Towards Handling General Purpose Topics for a Conversational Character

Manish Mehta[1], Andrea Corradini[2]

1. College of Computing	2. Department of Linguistics
Georgia Institute of Technology	University of Potsdam
Atlanta, Georgia.	Germany.
mehtama1@cc.gatech.edu	andrea@ling.uni-potsdam.de

Abstract

We discuss a new method to address general purpose topics for a conversational agent that utilizes web directories for categorization of multiple general purpose topics. This categorized is combined with the existing ontological properties and dialog acts to create an automated meaning representation. To generate a response on questions and utterances on topics not covered by our system we resort to existing question-answering systems and resources freely available from the web.

1. Introduction

Embodied conversational agents (ECAs) aim to use and realize cues inherently peculiar to human-human communication, such as sense of presence, mixed initiative and non-verbal behaviors to hold up their end of the dialog with the user. As we start developing ECAs with the clear goal in mind to provide the user with a rich social experience and act as a conversational partner rather than a mere computational task solver, effective handling of out of domain input becomes an important issue. We believe that simplistic approaches could hamper the overall player experience rather than enhancing it and would thus contradict with our original ideas and intentions.

Current work on developing programs that can simulate typed conversation mostly rely on a template based approach to generate answers to multiple general purpose topics [Mauldin, 1994]. Other approaches have used similar template based approach to address out of domain [1] topics [Mori et. al., 2003] and engage in small talk [Bickmore & Cassell, 1999]. The range of discussion topics is still limited since it is dependent on the amount of templates that can be created off-line. Moreover, creating these set of templates requires hand crafted answers to all the possible imaginable discussion topics. In our approach, we want to reduce the authorial burden of content creation for different general purpose discussion topics.

This paper presents an approach to address general purpose topics such as movies, games, current news, food, famous places and personalities by a) using web resources, notably

Google's ontological resources to enhance the understanding capabilities of the system, and b) utilizing existing question-answering (QA) systems and resources freely available on websites to address these topics. Our framework is a computer game where a player can interact with an embodied character in a 3D world, using spoken conversation as well as 2D gesture with fairytale author Hans Christian Andersen (HCA) to learn about the writer's life, historical period and fairy tales.

2. Approach

In our system the Natural Language Understanding (NLU) module has generic rules for detecting dialog acts present in the user utterance. These dialog acts provide a representation of user intent like types of question asked (e.g., asking about a particular place or a particular reason), opinion statements (like positive, negative or generic comments), greetings (opening, closing) and repairs (clarification, corrections, repeats). These dialog acts are reused across different domains of conversation. Moreover, generic rules are used to detect the domain independent properties (e.g., dislike, like, praise, read, write etc). [Mehta & Corradini, 2006] provides more implementation details on the NLU.

During its working, the NLU categorizes the word(s) that are not processed internally into an unknown category. The longest unknown sequence of words is combined into a single phrase. These words are then sent to the web agent, which uses Google's directory structure to find out whether the unknown words refer to a name of a movie, game, or a famous personality and the corresponding category is returned to the NLU. To illustrate the processing let us assume the user asked "*do you like quake*". In this case, the NLU marks the word *quake* as an unknown category that, as such, needs further resolution. The temporary output of the NLU is thus a *yes/no-question* as dialogue act, a property of the kind *like* and an *unknown* category. The *unknown* category is resolved by the web agent into the category *game* using Google's directory.

At the next stage, inside the dialog module, the output representation from the NLU is used to reason about the next conversational move of the character. This stage of processing is performed inside a module called the conversational mover. For each conversational move of the

character, rules are defined using the concept(s)/sub concept(s), property(s)/property type and dialog act/dialog act type pairs delivered by the NLU. This provides a systematic way to connect the user intention to the characters output move.
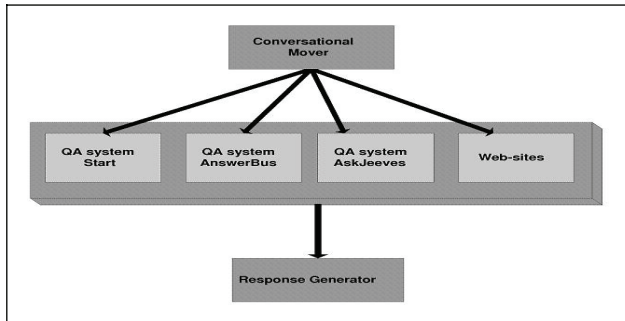
Figure 1: Output retrieval using QA systems and web-sites

When the conversational mover classifies the NLU output representation into a conversational move whose output is to be retrieved from the web, the request is sent to the web agent. The web agent, depending upon the type of move, finds a quick and concise output using three freely available open-domain QA systems: AnswerBus (Zheng, 2002), Start (Katz, 1997), and AskJeeves[2] or the web page at specific game[3] and movies[4] websites (Figure 1). The web agent employs a set of heuristics, such as removing output with certain stop words, to pick one single reply. Once a sentence is selected, we remove control/graphical characters to get a plain string that can be played by the Text to Speech component.

User:	I like My cousin vinnie
NLU:	<dialogact:user_opinion><dialogacttype:general > <concept:unknown><subconcept:cousin_vinnie > <property:like>
Google Class. :	<dialogact:user_opinion><dialogacttype:general> <concept:movie><subconcept:cousin_vinnie > <property:like>
C Mover :	user_opinion_movie
WebAgent :	My Cousin Vinny makes some good points about pointing fingers. But mostly it'll make you laugh

Table 1: The table shows the output produced by the main system components for an example utterance

At time, our classification approach faces problems when the group of words overlaps with the words in the lexicon. For example, when the user says "*Do you like the movie the Lord of the Rings*" where the words '*of*' and '*the*' have a lexical entry, their category is retrieved from the lexicon and the only unknown words remaining are "*Lord*" and "*Rings*" and the web agent is not able to find the correct category for these individual words. One solution would be to automatically detect the entries, which overlap with the words in the lexicon by parsing the Google's directory structure offline and having these entries made in the key phrase spotter. We plan to solve these issues in the future. This issue however arises only when important keywords

---

[2] www.askjeeves.com

[3] www.game-revolution.com

[4] www.rottentomatoes.com

are removed from the names of the movie, game or famous personality due to lexicon overlap. For instance, in example in Table 1, even though the word "*my*" has been removed, Google's directory structure is still able to provide appropriate categorization due to an existing lexical entry as sufficient keywords are present.

4. Conclusion & Future Steps

We presented an existing running prototype to handle general purpose topics by using data available on the websites and QA systems. On the understanding side, we have used Google's directory classification mechanism along with existing domain independent dialog acts and properties to understand these general purpose topics. In the current implementation, HCA is not able to continue a contextual conversational exchange beyond the sentence he selects as reply on these topics. Having the knowledge of topics addressable through QA systems and web-sites, we aim to conduct a mixed initiative dialog on these topics along with the normal domain-oriented conversation. The system parsing the data available on QA system and web-sites is susceptible to changes in website formats. Our development of the parsing algorithm has been designed so that HCA doesn't utter anything non-sense, however, we have not experienced any formatting issues until now.

References

Bickmore, T. and Cassell, J. 1999: Small talk and conversational storytelling in embodied conversational interface agent. *AAAI fall symposium on narrative intelligence*, pp. 87–92,.

Katz, B. 1997. Annotating the World Wide Web using natural language, In *Proceedings of RIAO Conference on Computer Assisted Information Searching on the Internet*

Mauldin, M. 1994. Chatterbots, TinyMUDs, and the Turing Test: Entering the Loebner Prize Competition, In *Proceedings of the 12th National Conference on Artificial Intelligence*, pp 16–21.

Mehta, M., and Corradini, A. 2006 Understanding Spoken Language of Children Interacting with an Embodied Conversational Character, In *Proceedings of the Combined Workshop on Language-Enabled Educational Technology and Development and Evaluation of Robust Spoken Dialog Systems at ECAI'06,* pp. 51-58.

Mori K., Jatowt A. and Ishizuka M., 2003. Enhancing Conversational Flexibility in Multimodal Interactions with Embodied Lifelike Agents, *In Proc. Int'l Conf. on Intelligent User Interfaces*, pp. 270–272.

Patel, R., Leuski, A., Traum, D. 2006. Dealing with Out of Domain Questions in Virtual Characters. In *Proceedings of the International Conference on Intelligent Virtual Agents*

Zheng, Z. 2002, AnswerBus Question Answering System, In *Proc. of the Human Language Technology Conference.*