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

Introduction

- 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** Σ *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 Π iff L must be* **satisfied** *(or* **executed***) at some point along all valid plans that achieve G from $\mathcal{I}$.*

Background: Plan Optimality Monitoring Problem

Definition (Plan Optimality Monitoring Problem)

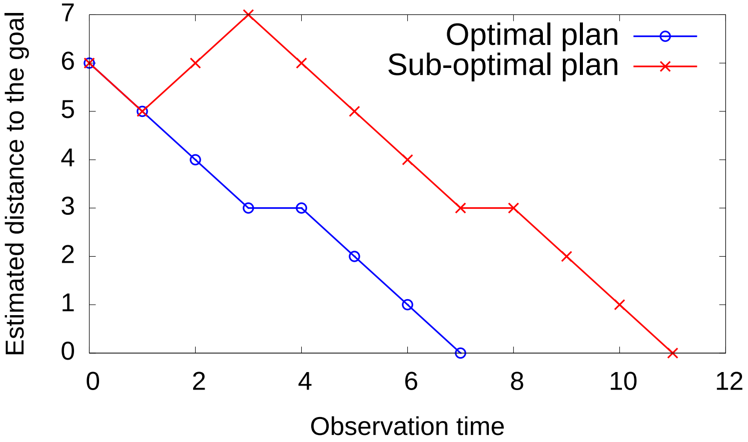
- Domain definition (Facts and Actions) $\Xi = \langle \Sigma, \mathcal{A} \rangle$;
- Initial state $\mathcal{I}$;
- A monitored goal G; and
- An observation sequence $O = \langle o_1, o_2, ..., o_n \rangle$, representing a full observable plan execution;

- 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 $h(s_{i-1}) < h(s_i)$.



Predicting Non-regressive Actions via Landmarks

- 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** $l \in$ **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** $pre(a) \in$ **current state and** $l \in$ **eff**$(a)^+$;
- These predicted actions **may reduce the distance** to the **monitored goal** and **next landmarks**.

Detecting Sub-Optimal Steps

- 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**: $o \notin$ set of predicted actions AND $(h(s_{i-1}) < h(s_i))$.

Detecting Sub-Optimal Steps (Monitor Plan Optimality)

Algorithm 2 Plan Optimality Monitoring.

Parameters: $\Xi = \langle \Sigma, \mathcal{A} \rangle$ *planning domain*, $\mathcal{I}$ *initial state*, G *monitored goal*, and O *observed actions*.
Output: $A_{SubOptimal}$ *as sub-optimal actions.*

function MONITORPLANOPTIMALITY(Ξ,$\mathcal{I}$,G,O)
    $A_{SubOptimal} \leftarrow \langle\rangle$ ▷ *Actions that do not contribute to achieve the monitored goal G.*
    $L \leftarrow$ EXTRACTLANDMARKS($\mathcal{I}$, G)
    $\delta \leftarrow \mathcal{I}$ ▷ *δ is the current state.*
    $\eta_{PActions} \leftarrow$ NONREGRESSIVEACTIONS(Ξ, δ, L)
    $D_G \leftarrow$ ESTIMATEGOALDISTANCE(δ, G) ▷ *A desired domain-independent heuristic to estimate goal G from δ.*
    for each observed action o in O **do**
        $\delta \leftarrow \delta$.APPLY(o)
        $D'_G \leftarrow$ ESTIMATEGOALDISTANCE(δ, G)
        if $o \notin \eta_{PActions} \wedge (D'_G > D_G)$ **then**
            $A_{SubOptimal} \leftarrow A_{SubOptimal} \cup o$
        $\eta_{PActions} \leftarrow$ NONREGRESSIVEACTIONS(Ξ, δ, L)
        $D_G \leftarrow D'_G$
    return $A_{SubOptimal}$

Experiments and Evaluation (1 of 2)

- 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	$\|O\|$	$\|\mathcal{L}\|$	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?