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
- \( f : S \times A \rightarrow S \) is deterministic transition function
- \( O \) is finite set of observation tokens
- \( \Omega : S \times A \rightarrow O \) 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 $b_\tau = \text{states deemed possible after seeing execution } \tau$:

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

Belief tracking on factored models is intractable!
Online Planner: Closed-Loop Controller

\[ \pi(\tau) = \pi(P, \tau) \]
Two Components in Online Planners

Planner $\pi$

Belief Tracking

Action Selection

$\tau \leq b^{\pi}_\tau$ approx.

$\pi(\tau)$
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 $b_\tau^\pi \subseteq S_G$ (i.e. agent isn’t sure of reaching goal) do
4. Run planner $\pi$ on input $\tau$ 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 $s' := f(\text{Last}(\lambda), a)$ and token $o := \Omega(s', a)$
8. Update $\lambda := \langle \lambda, s' \rangle$ and $\tau := \langle \tau, a, o \rangle$

where $b_\tau^\pi$ is approximation of $b_\tau$ 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* $\pi$ *is complete on task* $P$ *if for each initial state* $s \in S_{init}$, *the protocol* $L(P, s)$ *terminates successfully* on $\pi$

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

- Is planner $\pi$ complete on $P$?
- Why isn’t $\pi$ complete on $P$?
- How do we make $\pi$ 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 $\tau$ 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 $b_\pi^\tau$ of $b_\tau$ 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 $\pi$ are exact; i.e., $b_\tau^\pi = b_\tau$ for each $\tau$

- **Monotone:** for every execution $\tau$ and prefix $\tau'$ of $\tau$, $|b_\tau^\pi| \leq |b_{\tau'}^\pi|$ (i.e. non-increasing “amount of uncertainty” along executions)

- **Asserting:** there is asserting inference for pair $(\tau, \tau')$ (where $\tau'$ is proper prefix of $\tau$) if $|b_\tau^\pi| < |b_{\tau'}^\pi|$ (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 $\tau$

First action on each sequence $\sigma$ must be applicable

Properties:

- **Committed**: by caching last computed sequences, the planner sticks to selected plan “as much as possible”

- **Weak**: for each approximation $b^\pi$:
  - each sequence $\sigma$ returned by $\pi$ is a plan for some state $s \in b^\pi$
  - if $b^\pi_\tau$ is non-empty, $\pi$ returns at least one sequence $\sigma$

- **Covering**: the first action in sequences returned by $\pi$ cover all applicable actions at exact belief $b_\tau$
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 $\pi$ is **weak**: given execution $\tau$, $\pi$ returns at least one plan $\sigma$ for some state $s \in b_\tau^{\pi}$ (state $s$ may not be in $b_\tau$)

- Plan $\sigma$ is applied while possible (i.e. **committed planner**)

- Belief tracking is **monotone**

- Planner is **effective**: if executed prefix of $\sigma$ doesn’t reach goal, planner $\pi$ has **asserting inference** for $(\tau[\sigma], \tau)$
Main Formal Result

<table>
<thead>
<tr>
<th>Theorem</th>
</tr>
</thead>
<tbody>
<tr>
<td>Let $P$ be a <strong>solvable task</strong> and $\pi$ be a <strong>committed planner</strong>. If $\pi$ is a <strong>weak and effective</strong>, and has <strong>monotone inference</strong>, then $\pi$ is <strong>complete</strong> for $P$.</td>
</tr>
</tbody>
</table>
Main Formal Result

**Theorem**

Let $P$ be a **solvable task** and $\pi$ be a **committed** planner. If $\pi$ is a **weak** and **effective**, and has **monotone inference**, then $\pi$ 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 $\pi$ is weak and committed, $\pi$ 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 $\pi$ be a planner. If $\pi$ is a **covering planner**, then $\pi$ 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 $\pi$ be a planner. If $\pi$ is a **covering planner**, then $\pi$ 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. 

**Remark:** there is no need for $\pi$ 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