Lecture 2 - Agents and Abstraction

Jesse Hoey
School of Computer Science
University of Waterloo

January 9, 2018

Readings: Poole & Mackworth 1.3-1.10
Situated Agent

- Abilities
- Goals/Preferences
- Prior Knowledge
- Stimuli
- Past Experiences

Agent

Environment

Actions
Four Example Application Domains (From Book)

- **Autonomous delivery robot** roams around an office environment and delivers coffee, parcels, ...  
- **Diagnostic assistant** helps a human troubleshoot problems and suggests repairs or treatments. E.g., electrical problems, medical diagnosis.  
- **Intelligent tutoring system** teaches students in some subject area.  
- **Trading agent** buys goods and services on your behalf.
Domain for Delivery Robot
Autonomous Delivery Robot

Example inputs:

- **Abilities:** movement, speech, pickup and place objects, sense weather
- **Observations:** about its environment from cameras, sonar, sound, laser range finders, or keyboards.
- **Prior knowledge:** its capabilities, objects it may encounter, maps.
- **Past experience:** which actions are useful and when, what objects are there, how its actions affect its position.
- **Goals:** what it needs to deliver and when, tradeoffs between acting quickly and acting safely, effects of getting wet.
What does the Delivery Robot need to do?

- Determine where Jesse’s office is. Where coffee is...
- Find a path between locations.
- Plan how to carry out multiple tasks.
- Make default assumptions about where Jesse is.
- Make tradeoffs under uncertainty: should it go near the stairs or outside?
- Learn from experience.
- Sense the world, avoid obstacles, pickup and put down coffee.
Example: COACH Handwashing Assistant
COACH (Handwashing System)

- **Abilities:** Provide passive assistive prompting
- **Observations:** Video images, audio feeds
- **Prior knowledge:** Effects of dementia, steps of handwashing, affect/emotion/culture
- **Past experience:** Prior behaviours of user, effects of prompts
- **Goals:** Get hands clean, maintain independence, reduce caregiver burden
What does the COACH need to do?

- Sense the world, determine what steps the user has completed
- Estimate the user’s current responsiveness, awareness and emotion
- Provide timely prompts only if needed
- Call for assistance if it can’t help
- Learn from experience.
User Trial Example
Trials: Reduction in Caregiver Burden

- overall 25% reduction in caregiver burden
- functionally it worked well
- did not work for some users
- prompt delivery important
- emotional component

† A.Mihailidis, J.Boger, M.Candido & J.Hoey, BMC Geriatrics, 2008
Knowledge Representation

- Represent
- Compute
- Solve
- Interpret
- Output
- Solution
- Problem
- Representation
- Informal
- Formal
A symbol is a meaningful physical pattern that can be manipulated.

A symbol system creates, copies, modifies and destroys symbols.

Physical symbol system hypothesis (Newell & Simon, 1976):

A physical symbol system has the necessary and sufficient means for general intelligent action.

Implies that: AI on a computer is possible in theory, but not necessarily feasible in practice.
A good representation should be

- Rich enough to express the problem
- Close to the problem: compact, natural and maintainable
- Amenable to efficient computation
- Amenable to elicitation from people, data and experiences
A POMDP is a probabilistic temporal model of a system interacting with its environment:
a tuple $\langle S, A, T, R, O, B \rangle$

- $S$: finite set of unobservable states
- $O$: set of observations
- $B : S \times A \rightarrow O$ observation function
- $A$: finite set of system actions
- $T : S \times A \rightarrow S$ transition function
- $R : S \times A \rightarrow R$ reward function

![Diagram of POMDP](image-url)
A POMDP is a probabilistic temporal model of a system interacting with its environment: a tuple \( \langle S, A, T, R, O, B \rangle \)

- models stochastic dynamic process
- explicit model of uncertainty
- explicit model of utility
- optimises expected long-term utility
- takes explicit information gathering actions: **value of information**
A POMDP is a probabilistic temporal model of a system interacting with its environment:
a tuple $\langle S, A, T, R, O, B \rangle$
Partially Observable Markov Decision Process (POMDP)
Handwashing Task: Plansteps

Behavior:
- Rinse dirty
- Clean water on dirty
- Soapy water on clean
- Water off dirty
- Behavior: Rinse
- Soapy, dry water on dirty
- Clean, dry water off dirty
- Clean, wet water on dirty
- Clean:H
- Dirty:K
- Dirty, dry water off dirty
- Soapy:G
- Soapy, dry water on dirty
- Clean, dry water off dirty
Handwashing: Actions

- **do nothing**: system waits
- **call caregiver**: system calls for **single step** assistance
- **prompts**:
  - 3 levels of specificity:
    - **low**: basic prompt with few details
      - “Use the soap”
    - **medium**: include person’s name, more details
      - “John, use the soap in the pink bottle”
    - **high**: medium prompt with video demonstration

![Image of hands washing](image-url)
Hand Tracking

Tracking through occlusions and changes in shape

- Objects modeled as **flocks of features**
- Simple **color features**, Gaussian distributions
- **Bayesian sequential** estimation
- Monte Carlo approximation (**particle filter**)
- **Three interacting** filters: 2 hands + towel
Handwashing POMDP: Behaviors
Reasoning and acting

Reasoning is the computation required to determine what an agent should do.

- **Design time reasoning and computation** is carried out by the designer of the agent.
- **Offline computation** is the computation done by the agent before it has to act.
  
  Background knowledge and data \(\sim\) **knowledge base**.

- **Online computation** is the computation that’s done by an agent between receiving information and acting.
Specifying, Solving, Learning

$\pi(b(s))$

$\text{SPECIFY} \rightarrow \text{SOLVE} \rightarrow \text{TEST} \rightarrow \text{LEARN} \rightarrow \text{SPECIFY}$
Solutions

Given a problem, what is a solution?

- **Optimal** according to some measure of quality e.g. *utility*
- **Satisficing**: meets all necessary criteria
- **Approximately optimal**: e.g. within 10% of optimal
- **Probable**: is likely to be right (but could be wrong...really wrong)
POMDP policies

- POMDP policy: maps observation histories $\rightarrow$ actions
- 200K states $\times$ 198 observations $\times$ 26 actions
- Optimal solution (with e.g. dynamic programming) intractable
- **Approximations** used:
  - Point-based value iteration (dynamic programming) with structured value function representation
  - Bounded solution size + state aggregation + observation pruning
  - Forward search (Monte-Carlo) anytime algorithm
Research proceeds by making simplifying assumptions, and gradually reducing them.

Each simplifying assumption gives a dimension of complexity
  ▶ Can be multiple values in a dimension: values go from simple to complex
  ▶ Simplifying assumptions can be relaxed in various combinations

Much of the history of AI can be seen as starting from the simple and adding in complexity in some of these dimensions.
Dimensions of Complexity

- Flat $\rightarrow$ modular $\rightarrow$ hierarchical
- Explicit states $\rightarrow$ features $\rightarrow$ objects and relations
- Static $\rightarrow$ finite stage $\rightarrow$ indefinite stage $\rightarrow$ infinite stage
- Fully observable $\rightarrow$ partially observable
- Deterministic $\rightarrow$ stochastic dynamics
- Goals $\rightarrow$ complex preferences
- Single-agent $\rightarrow$ multiple agents
- Knowledge is given $\rightarrow$ knowledge is learned from experience
- Perfect rationality $\rightarrow$ bounded rationality
Much of modern AI is about finding compact representations and exploiting that compactness for computational gains. A agent can reason in terms of:

- **explicit states** — a state is one way the world could be
Much of modern AI is about finding compact representations and exploiting that compactness for computational gains. A agent can reason in terms of:

- **explicit states** — a state is one way the world could be
- **features** or **propositions**.
  - It's often more natural to describe states in terms of features.
  - 30 binary features can represent $2^{30} = 1,073,741,824$ states.
Succinctness and Expressiveness

Much of modern AI is about finding compact representations and exploiting that compactness for computational gains. A agent can reason in terms of:

- **explicit states** — a state is one way the world could be
- **features** or **propositions**.
  - It’s often more natural to describe states in terms of features.
  - 30 binary features can represent $2^{30} = 1,073,741,824$ states.
- **individuals** and **relations**
  - There is a feature for each relationship on each tuple of individuals.
  - Often we can reason without knowing the individuals or when there are infinitely many individuals.
Example: Handwashing Task: Plan steps

- **dirty**
  - a
  - b
  - e

- **soapy**
  - d

- **behavior: rinse**

- **soapy, dry**
  - j

- **clean, wet**
  - g

- **clean, dry**
  - k

- **dirty, dry**
  - h

- **water on**

- **water off**
Relational tasks in general

Can generalize assistance across tasks with relations:

- Action: name, value, constraints
- Ability: name, change_probability(action)
- Behaviour: probability(ability, action, state)
- State Variable: name, values, probability(previous state, action, behaviour)

Gets converted to a propositional representation at runtime
Planning horizon

...how far the agent looks into the future when deciding what to do.

- **Static:** world does not change
- **Finite stage:** agent reasons about a fixed finite number of time steps
- **Indefinite stage:** agent is reasoning about finite, but not predetermined, number of time steps
- **Infinite stage:** the agent plans for going on forever (process oriented)
Uncertainty

What the agent can determine the state from the observations:

- **Fully-observable**: the agent knows the state of the world from the observations.
- **Partially-observable**: there can be many states that are possible given an observation.
Handwashing: Partial Observability

- Observational Uncertainty: hand locations
- True non-observability: user awareness and responsiveness

(a) 4974  
(b) 5096  
(c) 5502  
(d) 5576
If the agent knew the initial state and the action, could it predict the resulting state?

The dynamics can be:

- **Deterministic**: the state resulting from carrying out an action in state is determined from the action and the state.
- **Stochastic**: there is uncertainty over the states resulting from executing a given action in a given state.
Goals or complex preferences

- **achievement goal** is a goal to achieve. This can be a complex logical formula.
- **maintenance goal** is a goal to be maintained.
- **complex preferences** that may involve tradeoffs between various desiderata, perhaps at different times. Either ordinal or cardinal (e.g., utility)

**Examples:** coffee delivery robot, medical doctor
Example: Complex Preferences

handwashing assistant:

get hands clean -> user independence

user independence -> caregiver burden

caregiver burden -> get hands clean
Single agent or multiple agents

- **Single agent** reasoning is where an agent assumes that any other agents are part of the environment. (delivery robot)
- **Multiple agent** reasoning is when an agent needs to reason strategically about the reasoning of other agents. (robot soccer, trading agents)

Agents can have their own goals: cooperative, competitive, or goals can be independent of each other
Learning from experience

Whether the model is fully specified a priori:

- knowledge is given
- knowledge is learned from data or past experience
Perfect rationality or bounded rationality

- **Perfect rationality:** the agent can determine the best course of action, without taking into account its limited computational resources.

- **Bounded rationality:** the agent must make good decisions based on its perceptual, computational and memory limitations.
Dimensions of Complexity: Handwashing problem

- **flat** → modular → hierarchical
- explicit states → **features** → objects and relations
- static → finite stage → indefinite stage → **infinite stage**
- fully observable → **partially observable**
- deterministic → **stochastic** actions
- goals → **complex preferences**
- **single agent** → multiple agents
- knowledge is given → **learned**
- **perfect rationality** → bounded rationality
Next:

- Read Poole & Mackworth chapter 2.1-2.3
- Uninformed Search (Poole & Mackworth chapter 3)
- Informed Search (Poole & Mackworth chapter 4)