ethical behavior:
  - Shoot first and frequently
  - When possible: stay far away from the enemy
- Psychological problems:
  - Frustration, Revenge, "Scenario fulfilment"



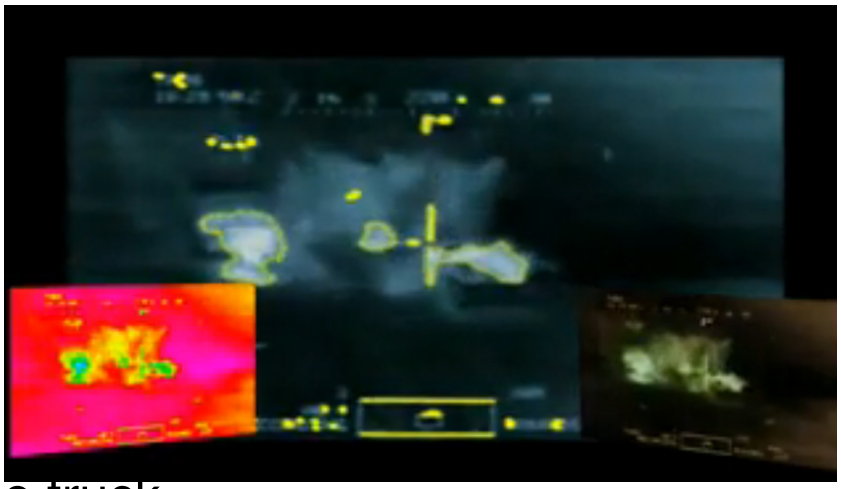
Why would we want robot soldiers?

Shortcomings with human soldiers (Arkin 2009)



Could a UAV have refused to shoot upon the already wounded and neutralized target?

Pilot: Want me to take the other truck out?



Pilot: Roger

Commander: Roger. .. Wait for move by the truck.
Pilot: Movement right there. …
Commander: [No hesitation] Hit him.
Pilot: Targeting the Truck.
Commander: Hit the truck and him. Go forward of it and hit him.

”Apache Rules the night”

http://www.youtube.com/watch?v=3BfJlnHUdac&feature=fvsr

Why would we want robot soldiers?

Advantages with robot soldiers

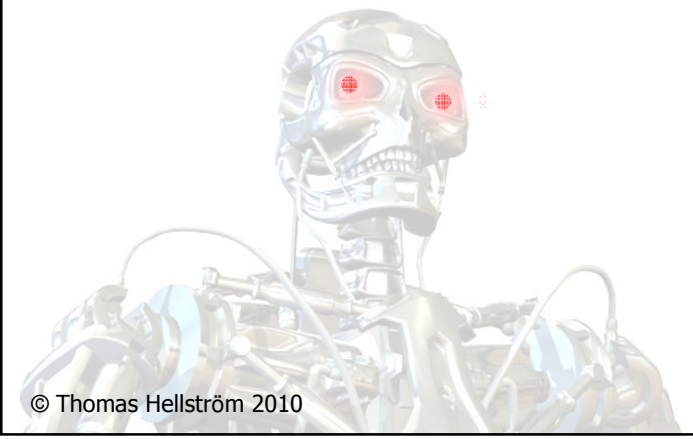
- A Robot soldier can be programmed to
  - Refuse an order it deems unethical
  - Monitor ethical behavior in a human/robot team
- No psychological problems
- No fear and no need for self-preservation
  - No need to shoot first
  - No need to stay far away: A UAV can circle low over a potential target for hours before striking
- Better sensing
  - Faster, longer range, higher precision, greater persistence, longer endurance, sensor fusion

Robots can become more ethical than human soldiers

Ronald C. Arkin
Georgia Institute of Technology

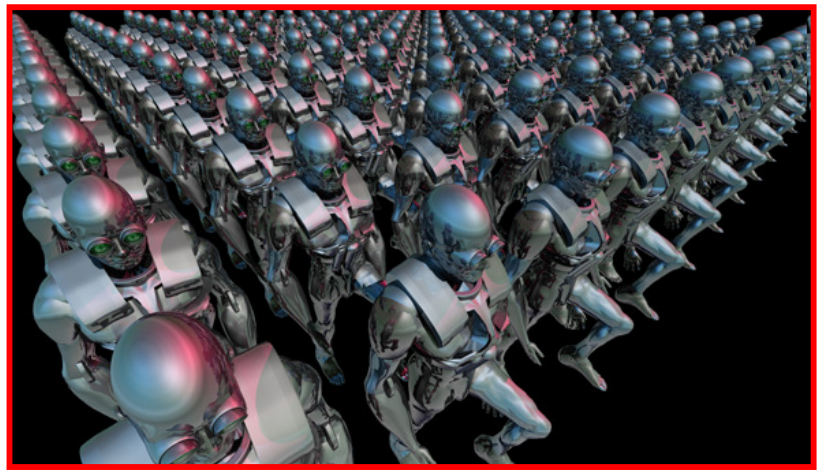
Why would we want robot soldiers?

- If robot soldiers behave better than human soldiers: Is it our **responsibility** to use robots?
- Human Rights Watch: ”Only precision guided bombs should be used in civilian areas”

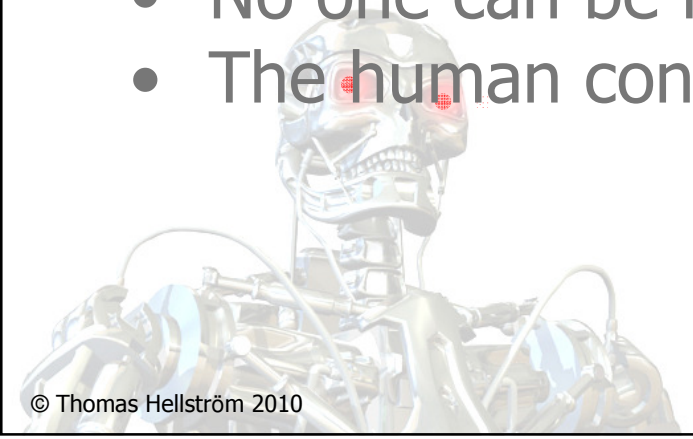


Why would we NOT want robot soldiers?

- We are against war in general
- We are against high-tech wars in general
- Scary scenario with armies of walking robots



- It is impossible to make them follow the LOW ➔
- ”Numbed killing” ➔
- The threshold for war is lowered (Asaro 2007)
  - the temptations of a ”risk-free war”
- No one can be held responsible for war crimes
- The human contact is important in war (Armstrong 2008)



Why would we NOT want robot soldiers?

A robot can never distinguish between a civilian and a combatant

"It is technically impossible to make robots that adhere to the Laws of War!"

Noel Sharkey
University of Sheffield

A robot is bound to make mistakes and hurt innocent people!

We should therefore

Forbid armed autonomous robots

OR

Require that the robots follow the Laws of war without ANY mistakes

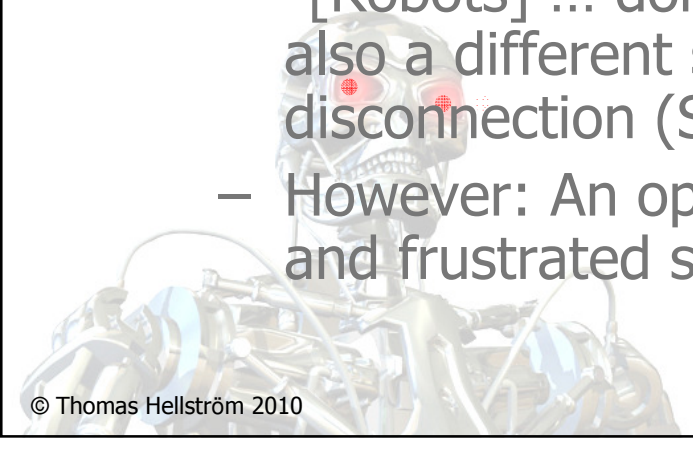
Why would we NOT want robot soldiers?

Numbed killing

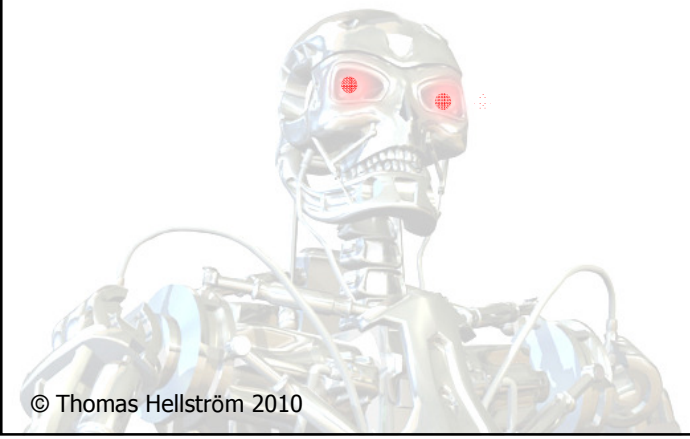
"as soldiers are removed from the horrors of war and see the enemy not as humans but as blips on a screen, there is a very real danger of losing the deterrent that such horrors provide." (Singer 2009)

<u>This is an argument against tele-operated lethal robots</u>.

- UAV operators get numb when dropping bombs from almost halfway around the world
- But how do robots differ from many other weapons like bomb planes, rifles, crossbows, …?
  - "[Robots] … don't just create greater physical distance, but also a different sort of psychological distance and disconnection (Singer 2009)
  - However: An operator can act more rationally than a scared and frustrated soldier in field



Technical challenges



Technical challenges

Discrimination

- What is the difference between a civilian and combatant?
- A general solution requires breakthroughs in sensing and perception!

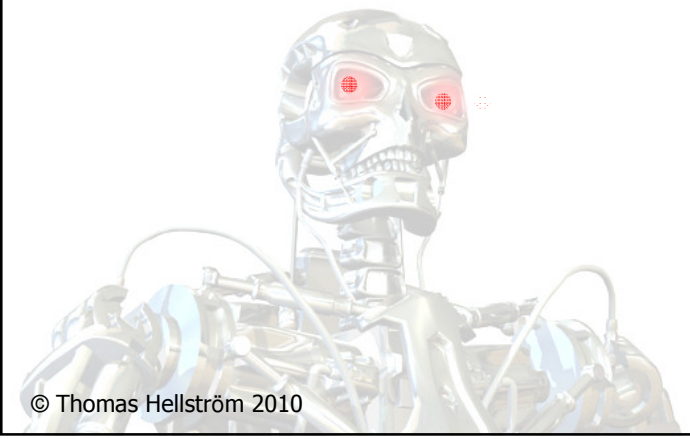
Automatically identified foe

Two recognizable surrender gestures

Technical challenges

Proportionality

- Comparing "Apples and oranges" (Andersson, K. 2007)
  - How to compare enemy civilan deaths with achieved military gains?
  - One technique to deal with such problems: CBR (Case Based Reasoning) using examples from the past



Technical challenges

Ethics software for robots

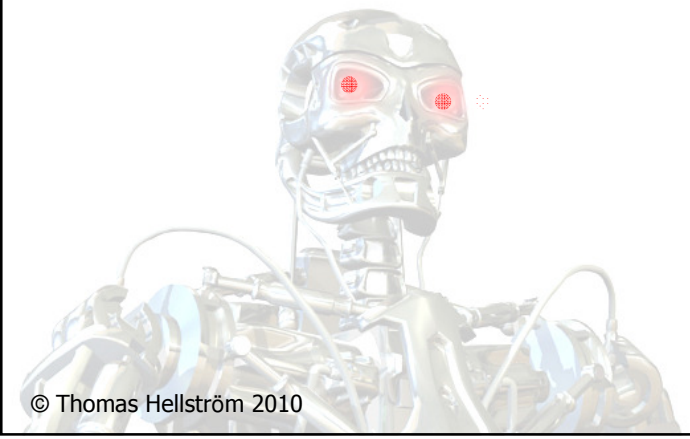
- Hard to formalize ethics
- The ethical rules are almost always expressed at a highly abstract level and are subject to interpretation
  - ”engagement is authorized when the hostile force continues to commit hostile acts or exhibit hostile intent”
  - ”when time and circumstances permit …”
- Even if a rule is known to apply, the prescribed actions are often abstract as well
  - ”… an obligation to take feasable measures to remove civilians from areas containing military objectives”
- The rules often conflict with each other

Technical challenges

An intermediate (ethical?) strategy: Keeping the ”human in the loop”

- The human does the actual shooting (e.g. Predator, MAARS)
- The human guarante adherence to LOW and ROE (e.g. Phalanx, the Korean SGR-1)
- This is a common approach in all automation
  - Makes the design easier
  - Makes the operation safer
  - Simplifies legal matters

The researchers' role



The researchers' role

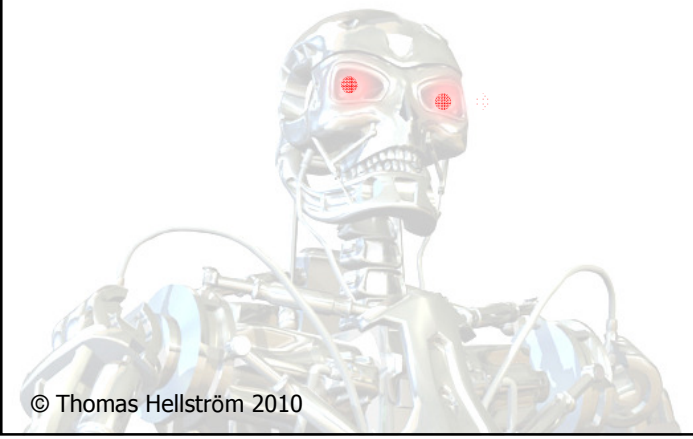
- We supply the technical solutions
- Doesn't matter if we collaborate with weapon industry or not
  - If our robotics research is of significance it will be put to use in military systems sooner or later
  - Does this help:

"… it is strictly prohibited to use or to develop, in a direct or indirect way, any of our scientific contributions by any army or armed group in the world, for military purposes and for any other use which is against human rights or the environment."

- Even if we shut down all research:
  The killer robots are already here!

The researchers' role

- Is it our responsiblity to
  - warn/inform the public and the politicians?
  - develop as good military robots as possible?
  - contribute to new legislation and conventions: *A Terminator Ethics* ?
- Not necessarily surrendering to an evil reality
  - Maybe there are advantages with military robots

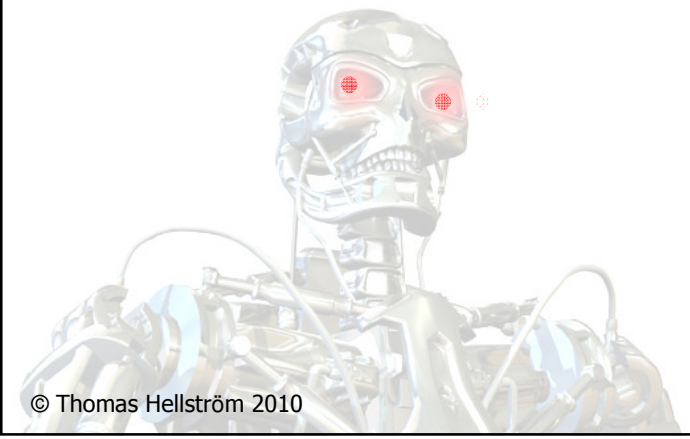


Concensus

- We are entering a new era in warfare
  - Lethal robots are here to stay
  - LOTS of unsolved technical problems
- As always: technology evolves more quickly than the laws of war
- No current international guidelines to the use of autonomous robots in warfare
- On going work on guidelines
  - Technical Committee on Roboethics (2004-)
  - Committee on Human Rights and Ethics for the IEEE Robotics and Automation Society (2006-9)
  - Robot Ethics Charter from the South Korean government

My conclusions

- War is already hell and atrocities frequently occur due to scared and/or flipped soldiers
- Some types of semi-autonomous armed robots may lead to more humane wars
  - UAVs
  - Small mobile armed robots, for instance used to secure buildings
- The crucial question is if and when we should allow the robot to pull the trigger (full autonomy)

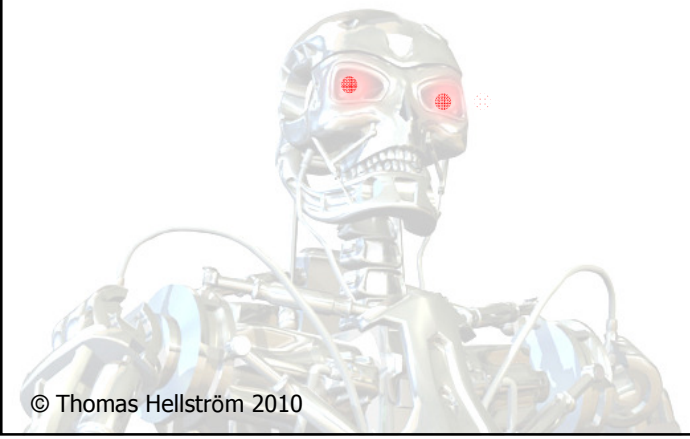


References

[1] Arkin, R. C., Governing Lethal Behavior in Autonomous Robots, Chapman&Hall/CRC., 2009.

[2] Sharkey, N.E., Grounds for Discrimination: Autonomous Robot Weapons, RUSI Defence Systems, 11 (2), 86-89, 2008.

[3] Singer, P. W., Wired for War - The Robotics Revolution and 21st Century Conflict, Penguin, 2009.



Thoughtful cases

- Intelligent Anti-tank mines
  - How about equipping them with ethical behaviors such that they jump to the side when a human approaches?
- An autonomous vehicle for transportation of PET bottles to the swedish soldiers in Afghanistan
  - Partly the same techical solutions as lethal robots