Going Beyond Literal Command-Based Instructions:
Extending Robotic Natural Language Interaction Capabilities

Tom Williams Gordon Briggs Bradley Oosterveld Matthias Scheutz
Human-Robot Interaction Laboratory, Tufts University, 200 Boston Avenue, Medford, MA, USA

The Problem

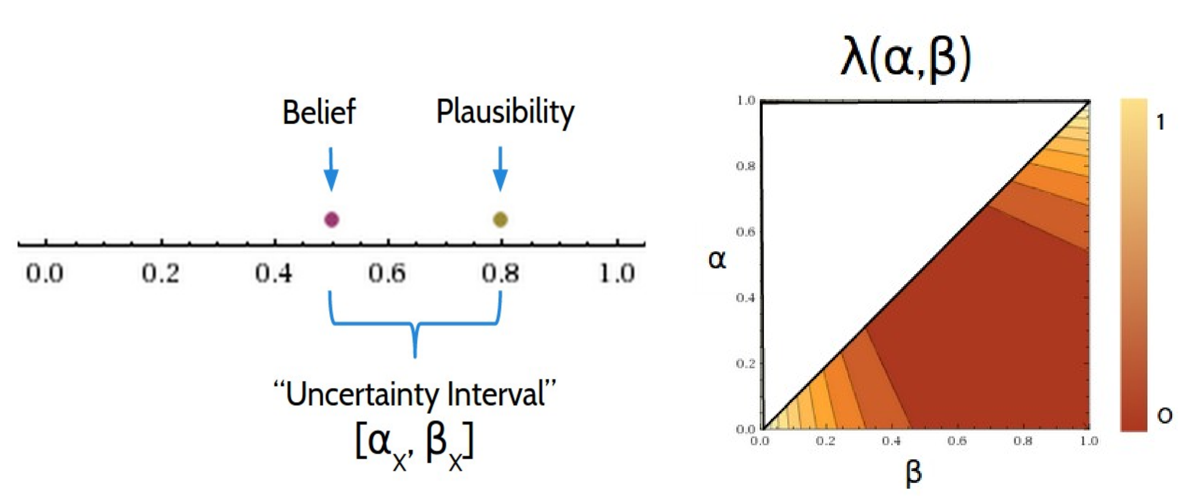
Robots should be able to understand (and generate) utterances whose meanings are not directly derivable from their semantics (as in the exchange to the right).

Robots must be able to understand these types of utterances because they comprise the majority of natural human dialogue for *social* reasons (e.g., politeness). Robots must be able to generate these types of utterances in order to be perceived as following those same social conventions.

Furthermore, this *pragmatic* understanding and generation must be possible in the face of *uncertainty*.

Our Approach

Our approach makes use of a set of **pragmatic rules** for both understanding and generation. These rules are represented using the following form:
(Utterance) U ^ (Context) C => (Intention) I
The **Dempster-Shafer Theory of Evidence is used** to represent and reason about the robot's uncertainty: the certainty of each Utterance, Context, Intention and Rule is represented by its associated **belief** (Bel(x)) and **plausibility** (Pl(x)) measures.



Left: A DS-Theoretic Uncertainty Interval
Right: Depiction of Nunez' Certainty Measure.

0 Initial context

The robot starts with some built in knowledge.
Example:
locationof(breakroom,medkit) [0.8,0.9]
bel(Jim,subordinate(self,Jim)) [0.5,0.6]
bel(Jim,subordinate(Jim,self)) [0.4,0.5]

1 Recognition and Parsing

When the robot hears a sentence, it is first **recognized** and **parsed.** If the robot's confidence in its recognition or parsing is too low (reflected through **Nunez' Certainty Measure (λ)**), the robot will ask for clarification. Otherwise the results are passed to **Pragmatic Inference (*PINF*).**

Statement(Jim,self, need(commander_z,medkit)) [0.95,1.0]
λ(0.95, 1.0) > 0.1, so semantics passed to PINF.

2 Rule Selection

PINF finds any rules that are applicable under the current utterance and context.

R1: If Bel(A,subordinate(B,A)): Stmt(A,B,needs(C,D))=>
Goal(B,bring(B,D,C))[0.8,0.9]]
not(ITK(A,locationof(E,D)))[0.8,0.9]
R2: If Bel(A,subordinate(A,B)): Stmt(A,B,needs(C,D))=>
not(Goal(B,bring(B,D,C)))[0.8,1.0]
ITK(A,locationof(E,D))[0.8,1.0]

3 Pragmatic Inference

Possible **Intentions** are **induced** by first applying **uncertain logical** AND and Modus Ponens and then fusing intentions that have the same semantic form, using Yager's Rule of Combination. The results are passed to the **Dialogue, Belief and Goal Manager (DBGM)**.

I1: Goal(slef,bring(self,medkit,commander_z))[0.47,0.67]
I2: ITK(Jim,locationof(X,medkit))[0.38,0.5]

4 Clarification Check

Nunez' uncertainty measure is used to check the produced intentions. If they are deemed too uncertain, a clarification request is generated. Otherwise, the Intentions are asserted into memory.

λ(0.47, 0.67) and λ(0.47, 0.67) both < 0.1,
so clarification request passed to PGEN with semantics:
ITK(self, or(ITK(Jim, locationof (X, medkit)),
Goal(self, bring(self, medkit, commander_z))))[1.0, 1.0].

5 Rule Selection

When the robot needs to communicate an intention of its own, **Pragmatic Generation (PGEN)** finds rules applicable under the current context and intention.

R1: AskWH(A,B,or(C',D'))=>ITK(A,or(C',D'))[0.95,0.95]
R2: Stmt(A,B,Want(A,Know(A,C)))=>ITK(A,C)[0.85,0.85]

6 Pragmatic Generation

PGEN then recursively applies **uncertain logical** AND and Modus Ponens to determine the degree to which various candidate utterance forms would communicate the desired intention. PINF is then used to check for any unwanted side effects of the resulting candidate utterances. The "best" utterance is then passed to NLG.

ITK(self, or(W ant(Jim, Know(Jim, locationof (X, medkit))),
Want(Y, bring(self, medkit, commander_z))))[0.95, 1.0].

7 Translation and Synthesis

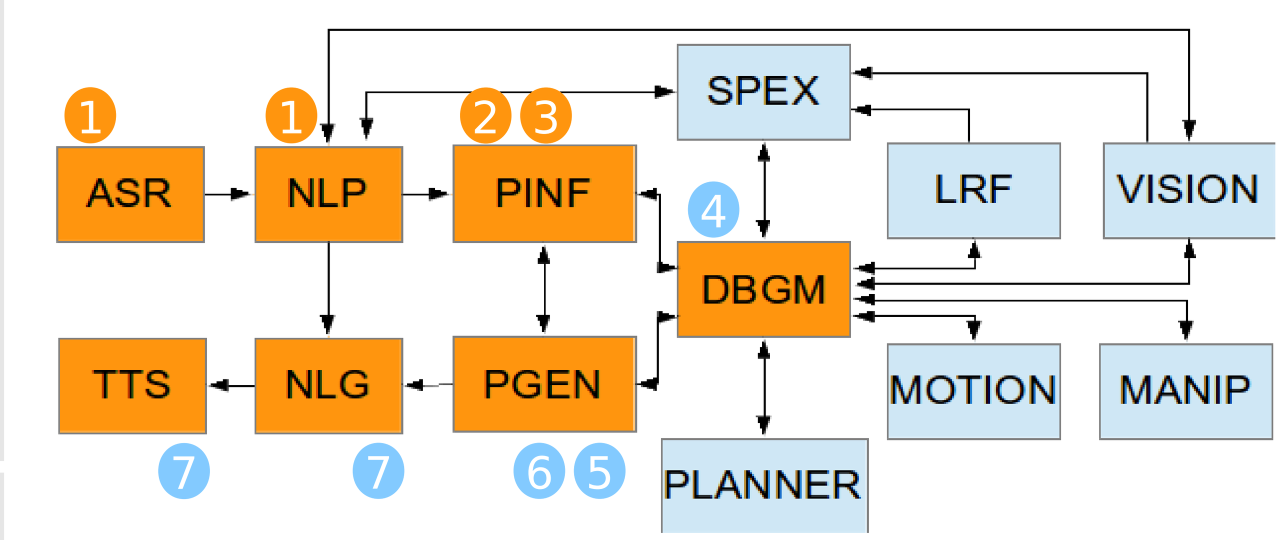
The chosen utterance is then translated by NLG and synthesized by TTS.

"Would you like to know where to find a medkit?
Or would you like me to bring commander Z a medkit?"

Integration and Conclusions

The capabilities presented here were fully integrated into the DIARC architecture, and the dialogue to the left was performed on a Willow Garage PR2. A video of this interaction can be viewed online at:
https://vimeo.com/106203678

These new architectural capabilities represent an advance in the state of the art of language-capable robot architectures, as the ability to understand human utterances with context-dependent implications brings robots closer to being able to engage in truly natural interactions with their human teammates.



Partial architecture diagram. Highlighted components form the natural language pipeline.

References

Tang, Y.; Hang, C.-W.; Parsons, S.; and Singh, M. P. 2012. Towards argumentation with symbolic dempster-shafer evidence. In COMMA, 462–469.
Williams, T.; Núñez, R. C.; Briggs, G.; Scheutz, M.; Premaratne, K.; and Murthi, M. N. 2014. A dempster-shafer theoretic approach to understanding indirect speech acts. Advances in Artificial Intelligence.
Scheutz, M.; Briggs, G.; Cantrell, R.; Krause, E.; Williams, T.; and Veale, R. 2013. Novel mechanisms for natural human-robot interactions in the DIARC architecture. In Proceedings of AAAI Workshop on Intelligent Robotic Systems.

Acknowledgments

This work was funded in part by ONR grants #N00014-11-1-0493 and #N00014-10-1-0140, and in part by NSF grants #1111323 and #1038257.