A Discourse System for Conversational Characters

Ron Zacharski

Computing Research Laboratory
New Mexico State University
Las Cruces, New Mexico U.S.
raz@nmsu.edu

Abstract. This paper describes a discourse system for conversational characters used for interactive stories. This system is part of an environment that allows learners to practice language skills by interacting with the characters, other learners, and native speakers using instant messaging and email. The dialogues are not purely task oriented and, as a result, are difficult to model using traditional AI planners. On the other hand the dialogues must move the story forward and, thus, systems for the meandering dialogues of chatterbots (for example, AliceBot) are not appropriate. Our approach combines two methods. We use the notion of dialogue game or speech act networks [1][2] to model the local coherence of dialogues. The story moves forward from one dialogue game to another by means of a situated activity planner [3].

1. Introduction

This paper describes a discourse system for conversational characters used for language learning. Language learners take part in an adventure story and must use dialogues similar to those found in basal language textbooks to move forward in the adventure. Successful completion of the adventures requires that the students collaborate on various tasks. Students communicate among themselves, advanced students of the language, and with conversational characters using a range of computer-mediated channels including email, instant messaging, audio and video files, and web pages. The conversational characters communicate among one another by structured instant messages. There is no central processor that guides the story; the progression of the story is accomplished by characters pursuing a reactive strategy.

Placing conversational characters in this adventure environment is more pedagogically sound than creating standalone conversational agents. The state of the art is currently not sophisticated enough to create interesting standalone conversational agents that can take part in creative dialogues. However, we can create conversational agents that can take part in highly predictable dialogues. These agents can be extremely useful, particularly for novice and intermediate students. Situating such an agent in this adventure has a number of important advantages. First, the adventure includes interacting with other people (fellow students and experts alike). This provides students with an opportunity to practice language skills in a range of contexts. Conversational characters augment this environment by being tireless discourse participants. An example of a dialogue (an instant message exchange) produced by the system is shown in (1):

(1) Maria: *hola, ¿ésta Rachel?*
Ann: *Sí*
Maria: *Qué bueno, ¿tienes el telefono de Marta?*
Ann: *Sí*
Maria: *¡Qué Bién! ¿Cuál es el número?*
Ann: *282-8992*
Maria: *Ok, ¡gracias!*

The implemented prototype system allows three students of Spanish to take part in a magical adventure centered on the famous Mexican polyglot and heroine, La Malinche.

The discourse system for these characters consists of three main components:

- A knowledge store representing the short and long term memory of the conversational character including what the character is currently focusing on. This component is based on file card pragmatics [4], a computational model of a theory of givenness described by Gundel, Hedberg, and Zacharski [5].
- A speech act network described in §2, which represents the coherence of short speech act sequences.
- A situated-activity-based planner described in §3, which determines the actions of the conversational character.

2. Speech Act Networks

As Austin [6], Searle [7], and others have noted, language is not simply a system for conveying thoughts and information. Language is a set of actions that enable people to cooperatively live in the world. These linguistic actions, such as promises, requests, assertions, occur within an implied background of standard practices based on the current context including what is normal in the particular speech community in which the act is situated. Some linguistic actions, like requests and promises, can be initial moves in a conversation and define a simple network of possible continuations For example, if someone makes a request, you can promise, decline, or make a counter-offer; if you then make a promise, the other person can accept or decline that promise; if that person accepts the promise you can report completion or revoke the promise. Any of these acts can be expressed in a number of different ways.

This notion of speech act networks has been used by Carletta [1] to develop agents that describe map routes. In the current system this notion of dialogue games is extended to include the standard 'micro' dialogues that occur in basal language learning textbooks.[1] One motivation for doing so is that the sentences that make up a standard dialogue typically are learned as one unit by the language learner. These speech act networks are modeled by finite state networks.

[1] An example of such a micro-dialogue is utterances 3-7 of (1) above.

It is important to distinguish these speech act networks from finite state models of discourse, which are typically used for spoken dialogue systems. In such dialogue systems a large finite state network is used to model the entire conversation. An extremely simplified version of a network for a flight reservation system might look like:

ask origin airport > ask destination airport > ask departure date > etc.

An agent based on this network would simply progress through the set of questions. If the user responded to the initial question *Where will you be flying from?* with *I'd like to fly from El Paso to Houston on December 8th* the system would only extract the information related to the city of origin, ignore the rest, and ask the redundant question *Where will you be flying to?* This problem can be fixed, to an extent, by creating a network with more transitions; however, the size of the network may become problematic. In addition, while this network approach may work for dialogues of simple inquiry (travel, financial, etc.) it will not scale up to model dialogues dealing with complex tasks—for example, the diagnosis and repair of equipment or dialogues dealing with negotiation. Other problems with finite-state dialogue models include difficulties with tasks that involve interdependencies (since backtracking is difficult) and difficulties with verification—asking for verification of each item of information tends to be unnatural and cumbersome.[2] These criticisms do not apply to the use of speech act networks. Speech act networks represent very short exchanges and represent regularities in language use.

3. Situated action planner

The most widely recognized theory of dialogue within NLP research uses inferential planning. This method is based on the work of Austin [5] who argued that uttering a sentence is a type of action that is performed by the speaker with the intention of having a specific effect on the addressee. Based on this theory, Cohen and Perrault [9] argued that a standard STRIPS-type AI planner can be used to model dialogue. This notion of plan-based dialogue has been used successfully in a wide range of applications and is still an active area of research. However, Bickmore and Cassell [10], among others, argue that this architecture is not well suited for dynamic, real-time systems such as conversational characters used for interactive, dramatic stories. Instead of STRIP-based planners researchers have used reactive systems (see, for example, [11]) or reactive systems combined with traditional planners (for example, [10]). Our approach uses a situational action planner along with the speech act networks described above. While the prototypical situated action planner has no internal representation and is stateless (see, [3]), this system violates these restrictions and has both an internal knowledge store and a sense of state (from the use of the speech act networks described above, and from the knowledge store). The operators of the system resemble those of a traditional planner. Each operator consists of a **situation**, which is a set of propositions that must be in the current utterance and/or knowledge store in order for the situation to be applicable, an **action**, which describes

[2] See [8] for a discussion on the limitations of finite state models for spoken dialogue systems.

the action to perform (for example, initiate an instant message exchange with Ann with the content speech-act-network-x), and **result**, which describes changes to the knowledge store. This is similar to the notion of strong autonomy described by Mateas and Stern [11]. An agent based on this notion of strong autonomy chooses its next action based on local perception of its environment plus internal state (knowledge store). At any point, multiple actions may be applicable. For example, the character might choose to continue in an activated speech act network or choose to interrupt that network in favor of pursuing a new action. Choosing among these alternatives is based on a notion of activation similar to that proposed by Maes [13].

4. Conclusion

In this paper I have described an dialogue architecture for conversational characters used for interactive stories. The innovation of the architecture is in its combined use of speech act networks and a situated action planner. We have implemented this architecture in a system that engages three language learners in a Spanish adventure.[3]

References

1. Carletta, J. 1990. Modelling variations in goal-directed dialogue. In Proceedings of the International Conference on Computational Linguistics 13.324-326.
2. Winograd, T. 1988. Where the action is. Byte 13(13).256a-258.
3. Agre, P, and Chapman, D. 1988. What are plans for? MIT AI Memo 1050a.
4. Zacharski, Ron. 1993. A discourse pragmatics model of pitch accent in English. University of Minnesotat dissertation.
5. Gundel, Jeanette, Nancy Hedberg, and Ron Zacharski. 1993. Cognitive status and the form of referring expressions in discourse. Language 69.274–307.
6. Austin, J. 1962. How to Do Things with Words. Cambridge, MA: Harvard University Press.
7. Searle, J. 1969. Speech Acts: An Essay in the Philosophy of Language. London: Cambridge University Press.
8. McTear, M. 1998. Modeling spoken dialogues with state transition diagrams. In Proceedings of the Fifth International Conference on Spoken Language Processing (ICSLP-98) 1223-1226.
9. Cohen, P, and Perrault, C. 1979. Elements of a plan-based theory of speech acts, Cognitive Science 3(3):177-212. Reprinted in Readings in Natural Language Processing, ed. By Grosz, B, Jones, K, and Webber, B. 1986. Los Altos, CA: Morgan Kaufmann Publishers.
10. Bickmore, Timothy, and Justine Cassell. 2000. “How about this weather?” Social dialogue with embodied conversational agents. Proceedings of the AAAI Fall Symposium on Socially Intelligent Agents.
11. Jordan, Pamela, Carolyn Rosé, and Kurt VanLehn. 2001. Proceedings of AI in Education. Austin, Texas, May 2001.
12. Mateas, Michael, and Andrew Stern. 2002. A behavior language for story-based believable agents. AAAI Working Notes of Artificial Intelligence and Interactive Entertainment.
13. Maes, Pattie. How to do the right thing. Connection Science Journal 1(3)

[3] More detailed information on this work can be found at http://www.zacharski.org/characters.