Semantic SLAM for Collaborative Cognitive Workspaces

Frank Dellaert
College of Computing
Georgia Institute of Technology

David Bruemmer
INEEL: Idaho National
Engineering and Environmental Laboratory

A Common Cognitive Workspace

In order to create a meaningful interface between robots and humans, we advance the notion that a "collaborative cognitive workspace" should form the basis for a dynamic information exchange between robots and their operators. One way to achieve this is through mapping the environment at the semantic level. Instead of traditional SLAM methods, which do not interpret sensor information other than at the geometric level, we propose that semantic mapping can provide the means to collaboratively build a shared understanding of the environment. By doing this, members of a heterogeneous robot-operator team can share information about the task and environment more naturally, using appropriate levels of abstraction that mitigate the perceptual and cognitive differences between team members.

Semantic SLAM

Semantic mapping, or "SemanticSLAM", is a proposed extension to current state of the art in simultaneous and mapping (SLAM) algorithms, which do not deal well with semantic information. In recent years, the robot localization and mapping literature has seen tremendous progress with the development of particle-filter based localization methods (Dellaert *et al.* 1999a; Fox *et al.* 1999; Dellaert *et al.* 1999b), and FastSLAM (Montemerlo *et al.* 2002; 2003). Monte Carlo localization drastically improved the representational power of localization methods by representing knowledge about the current location of a robot as a set of particles, and updating this set over time. Its power and simplicity made it the most popular localization method in mobile robotics today. FastSLAM is based on the idea that, in addition to keeping track of the robot's trajectory, each particle could additionally store a map of its environment.

Since this map is very easily computed when conditioned on the particle's associated trajectory, the resulting SLAM method is quite fast, hence FastSLAM.

In SemanticSLAM we propose to extend the FastSLAM approach by adding semantic information about the environment to each particle's map. Intuitively, instead of keeping a large occupancy map of the environment, we will now indicate the number and type of a set of parameterized environment elements. These could be spatial in nature, such as "there are 10 rooms and 8 corridors with these shapes", or yield information about interesting objects in the environment, i.e. "there is a cart and two pallets with parts in area 7". As with FastSLAM, the exact dimensions and locations of these elements can be calculated efficiently when conditioned on the particle trajectory, and each particle can be given a score as to how well the proposed map matches the current sensor readings. In terms of algorithm development, the main research challenges to be addressed are those of developing an appropriate representation, and computationally tractable algorithms.

Interacting through the Semantic Map

As the robot identifies environmental features and locations, these could appear automatically as icons within the map. How to present a probabilistic understanding of the environment at the semantic level to a user is by no means evident, however. The use of relative transparency is an obvious candidate for displaying two competing interpretations of raw data onto the map. The user should be able to click on certain entities and query the "state of mind" of the robot, which responds by presenting alternative interpretations of the environment and their relative probabilities, e.g. "with 80% probability this is a cart with dimensions X, but there is a 20% chance that this is a pallet with dimensions Y".

The human should be given the capability to identify and

classify targets in the environment using an interactive visual display, which could be used as additional constraints by a semantic mapping approach. If the operator has more information available than the robot, the user could add, verify, remove, or annotate abstractions within the map. Collaborative construction of the map would enhance each individual team member's understanding of the environment and provides a shared cognitive workspace for the human-robot team to communicate about the environment. The robot could use the map to communicate about the task and environment both graphically (e.g., "The highlighted area has been searched!") and verbally ("Rock 2 found near 'Obstacle 3'") using the semantic names designated within the shared cognitive workspace. The human team members could use the map abstractions for tasking either graphically (e.g., "Search this region for temperatures above 80 0F!") or verbally ("Go to Rock 5!").

Supporting Planning and Decisions

In addition to communication, the probabilistic semantic information could also be used in planning and decision-making. Expanded planning capabilities based on user preferences and semantic maps could support semi-autonomous navigation that alleviates the problems arising from communication delays. The capability of representing a probability distribution over semantic interpretations of the environment opens up novel ways to plan and communicate about plans. The user could provide input about risk attitude, e.g. "avoid this outcome at all cost", "be conservative", or "be optimistic", and high level goals that can be communicated at a semantic level to the robot, e.g. "go to the red room". Plans can be constructed on the expected utility of a certain path or behavior, where now the utility can depend on semantic properties of the environment. As a concrete example, if a certain path is chosen that with 10% probability has a blockage (closed door, large obstacle), the expected time to complete a task ("get out of the building") could be unacceptably long. The proposed work will provide both the means to communicate about risk and plan goals, based on the semantic map representation afforded by Semantic-SLAM, as well as adapt existing utility-based planning algorithms to deal with such (probabilistic) semantic maps.

Conclusion

By marrying probabilistic SLAM methods and object recognition, "semantic SLAM" will bring map building for robots to the next level, enabling meaningful interactions between humans and robots, and supporting planning and decision-making at the mission-level. These capabilities are essential as we approach an era in which robots will play an increasingly important role in both military and non-military applications of the future.

References

Dellaert, F.; Fox, D.; Burgard, W.; and Thrun, S. 1999a. Monte Carlo Localization for mobile robots. In *IEEE Intl. Conf. on Robotics and Automation (ICRA)*.

Dellaert, F.; Fox, D.; Burgard, W.; and Thrun, S. 1999b. Using the condensation algorithm for robust, vision-based mobile robot localization. In *IEEE Conf. on Computer Vision and Pattern Recognition (CVPR)*.

Fox, D.; Burgard, W.; Dellaert, F.; and Thrun, S. 1999. Monte Carlo Localization – Efficient position estimation for mobile robots. In *AAAI Nat. Conf. on Artificial Intelligence*.

Montemerlo, M.; Thrun, S.; Koller, D.; and Wegbreit, B. 2002. FastSLAM: A factored solution to the simultaneous localization and mapping problem. In *AAAI Nat. Conf. on Artificial Intelligence*.

Montemerlo, M.; Thrun, S.; Koller, D.; and Wegbreit, B. 2003. FastSLAM 2.0: An improved particle filtering algorithm for simultaneous localization and mapping that provably converges. In *Intl. Joint Conf. on Artificial Intelligence (IJCAI)*.