Emotionally Driven Natural Language Generation for Personality Rich Characters in Interactive Games

Christina R. Strong, Manish Mehta, Kinshuk Mishra, Alistair Jones and **Ashwin Ram**
Cognitive Computing Lab (CCL)
College of Computing, Georgia Institute of Technology
Atlanta, Georgia, USA
{cstrong,mehtama1,kinshuk,ashwin}@cc.gatech.edu

Abstract

Natural Language Generation for personality rich characters represents one of the important directions for believable agents research. The typical approach to interactive NLG is to hand-author textual responses to different situations. In this paper we address NLG for interactive games. Specifically, we present a novel template-based system that provides two distinct advantages over existing systems. First, our system not only works for dialogue, but enables a character's personality and emotional state to influence the feel of the utterance. Second, our templates are resuable across characters, thus decreasing the burden on the game author. We briefly describe our system and present results of a preliminary evaluation study.

Introduction

Believable agents are autonomous agents that display distinct personalities, similar to traditional characters in animation, film, or literature. While there has been considerable research in developing believable agents (Reilly 1996; Mateas & Stern 2004), there have been very few attempts to generate their verbal output (Loyall & Bates 1997). Traditional approaches have required hand-authored responses for all possible situations (Mateas & Stern 2003; Corradini *et al.* 2005), making believable natural language for personality rich agents in a real time game domain difficult.

In this work, we address this problem and present an approach to natural language generation (NLG) that focuses on the personalities and emotions of the believable characters that interact with each other and human players in the real time game of Tag (Zhang *et al.* 2007). Our focus is not to create a standalone technique for NLG, but one that exists and works within the context of a real time interactive domain. This motivation is driven by the knowledge that human conversation is constantly changing and evolving as time passes. In order to make computer generated conversation believable, this same property must hold true.

Interactive domains inhabited with believable characters provide rich opportunities and present challenging requirements for natural language generation. This paper describes an approach to overcoming these challenges.

Natural Language Generation Approach

Our real time domain consists of two embodied characters named Jack and Jill who are involved in a game of Tag where the character who is IT chases other players around the game area. The behaviors authored for each character reflect their personalities. Our natural language generation system takes certain parameters (such as character personality) from the Tag environment and creates single sentence utterances based on these parameters.

Our approach to emotion tracking follows existing work (Zhang *et al.* 2007). Emotion values serve as compact representations of long-term behavior. At runtime, a character's emotional state is incremented when specific behaviors, annotated by the author, succeed or fail. The emotion increment value per behavior is defined by the author as part of specifying the character personality.

In our current implementation, we have defined four emotions (angry, relaxed, sad, happy), where the overall emotional state of the character is a combination of these four emotions. In order to produce a single emotional state, the variable with the highest value is used. A more complicated model of emotion mixing is beyond the scope of the prototype system described in this paper.

The Natural Language Generation Module (NLGM) takes the following output from the game engine: the particular game event to which the character is responding, the current emotional state of the character who is about to speak, and the author defined personality of the character who is about to speak. Using these three parameters, the NLGM outputs a sentence that reflects the events, emotions, and personalities occurring in the game.

In general, there is a lack of reusability in NLG tools. Thus, we decided to build a system that is reusable across game domains. Our system uses a set of author defined templates that can be reused across different characters and emotional states. We defined these templates as a set of sentence structures that require words and phrases from a lexicon to be filled in. The words and phrases that are used are dependent upon the emotions and personality of the speaker.

The templates we used are structured around ten major concepts. Several of these concepts are then further divided into subconcepts which allow for more variety in the kind of sentences that can be generated. This approach is extensible—new concepts can easily be added to increase the number of templates that can be created. Each concept has a word or phrase, or often a set of different words or phrases, associated with it. This provides the freedom for an author to associate arbitrary words with a concept.

These major concepts are structured as large segments of a sentence. For example, one concept corresponds to a positive phrase. This concept then has three different subconcepts: “proficient”, “knowledgeable”, and “likeable”. Each subconcept has a set of words or phrases associated with it. This is easily expandable by either adding additional subconcepts, or adding new words or phrases associated with a particular subconcept.

Illustrative Example

In the Tag domain, consider the situation where Jill has been tagged. When Jill is tagged, the game event causes her emotional state to change and she becomes happy. These changes are passed to the NLGM. In addition, the NLGM receives information about Jill’s personality (the value tempermental) and from these three parameters creates a sentence consistent with them.

For the purposes of this example we assume a set of templates is found that are representative of all the parameters, and a random template is chosen from that set. Let us say that the template chosen was $<subj - phrase><intrans - positive>$. Eventually, the system will have to choose (based on the character’s personality) where the character’s emotions should be directed, toward the character or toward someone else.

Completing the sentence is slightly more complex. In this case, the concept “intrans-positive” has three different sub-concepts: knowledgeable, proficient, and likeable. Each sub-concept has to be queried again in order to retrieve the list of words or phrases that are associated with it. This list of all possible intransitive negative words or phrases is then returned. The NLG system randomly selects one of the words or phrases to complete the sentence.

The words and phrases retrieved by queries to the database are then concatenated together to form one string, which can be output as a sentence of text to the screen. So Jill can now say something such as “Hey! I am brilliant” to whoever tagged her, without the author of the game predefining the actual sentence. The same template could result in different output depending upon a different personality type and emotional state. This relieves the author of the necessity of generating every possibility by hand and makes authorship of complex interactive experiences far more feasible.

Initial Evaluation

In order to evaluate our approach, we asked a group of seven participants (six men and one woman) to consider whether the NLG system was producing verbal output in accordance with the emotional state of the character. The question was structured as follows: the participant was given the character’s personality and a set of eight to ten sentences; they were asked to determine an emotional state for each sentence given that specific personality.

Emotion	Bad	Good	Undecided
sad	12	2	0
angry	10	4	0
happy	2	9	3
relaxed	2	12	0

Table 1: Frequency with which participants classified sentences as generally good, generally bad, or undecided.

As it is very difficult to come up with emotional states and personality types, we devised a list of emotional states and a list of personality types for players to label each example utterance. This list was an extended version, including two examples that are not actually represented in our system, of the personality types and emotions used in our system.

Due to the subjective nature of the evaluation, we did not expect the participants to correctly identify the emotional states with any significant accuracy. Participants correctly identified the emotional state portrayed with only 20% accuracy. That is, out of 56 possible sentences, 11 were exactly matched to the correct emotional state.

We asked an additional participant who had not taken part in the first two questions to classify each of the emotional states identified by the original seven participants as “generally good” or “generally bad”. When the answers were categorized in this way, the accuracy increased to 71% (40 out of the 56 sentences were correctly categorized). Table 1 provides a more detailed breakdown for the newly classified emotional states. These results are particularly encouraging. Emotional states which could have been placed in either category, such as “surprised”, contributed to some of the false positives and false negatives that occurred.

We feel that the difficulty participants encountered in distinguishing between more precise emotions is due in part to the limited vocabulary in our prototype implementation. However, we also feel it is influenced by the fact that in real conversation these emotions can be difficult to differentiate as well. This raises the question whether it increases the believability of the sentence if the emotion is exactly determined by the player. We leave this evaluation for future work.

Related Work

One of the earliest examples of natural language generation is PAULINE. With PAULINE, Hovy (1988) attempted to address the problem of saying the same thing in different ways to achieve a different effect, or incorporating pragmatics into language generation. Initial pragmatic constraints define rhetoric goals that guide planning and text realization in the natural language generator, by suggesting which wording to use when more than one is available. The concern of pragmatics is closer to the personality rich language generation we are looking to create, but is still lacking.

MINSTREL (Turner 1994) created novel stories in a medieval setting. It was given several small story ideas to populate the episodic memory, and then used a modified case-based reasoner to tell creative stories taking place in a world similar to that of King Arthur. What is unique about these stories is that they have a "moral" so to speak, indicating the reason for telling the story. While MINSTREL illustrated several interesting ideas, it is difficult to use it in a larger domain, let alone in an interactive setting such as a game.

STORYBOOK comes closer to our system in that it functions in real time, but the focus of Callaway's work (2000) was, similar to the previous work we discussed, to create a story. STORYBOOK has the distinction of being able to integrate different elements of language generation that have previously been studied separately and can produce varying styles of prose depending on the initial configuration. However, while it constructs the story in real time, it does not adapt as things change in the environment. Once the initial configuration has been set, STORYBOOK creates the entire story. This makes STORYBOOK inextensible for parameters such as emotions, which change over time.

There has been work in natural language generation for characters with rich personalities and distinctive behaviors in the Oz group at Carnegie Mellon University, by Loyall & Bates (1997), and Kantrowitz & Bates (1990). GLINDA (Kantrowitz 1990) is incorporated with the Oz architecture and explored modifying the style of output depending on the dramatic content in the story. At a high level, the work done by Loyall & Bates appears to be similar to what we are trying to accomplish. They extend GLINDA in order to give personality rich characters natural language abilities. However, looking beneath the surface we find that Loyall & Bates (1997) touch on changing the output based on the emotional state of the character, but ignores the idea of personalities being a factor in generation as well.

Future Work

Due to the lack of an explicit grammar along with a limited ontology, our system can be *overgenerative*—it will produce very similar if not identical sentences within a small span of time. One way we can solve this problem is to expand the knowledge base to use an existing system (OHara *et al.* 2004), to increase the vocabulary and understanding capability of the characters.

Finding the right way to indicate emotional states is closely related to adding character-specific parameters. In order to communicate being upset, there are some elements of sad, angry, and frustrated. The simple "winner takes all" method we currently use does not capture this distinction. A different approach that combines the values for each emotion could potentially allow for more expression.

Additionally, we would like to analyze the effect of a character's goals on her conversation. For example, if the character had the goal to become IT, then her conversation may involve hints as to where she is hiding or taunting the person who is IT. However, if the character does not want to be IT, she will most likely steer clear of those tactics and employ ones which allow her to avoid being tagged.

Concluding Thoughts

In this paper, we present a novel approach to natural language generation that focuses on the personalities and emotions of the characters, as well as the events in the game, in order to construct sentences. We have presented a prototype system that works in a real time domain to generate natural language based on the personality and emotional state of the character speaking and explained how this system reduces the burden of authoring dialog for believable characters in games. We see these two factors as necessary parameters in generating believable language in characters, and have shown that the general emotional state of a character is easily identifiable by human observers. We believe this work provides a solid foundation for the continued exploration of natural language generation for dialogue in believable agents in games.

References

Callaway, C. B. 2000. *Narrative Prose Generation*. Ph.D. Dissertation, North Carolina State University.

Corradini, A.; Mehta, M.; Bernsen, N. O.; and Charfuelan, M. 2005. Animating an interactive conversational character for an educational game system. In *Proceedings of Conference on Intelligent User Interfaces*, 183–190.

Hovy, E. H. 1988. *Generating Natural Language Under Pragmatic Constraints*. Hillsdale, New Jersey: Lawrence Erlbaum Associates.

Kantrowitz, M., and Bates, J. 1990. Integrated natural language generation systems. Technical report, Carnegie Mellon University.

Kantrowitz, M. 1990. Glinda: Natural language text generation in the oz interactive fiction project. Technical report, Carnegie Mellon University.

Loyall, A. B., and Bates, J. 1997. Personality-rich believable agents that use language. In *Autonomous Agents 97 Marina del Rey CA USA*.

Mateas, M., and Stern, A. 2003. Facade: An experiment in building a fully-realized interactive drama. In *In Game Developer's Conference: Game Design Track*.

Mateas, M., and Stern, A. 2004. A behavior language: Joint action and behavioral idioms. In Prendinger, H., and Ishizuka, M., eds., *Life-like Characters. Tools, Affective Functions and Applications*. Springer.

OHara, T.; Bertolo, S.; Witbrock, M.; Aldag, B.; Curtis, J.; Panton, K.; Schneider, D.; and Salay, N. 2004. Inferring parts of speech for lexical mappings via the cyc kb. In *Proceedings of the 20th International Conference on Computational Linguistics*.

Reilly, S. 1996. Believable social and emotional agents. Technical report, Carnegie Mellon University.

Turner, S. R. 1994. *The Creative Process: A Computer Model of Storytelling and Creativity*. Hillsdale, New Jersey: Lawrence Erlbaum Associates.

Zhang, P.; Mehta, M.; Mateas, M.; and Ram, A. 2007. Towards runtime behavior adaptation for embodied characters. In *International Joint Conference on Artificial Intelligence*.