Goal Recognition: Models, Algorithms, Challenges

Hector Geffner
ICREA and Universitat Pompeu Fabra
Barcelona, Spain

Planning

- Planning is the **model-based** approach to action selection: behavior obtained from **model** of the **actions**, **sensors**, **preferences**, and **goals**

$$\textit{Model} \Longrightarrow \boxed{\textit{Planner}} \Longrightarrow \textit{Controller}$$

- Many **planning models**; many **dimensions**: uncertainty, feedback, costs, . . .

Basic Model: Classical Planning

- finite and discrete state space S
- a **known initial state** $s_0 \in S$
- a set $S_G \subseteq S$ of goal states
- actions $A(s) \subseteq A$ applicable in each $s \in S$
- a **deterministic transition function** $s' = f(a, s)$ for $a \in A(s)$
- positive **action costs** $c(a, s)$

A **solution** is a sequence of applicable actions that maps s_0 into S_G, and it is **optimal** if it minimizes sum of action costs ($\#$ of steps)

Other **models** obtained by relaxing assumptions in **bold** . . .

Uncertainty and Full Feedback: Markov Decision Processes

Goal MDPs are **fully observable, probabilistic** state models:

- a state space S
- initial state $s_0 \in S$
- a set $S_G \subseteq S$ of goal states
- actions $A(s) \subseteq A$ applicable in each state $s \in S$
- **transition probabilities** $P_a(s'|s)$ for $s \in S$ and $a \in A(s)$
- action costs $c(a, s) > 0$

– **Solutions** are **functions (policies)** mapping states into actions

– **Optimal** solutions minimize **expected cost** to goal

Partial Feedback: Partially Observable MDPs (POMDPs)

Goal POMDPs are **partially observable, probabilistic** state models:

- states $s \in S$
- actions $A(s) \subseteq A$
- transition probabilities $P_a(s'|s)$ for $s \in S$ and $a \in A(s)$
- observable goal states $S_G \subseteq S$

- initial **belief state** b_0
- **sensor model** given by probabilities $P_a(o|s)$, $o \in O$, $s \in S$

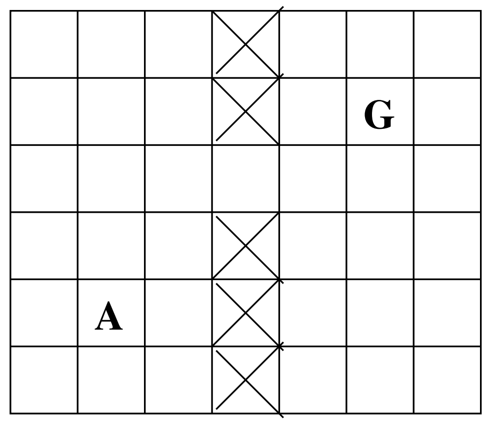
– **Belief states** are probability distributions over S

– **Solutions** are policies that map belief states into actions

– **Optimal** policies minimize **expected** cost to go from b_0 to S_G

Example

Agent **A** must reach **G**, moving one cell at a time in **known** map



- If actions deterministic and initial location known, planning problem is **classical**
- If actions stochastic and location observable, problem is an **MDP**
- If actions stochastic and location partially observable, problem is a **POMDP**

Three problems, three models, three solution forms

From Planning to Plan/Goal Recognition

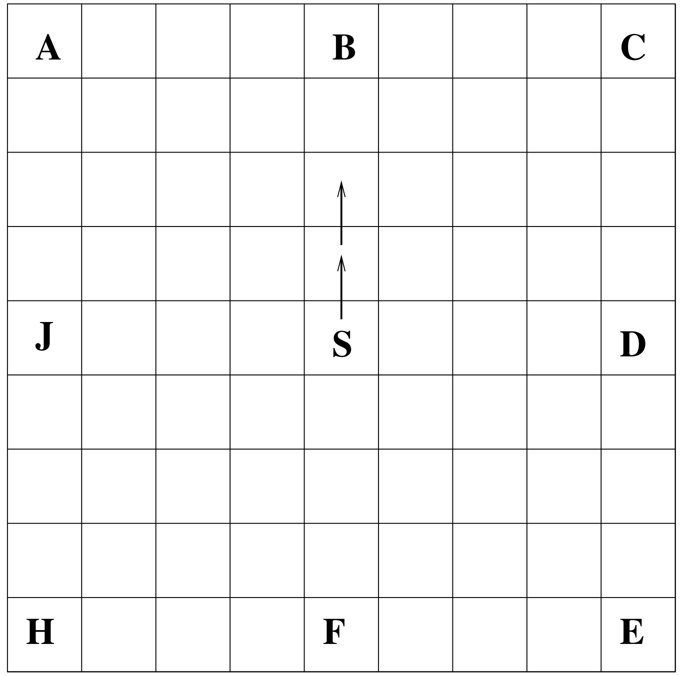
- **General idea:** solve **plan recognition** problem over **model** (classical, MDP, POMDP) using **planner** for that **model**.

- Early work in this direction using **classical** models, **MDPs**, and **POMDPs**:

  ▷ *Plan Recognition as Planning*, M. Ramirez and H. G., Proc. IJCAI-2009
  ▷ *Probabilistic Plan Recognition using off-the-shelf Classical Planners*, M. Ramirez and H. G., Proc AAAI-2010
  ▷ *Goal recognition over POMDPs: Inferring the intention of a POMDP agent*. M. Ramirez and H. G., Proc IJCAI-2011

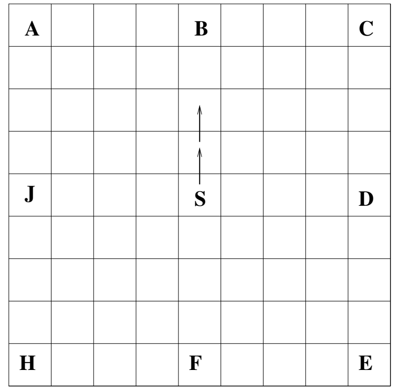
  ▷ *Goal Inference as Inverse Planning*, C. Baker, J. Tenenbaum, R. Saxe. Proc. Cog-Sci 2007
  ▷ *Action Understanding as Inverse Planning*. C. Baker, R. Saxe, and J. Tenenbaum. Cognition, 2009
  ▷ *Bayesian theory of mind: Modeling joint belief-desire attribution*. C. Baker, R. Saxe, J. Tenenbaum, Proc. Cog Science 2011

Example: Classical Setting



- Agent can **move** one unit in the four directions
- Possible **targets** are A, B, C, . . .
- Starting in S, he is **observed** to move up twice
- **Where** is he going? Why?

Example (cont'd)



- From Bayes, **goal posterior** is $P(G|O) = \alpha\, P(O|G)\, P(G)$, $G \in \mathcal{G}$
- If **priors** $P(G)$ given for each goal in $\mathcal{G}$, the question is what is $P(O|G)$
- $P(O|G)$ measures **how well goal** G **predicts observed actions** O
- In **classical** setting,
  - ▷ G predicts O **worst** when needs to get off the way **to comply with** O
  - ▷ G predicts O **best** when needs to get off the way **not** to comply with O

Posterior Probabilities from Plan Costs

- From Bayes, **goal posterior** is $P(G|O) = \alpha\, P(O|G)\, P(G)$,
- If **priors** $P(G)$ given, set $P(O|G)$ to monotonic function

$$\textbf{function}(c(G+\overline{O})\; - \; c(G+O))$$

  - ▷ $c(G+O)$: cost of achieving G **while complying with** O
  - ▷ $c(G+\overline{O})$: cost of achieving G **while not complying with** O

– Costs $c(G+O)$ and $c(G+\overline{O})$ **computed** by **classical planner**

– Goals of **complying** and **not complying** with O translated into normal goals

– **Function** of cost difference set to **sigmoid**; follows from assuming action selected with Boltzmann distributions

– Posterior probabilities $P(G|O)$ **computed in** $2|\mathcal{G}|$ **classical planner calls**, where $\mathcal{G}$ is the set of possible goals (Ramirez and G. 2010)

Goal Recognition in other Settings: Example

- Two objects A and B: A can be in drawers 1 or 2; B can be in 1, 2, or 3
- Agent doesn't know where A and B are but has **priors** $P(A@i)$, $P(B@i)$
- She can move around, open and close drawers, **look** for object in open drawer, and grab object from drawer if known to be there
- The **sensing** action is not perfect, and agent may fail to see object in drawer
- **Agent observed to do the actions:**

$$O = \{open(1), open(2), open(1)\}$$

- **What's the agent goal?** Is she looking for object A or object B?

Unified Formulation for Classical, MDP, and POMDP Models

$O_i = \langle a_i, o_{i+1} \rangle$: action a_i, observation token o_{i+1} by (observed) actor

Posterior probabilities: using Bayes' rule

$$P(G|O_1, \ldots, O_n) \; = \; \alpha P(O_1, \ldots, O_n|G) P(G)$$

Likelihood $P(O_1, \ldots, O_n|G)$ from

$$
\begin{aligned}
&= P(O_n|O_1, .., O_{n-1}, G)\, P(O_1, .., O_{n-1}|G) \\
&= P(o_{n+1}|a_1, .., a_n, o_1, .., o_n, G)\, P(a_n|a_1, .., a_{n-1}, o_1, .., o_{n-1}, G)\, P(O_1, .., O_{n-1}|G) \\
&= P(o_{n+1}|a_n, bel_n)\, P(a_n|bel_n, G)\, P(O_1, .., O_{n-1}|G)
\end{aligned}
$$

; bel_i is **belief** at time i

- $P(o_n|a_n, bel_n)$ computed from POMDP parameters
- $P(a|b, G) = \alpha' exp\{-\beta\, Q_G(a, b)\}$ is prob of selecting action a for G in b:
  - ▷ $Q_G(a, b) = c(a, b) + \sum_{o \in O} b_a(o) V_G(b_a^o)$, and
  - ▷ $V_G(b)$ computed by **planner** represents expected cost from b to G

Assumptions and Special Cases

$$P(O_1, \ldots, O_n|G) = P(o_n|a_n, bel_n)\, P(a_n|bel_n, G)\, P(O_1, \ldots, O_{n-1}|G)$$

Assumptions

Observer can track beliefs of **actor**: **model**, **priors**, **actions**, and **observations**

Special Cases

- For **MDPs**, **beliefs** b_i and observations o_i replaced by **states** s_i:

$$P(O_1, \ldots, O_n|G) = P(s_{n+1}|a_n, s_n)\, P(a_n|s_n, G)\, P(O_1, \ldots, O_{n-1}|G)$$

- For **classical model** (**deterministic**), no need to observe actions **and** states:

$$P(O_1, \ldots, O_n|G) = P(a_n|s_n, G)\, P(O_1, \ldots, O_{n-1}|G)$$

Limitations and Challenges

- **Scalability:**
  - ▷ precompute v-functions $V_G(b)$ ($V_G(s)$) for all b (s) or call planner as needed
  - ▷ $2|\mathcal{G}|$ planner calls in classical setting approx'ed by single **poly IW(2)** call?
- **Lack of knoweledge**
  - ▷ what if no access to actions, observations, model, or priors of **actor**
  - ▷ what if **observer** has to act to get such observations (active goal recognition)
- **Role of intention recognition for general planning agent**
  - ▷ why intentional agent needs to infer intentions of others?
  - ▷ e.g., MDP agents with common goal: why infer other's subgoals?
- **From goal recognition to story understanding**
  - ▷ **general formulation** to explain children stories (Little Red Riding Hood)?
  - ▷ It's all about plans, intentions, and interactions unfolding . . .
  - ▷ why so simple for people . . . and children (!), and not yet in reach?

Little Red Riding Hood

Characters: Little Red Riding Hood (good, kind, skipping), Wolf (bad, scary, gobbles), Grandma (good), Woodcutter (good, strong, brave).
Setting: Woods (birds singing, sunny), Grandma's house (small, tidy, quaint)
Introduction: LRRH walking through woods on way to Grandma's – taking cakes (fairy, delicious) Meets wolf asks where going (growls, frightens)
Build-Up: Wolf goes to Grandma's house (knocks on wooden door). Gobbles her up (tasty, licks lips). Dresses in her clothes, gets into bed. Waits for LRRH
Climax: LLRH arrives – door open (cautiously, carefully). Talks to the wolf - eyes, hands, teeth (dialogue). He jumps out and chases her into the woods. (running, tripping, screaming).
Resolution: Woodcutter hears her cries and kills the wolf with axe. (striding, swings axe, happily ever after).

Summary

- **Model-based approach to goal recognition:** use planning model and planners to infer goals of an agent
- Many models depending on **uncertainty** and **feedback** like classical, MDPs, and POMDPs
- Most work assumes that **model of observer and actor suitably aligned**; convenient but unrealistic
- Many open questions:
  - ▷ what role for intention recognition for a general planning agent?
  - ▷ what general and effective model for understanding simple multigent stories?