CSC384: Intro to Artificial Intelligence

Paulo Pacheco
- ppacheco@cs.toronto.edu
- Office Hours: TBA (watch Web page), by appt.
Artificial Intelligence (AI)

- What is AI?
- What is intelligence?
- What features/abilities do humans (animals? animate objects?) have that you think are indicative or characteristic of intelligence?

- abstract concepts, mathematics, language, problem solving, logical reasoning, emotions, morality, ability to learn/adapt, etc…
## Some Definitions (Russell + Norvig, 1995)

<table>
<thead>
<tr>
<th>The exciting new effort to make computers that think... machines with minds in the full and literal sense [Haugeland 85]</th>
<th>The study of mental faculties through the use of computational models [Charniak &amp; McDermott 85]</th>
</tr>
</thead>
<tbody>
<tr>
<td>[The automation of] activities that we associate with human thinking, such as decision making, problem solving, learning [Bellman 78]</td>
<td>The study of computations that make it possible to perceive, reason and act [Winston 92]</td>
</tr>
<tr>
<td>The art of creating machines that perform functions that require intelligence when performed by a human [Kurzweil 90]</td>
<td>A field of study that seeks to explain and emulate intelligent behavior in terms of computational processes [Schalkoff 90]</td>
</tr>
<tr>
<td>The study of how to make computers do things at which, at the moment, people are better [Rich&amp;Knight 91]</td>
<td>The branch of computer science that is concerned with the automation of intelligent behavior [Luger&amp;Stubblefield93]</td>
</tr>
</tbody>
</table>
## Some Definitions (Russell + Norvig, 1995)

<table>
<thead>
<tr>
<th>Systems that think like humans</th>
<th>Systems that think rationally</th>
</tr>
</thead>
<tbody>
<tr>
<td>Systems that act like humans</td>
<td>Systems that act rationally</td>
</tr>
</tbody>
</table>
Computational Intelligence

- Notice the term **systems** in each box
  - most AI researchers build **systems** to do these things
  - *computational theories* or **models** of intelligence (and systems that implement them) our goal
  - **computational**, not “artificial”, intelligence

- But it still comes down to intelligence: what is it?

- Our defns break down as follows:
  - Two in terms of **thinking**, two acting
  - Two in terms of **what humans do**, two **rationality**

- Consider relative advantages of each
Aside on Rationality

What is **rationality**?
- A precise mathematical notion of what it means to *do the right thing* in any particular circumstance (see Ch.10.4 of our text)
- If you find yourself in a situation where you have several courses of action, choose one that’s best for you: one that has the best chance of achieving your goals (or furthering your interests)… balancing short and long term objectives

Does this sound selfish? Where do these interests come from?
Agency

- We’ll focus on acting rationally
  - which has implications for thinking/reasoning

- Our aim is to build **agents**, either:
  - with their own goals
  - or that act on behalf of someone (a “user”)

- An *agent* is an entity that that exists in an *environment* and *acts* on that environment based on its *perceptions* environment

- An *intelligent agent* acts to further its own interests (or those of a user)
Agent Schematic (I)

- Thermostat, ROSI, Web browser, Craig, you, software agent, etc.
Agent Schematic (II)

First diagram is misleading
• ROSI knew you were a student before interacting with you, and remembers stuff too!
Monitoring Intensive-Care Patients

The “alarm” network

37 variables, 509 parameters (instead of \(2^{52}\))
ALVINN

- Pomerleau, et al. NAVLAB Group (CMU)
NAVLAB
Autoclass (Cheeseman et al.)
Other Examples

- credit card fraud detection
- printer diagnostics, help in Windows, spam filters
- medical diagnosis, teleoperated/micro surgery
- information retrieval, Google
- TAC
- Cobot
- scheduling, logistics, etc.
- aircraft, pipeline inspection
- speech understanding, generation, translation
- Mars rover, DS1
- and, of course, cool robots (let’s go to video)
Degrees of Intelligence

- Range of circumstances in which agent can act
- E.g., a software agent
  - gets Web page you type in
  - does URL completion of unambiguous match
  - does URL completion of most frequent/likely match
  - completes URL based on context (time, history, etc.)
  - notices trends and prefetches Web pages, clips, etc.
  - downloads new music, makes new playlist for gym
  - buys your clothes at Eddie Bauer online
    - charges it to stolen credit card… just kidding
  - schedules car for tune ups, plans your trip to Italy this summer, makes a fortune daytrading, etc…
Degrees on Intelligence

- A rational agent’s degree of intelligence might be characterized by

1. ability to act rationally in a *variety of circumstances*

2. ability to *adopt complex goals* and balance their achievement

3. ability to *adapt to new circumstances*
Agent Properties (Wooldridge, Jennings)

- **Autonomy**: needs no direct intervention to perform duties
- **Reactivity**: perceives its environment and reacts appropriately to it
- **Proactivity**: exhibits goal-directed behavior
- **Sociability**: interacts with other agents
Areas of AI

- Perception: vision, speech understanding, etc.
- Robotics
- Natural language understanding
- Reasoning and decision making (our focus)
  - Knowledge representation
  - Reasoning (logical, probabilistic)
  - Decision making (search, planning, decision theory)
  - Machine Learning
Topics We’ll Cover

- What we’ll cover in this class
  - logical knowledge representation and reasoning
  - problem solving; graph-based search (AI-style)
  - game tree search
  - planning
  - probabilistic reasoning
  - Bayesian networks
  - utility theory
  - decision making under uncertainty

- Lots of other advanced AI courses in other areas
Organization

- Check Web page!
  - lots of info will be found there, including lecture slides, references, announcements, etc.
  - Newsgroup important! -- ut.cdf.csc384h
- Text: *Computational Intelligence, A Logical Approach*; Poole, Mackworth, Goebel
- Classes: Keep up with the readings!
  - see Web page for schedule, readings, and slides
- Tutorials: Thursday 6PM (just before class)
  - sections to be assigned; see Web page
- Five assignments, two midterms, one exam
  - 35% / 20% / 45% / *approximate* dates on Web pg
### General Lecture Announcements

- **Last time**
  - nothing

- **Today (rest of class)**
  - RRSs and DCL

- **Readings:**
  - Today ✅: Ch.1, Ch.2.1-2.4;
  - Next week: Ch.2.5, 2.6, 2.7 (excl. SLD/top-down proofs)

- **Announcements:**
  - See me immediately if you don’t have prereqs
  - Course Accounts created, see Web page
  - See Web page for tutorial section assignment
Logical KR and Reasoning

- We start with a specific way of doing logical KR and logical reasoning
  - since FOL and definite clauses (Prolog) familiar: fast
- Representation and Reasoning Systems (RRS)
  - how do we represent knowledge about the world in a computer/agent
  - how can the agent reason with (draw conclusions from) that knowledge
  - initial focus on static environments (no uncertainty)
  - we’ll take a formal, logical approach
RRSs require three components

**Specification:**
- provides firm, formal foundations for sentences we use to express knowledge
- the meaning of those sentences (what facts do they correspond to)
- how we draw conclusions (derive new facts) from the initial set of facts (e.g., answer questions)

**Implementation:**
- how we implement the spec. on a computer (later…)

**Representational Methodology:**
- how we use the system to represent specific types of knowledge
- some ways more natural, compact, efficient
Specifications

- Specifications require three components
  - **(Logical) Representation Language:**
    - syntax used to express sentences (knowledge)
  - **Semantics:**
    - method for determining the meaning of sentences
  - **Proof Procedures:**
    - how to answer questions, derive new facts
- **Our language: Definite Clause Language (DCL)**
  - subset of FOL (no disjunction, negation; tricks later)
  - forms the basis of logic programming (e.g., Prolog)
  - restrictive, but very powerful (will do all we need)
Assumptions

Assumptions we’ll start with (some justify DCL)

(1) IR: Agent’s world usefully described in terms of individuals and relations (properties) of them

Example: Computer travel agent domain

- Individuals: clients, destinations, hotels, airflight segments, airlines, prices, dates, …
- Properties: cost (hotel/segment), reliability (airline), rating (hotel), satisfied (client), carrier (segment) …
- Relations: desirable (dest’n for client on date); available (hotel, date); location (hotel, city), etc.
Assumptions

(2) **DK:** Agent’s knowledge is *positive* and *definite*
- no imprecise or negative knowledge
- OK: `home(craig, toronto)`
- not OK: `home(craig, toronto) OR home(craig, halifax)`
- not OK: `NOT home(craig, sherbrooke)`
- definiteness not the same as completeness

(3) **SE:** Agent’s environment is static
- this is only temporary (once we get to problem solving, planning, decision making, we’ll relax this)
DCL Formally

- Defined structurally (familiar from FOL)
- We’ll use Prolog-like notation

- A constant is a (lowercase) symbol
  - begins with lowercase letter, or a number
- A function symbol is a (lc) symbol
- A predicate symbol is a (lc) symbol
- A variable is a (uppercase) symbol

Each fctn and pred symbol has a specific arity (number of arguments)
DCL Formally

- A **term** is either:
  - a variable
  - a constant
  - an expression of the form \( f(t_1, \ldots, t_k) \) where (a) \( f \) is a function symbol; (b) \( k \) is its arity; (c) each \( t_i \) is a term

- An **atom** is an
  - expression of the form \( p(t_1, \ldots, t_k) \) where (a) \( p \) is a predicate symbol; (b) \( k \) is its arity; (c) each \( t_i \) is a term

- **Note:** if \( p \) takes zero args, we write “\( p \)”, not “\( p() \)”
  - if \( f \) takes zero args, it is a constant
Intuitions

- Terms denote individuals: constants denote individuals; functions build up ind’s out of others
  - bill dick jane father(jane) father(father(jane))
  - X father(X) hotel7 rating(hotel7) cost(hotel7)

- Atoms denote facts that can be true of false about the world
  - father_of(jane, bill) female(jane) system_down
  - satisfied(client15) satisfied(C)
  - desires(client15, rome, week29) desires(X,Y,Z)
  - rating(hotel7, 4) cost(hotel7, 125)
DCL Formally

- A **body** is $a_1 \& a_2 \& \ldots \& a_n$ where each $a_i$ is an atom
  - a *conjunction* of atoms (denotes that each is true)
- A **rule** is $a \leftarrow <\text{body}>$ where $a$ is an atom
- A **fact** is $a$. where $a$ is an atom
  - note the period
  - it’s a rule with an empty body
- A **definite clause** is a fact or a rule
- A **knowledge base (KB)** is a set of definite clauses
Example Clauses

\begin{align*}
\text{happy(client17)} & \leftarrow \text{desires(client17, rome, week29)} \land \\
& \quad \text{available(hotel7, week29)} \land \text{location(hotel7, rome)} \\
& \quad \land \text{rating(hotel7)} > 4.
\end{align*}

\begin{align*}
\text{happy(C)} & \leftarrow \text{desires(C, Dest, Date)} \land \text{available(H, Date)} \\
& \quad \land \text{location(H, Dest)} \land \text{rating(hotel7)} \geq R \\
& \quad \land \text{minQuality(C)} = R.
\end{align*}

\begin{align*}
\text{happy(C)} & \leftarrow \text{desires(C, Dest, Date)} \land \text{available(H, Date)} \\
& \quad \land \text{location(H, Dest)} \land \text{rating(hotel7)} < R \\
& \quad \land \text{minQuality(C)} = R \land \text{offTravelMug}(C).
\end{align*}

\begin{align*}
\text{desires(client17, rome, week29).} \\
\text{desires(client17, rome, week29).} \\
\text{location(hotel7, rome).}
\end{align*}

\ldots \ldots