Title: Historical notes & Intelligent Agents
AIMA: Chapters 1 & 2

Introduction to Artificial Intelligence
CSCE 476-876, Spring 2000
`www.cse.unl.edu/~cse476`

Berthe Y. Choueiry (Shu-we-ri)
`choueiry@cse.unl.edu, (402)476-5444`

Announcements

- Course: Philosophy of Psychology, PHIL 923
  Instructor: Dr. Barbara Von Eckardt
  Textbooks: What is Cognitive Science? (Von Eckardt)
  Understanding Cognitive Science (Bickle?)
  Psychoneural Reduction (Bickle)
  Time/Place: Monday 3:30-5:35, Oldfather 1007
  $\longrightarrow$ Corey Maley
- Dropping/joining CSCE476/876: cse476@cse.unl.edu
- Reading material: read before *and* after the class
- Homework: start as early as possible, acknowledge help
- Glossary: any reference you like, *always* cite your sources

Summary

- Goal of AI:
  - Understand intelligent entities
  - Build intelligent entities
- Multidisciplinary:
  - philosophy, psychology, cognitive science
  - mathematics, engineering, computer science
- Study:
  - general mechanisms
  - specific tasks
- Dimensions:
  - Concern: thought+mental processes vs. action+behavior
  - Performance: human-like vs. rational

Quick historical note (I)

1956: McCarthy organizes a two-month workshop in Dartmouth
no breakthrough, united major players, term coined

1952-1969: Early enthusiasm and great expectations
General Problem Solver (Newell & Simon),
chess program (learning disproved "computers do what they are told to do"),
LISP in 1958, time sharing, DEC, principles of knowledge representation and reasoning,
Split: logic (neat) vs. anti-logic (scruffy, clumsy),
Microworlds (*e.g.*, block world: → vision, constraint propagation, NL understanding, planner),
Neural Nets, etc.

Quick historical note (II)

1966-1974: hard reality. Too big claims, wildly optimistic

1- Systems work on 1 or 2 examples, failed otherwise

NLP: Russian → English. The spirit is willing but the flesh is weak → the vodka is good but the meat is rotten

2- Intractability: difficulty to scale up, handle combinatorial explosion

UK → report Lighthill in 1973, etc..

Neural Net almost disappears

1969-1979: Knowledge-based systems

Knowledge-based system (DENDRAL) expert knowledge

expert systems (MYCIN), certainty factors, frames (OO!)

Quick historical note (III)

1980-1988: AI becomes an industry
R1 at DEC, Fifth Generation project.
S/W Carnegie Group, Inference, Intellicorp, Teknowledge
H/W: Lisp Machines, TI, Symbolics, Xerox.

1986: Expert Systems failed; Neural Networks returned. Late 80's-early 90's: AI winter

1988-: Resurgence of probabilistic & decision theory (UAI).
"Nouvelle AI:" ALife, GAs, soft computing

Present: Big changes. Claims more rigorously supported: empirically or theoretically

Intelligent Agents

- Definition and components (PAGE)
- Types of Intelligent Agents
- Types of environments

Agent

$$\text{Anything that} \begin{cases} \text{perceives its environment through sensors} \\ \text{acts upon its environment through effectors} \end{cases}$$

Include 8 definitions in Fig. 2.1

Human: sensors? effectors?

Robot: sensors? effectors?

Software: sensors? effectors?

Goal of AI: build <u>rational</u> agents. Rational = ?

What is "rational" depends on:

1. Performance measures (how, when)
2. Perceptual history: everything agent has perceived so far
3. What the agent knows about the environment
4. Actions the agent can perform

Ideal rational agent: PAGE

- For each perceptual history, **P**ercepts
- chooses an action **A**ctions
- to maximize performance measures ($\rightarrow$ key!) **G**oals
- given perceptual history and built-in knowledge **E**nvironment

Examples

Include Fig. 2.3

PAGE

The task of designing an automated taxi:

Percepts:

Actions:

Goals:

Environment:

PAGE: automated taxi

Percepts: video, accelerometers, engine sensors, keyboard, microphone, GPS, ...

Actions: steer, accelerate, brake, horn, speak/display, ...

Goals: safety, reach destination, maximize profits (fuel, tire wear), obey laws, passenger comfort, ...

Environment: US urban streets, freeways, traffic, pedestrians, weather, customers, ...

$\longrightarrow$ Different aspects of driving, different types of agent programs

Types of Agents

Four, in order of increasing generality:

1. Simple reflex agent
2. Agents that keep track of the world
3. Goal-based agents
4. Utility-based agents

Simple reflex agent

• Simple look-up table, mapping percepts to actions, is out of question (too large, too expensive to build)

• Many situations can be summarized by condition-action rules (humans: learned responses, innate reflexes)

Include Fig. 2.7

Rectangles: agent's internal state Ovals: background information

Implementation: easy; Applicability: narrow

Agents that keep track of the world

- Sensory information alone is not sufficient
- Need to keep track of how the world evolves

(evolution: independently of agent, or caused by agent's actions)

Include Fig. 2.7

Goal-based agents

- State & actions don't tell <u>where</u> to go
- Need goals to build sequences of actions (planning)

Include Fig. 2.11

Goal-based: uses the same rules for different goals

Reflex: will need a complete set of rules for each goal

Utility-based agents

- Several action sequences to achieve some goal (binary process)
- Need to <u>select</u> among actions & sequences. Preferences.
- Utility: State $\rightarrow$ real number (express degree of satisfaction, specify trade-offs between conflicting goal)

Include Fig. 2.11

Environment (I)

typical quizz question

Accessible: sensors can detect all aspects of the world

Effectively accessible: relevant aspects

Deterministic:

Next state determined by current state and agents' actions

Inaccessible + deterministic appears nondeterministic from the agent's view point

Episodic: Agent's experience divided into episodes; subsequent episodes do not depend on actions in previous episodes

Static vs. dynamic:

Dynamic: Environment changes while agent is deliberating

Semi-dynamic: environment static, performance scores dynamic

Discrete vs. continuous: Finite number of precepts, actions

Chess? Taxi driving? hardest case?

Environment (II)

Hardest case: inaccessible, nondeterministic, nonepisodic, dynamic, continuous

	Solitaire	Backgammon	Internet shopping	Taxi
Accessible				
Deterministic				
Episodic				
Static				
Discrete				

Answers depend on how you define/interpret the case

Episodic: chess tournament

List of terms for glossary:

Due date: Wed, Jan 19, 2000

Agent
Autonomy (autonomous agent)
Chinese room experiment
Condition-action rule
Occam's (or Ockham's) razor
Planning
Restricted rationality
Total Turing test
Turing machine
Turing test
Syllogism
Utility