Situated Natural Language Interaction in Uncertain and Open Worlds

Tom Williams (Tufts University; williams@cs.tufts.edu)
DOI: 10.1145/3098888.3098896

As intelligent robots become integrated into society, it becomes important for them to be capable of natural, human-like human-robot interaction (HRI). While there has been some progress on enabling natural-language based HRI (Mavridis, 2015), most natural language enabled robots rely on highly scripted interactions, keyword spotting, and shallow natural language processing techniques. For many applications, these methods may be sufficient to achieve the desired behavior, which may be restricted to a small class of tasks. Such methods, however, are not helpful for the development of robots that are generally and flexibly taskable, that can learn about new entities and concepts on the fly, and that are capable of engaging in truly natural *human-like* HRI.

What is more, even natural-language enabled robots designed to handle more natural, flexible dialogue typically operate under a set of assumptions that severely restrict the types of language they are prepared to handle. Specifically, many language-enabled robots assume that (1) their knowledge is certain, (2) they operate in a closed world, (3) only entities from a single domain will be referred to, (4) knowledge is centralized, and homogeneous in representation, (5) humans' utterances should be understood as commands or requests, (6) humans' utterances will be expressed directly, and/or that (7) the meaning of humans' utterances will not vary with context.

To advance the state of the art of natural language based HRI, we must develop natural language enabled robots that challenge these assumptions, that is, robots which are able to (1) handle uncertain and open worlds; (2) make use of distributed knowledge that is heterogeneous in domain and in representation; (3) process a wide variety of utterance forms and referring expression forms; and (4) process such utterances in a context sensitive manner.

In this dissertation, I describe algorithms I have developed in service of these goals, and the experimental and theoretical work I have performed which informs those algorithms and mechanisms.

Copyright © 2017 by the author(s).

Figure 1: The *Vulcan* Intelligent Wheelchair: one of the robot platforms used in the presented work.

I first present a set of algorithms that provide **reference resolution** and **referring expression generation** capabilities: *SPEX*, the *Spatial Expert*, an architectural component responsible for performing spatial reference resolution in open worlds (Williams, Cantrell, Briggs, Schermerhorn, & Scheutz, 2013); *REX*, the *Referential Executive*, an architectural component responsible for a broader class of referential activities, including domain-independent reference resolution of definite noun phrases in uncertain and open worlds (Williams & Scheutz, 2015a,b, 2016a); *GH-POWER*, an algorithm which incorporates *REX* into a broader *Givenness Hierarchy*-theoretic (Gundel, Hedberg, & Zacharski, 1993) framework in order to additionally resolve anaphoric and deictic expressions in a context sensitive manner (Williams, Acharya, Schreitter, & Scheutz, 2016; Williams & Scheutz, 2017); and *PIA*, an algorithm which uses *REX* for the purposes of referring expression *generation*.

Next, I move on to discuss **pragmatic reasoning**. I begin by presenting experimental evidence demonstrating the extent of *indirect speech act* use in HRI (Briggs, Williams, & Scheutz, 2017), and then present a Dempster-Shafer theoretic framework for both understanding and generating indirect speech acts in a context sensitive manner under uncertainty and ignorance (Williams, Briggs, Oosterveld, & Scheutz, 2015). Next, I demonstrate how this framework can be used to generate clarification requests to resolve pragmatic and referential ambiguity (Williams & Scheutz, 2016b). Finally I move beyond the pragmatics of *human*-robot communication, and discuss the pragmatics of *robot*-robot communication (Williams, Briggs, & Scheutz, 2015).

Finally, I discuss the application of the presented algorithms to assistive robotics, by providing a comprehensive survey of natural language enabled wheelchairs, and then demonstrating how the use of the presented algorithms on the University of Michigan's Vulcan intelligent wheelchair (Figure 1) (Murarka, Gulati, Beeson, & Kuipers, 2009) advances the state of the art of such wheelchairs (Williams, Johnson, Scheutz, & Kuipers, 2017).

Acknowledgments

This doctoral work was funded in part by ONR grants #N00014-11-1-0289, #N00014-11-1-0493, #N00014-10-1-0140 #N00014-14-1-0144, #N00014-14-1-0149, #N00014-14-1-0751, and NSF grants #1111323, #1038257.

References

Briggs, G., Williams, T., & Scheutz, M. (2017). Enabling robots to understand indirect speech acts in task-based interactions. *Journal of HRI*.

Gundel, J. K., Hedberg, N., & Zacharski, R. (1993). Cognitive status and the form of referring expressions in discourse. *Language*, 274–307.

Mavridis, N. (2015). A review of verbal and non-verbal human–robot interactive communication. *Robotics and Autonomous Systems*, *63*, 22–35.

Murarka, A., Gulati, S., Beeson, P., & Kuipers, B. (2009). Towards a safe, low-cost, intelligent wheelchair. In *PPNIV.*

Williams, T., Acharya, S., Schreitter, S., & Scheutz, M. (2016). Situated open world reference resolution for human-robot dialogue. In *ACM/IEEE HRI.*

Williams, T., Briggs, G., Oosterveld, B., & Scheutz, M. (2015). Going beyond command-based instructions: Extending robotic natural language interaction capabilities. In *AAAI.*

Williams, T., Briggs, P., & Scheutz, M. (2015). Covert robot-robot communication: Human perceptions and implications for human-robot interaction. *Journal of HRI*.

Williams, T., Cantrell, R., Briggs, G., Schermerhorn, P., & Scheutz, M. (2013). Grounding natural language references to unvisited and hypothetical locations. In *AAAI.*

Williams, T., Johnson, C., Scheutz, M., & Kuipers, B. (2017). A tale of two architectures: A dual-citizenship integration of natural language and the cognitive map. In *AAMAS.*

Williams, T., & Scheutz, M. (2015a). A domain-independent model of open-world reference resolution. In *COGSCI.*

Williams, T., & Scheutz, M. (2015b). POWER: A domain-independent algorithm for probabilistic, open-world entity resolution. In *IEEE/RSJ IROS.*

Williams, T., & Scheutz, M. (2016a). A framework for resolving open-world referential expressions in distributed heterogeneous knowledge bases. In *AAAI.*

Williams, T., & Scheutz, M. (2016b). Resolution of referential ambiguity using Dempster-Shafer theoretic pragmatics. In *AI-HRI.*

Williams, T., & Scheutz, M. (2017). Reference resolution in robotics: A givenness hierarchy theoretic approach. In J. Gundel & B. Abbott (Eds.), *The Oxford handbook of reference.*

Tom Williams completed a joint PhD in Computer Science and Cognitive Science in May 2017 with Professor Matthias Scheutz at Tufts University, and will be joining Colorado School of Mines in August 2017 as an Assistant Professor of Computer Science. Tom's research focuses on allowing robots to communicate in natural language in uncertain and open worlds, with applications to assistive and search-and-rescue robotics.