LAW4039 Notes

Adithya C. Ganesh

September 27, 2019

Contents

1 Lecture 1: Society, Institutions, Risk, and AI

1.1 R1: Why the US Went to War in Vietnam

https://www.fpri.org/article/2017/04/united-states-went-war-vietnam/

1.2 R2: Five years after Deepwater Horizon, what has the disaster taught us?

https://www.minnpost.com/earth-journal/2015/04/five-years-after-deepwater-horizon-what-has-disaster-taught-us/

1.3 R3: A Declaration of the Independence of Cyberspace (1996)

1.4 R4: The Great AI Awakening, N.Y. Times Magazine (Dec. 14, 2016)

1.5 R5: Regulating AI Systems: Risks, Challenges, Competencies, and Strategies

1.6 R6: Russell and Norvig, AI (2015)

1.7 R7: Suleyman and Musk: Humans Must Become Cyborgs to Avoid AI Domination

1.8 R8: Chui: Applying AI for Social Good, Discussion Paper, McKinsey and Co

1.9 Lecture Notes

The challenge for technologists: learning a little bit of law during these 10 weeks.

1.9.1 A tragic instrusion

The hardest cases to discuss and analyze: the death penalty cases (these ones go to California Supreme Court). Example: intrudred enters the home of a woman who's home with her four year old son. Mother was stabbed, and bleeds to death.

Important questions:

- How to structure a six-picture "show-up" for the child at the push of a button?
- Do we use pervasive surveillance capacity into an ML-enabled infrastructure for security?

Death penalty cases are not-unlike analyzing the ethics of AI systems, where one must weigh the balance between different foundational ethical principles.

Two AI based approaches:

- GANs to generate images based on textual descriptions.
- Convolutional neural networks for facial recognition / person tracking.

But there are concerns:

- How accurate is the technology?
- What is the line between police and non-police?

1.10 Why are we here? Because AI is not just theory…

AlphaZero works across games. 5-7 years ago, most of these feats were highly domain specific.

Kasparov: "I can't disguise my satisfaction that it plays with a very dynamic style…much like my own."

1.11 What is AI?

- Capacity to undertake functions that, if performed by a human, would generally be understood to require intelligence" (Russell and Norvig)

- Definition of "intelligence" "Ability to learn or understand or to deal with new or trying situations"

It's worth differentiating between a few types of AI: *domain-specific AI*, which is supervised ML attached to some application; *simulated AGI*, (e.g. Echo / Google Home, etc.).

1.12 What is regulation?

For instance: how fast can you go on 280? While the Autobahn has no speed limits, 70-90 is a reasonable speed when

- We define the term broadly: How society defines legitimate authority to structure relations among people, organizations, information, and the physical world.

- Not just specific rules, imposed by agencies, governing the design, operation, or use of technology.

Governance intersects with politics, socioeconomic behavior, specifically including:

- Constitutions

- Statutes

- Judge-made law (e.g. the common law)

- Agency decisions

- Norms

1.13 A (straightforward) problem AI can help address

- 210 languages spoken in California courts.

- Interpreters provide services in courtrooms, but not enough multi-lingual staff over the counter.

- 2018: We began a pilot project to provide translation over-the-counter to limited-English court users, using tablets

- AI could help enhance: access to justice (but als omedical care, education).

1.14 Deepwater Horizon: disaster

- Offshore drilling: Hole, metal casing, cement "slugs," drilling fluid pumped down to balance pressure; drilling operators cement well, then replaced by rig to extract oil.

- But problems lurked: Well had a fracture, troubling readings ignored, too much mud displaced, blowout preventer failed.

- 9:49pm, April 20, 2010: First explosisons occurred.

- Cuellar at the WH, helping to reconcile work of NOAA and FDA to assess and mitigate impact of the disaster on seafood.

Importantly, note that:

- Blowout preventer failed; no clear solution other than to drill a replacement well, which would take months

- Hard to understand who was responsible.

- 5M gallons of oil released over 87 days.

- 11 lost lives and many more sickened

- Oil spill affected water quality, fisheries, etc.

- Financial impact as of 2018: circa $65B (nearly $9B impact on fisheries alone).

1.15   Deepwater Horizon: Aftermath

- "oil and gas industry has not retreated to safety. Instead, it has expanded its technological horizon in ways that make it harder to foresee the complex interactions between drilling technologies, inevitable human errors, and the ultra-deepwater environment"

- More layers of abstraction? Seems like sketchy incentives.

1.16   Deepwater Horizon: Why does this matter?

- Could some kind of AI system have helped prevent the disaster?

- Maybe: accidents and the result harms can be reasons to innovate.

- Can we at least expect govt. to handle thoughtfully the biggest risks they perceive?

1.17   Second case study: Vietnam War - Puzzles

- Why did the U.S. intervene on the other side of the world? (Where the French had failed)

- Why did the U.S. persist in a course of action so likely to result in failure?

- How could Vietnam win against a technologically and economically superior adversary?

- How does this episode tell us about how governments make decision (even when they have so much at stake) and about geopolitics?

1.18   Vietnam: French influence and internecine struggles

- French influence: By 1860s, French troops engage in battle for control of "Indochina," are awarded 3 provinces by Emperor Tu Dug as a concession

- French power consolidated: By 1879, French (having taken over Cambodia and Vietnam) name first colnial governon and claim "Annam and Tonkin" as protectorate.

1.19   U.S. interlude: Roosevelt and anti-colonialism

- U.S. strategic interest not served by advancing colonialism

- Franklin Roosevelt had personal concerns   `Finish this`

1.20   Vietnam War: A tenuous post-colonial compromise (1954)

- Peace conference in Geneva resulted in the agreement that the French positions would be evacuated completely

1.21   The Geneva Accords break down

- Accords quickly unravel:

  - French encourage independence of South Vietnam, using Bao Dai.

  - Bao Dai (with advice of the French) appoints Diem premier, then President. Rejects Geneva Accords.

  - U.S. did not sign and did not feel bound to them; intense domestic concern over the victory of Mao and allies in China.

1.22 Kennedy & Ho: Evolving priorities

- Kennedy - concern with 1964 election.
- Ho - consolidation, Internal political divisions.

1.23 From Kennedy to Johnson: Change & continuity

- Continuity: powerful faction concerned about U.S. credibility
- Change: different political calculations (Johnson v. Kennedy): "Perhaps the most complicated figure in 20th century history."

1.24 Johnson's political & org constraints

- Expansive domestic program
- Looming presidential election
- Key advisers preferred expanding the war
- Political opponents sought to target Johnson Administration

1.25 Self-reinforcing strategic traps

- North Vietnamese and VC strategy: Inflict losses on the U.S. sufficient to persuade it to leave or negotiate.
- Washington D.C. strategy: pursue a course minimizing risks that U.S. will lose credibility; leaving in the wake of attacks on troops would all but guarantee such a loss in credibility

1.26 Johnson's decision

- Johnson's domestic program
- Why does Johnson accept?

1.27 U.S. context: Domestic impact

- Draft
- Economic impact
- Mounting casualties
- Growing social unrest

1.28 By 1966: McNamara relents

McNamara: a complex figure, who did math. The 8th U.S. Secretary of Defense. He would be able to understand the mathematical basis of AI that we're talking about here.

- Former architect of war begins to see no change.
- McNamara writes key memo, but keeps it secret.
  - "The picture of the world's greatest superpower killing or seriously injuring 1000 noncombatants a week, while trying to …is not a pretty one.
- Out of loyalty to LBJ, continues to defend war.
- Personal dimension.

1.29 Mid-1970s: Saigon Falls

- Saigon's deterioriating position.

- Chaos as Saigon faills in 1975

1.30 Why the delay in stopping? Countries are a "they," not an it

1.31 Vietnam War: Implications

- Severe economic costs and political isolation for Vietnam.

- Collapse of South Vietnam.

- Political turmoil and devastation in Southeast Asia. Laos is perhaps the most heavily bombed country in history.

- Deaths: up to 2 million Vietnamese civilians, 1.1 million N. Vietnamese military, 250K S. Vietnamese military, 58K U.S. military.

Operation Igloo White: an early military operation of AI ($1B).

1.32 Responding to risk, conflict, and innovation: Preliminary implications

- Deepwater Horizon
  - Harm from neglect of systemic safety issues
  - Overlapping responsibility and lack of clarity in accountability, despite laws
  - Deep contradictions in how society benefits from and is harmed by technology
  - Political economy lurks in the background

- Vietnam war: interaction of history and geopolitical circumstances, technology, domestic political pressures, and organizational realities = an outcome no one person or organization wanted.

1.33 Preliminary implications of case studies (cont…)

- Early internet

- Explosive growth

1.34 Simple summary of case study implications

- Small choices have big consequences

- Geopolitics lurk in the background at nearly every turn

- "Emergence" is a crucial factor

- Virtually all big societal dilemmas have both a legal and an organizational dimension (e.g. Vietnam).

Themes make the case for "regulatory thinking"…

1.35 Defining AI: "The behavior of human beings"

There are two broad kinds of AI problems: *recognition*, vs. *cognition.* The former is less "explainable" and the latter is more "explainable." Slowing down deliberately can enable "system 2" modes of thinking.

Alpha Zero is definitely both

1.36   Execute a class w/ these goals: Four legs of a stool

- Lectures
- Speakers
- Group projects
- Your own work
- Attendance

1.37   Preview: substantive topics

- Key people: Herbert Simon (AI as decision science), Frank Fukuyama (trust and power are central to politics and the state), Ed Feigenbaum (Cognition, recognition; not just machine learning)
- Intellectual influences: Judith Shklar, Charles Perrow, Arlie Hochschild