Enriching Chatter Bots With Semantic Conversation Control

Chayan Chakrabarti

Department of Computer Science, University of New Mexico, Albuquerque, NM 87131

cc@cs.unm.edu

Abstract

Businesses deploy chatter bots to engage in text-based conversations with customers that are intended resolve their issues. However, these chatter bots are only effective in exchanges consisting of question-answer pairs, where the context may switch with every pair. I am designing a semantic architecture that enables chatter bots to hold short conversations, where context is maintained throughout the exchange. I leverage specific ideas from conversation theory, speech acts theory, and knowledge representation. My architecture models a conversation as a stochastic process that flows through a set of states. The main contribution of this work is that it analyses and models the semantics of conversations as entities, instead of lower level grammatical and linguistics forms. I evaluate the performance of the architecture in accordance with Grice's cooperative maxims, which form the central idea in the theory of pragmatics.

Motivation and Problem Statement

Communicating with chatter bots has come a long way, from pioneering AI demonstrations like ELIZA, to modern software like Siri. Many businesses also realize their customer service operations using chatter bots as virtual representatives (www.goarmy.com, www.alaskaair.com). But communication with these chatter bots takes the form of successive question-answer pairs, where the context may switch with every pair. The main goal of this work is to design an integrated architecture that enables the chatter bot go beyond mere question-answer exchanges, to hold a short conversation, where the context is maintained throughout the exchange. The domain is restricted to customer service situations through text-based chat. The chatter bot answers FAQ type questions, resolves customer service issues, spots opportunities during the conversation to disseminate unsolicited information (information about related services and promotions), and evaluates the semantic flow of the conversation. If the flow of the conversation requires the chatter bot to pursue a course of action beyond its programmatic capabilities, it realizes this and transfers the conversation to a human representative.

Conversation theory defines a formal framework for shared construction of knowledge between multiple conversationalists (Pask 1976). A conversation is a process that flows through a set of states (Winograd and Flores 1986). The set of states represents a speech act (GoldKuhls 2003). A conversation can have several semantic states (Ginsburg 2008). The states are proxy indicators of customer satisfaction (Stolcke et al. 1998). Sentiment detection (Pang and Lee 2008) is also another indicator that influences conversations. Hence customer satisfaction can be measured by observing the flow of a conversation through various states (Twitchell et al. 2004).

While the proposed research borrows ideas from other works, it solves a distinct problem. I am not modeling the low level abstractions of sentences, phrases, and words, or linguistic artifacts of grammar, discourse resolution, and parts of speech. I consider conversations to be the unit of analysis. Each conversation segment is a data point. I am analyzing and modeling the conversation itself, and not the lower level grammatical minutia that form individual components of the conversation. Furthermore, I am exploring well-structured conversations in a fixed domain.

Semantic Conversation Architecture

The key aspects of holding a conversation are *what to say*, and *how to say* it, which are handled by the Knowledge Engine (KE), and the Conversation Engine (CE) in the chatter bot architecture. The Chat Interface (CI) directly interfaces with the user, and has modules for pre-processing the raw text input, identifying the Speech Act, Sentiment, and Topic associated with the utterance, and passing this information to the CE and KE. See Figure 1.

The KE contains a Speech Acts Hash Set (SAHS) and a Topic Hash Set (THS). The SAHS is a data structure that

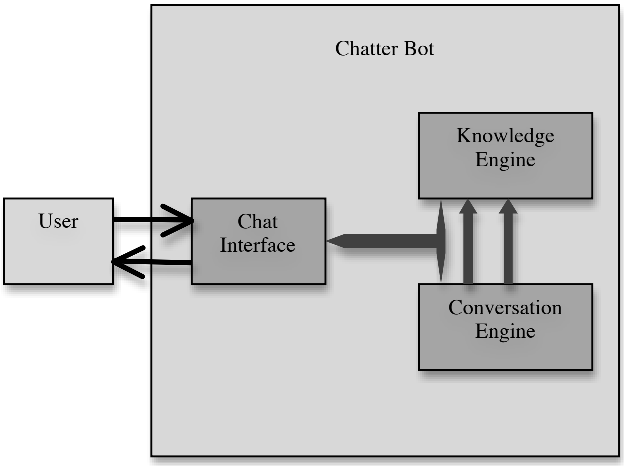

Figure 1: Overview of Chatter Bot Architecture.

stores the Speech Acts in the form of probabilistic finite state automata. The probabilities are learned from a corpus

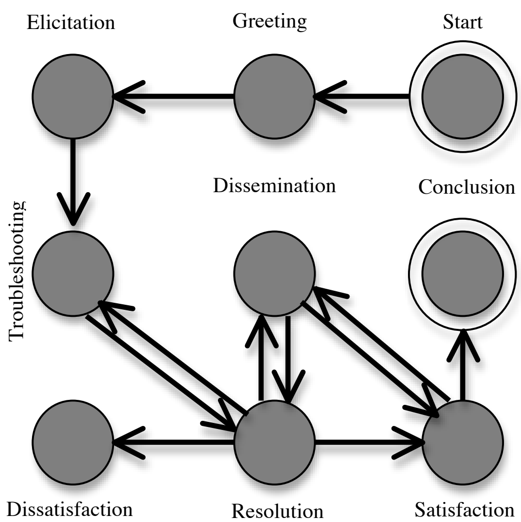

Figure 2: Conversation Speech Act represented by a probabilistic finite state automaton.

of conversations. The THS is a data structure containing specific information about the domain. The information is organized in the form of Goal-Fulfillment Maps (O'Shea et al. 2010) and ontologies.

The CE contains the probabilistic finite state representation of the conversation Speech Act. One such example, shown in Figure 2, consists of 8 states: *Greeting* (includes small talk), *Elicitation* (trying to understand issue), *Troubleshooting (working through steps to understand the issue)*, *Resolution*, *Dissatisfaction*, *Dissemination* (give unsolicited information such as possible promotions), and *Conclusion*. The CE also has a conversation planner (CP) containing a simple decision tree that engineers the conversation towards satisfactory states, guided by the transition matrix of the automaton.

Hypotheses, Datasets, and Validation

There are two specific hypotheses. First, the knowledge representation framework will have sufficient representation power to model domain knowledge. Second, the semantic conversation control algorithm will be able to detect transitions in conversation states, predict outcomes, and use this knowledge to control the conversation.

I am using four corpora of logged chat sessions. Each corpus has several million lines of chat, split in to distinct chat sessions. Two contain chat logs from human-led sessions, and two contain chat logs from chatter bot led sessions in the same domain.

The performance of my chatter bot is evaluated according to Grice's maxims of cooperation (Grice 1989): *quality* (bot gives correct information), *quantity* (bot is informative as needed), *relation* (conversation is relevant to discussion), and *manner* (conversation is unambiguous).

Current and Future Work

I had an internship at NextIT in Spokane, WA, where I had access to real-world corpora for testing my approach. I have implemented a prototype of the CE. Future work includes implementing the KE, the CI, and experimenting with conversation strategies in the CP.

References

Ginzburg, J. 2008. *Semantics for Conversation*. King's College, London. CSLI Publications.

GoldKuhl, G. 2003. *Conversation Analysis as a Theoretical Foundation for Language Action Approaches?* 8th Int'l Working Conf. on the Lang. Action Perspective on Comm. Modeling.

Grice, P. 1989. *Studies in the Way of Words*. Harvard Univ. Press.

O'Shea, K., Bandar, Z., and Crockett, K. 2010. *A Conversational Agent Framework using Semantic Analysis*. Int'l Journal of Intelligent Computing Research. Vol 1, Issue 2, June 2010.

Pang, B., and Lee, L. 2008. *Opinion Mining and Sentiment Analysis*. Foundations and Trends in Information Retrieval. Vol. 2, Nos 12 (2008) 1135.

Pask, G. 1976. *Conversation Theory*. Elsevier Press.

Stolcke, A., Shriberg, E., Bates, R., Coccaro, N., Jurafsky, D., Martin, R., Meteer, M., Ries, K., Taylor, P., and EssDykema, C. 1998. *Dialog Act Modeling for Conversational Speech*. AAAI 98.

Twitchell, D., Adkins, M., Nunmaker, J., and Burgoon, J. 2004. *Using Speech Act Theory to Model Conversations for Automated Classification and Retrieval*. 9th Int'l Working Conf. on the Lang. Action Perspective on Comm. Modeling.

Winograd, T., and Flores, F. 1986. *Understanding Computers and Cognition. A New Foundation for Design*. Norwood, NJ. Ablex Publishing Corp.