Human Robot Interaction through TCL

John R. Lee
University of Iowa
Dept of Electrical and Computer Engineering
4016 Seamans Center
Iowa City, Iowa 52242-1595
AIResearcher@gmail.com

Andrew B. Williams
Spelman College
Dept of Computer and Information Science
350 Spelman Lane SW
Atlanta, GA 30314
Williams@Spelman.edu

ABSTRACT
Human language and communication technologies play a key role in enabling humans to interact with robots in a natural and intuitive way. This is essential for the usability and adoption of robots into every day life. Task Communication Language attempts to be a unifying architecture for the representation and incorporation of communication within the behavior of an intelligent agent or robot. This paper provides a general overview of TCL as well as techniques and methodologies currently under development.

Categories and Subject Descriptors
I.2.7 [**Artificial Intelligence**]: Natural Language Processing – *Discourse, Language models.*

Keywords
Human Robot Interaction, Natural Language Understanding, Task Communication Language.

1. INTRODUCTION AND MOTIVATION
Current research in robotics has been dealing with very specific robotics issues, such as navigation, grasping, moving obstacle detection and avoidance, object and facial recognition, even multiple robot or human/robot object balancing. There has been a great push for a humanoid service robot capable of interacting directly with humans in daily life. If these robots are expected to interact with humans in activities of every day living, much greater communication and interaction technologies are needed.

Human language and communication technologies play a key role in enabling humans to interact with robots on a natural and intuitive level. Although many research groups have great ambitions for human robot communication, there is a lack of complex language skills in multi-modal implementations. Currently, most assistive robotics systems recognize specific keywords in incoming speech that then lead to the execution of predefined actions and procedures. Some allow the imposing of constraints or beliefs or planning and others allow multimodal interaction to allow gestures and pointing to accompany commands.

Although there has been a great deal of dialogue models which allow for action and plan recognition, procedure learning, intention recognition, collaborative planning, mixed initiative behavior, negotiation and more; these models have not been successfully integrated into working human-centered applications, other than a few domain-specific research toys or telecommunication based systems. This is somewhat due to the lack of implementation platforms as well as current deficiencies of speech-act and dialogue-act recognition.

Task communication language (TCL) attempts to be a unifying architecture for the representation and incorporation of communication within the behavior of an intelligent agent. This produces a foundation on which many of the communication models in the preceding paragraph can be easily modeled, expanded and integrated together into one implementation.

2. TASK COMMUNICATION LANGUAGE
TCL is based on the *practical communication language hypothesis* [4], as well as the *practical dialogue hypothesis* and the *domain-independence hypothesis* [1]. The former is provided below.

> *The Practical Communication Language Hypothesis*: There exists a language between that of a human conversational participant and that of an intelligent agent. This language is capable of abstracting away the complexity of human language while yet maintaining the practical information of the conversation.

Many natural language understanding systems have some form of semantics between a human and its programming, whether it is a language, protocol, or programming interface. However, TCL is deeply founded in this hypothesis, which provides a common foundation upon which dialogue systems can be unified.

The true practical communication language is ideal and volatile. This is due to the ever-evolving notion of what is 'practical'. It should be abstracted of all region and dialect aspects of language, informal, colloquial, slang, idiomatic expressions and even modality. However, this information should not simply be discarded. For example, [3] demonstrates that modality can influence the establishment of common ground. These nuances of interaction should be understood through the translation rather than relying on the existence of corresponding concepts.

Human Robot Interaction'06, March 2-3, 2006, Salt Lake City, Utah USA.

2.1 Approach

Only a general overview of TCL is provided. Interested readers are referred to [4] for more details.

TCL consists of a series of messages. Each message is wrapped into a well-structured outer layer that allows for the modeling of non-traditional conversational paradigms as well as learning through feedback of any interpretation information collected. The content of each message is composed of nested meaning-action concepts (MAC). A MAC is a particular meaning of an utterance or gesture, which may or may not have an associated action. They are based on the task-model below.

2.2 Methodology

Associated with the language of TCL are techniques and methodologies of modeling the interaction between communication and behavior and integrating them into an intelligent agent implementation. The details of this section represent work in progress.

2.2.1 Dialogue Models

Various dialogue models obtained from examples of natural human interaction are built. These models represent higher-order communicative abilities such as negotiation, team formation, persuasion, collaborative problem solving, command and control, learning by description and adjusting autonomy.

2.2.2 Task Modeling

The task model consists of a set of core atomic concepts as well as a set of operators that act upon those concepts. Although generalized to most conceptual models, the original inspiration came from the problem solving concepts of [2]. The subset of core concepts includes objectives, recipes, actions, resources, situations, states, constraints, beliefs, intentions, metrics and priorities. The set of task operators include adoption, selection, deferment, abandonment, release, identification, evaluation and modification.

Although rationalists will demand that each concept is explicitly stated with every possible consequence in order to integrate these concepts within a mathematical logic, the irrational properties of human language and behavior force the model to take an empirical approach.

2.2.3 Task-Communication Model

The task-communication model interleaves the concepts and operators of the task model into a communicative and behavioral model of the agent. For example, take the core concept 'action' and the communicative acts *propose(action), accept(action)* and *reject(action)* as well as the behavioral act *evaluate(action)*. A simple task-communication model would then describe the following. The input communicative act of the proposal of an action leads to the evaluation of that action. If the evaluation is good, then the action is accepted, if not it is rejected. Complications can then be introduced such as the misunderstanding of an action, which leads to a clarification dialogue rather than an immediate evaluation or acceptance.

Hierarchies in the task-communication model allow for the layering of communicative behavior. An example may include persuasion using the lower elements of formal and informal argumentation.

2.2.4 Interaction Modeling

The interaction model is built from the task-communication model by dropping behavioral acts and keeping communicative acts. The purpose of the interaction model is to incorporate and evaluate the composition of many task-communication models.

2.3 Implementation

The task concepts and operators should be implemented directly within the agent's core facilities such as reasoning, planning and knowledge. The communicative acts correlate to messages of TCL. Thus, the task-communication model demonstrates how communication brings about a change in behavior as well as how changes in beliefs or rationality bring about communication.

2.4 Validation

Annotated corpuses of dialogue sequences can be applied against the interaction model. If the dialogue sequence is covered within the interaction model, then that particular sequence is validated. If the dialogue sequence is not present, then the interaction model should be examined to see if the task-communication model could be expanded to include the cases that the dialogue sequence reveals are missing.

3. DISCUSSION

The dialogue manager must shift roles from being the center of communication for an intelligent agent to one that translates human centric communication. Dialogue moves must now be encapsulated within the framework of a model integrating communication with behavior. The notions of a dialogue stack and context must be shared between the manager and the model. The task-communication model will have to reflect notions of turn-taking and mixed initiative control. Due to length considerations, these types of issues have not been discussed.

We plan to run human investigations with one of our agent-based implementations. This will be done with a wizard-of-oz based technique, employing a human operator to translate the subject's commands, into meaning-action concepts and entering those into the system. The subject will be unaware of this process.

4. REFERENCES

[1] James Allen, Donna Byron, Myroslava Dzikovska, George Ferguson, Lucian Galescu and Amanda Stent. "An architecture for a generic dialogue shell." *Natural Language Engineering*, 6(3-4):213-228, December 2000.

[2] Nate Blaylock, John Dowding and James Allen. "A dialogue model for interaction with planers, schedulers and executives." In *Proceedings of the 3rd International NASA Workshop on Planning and Scheduling for Space*, Houston, Texas, October 27-29, 2002.

[3] Herbert H. Clark and Susan E. Brennen. "Grounding in Communication." In L. B. Reskick, J. M. Levine and S. D. Teasley (Eds.) Perspectives on Socially Shared Cognition (pp.127-149). Washington, DC: American Psychological Association, 1991.

[4] John R. Lee and Andrew B. Williams. Towards a Theoretical Framework for the Integration of Dialogue Models into Human-Agent Interaction. *To Appear AAAI Fall 2005 Symposium on Mixed Initiative Planning*, Washington D.C. November 2005