Integrating Open-ended Learning and Planning for Long-Term Autonomy

(Article begins on next page)

Integrating Open-ended Learning and Planning for Long-Term Autonomy

Gabriele Sartor[1]

Abstract. Classical planning is still a powerful tool able to perform rather complex reasoning on domains defined by a high-level representation. However, its main problem is the lack of flexibility in the definition of the domain. Once the representation of the world is defined by the expert, the capabilities of the agent are fixed and, consequently, also its potentially achievable goals. For this reason, many researchers have shifted their attention on developing systems able to produce autonomously a high-level representation of the world, resulting from the experience gathered during the interaction with the surrounding environment. IMPACT (Intrinsically Motivated Planning Architecture Curiosity-driven roboTs)[2] has been our first attempt to implement a software architecture using high-level planning and able to extend its operational capabilities.

1 INTRODUCTION

In the last decades, traditional planning robotic systems have been developed providing the agent the knowledge necessary to perform tasks defined at design time. However, in some cases, particularly in space exploration missions [6], the agent could need to deal with unforeseeable situations or simply detect a variation in the environment dynamics without the human intervention.

Consequently, many researchers have started to study new methods to abstract knowledge learned during the interaction with real world in order to encapsulate the complexity under an easier representation, suitable for performing high-level planning. Recently, some methods have been proposed to translate the agent's experience in PDDL representation [7]. For example, it has been proposed an algorithm translating low-level information about the initial and final states of each action, called options, into a fully working propositional symbolic planning domain [5]. Some works have also started to reconcile deep learning with more abstract representations [4]. For instance, a system demonstrated the possibility to autonomously generate a first-order logic (FOL) representation compatible with symbolic classical planning, using neural networks [1]. In this case, the most important element developed has been a particular autoencoder, able to transform the feature vectors of the objects visualized in images into a FOL description, and vice versa.

The project IMPACT [9, 8] has extended the previous work creating an open-ended learning system [3] able to learn new capabilities interacting with a simulated environment, abstract this knowledge into propositional PDDL as in [5] and plan on it.

[1] Università degli Studi di Torino, Italy, email: gabriele.sartor@unito.it

[2] This research has been supported by the European Space Agency (ESA) under contract No. 4000124068/18/NL/CRS, project IMPACT (Intrinsically Motivated Planning Architecture for Curiosity-driven roboTs). The view expressed in this paper can in no way be taken to reflect the official opinion of the European Space Agency.

2 CURRENT RESEARCH

In the last years, the importance of space missions is increasing. In particular, the next most challenging missions are focusing on the exploration of Mars. The robots designated to perform the experiments on-site will have to be equipped with all the knowledge necessary to deal with their duties with a high degree of autonomy because of the latency of communications between the Earth and Mars. For this reason, it is important to build systems as autonomous as possible, but also able to extend their operational capabilities on-site in order to face changes in the environment, unforeseeable events and increase the utility of the mission discovering new aspects of the world and achieve goals unknown at design time. IMPACT has been our first try to develop such software architecture.

This system is implemented as a three-layered architecture integrating the following main modules:

1. *Planning*, reasoning on the initial knowledge of the world and operational capabilities included in the agent at design time used to reach the mission goals;
2. *Abstraction*, translating the low-level data gathered during the mission into a propositional domain of the environment;
3. *Learning*, responsible for learning new skills, triggered by the robot's curiosity [10].

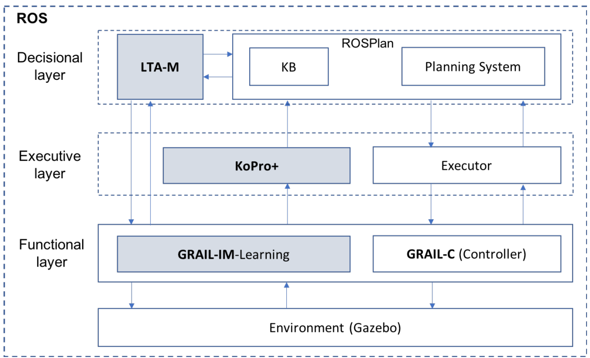

Figure 1. IMPACT high-level software architecture.

In figure 1 it is shown the architecture of our software system, divided in three layers. The lower one, *the Functional layer*, contains the modules dealing with the low-level information of the environment. It implements the controller of the robot and the learning ability integrating the GRAIL system [10], an intrinsically motivated reinforcement learning (IM-RL) component able to extend the competences of the agent. This component can autonomously discover interesting states of the environment and, using a competence-based

Proceedings of the 1st Doctoral Consortium at the European Conference on Artificial Intelligence (DC-ECAI 2020), pages 27–28, Santiago de Compostela, Spain, August 29-30, 2020.

IM-RL algorithm, acquires a new skill to reach that significant configuration.

Instead, the *Executive layer* is a mapping level in which (i) the high-level planning operators are translated into commands for the controllers and (ii) low-level information is transformed into a PDDL representation of the world [5], suitable for planning. In particular, the abstraction of the low-level information is crucial for such systems, because it defines the type of representation and, consequently, the limitations of the higher level. For example, despite its simplicity, the propositional representation can limit the expressiveness of the higher level, given that it is not possible to generalize similar actions and features.

The high-level planning domain and the planning system, implemented using the ROSPlan framework [2], are included inside the *Decisional layer*. This level is responsible for deciding how the high-level goals have to be achieved and controlling the execution feedbacks received from the lower layers, in order to potentially replan its activities. The top layer has also to decide the schedule of the goals to be achieved. For this purpose, it has been developed also a simple component called *Long-Term-Autonomy Module* (LTA-M). The idea of such component is to create a decisional module selecting the next tasks to be performed in the future, in order to increase as much as possible its satisfaction in terms of mission goals and extension of operational capabilities. In other words, it implements strategies to alternate the achievement of intrinsic and extrinsic goals.

The system described in this paper has been tested in two different space exploration scenarios. In particular, we simulated a scenario in which a robotic arm is sent to Mars in order to pickup some stone samples necessary for its experiments [8]. In the simulation, we assume that the robot is designated to explore two interesting valleys in which the scientists of the mission control center want to perform analysis of the ground. In the first valley, the agent is able to pickup samples of stones without any difficulty, while in the second one the stones have a particular concave shape. During its activities, the agent tries to pickup a vase-shaped stone in the second valley, discovering that it is not able to grasp with its current capabilities because of its dimensions, failing to reach its high-level goal. The failure triggers its curiosity towards the unexpected situation and tries to learn a new way to grasp the unknown object. The intrinsic motivated component GRAIL [10] detects a curious state and stores it as intrinsic goal. When the agent decides to focus on that interesting state, it will try to reach again that situation with a high level of confidence. After several attempts, the system learns to grasp the new object and, exploiting the execution layer, to abstract this new discovered aspect of world creating a new symbol and operator to deal with it. In this way, it has been demonstrated that the architecture has been able to extend its operational capabilities and the potential of this system.

3 FUTURE WORKS

The architecture presented in the paper could be still improved in different aspects. One of the most important limitations of the system is the use of a high-level planning propositional representation. This approach is not effective in terms of scalability, because each aspect of the world has to be represented with a specific symbol. Instead, first-order logic representation can use parametrized symbols and operators to generalize some aspects in a more compact description. Given the potential of the autoencoders in generating a representation from the world [4], we will examine the possibility of combining them with open-ended learning to deal with such limitation.

Another component to be examined in depth is the LTA-M [6], which still presents a simple form. We will examine the possible strategies to manage the long-term needs of the agent based on the goal required by the user and its curiosity. A possible extension could be exploiting the high-level planning to reach less explored states in order to increase the possibility of encountering new intrinsic goals and, consequently, learning new aspects of the world. Lastly, our intention is to generalize this architecture in order to extend the range of applicability of this methodology.

4 CONCLUSION

We proposed an open-ended learning architecture able to learn new competences, abstract this knowledge and reasoning over this new representation extending its operational capabilities. Robotics system autonomy is a branch of research of increasing importance. However, in this field, autonomy and interpretability seem to be two orthogonal aspects[1]. We think that an abstract representation of the domain will be fundamental for keeping track of the agent's improvement caused by its learning components. Consequently, we intend to improve the robustness of this architecture and generalize it for a wider range of applications.

ACKNOWLEDGEMENTS

For more details on the system implementation, please see the main conference article *"Integrating open-ended learning in the sense-plan-act robot control paradigm"* [8].

REFERENCES

[1] Masataro Asai, 'Unsupervised grounding of plannable first-order logic representation from images', in *Proceedings of the International Conference on Automated Planning and Scheduling*, volume 29, pp. 583–591, (2019).

[2] M. Cashmore, M. Fox, D. Long, D. Magazzeni, B. Ridder, A. Carrera, N. Palomeras, N. Hurtos, and M. Carreras, 'Rosplan: Planning in the robot operating system', in *Twenty-Fifth International Conference on Automated Planning and Scheduling*, (2015).

[3] S. Doncieux, D. Filliat, N. Díaz-Rodríguez, T. Hospedales, R. Duro, A. Coninx, D. M. Roijers, B. Girard, N. Perrin, and O. Sigaud, 'Open-ended learning: a conceptual framework based on representational redescription', *Frontiers in neurorobotics*, **12**, 59, (2018).

[4] M. Garnelo and M. Shanahan, 'Reconciling deep learning with symbolic artificial intelligence: representing objects and relations', *Current Opinion in Behavioral Sciences*, **29**, 17 – 23, (2019). SI: 29: Artificial Intelligence (2019).

[5] G. Konidaris, L. P. Kaelbling, and T. Lozano-Perez, 'From skills to symbols: Learning symbolic representations for abstract high-level planning', *Journal of Artificial Intelligence Research*, **61**, 215–289, (2018).

[6] L. Kunze, N. Hawes, T. Duckett, M. Hanheide, and T. Krajník, 'Artificial intelligence for long-term robot autonomy: A survey', *IEEE Robotics and Automation Letters*, **3**(4), 4023–4030, (2018).

[7] D. Mcdermott, M. Ghallab, A. Howe, C. Knoblock, A. Ram, M. Veloso, D. Weld, and D. Wilkins, 'PDDL - The Planning Domain Definition Language', Technical report, CVC TR-98-003/DCS TR-1165, Yale Center for Computational Vision and Control, (1998).

[8] A. Oddi, R. Rasconi, V.G. Santucci, G. Sartor, E. Cartoni, F. Mannella, and G. Baldassarre, 'Integrating open-ended learning in the sense-planact robot control paradigm', *24th European Conference on Artificial Intelligence, ECAI 2020*, (2020).

[9] A. Oddi, R. Rasconi, V.G. Santucci, G. Sartor, E. Cartoni, F. Mannella, and G. Baldassarre, 'An intrinsically motivated planning architecture for curiosity-driven robots', *CEUR Workshop Proceedings*, **2594**, 19–24, (2020).

[10] V. G. Santucci, G. Baldassarre, and M. Mirolli, 'Grail: a goal-discovering robotic architecture for intrinsically-motivated learning', *IEEE Transactions on Cognitive and Developmental Systems*, **8**(3), 214–231, (2016).

Proceedings of the 1^{st} Doctoral Consortium at the European Conference on Artificial Intelligence (DC-ECAI 2020), pages 27–28, Santiago de Compostela, Spain, August 29-30, 2020.