Interactive On-line Semantic Mapping

Emanuele Bastianelli, Domenico Bloisi, Roberto Capobianco, Guglielmo Gemignani,
Luca Iocchi, Daniele Nardi

Robots are expected to be the next technological frontier, especially for the impact that they will have on our society. In particular, they are expected to become consumer products, massively entering our homes to be part of everyday life. The first robotic platforms have already been released on the market for consumer applications, such as cleaning or entertainment. Such robots however still lack the intelligence expected by most of the users.

One of the key problems preventing the acceptance of robotic platforms in domestic environments is their inability to know and be aware of their surrounding. Although their perception capabilities are rapidly improving, they are still significant limited in categorizing environments and in recognizing the objects therein. In other words, the world models that robots are able to build do not support even simple forms of common sense knowledge and reasoning. These particular difficulties arise from several sources, such as the limitation of perception systems, the difficulties in communicating with humans, or the inability to acquire, maintain, and use symbolic knowledge. Our long term research goal is to improve the performance of robotic systems by building an explicit representation of symbolic knowledge and, in particular, to devise *knowledgeable robots* [1].

One of the main problem that must be faced in developing knowledgeable robots is *symbol grounding*. The symbols used in the internal representation of a robot need in fact to be connected to the proper elements in the operational scenario. In the case of robots operating in a physical environment, symbol grounding means to "close the loop" with perception, thus defining the relationship between numeric data coming from the sensors and the symbolic representation.

In order to enable symbol grounding, knowledge about the world must be encoded inside the robots. The acquisition of knowledge about the environment is the goal of the so-called *semantic mapping* process [7]. The resulting representation of such a process is typically a map, called *semantic map*, that integrates the information needed for navigation with labels that allow to semantically characterize places and objects in the environment. Semantic maps however currently suffer from the limitations of robotic perception capabilities. Moreover, current methods for semantic mapping are usually designed to work off-line, focusing on the process that enables the robot to build a suitable representation before operating in a given environment.

Approaches to acquire and build semantic maps usually belong to two different categories. The first category mainly relies on *fully automatic* methods, where human interaction is not taken into account. These methods aim at acquiring features of the environment and automatically segmenting the mapped rooms, generally relying on different set of techniques, such as classification and clustering, or exploiting visual features[2, 3, 10]. Although significant progress has been made in fully automated semantic mapping, even the most recent approaches still lack robustness and generality.

The second category includes instead *human augmented mapping* techniques, where the user actively supports the robot in acquiring the required knowledge about the environment. In such approaches, the user is exploited for grounding symbols to objects that are still autonomously recognized by the robotic platform. This is coherent with the paradigm of *symbiotic autonomy* [9], in which robots are aware of their limitations and proactively ask humans for help. In such approaches, the interaction is usually unimodal and is typically achieved through natural language [11], [8], [6]. Few approaches aim at a multi-modal human-robot collaboration, where the interaction is used not only for place categorization and labeling, but also for object recognition and positioning. Such kind of interactions are more complex and require natural paradigms to avoid a tedious effort from the users [5], [4].

In our work, the idea is to exploit a laser-pointer together with a natural language interaction with the user to overcome the limitations of the robot. As compared with previous work, our semantic mapping approach can be seen as an *on-line* process, in which a rich and detailed representation of the operative scenario is incrementally built with the help of the user. Our goal is to build an accurate representation of the specific environment, which does not necessarily need to be coherent with an *a priori* defined general knowledge. The resulting integrated representation enables the robot to understand the position of the objects found in the environment and to robustly navigate by exploiting a specifically built topological graph. The approach described in this paper is based on two main components: *(i)* a component for Simultaneous Localization And Mapping (SLAM), that provides a metric map of the environment; *(ii)* a multi-modal interface that allows the user to point at the elements of the environment and to assign them a semantic role.

Our work presents different novel contributions. First, the process of building and updating semantic maps, which is done incrementally during the deployment of a robot through a continuous interactive process, which exploits several forms of complex multi-modal interaction. Second, we propose a new way of representing both the actions of the robot and the dialogs with the user, relying on the Petri Net Plans (PNPs) formalism. Third, we handle object disambiguation through the aid of a spatial reasoner, needed to understand spatial referring expression used in the dialogs. Finally, we propose a multi-layered representation of the environment, including a *Semantic Grid Map* and a *Topological Graph* which are automatically extracted from the labeled semantic map and that can be used to enhance navigation.

The above listed features have been embedded into a prototype system that has been extensively used over months to validate the proposed approach. This system has been deployed in substantially different environments and has been used by multiple users that helped to acquire new knowledge about the environment. In the long term, we believe that the proposed approach will resemble a first step towards the implementation of long-life learning robots, that improve their performance through living with the humans.

References

1. Aiello, L.C., Nardi, D., Randelli, G., Scalzo, C.: Suppose you have a robot. Gerhard Lakemeyer and Sheila A. McIlraith (Eds.) Knowing, Reasoning, and Acting, Essays in Honour of Hector J. Levesque pp. 27–46 (2011)

2. Galindo, C., Saffiotti, A., Coradeschi, S., Buschka, P., Fernández-Madrigal, J., González, J.: Multi-hierarchical semantic maps for mobile robotics. In: Proceedings of the IEEE/RSJ Int. Conf. on Intelligent Robots and Systems (IROS). pp. 3492–3497 (2005)
3. Goerke, N., Braun, S.: Building semantic annotated maps by mobile robots. In: Proceedings of the Conference Towards Autonomous Robotic Systems. pp. 149–156 (2009)
4. Hemachandra, S., Kollar, T., Roy, N., Teller, S.: Following and interpreting narrated guided tours. In: Proceedings of the IEEE International Conference on Robotics and Automation. pp. 2574–2579. Shanghai, China (2011)
5. Kruijff, G., Zender, H., Jensfelt, P., Christensen, H.: Clarification dialogues in human-augmented mapping. In: Proceedings of the 1st Annual Conference on Human-Robot Interaction (HRI'06). pp. 282–289 (2006)
6. Nieto-Granda, C., III, J.G.R., Trevor, A.J.B., Christensen, H.I.: Semantic map partitioning in indoor environments using regional analysis. In: IEEE/RSJ International Conference on Intelligent Robots and Systems. pp. 1451–1456 (2010)
7. Nüchter, A., Hertzberg, J.: Towards semantic maps for mobile robots. Robot. Auton. Syst. 56(11), 915–926 (2008)
8. Pronobis, A., Jensfelt, P.: Large-scale semantic mapping and reasoning with heterogeneous modalities. In: Proceedings of the 2012 IEEE International Conference on Robotics and Automation (ICRA'12). pp. 3515–3522 (2012)
9. Rosenthal, S., Biswas, J., Veloso, M.: An effective personal mobile robot agent through symbiotic human-robot interaction. In: Proceedings of 9th International Joint Conference on Autonomous Agents and Multi-Agent Systems (AAMAS). pp. 915–922 (2010)
10. Wu, J., Christenseny, H.I., Rehg, J.M.: Visual place categorization: Problem, dataset, and algorithm. In: Proceedings of IEEE/RSJ Conference on Robots and Systems (IROS). pp. 4763–4770 (2009)
11. Zender, H., Martínez Mozos, O., Jensfelt, P., Kruijff, G., Burgard, W.: Conceptual spatial representations for indoor mobile robots. Robotics and Autonomous Systems 56(6), 493–502 (2008)