Autonomous and Pervasive Technology

Professional Ethics in Computing

Lecture 11

(Largely based on lecture by J. Cartlidge)

Module Plan & Progress

Teaching Week	Date	*Contents*
2	*26 Sep*	*Lecture 1: Introduction and Administration*
4	*8 Oct*	Lecture 2: Critical Reasoning and Moral Theory 1
5	*15 Oct*	Lecture 3: Critical Reasoning and Moral Theory 2
6	*22 Oct*	Lecture 4: Computing Professionals and Professional Ethics
7	*29 Oct*	Lecture 5: Privacy
8	*5 Nov*	Lecture 6: Intellectual and Intangible Property
9	*12 Nov*	Lecture 7: Critical Thinking
10	*19 Nov*	Lecture 8: Trust, Safety and Reliability
11	*26 Nov*	Lecture 9: How Computing is Changing Who We Are
12	*3 Dec*	Lecture 10: Computing and Vulnerable Groups
13	*10 Dec*	Lecture 11: Autonomous and Pervasive Technologies

1.	Theory	DONE
2.	Professional codes of conduct	DONE
3.	Real World issues	To do

Coursework 2 Deadline: **23:59** on **12/12/2018** (submitted by Moodle)

Notes: Group 1

Next week – talk about exam / revision [useful!]

Autonomous & Pervasive

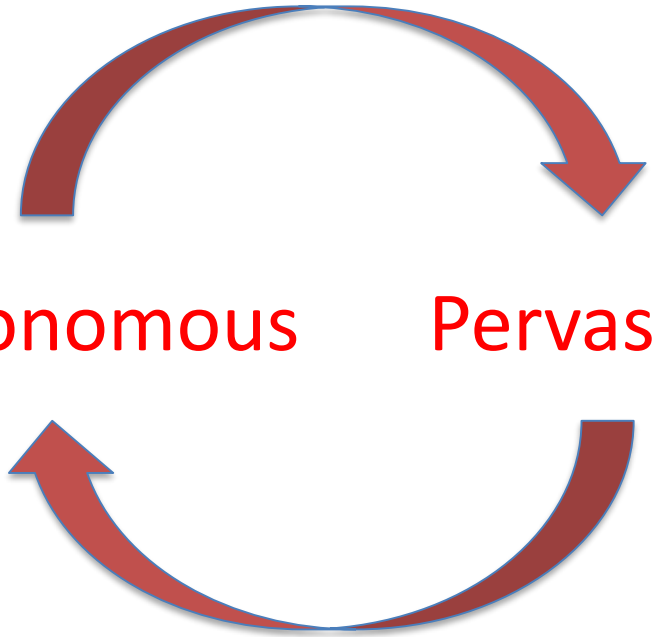
- Modern technologies force us to make serious and complicated ethical decisions
- Two qualities that are especially problematic:
  1. **Pervasive**: a technology that has spread widely through society
  2. **Autonomous:** freedom to make decisions without outside constraints or interference
- We also consider two related concepts:
  - **Emergence**—when software behaves in a way that was not directly programmed
  - **Moral responsibility** for autonomous technology

Today's Topic

- Most ethically difficult
- Most interesting(?)
- Most likely to play an increasingly important role during your (professional) lifetimes!

Easier to use, cheaper to operate and maintain

People use more often and in new situations



Make decisions without direct human control or supervision

Autonomous

Pervasive

Used widely through all society

As tech becomes more used, financial incentives to automate it increase

CASE: CAMERA-PHONES

Digital Cameras & Cell-phones

- **Qu: How many cameras do you use everyday?**
  - Nintendo 3DS, three cameras
  - Smart phone, two cameras
  - Laptop, one integrated webcam
  - High-quality digital camera, for holiday snaps
- **Pervasiveness** of cell phones, people have cameras on them at all times
  - International Communications Union (2013)
  - 96% world mobile-cellular penetration (Africa 63%)

http://www.itu.int/en/ITU-D/Statistics/Documents/facts/ICTFactsFigures2013-e.pdf

Before: Cameras & Film

Camera film

Disposable camera

Analogue (film) camera

Digital Photographs: Autonomy

- Past: film cameras—film developed in a lab
  - Film/development expensive, so photo taking not wasted
  - Workers in lab can see/steal/copy photos
  - People afraid to take obscene or damaging photographs
- Modern cameras—these issues no longer relevant
  - Side effect: child predators taking inappropriate pictures
  - "Camera Phone Predator Alert Act" King's bill (New York); requires camera phone to sound tone when picture taken

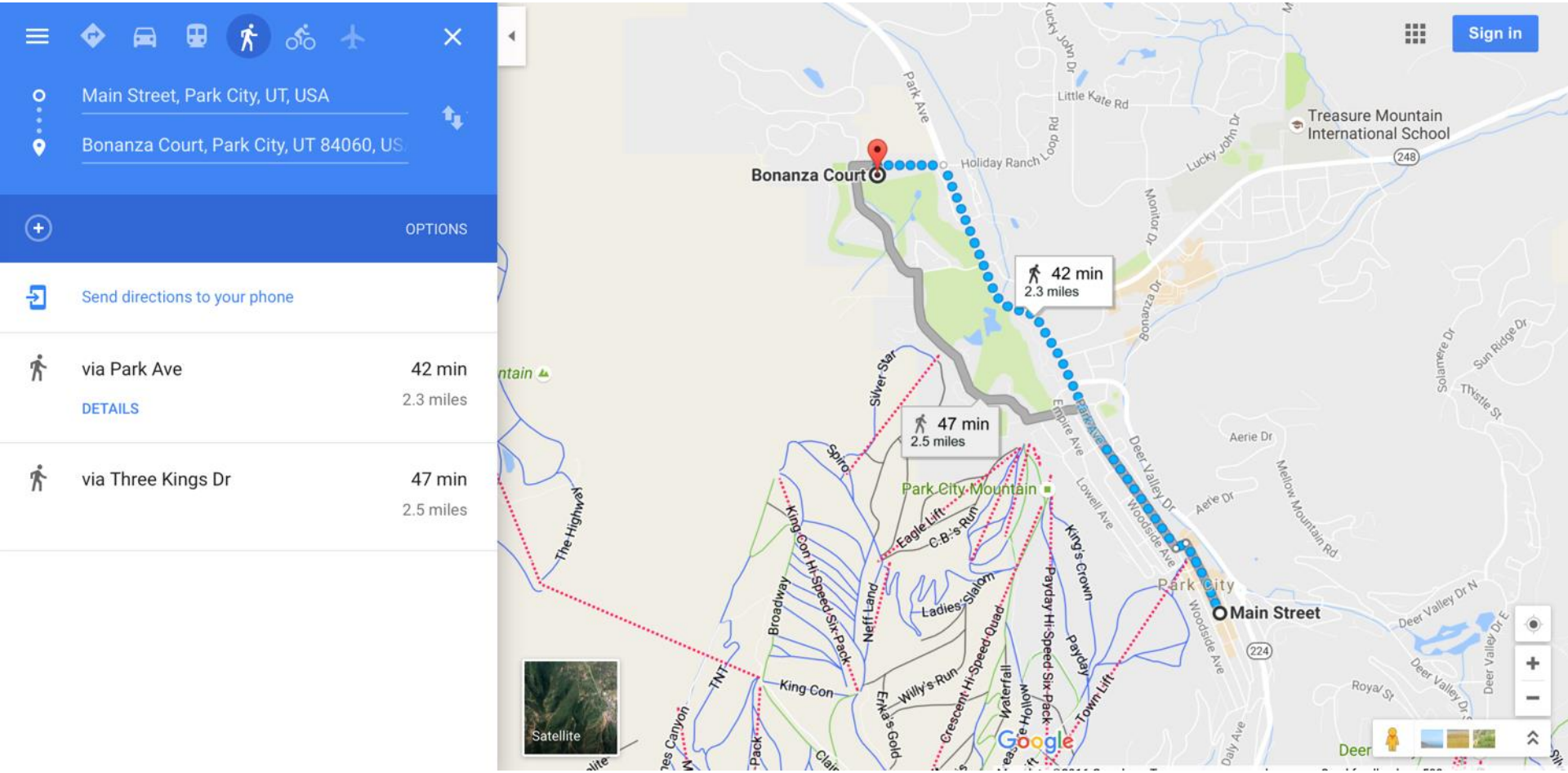
https://www.congress.gov/bill/111th-congress/house-bill/414

CASE: INJURED BY GPS

Case: Injured by GPS

- Jan 2010, Lauren Rosenberg using Google for walking directions
- Rosenberg walked along Utah State Route 224
  - No sidewalks, and high speed limits
  - Rosenberg hit and injured by car
  - Rosenberg sued Google and car driver for negligence

Injured by GPS



Qu: Who is most to blame?

- Ms Rosenberg?
- Google?
- The Driver?

Utah State Route 224

Qu: Would you walk a road like this?

Qu: How about at night?

Trust

- Remember week 8?
- Humans have a tendency to trust computers such as GPS to do certain kinds of tasks
- Computer generated directions often superior to human-generated
  - But, we should still use *common sense*!

Court Ruling: Rosenberg's Fault

- Google didn't have a direct legal relationship with Rosenberg
- Google's mapping services offer considerable value to the public and that allowing the litigation to go forward could open the door to "*nearly unlimited liability*" for Google.

Qu: Which moral theory do the courts appear to be applying here?

http://searchengineland.com/court-dismisses-google-walking-directions-lawsuit-claims-82312

http://www.leagle.com/decision/In%20FDCO%2020110617C10/ROSENBERG%20v.%20HARWOOD?

CASE: MORE GPS, MORE RISKS

The *"Yuppie 911 Device"*

- Block, Melissa, "With new GPS devices, hikers taking more risks", *All Things Considered*, 26 Oct 2009 [Radio Broadcast]
- http://moodle.nottingham.ac.uk/mod/resource/view.php?id=2344897
- SARSAT (Search & Rescue Satellite Aided Tracking System)
- Cost around $250; popular with hikers/skiers
- People using SARSAT for trivial reasons
  - to avoid bad weather/get home quickly
- Coast guard search and rescue costs $15,000

Qu: Has pervasiveness of technology resulted in riskier behaviour?

Qu: Is it morally permissible to charge people to be rescued when they really are in danger?

DEFINING: “AUTONOMOUS”

	Autonomous	Intelligent	Robotic
Roomba, vacuum cleaner sold by iRobot			
Bomb-defusing robots			
Stock-trading software using in algorithmic trading			
TD-Gammon algorithm for playing backgammon			
A marble rolling down a hill			

	Autonomous	Intelligent	Robotic
Roomba, vacuum cleaner sold by iRobot	Yes. Traverses the room automatically, without human supervision	No. The robot simply wanders, and turns at obstacles. No learning or reasoning	Yes
Bomb-defusing robots	No. Remotely operated by human	No	Yes
Stock-trading software using in algorithmic trading	Yes. Automatically chooses what shares to buy/sell (much faster than humans)	Varies. Some follow simple rules, others learn from experience	No. Purely software. No physical/mechanical parts
TD-Gammon algorithm for playing backgammon	Yes. Learned to how play backgammon by playing itself. No human supervision needed	Yes. Invented a more conservative play style that has changed human strategy	No. Purely software. No physical/mechanical parts
A marble rolling down a hill	No. It is unsupervised, but not making decisions	No	No

- The synergy between autonomy and pervasiveness

CASE: GOOGLE VS. YAHOO!

Google vs. Yahoo!

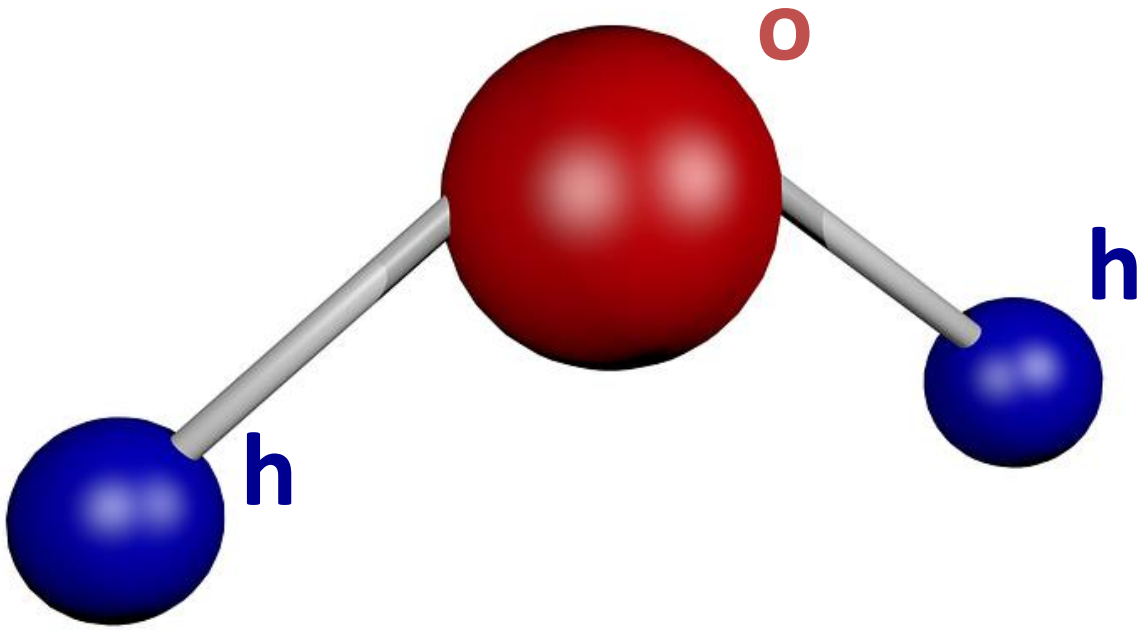
- 1991 Tim Berners-Lee launches WWW
  - Soon WWW too big to find what you want
  - Solution: Web-crawlers using keyword matches (too simplistic)
- 1994 Yahoo! Launched by 2 Stanford PhD students
  - Have humans evaluate quality of every Web Page
- 1996 Google launched by 2 Stanford PhD students
  - PageRank algorithm: automate process using Web links (workshop 6)
  - Automation helped Google scale & beat Yahoo!
  - Google became biggest company on the planet
  - Now "to google" is listed as a verb in *Oxford English Dictionary*
- **Pervasiveness** of Web necessitated autonomy of search
- **Autonomy** of search led to pervasiveness of Google

AUTONOMOUS SYSTEMS: WHO IS RESPONSIBLE?

Software with Emergent Behaviours

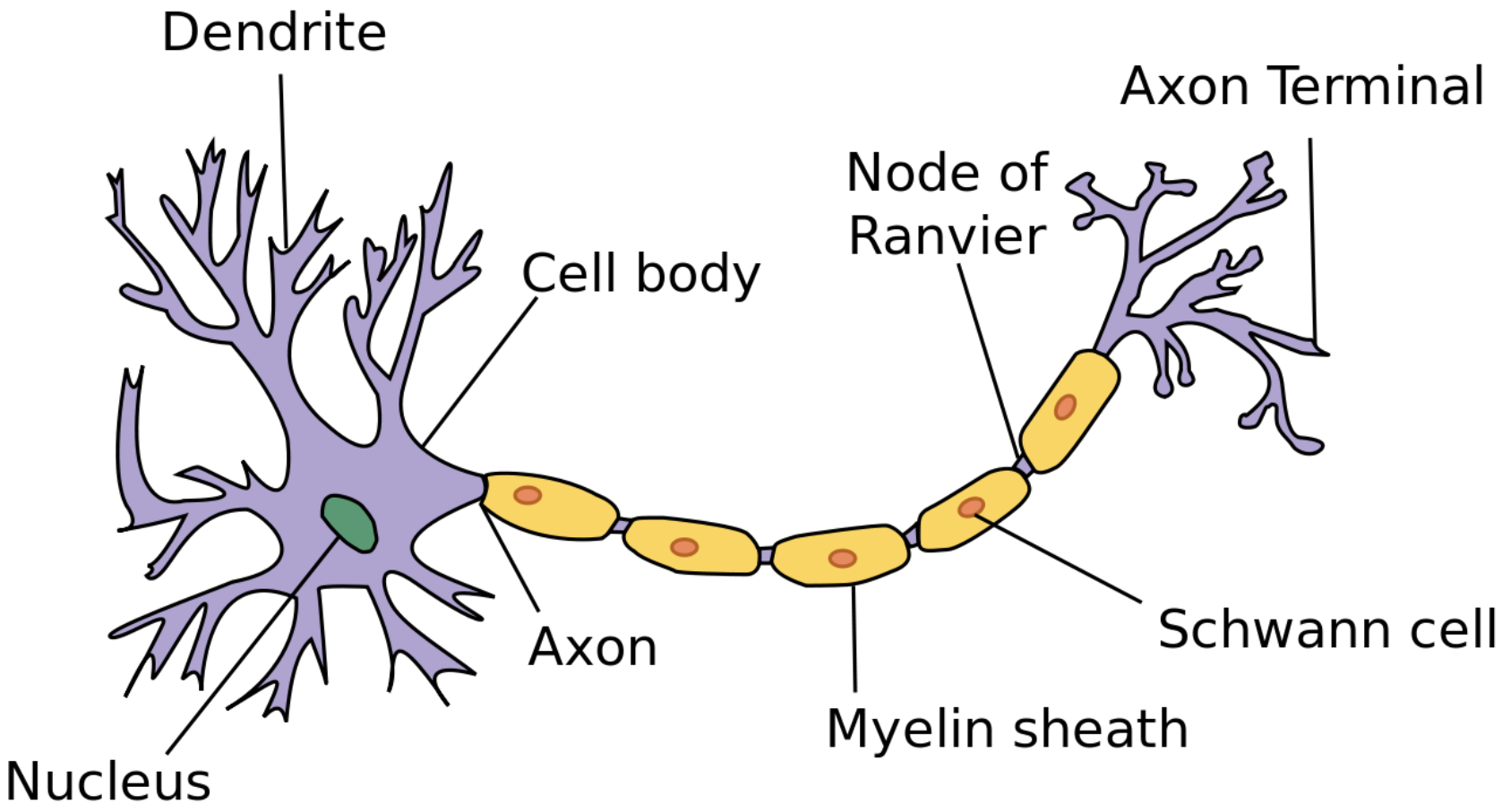
- **Common misconception**—for any piece of software, a programmer wrote and understands the software
- This is not true for machine learning / artificial intelligence applications
  - Often program code is “learned” / “evolving” / “changing” with experience
  - (Socio-technical) Systems of intelligent agents / humans interact with complex emergent behaviours

https://pixabay.com/en/h2o-water-molecule-580992/



Qu: Where's the tsunami?

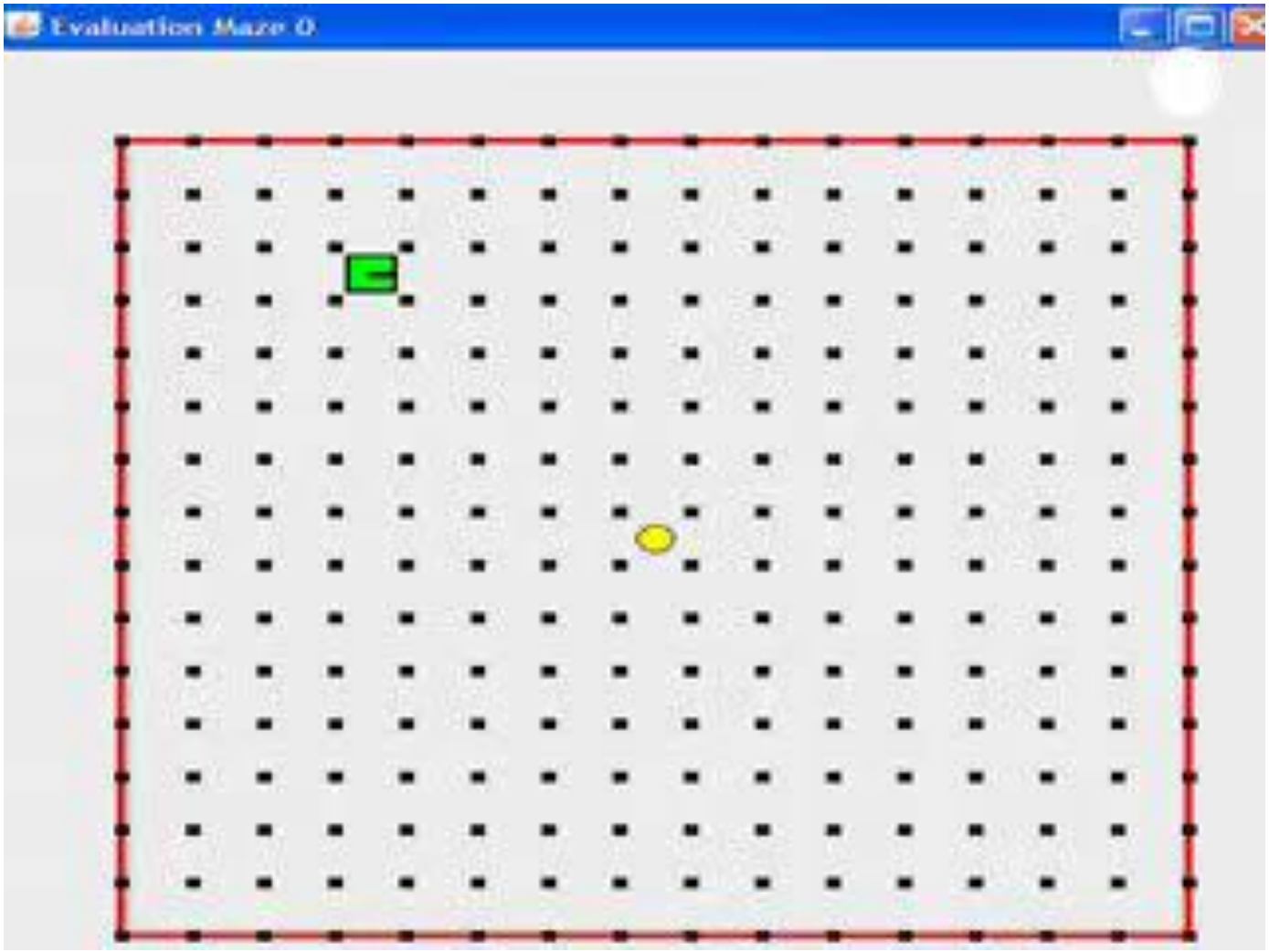
http://shajitheodore.com/wp-content/uploads/2015/09/Neuron.png



Qu: Where's consciousness?

Emergence

Following Rules That *Nobody* Programmed



Cartlidge & Ait-Boudaoud, (2011), "**Autonomous virulence adaptation improves coevolutionary optimization**," IEEE Trans Evolutionary Computation, 15:2, pp. 215–229

Human Readable: Human Understandable?

1. [Fire if: 'right is not blocked' 'behind is not blocked' 'target is right' 'target is not left' 'have not visited forward' 'have visited behind'][Outputs: 'move ahead with prob=0/10' 'move right with prob=5/10' 'move behind with prob=0/10' 'move left with prob=0/10' 'NEXT RULE with prob=5/10']

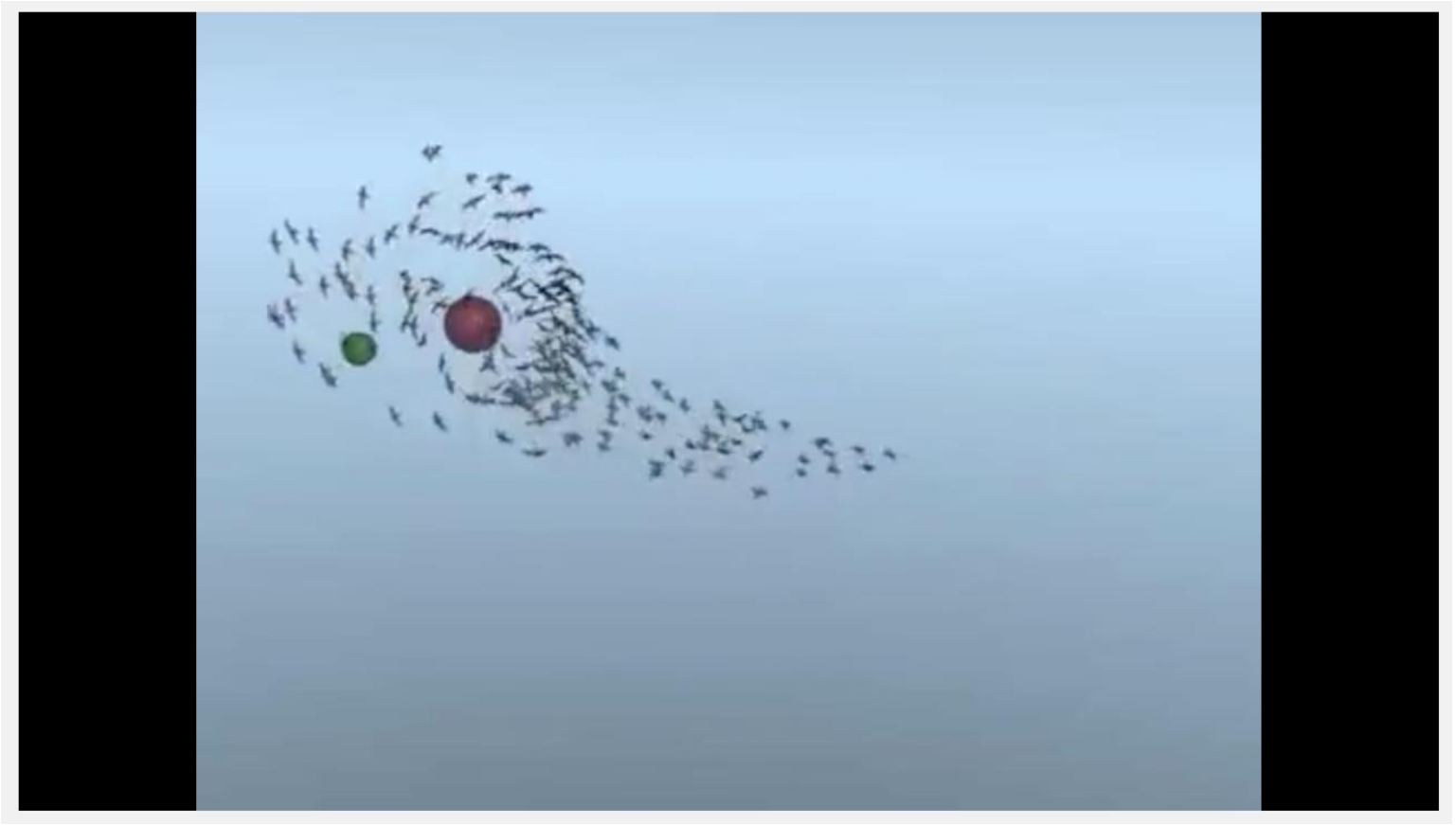
2. [Fire if: 'forward is not blocked' 'target is not behind' 'have not visited left'][Outputs: 'move ahead with prob=3/10' 'move right with prob=0/10' 'move behind with prob=0/10' 'move left with prob=0/10' 'NEXT RULE with prob=7/10']

1. …

Qu: Can you understand how "robot" navigates?

Qu: Would it worry you if a military drone (UAV) navigated this way?

Flocking Boids



https://www.youtube.com/watch?v=GUkjC-69vaw

- **Separation**—Steer to avoid crowding local boids
- **Alignment**—Steer towards average heading of local boids
- **Cohesion**—Steer towards average position (centre of mass) of local boids

- How do we apportion blame for harm caused by complicated computing systems?

MORAL RESPONSIBILITY FOR COMPUTING ARTEFACTS: 5 RULES

Ad Hoc Committee for Responsible Computing

- If a system is so complicated, how do we decide who (or what) is to blame?
- We do not want people to avoid responsibility for **autonomous systems** they have created
- A committee of leading ethicists and CS researchers formulated a set of 5 rules…

Five rules: moral responsibility

1. People who make computing artefacts are **morally responsible for the foreseeable effects** of those artefacts
2. The **amount of blame** assigned to a person **does not depend on the number of people involved** (i.e., we cannot split blame up like we split a cake; each person 100% to blame)

Five rules: moral responsibility

3. The **user** of a computing artefact is **morally responsible** for the consequences of that use

4. People must always make a reasonable effort to **foresee the societal effects** of a technology they design or use

5. People who design computing artefacts must **provide enough information** about them so that others can foresee the consequences of using them

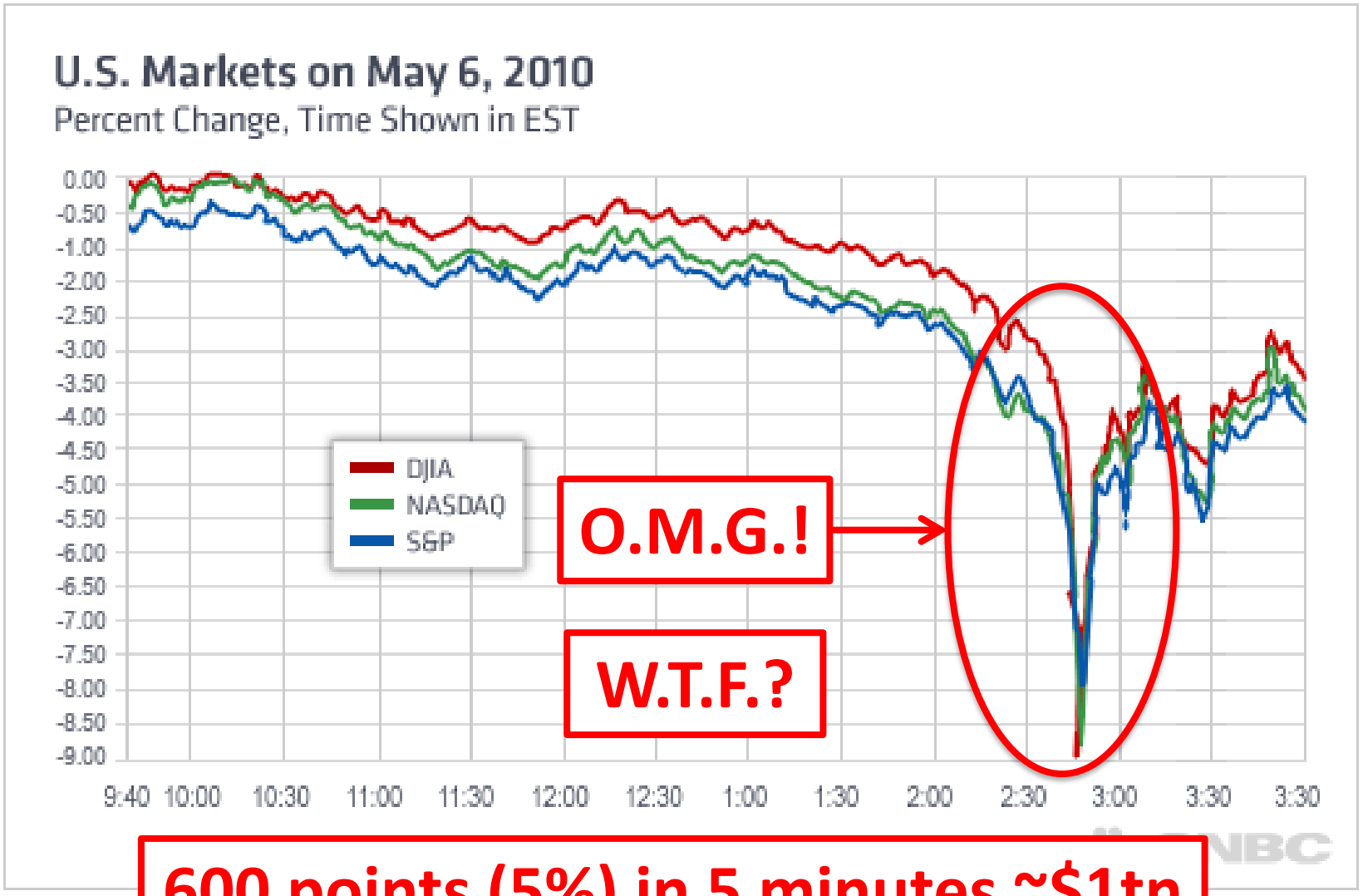
Caveats

- No matter how sophisticated computing artefacts become**, the rules still hold**:
  - E.g., artefact uses a neural net and designers are subsequently surprised by **emergent** effects
  - E.g., artefact **self-modifying** and behaviour changes
  - E.g., person launches artefact A, which is designed to launch artefact B, the **rules apply to both** (i.e., don't launch A if you are not prepared to take moral responsibility for B)
- If you cannot accurately predict what a system will do, then you should not allow it to be used!

Qu: Is this similar/different to Birsch's criteria for moral responsibility?

- Ignore what the textbook says – it is out of date!

CASE: THE FLASH CRASH


U.S. Markets on May 6, 2010
Percent Change, Time Shown in EST
0.00
-0.50
-1.00
-2.00
-2.50
-3.00
-3.50
-4.00
-4.50
-5.00
-5.50
-6.00
-6.50
-7.00
-7.50
-8.00
-8.50
-9.00
9:40
10:00
10:30
11:00
11:30
12:00
12:30
1:00
1:30
2:00
2:30
3:00
3:30
3:30
DJIA
NASDAQ
S&P
O.M.G.!
W.T.F.?
CNBC
600 points (5%) in 5 minutes ~$1tn

Sotheby's worth more than entire Chinese economy!

Accenture: **$40 to 1¢**

Buy a $25bn company for $6million!

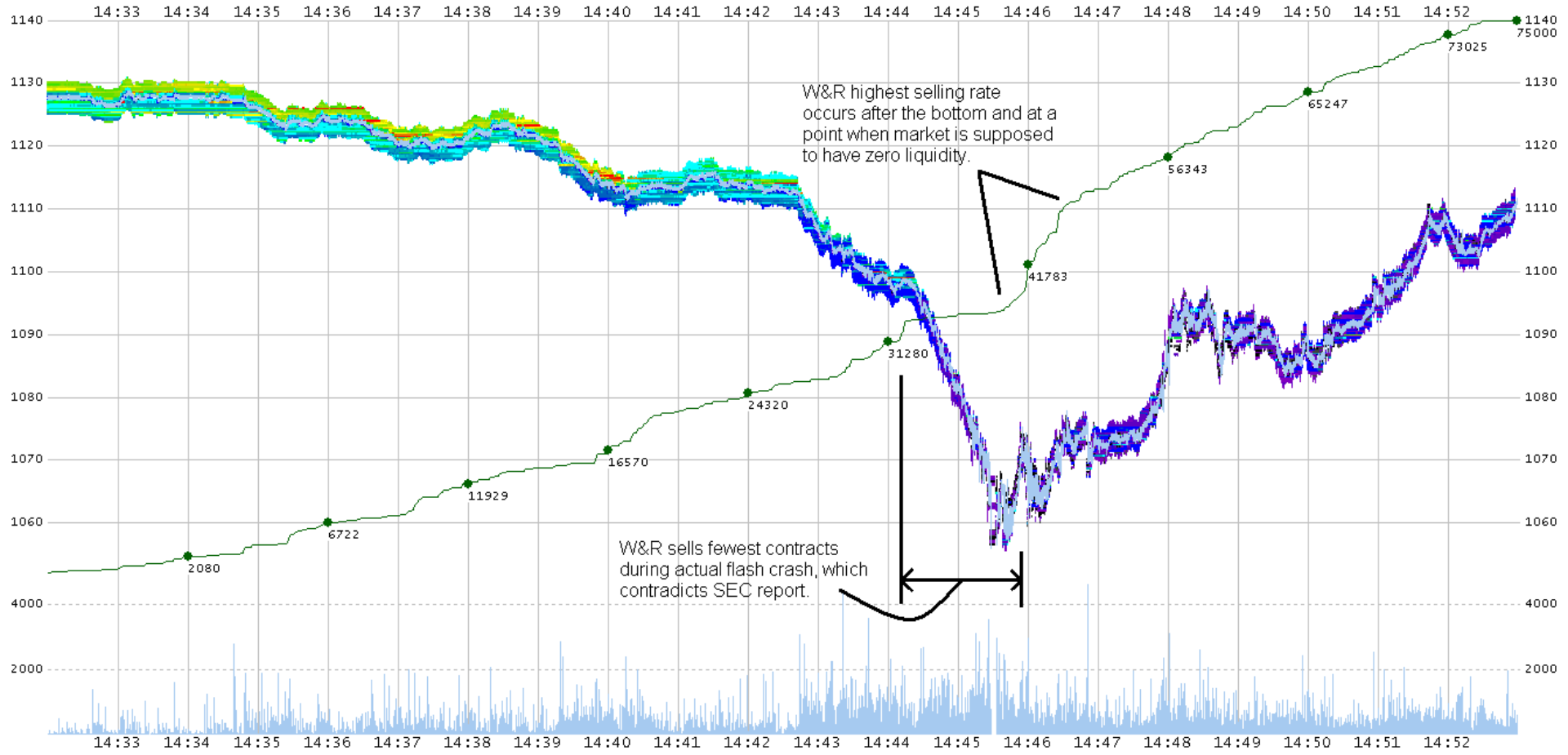
SEC/CFTC Report

- US Securities and Exchange Commission & the Commodities Futures Trading Commission launched and enquiry
- Released on 30th Sep 2010
  - After 5 months of data analysis / investigation
- Identified the cause of this unprecedented and potentially world-shattering event
  - Which was…

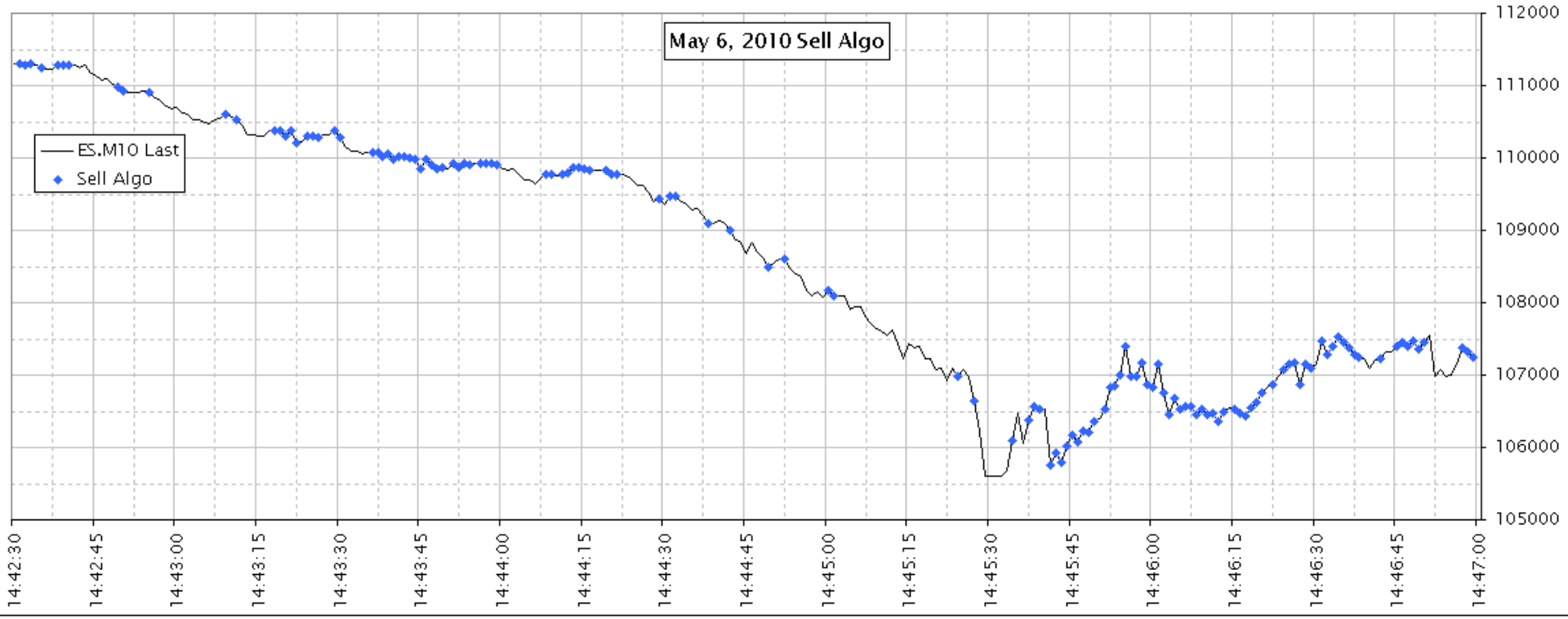
https://pixabay.com/en/fan-foam-hand-characters-finger-1027184/

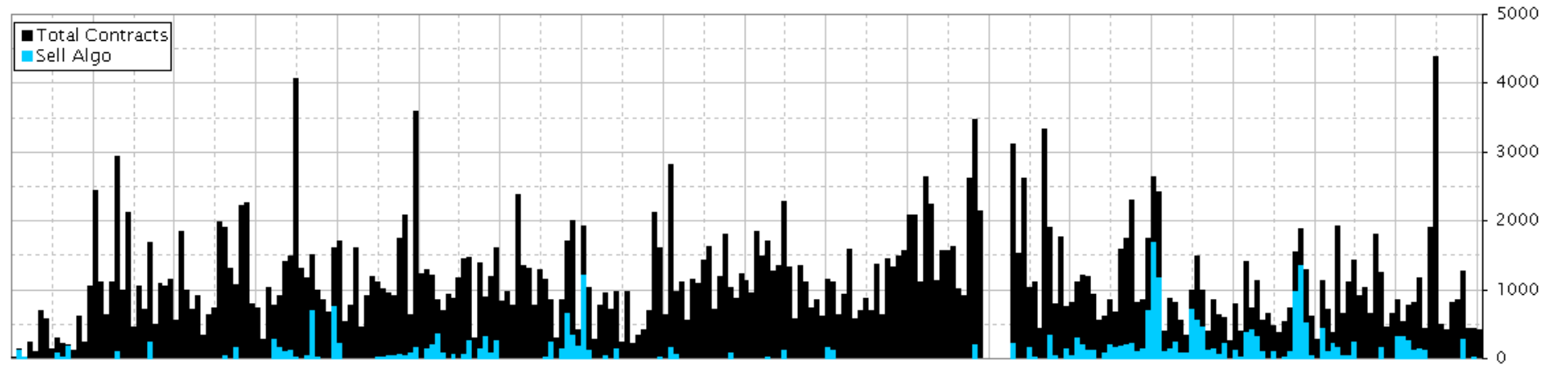
A fat finger!

Nobody likes being called fat!



- Hedge fund managers don't like it either
- Waddell & Reed released their trade data



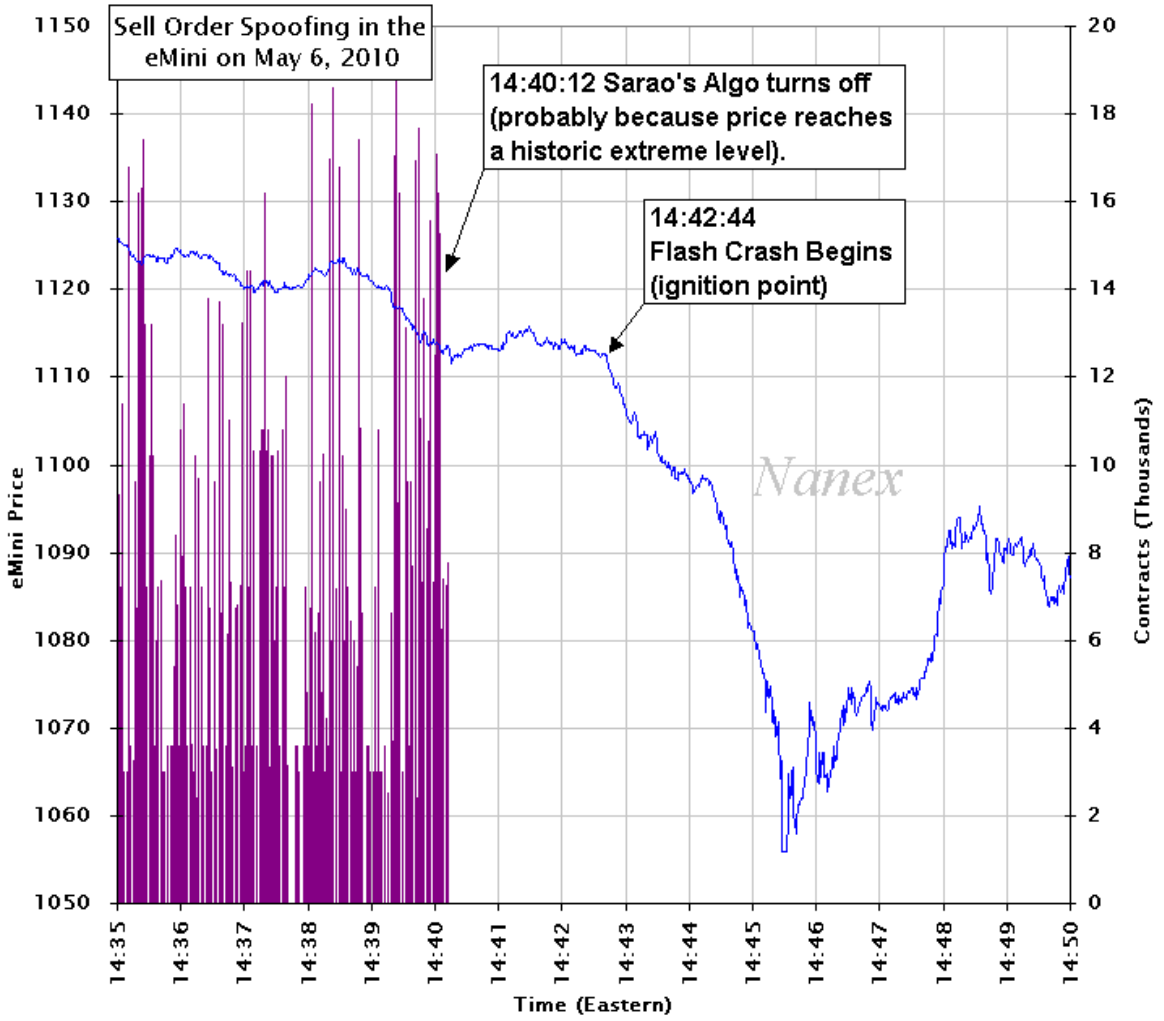


Selling mainly on up-tick: It was ***not*** their fault

Update: 2016

- Flash crash caused by a lone trader (Sarao)
  - The *Hound of Hounslow*—trading manually (using hand and mouse) from his parent's house
  - Spoof messaging (adding and removing orders before execution to unbalance the book)
  - [He has] "*Always been good with reflexes*"
    - So this was no fat-finger error, this was fraud!
  - Made £20-30 million profit. Extradited to USA
  - Admitted guilt. Now awaiting 30 year jail sentence

https://twitter.com/nanexllc/status/591648730501156865/photo/1?ref_src=twsrc%5Etfw



Qu: Sarao's guilty, right?


https://pixabay.com/en/binary-man-face-despair-715831/

Flash Crash: ***Why?***

- High Frequency Trading (HFT) algorithms
  - Automatically buy/sell at ultra-high speeds
  - Now account for more than 50% of trades
- Complex emergent interactions
  - Result in (perhaps) unanticipated system behaviours

Qu: Are HFT programmers morally responsible?

Summary

- Autonomy and pervasiveness go hand in hand
  - Greater autonomy makes tech cheaper, so tech becomes more pervasive
  - As tech becomes pervasive, financial incentives for automating it increase
- Autonomous systems have emergent behaviours
  - Machine learning/AI systems “learn”/”adapt”, so behaviours not fully understood by the original programmers
- Who is morally responsible for autonomous systems?
  - We can use the Ad Hoc Committee’s five rules
- We will apply these concepts in **tomorrow’s workshop**

Related Reading

- Ad Hoc Committee for Responsible Computing (2010), *"Moral responsibility for computing artifacts: five rules",* Version 27, 30 Oct 2010. [Link to Article]
- Bartz, Daniel, "Toyota sees robotic nurses in your lonely final years", *Wired.com*, 19 Jan 2010 [Link to Article]
- Block, Melissa, "With new GPS devices, hikers taking more risks", *All Things Considered*, 26 Oct 2009 [Link to Radio Broadcast]
- Chopra, (2010), "Rights for autonomous artificial agents?", *Communications of the ACM*, Vol. 53, No. 8, pp. 38-40. [Link to ACM Access]
- Chopra & White, (2011) "A legal theory for autonomous artificial agents", University of Michigan Press [Link to Publisher]
- Lovgren, Stefan, "Robot code of ethics to prevent android abuse, protect humans", *National Geographic*, 16 Mar 2007 [Link to Article]

Related Reading [Books]

- Orwell, George. *Nineteen Eighty-Four*. New York: Plume, 2003. Print.

  Orwell's famous novel, written in 1948, is a hugely influential tale of the perils of pervasive surveillance technology. The concept of two-way tele screens might seem to be a more natural fit for week 5 (privacy), but the pervasiveness of tele screens is key to understanding their power. You might also want to think about this book in relation to panopticism, as discussed in week 5.

- Asimov, Isaac. *I, Robot.* New York: Gnome Press, 1950. Print.

  Asimov's "three laws of robotics" are one of the earliest and most influential attempts to define a type of moral code for autonomous machines. For example, in 2011, Motorola purchased Three Laws Mobility (3LM), a company that makes mobile-phone security software. Their design philosophy is based on a version of Asimov's three laws.

- Wallach, Wendell, and Colin Allen. *Moral Machines: Teaching Robots Right from Wrong.* New York: Oxford University Press, 2010. Print.

  Yale ethicist Wendell Wallach and Indiana University cognitive scientist Colin Allen explore the possibility of engineering morality and creating moral autonomous agents.