ENSI CAEN
ÉCOLE PUBLIQUE D'INGÉNIEURS
CENTRE DE RECHERCHE

01
Chapiter

Intelligent Agents

2I1AE1: Artificial Intelligence

Régis Clouard, ENSICAEN - GREYC

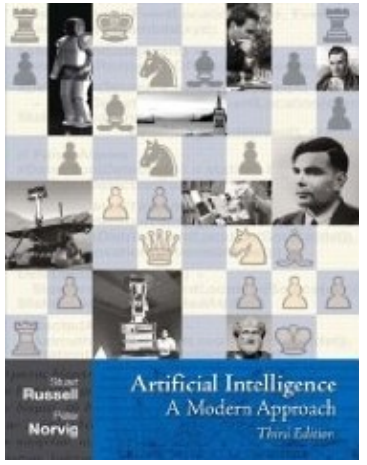
"There is not a solid definition of intelligence.
The problem is that we cannot yet characterize what kinds of computational procedures we want to call intelligent."
John McCarthy

Course Information

- Workload
  - Lectures 8hrs
  - Tutorials 6hrs
  - Labs 12hrs
- Work and Grading
  - 4 labs (Single – No mark)
  - 1 homework assignment (50%):
    - Pair
    - Defense
    - **Academic honesty and plagiarism**: We trust you all to submit your own work only; please don't let us down. If you do so, we will pursue the strongest consequences available to us.
  - 1 final exam (50%)

Textbook

- Handout on Moodle (course #47):
  - https://foad.ensicaen.fr/course/view.php?id=47
- Course textbook:
  - Stuart J. Russell, Peter Norvig, "*Artificial intelligence: a modern approach*", Prentice Hall, 2020 (4th edition).
- Other widely used AI textbooks
  - P. Marquis, O. Papini, H. Prade, "*Panorama de l'intelligence artificielle : ses bases méthodologiques, ses développements*", Cépaduès éditions, 2016 (3 volumes).
  - Peter H. Winston, "*Artificial Intelligence*", Addison-Wesley, 1992.
  - Jean-Paul Haton et al., "*Le raisonnement en Intelligence Artificielle*", InterEditions, 1991.
  - Jean-Louis Laurière, "*Intelligence artificielle : Résolution de problèmes par l'homme et la machine*", Eyrolles, 1987. http://code.google.com/p/jll-ia-1987/downloads/list.



[Moodle]

In this chapter

- What is Artificial Intelligence?
- Designing Intelligent Agent

- Preliminary questions:
  - Do you believe in IA (ie, do you consider ChatGPT to be intelligent)?

What is Artificial Intelligence?

- Artificial Intelligence is one of scientific disciplines, with genetics, that raises the most fears and fantasies.
  - Reason: The monopoly on human intelligence is being challenged.

- What is Artificial Intelligence (**AI**)?
  - Artificial intelligence is based on the assumption that the process of human thought can be mechanized.
    - If nature can be written by mathematical language, why not intelligence?
- First definition attempt:
  - Artificial Intelligence is the branch of computer science concerned with making machines capable of imitating and performing human cognitive functions.
    - Aporetic definition

Human vs. Computer

- Collective unconscious:
  - Intelligence is what a machine cannot do:
    "I would be dehumanized if it were proved that an artificial intelligence exists"-- Jacques Arsac, 1987.
  - **Human being is:**
    - Intelligent, fallible, abstraction ability.
  - **Computer is:**
    - Dummy, rigorous, calculation capability.

Human vs. Computer

- Two examples in pattern recognition
- Trivial problem for human: What is the paragraph about?

  Sleon une édtue de l'uvinertisé de Cmabrigde, l'odrre des ltteers dans un mto n'a pas d'ipmrotncae, la suele coshe ipmrotnate est que la pmeirère et la drenèire soit à la bnnoe pclae. Le rsete peut êrte dnas un dsérorde ttoal et vuos puoevz tujoruos lrie snas porlblème. C'est prace que le creaveu hmauin ne lit pas chuaqe ltetre elle-mmêe, mias le mot cmome un tuot.
- Trivial problem for computer: How many 'F' in this sentence?

FINISHED FILES ARE THE RESULT
OF YEARS OF SCIENTIFIC
STUDY COMBINED WITH THE
EXPERIENCE OF YEARS

Human vs. Computer

- Trivial problem for computer:

$$\sqrt[13]{\begin{array}{l}858990809132578040222986483937114579787851376179717518054315065077274\\063859398978034751926880410465769118780136223581696838195379767325980\\998981723892608905338894073093288513224763223516647191959469 31\end{array}}$$

  - Absolutely impossible for human.
  - However

    1990, Alexis Lemaire, a french 24-year-old student in computer science, successfully performed this most difficult mental calculation in the world: in 4 minutes and 27 seconds.

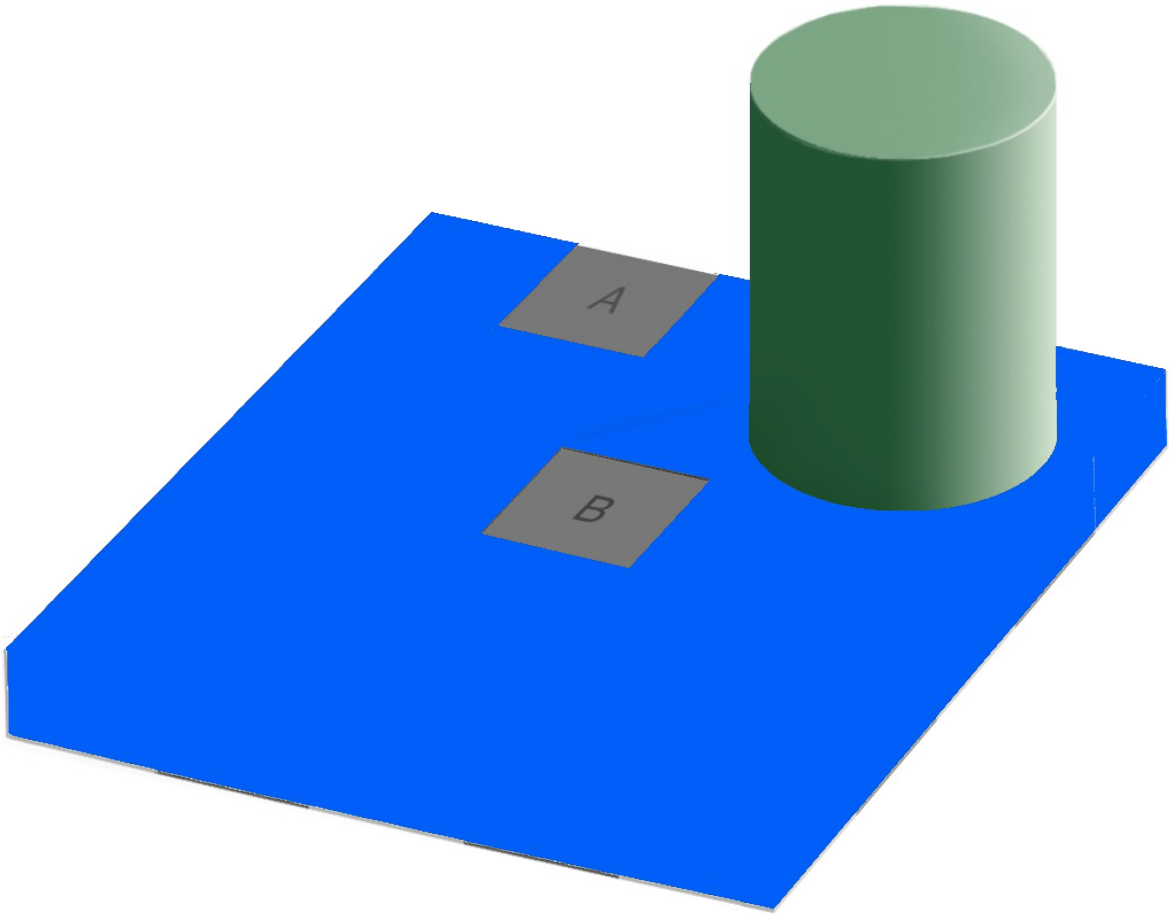
    Solution: 2,396,280,083,911,011

Human vs. Computer

- Trivial problem for human:
  - Where is the dog?

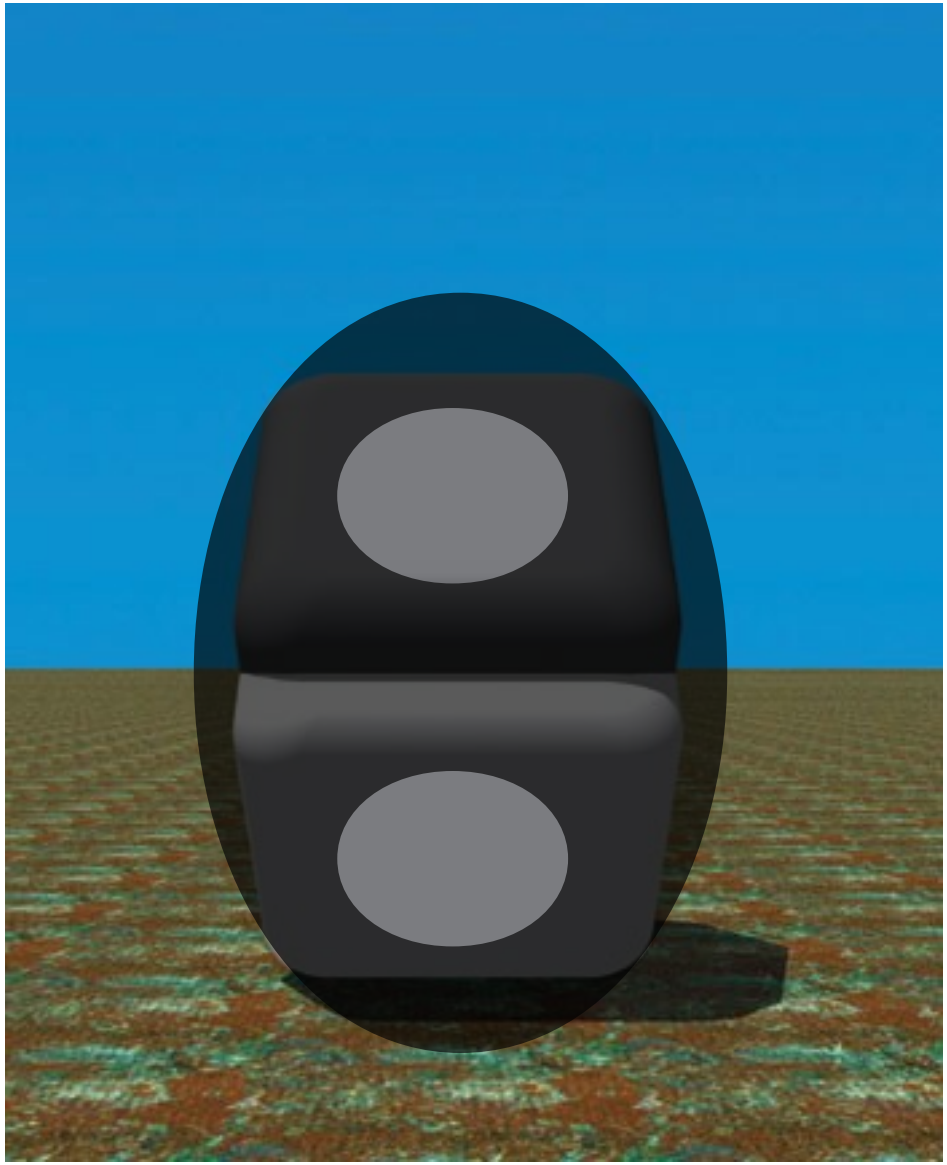
Human vs. Computer

- Trivial problem for computer:
  - What is the clearest square A or B?



Adelson's chessboard

Human vs. Computer

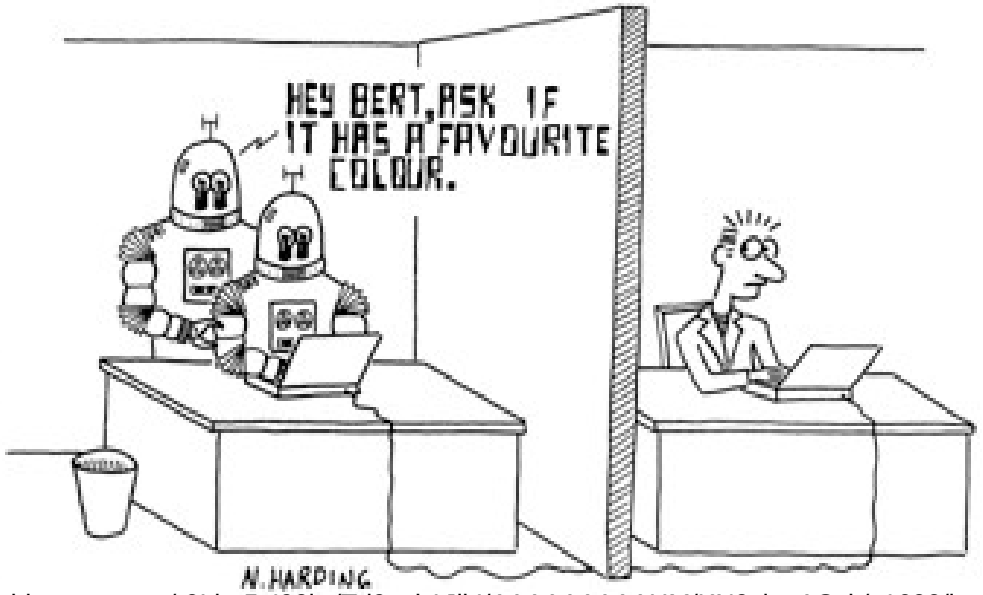


Human vs. Computer

- Preliminary conclusion:
  - Human and computer have opposite abilities.
    - Human is mostly irrational and fast at recognizing a solution.
    - Computer is rational and fast at calculating a solution.
  - Conclusion: AI is impossible!
- However, computer solves problems that are considered as human's prerogative
- So the question remains: what is AI?
  - The previous definition is tied to the definition of human cognitive functions (intelligence).
    - But, intelligence is impossible to define.

The Turing's Test

- If we fail to define intelligence, why not try to measure what is intelligent?
  - Turing test : "*Computing Machinery and Intelligence*", 1950.
    - *Loebner price since 1991 → 2019.*

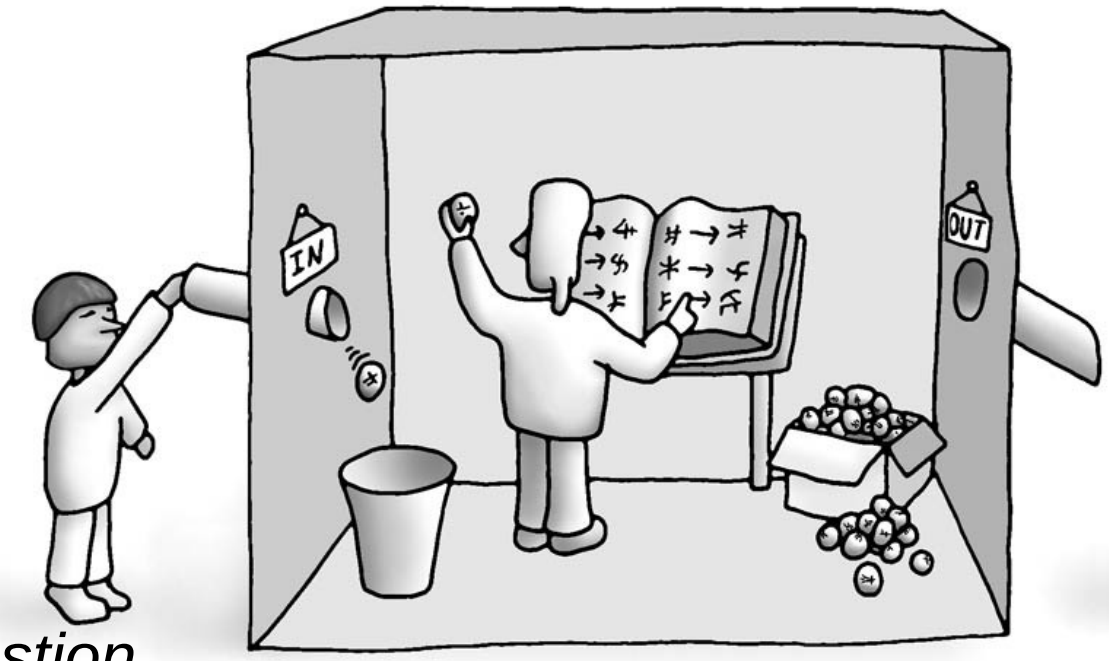


https://2.bp.blogspot.com/-0Lbu5vI90iw/Td9ovk1JjkI/AAAAAAAAAHM/KN0clnoACck/s1600/images.jpg

- Since 2011, multiple programs pass the test
  - Bias: intelligence refers to conversational skills.
  - But is it necessary that AI system also thinks humanly?

Turing's Test Criticisms

- John Searle's Chinese Room.



Question (Chinese)

Answer (Chinese)

Credit: jolyon.co.uk

- By mechanically following these rules, the person produces English translations of the Chinese text.
- But : is it intelligent behavior? → polling

Two points of view

- IA opponents (Searle) say **no**:
  - It cannot be intelligence without understanding?
  - Semantics cannot be reduced to symbol manipulation according to syntactic rules.
- IA partisans (Turing) say **yes**:
  - The system composed of the person and the book is intelligent because it answers the questions.
  - *"Saying Deep Blue doesn't really think about chess is like saying an airplane doesn't really fly because it doesn't flap its wings."* Drew McDermott.
- ChatGPT says:
  - "I wouldn't say I'm intelligent in the human sense. I process and generate text based on patterns in the data I was trained on, which allows me to provide information and engage in conversation. But I don't have understanding, emotions, or consciousness like a person does. So, while I can simulate intelligent responses, it's not the same kind of intelligence you might think of!"

Tentative Definition

- Previous discussion leads to two definitions of AI:
  - **AI is the science of making machines that:**

Think like people	Think rationaly
Act like people	Act rationaly

Strong IA Modeling human behavior	**Weak IA** Imitating human behavior

  - Searle has no quarrel with weak AI which he regards as essentially a technological undertaking. He does not believe in Strong AI.

Weak IA: Rational Decisions

- We'll use the term rational in a very specific technical way:
  - Rationality only concerns what decisions are made not the thought process behind them.
  - Goals are expressed in terms of the **utility of outcomes** (an objective function).
  - Being rational means maximizing your expected utility.

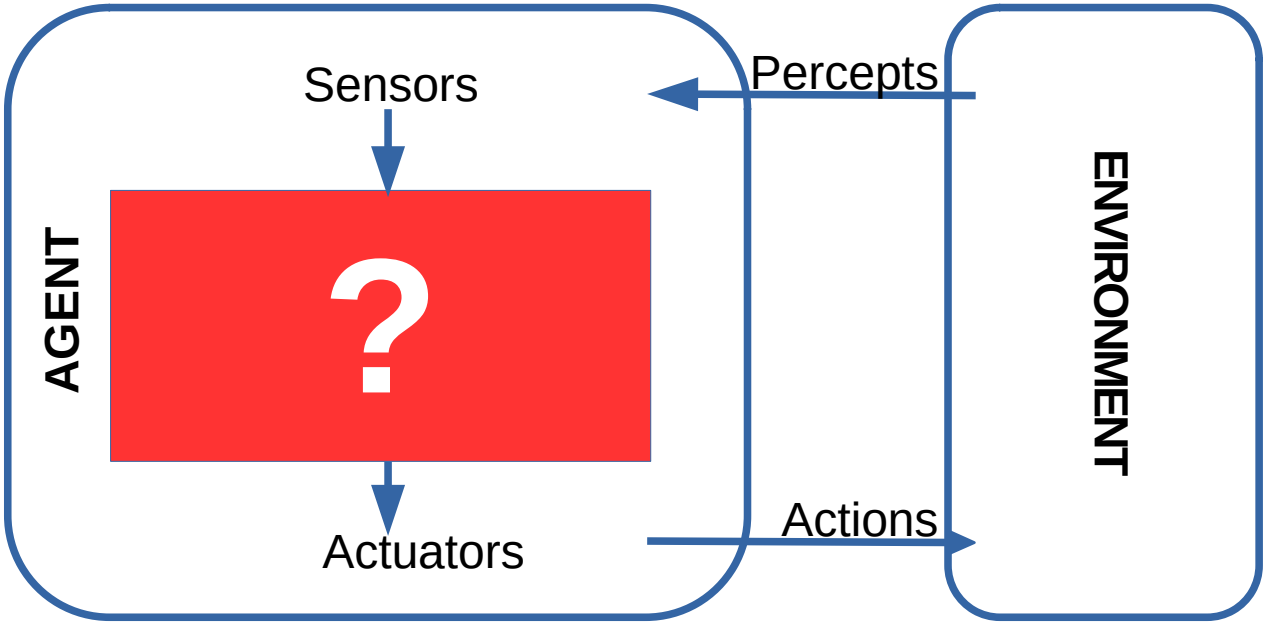
- A better title for this course would have been:
  - **Computational Rationality**
    - Study of algorithms for maximizing the expected utility
    - Differ from classical algorithm : in most problem, computational limitations make **perfect rationality unachievable**.

A (Short) History of AI

- 1940: Early days – Mathematical models of reasoning and thought
  - 1943: Mc Colloch & Pitts: boolean circuit model of brain
  - 1945: ENIAC first computer
  - 1950: Turing's: "Computing Machinery, and Intelligence"
- 1950: Golden years
  - 1950's: Early AI programs, including
    - Samuel's checkers program: learning by reinforcement
    - Newell & Simon's Logic Theorist (proved 38 out of 52 of Russell's theorems)
  - 1956: Darmouth's meeting: birth of "Artificial Intelligence" (John Mc Carthy)
  - 1959: Newell & Simon General Problem Solver
  - 1961: Rosenblatt, neural networks
- 1970: Knowledge-based approaches
  - 1969-79: Early development of knowledge-based systems (MYCIN, DENDRAL)
  - 1980-88: Expert systems industry booms (eg. DEC: 40 experts systems)
  - 1988-93: Expert system industry busts
- 1990: Statistical approaches
  - Learning approaches : Resurgence of neural networks
  - Focus on uncertainty (probabilities)
- 2000: Distributed and enactive approaches
  - Multi-agent systems
  - Interactive systems
- 2016: Deep learning (neural networks)

General goal: Designing Rational Agent

- An agent is an entity that **perceives** and **acts** in an **environment**.
- A rational agent selects actions that maximize its (expected) **utility**.
  - Utility is based on **performance measure**
- PEAS orients towards techniques for selecting rational actions
  - **P**erformance measures
  - **E**nvironment
  - **A**ctions
  - **S**ensors



Example of Intelligent Agent. Self-driving car

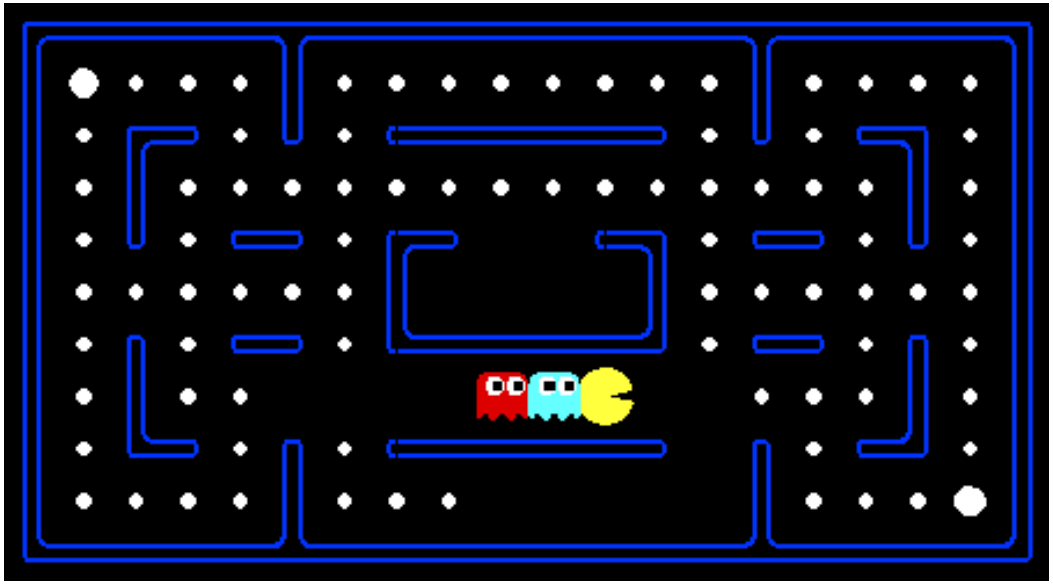
- Environment
  - France urban streets, freeways, traffic, pedestrians, weather, …
- Sensors
  - Video, sonar, speedometer, odometer, engine sensors, keyboard input, microphone, GPS, …
- Actuators
  - Steer, accelerate, brake, horn, speak/display, …
- Performance measure
  - Maintain safety, reach destination, maximize profits (fuel, tire wear), obey laws, provide passenger comfort, …

Example of Intelligent Agent. Medical Diagnosis System

- Environment
  - Patient, hospital, staff.
- Sensors
  - Keyboard (entry of symptoms findings, patient's answers).
  - Data files (patient files).
- Actuators
  - Screen display (questions, tests, diagnoses, treatments).
- Performance measure
  - Healthy patient, minimize costs, lawsuits.

Example of Intelligent Agent. Pac-Man Agent

- Environment
  - Maze: array of cells.
- Sensors
  - Location.
  - Cell contents.
  - Ghosts location.
- Actuators
  - Go north, south, east, or west.
- Performance measure
  - Number of points, death.



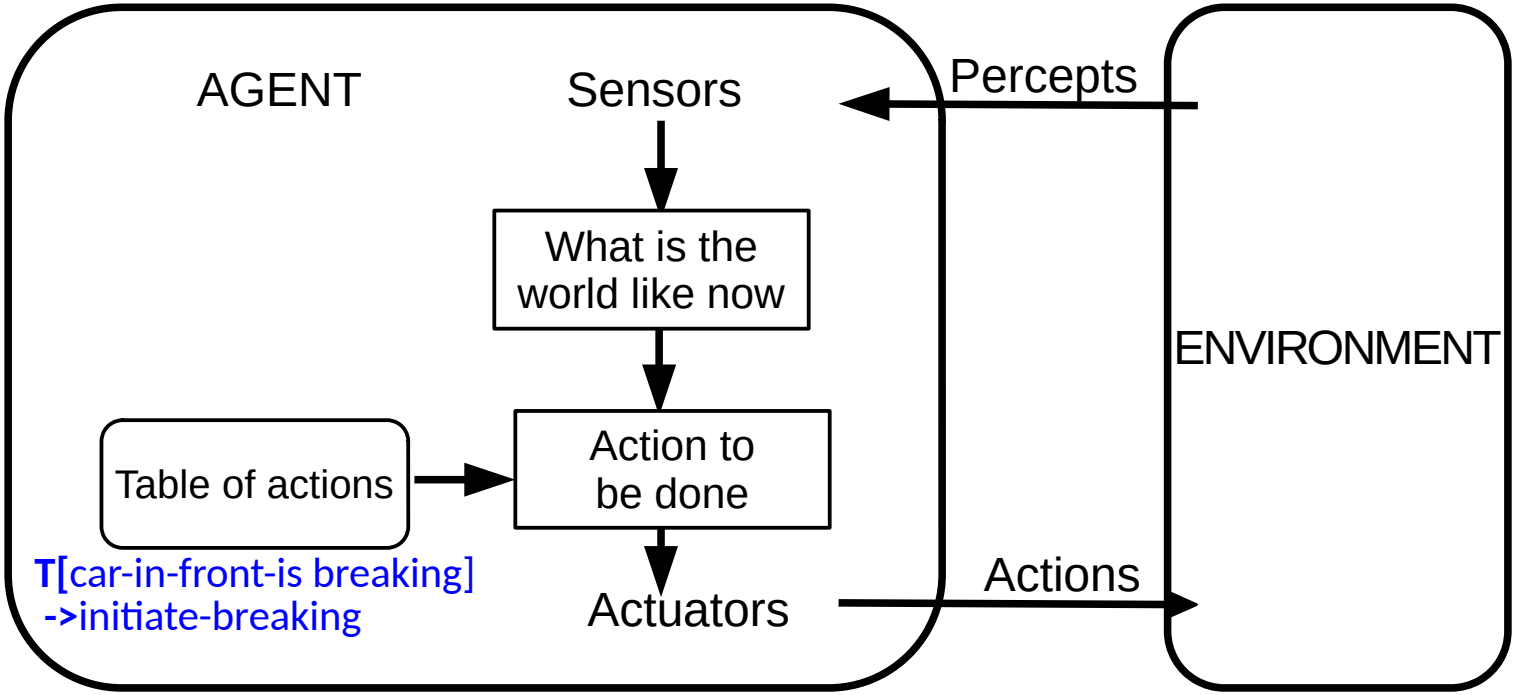
[pacman.mp4]

2 classes of agents

- Two classes of rational agents:
  - Reflex agent
  - Utility-based agents

1. Reflex Agent

- Rather dummy agent.
  - Selects actions based on current percept.
    - Ignoring rest of percept history.
  - May have memory or model of the world's current state.
- Acts on how the world IS.
  - Do not consider the future consequences of their actions.



Reflex Agent

- A reflex agent works by finding a rule whose condition matches the current situation (as defined by the percept) and then doing the action associated with that rule.

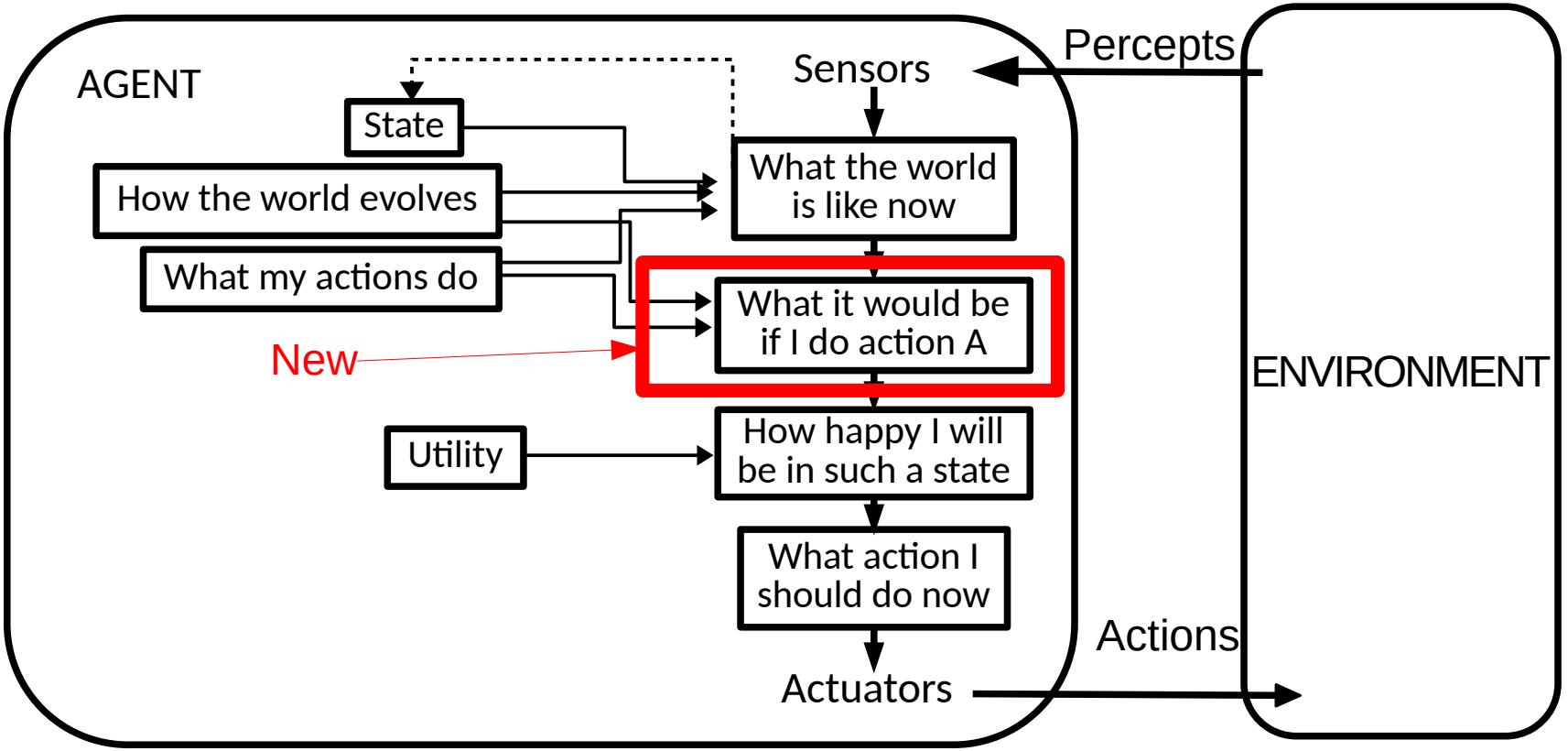
```
function REFLEX-AGENT(percept) returns an action
   static: rules, a set of condition–action rules

   state ← INTERPRET-PERCEPTS(percept)
   action ← CHOOSE-BEST-ACTION(state, rules)
   return action
```

- Applicability: very low.

[pacman-reflex-agent.sh]

2. Utility-based Agent

- Plans ahead.
  - Decisions based on *hypothesized* consequences of actions.
- **Acts on how the world WOULD BE : ask "What if".**
  - Must have a model of how the world evolves in response to actions.
  - Always choose action that maximizes its own utility function.



Utility-based Agent


```
function INTELLIGENT-AGENT(percept, goal) returns an action
   static: state, the agent's memory of the world state

   state ← UPDATE-STATE-FROM-PERCEPTS(state, percept)
   action ← CHOOSE-BEST-ACTION(state)
   state ← UPDATE-STATE-FROM-ACTION(state, action)
   return action
```

[pacman-utility-based-agent.sh]