Process Mining Conference 2023
5th International Conference on Process Mining, 23-27 October, 2023

Plan Recognition as Probabilistic Trace Alignment

Jonghyeon Ko, Fabrizio Maria Maggi, Marco Montali, Rafael Peñaloza, Ramon Fraga Pereira

Freie Universität Bozen
Libera Università di Bolzano
Free University of Bozen · Bolzano
Università degli Studi di Milano Bicocca
Manchester 1824
The University of Manchester

Problem (1/2)

- Plan/Goal Recognition
  - Recognize plans and higher level goals of an agent from partial observations of the agent's behavior
  - Applications: strategic planning, intelligent user interfaces, story understanding, …

(Example)

Goal: handle first visit	**Goal: handle emergency patient**
Plans: - Reservation >> Reception Department >> Wait >> … - Reservation >> Reception Department >> Consultation >> … - ***Registration >> Reservation*** >> Reception Department >> … - …	Plans: - Severity classification >> Blood test >> … - Severity classification >> X - ray >> … - Severity classification >> CT/MRI >>… - …

Given a partial observation [***Registration >> Reservation***], what is the plan/goal of the agent?

Problem (2/2)

- Plan/Goal Recognition
  - input
    - ✓ $\boldsymbol{M}$: a reference model (e.g., a pddl model) describing possible behaviors of agents
    - ✓ $\boldsymbol{Obs}$: a partial sequence of events executed by an agent
  - output
    - ✓ **a set of plans** that “best explain” an observed partial sequence $\boldsymbol{Obs}$
    - ✓ **the most likely goals** an agent is aiming to achieve through the observed behavior

Existing solution (1/2)

- **Probabilistic Plan/Goal Recognition using classical planners**
  - M. Ramirez and H. Geffner, “Plan recognition as planning,” in IJCAI, 2009.
  - M. Ramirez and H. Geffner, “Probabilistic plan recognition using off the-shelf classical planners,” in AAAI, 2010.
- **Probabilistic Plan/Goal Recognition using process mining techniques**
  - A. Polyvyanyy, Z. Su, N. Lipovetzky, and S. Sardina, “Goal recognition using off-the-shelf process mining techniques,” in AAMAS, 2020

-> Assumption: closest plans are the most likely ones.

-> **Do not** consider probabilistic perspective of observations to select most probable plans

Existing solution (2/2)

- **Objective**: to find the most likely goal given an observed sequence
- **Solution:** $\underset{G}{argmax}$ ***Pr(G |σ)***

(the goal with maximum probability when given a partial sequence σ)

-> *Pr(G |σ) = α Pr(σ|G) Pr(G)* (Bayesian theorem)

$$\approx \frac{e^{-\beta \times costdiff(\sigma, G)}}{\sum_{G'} e^{-\beta \times costdiff(\sigma, G')}}$$

(Assumption: closest plans are the most likely ones)

where $\boldsymbol{costdiff(\sigma, G)}$ is distance between the optimal plan in $\boldsymbol{G}$ and $\boldsymbol{\sigma}$

(Example)

- Observed events: A-B-C
- Possible Plans in G1: {(A-B-C-D-E) $^?$, (A-B-C-D-F) $^?$}
- Possible Plans in G2: {(A-B-C-X) $^?$, (A-B-C-X-Y-Z) $^?$}

∴ **optimal goal = G2**

Our approach: Probabilistic Trace Alignment

- Plan/Goal Recognition
  - Input
    - ✓ $\boldsymbol{M}$: a **process model (or a set of model traces)** describing possible behaviors of agents
    - ✓ $\boldsymbol{Obs}$: a sequence of observed events, i.e., a trace prefix
  - Output
    - ✓ **a set of the closest model traces (=plans)** to the observed sequence $\boldsymbol{Obs}$
    - ✓ **the most likely goals** an agent is aiming to achieve through the observed behavior

    - ✓ (new) **a ranked list** of model traces
    - ✓ (new) the ranked list is built using both **probability** of model traces and **alignment cost**
    - ✓ (new) can handle observations with **faulty events**

Our approach: how to compute Pr(G| σ)

- Probabilistic Plan/Goal Recognition (computing $\underset{G}{argmax}\, Pr(G \mid \sigma)$)

-> $Pr(G \mid \sigma) = \alpha\, Pr(\sigma \mid G)\, Pr(G)$ (Bayesian theorem)

*(New assumption: plans that are **not only closer but also more frequent** are better predictors of G)*

$$\approx \alpha\, \widehat{Pr}(\sigma \mid G)\, \widehat{Pr}(G) \times \frac{e^{-\beta \times costdiff(\sigma, G)}}{\sum_{G'} e^{-\beta \times costdiff(\sigma, G')}}$$

(Example)

- Observed events: A-B-C
- Possible Plans in G1: {**(A-B-C-D-E)** ^{15}, (A-B-C-D-F) ^{5}}
- Possible Plans in G2: {(A-B-C-X) ^{3}, (A-B-C-X-Y-Z) ^{1}}

∴ **optimal goal = G1** ($\widehat{Pr}(\sigma \mid G1)$ = 1.0 (all plans start with A-B-C), $\widehat{Pr}(G1) = 20/24$)

Our approach: how to compute $\widehat{Pr}(\sigma|G)\,\widehat{Pr}(G)$

-> $Pr(G\,|\sigma) \approx \alpha\ \widehat{Pr}(\sigma|G)\ \widehat{Pr}(G) \times \frac{e^{-\beta \times costdiff(\sigma,G)}}{\sum_{G'} e^{-\beta \times costdiff(\sigma,G')}}$

If σ *= [A , B, C], then:*

(1)$\widehat{Pr}(\sigma|G) = \widehat{Pr}(B|A, G) \times \widehat{Pr}(C|B, G)$

(2)$\widehat{Pr}(G) = \frac{|\pi \in G|}{|\sum_{G'} \pi \in G'|}$

! But, there still exists a challenge in $\widehat{Pr}(\sigma|G)$.

Our approach: dealing with faulty events

- But, what if there exists a faulty event execution in σ? (e.g., σ = [a, b, c, ***X***])

-> ***Pr(σ|G) becomes 0.***

- To avoid it, we define :

-> ***Pr(σ|G)*** = ***Pr(σ^i|G)*** $\times$ $1/_{2^{|\sigma|-i}}$ *(= penalty), where* $\boldsymbol{\sigma^i}$ *is the largest compliant prefix of* $\boldsymbol{\sigma}$ *with length i.*

(Example)

- *Pr(σ|G)* = *Pr(σ^3|G)* $\times$ $1/_{2^{4-3}}$ with σ = [a, b, c, X].

Our approach: how to compute $costdiff(\sigma, G)$

-> $Pr(G \mid \sigma) \approx \alpha\, \widehat{Pr}(\sigma \mid G)\, \widehat{Pr}(G) \times \frac{e^{-\beta \times costdiff(\sigma,G)}}{\sum_{G'} e^{-\beta \times costdiff(\sigma,G')}}$

(3) $costdiff(\sigma, G) = levenshtein(\sigma, \pi^*)$ where $\pi^* = \underset{\pi \in G}{argmax}\ \mathcal{R}(\sigma, \pi)$.

i.e., π^* is the plan maximizing the ranking score $\mathcal{R}$ that takes into consideration both the probability of the plan given σ and the distance between the two.

-> $\mathcal{R}(\sigma, \pi) = Pr(\pi \mid \sigma) \times \frac{e^{-\beta \times levenshtein\,(\sigma,\pi)}}{\sum_{G'} e^{-\beta \times levenshtein\,(\sigma,\pi)}}$

(i.e., we consider both perspectives of probability and distance)

Experimental setting

➢ Datasets

Datasets	*# Goals*	*# Plan traces*	*# Variants*	*# Actions*
DAILY LIVING	8	27∼243	6∼43	14∼19
GRID NAVIGATION	9	47∼609	6∼102	12∼94
BLOCKS-WORLD	21	1000 (uniform)	1000	56∼94

- We split all logs into **60% of plan traces** and **40% of observations.**
- For experiment in stochastic setting, we injected frequency of traces following an exponential distribution.
  - Datasets for stochastic setting: DAILY LIVING, GRID NAVIGATION
  - Datasets for non-stochastic setting: BLOCKS-WORLD

Experimental setting

- **Baselines**

  (process mining approach)
  - TA 2020 [1] [**stochastic**, non-stochastic setting]

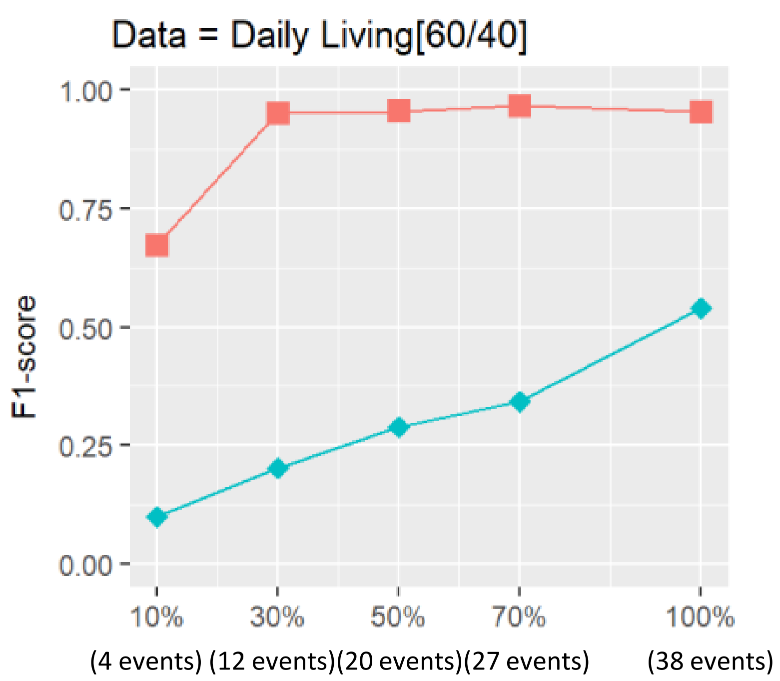
  (classic plan recognition)
  - R&G 2009 [2] [non-stochastic setting]
  - R&G 2010 [3] [non-stochastic setting]
  - POM 2017 [4] [non-stochastic setting]
  - LP 2021 [5] [non-stochastic setting]

[1] A. Polyvyanyy, Z. Su, N. Lipovetzky, and S. Sardina, "Goal recognition using off-the-shelf process mining techniques," in AAMAS, 2020.
[2] M. Ramirez and H. Geffner, "Plan recognition as planning," in IJCAI, 2009
[3] M. Ramirez and H. Geffner, "Probabilistic plan recognition using off the-shelf classical planners," in AAAI, 2010
[4] R. Pereira, N. Oren, and F. Meneguzzi, "Landmark-based heuristics for goal recognition," in AAAI, vol. 31, 2017.
[5] L. R. d. A. Santos, F. Meneguzzi, R. F. Pereira, and A. Pereira, "An LP-Based Approach for Goal Recognition as Planning," in AAAI, 2021.

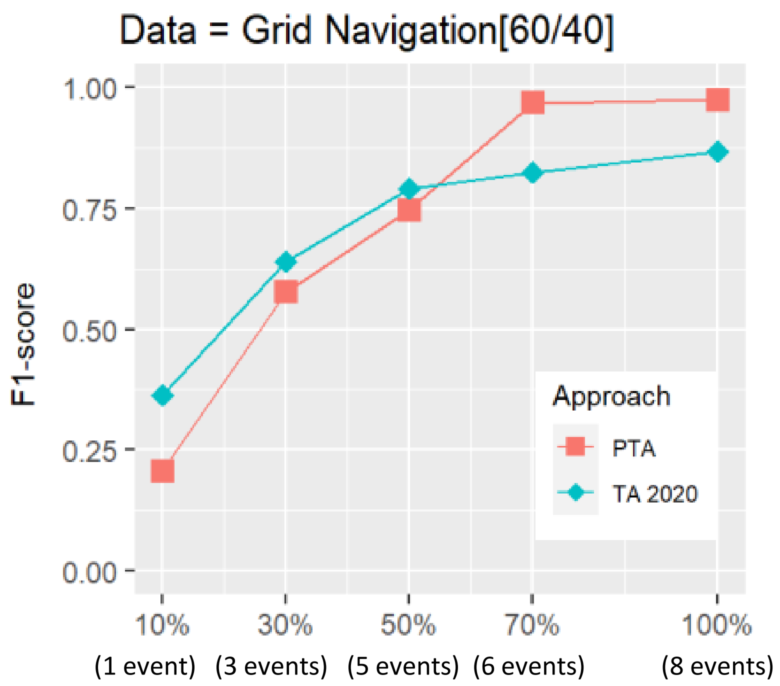
Experimental results (baseline comparison)

1. Performance of goal recognition in the stochastic setting (main contribution)

- Overall, PTA performs better **in terms of F1-score** than TA2020
- TA 2020 sometimes performs better in GRID NAVIGATION dataset



% observation percentage
(# events: average prefix length)



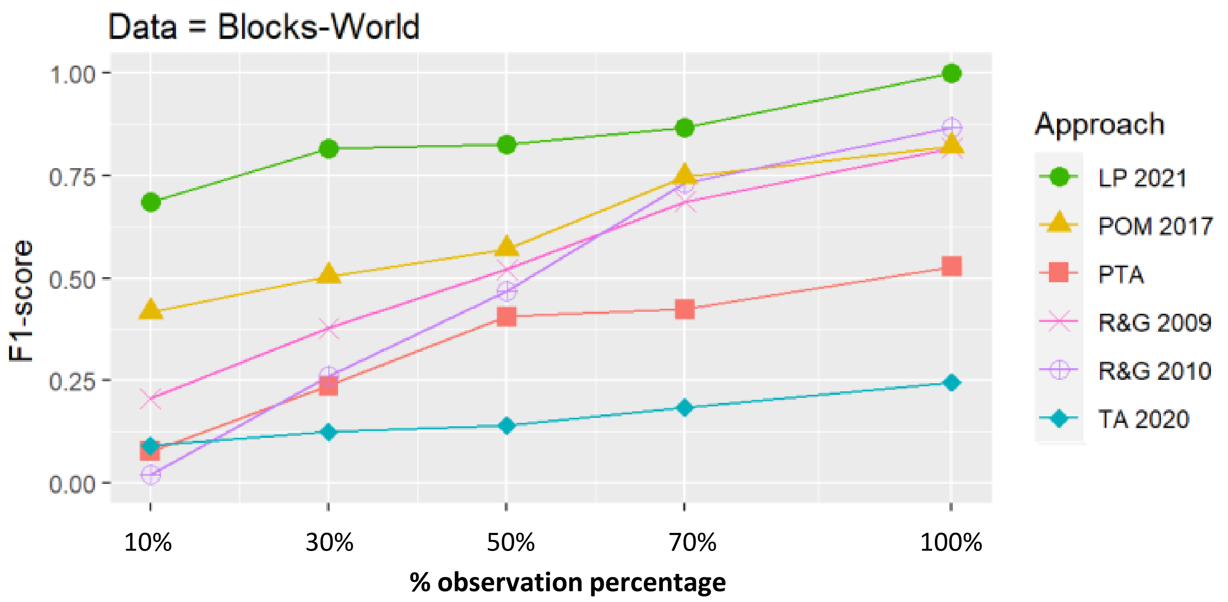
% observation percentage
(# events: average prefix length)

Experimental results (baseline comparison)

2. Performance of goal recognition in the non-stochastic setting

- PTA performs better in terms of **F1-score** than TA2020
- Overall, **classic planning approaches perform better than process mining based approaches**

- However, classic planning approaches
 1. require a pddl model which is not possible to be used in stochastic setting
 2. do not consider faulty events



Conclusion

➢ **Plan/Goal Recognition using Probabilistic Trace Alignment**

- Process mining approaches can be used effectively when PDDL model is not available.
- Our approach can consider **the probabilistic perspective of observations**, which has not been considered in classical planning problem.
- Our approach considers **faulty observations**.

➢ **FUTURE WORK**

- Consider the case that an agent adopts a stochastic policy to choose its next action.
- Consider the case of non-deterministic actions for agents.
- Apply the approximate probabilistic trace alignment presented in [6] to improve the execution time.

[6] Giacomo Bergami, Fabrizio Maria Maggi, Marco Montali, Rafael Peñaloza: Probabilistic Trace Alignment. ICPM 2021: 9-16