CSCI 215
Social and Ethical Issues in Computer Science

artificial intelligence

Recitation week 5

- Any questions?

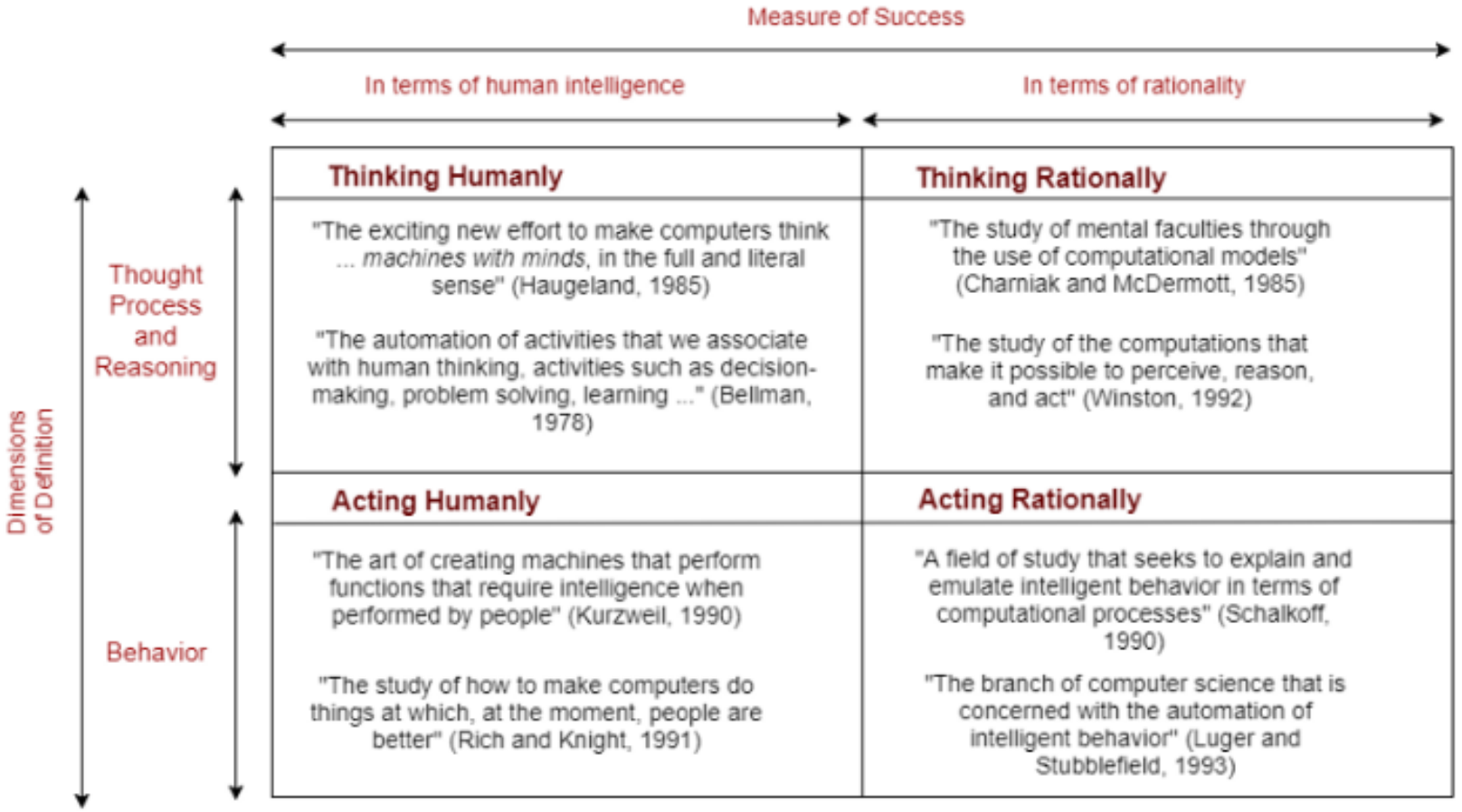
What is artificial Intelligence?

- What is a definition of AI?
- Give us some examples of AI
- Why are we talking about this in this class?

What is artificial Intelligence?

- Definition: the theory and development of computer systems able to perform tasks that normally require human intelligence, such as visual perception, speech recognition, decision-making, and translation between languages.
- Agree or disagree?

From CSCI 446 Book



Credit—Stuart Russell and Peter Norvig

Approaches

- Human centered
  - Involves observations and hypotheses about human behavior
  - Empirical
    - based on, concerned with, or verifiable by observation or experience rather than theory or pure logic

VERSUS

- Rationalist
  - Uses math and engineering
  - reasoning
  - Independent of experience

Rational Thinking

- Rational thinking is the ability to think out of the box, to think critically, to produce “good quality of thoughts” in certain situations to arrive at a rational decision. – quora.com

- In different situations, rational thinking encompasses our ability to draw justifiable conclusions from data, rules and logic – scientificamerican.com

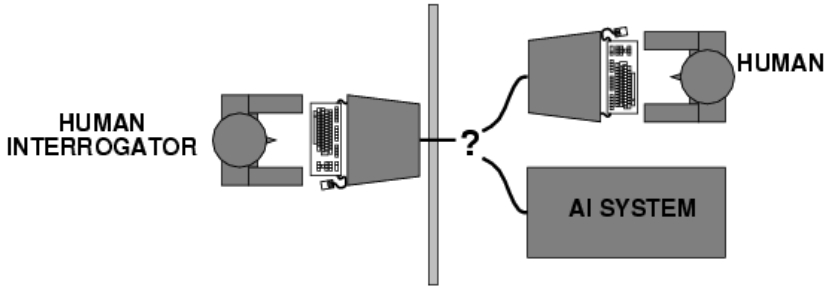
- A rational process means always doing the right thing based on current information, given an ideal performance measure
  - They go by the book and assume the book is correct

Acting rationally: Rational agent

- A rational agent is one that acts to achieve the best expected outcome
  - Goals are application-dependent and are expressed in terms of the utility of outcomes
  - Being rational means maximizing your goal achievement
- Computer that acts rationally relies on the recorded actions to interact with the environment based on conditions, environmental factors, and existing data.

Acting humanly

- Turing (1950) "Computing machinery and intelligence"
- The Turing Test



- What capabilities would a computer need to have to pass the Turing Test?
  - Natural language processing
  - Knowledge representation
  - Automated reasoning
  - Machine learning
- Turing predicted that by the year 2000, machines would be able to fool 30% of human judges for five minutes

Artificial Intelligence

- The big problem is that the complexity of the software often means that it is impossible to work out exactly why an AI system does what it does. With the way today's AI works – based on a massively successful technique known as machine learning – you can't lift the lid and take a look at the workings. So we take it on trust. The challenge then is to come up with new ways of monitoring or auditing the very many areas in which AI now plays such a big role.

Taken from http://www.bbc.com/future/story/20170307-the-ethical-challenge-facing-artificial-intelligence

Question

- How would you certify these systems as safe?

Artificial Intelligence

- We can't get so excited about what we are able to do that we forget to ensure that it is ethical

- Ian Malcolm: "Your scientists were so preoccupied with whether they could, they didn't stop to think if they should."

Case Study

fairness and neutrality

- Imagine, in the near future, a bank using a machine learning algorithm to recommend mortgage applications for approval. A rejected applicant brings a lawsuit against the bank, alleging that the algorithm is discriminating racially against mortgage applicants. The bank replies that this is impossible, since the algorithm is deliberately blinded to the race of the applicants. Indeed, that was part of the bank's rationale for implementing the system. Even so, statistics show that the bank's approval rate for black applicants has been steadily dropping. Submitting ten apparently equally qualified genuine applicants (as determined by a separate panel of human judges) shows that the algorithm accepts white applicants and rejects black applicants. What could possibly be happening?

Taken from https://intelligence.org/files/EthicsofAI.pdf

Ethical Issues

Safety of AI from our enemies

- A machine vision system to scan airline luggage for bombs must be robust against human adversaries deliberately searching for exploitable flaws in the algorithm— for example, a shape that, placed next to a pistol in one's luggage, would neutralize recognition of it. Robustness against manipulation is an ordinary criterion in information security; nearly the criterion. But it is not a criterion that appears often in machine learning journals, which are currently more interested in, e.g., how an algorithm scales up on larger parallel systems

Taken from https://intelligence.org/files/EthicsofAI.pdf

Ethical Issues

Guarding against mistakes

- Another important social criteria for dealing with organizations is being able to find the person responsible for getting something done.
- When an AI system fails at its assigned task, who takes the blame? The programmers? The end-users?
- Modern bureaucrats often take refuge in established procedures that distribute responsibility so widely that no one person can be identified to blame for the catastrophes that result (Howard 1994). The provably disinterested judgment of an expert system could turn out to be an even better refuge. Even if an AI system is designed with a user override, one must consider the career incentive of a bureaucrat who will be personally blamed if the override goes wrong, and who would much prefer to blame the AI for any difficult decision with a negative outcome.

Taken from https://intelligence.org/files/EthicsofAI.pdf

Ethical Issues

Loss of jobs

- One area fraught with ethical issues is the workplace. AI will let robots do more complicated jobs and displace more human workers. For example, China's Foxconn Technology Group, which supplies Apple and Samsung, has announced it aims to replace 60,000 factory workers with robots, and Ford's factory in Cologne, Germany puts robots right on the floor alongside humans.

What should we do with these displaced workers?

Taken from http://www.bbc.com/future/story/20170307-the-ethical-challenge-facing-artificial-intelligence

Ethical Issues

Affect on Mental Health

- If increasing automation has a big impact on employment it could have a knock-on effect on people's mental health. "If you look at what gives people meaning in their lives, it's three things: meaningful relationships, passionate interests, and meaningful work," says Ezekiel Emanuel, a bioethicist and former healthcare adviser to Barack Obama. "Meaningful work is a very important element of someone's identity." He says that in regions where jobs have been lost when factories close down can face increased risk of suicide, substance abuse and depression.

Taken from http://www.bbc.com/future/story/20170307-the-ethical-challenge-facing-artificial-intelligence

Ethical Issues

- Disbursement of wealth generated by machines
- Affecting our behavior and interaction
- Artificial Stupidity - Guarding against mistakes
- Unintended consequences
- Possible loss of human control
- Do we have to worry about humane control of AI/robots?
  - Example: ability to inflict pain - futuristic

https://www.weforum.org/agenda/2016/10/top-10-ethical-issues-in-artificial-intelligence/