Failure Avoidance through Learning Guided Probabilistic Planning

Melis Kapotoglu, Cagatay Koc, and Sanem Sariel

***Abstract*— Robots need to gain experience online, and use this experience to improve their task execution performance to prevent undesired outcomes. In this work, we present an adaptive planning method for robots that uses the outcomes of a real-world learning process during probabilistic planning to achieve this objective. We analyze this method on a case study with our autonomous mobile robot in a multi-object manipulation domain. The results show that the presented solution ensures efficiency in mobile manipulation performance.**

I. INTRODUCTION

The main objective of this work is to develop an experience-guided planning method for a robot to improve its performance on manipulation tasks. In this method, the robot builds its experience by learning through observations made after each action execution. We selected a mobile manipulation case study where the focus is on optimizing the number of manipulated objects in the face of failures. This can be achieved by using the built experience from previous failure cases to guide future planning tasks of the robot.

Our prior work involves an experimental learning process [1] using Inductive Logic Programming (ILP) and a deterministic planner that uses the built experience [2]. The previous deterministic planner takes contexts of hypotheses into account to provide feedback to planning. The method that we propose here differs from our earlier study with the development of a probabilistic planner framework to use hypotheses associated with probabilities built by learning. Our probabilistic framework uses a Partially Observable Markov Decision Process (POMDP) model to create an adaptive policy for the robot to deal with uncertainties. We use a new point-based algorithm named SARSOP [3] as a POMDP planner in our learning-guided planning framework.

In the literature many studies can be found in which POMDPs have been investigated for the object manipulation task. In [4] and [5], manipulation of multiple objects is selected as the main challenge, and an online POMDP planning approach is proposed to change system dynamics according to action performances. Our work differs from these works in the way we guide the planning process with experience and apply a generic method that can be applicable for all types of actions in a planning domain without changing state definitions. Some studies [6], [7] aim to generate planning operators which are represented as probabilistic relational rules to reduce or exclude the possibility of failures by applying adaptive planning algorithms. Contrary to these studies, our system makes use of real-world observations in the learning process. We demonstrate that our system on an autonomous mobile robot can successfully use the real-world experience on guiding the planning process.

This research is funded by a grant from the Scientific and Technological Research Council of Turkey (TUBITAK), Grant No. 111E-286.

Melis Kapotoglu, Cagatay Koc, and Sanem Sariel are within the Department of Computer Engineering, Istanbul Technical University, Istanbul, Turkey. {kapotoglu,kocca,sariel}@itu.edu.tr

II. EXPERIENCE-BASED GUIDANCE IN MOBILE MANIPULATION

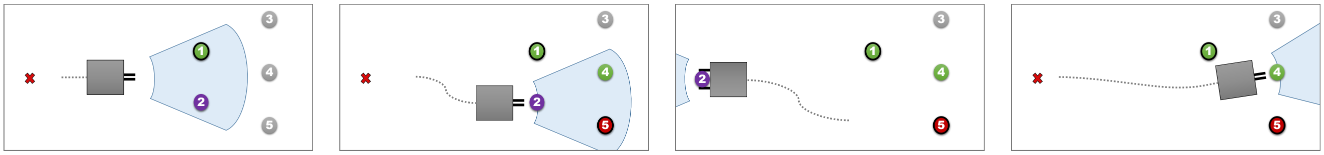
Our case study involves mobile manipulation scenarios by an autonomous robot whose goal is to maximize the number of objects transported to a destination while avoiding failures in a given time period. In our system, a consistent world model is maintained during runtime in the face of the challenges of noise in sensory data, partial observability and unexpected situations by a scene interpretation process [8]. Additionally, action execution is monitored by another process in which metric temporal formulas are defined specifically for each action and, failures occurred during execution are detected simultaneously [9].

In our study, the robot selects the preference order of objects to be manipulated and plans to execute $moveTo$, $pickUp$ and $transport$ actions in sequence. Whenever a failure is detected during each transportation sequence, the corresponding observation along with its related context is encoded in the knowledge base of the robot. After a certain number of observations, the robot builds its experience by learning to determine its abilities on object manipulation, and derive hypotheses on failure cases. Each hypothesis relates a failure context (the antecedent part) to the outcome of an action (the conclusion). An example hypothesis is given as:

$$category(cylindricalObj) \wedge color(green) \Rightarrow failure(pickUp) \text{ (P: 0.22)}$$

where category and color attributes with $cylindricalObj$ and $green$ values respectively represent the failure context for the failure of action $pickUp$, and this rule is associated with probability 0.22.

Our main contribution lies in the POMDP planning formulation developed to use learning outcomes. In our formulation, an object composition is defined to represent encapsulation of qualitative and spatial information about an object. Qualitative information specifies the predetermined attributes (type, color, material, height, width, etc.) and the spatial information represents the semantic location of an object. For all combinations of object attributes and semantic locations, different states are created. The combination sets of all possible object compositions construct the state space of POMDP formulation. The state definition is expanded with two more states, one of which represents the possible failure cases during action execution. A transition to this state takes place whenever a failure is detected in the action



(a) (b) (c) (d)

Fig. 1. Snapshot sequence of a plan execution. Objects are represented as circles which are enumerated respectively for the green bowling pin, the purple ball, the red big ball, the green cylinder and the red bowling pin. Red cross represents the destination. The blue region indicates the field of view of the robot where the objects can be recognized. The objects colored gray indicate unseen objects. The circles with bold edges denote the objects with lower probability of $pickUp$ success. The recognized objects are shown in their respective colors.

execution. The other state specifies that the robot is holding an object. The purpose of separating these states is creating an abstraction from irrelevant state components to decrease the number of states generated by the POMDP.

The preference orders of objects can be determined by either prioritizing objects based on their locations in order to minimize the travelled distance or primarily targeting objects on which the robot has the best manipulation performance. Success probabilities for action $pickUp$ related to different contexts are determined by learning and used while assigning probabilities to state transitions in the POMDP model. The distance between the location of the robot and a corresponding object to be manipulated is another factor which affects the manipulation timing and performance. Therefore, semantic locations of the objects are also used to decide on the next object to manipulate by determining probabilities of action $moveTo$ according to distances.

III. EXPERIMENTS

We evaluate the performance of our system on our Pioneer 3-DX mobile robot in a dynamic environment with varying changing illumination conditions. The robot is equipped with a laser rangefinder and an RGB-D camera to perceive its environment and a 2-DOF gripper to manipulate objects. The performance of the robot is tested on a set of five different objects from various colors and categories: two plastic bowling pins whose colors are green and red, a green plastic cylindrical object, a small purple ball and a big red ball. The actions $moveTo$, $pickUp$ and $transport$ are executed ten times for the object set randomly cluttered in the environment to measure the overall performance of the robot. After preliminary results, an experiment is performed by injecting failures, and the performance of the system is analyzed with and without learning on the same failure probability distribution which is determined randomly.

In the experiment, objects with $category(bowlingPin)$ are externally taken away from the environment with a predetermined probability at the time of action $pickUp$. The graphical illustration of this scenario is given in Fig. 1. After executing action $search$ to find objects in the environment, the robot detects the purple ball and the green bowling pin (Fig. 1(a)). According to the success probabilities of action $pickUp$ for different objects that are determined with experience, the robot selects the purple ball to move to, since its $pickUp$ success probability is higher (Fig. 1(b)). The robot detects farther objects while moving to the purple ball. After transporting the purple ball successfully (Fig. 1(c)), the robot selects the cylindrical object with the higher success probability instead of the closer bowling pin (Fig. 1(d)). If there is enough remaining time after transporting the objects with higher success probability, the robot tries to manipulate the objects with low success probability.

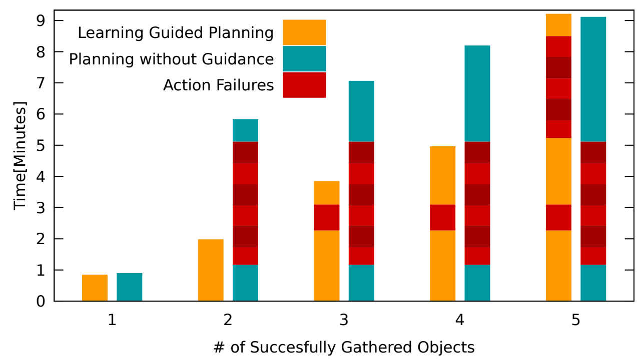


Fig. 2. Time comparison between planning with and without learning guidance in accordance with the number of successfully gathered objects

The described experiment is also performed without learning and the comparison of the planning with and without learning is given as a histogram chart in Fig. 2. Here, the number of successfully transported objects is shown in relation to time. The red intervals in the histogram bars indicate delays due to failures. Although the total required time to manipulate all objects scattered in the environment are approximately the same and the number of manipulation trials are equivalent in both cases, learning guided planner postpones the manipulation of the objects with lower success probabilities until no other objects are available. On the other hand, the planner without guidance chooses the closest object to manipulate, since the success probabilities of all the observed objects are believed to be equal. Therefore, the first failure is encountered earlier in the plan without guidance than that of the learning-guided planning system.

IV. CONCLUSION

We presented our adaptive probabilistic planning framework that uses the outputs of an experimental learning process by an autonomous robot. Our results show that the probabilistic guidance in planning achieves the best manipulation order of the objects to maximize the transportation success over time in a real world scenario.

REFERENCES

[1] S. Karapinar, D. Altan, and S. Sariel-Talay, "A robust planning framework for cognitive robots," in *Proceedings of the AAAI-12 Workshop on Cognitive Robotics (CogRob)*, 2012.

[2] P. Yildiz, S. Karapinar, and S. Sariel-Talay, "Learning guided symbolic planning for cognitive robots," in *The IEEE International Conference on Robotics and Automation (ICRA), Autonomous Learning Workshop*, 2013.

[3] H. Kurniawati, D. Hsu, and W. S. Lee, "Sarsop: Efficient point-based pomdp planning by approximating optimally reachable belief spaces." in *Robotics: Science and Systems*, vol. 2008. Zurich, Switzerland, 2008.

[4] J. Pajarinen and V. Kyrki, "Robotic manipulation of multiple objects as a pomdp," *arXiv preprint arXiv:1402.0649*, 2014.

[5] ——, "Robotic manipulation in object composition space," in *Intelligent Robots and Systems (IROS 2014), 2014 IEEE/RSJ International Conference on*. IEEE, 2014, pp. 1–6.

[6] H. Pasula, L. S. Zettlemoyer, and L. P. Kaelbling, "Learning probabilistic relational planning rules." in *ICAPS*, 2004, pp. 73–82.

[7] T. Lang and M. Toussaint, "Planning with noisy probabilistic relational rules," *Journal of Artificial Intelligence Research*, vol. 39, no. 1, pp. 1–49, 2010.

[8] M. D. Ozturk, M. Ersen, M. Kapotoglu, C. Koc, S. Sariel-Talay, and H. Yalcin, "Scene interpretation for self-aware cognitive robots," in *AAAI-14 Spring Symposium on Qualitative Representations for Robots*, 2014.

[9] M. Kapotoglu, C. Koc, S. Sariel, and G. Ince, "Action monitoring in cognitive robots (in turkish)," in *The 22nd Signal Processing and Communications Applications Conference (SIU)*, 2014, pp. 2154–2157.