

Influence of Personality-based Features for Dialogue Generation in Computational Narratives

Weilai Xu¹ and Fred Charles and Charlie Hargood and Feng Tian and Wen Tang

Abstract. In this paper, we present an approach for generating dialogues for characters within the context of computational narratives using personality-based features for deep neural networks. The approach integrates the requirements of both narrative genres and personality traits for the definition of character-based stylistic models. The modelling of characters’ features from existing datasets of complete stories permits the generation of personality-rich character dialogues. We present early results from an evaluation based on a sample of characters’ personality traits across different narrative genres, demonstrating variability in the resulting dialogues.

1 Introduction

Dialogues in computational narratives are a strong mean of communicating author-driven story elements through the characters themselves. Computational narratives have been well researched over the last 20 years, but mostly around exploring narrative structure(s) (e.g. [6]), whilst the dialogue-driven narratives are still hindered due to the limitations of dialogue generation. In recent years, research in natural language generation has focused on content rather than on *style*, even though style is a crucial part of narrative dialogue and its rich contents [3, 8]. Existing works incorporate stylistic information by using additional features, such as gender [2], sentiment [1], and theme [1], as well as controlling the style of generated sentences by altering these features. Most of these works only consider the local features, i.e. the features pertaining to individual sentences affecting expression alteration in the scope of each individual sentences. However, it is necessary to use higher level knowledge for generating narrative-based dialogues, which represents the authorial intent and provides consistency over the story generated. Based on the consideration above, our motivation is to generate dialogues with features derived from the story, with which the generated dialogues can reflect some styles in the context of story design and authoring.

2 Methodology

McKee [5] points out that characters are assigned various characteristics based on their functions within a story. This was the basis of our approach where the variety of characters’ personality traits has the potential to generate impact on the utterances whilst retaining the authorial story intents. The main processes of the methodology are listed as following:

1. **Data collection and process:** The dialogues are collected from IMSDb since [7] demonstrated that rich narrative knowledge can be extracted from screenplay. All screenplays are processed by removing all elements other than dialogue, and the dialogues are

Table 1: Generated samples with personality traits combinations based on an input utterance collected from a movie of Drama genre.

Trait Combination	Input: Man, I just been sitting here thinking.
Ex: <i>H</i> , Em: <i>H</i> , Ag: <i>H</i>	Output: Oh, my God, what are we doing here?
Ex: <i>H</i> , Em: <i>H</i> , Ag: <i>M</i>	Output: You know, it ’s been a long time since we got out of the car.
Ex: <i>H</i> , Em: <i>L</i> , Ag: <i>M</i>	Output: How do you know I ’m here?
Ex: <i>M</i> , Em: <i>L</i> , Ag: <i>H</i>	Output: How long does it take?
Ex: <i>M</i> , Em: <i>L</i> , Ag: <i>L</i>	Output: No, we ’re not.
Ex: <i>L</i> , Em: <i>H</i> , Ag: <i>L</i>	Output: No way!
Ex: <i>L</i> , Em: <i>M</i> , Ag: <i>H</i>	Output: You ’re right, sir.

- parsed into utterance pairs (partial dialogue) according to character (speaker) names by movie and genre.
2. **Label:** We specifically select three primary traits (extraversion, emotional stability, and agreeableness, or abbr. Ex, Em and Ag) using principal component analysis of training datasets. Each response of the utterance pairs is labelled with *High*, *Medium*, or *Low* on these three traits according to their character’s personality trait score using Mairesse’s personality tool [4] based on the Big-Five personality model. This score is here calculated from all the sentences uttered by a single character from a complete movie, representing an *overall* personality score of this character. This score is in the range 1 to 7, which is then divided into 3 sub-ranges: *Low* in the range lower than 3.5, *Medium* in 3.5 - 4.5, and *High* in the range greater than 4.5. For instance, for the extraversion trait, the label *Low* denotes introvert and the label *High* denotes extravert.
 3. **Train Seq2Seq dialogue system:** The LSTM Seq2Seq model is pre-trained on Cornell’s² and DailyDialog³ corpora, which only contain dialogue pairs. The pre-trained model then is fine-tuned on the IMSDb corpus with the personality labels described above. During the encoding phase, the unlabelled utterances are processed through the encoder for generating the context vectors. Then the corresponding responses with the labels representing the personality of characters are used as input to the decoder for the supervised learning stage.
 4. **Generate responses:** The fine-tuned model tends to generate possible responses in variety according to different target personality trait combinations. We randomly selected 30 utterances from the test dataset for each genre out of 4 major narrative-rich genres (Drama, Romance, Action, and Thriller). Therefore, each test utterance with 28 different combinations ($3^3 + 1$ blank control) is processed and 100 possible responses for each combination (2,800 in total) are generated using beam search with beam size 100.

¹ Bournemouth University, Faculty of Science and Technology, UK, email: wxu@bournemouth.ac.uk

² https://www.cs.cornell.edu/~cristian/Cornell_Movie-Dialogs_Corpus.html

³ <http://yanran.li/dailydialog>

