Completeness of Online Planners for Partially Observable Deterministic Tasks

Blai Bonet Gabriel Formica Melecio Ponte

Universidad Simón Bolívar, Venezuela

ICAPS. Pittsburgh, USA. June 2017.

Motivation

Many online planners for **partially observable deterministic** tasks (e.g. Brafman & Shani 2016, B. & Geffner 2014, Maliah et al. 2014, ...)

Some planners offer **guarantees** over classes of problems

But theoretical analyses are often overly complex and specific to the planners and tasks

Want to develop **general framework** for analysis of online planning

Model for POD Tasks

Partially observable deterministic tasks correspond to tuples $P = (S, A, S_{init}, S_G, f, O, \Omega)$ where:

- S is finite state space
- A is finite set of actions where $A(s)$ is set of actions applicable at s
- $S_{init} \subseteq S$ is set of possible initial states
- $S_G \subseteq S$ is set of goal states
- $f : S \times A \to S$ is **deterministic transition function**
- O is finite set of observation tokens
- $\Omega : S \times A \to O$ is **deterministic sensing model**

Executions and Belief States

Agent sees **observable executions**; an observable execution is a **finite interleaved sequence** of actions and observations:

$$\tau \; = \; \langle a_0, o_0, a_1, o_1, \ldots \rangle$$

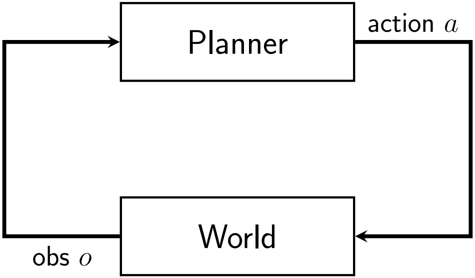
Belief b_τ = states deemed possible **after seeing execution** τ:

- $b_{\langle\rangle} \; = \; S_{init}$
- $b_{\langle \tau, a \rangle} \; = \; \{\, s' \in S \,:\, \text{there is } s \in b_\tau \text{ and } s' = f(s,a) \,\}$ (progression)
- $b_{\langle \tau, a, o \rangle} \; = \; \{\, s' \in b_{\langle \tau, a \rangle} \,:\, \Omega(s', a) = o \,\}$ (filtering)

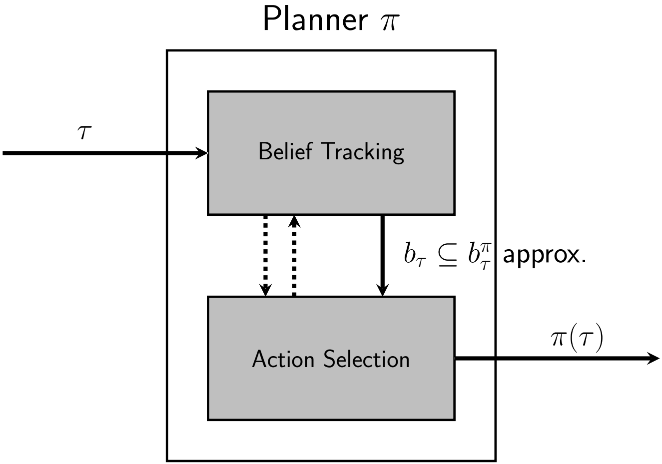
$$b_\tau \; \xrightarrow{a} \; b_{\langle \tau, a \rangle} \; \xrightarrow{o} \; b_{\langle \tau, a, o \rangle}$$

Belief tracking on factored models is intractable!

Online Planner: Closed-Loop Controller



Two Components in Online Planners



Online Protocol

Use of planner in online setting normed/modeled by **protocol**

Protocol $L = (P, s)$ **determined** by task P and initial state s:

1. Let $\lambda = \langle s \rangle$ be initial **state trajectory** seeded at s
2. Let $\tau = \langle \rangle$ be empty **execution**
3. While $b^\pi_\tau \subseteq S_G$ (i.e. agent isn't sure of reaching goal) do
4. **Run** planner π on input τ to obtain set of applicable actions $\pi(\tau)$
5. If $\pi(\tau)$ is empty, terminate with **FAILURE**
6. **Non-deterministically choose** action $a \in \pi(\tau)$
7. Let $s' := f(Last(\lambda), a)$ and token $o := \Omega(s', a)$
8. **Update** $\lambda := \langle \lambda, s' \rangle$ and $\tau := \langle \tau, a, o \rangle$

where b^π_τ is **approximation** of b_τ computed by agent

Main Goal

Formulate **formal properties** of components and their relation in order to guarantee **completeness** over **solvable tasks**

Definition (Completeness)

Online planner π is complete on task P if for each initial state $s \in S_{init}$, the protocol $L(P, s)$ **terminates successfully** *on π*

We would like to reason about completeness; e.g.

- Is planner π complete on P?
- Why isn't π complete on P?
- How do we make π complete on P?
- ...

Solvable Tasks

Two definitions:

Definition (Solvable Tasks)

Task P is **solvable** *(or goal connected) if there is a plan for each state s in P*

Definition (Strongly Solvable Tasks)

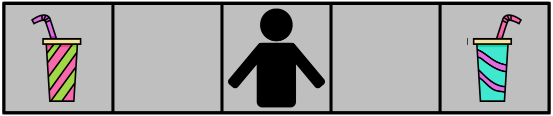
Task P is **strongly solvable** *(or goal connected in belief space) if for each initial state s and execution τ compatible with s, there is an extension $\tau' = \langle \tau, \tau'' \rangle$ compatible with s such that $b_{\tau'}$ is a goal belief*

Definitions are **incomparable:** there are tasks that are solvable but not strongly solvable, and vice versa

Reasons for Incompleteness

- Belief tracking is too weak; i.e. approximation b_τ^π of b_τ is too coarse
- Action selection is bad or **uncommitted**
- Combination of belief tracking and action selection isn't good enough

Uncommitted Planner Fails in Simple Example



- Agent is thirsty and wants a drink; it can move and gulp a drink
- There are two drinks
- No need for belief tracking as state is always known
- Agent may loop even if selected action always moves "toward goal" (e.g. Left, Right, Left, Right, . . .)

Properties for Belief Tracking

- **Exact:** beliefs computed by π are **exact;** i.e., $b^{\pi}_{\tau} = b_{\tau}$ for each τ
- **Monotone:** for every execution τ and **prefix** τ' of τ, $|b^{\pi}_{\tau}| \leq |b^{\pi}_{\tau'}|$ (i.e. non-increasing "amount of uncertainty" along executions)
- **Asserting:** there is asserting inference for pair (τ, τ') (where τ' is **proper prefix** of τ) if $|b^{\pi}_{\tau}| < |b^{\pi}_{\tau'}|$ (uncertainty decreases)

Exact inference $\implies$ monotone inference (because determinism)

Properties for Action Selection

For handling commitment, we do a slight reformulation and consider planners that return set of **action sequences (plans)** on input τ

First action on each sequence σ **must be applicable**

Properties:

- **Committed:** by caching last computed sequences, the planner sticks to selected plan "as much as possible"
- **Weak:** for each approximation b^π:
  - each sequence σ returned by π is a **plan for some state** $s \in b^\pi_\tau$
  - if b^π_τ is non-empty, π returns at least one sequence σ
- **Covering:** the first action in sequences returned by π **cover all** applicable actions at **exact belief** b_τ

Relation between Components

Do we need **exact but intractable** belief tracking for completeness?

Relation between Components

Do we need **exact but intractable** belief tracking for completeness?

Fortunately not!

A **sufficient** condition:

– Planner π is **weak:** given execution τ, π returns at least one plan σ for some state $s \in b_\tau^\pi$ (state s may not be in b_τ)

– Plan σ is applied while possible (i.e. **committed planner**)

– Belief tracking is **monotone**

– Planner is **effective**: if executed prefix of σ doesn't reach goal, planner π has **asserting inference** for $(\tau[\sigma], \tau)$

Main Formal Result

Theorem

Let P be a **solvable task** *and π be a* **committed** *planner. If π is a* **weak** *and* **effective,** *and has* **monotone inference**, *then π is* **complete** *for P.*

Main Formal Result

Theorem

Let P be a **solvable task** *and π be a* **committed** *planner. If π is a* **weak** *and* **effective,** *and has* **monotone inference**, *then π is* **complete** *for P.*

Sketch: For each protocol $L = (P, s)$, planner in worst case generates a **sequence of beliefs** (associated to ongoing execution):

$$b_0^\pi \supseteq b_1^\pi \supseteq b_2^\pi \supseteq \cdots \supseteq b_n^\pi = \{s^*\}$$

that ends at **singleton**. Once there, since π is weak and committed, π generates and applies a plan for the current hidden state s^* QED

Another Result

Under **randomized protocols** where action selection is **stochastic** instead of just **non-deterministic:**

Theorem

Let P be a **strongly solvable** *task with* **observable goals** *and π be a planner. If π is a* **covering planner**, *then π is complete under randomized protocols*

Another Result

Under **randomized protocols** where action selection is **stochastic** instead of just **non-deterministic:**

Theorem

Let P be a **strongly solvable** *task with* **observable goals** *and π be a planner. If π is a* **covering planner**, *then π is complete under randomized protocols*

Sketch: Since task is strongly solvable, there is always a plan from current belief. Under assumptions, this plan can be "followed" with **non-zero probability.** Upon reaching a goal state, the agent will know it since goals are observable QED

Remark: there is no need for π to be weak or committed, or to have exact inference; it has to be covering though!

Experimental Results

See paper for details and experimental results on benchmarks

Wrap Up

- Framework for understanding and reasoning about online planning
- Preliminary theoretical results
- Played with planner LW1
- Future work:
  - Study necessary conditions for completeness
  - "Effectiveness" cannot be tested in an efficient manner
  - Novel action selection mechanisms
  - Novel tractable belief tracking methods

Lot of ground breaking work to be done in the area