

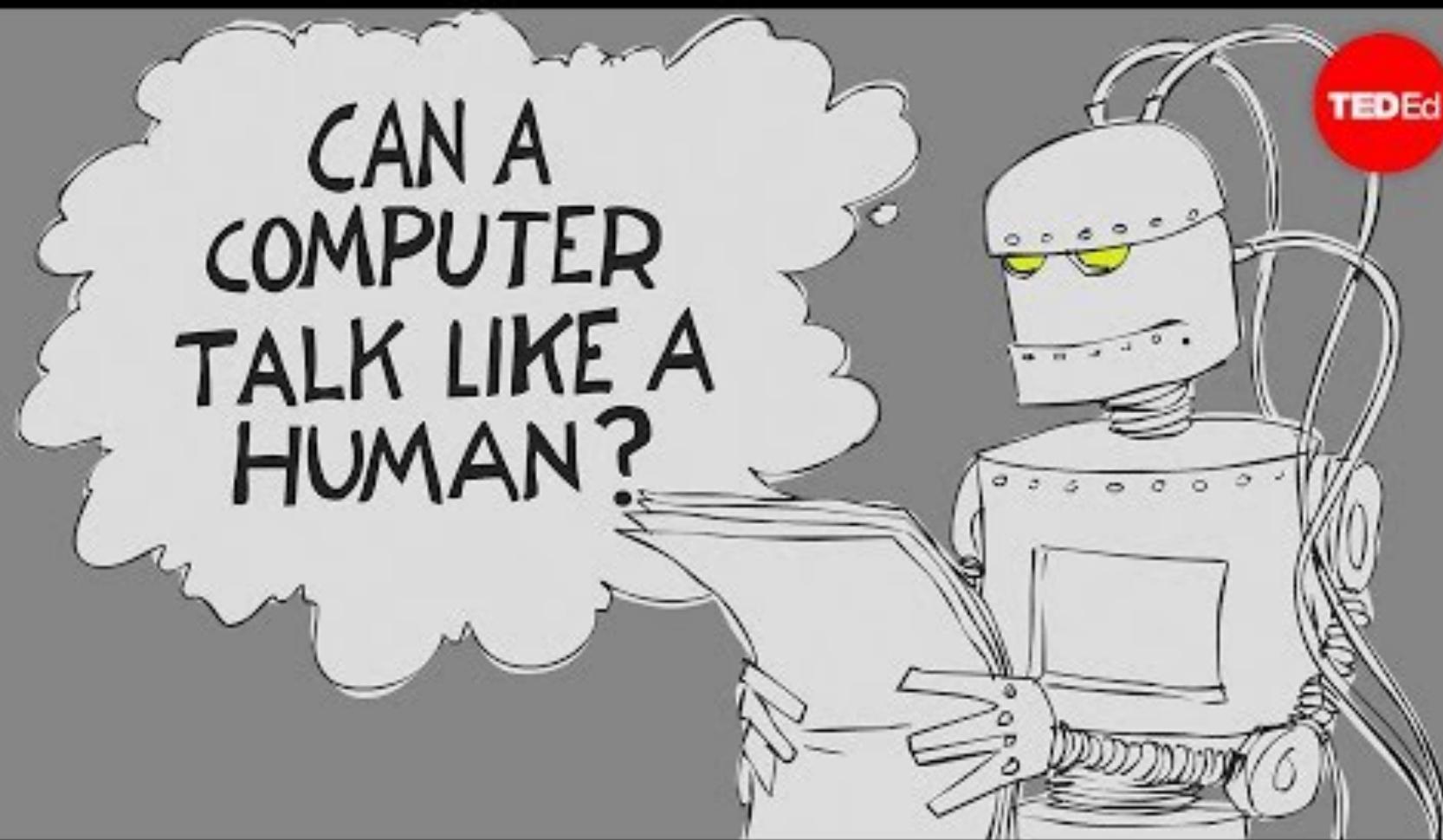
# The four approaches towards AI

Pantelis Monogioudis,  
Jan 2020

# The four approaches towards AI



# Turing Test Approach



- A 5-min behavioral intelligence test, where an interrogator chats with the player and at the end it guesses if the conversation is with a human or with a programmed machine.
- What capabilities we need to have to pass a turing test.
  1. Natural Language Processing
  2. Knowledge Representation
  3. Automated Reasoning
  4. Machine Learning
  5. Computer Vision
  6. Robotics



# Cognitive Models

- A cognitive model is a theory on how the mind works. These theories of cognitive processes are tested via various cognitive *architectures*.
- Historically, there are many architectures and open source implementations available such as ACT-R, SOAR (U of Michigan).
- Its important to realize that much of todays' debate about the path ahead in AI maps to the few different architectures. The *hybrid* architecture (symbolic and connection-oriented) is what is being investigated today by many research institutions. Many publications from big AI shops (Deep Mind, OpenAI, Facebook, etc.) tests at least an element of a theory of cognition and suggest a corresponding architecture.

# The Syllogism-based approach

- The translation from Greek of the word *syllogism* is to support logic.
- This approach emphasizes how we think the right way and encodes the pattern of logical arguments that can reach correct conclusions from a set of propositions (*premises*). Problem solving where the problem statement is expressed in a logic notation, matured in the 60s. Such rule-based systems are still with us and for small and deterministic state spaces can provide a very compact and elegant way to inference.
- Logic-based reasoning is coming back to fashion. One of the most promising areas is **their application** to *interpretability* (also known as *explainability*) of deep learning based methods for e.g. classification in medical diagnosis. Probabilistic Logic Networks (PLN) are extensions of this approach to address problems with uncertainty.

# The Rational Agent approach

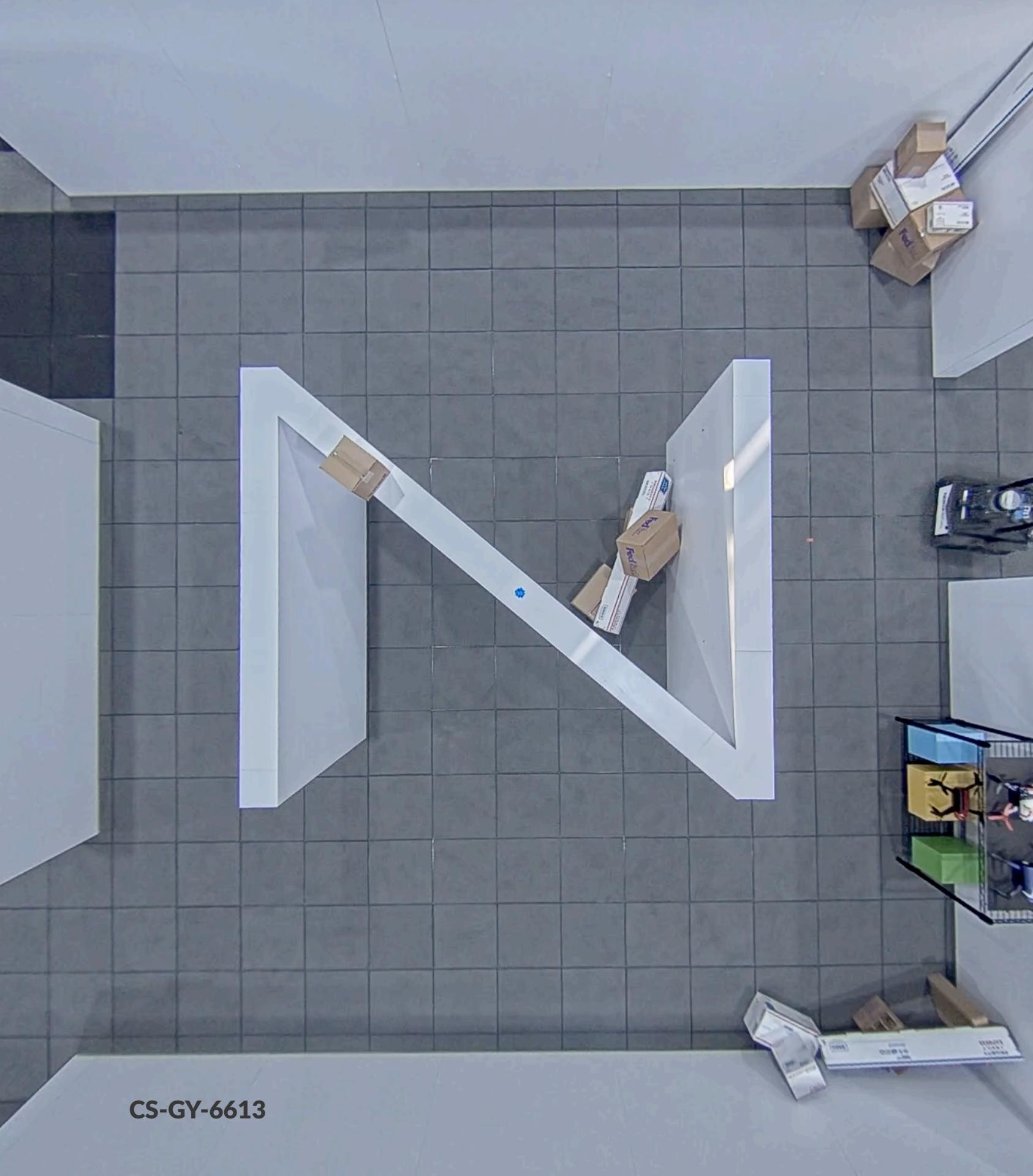
- A rational agent acts to achieve the best outcome according to an internal performance measure given the evidence provided by the percept sequence and prior knowledge the agent has. Does this definition remind you of something ?
  - Bayes Theorem
- The rational approach encompasses the syllogism and Turing-test approaches. We still need provably correct inference and the formal representations of logic as well as the ability to perceive, communicate, and learn to achieve a good outcome.
- But we need to **generalize these approaches to include "good-enough" inference and adaptation to the changing environment contexts** that the agent is facing without giving up on the mathematical formality that *utility theory* allows us to design such agents.
- The agent facing a fire is an instructive example. There maybe no time for optimal reasoning to a conclusion (e.g. run) but a simple reflexive plan can offer the best outcome.

# AI Problem Specification for Rational Agents

- Performance
- Environment
- Actuators
- Sensors
- Sometimes we have no physical environments but virtual ones. Its not only the various chatbots and smart speakers of today - its also simulated task environments where we host the *digital twin* models of everything we care about.

# The Bell Labs AI Research Environment

- This maze was constructed in Bell Labs for the [Unix 50 competition](#) to demonstrate the ability of intelligent robotic agents to search for unknown objects in it.
- It is constantly enhanced with more challenging parts such as collaboration between robots.
- A digital twin of the robot and the environment is modeled inside Unity.



# A real self-driving car circa 2018 (Udacity's CARLA)

