

Human Rights for Machines



The Last Revolution

Human Rights for Machines

- Example of IT getting uncontrollable
- Problematic in several areas in society
- Why is this unavoidable?
- Possible solution

Technology related, 1987-2011

- Build software since 1987
- Write and publish about technology since 1992, Future AI since 2000.
- Publish Venture Capital Guide (*Venture Capital is a selection mechanism for new technology*)
- Reconstructed Jan Sloot compression technology (patent granted)
- Built: Software, like word processor, browser.

May 6, 2010: The Flash Crash



- . \$ 700 billion dollar disappears into thin air
- . 1 share Procter & Gamble: 1 cent.

What was this about?

- One contract E-mini futures on the S & P-index
- 1. Share: derivative of a company
- 2. Futures: derivative of shares
- 3. E-Mini's: derivative of a future
- 4. Standard & Poors-index: derivative of stock prices

Domino's on the stock markets

- Algo of investment fund in derivatives start automated selling against any price
- Algo's on NYSE stop trading through built in safety mechanism
- Free fall Dow Jones Industrials Average
- Circuit breaker interrupts trading: 5 seconds
- Prices recover to starting point: 5 minutes
- Total duration: 20 – 30 minutes.



Official investigation: 5 months



- May 6 - September, 30, 2010
- “Flash crash caused by a single large order. No change in regulation. Perhaps more circuit breakers.”

Problem solved?

- It can happen again every day.
- Circuit breakers are active on one exchange at a time
- All exchanges are interlinked
- Crash was partly due to safety mechanisms built into trading algo's

Financial weapons of mass destruction

- Warren Buffet: “*Such derivatives are financial weapons of mass destruction*”



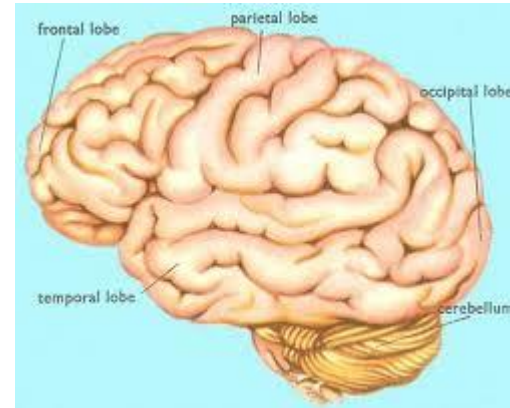
Next: Designed to learn



- Neural networks + Genetic programming
- In future, we will never know what happened

Human AI Parity: when?

- Human brain:
- 18 000 000 MIPS



- 2012: IBM's BlueGene/Q
- 9 000 000 MIPS



- Hardware is almost halfway

Machines are getting uncontrollable I

- Financial world: not **who** but **what** does the trading?
- Military: not **who** but **what** fires a rocket?
- Health care: not **who** but **what** unplugs the cord?

Robot platoon



Machines are getting uncontrollable II

- Disease control: not **who** but **what** confirms epidemics?
- Search engines: not **who** but **what** is giving me information?
- Social networks: not **who** but **what** is giving me advice?

No going back

- **Companies** will need smarter algorithms to remain competitive.
- **Citizens** will ask for smarter solutions for mobility, health care robots, a.s.o.
- **Society** needs higher productivity, less KIA's

- **Problem:** We can never control learning machines

Solution: Machine rights

- Smart machines will be as good as the rights that they have
- Learn, inquire and collect information
- Contact people
- Decide
- Spend money, etcetera.
- Machine rights can be tied to obligations

Pandora's box

- A machine's duty to perform will require its right to learn and develop ...
- ... which will require the right to spend and upgrade ...
- ... which will require an income, a bank account, insurance, bankruptcy protection, etc. ...
- ... which will require the right to privacy ... etc. etc.

Responsibility?



- The builder cannot anticipate the outcome of a learning machine's behavior

Solution?

- Disconnect builder from malicious algorithm
- Compensate builder
- Create technological entity responsible to independent human oversight authority

Larry Page, co-founder of Google



- *“If you had all the world’s information directly attached to your brain, or an artificial brain that was smarter than your brain, you’d be better off.”*

Human Rights for Machines



Our greatest challenge