Monitoring Plan Optimality using Landmarks and Domain-Independent Heuristics

Ramon Fraga Pereira†, Nir Oren‡, and Felipe Meneguzzi†

†Pontifical Catholic University of Rio Grande do Sul, Brazil
ramon.pereira@acad.pucrs.br
felipe.meneguzzi@pucrs.br

‡University of Aberdeen, United Kingdom
n.oren@abdn.ac.uk

February, 2017
Agents often **deviate from the optimal plan**, either because they have concurrent/multiple goals or because they are not perfect optimizers;

In this work, we develop an approach to **detect which actions during a plan execution do not advance (are non-optimal) to perform an optimal plan for achieving a monitored goal**;

Our contribution is twofold:

- We formalize this problem using **planning domain definition**; and
- We combine two planning techniques to solve this problem: **landmarks** and **domain-independent heuristics**.

We evaluate our approach using several planning domains, and show that our approach yields **high accuracy at low computational cost**.
Background: Planning, Heuristics, and Landmarks

Definition (Planning)

A planning instance is represented by a triple $\Pi = \langle \Xi, \mathcal{I}, G \rangle$, in which:

- $\Xi = \langle \Sigma, \mathcal{A} \rangle$ is the domain definition, and consists of a finite set of facts $\Sigma$ and a finite set of actions $\mathcal{A}$ (action costs $= 1$);
- $\mathcal{I}$ and $G$ represent the planning problem, in which $\mathcal{I} \subseteq \Sigma$ is the initial state, and $G \subseteq \Sigma$ is the goal state.

Heuristics are used to estimate the cost to achieve a particular goal. In this work, we use domain-independent heuristics;

Definition (Landmarks)

Given a planning instance $\Pi = \langle \Xi, \mathcal{I}, G \rangle$, a fact (or action) $L$ is a landmark in $\Pi$ iff $L$ must be satisfied (or executed) at some point along all valid plans that achieve $G$ from $\mathcal{I}$. 

Pereira, Oren, and Meneguzzi
Monitoring Plan Optimality
February, 2017 3 / 13
Definition (Plan Optimality Monitoring Problem)

- Domain definition (Facts and Actions) \( \Xi = \langle \Sigma, A \rangle \);
- Initial state \( I \);
- A monitored goal \( G \); and
- An observation sequence \( O = \langle o_1, o_2, ..., o_n \rangle \), representing a full observable plan execution;

The solution to a plan optimality monitoring problem is the set of observations (non-optimal actions) that do not advance an optimal plan that the agent may be following.
Plan Optimality Monitoring Approach

- Our approach combines planning techniques, i.e., landmarks and domain-independent heuristics.
- We use landmarks to obtain information about what cannot be avoided to achieve a monitored goal $G$; and
- We use heuristics to analyze possible plan execution deviation.
If an observation $o_i$ results a state $s_i$, we consider a deviation from a plan to occur if $h(s_{i-1}) < h(s_i)$.

![Graph](image.png)

- Estimated distance to the goal
- Observation time
- Optimal plan
- Sub-optimal plan
To predict which actions could be executed in the next observation, we analyze the closest landmarks by estimating the distance (using $h_{max}$) from the current state to the extracted landmarks $\mathcal{L}$, namely:

- For every fact landmark $l \in \mathcal{L}$ in which the estimated distance is 0, we select those actions $a \in \mathcal{A}$ such that $l \in \text{pre}(a)$; and
- For every fact landmark $l \in \mathcal{L}$ in which the estimated distance is 1, we select those actions $a \in \mathcal{A}$ such that $\text{pre}(a) \in \text{current state}$ and $l \in \text{eff}(a)^+$;

These predicted actions may reduce the distance to the monitored goal and next landmarks.
To detect sub-optimal steps (actions) in observation sequence $O$ for a monitored goal $G$, we combine the techniques we developed and filter with the following condition:

An observed action $o \in O$ is considered sub-optimal if:

$$o \notin \text{set of predicted actions AND } (h(s_{i-1}) < h(s_i)).$$
Algorithm 2 Plan Optimality Monitoring.

Parameters: $\Xi = \langle \Sigma, A \rangle$ planning domain, $I$ initial state, $G$ monitored goal, and $O$ observed actions.

Output: $A_{SubOptimal}$ as sub-optimal actions.

1: function MONITORPLANOPTIMALITY($\Xi, I, G, O$)
2:     $A_{SubOptimal} \leftarrow \langle \rangle$ ▶ Actions that do not contribute to achieve the monitored goal $G$.
3:     $L \leftarrow$ EXTRACTLANDMARKS($I, G$)
4:     $\delta \leftarrow I$ ▶ $\delta$ is the current state.
5:     $\eta_{PActions} \leftarrow$ NONREGRESSIVEACTIONS($\Xi, \delta, L$)
6:     $D_{G} \leftarrow$ ESTIMATEGOALDISTANCE($\delta, G$) ▶ A desired domain-independent heuristic to estimate goal $G$ from $\delta$.
7:     for each observed action $o$ in $O$ do
8:         $\delta \leftarrow \delta$.APPLY($o$)
9:         $D'_{G} \leftarrow$ ESTIMATEGOALDISTANCE($\delta, G$)
10:        if $o \notin \eta_{PActions} \land (D'_{G} > D_{G})$ then
11:           $A_{SubOptimal} \leftarrow A_{SubOptimal} \cup o$
12:        $\eta_{PActions} \leftarrow$ NONREGRESSIVEACTIONS($\Xi, \delta, L$)
13:        $D_{G} \leftarrow D'_{G}$
14:     return $A_{SubOptimal}$
We evaluate our approach over 10 planning domains;
- Precision: percentage of correctly detected sub-optimal steps;
- Recall: percentage of true sub-optimal steps, actually detected.

We use 6 domain-independent heuristics:
- $h_{adjsum}$, $h_{adjsum2}$, $h_{adjsum2M}$, $h_{combo}$, $h_{ff}$, and $h_{sum}$;

To extract landmarks and their ordering, we use an algorithm developed by Hoffman et al. (Ordered Landmarks in Planning. JAIR, 2004);

We manually generate the dataset from medium and large planning problems, containing both optimal and sub-optimal plan execution.
| Domain        | $|O|$ | $|L|$ | Heuristic         | Time   | Precision / Recall / F1 |
|---------------|-----|------|-------------------|--------|-------------------------|
| Blocks-World  | 15.2| 20.1 | $h_{adjsum2} / h_{ff}$ | 0.19 / 0.21 | 100% / 74.2% / 85.2% |
| Driver-Log    | 20.1| 53.6 | $h_{adjsum2M}$     | 1.33    | 100% / 100% / 100%     |
| Depots        | 16.7| 64.7 | $h_{adjsum2} / h_{ff}$ | 1.22 / 1.43 | 81.2% / 100% / 89.6% |
| Easy-IPC-Grid | 14.1| 48.5 | $h_{adjsum2} / h_{ff}$ | 0.77 / 0.86 | 100% / 100% / 100%     |
| Ferry         | 13.8| 18.1 | $h_{adjsum} / h_{sum}$ | 0.23 / 0.19 | 88.8% / 78.5% / 83.1% |
| Logistics     | 20.8| 24.0 | $h_{adjsum2} / h_{ff}$ | 0.35 / 0.55 | 100% / 91.3% / 95.4%  |
| Miconic       | 18.1| 19.4 | $h_{adjsum} / h_{sum}$ | 0.29 / 0.21 | 100% / 86.9% / 93.1%  |
| Satellite     | 25.7| 60.8 | $h_{adjsum2M}$     | 9.58    | 88.8% / 53.3% / 66.6% |
| Sokoban       | 24.0| 76.5 | $h_{combo}$        | 4.28    | 90.9% / 83.3% / 86.9% |
| Zeno-Travel   | 12.2| 38.7 | $h_{adjsum2} / h_{ff}$ | 0.86 / 0.99 | 100% / 92.8% / 96.2%  |

Table: Plan Optimality Monitoring experimental (best) results.
Conclusions

- **Contribution:**
  - Formalized plan optimality monitoring problem as planning;
  - Developed an approach based on landmarks and heuristics;
  - We show that our approach has high accuracy in almost all domains (besides SATELLITE).

- **Limitations:**
  - We do not yet deal with partial observability;

- **Future Work:**
  - Evaluate our approach using more modern domain-independent heuristics;
  - Try/use different landmark extraction algorithms; and
  - Apply our approach to goal recognition (online and offline).
Thank you!
Questions?