Verb Semantics for Robot Dialog

Brian J Thomas and Odest Chadwicke Jenkins
Department of Computer Science
Brown University
Providence, Rhode Island 02912
Email: {brian.cjenkins}@cs.brown.edu

Abstract—Advancements in robotics have led to an ever-growing repertoire of software capabilities (e.g., recognition, mapping, and object manipulation). However, robotic capabilities grow, the complexity of operating and interacting with such robots increases (such as through speech, gesture, scripting, or programming). Language-based communication can offer users the ability to work with physically and computationally complex robots without diminishing the robot’s inherent capability. However, it remains an open question how to build a common ground in human language that will scale with the growth of robot capabilities, for instance within development environments such as ROS (the Robot Operating System). We examine this scaling problem through large-scale symbol grounding for robot dialog. We explore this problem in two parts through the development of: (1) a generic software framework for ROS that grounds parts of speech (verbs for now) into robotic capabilities using the proposed action hierarchy model and (2) a dialog interface for human-robot interaction through an expressive subset of natural language. We will evaluate the framework and interface through mobile manipulation experiments with a PR2, with consideration of the future scalability of our approach.

I. BACKGROUND AND RELATED WORK

Winograd [8] developed SHRDLU, a system which processed natural-language instructions and performed actions in a virtual environment. From this, researchers pushed forward trying to extend SHRDLU’s capabilities into real-world environments and soon branched into tackling various sub-problems, including NLP and robotics systems. Research conducted on the robotics systems side has resulted in frameworks such as ROS, developed by Quigley et al. [5], which has been used in several domains of modern robotics research. NLP research including robotic components has also lead to advancements. Notably, Kollar et al. [3] and MacMahon et al. [4] have developed methods of following route instructions given in natural language. Dzifcak et al. [2] studied translating natural language instructions into goal descriptions and actions. Chernova et al. [1] implemented natural-language and action-oriented human-robot interaction with humans in a task by data-mining previous human-human interactions of the same task. However, the scalability of these solutions outside of their test domain remains open. Attempts have been made recombining these fields. For instance, Tenorth et al. [2] has developed robotic systems capable of inferring and acting upon implicit commands using knowledge databases. Finally, linguists have also examined problems related to understanding verbs; for instance, Ruppenhofer et al. [6]’s FrameNet portrays verbs as a key role of a “scene” or semantic frame.

II. CONTRIBUTIONS AND IMPLEMENTATION

We approach the problem using an action hierarchy model, which binds verbs in input dialog to actions in ROS. Dialog is defined an expressive subset of natural language and is a starting point for more sophisticated language processing. Here, we use dialog to establish communication patterns that both robots and humans can understand, establishing a common ground between humans and robots. As NLP continues to improve, common ground can be established in ways even more similar to natural language than dialog. Further, grounding can be established not only for verbs, but for other parts of speech such as nouns, prepositional phrases, and the like.

We choose ROS to implement this approach because of its community support for a multitude of platforms and localization packages. Further, code developed for ROS is already modularized into constructs called nodes, which interact with each other via inter-process communication. The combination of community and modularization provides us with a plethora of code chunks that verbs can ground themselves in. Because of ROS’s community’s size and heterogeneity, we need to also consider how the hierarchy will be stored, who will have permission to modify it, and how. Centralized control (in the form of, e.g., a “blessed” verb grounding package) offers better reliability and control of user experience than community control but at the cost of the community being able to contribute. To get the best of both worlds, we will develop both a “blessed” verb grounding package and a patch system, the former creating a positive user experience and the latter enabling the community to link their work to ours without requiring they go through a formal appeals process.

III. ACTION HIERARCHY MODEL

A. Concept

The action hierarchy model is a four-level model which binds verbs in dialog to actions in ROS. Namely, the four levels (from highest to lowest) are instructions, verbs, action classes, and action instances. Instructions are dialog language input and can be thought of as a state machine that the robot should execute, with each verb in the dialog representing a single state. Instructions can turn into verbs by creating this state machine to structure the dialog-implicit transitions between verbs. Each verb is associated with a single action class, so verbs turn into action classes by simply referring
An example of grounding verbs in dialog into ROS actions using the action hierarchy model.

Fig. 1. An example of grounding verbs in dialog into ROS actions using the action hierarchy model.

```
PathPlan - Navigate - Move
or
Wander
```

...
will continue in an attempt to find another match. It may also be possible in the future to learn the best combinations and attempt them first.

IV. EVALUATION

We will demonstrate the framework’s capability of scaling to all possible ROS functionality by grounding a multitude of verbs using our action hierarchy model. To demonstrate this grounding, experiments will be performed in areas such as structure assembly, object retrieval, and place setting. We will use the PR2 platform to accomplish these tasks. Success will be measured by the scope of verbs and instructions grounded in our model and the variety of tasks successfully performed by our model.

V. CONCLUSION

Communication with robots through dialog allows users to interact with complex robots without diminishing either the necessary expressiveness of the commands or the abilities of the robot. The proposed action hierarchy model will facilitate this communication by enabling verbs to be grounded into concrete robot actions. Due to its domain agnosticism, this model can scale to ROS’s capabilities. ROS offers community code support and provides a framework upon which to implement this model. We plan to develop our framework and evaluate its abilities through mobile manipulation experiments using a PR2.

This work was supported in part by NSF CAREER Award IIS-0844486 and by ONR PECASE Award N000140810910.

REFERENCES


