AI Planning for Autonomy

1. Plan & Goal Recognition

Speak Up Room: ‘TBD”, PIN: TBD

Miquel Ramirez

THE UNIVERSITY OF MELBOURNE

August, 2018

Outline of the Lecture

1. Perceiving and Interpreting the Behavior of Others
2. Plan and Goal Recognition in AI
3. Plan and Goal Recognition and Classical Planning

The Heider-Simmel Experiment

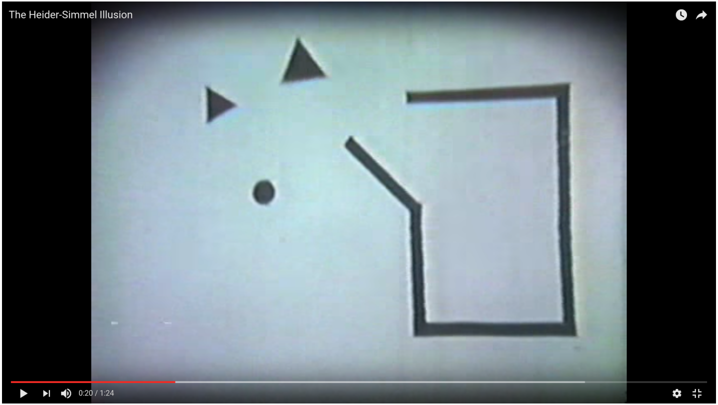


Figure: *An Experimental Study of Apparent Behavior*. F. Heider, M. Simmel. The American Journal of Psychology, Vol. 57, No. 2, April 1944

Link to video (YouTube)

Parsing the Big Triangle

Figure: The BIG triangle T.

Speak Up Room: "TBD", PIN: TBD

Question!

What kind of person is the Big Triangle?

(A): Aggressive, mean, angry. (B): Strong, powerful.

(C): Dumb, stupid. (D): Ugly, sly.

what about the Smaller one...

Figure: The small triangle t.

Question!

What kind of person is the Small Triangle?

(A): Fearless, defiant, cocky.
(B): Passive–aggressive.
(C): Clever, weak.
(D): Protective, loyal, devoted.

and about the circle...

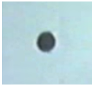


Figure: The circle c.

Question!

What kind of person is the Circle?

- (A): Frightened, fearful, helpless.
- (B): Fidgety, playful, nervous.
- (C): Clever, smart.
- (D): Courageous.

Significance of Heider & Simmel Results

Leaving aside issues with *priming* experimental subjects...

It does seem that

1. humans *tend* to **ascribe intentions** to *anything* that changes over time,
2. this rests on deeply rooted **assumptions**.

Heider & Simmel results are the *first* quantitative characterization of:

Folk Psychology

Human capacity to **explain** and **predict** behavior and mental state of others

. . . we're usually very good at it, but we fail often!

A Theory of Common Sense

The Intentional Stance, **Daniel Dennett** (1988)

1. **Decide** to consider the object being observed as *rational*.
2. Work out its **beliefs** and **goals** based on its *place* and *purpose* in the world.
3. Use **practical reasoning** to assess what the agent *ought to do* to pursue its **goals**.

The above provides a *systematic*, *reason–giving explanation* for actions, based on deeply embedded beliefs about the agent.

So Deep we are Barely Aware of Them

Typical Assumptions

a) Agent's actions are *entirely* rational
   - → cares about some measure of value, *but* may be bad at appraising it

b) Actions are reasonable or follow some law
   - → e.g. effects as per known laws of physics, or economic theories
   - → *all models are wrong*

c) Beliefs and Goals held for the duration of the performance
   - → may change their mind, we may be observing reflex action

d) Future actions can be systematically predicted from beliefs and goals *ascribed*
   - → we may be missing possible and relevant goals,
   - → we may lack enough time to test plausible explanations exhaustively.

Several of them false **most of the time**, yet *rarely* **all** of them are.

Plan and Goal Recognition in Artificial Intelligence

Key Idea: use generative models of behavior to predict actions.

Plan Recognition (PR) is Planning in reverse.

- Planning – we seek *plans* π to *achieve* goals G.
- PR: find goals G *accounting for* partially *observed* plan π.

Multi–Agent setting

Two possible *roles* for each agent:

- **Actor** – *performs* actions to change the state of the world.
- **Observer** – *perceives* actions and updates its beliefs on the **Actor** intentions.

and three possible *stances* for the **Actor**:

- *Adversarial* – obfuscates deliberately its goals.
- *Cooperative* – tries to tell the **Observer** what she is up to.
- *Indifferent* – does not care about the **Observer**.

Open Challenge→Stances could be *changing over time*

The Elements of PR

Actions describe what the **Actor** does

- Walking from X to Y, opening a door, using a credit card...

Goals describe what the **Actor** wants

- To have breakfast, Plant a bomb, Troll the Interwebs...

Plans describe how goals can be achieved

- Ordered sequences of actions
- These can be ranked according to **cost** or **efficiency**

Sensor Model describes what does the **Observer** perceives

- Does it always see every action done by the **Actor**?
- Are actions observed *directly*? Or only their *effects* are?
- Does it know exactly *where* in the world the **Actor** is?

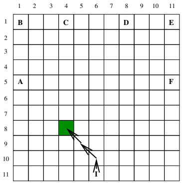
Example: Agent on a Grid World

	1	2	3	4	5	6	7	8	9	10	11
1	**B**			**C**				**D**			**E**
2											
3											
4											
5	**A**										**F**
6											
7											
8											
9											
10											
11						**I**					

- **starts** in "I", may be heading to "A", "B", . . ., "F".
- **moves along compass directions** *North*, etc. with cost 1 and *North West*, etc. with cost $\sqrt{2}$.

Example

Actor now at $(4, 8)$ after going N once, and twice NW.



Question!

Assuming the Actor prefers CHEAPEST plans which goals are most likely?

(A): A & B.

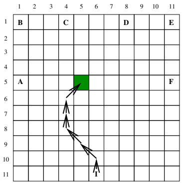
(B): C.

(C): D

(D): E & F

Example

Actor now at $(5, 5)$ after going N twice and once NE.



Question!

For which goal(s) observed actions are in a CHEAPEST plan?

(A): A & B.

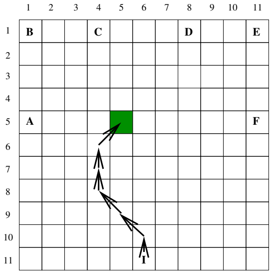
(B): C.

(C): D, E & F

(D): None

So Folk Psychology is Useless?

Perhaps, but as a framework is useful we can extend it.



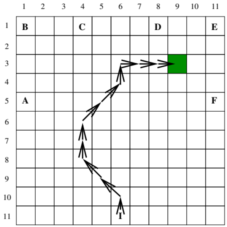
Counterfactual Reasoning (Pearl, 2001) to Handle Non–Optimal Behaviour

Compare best plans that comply with observed actions, with best plans that do not.

→ Then it follows B and C *more likely* than A or the rest.

Example

Actor at $(9, 3)$, has gone NE, N and three times E.



Question!

Now, which are the most likely goals?

(A): A & B.

(B): C & D.

(C): E

(D): F

Plan Recognition over Plan Libraries

Definition

$\textsc{Pr}$ problem defined as tuple $T = \langle \mathcal{G}, \mathcal{L}, A, O \rangle$ where

- $\mathcal{G}$ is the set of **possible goals** G,
- $\mathcal{L} = \bigcup \mathcal{L}(G)$ set of **possible plans** π for G, $G \subseteq \mathcal{G}$,
- O is **observation sequence** $(a_1, \ldots, a_i, \ldots, a_n)$, $a_i \in A$

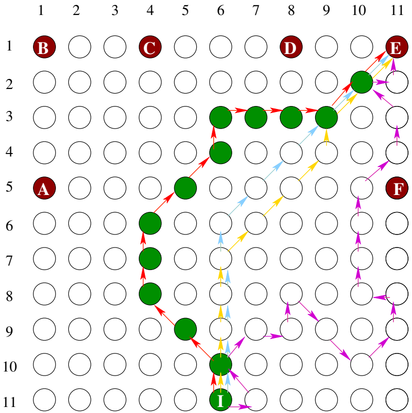
Key Property

An action sequence π satisfies O if O embedded in π.

Solution

A possible goal $G \in \mathcal{G}$ is plausible if $\exists$ plan π in $\mathcal{L}(G)$ that satisfies O.

Plan Recognition over Plan Libraries: Limitations



4 plans:

- two optimal
- two valid

What plans get into $\mathcal{L}$?

Why?

$|\mathcal{L}|$ **exponential** on **size** of grid to be **complete**.

Key Facts of the Model–Based Approach

1. $\mathcal{L}$ given **implicitly**, requires to **solve** $|\mathcal{G}|$ planning tasks
2. Plans "**extracted**" with **off–the–shelf** planning algorithms.
3. **Plausibility** of goals $\mathcal{G}$ given as a **probability distribution**
   - Goals are *plausible* when motivate plans *consistent* with O,
   - **and** when O is *necessary* to achieve goals *efficiently*.

Plan Recognition As Planning (Ramirez & Geffner, 2009-10)

Define possible agent behavior implicitly, according to general principles:

- A STRIPS planning domain $P[\cdot]$,
- Cost function over actions and/or states,
- Hypothetic set of goals $\mathcal{G}$,
- General Knowledge given as prior probability distribution over $\mathcal{G}$.

Assume the **Actor** prefers lower cost plans

- Rationality Assumption in *The Intentional Stance* (Dennett, 1987).

and is indifferent towards the **Observer**

Use generic planners to **compute** $P(G|O)$.

STRIPS: A Factored Representation of Planning Models

A STRIPS planning domain is the tuple $P[\cdot] = \langle F, I, A \rangle$ where

- F is the set of fluents.
- $I \subseteq F$ is initial situation.
- A set of STRIPS actions, a with $Pre(a)$ and conditional effects e_i,
  - $\langle\ Cond(e_i),\ Add(e_i),\ Del(e_i)\ \rangle$

  all *subsets* of F.

A planning problem $P[G]$ is domain P and goal $G \subseteq F$.

Solution of $P[G]$ is a valid sequence of actions, or *plan*

$$\pi = (a_1, \ldots, a_n)$$

Defining Behaviour Implicitly via Constraints

Hard Constraints: Valid plans for $P[G]$, is sequence $\pi = (a_1, \ldots, a_n)$ s.t.

1. $Pre(a_1) \subseteq I$,
2. $Pre(a_i) \subseteq I[a_1, \ldots, a_{i-1}]$ for $i > 1$,
3. $G \subseteq I[a_1, \ldots, a_n]$.

Soft Constraints: The cost of a plan π is

$$c(\pi) = \sum_{a \in \pi} c(a)$$

Valid plans are optimal when they have minimum cost.

The cost of $P[G]$, $c^*(P[G])$, is the cost of an optimal plan.

Roadmap

1. Make off-the-shelf planners compute plans constrained w.r.t. O,
2. Derive $P(G|O)$ from best plans that comply with *and* work around O.

Computing Plans that Satisfy O

Problem

To find efficiently what plans satisfy O with an unmodified planner.

Idea #1: Plan Constraints

Constrain plans to embed (or avoid so) O requiring extra dummy goals through changing preconditions and effects

Idea #2: Implicit Sorting of Plans

For each $G \in \mathcal{G}$, formulate two planning tasks, splitting set of valid plans

1. π plan for $P[G]$ that satisfies O **iff** π plan for $P'[G + O]$
2. π plan for $P[G]$ that doesn't satisfy O **iff** π plan for $P'[G + \overline{O}]$

PR as planning: Inferring the Goal Probabilities

Goal

Obtain **probability distribution** $P(G|O)$, $G \in \mathcal{G}$.

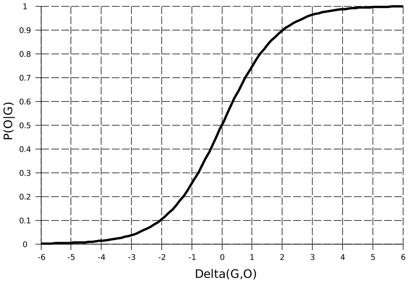
Outline of Approach

From **Bayes' Rule** $P(G|O) = \alpha\, P(O|G)\, Prob(G)$, where

- α norm. constant
- $Prob(G)$ given in problem specification
- $P(O|G)$ function of extra cost needed to not comply with O

$$P(O|G) = \text{function}(c^*(P'[G + \overline{O}])) - c^*(P'[G + O])) \tag{1}$$

$P(O|G)$ and $P(\overline{O}|G)$



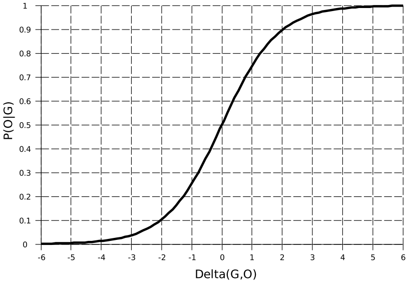
Ramirez & Geffner Observer Model (AAAI, 2010)

$$P(O|G) = sigmoid(\beta\, \Delta(G,O))$$

where

$$\Delta(G,O) = c^*(P'[G+\overline{O}]) - c^*(P'[G+O])$$

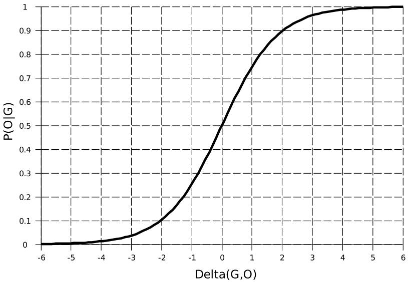
Goals as Predictors for O (informally)



Properties

1. G predicts O badly when it would be more efficient to deviate from O.
2. G predicts O **perfectly** when G unfeasible if not doing O.

Goals as Predictors for O (formally)



Properties

1. G predicts O badly when $P(O|G) < P(\overline{O}|G)$ since

$$c^*(P'[G+\overline{O}]) < c^*(P[G+O])$$

2. G predicts O perfectly when $P(O|G) = 1$ since

$$c^*(P'[G+\overline{O}]) = \infty$$

Demo: A Slightly More Interesting STRIPS Model

Fluents: *facts about the world*

- Locations of people
- State of appliances
- Locations of objects

Actions: *stuff people may do*

- Move across the place
- Interaction with objects & appliances

Goals: *why people do stuff*

- Cook some foodstuff
- Watch a movie
- Listen to a record
- Go to sleep
- Get ready to leave for work

Unitary action costs (to keep it simple)

GitHub Repo Pull Requests Welcome!

Anyone looking for a Masters' project? Revamp needed.

Further Reading or Watching

Article *An Experimental Study of Apparent Behavior*. F. Heider, M. Simmel. The American Journal of Psychology, 57(2), 1944

A Probabilistic Plan Recognition Algorithm based on Plan Tree Grammars C. Geib, R. Goldman, Artificial Intelligence 173(11), 2009

Probabilistic Plan Recognition using off-the-shelf Classical Planners. M. Ramirez and H. Geffner. Proceedings AAAI, 2010.

Landmark-Based Heuristics for Goal Recognition. R. Pereira, N. Oren and F. Meneguzzi. Proceedings AAAI, 2017.

Heuristic Online Goal Recognition in Continuous Domains, M. Vered and G. Kaminka. Proceedings IJCAI, 2017.

Book Chapter 4, Section 4.3 *A Concise Introduction to Models and Methods for Automated Planning*. B. Bonet & H. Geffner, Morgan Claypool, 2013.

Video Lecture *Engineering & Reverse-engineering Human Common Sense*, J. Tenenbaum, Allen Institute for AI, 2015.

Prob. PR as Planning: Cheat Sheet

A **plan recognition problem** is a tuple $T = (P, \mathcal{G}, O, Prob)$ where

- P is a planning domain $P = (F, I, A)$
- $\mathcal{G}$ is a set of possible goals G, $G \subseteq F$
- O is the observation sequence $(a_1, \ldots, a_n)$, $a_i \in O$
- $Prob$ is prior distribution over $\mathcal{G}$

Posterior distribution $P(G|O)$ obtained from

- Bayes Rule $P(G|O) = \alpha P(O|G) Prob(G)$ and
- Likelihood $P(O|G) = sigmoid\{\beta\,[c^*(P'[G+\overline{O}]) - c^*(P'[G+O])]\}$

$c^*(P'[G+O])$ and $c^*(P'[G+O])$ **computed**

- exactly with an optimal planner
- approximately with a sub–optimal planner

In either case, $2 \cdot |\mathcal{G}|$ **planner calls are needed** for $P(G|O)$