Toward More Natural Human-Robot Dialogue

Tom Williams
Human-Robot Interaction Laboratory, Tufts University
Medford, MA, USA
williams@cs.tufts.edu

Categories and Subject Descriptors

I.2.7 [**Artificial Intelligence**]: Natural Language Processing

Keywords

Speech Act Theory; Intention Understanding; Dempster-Shafer Theory

1. INTRODUCTION

A primary goal of the field of Human-Robot Interaction is to allow for *natural* human-robot interactions, and thus robot architectures must eventually be able to understand truly natural human speech. And yet, despite the abundance of research devoted to language understanding, most robots capable of participating in linguistic interactions are only able to understand relatively simple utterances (e.g., commands), and do not consider those utterances' deeper implications.

We believe that this suggests a shortcoming of the current state of the art: humans do not typically restrict themselves to commands, and humans' intentions are often not derivable from the semantic content of the utterances they employ. Indeed, most human language is intentionally indirect and ambiguous so as to conform with social conventions (e.g., politeness). If we desire truly natural human-robot interactions, we must thus go beyond the command-based paradigm characterizing most current robot architectures.

While a few architectures have made first steps toward a deeper understanding of human utterances, these have not attempted to represent a robot's *certainty* in its beliefs or perceptions. As human utterances are rife with both intentional and incidental ambiguity, we believe such systems are ill-equipped for use in the real world.

Our research seeks to address the shortcomings of current architectures by developing mechanisms for natural language understanding and generation. These mechanisms use the robot's goal-based, social, and environmental knowledge to deeply understand human utterances and to generate utterances that adhere to the social conventions of human interactions, exploiting the robot's knowledge of its own ignorance to achieve robustness to uncertainty and to appropriately generate clarification requests.

HRI'15 Extended Abstracts, March 2–5, 2015, Portland, OR, USA.
ACM 978-1-4503-3318-4/15/03.
http://dx.doi.org/10.1145/2701973.2702703 .

2. PRIOR WORK

Our work follows in the tradition of Searle's speech act theory [6], and extends the capabilities of several recent language-capable robot architectures [2, 1, 5, 4]. Of these architectures, few enable robots to reason about the certainty of their own beliefs, and those that do [2, 1] are unable to infer the intentions behind non-literal utterances such as *indirect speech acts (ISAs)*. On the other hand, some architectures allow robots to interpret a variety of ISAs [5, 3], but do not explicitly represent the uncertainty of the robot's beliefs and thus are not robust to uncertain context. To the best of our knowledge, only one approach has both represented uncertainty and afforded the ability to interpret non-literal utterances [9], but this approach is limited to understanding indirect *commands* and uses a rudimentary representation of uncertainty that is only applied to the rules used by the robot, and not to the robot's knowledge itself.

3. METHODOLOGY

Our proposed robot architecture is best characterized by its use of Dempster-Shafer (DS) theoretic representations of uncertainty, and its symmetric pragmatic reasoning components. I will now briefly discuss each of these features.

3.1 A Dempster-Shafer Theoretic Approach

The Dempster-Shafer theory of evidence is a generalization of the standard Bayesian probability system which assigns degrees of belief to sets of mutually exclusive hypotheses rather than specifying probability distributions to characterize individual events. The uncertainty associated with a particular hypotheses can then be represented using a two-valued probability interval calculated using the degrees of belief of hypotheses that support and refute that hypothesis. The use of a DS-theoretic approach provides an elegant way to represent and reason about the uncertainty and ignorance of a robot's beliefs, without committing to a particular probability distribution which may or may not be justified.

3.2 Pragmatic Reasoning Components

Our architecture introduces a pair of pragmatic reasoning components: a pragmatic inference component which uses a table of pragmatic rules and knowledge of the current context to deduce a set of intentions from a given utterance, and a pragmatic generation component which uses the same rules and context to abduce the best utterance to communicate a given intention.

Each rule representation contains an utterance form, a set of contextual items which must be true for the rule to be applicable, a set of intentions to infer from the rule if it is applied, and a DS-theoretic uncertainty interval representing the robot's confidence in the rule itself. Representing this uncertainty explicitly makes the robot more robust to uncertain contextual factors, and allows us to write rules that are more general and versatile than the pragmatic rules used in other proposed architectures (e.g., [5]). We will now briefly describe the inference algorithms utilized by our pragmatic inference and generation components.

3.2.1 Pragmatic Inference

The Pragmatic Inference component takes as input the semantic content of an utterance and the DS-theoretic uncertainty interval reflecting the robot's confidence in those semantics. It then considers each pragmatic rule applicable for the given utterance in the current context (as determined by the items residing in the robot's knowledge base, each of which has an associated interval reflecting the robot's confidence in the truth of that item). For each rule, the algorithm combines its representations of the utterance and contextual item associated with that rule, by way of a DS-theoretic AND operation. A DS-theoretic Modus Ponens operation is then used to obtain a representation of the intention associated with the rule. The uncertainty interval associated with this intention captures the robot's degrees of uncertainty and ignorance in its belief that that intention was intended by the interlocutor. DS-theoretic fusion operators are then used to combine the results of all rule applications into a set of inferred intentions. This algorithm is detailed in [8].

3.2.2 Pragmatic Generation

When the robot must communicate its intentions, it must choose an appropriate surface realization of those intentions. This is accomplished using an abductive algorithm which uses the same set of pragmatic rules used for inference. This algorithm chooses as the "best" utterance the utterance most likely to appropriately communicate the robot's intentions, without also communicating anything the robot does not actually believe. The DS-theoretic approach becomes particularly useful here; since the rules used for inference are essentially equations that relate premise and rule to consequent, they can be used for both deductively and abductively. This algorithm is detailed in [7].

3.3 Integration

These components were integrated into the DIARC architecture and implemented on a Willow Garage PR2. This integration demonstrated several benefits of our approach: by explicitly representing the uncertainty of the robot's knowledge and perceptions, the robot was able to use DS-theoretic uncertainty assessment mechanisms to ask for clarification when high uncertainty or ignorance was reflected in the semantics of recognized speech or in the intentions produced by pragmatic inference. This integration is detailed in [7], and can be observed in a video at https://vimeo.com/106203678. While an evaluation will eventually be critical, we believe it would be premature, as it is not yet clear how best to evaluate such an integrated system. For instance, it is unclear how many scenarios and rules must be examined, and how to be sure that such data is sufficient. Instead, our approach should be viewed as a real-time, integrated, proof-of-concept demonstration.

4. FUTURE WORK

This work will be extended in several ways. First, we seek to use DS-theoretic adaptation mechanisms to automatically learn new rules and to adapt the uncertainties of learned rules. We also plan to extend our approach to maximally leverage the set-reasoning capabilities afforded by the DS-theoretic approach. Finally, research is needed to determine how best to evaluate dialogue systems in the context of integrated robot architectures.

5. ACKNOWLEDGMENTS

This research was funded in part by ONR grants #N00014-11-1-0493, #N00014-14-1-0751, and #N00014-14-1-0144.

6. REFERENCES

[1] J. Y. Chai, L. She, R. Fang, S. Ottarson, C. Littley, C. Liu, and K. Hanson. Collaborative effort towards common ground in situated human-robot dialogue. In *HRI*, pages 33–40, 2014.

[2] R. Deits, S. Tellex, T. Kollar, and N. Roy. Clarifying commands with information-theoretic human-robot dialog. *JHRI*, 2013.

[3] E. A. Hinkelman and J. F. Allen. Two constraints on speech act ambiguity. In *Proceedings of the 27th annual meeting on Association for Computational Linguistics*, pages 212–219, 1989.

[4] G.-J. M. Kruijff, P. Lison, T. Benjamin, H. Jacobsson, H. Zender, I. Kruijff-Korbayová, and N. Hawes. Situated dialogue processing for human-robot interaction. In *Cognitive Systems*, pages 311–364. 2010.

[5] M. Scheutz, G. Briggs, R. Cantrell, E. Krause, T. Williams, and R. Veale. Novel mechanisms for natural human-robot interactions in the diarc architecture. In *Proceedings of AAAI Workshop on Intelligent Robotic Systems*, 2013.

[6] J. R. Searle. Indirect speech acts. *Syntax and semantics*, pages 59–82, 1975.

[7] T. Williams, G. Briggs, B. Oosterveld, and M. Scheutz. Going beyond command-based instructions: Extending robotic natural language interaction capabilities. In *Proceedings of Twenty-Ninth AAAI Conference on Artificial Intelligence*, 2015.

[8] T. Williams, R. C. Núñez, G. Briggs, M. Scheutz, K. Premaratne, and M. N. Murthi. A dempster-shafer theoretic approach to understanding indirect speech acts. *Advances in Artificial Intelligence*, 2014.

[9] S. Wilske and G.-J. Kruijff. Service robots dealing with indirect speech acts. In *IROS*, pages 4698–4703. IEEE, 2006.