Towards domain-independent, task-oriented, conversational adequacy

Darsana P. Josyula[1], Michael L. Anderson[2], and Don Perlis[1,2]
(1) Department of Computer Science
(2) Institute for Advanced Computer Studies
University of Maryland, College Park, MD 20742
{darsana, mikeoda, perlis}Ccs.umd.edu

Humans are often able to communicate effectively to solve a joint task, *even* when they do not speak much of a common language, for instance, when buying a train ticket in a foreign country. Although the conversation might not be fully fluent and error-free, the communication does not break down; instead, through clarifications, confirmations and inferences, the dialog can proceed towards solving the common task. Key to this is the ability to note an error (or miscommunication), assess it, and address it. We believe that these principles of human dialog repair can be used to enhance the conversational abilities of task oriented interactive systems, allowing them to interact more smoothly with human users (Perlis, Purang, & Andersen 1998). To this end, we are developing an agent that can act as an interface between the user and a task-oriented system, providing a bed for dialog correction and repair.

Our agent, ALFRED (Active Logic for Reason Enhanced Dialog) is built on a time-tracking, logical framework called active logic (Elgot-Drapkin *et al* 1993; Elgot-Drapkin & Perlis 1990). Active logic works by combining inference rules with a constantly evolving measure of time (a "Now") that can itself be referenced in those rules. As an example, from Now(t)-the time is now "t"-one infers Now(t+I). Special rules governing the persistence of formulas in the database, including rules for disinheriting direct contradictions, along with a quotation mechanism which allows formulas to refer to one another, give active logic the expressive and inferential power to monitor its own reasoning, allowing it to watch for errors (such as mismatches between conveyed and intended meanings), note temporal conversational cues such as pauses that may signal a turn change, and to exert reasoned control over its past and upcoming inferential processes, including re-examination and alteration of beliefs and inferences.

Capabilities

Some of the capabilties that ALFRED currently has are as follows.

1. Understanding the use-mention distinction

   ALFRED is sensitive to whether a word is being used or mentioned, and interprets the relevant utterance accordingly. Following (Saka 1998), we have chosen to characterize the use-mention distinction in terms of the possible ostensions of words. That is, we consider X is being "used", if the speaker intends to direct the thoughts of the audience to the extension of X; and X is being "mentioned", if he intends to direct the thoughts of the audience to some item associated with X *other than its extension.* See (Anderson *et al* 2002) for more details. The features that ALFRED uses to make this distinction are context, cues and meta-dialog.

2. Using meta-dialog

   ALFRED is capable of engaging in meta-dialog with the human user when necessary, in order to identify miscommunication problems and perform dialog repairs.

3. Learning new words

   ALFRED can learn new names that refer to already existing objects in the domain. It can also learn new ways to accept known commands. This feature can allow a given user to negotiate a more comfortable vocabulary for interaction, and helps to increase the vocabulary of the system as a whole. We are working on methods of structuring and manipulating conceptual relations to allow ALFRED to learn genuinely new objects and concepts.

4. Maintaining context

   ALFRED maintains the context of the ongoing conversation by keeping track of user intentions (interpretations of past utterances), needs and expectations. A "need" is created when there is some requirement (like determining the meaning of a word) that has to be met before the system can interpret a user utterance. An "expectation" is created when the system expects a particular kind of response from the user.

5. Identifying miscommunication

   ALFRED recognizes miscommunication problems by looking for contradictions in its interpretations of the user's intentions (Traum *et al.* forthcoming).

6. Connecting to different domains

   ALFRED functions as a kind of translator between human natural language and the specialized language of the task-oriented domain, and enhances the performance of the original system interface by incorporating a suite of dialog error detection and repair strategies. These can allow a dialog to continue even in the presence of various kinds of miscommunications.

   The upshot of this is that ALFRED can act as a natural language interface between a user and *any* task-oriented system as long as the specialized language of the task-oriented system is specified explicitly to ALFRED. This specialized language can vary from menu driven commands in the simplest case to natural language-like commands in a more complicated scenario. However, since the same principles and reasoning strategies are being used in each case to enhance the dialog, orienting ALFRED to a different domain is as easy as changing the files specifying the specialized language for that domain.

Example

Consider a task-oriented system whose domain is defined by the following objects and commands.

- Domain objects

  train: Metroliner, Northstar, Bullet; city: Baltimore, Richmond, Buffalo
- Domain commands

  send: [send train city]

A stepwise example of command translation and intentional discernment follows:

1. User types "Send Metro to Baltimore"; the parser sends output to ALFRED
2. Since it is a request for an action, ALFRED first checks whether it is a command for the domain or for itself.
3. Since it is a domain command, ALFRED will get the structure of the command from the domain commands knowledge base. In this case, the command structure is *send - [send train city],*
4. ALFRED then tries to disambiguate the object references in the user utterance by finding the mapping from these object references to known domain objects of the types specified in the structure.
5. Since there is no direct mapping between "Metro" and any existing train in the domain, ALFRED now "needs" the meaning of the word "Metro". This "need" causes ALFRED to temporarily stop further interpretation of the utterance until this "need" gets satisfied.
6. This "need" then causes ALFRED to introspect (to examine its existing knowledge base or try to reason) about the meaning of the word "Metro". Since it does not get an answer after introspection, it asks the user: "What does 'Metro' mean?", and "expects" the user to provide the answer. This "expectation" helps in keeping track of the context of the conversation, which is essential for understanding the dialog state, and has particular use in helping distinguish whether a word is being used or mentioned in an utterance.
7. The user says "Metroliner". The expectation, above, helps ALFRED to interpret this elliptical utterance as " 'Metro' means Metroliner". Hence ALFRED would learn this meaning so that at any later point in time, if the user says "Metro", ALFRED would be able to identify the correct referent by introspection.
8. Once ALFRED gets the meaning of "Metro", the "need" created above is satisfied, which in turn causes the further interpretation of that user utterance.

Conclusion

ALFRED can enhance the interactive capability of a task-oriented computer system by adding the ability to detect and recover from miscommunication problems, including ambiguous references, incompatible or contradictory user intentions (Traum *et al.* forthcoming), and the use of unknown words.

This technology can already be usefully applied to current application domains, such as home-control software, and we expect that the techniques employed can be refined and extended to handle more sophisticated domains.

References

Anderson, M. L.; Okamoto, Y.; Josyula, D.; and Perlis, D. 2002. The use-mention distinction and its importance to HCI. In *Proceedings of the Sixth Workshop on the Semantics and Pragmatics of Dialog.*

Elgot-Drapkin, J., and Perlis, D. 1990. Reasoning situated in time I: Basic concepts. *Journal of Experimental and Theoretical Artificial Intelligence* 2(l):75-98.

Elgot-Drapkin, J.; Kraus, S.; Miller, M.; Nirkhe, M.; and Perlis, D. 1993. Active logics: A unified formal approach to episodic reasoning. Technical Report UMIACS TR # 99-65, CS-TR # 4072, Univ of Maryland, UMIACS and CSD.

Perlis, D.; Purang, K.; and Andersen, C. 1998. Conversational adequacy: mistakes are the essence. *Int. J. Human-Computer Studies* 48:553-575.

Saka, P. 1998. Quotation and the use-mention distinction. *Mind* 107:113-135.

Traum, D.; Andersen, C; Chong, Y.; Josyula, D.; O'Donovan-Anderson, M.; Okamoto, Y.; Purang, K.; and Perlis, D. forthcoming. Representations of dialogue state for domain and task independent metadialogue. *Electronic Transactions on Artificial Intelli-*