Towards a Theoretical Framework for the Integration of Dialogue Models into Human-Agent Interaction

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 Sciences
350 Spelman Lane SW
Atlanta, GA 30314
Williams@Spelman.edu

Abstract

This paper describes ongoing research in the development of formal techniques and methods for human agent interaction using a layered protocol based approach and a task communication language (TCL). This approach focuses upon the interaction between a dialogue manager and the behavioral aspects of an intelligent agent. This separation allows parallel development of each as well as interoperability through a flexible, dynamic and adaptive interface. We describe the proposed techniques for creating the semantics for a task communication language and modeling methodology and providing empirical evaluations. We show how this will allow for the merging of a variety of existing dialogue and interaction models into a single integrated system.

Introduction and Motivation

Research in task-oriented dialogue understanding and modeling has produced a plethora of models and ideas on how an intelligent agent should understand, interact with, and incorporate communication from a human conversational participant. However, these models have not been integrated together within an intelligent agent implementation. In fact, some of the models are merely theoretic in nature, based on a few dialogue examples and have not been implemented at all.

A unified architecture for the representation and incorporation of communication within the behavior of an intelligent agent would produce a foundation on which many of the behavioral aspects could be easily modeled and expanded. Furthermore, this architecture should be accompanied by formal techniques for merging into a single integrated model. The goal of our research is to create this unified representation along with the formal techniques to operate upon and prove properties of models built on this representation.

(Allen 2000) presented *The Practical Dialogue Hypothesis* and *The Domain-independence Hypothesis*. Although both of these hypotheses allow for the development of generic dialogue systems, they do not provide a common foundation upon which dialogue systems can be unified. It is essential that this common foundation conceptualize notions of communication and interaction. This leads us to form a third hypothesis.

> *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.

The practical communication language (PCL) hypothesis is built on the idea that human-agent communication and interaction can be modeled as a protocol. (Duke 2000) provides cognitive and psychological justification for viewing the interaction of humans and devices as a hierarchy of protocols. In addition, spoken language interpretation is performed as a layer of protocols.

The true practical communication language itself is ideal and volatile. This is due to the definition of 'practical' and its ability to continually evolve and expand. For example, the PCLs of the past may have been first order logic semantics for command and control, but recent developments in modeling today have a greatly expanded vocabulary of speech-acts and require aspects like prosody for detecting notions such as sarcasm, levels of commitment, or knowledge certainty.

We hold the following beliefs in our pursuit of the ideal PCL.

1) PCL should be abstracted of all region and dialect aspects of language. (e.g. English, German…)
2) PCL should be abstracted of informal, colloquial, slang and idiomatic expressions.
3) PCL should be abstracted of all modality. (e.g. Spoken, written, gestural, specialized-GUI…)

The idea behind PCL is to carry aspects of meaning and attempt to handle the majority of anamorphic resolution and sub-clarification dialogues at lower levels. PCL should be abstracted of things listed above, but this information should not simply be thrown away. For example, (Clark 1990) demonstrates that modality can influence the establishment of common ground.

The rest of the paper is organized as follows. We describe the background to generic dialogue systems and artificial discourse languages. We then describe our approach for task communication semantics and language. Then we describe how we intend to evaluate our approach. Finally, we discuss our current research and future work.

Background

In work on generic dialogue systems (Allen 2000) mentions the communication between a discourse manager and a behavioral agent. (Blaylock 2002 and Allen 2001) expand this by discussing a task-model inside the behavioral (domain) agent containing various problem solving concepts of objectives, recipes, actions, resources and situations and discussing acts such as adoption, selection, deferment, abandonment, release, identification, evaluation and modification, that operate directly upon those problem-solving concepts. The communication between the dialogue manager and the domain agent is referred to as the *interaction act*.

Other work (Sidner 2002) describes use of an artificial discourse language for collaborative negotiation (Sidner 1994) along with an utterance interpretation module and an utterance generation module. However, as opposed to (Allen 2000) which places the interaction acts after the dialogue manager, Sidner places the utterance intention language before the discourse manager. This leads us to believe that it may be possible to build a practical communication language on either side of the dialogue/discourse manager.

Approach

We are developing a semantic representation for task communication language (TCL). Again, because of the volatile nature and ever expansion of what is deemed 'practical' and what can currently be modeled, we have developed TCL to allow for constant expansion. In this section, we provide a brief look into our TCL semantics as well as the entire modeling methodology. We describe current work in implementing our TCL and running empirical evaluations.

Task Communication Language Semantics

TCL is first wrapped in an outer layer that allows various social aspects of the conversation to be defined for later expansion of conversation paradigm. Other possible conversation paradigms that this will allow include multi-human or agent, participant chaining, and multiple conversations. Also included are interpretation aspects, which allow for adaptive learning on the interpretation mechanism through direct feedback from the discourse or behavior model.

Agent Communication Language Overhead: The current overhead consists of a header that includes a series of fields and their respective values. The *Generator* field represents the agent, human or otherwise, that generated the utterance or gesture. The *Addressee* field represents the direct intended receiver(s) of the utterance or gesture and the *Receiver* field represents all agents who received the utterance or gesture. Uncertainty is introduced by allowing the values of these fields to contain a first-order logic expression of various confidence ratings in addition to agent identifiers. For instance, it may not be directly known who generated a message or if a particular agent heard a message. The *Interpretation Stack* holds all information obtained at all levels of translating the original perception, such as sound, text, or movement, to the TCL representation. The interpretation stack can later be used with a feedback mechanism for improving the accuracy of the interpretation mechanisms once the actual meaning has been confirmed. The *Content* represents the TCL message itself, which will be further expanded.

Meaning-Action Concept: A meaning-action concept (MAC) is a particular meaning of an utterance or gesture, which may or may not have an associated action. It is the basic vocabulary block of TCL. Examples in the Stratagus (Stratagus 2005) domain include the concept of an engineer "engineer_01" or perhaps any engineer "engineer", the building of a new bunker "build(builder:engineer_01, building:bunker)" or even the proposal of building a new bunker "propose(action:build(builder:engineer_01, building:bunker))".

Meaning-action concepts are divided into layers according to the task model, and may be nested in definition. Examples of generic definitions in the task-domain ontology include: "propose(action:)", "propose(goal:)", "reject(goal:)", "counter-propose(action:)", "query(justification(action))". In these examples, action is defined in a separate domain-dependent ontology and accompanying grammar.

Meaning-action concepts are also defined in an ontological format, which allows for rollback to known concepts. For example, a counter-proposal is a child of proposal. Therefore, if a particular agent did not know how to handle a counter proposal, the agent may treat it as a proposal. The distinction of concepts is made for intelligent protocol modeling systems as well as for generation mechanisms, such as "instead why don't we…" or "nah, how about…". As another example, confidence ratings may vary. For example, "I'll get on that right away!" may correlate to "commit(confidence:100)" while "Well, I don't know… I'll see what I can do" may correlate to "commit(confidence:15)". The determination of these values is left to user modeling. If a particular agent implementation did not know how to handle confidence values, then it may treat both as just "commit( )", which will commit to what is in context.

The ontological organization also allows for the mappings of meaning-action concepts to a root dialogue tag for the incorporation of dialogue tag-sets and associated benefits into the dialogue manager.

Following the nature of the complexity of human language, the vocabulary space of meaning action concepts will explode. Using ontological structures is essential in the management of future vocabulary spaces.

Task Communication Expression: The task communication expression, representing the guts of the utterance or gesture with respect to communication, is a first-order logic based expression of various meaning-action concepts. This representation allows us to account for conjunction, where a particular utterance can carry multiple meanings. For example, "Don't touch that!" in a harsh voice can be used for both a constraint and reinforcement learning. This representation also allows for us to account for disjunction in which a particular meaning is ambiguous. It also allows us to express a variety of utterances such as "upgrade all soldiers," "if there are any enemy soldiers, attack them."

Task Communication Methodology

In the previous section, the vocabulary and semantics of TCL was described. This section outlines the design methodology used to incorporate existing dialogue models into the TCL framework and integrate it successfully into a working agent implementation. The details of this section represent work in progress.

Dialogue Models: Various dialogue models, such as negotiation, and collaborative problem solving can be obtained through a number of areas. The models can be human-human, agent-agent or mixed. However, it is recommended to use models directly derived from natural human interaction, otherwise the resultant model may not be readily useable by human participants.

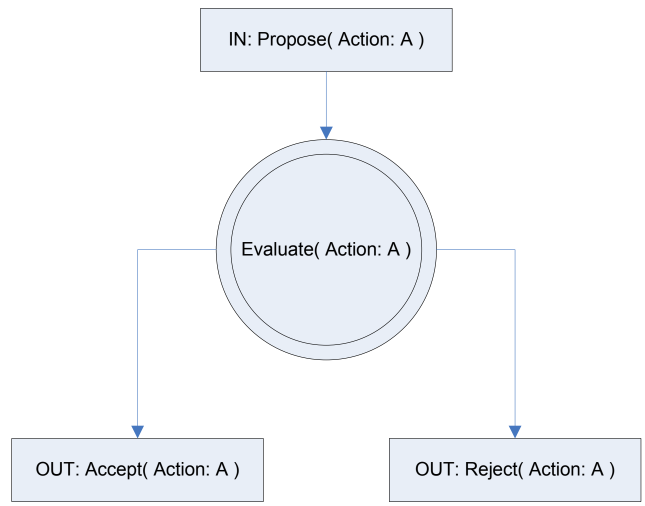
Many models of agent-communication including negotiation, team formation, and such may easily be adapted to our framework. Other agent interaction models and behavioral models, including learning by description, adjustable autonomy, command and control may also be adapted.

Task Modeling: The first step is to use and expand the set of task concepts as necessary. Task concepts are borrowed from problem solving concepts (Blaylock 2002). A subset may include objectives, recipes, actions, resources, situations, states, constraints, beliefs, intentions, metrics and priorities. In addition, a set of task operators should also be defined. A subset may include adoption, selection, deferment, abandonment, release, identification, evaluation and modification.

Task-Communication Modeling: The second step is to interleave the task concepts, task operators and communicative concepts into a behavioral model of the agent. There is ongoing work on the representation of the task-communication model, it is hoped to be able to leverage the protocol modeling and formal technique based properties of hierarchical colored Petri nets. A conceptual model is provided below.

The communicative acts are presented in rectangles and the task operators are presented in circles. The task concepts (action in this case) connect the operators and communicative acts together.

There is ongoing research at this point to apply protocol engineering (specification, verification and testing) to prove various properties about this model.



Interaction Modeling: Using the task-communication model, one can extract the possible input-output sequences.

In this case, a proposal is followed by either an acceptance or rejection of the proposed item. These models should be much more complex. For example, a proposal might demonstrate misinformation on the generators behalf and result in statements, which render the proposal useless, or perhaps if the proposal itself was not understood there may be a sub-clarification dialogue in resolving the proposal.

The interaction model can then be applied against known dialogue sequences. 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 it could be expanded to include the cases that dialogue sequence reveals.

The purpose of the interaction model is to incorporate and evaluate the composition of many task-communication models. It is yet to be seen whether protocol-engineering approaches may assist in this integration or evaluation process.

Human to Task Communication Language

We are currently developing a full natural language front-end similar to natural language translation by treating our task communication language as such. We have been generating grammars accordingly and have discovered that by abstracting the grammars into separate layers, the task-model and the domain-implementation can successfully be domain-independent.

Proposed Evaluation

We are in the process of implementing an intelligent agent in Stratagus (Stratagus 2005) with the theoretical modeling techniques described in this paper. Stratagus is an open-source real-time strategy game engine hosted on SourceForge. We are currently working with the Battle of Survival data set but hope to expand to multiple game implementations by using the Wargus and the Magnant data sets as well.

We will run a series of human investigation trials, asking the user to communicate with the system, and then comment on the naturalness of their experience.

Our hope is to not only demonstrate the feasibility and power of this approach to dialogue management, but also empirically evaluate many ACL-based dialogue models with real human investigations.

Discussion

This work represents a shift in development where a dialogue manager is no longer the center of control for the communication of an intelligent agent, but rather acts as a pass-through creating TCL for every user utterance or gesture and generating output (of whatever modality) when receiving TCL. The task-communication model is now the driving force in the communications aspect of the agent.

Mixed-Initiative Control

The task-communication model will have to model mixed-initiative interaction, adjustable autonomy and turn taking accordingly. It is not yet clear whether this will be done through direct task operators or through the introduction of inhibitors on operations.

There is ongoing evaluation into protocol verification techniques to prove or disprove various mixed-initiative qualities of a given interaction model.

Task-Communication and Interaction Hierarchy

It is hoped that mixed-initiative will be controlled using nested task operators, represented in a hierarchical form. This would yield the layering of communication models. Proposed low-level modes of operation include basic dialogue types (persuasion, inquiry, deliberation, formal argumentation, information argumentation, clarification, information absorption, active listening) and higher-order models would then be built on those modes.

Related Work

There has been much work in generic dialogue systems, especially by (Allen 2000) and (Sidner 2002). However, these focus more on the dialogue management rather than the communications between the dialogue manager and the behavioral agent.

There has also been work in natural language itself as an agent communication language (March 2004), where one agent generates natural language output for another agent to parse and interpret accordingly.

(Montesco 2005) describes work on a Universal Communication Language (UCL) based on Universal Network Language (UNL), which is used to allow communication among people of different languages. UNL comes with a library of universal words (translatable into every language) as well as relation and attribute labels. Although the work demonstrates the ability of an agent to understand universal concepts and relations, it is unclear how this leads to change in the agent's behavior, or communicative modeling such as negotiation or coordination.

(Johannesson 1998) has done work modeling agent communication languages in first order logic. Specifically, creating a language for specifying, creating and monitoring obligations, called First order Action Logic (FAL).

Our earlier work on complex command and control has yielded results in real-time behavior development and management (Lee 2004).

Conclusion and Future Work

We have presented our current theoretical methodology towards incorporating a variety of behavior based dialogue models into a single working implementation. We have proposed the use of a practical communication language, TCL with which we hope to demonstrate the feasibility of this approach.

Much work is to be done in developing semantics for the representation of meaning-action concepts as well as the task-models, the task-communication models and the interaction models.

It is hoped that layering will be performed within the task-communication and interaction models, creating an interaction core, to facilitate the development of many future dialogue theories.

References

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.

James Allen, George Ferguson, and Amanda Stent. "An architecture for more realistic conversational systems." In *Proceedings of the 6th International Conference on Intelligent User Interfaces*, Santa Fe, NM, Jan 2001, pp. 1-8.

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.

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.

D. J. Duke, P. J. Barnard, D. A. Duce, I. Herman, and J. May. "Human-Computer Protocols." In *Proceedings of the Workshop on Continuity in Human Computer Interaction*, Scheveningen, The Netherlands, April 2-3, 2000.

Paul Johannesson, and Petia Wohed. "Modeling agent communication in a first order logic." *Accounting Management and Information Technologies*, No. 8, Elsevier Science Ltd., pp 5-22, 1998.

John R. Lee, and Andrew B. Williams. "Behavior development through task oriented discourse." *Computer Animation and Virtual Worlds*. Wiley Interscience, Vol. 15, Is. 3-4, pp 327-337, June 2004.

Olivia C. March. *Natural language as an Agent Communication Language*. Honours thesis, Department of Computer Science and Software Engineering, University of Melbourne.

Carlos A. Estombelo Montesco, and Dilvan de Abreu Moreira. "UCL – Universal Communication Language." to appear in the *Journal of Research on Computer Science*, ISSN 1665-9899.

James Odell, H. Van Dyke Parunak, and Bernhard Bauer. "Representing Agent Interaction Protocols in UML." In *Agent-Oriented Software Engineering*, P. Ciancarini, and M. Wollridge, Eds. Springer, pp. 121-140, Berlin, 2001.

C. L. Sidner. "An artificial discourse language for collaborative negotiation." In *Proceedings of the Twelfth National Conference on Artificial Intelligence*, MIT Press, Cambridge, MA, Vol.1: 814-819, 1994.

C. L. Sidner. "Building Spoken Language Collaborative Interface Agents." *Technical Report, Mitsubishi Electric Research Laboratories, TR2002-038*. August 2002.

Stratagus, A Real Time Strategy Game. SourceForge Website. http://stratagus.sourceforge.net/