Artificial intelligence and systemic risk

Jón Daníelsson

Systemic Risk Centre
London School of Economics

modelsandrisk.org/AI

24 October 2019

From

- Artificial intelligence and Systemic Risk. Jón Daníelsson, Robert Macrae and Andreas Uthemann
- Available at SSRN: https://ssrn.com/abstract=3410948
- VoxEU Artificial intelligence and- the stability of markets https://voxeu.org/article/artificil-intelligence-and-stability-markets

What AI can and cannot do

- AI can master any decision process with a *defined action space* better than any human
  - chess, go, computer games,...
- If the action space is ill defined (like all human endeavours), not so easy
- AI today is *unable to reason about things it has not seen*
- It can generalise within a local problem but cannot apply experiences from one domain to another
- Because it does not understand the global problem in which the local one is embedded
- It can handle decisions to the extent they can be mapped onto a *contained local problem*
  - driving a car, medical diagnosis, allocation of credit

Systemic risk

Probability of an unlikely financial crisis causing a severe economic recession

- Systemic risk can not be eliminated
- Happens on the *boundaries* of silos
- One year out of every 43 for OECD countries
- Does generally not arise from the behaviour *or failure* of any individual financial institution
- For the United States it is not an event with costs in the USD billions or tens of billions or hundreds of billions
- Instead, *several trillions*
- Perhaps more than 10% of GDP

Systemic risk is all about the unknown unknowns

- The US stock market goes down by $200 billion in one day and nobody cares
- Potential subprime losses of less than $200 billion and a global crisis happens
- Risk we know we prepare for — *known unknowns*
- Risk we don't know is the dangerous type
- *Unknown unknowns* are most damaging
- Almost axiomatic that the next crisis will *happen* where the *nobody* looking

Bob — The Bank of England Bot

- Suppose we use BoB for supervision
- He will be very useful for micropru
- Not as much for macropru, where he can be dangerous

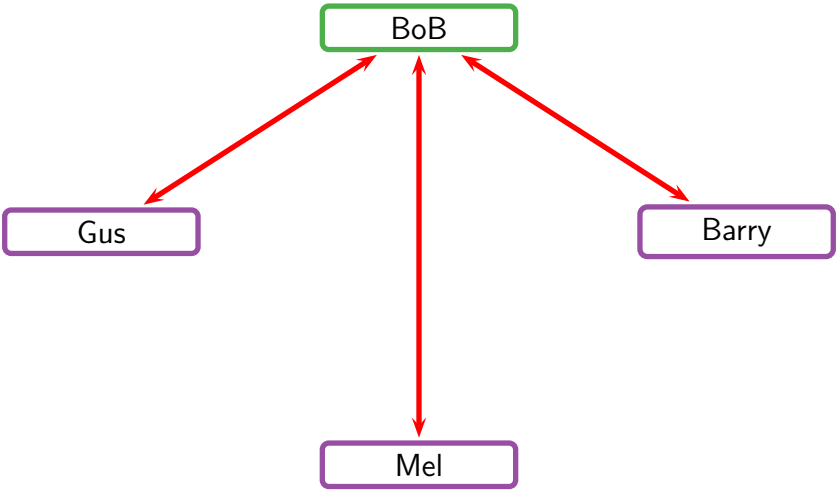
Risk management, compliance and micropru

- There is no *technological reason* why AI cannot play a major role in most
  - risk management functions in financial institutions
  - microprudential supervision
- Job of the supervisor and the risk manager will become high level interaction with their respective AIs
- BoB talks to the banks' bots
- Passing data, model, rules, questions and decisions
- Objections are *cultural, political, legal*
- Short-term investment in technology versus long-term savings in human capital
- Technology is mostly here
- The financial case is clear

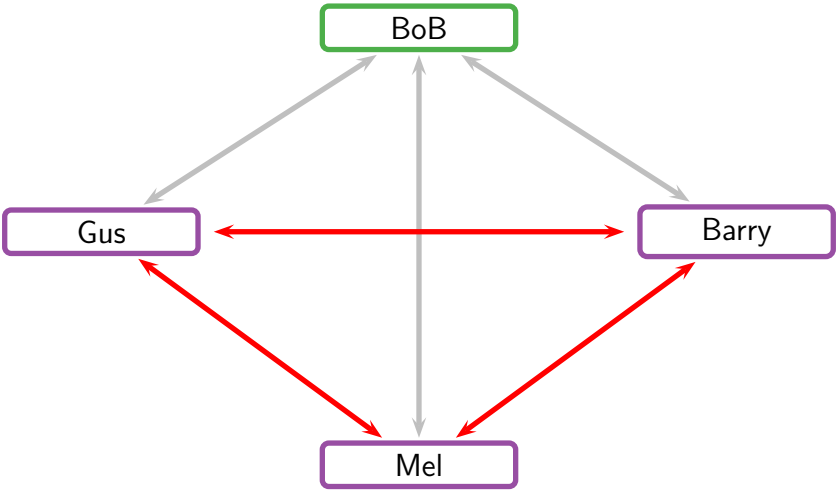
Macro. Case not so clear

- There are limits to AI
- Some can be overcome with technical developments
- Others at present can not
  1. understanding reasoning
  2. procyclicality
  3. unknown unknowns
  4. optimisation against the system

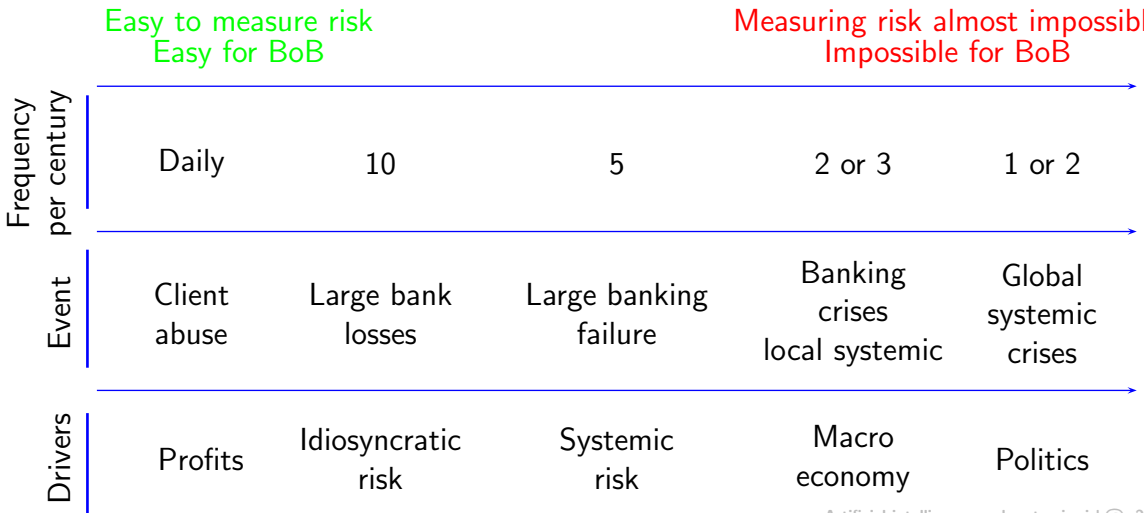
Bob, the Bank of England Bot, and friends



Bob, the Bank of England Bot, and friends



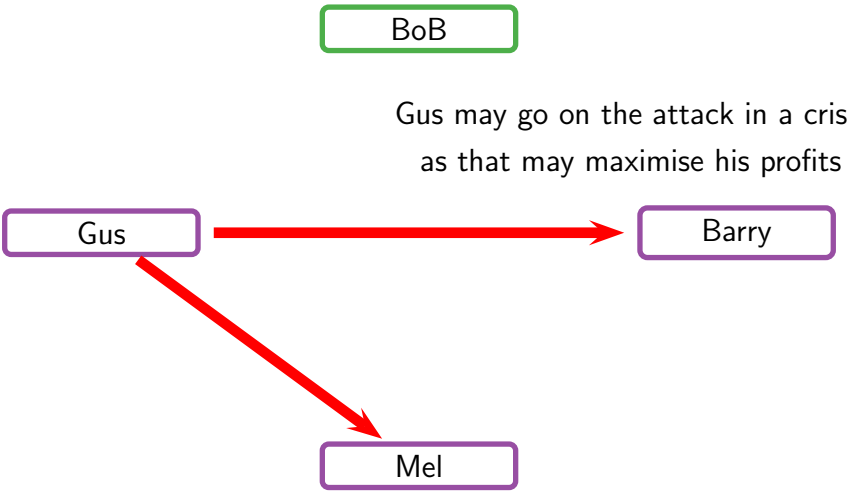
The time dimension of risk



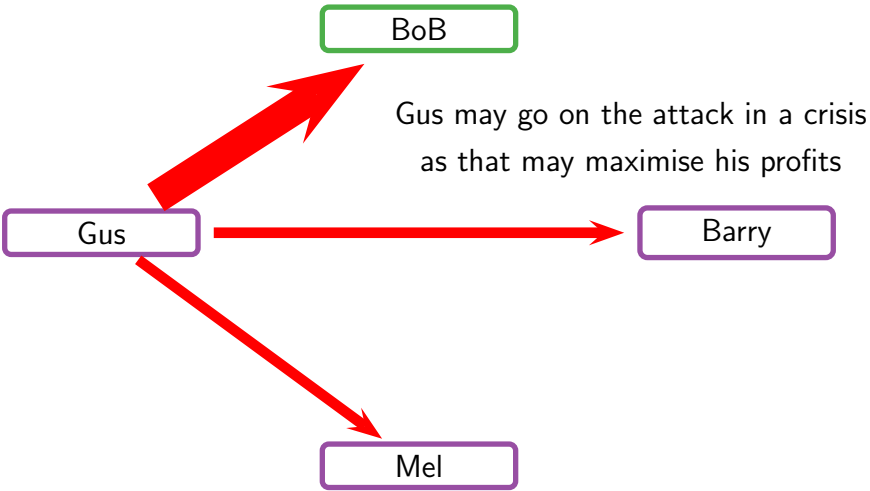
Trust: Inability to do causality and reason

- A 1980s AI, EURISKO, played a naval wargame
- It found the best solution was to sink its own slowest ships
- It is impossible to specify all eventualities
- Humans can reason about unseen things, AI will not
- Can ask a human how she will reason before employing her
- But AI will make decisions, so it will need a *kill switch to prevent it from doing something catastrophic*

The need for a kill switch



The need for a kill switch



Procyclicality

- BoB will favour homogeneous best–of–breed methodologies and standardised processes
- Even stronger than human authorities
- In-breeding and homogeneity make behaviour more procyclical
- Which increases systemic risk

BoB cannot find unknown–unknowns

- Systemic vulnerabilities tend to happen on the boundaries of areas of responsibilities (the silos)
  - subprime mortgages put into structured credit products with hidden liquidity guarantees
  - crossing multiple jurisdictions, institutional categories and countries
- Where humans and AI are least likely to look
- Current AI can easily be trained on events that have happened: risk
- It can perhaps be trained on events that have been imagined: known-unknowns
- Our system is *endogenously infinitely complex*
- It will always miss unknown-unknowns.

Optimise against the system

- BoB's optimisation is harder than that of a malicious actor because BoB faces an infinitely complex computational problem
- A malicious actor only has to optimise against very small part of that domain
- A human regulator provides a natural defence because they create randomness, nuance and interpretation which varies across individuals and time
- Regulators also use common sense and understand out-of-domain constraints such as the limits to regulator powers
- For AI, such randomised responses, would have to be programmed in, and hence would not be acceptable

To be effective, the macroprudential AI needs to

1. control across borders
2. control across silos
3. share data across borders and silos
4. randomise responses
5. create rules in a nontransparent way

6. understand causality in in unforeseen cases
7. reason on a global rather than local basis
8. identify threats that have not yet had bad outcomes

The first 5 are unacceptable; the last 3 are beyond current capabilities

So...

- BoB and his friends will become increasingly useful to microprudential regulators and risk managers
- Reduce costs for financial institutions and supervisors
- Change the job of the supervisor
- Increase systemic risk
- Reduce volatility and fatten tails
- Need a kill switch