Introduction to Artificial Intelligence & Intelligent Agents

Tian-Li Yu

Taiwan Evolutionary Intelligence Laboratory (TEIL)
Department of Electrical Engineering
National Taiwan University
tianliyu@ntu.edu.tw

Readings: AIMA Chapters 1 & 2

Outline

What Is AI?

Definitions of AI (textbook)

	Human	**Rational**
Think	Thinking humanly	Thinking rationally
Act	Acting humanly	Acting rationally

- The textbook advocates acting rationally.

Strong AI vs. Weak AI (John Searle)

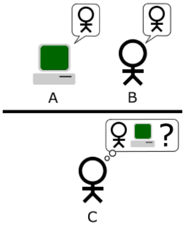
Strong AI : A physical symbol system can have a mind and mental states.

Weak AI : A physical symbol system can act intelligently.

- The textbook advocates weak AI.

Acting Humanly: Turing Test

- Turing (1950) Computing machinery and intelligence: "Can machines think?" $\rightarrow$ "Can machines behave intelligently?"
- Operational test for intelligent behavior:

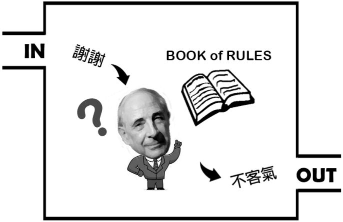


  - ELIZA (1965).
  - Mitsuku (2013 Loebner prize winner).
  - CAPTCHA.
- Involving
  - Natural language processing
  - Knowledge representation
  - Automated reasoning
  - Machine learning

Thinking Humanly: Cognitive Modeling

- 1960s cognitive revolution: Information-processing psychology.
- Requires scientific theories of internal activities of the brain.
- How to validate? Requires
    1. Predicting and testing behavior of human subjects (top-down), or
    2. Direct identification from neurological data (bottom-up)
- Both approaches (roughly, cognitive science and cognitive neuroscience) are now distinct from AI.

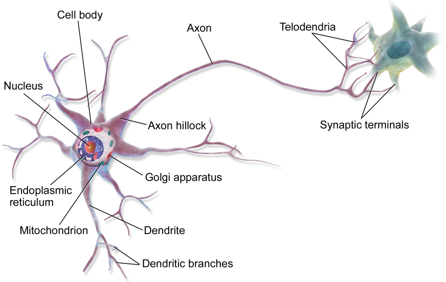
Chinese Room Argument by John Searle



- Suppose it is possible to have a program $\mathcal{P}$ that is sufficient for understanding of Chinese.
- In principle a person in the Chinese room can carry out $\mathcal{P}$.
- But such a person would not understand Chinese.
- So $\mathcal{P}$ is not sufficient for producing understanding of Chinese.
- So there is no program sufficient for producing understanding of Chinese.

Neuroscience

	Supercomputer	Personal Computer	Human Brain
Computational units	10^{12} transistors	10^{9} transistors	10^{11} neurons
Storage units	10^{14} bits RAM	10^{11} bits RAM	10^{11} neurons
	10^{15} bits disk	10^{15} 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: Laws of Thought

- Aristotle: What are correct arguments/thought processes?
- Several Greek schools developed various forms of logic.
- Direct line through mathematics and philosophy to modern AI.
- Problems:
  1. Not all intelligent behavior is mediated by logical deliberation.
  2. What is the purpose of thinking? What thoughts should I have?

Acting Rationally: Rational Agent

- Rational behavior: Do the right thing.
- The right thing: Expected to maximize goal achievement, given the available information.
- Doesn't necessarily involve thinking – *e.g.*, blinking reflex – but thinking should be in the service of rational action.

Rational Agent

- An agent is an entity that perceives and acts.
- Abstractly, an agent is a function from percept histories to actions.
- For any given class of environments and tasks, we seek the agent (or class of agents) with the best performance.
- Computational limitations $\rightarrow$ limited rationality.

Foundations of AI

Philosophy	logic, reasoning, learning, language, rationality
Mathematics	formal representation of algorithms, decidability, tractability.
Economics	utility, decision theory
Neuroscience	physical substrate for mental activity
Psychology	phenomena of perception and experimental techniques
Computer engineering	building fast computers
Control theory	design systems that maximize an objective function over time
Linguistics	knowledge representation, grammar

AI History

1943	Boolean circuit model of brain.
1950	Turing's computing machinery and intelligence.
1956	Dartmouth meeting: artificial intelligence.
1952-69	Look, Ma, no hands! Perceptron convergence theorem.
1950s	Early AI programs.
1965	Robinson's complete algorithm for logical reasoning.
1966-73	AI discovers computational complexity. Neural network research almost disappears. Machine evolution.
1969-79	Early development of knowledge-based systems
1980–	AI becomes an industry
1986–	Return of neural networks
1987–	AI adopts scientific methods.
1995–	The emergence of intelligent agents
2001–	Availability of large data sets

State of the Art

Game Playing: DEEP BLUE defeated the reigning world chess champion Garry Kasparov in 1997.

Spam Filtering: Mail clients have embedded spam filtering since 2000.

Robotic Vehicles: CMU's BOSS won the Urban Challenge in 2007.

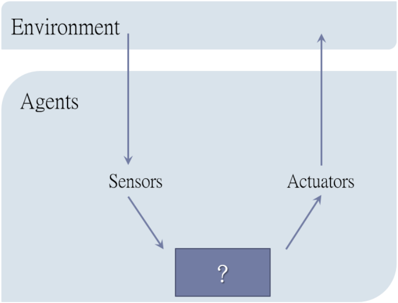
Speech Recognition: WATSON defeated *Jeopardy!* champions in 2010.

Robotic: ROBO-ONE, RoboCup

Agents

- An agent is anything that can be viewed as perceiving its environment through sensors and acting upon that environment through actuators.
- Human agents: eyes, ears, and other organs for sensors; hands, legs, mouth, and other body parts for actuators.
- Robotic agents: cameras and infrared range finders for sensors; various motors for actuators.

Agents and Environments

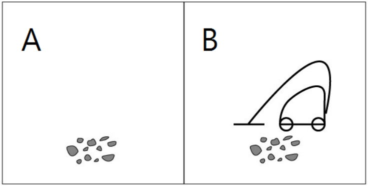


- Agents include humans, robots, softbots, thermostats, etc.
- The agent function maps any percept sequences to an action.

$$f : \mathcal{P}^* \rightarrow \mathcal{A}$$

- An agent program is the implementation of an agent function.

Vacuum Cleaner



Percepts: Location and Contents, *e.g.* [B, Dirty].

Actions: Left, Right, Suck, NoOp.

Reflex-Vacuum-Agent(*location*, *status*)

```
if status == Dirty
        return Suck
elseif location == A
        return Right
elseif location == B
        return Left
```

A Vacuum-Cleaner Agent

Percept Sequence	Action
[A,Clean]	Right
[A,Dirty]	Suck
[B,Clean]	Left
[B,Dirty]	Suck
[A,Clean],[A,Clean]	Right
[A,Clean],[A,Dirty]	Suck
⋮	⋮
[A,Clean],[A,Clean],[A,Clean]	Right
[A,Clean],[A,Dirty],[A,Dirty]	Suck
⋮	⋮

- What is the right way to fill a table or to write a function?
- What makes an agent good or bad, intelligent or stupid?

Rationality

- Performance measure according to
  - What one actually wants in the environment, rather than
  - How one thinks the agent should behave.
- A rational agent maximizes the expected performance.
- Rational $\neq$ Omniscience
  - Percepts may not supply all relevant information.
- Rational $\neq$ Clairvoyant
  - Action outcomes may not be as expected.
- Rational $\Rightarrow$ information gathering, exploration, learning, autonomy.

PEAS

- To design a rational agent, we must specify the task environment.

Agent Type	**P**erformance Measure	**E**nvironment	**A**ctuators	**S**ensors
Taxi driver	safety, destination, profits, legality, comfort	Roads, pedestrians, customers	Steering, accelerator, brake, horn	Accelerometers, camera, engine sensors, GPS

Environment Types

	Chess	Poker	Taxi	Image analysis
Observability				

Environment Types

	Chess	Poker	Taxi	Image analysis
Observability	Fully	Partially	Partially	Fully
Agents				

Environment Types

	Chess	Poker	Taxi	Image analysis
Observability	Fully	Partially	Partially	Fully
Agents	Multi competitive	Multi competitive	Multi cooperative	Single
Deterministic /stochastic				

Environment Types

	Chess	Poker	Taxi	Image analysis
Observability	Fully	Partially	Partially	Fully
Agents	Multi competitive	Multi competitive	Multi cooperative	Single
Deterministic /stochastic	Deterministic	Stochastic	Stochastic	Deterministic
Episodic /sequential				

Environment Types

	Chess	Poker	Taxi	Image analysis
Observability	Fully	Partially	Partially	Fully
Agents	Multi competitive	Multi competitive	Multi cooperative	Single
Deterministic /stochastic	Deterministic	Stochastic	Stochastic	Deterministic
Episodic /sequential	Sequential	Sequential	Sequential	Episodic
Static /dynamic				

Environment Types

	Chess	Poker	Taxi	Image analysis
Observability	Fully	Partially	Partially	Fully
Agents	Multi competitive	Multi competitive	Multi cooperative	Single
Deterministic /stochastic	Deterministic	Stochastic	Stochastic	Deterministic
Episodic /sequential	Sequential	Sequential	Sequential	Episodic
Static /dynamic	Static/Semi	Static	Dynamic	Semi
Discrete /continuous				

Environment Types

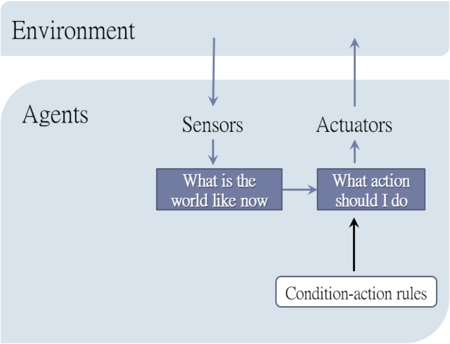
	Chess	Poker	Taxi	Image analysis
Observability	Fully	Partially	Partially	Fully
Agents	Multi competitive	Multi competitive	Multi cooperative	Single
Deterministic /stochastic	Deterministic	Stochastic	Stochastic	Deterministic
Episodic /sequential	Sequential	Sequential	Sequential	Episodic
Static /dynamic	Static/Semi	Static	Dynamic	Semi
Discrete /continuous	Discrete	Discrete	Continuous	Continuous

- The real world is partially observable, stochastic, sequential, dynamic, continuous, multi-agent.

Agent Types

- Four basic types in order of increasing generality
  1. Simple reflex agents.
  2. Model-based agents.
  3. Goal-based agents.
  4. Utility-based agents.

Simple Reflex Agents



SIMPLE-REFLEX-AGENT(*percept*)

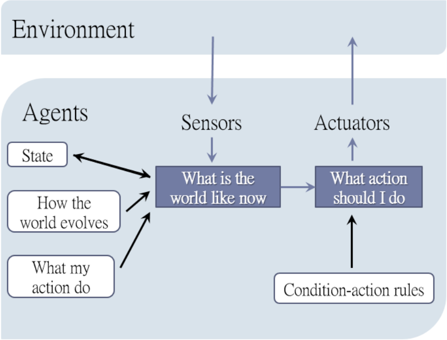
state = INTERPRET-INPUT(*percept*)
rule = RULE-MATCH(*state*, *rules*)
action = *rule.action*
return *action*

Simple Reflex Agents

- REFLEX-VACUUM-AGENT is a simple reflex agent.
- Actions rely purely on condition-action rules:

  if condition **then** action.
- Also called memory-less or state-less
- Works *only if the correct decision can be made on the basis of only the current percept*.
- Works only if the environment is fully observable.
- Often trapped in infinite loops if the environment is partial observable.

Model-Based Reflex Agents



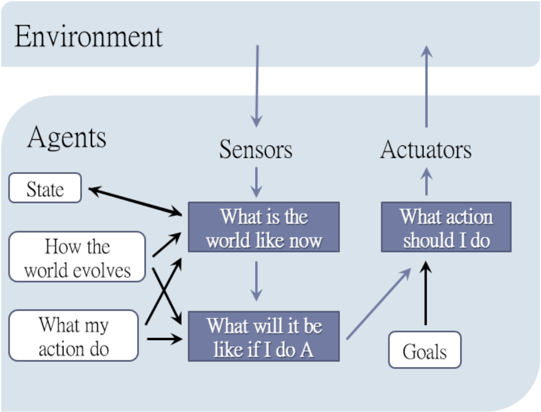
MODEL-BASED-REFLEX-AGENT(*percept*)

$state = \textsc{Update-State}(state, action, percept, model)$
$rule = \textsc{Rule-Match}(state, rules)$
$action = rule.action$
return *action*

Model-Based Reflex Agents

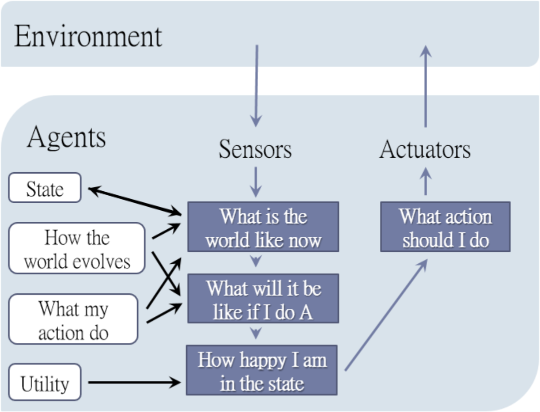
- Handle partial observability by keeping track of the part of the world it can't see now.
- Maintain internal states to model the world.
- The model of the world represents the agent's best guess(es), can't be exact.
- Internal states can also be used to maintain the status of the agent instead of the world.

Goal-Based Agents



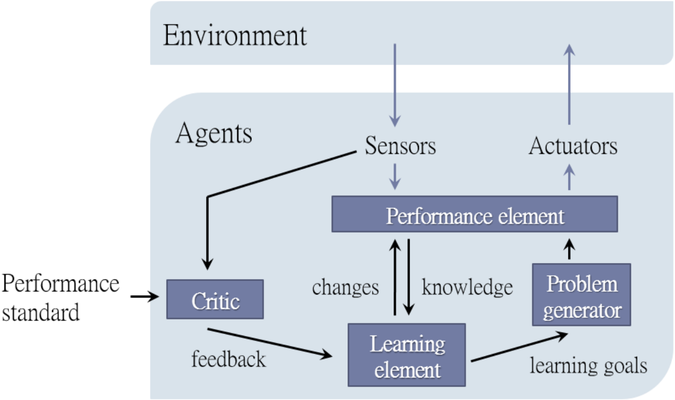
- Instead of using condition-action rules, the agent uses goals to decide what action it does.
- Search (Chapters 3, 4 and 5) and Planning (Chapters 10 and 11).

Utility-Based Agents



- Utility function: Happiness of the agent.
- Maximizing the expected utility.

Learning Agents



- Learning from rewards (or penalty).
- Learning techniques form another field called machine learning.

Summary

- Strong vs. weak AI.
- Agents interact with environments through actuators and sensors.
- The agent function describes what the agent does in all circumstances.
- The performance measure evaluates the environment sequence.
- A perfectly rational agent maximizes expected performance.
- Agent programs implement (some) agent functions.
- PEAS descriptions define task environments.
- Different environments & agent types.
- All agents can improve their performances through learning.