Introduction to Artificial Intelligence
Lecture 1

What is AI and why is it worthy of study?

What does it mean to think and could/should artifacts do so?

Agenda

- What is AI?

- Foundations

- History

- State of the art

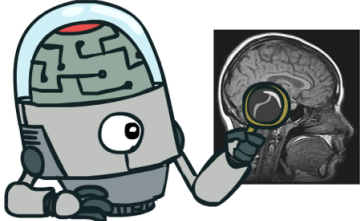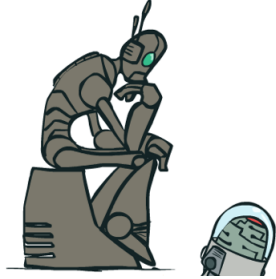- Philosophy: Weak vs. Strong AI
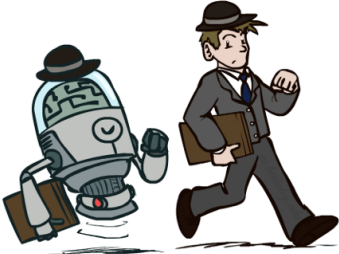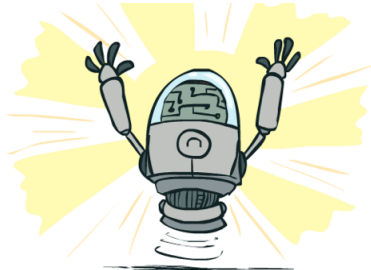
- Ethical considerations

Artificial Intelligence

- Various fields of study attempt to understand intelligence

- Artificial Intelligence (AI) attempts not just to understand, but to build intelligent entities/systems (known as **agents**)

- But what does that mean?

Approaches to AI

	Humanly	Rationally
Thinking	**Cognitive Modeling**	**"Laws of Thought"**
Acting	**Turing Test**	**Rational Agent (this course)**

Introduction to Artificial Intelligence

Acting Humanly

Vol. LIX.   No. 236.]   [October, 1950

MIND

A QUARTERLY REVIEW

OF

PSYCHOLOGY AND PHILOSOPHY

I.—COMPUTING MACHINERY AND
INTELLIGENCE

By A. M. TURING

1. *The Imitation Game.*

I PROPOSE to consider the question, 'Can machines think?'
This should begin with definitions of the meaning of the terms
'machine' and 'think'. The definitions might be framed so as to
reflect so far as possible the normal use of the words, but this
attitude is dangerous. If the meaning of the words 'machine'
and 'think' are to be found by examining how they are commonly
used it is difficult to escape the conclusion that the meaning
and the answer to the question, 'Can machines think?' is to be
sought in a statistical survey such as a Gallup poll. But this is
absurd. Instead of attempting such a definition I shall replace the
question by another, which is closely related to it and is expressed
in relatively unambiguous words.

The new form of the problem can be described in terms of
a game which we call the 'imitation game'. It is played with
three people, a man (A), a woman (B), and an interrogator (C) who
may be of either sex. The interrogator stays in a room apart
from the other two. The object of the game for the interrogator
is to determine which of the other two is the man and which is
the woman. He knows them by labels X and Y, and at the end
of the game he says either 'X is A and Y is B' or 'X is B and Y
is A'. The interrogator is allowed to put questions to A and B
thus :

C : Will X please tell me the length of his or her hair ?
Now suppose X is actually A, then A must answer. It is A's

28   433

THE
IMITATION
GAME
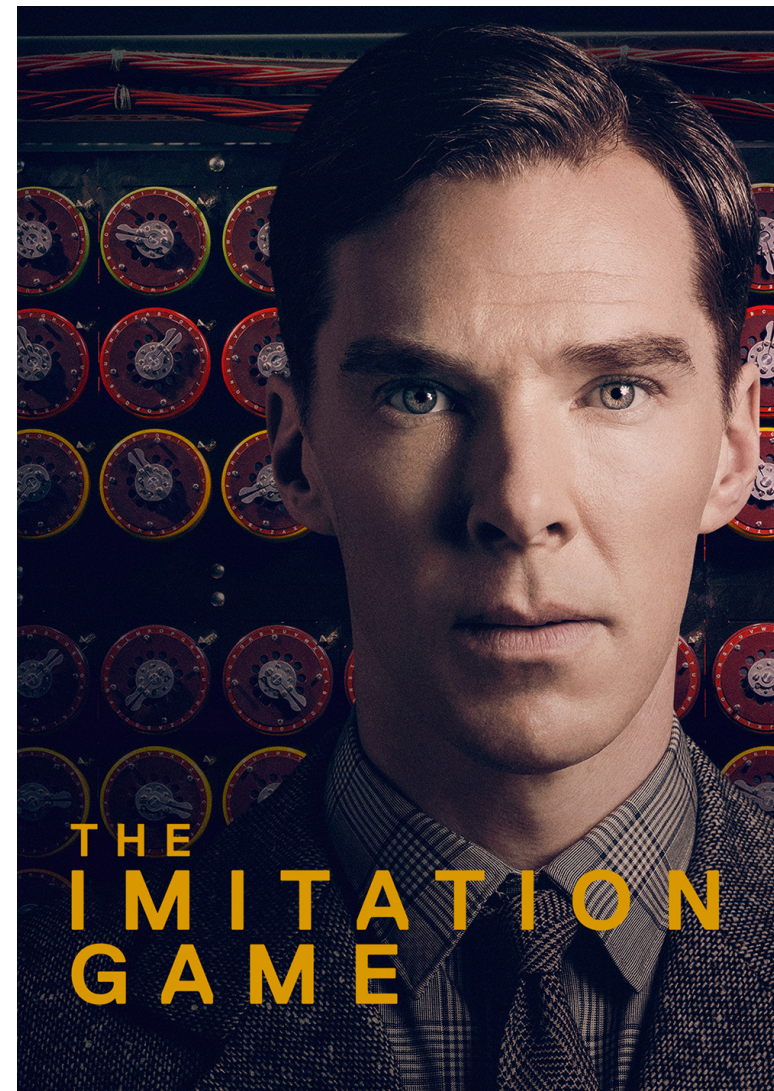
The Turing Test

- Allow a human to determine if a responder is human/AI

- Requires natural language processing (NLP), knowledge representation and reasoning (KRR), learning (ML)
  - A *total* variant incorporates video, and would thus require perception (vision), robotics, [e]motion modeling

- Issues: forces us to focus on minutia (e.g. speed of response, having favorite everything, etc.); must we convince pigeons that we fly like them in order to fly airplanes… rockets?
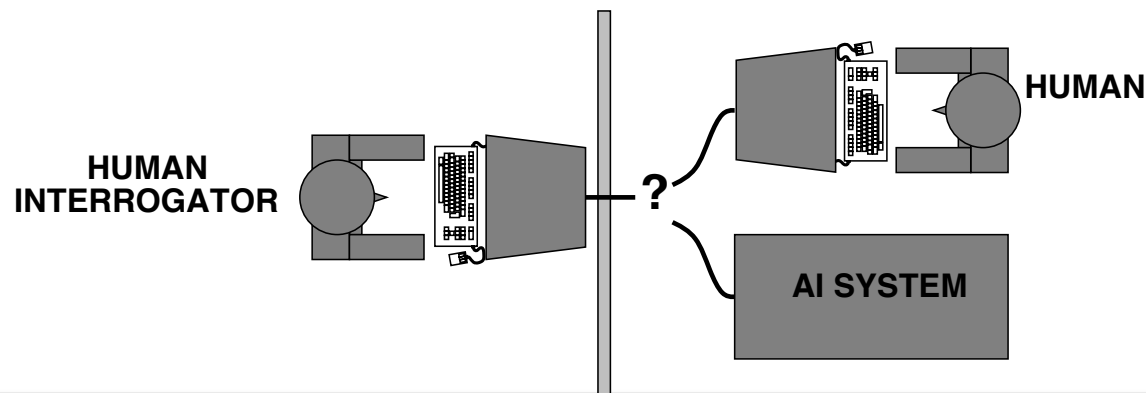
Thinking Humanly

- In the 1960s "cognitive revolution," information-processing psychology replaced prevailing orthodoxy of behaviorism

- So then there was a question of how to develop/validate theories of the brain
  – Cognitive science/modeling: knowledge, human/animal experiments
  – Cognitive neuroscience: circuits, traces/scans

- Issues: difficult to scale up, fly like a pigeon?
  – But fields cross-fertilize

Neuroscience 101
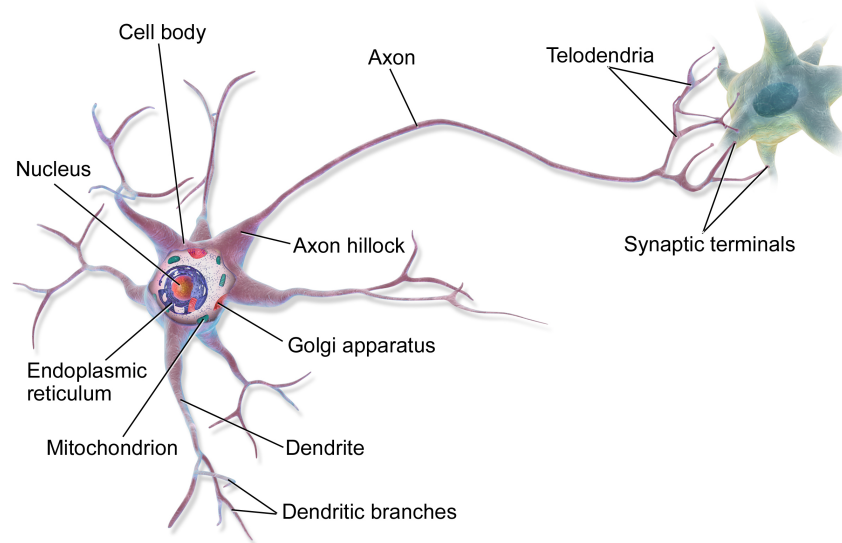
	Supercomputer	Personal Computer	Human Brain
Computational units	10^4 CPUs, 10^{12} transistors	4 CPUs, 10^9 transistors	10^{11} neurons
Storage units	10^{14} bits RAM	10^{11} bits RAM	10^{11} neurons
	10^{15} bits disk	10^{13} bits disk	10^{14} synapses
Cycle time	10^{-9} sec	10^{-9} sec	10^{-3} sec
Operations/sec	10^{15}	10^{10}	10^{17}
Memory updates/sec	10^{14}	10^{10}	10^{14}

Thinking Rationally

- Long history: Aristotle & syllogisms
  - "Socrates is a man, all men are mortal, therefore Socrates is mortal."

- Complex systems have existed for decades that can deduce facts from logical representations

- Issues: world->formal description is difficult (particularly uncertain); many facts = massive computational costs; seemingly not all actions can/should be mediated by logic
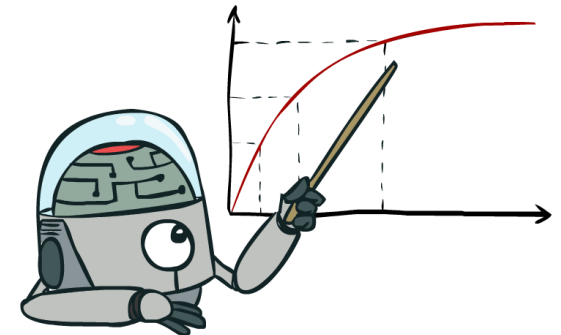
Acting Rationally

- **Rational**: maximally achieving goals
  - Only concerns what decisions are made (not thought process behind them) – mathematically appealing
  - Goals are expressed in terms of the utility of outcomes

- An **agent** perceives and acts
  - Maps percept histories to actions

$$f : P^* \rightarrow A$$

- A **rational agent** acts to maximize <u>expected</u> utility
  - Given limited time/resources, still acts appropriately

AI Foundations

- ## Philosophy
  - – Mind/brain duality, empiricism, induction
- ## Mathematics
  - – Gödel incompleteness, tractability, NP, probability
- ## Economics
  - – Decision/game theory, MDPs, satisficing
- ## Neuroscience, [Cognitive] Psychology
  - – Many neurons -> mind, physical computation
- ## Computer Engineering
- ## Control Theory
  - – Objective function
- ## Linguistics

A Brief History of AI

1940s	Binary model of neurons, Hebbian learning
1950	Turing's "Computing Machinery and Intelligence"
1956	McCarthy, Dartmouth workshop: "Artificial Intelligence" coined
1952-1974	"Look, Ma, no hands!" (Computers can do X!): GPS, checkers (learning!), vision, CSPs, NLP Complexity issues, ANNs disappear
1969-1988	Knowledge-based/expert systems developed, boom!
1988-1993	Expert systems bust, "AI Winter"
1986-	Neural networks reborn (back-propagation), industry investment, resurgence of probabilistic methods, "return to" scientific method
1995-	Refocus on agents, AGI
2001-	Big data

TM-1950

Introduction to Artificial Intelligence

State of the Art

Almost got it!

- Table tennis
- Jeopardy
- Rural driving
- Fold [some] laundry
- Buy groceries on the web
- Real-time translation
- Formulaic journalism

Much work to be done…

- Urban driving
- Buy groceries in store
- Real-time conversation
- Discovery/proof
- Intentional humor

Some Demos

📹 NLP-ASR (Automatic Speech Recognition)

📹 Vision-Object-Recognition

📹 Robotics-Soccer

📹 Robotics-Laundry

Can We Achieve AI?

- Important distinction
  - **Weak AI**. Machines that act is if they are intelligent
  - **Strong AI**. Machines that actually are thinking (not just simulating thought)

- Most assume weak AI is possible, so we focus on the philosophical question…
  "Can machines think?"
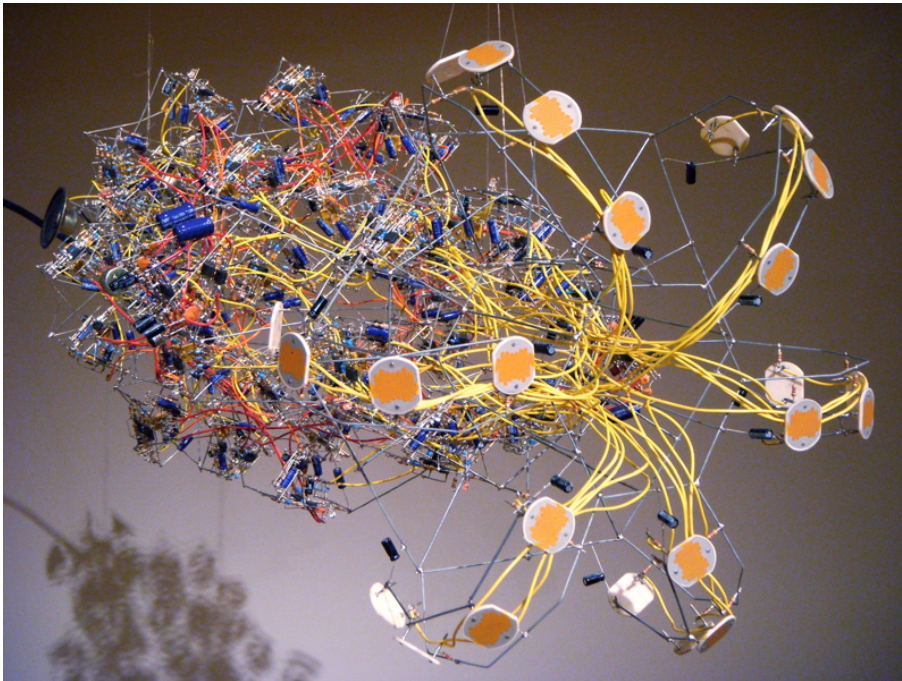 - Turing: "polite" assumption that humans can think

Mental States, Brain in a Vat
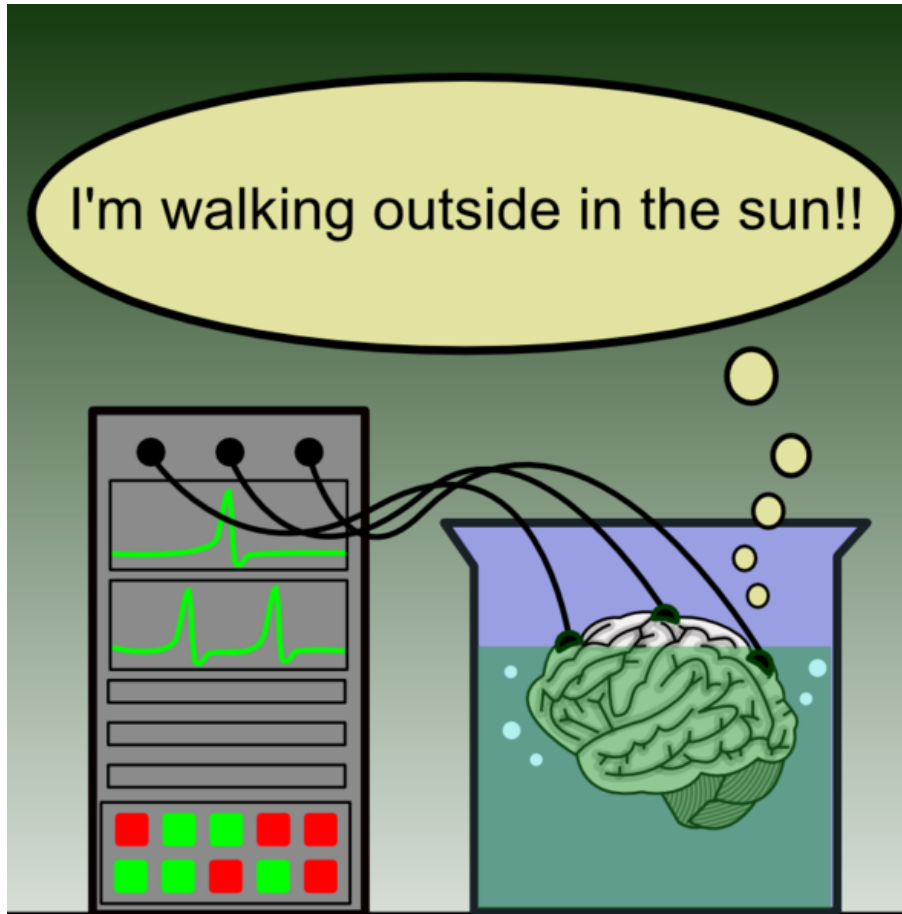
- **Wide content**: omniscient view
- **Narrow content**: consider only brain state

- For purposes of AI, we consider narrow
  - What matters about brain state is its functional role within the operation of the entity

Functionalism, Brain Replacement

- **Functionalism**: mental state is any intermediate causal condition between input and output
  - Isomorphic processes would have same mental states

- If you believe that the replacement brain is conscious, then we could replace the system with a lookup table of states + circuitry

Biological Naturalism, Chinese Room

- Typically seen as an **intuition pump**
  - Amplifies prior intuition without changing anyone's mind

- What would the output be if asked "do you understand Chinese?" What would a human respond?
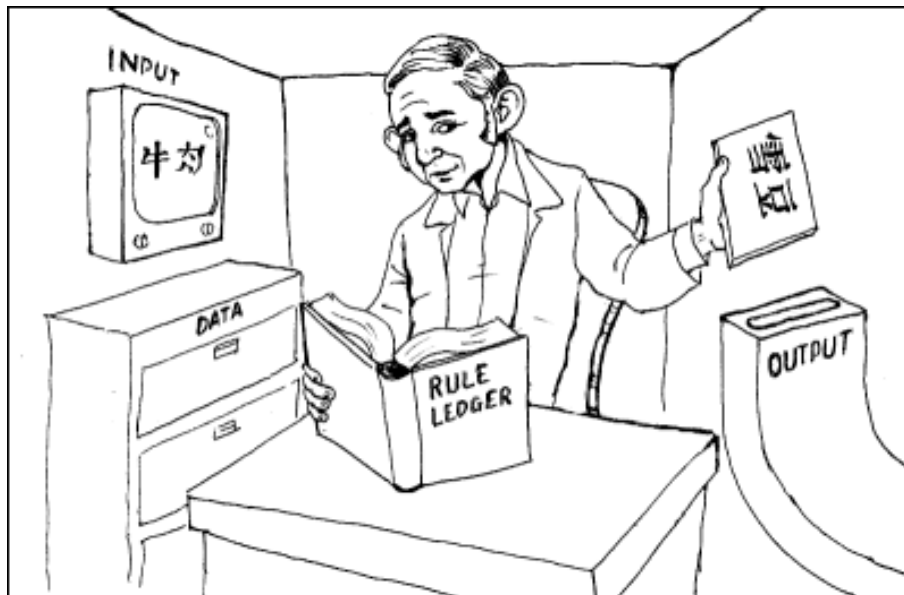
Should We Develop AI?

- In recent years, a popular topic, for politicians, media, and researchers

- Let us consider some issues…

Unemployment

- Generally IT (including AI) has created more jobs than it has eliminated

- There is a trend today towards humans as managers/directors, and human/computer teams

Too Much/Little Leisure

- AI could lead to not enough need for human thought/labor

- Presently, AI amplifies rate of innovation, which increases pressure for work

Losing Sense of Uniqueness

- May lead to questioning foundational moral assumptions

- Consider the current controversy over Darwinism

Undesirable Ends

- There is a need for deliberate policies to balance public/private interests, privacy vs. security

- These discussions need to happen within research areas, as well as in public policy

Loss of Accountability

GOOGLE SELF-DRIVING CAR GETS INTO AN ACCIDENT INVOLVING INJURIES

Robot kills worker at Volkswagen plant in Germany

Contractor was setting up the stationary robot when it grabbed and crushed him against a metal plate at the plant in Baunatal

📷 An investigation is under way into whether human error was to blame for the death of a contractor at the of a robot at a Volkswagen production plant. Photograph: Joerg Sarbach/AP

A robot has killed a contractor at one of Volkswagen's production plants in Germany, the automaker has said.

The man died on Monday at the plant in Baunatal, about 100km (62 miles) n of Frankfurt, VW spokesman Heiko Hillwig said.
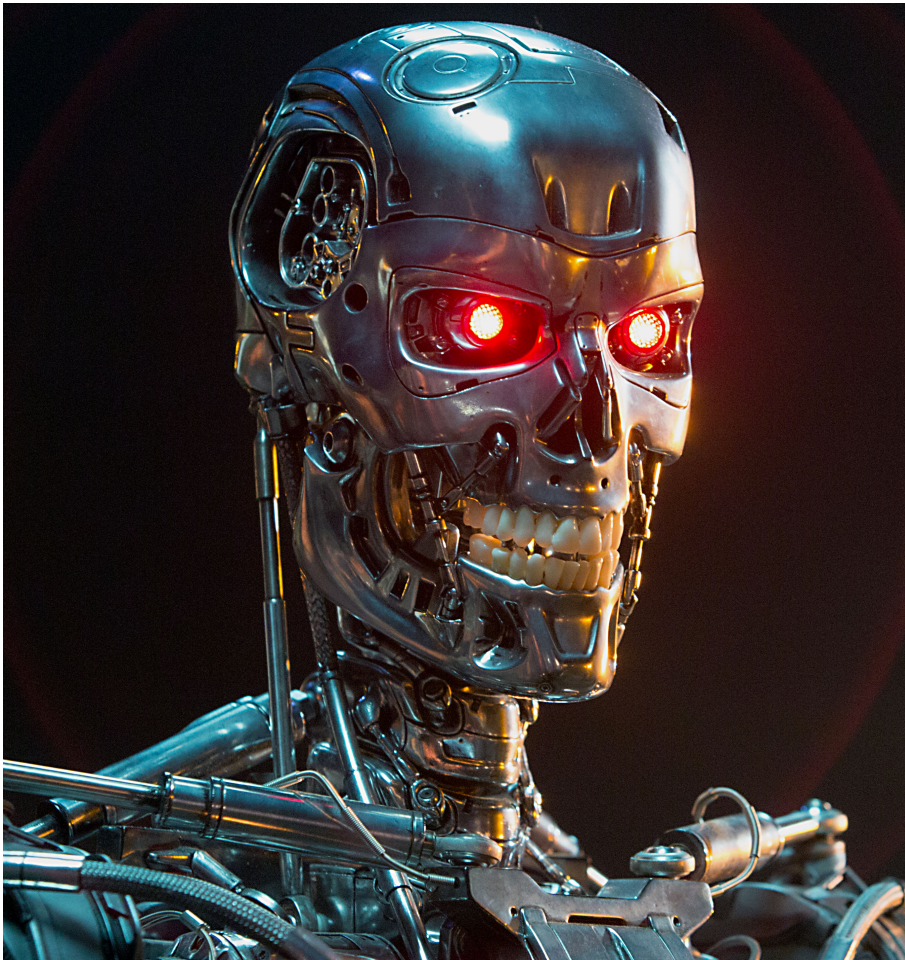
The 22-year-old was part of a team that was setting up the stationary robot w it grabbed and crushed him against a metal plate, Hillwig said.

- The law has yet to catch up with modern developments in the areas of AI, and particularly machine learning

- There is a balance to be struck between hampering innovation and adapting to new technologies

End the Human Race

- Incorrect state estimation
  - Could happen by a human
  - Need checks and balances

- Utility function is hard
  - Minimize human suffering ?= kill humans

- Unintended evolution
  - Singularity
  - Need to consider morality towards AI

Summary

- In this class we will study how to build rational agents, those that maximize expected utility

- AI is an interdisciplinary field that has rich foundations, promising achievements, and a bright future

- As practitioners/researchers, we need to consider the philosophical and ethical implications of AI

Introduction to Artificial Intelligence