Towards Generating Stylistic Dialogues for Narratives using Data-Driven Approaches

Weilai Xu, Charlie Hargood, Wen Tang, and Fred Charles

Bournemouth University, Faculty of Science and Technology,
Creative Technology Department, Poole, BH12 5BB, UK
{wxu,chargood,wtang,fcharles}@bournemouth.ac.uk

Abstract. Recently, there has been a renewed interest in generating dialogues for narratives. Within narrative dialogues, their structure and content are essential, though *style* holds an important role as a mean to express narrative dialogue through telling stories. Most existing approaches of narrative dialogue generation tend to leverage hand-crafted rules and linguistic-level styles, which lead to limitations in their expressivity and issues with scalability. We aim to investigate the potential of generating more stylistic dialogues within the context of narratives. To reach this, we propose a new approach and demonstrate its feasibility through the support of deep learning. We also describe this approach using examples, where story-level features are analysed and modelled based on a classification of characters and genres.

Keywords: Dialogue Generation · Interactive Narratives · Dialogue Style · Neural Networks

1 Introduction

Stories can be represented through plot structures [20,21,24], and are often equated to a sequence of narrative actions [22], normally represented as objects, subjects, actions and other features. Narrative actions can be presented through discourse in different ways: visual [14], audio, and linguistic [10].

Dialogue plays an essential role in narratives, by which the story or plot can be conveyed, and characters are able to express their intentions, emotions, and personalities. In recent years, dialogue generation has become ever more popular as evidenced through an increase in research in natural language generation (NLG) [4,7,9,26]. However, most of these works concentrate on task-oriented dialogue generation and "chit-chat" conversations, along with only few works specifically undertaken within the context of interactive narratives.

Style is a crucial part of narrative dialogue, responsible for telling stories vividly and lively. Lin and Walker [14] point out that many linguistic styles can be designed to express dramatic characters in dialogue from television. Existing research mostly relies on predefined hand-crafted rules or plans [2,5] to generate dialogue in narratives. These approaches are supposed to generate grammatically correct dialogues with the empirical knowledge, but in the meantime, they suffer

from several limitations: a lack of style, difficulty to scale-up, and often a very onerous manual labour.

In dialogue generation, styles are usually described using linguistic features (e.g. sentence length, number of nouns/verbs) [14,30] or speaker profile features (e.g. age, occupation) [12]. To some extent, these features can represent the dialogue content style by changing the expression of the utterances (e.g. using interjections). While in narratives, those types of stylistic dialogues do not have a strong relationship with the storyline (i.e. the story plot has little effect on the dialogue generation). Here, we propose to introduce different characters' relationship types and genres as a measure of variability in dialogue generation.

In this paper, we propose a data-driven approach to generate story-level stylistic dialogues for narratives. To address these challenges, we investigate the potential of using neural networks to generate dialogues from input based on narrative actions, by analysing the story-level features which are able to provide dialogue styles with different aspects. Fortunately, the progressive development of neural network models [12,16,28] for dialogue generation provides the mechanism, which is a solid foundation to implement this theoretical approach through building this dialogue generator and incorporating the vectorised story-level features as dialogue styles.

2   Related Work

As we are working on generating stylistic dialogues for narratives using data-driven approaches, we concentrate on four relevant fields of prior work.

Narrative Representations In narrative research, narrative action is strongly related to natural language generation because it provides all main semantic elements in a story through abstract representation. Previous research proposed different formalisms for the representation of narrative actions. Pichotta and Mooney [18] developed a 5-tuple event representation of (v, e_s, e_o, e_p, p), where v is the verb, p is a preposition, and e_s, e_o, and e_p are nouns representing the subject, direction object, and prepositional object, respectively. Chambers and Jurafsky [6] proposed unsupervised induction of similar schemata called narrative event chains which are centred around the protagonist. These narrative chains are related to structured sequences of participants and events called scripts [25].

NLG in Interactive Narratives Storytelling through NLG is an important part in our work since the story should be represented in comprehensible natural language. Cavazza and Charles [5] generated natural dialogue through semantic templates using predetermined modalities of expression, which can be regarded as a rule-based strategy. Other systems using plan-based strategies can develop a plan composed of a series of (dialogue) actions, to reach goals in hierarchical tasks [2]. Bowden et al. [3] present algorithms for generating dialogues by converting a deep representation made by annotating and processing a story semantically [8,23].

Neural Networks and Dialogue Generation Progressively, with the development of these data-driven methods, several approaches to dialogue systems and dialogue generation are inclined to apply different types of neural networks, e.g. Recurrent Neural Networks (RNN). Currently, many works [16,28,32] build dialogue systems using data-driven methods to provide an automated generation process. The Seq2Seq [28] model, an encoder-decoder RNN model for generating sequences of unfixed length, is well-suited for generating utterances. The efficiency of this model is improved with the assistance of both Long Short Term Memory (LSTM) [11] and Attention Mechanism [1]. LSTM is an RNN model for solving long term dependencies problem (i.e. it is capable of using context to improve language coherence) and Attention improves the decoding precision by increasing weight for information of interest. Martin et al. [16] present two Seq2Seq models for generating event representations and sentences. In our work, we aim to find the relationship between narrative actions and utterances and use their relation to generate dialogue. Both of these can be regarded as word sequences and processed by Seq2Seq with LSTM and Attention rationally following certain aspects (e.g. grammar accuracy, content diversity, context coherence).

Style Representation Beyond the challenge of dialogue generation within a story context, the style of the dialogue (i.e. the ways to present the variant dialogues), is another key aspect this research focuses on. Walker et al. [29,30] focus on generalizing character style from films on the basis of qualitative statistical analysis. Li et al. [12] created two persona-based models by generating different word embeddings for different speakers and addressees. Some works on narrative generation have strategies based on characters point of view (PoV) [20] and relationship between characters [21], which could be means to specify styles for dialogue generation.

3   Proposed Approach

In this section, we introduce our proposed approach for dialogue generation in narratives.

A story consists in a sequence of narrative actions, which are the backbone of the whole storyline. As dialogue is a recognised manifestation of narrative discourse, it is therefore reasonable for us to hypothesise that the narrative actions representation can be leveraged for the purpose of generating narrative dialogue efficiently.

Currently, narrative generation is a more mature field than narrative dialogue generation. Martin et al. [16] postulate that narrative actions and natural language can be equated to each other. They define a story as a sequence of events. Each event is formalised as a 4-tuple event representation (s, v, o, m), where v is a verb, s and o are the subject and object of the verb, and m indexes a modifier, which can be any other grammatical parts in a sentence besides the above three. They use an event representation to generate the next event
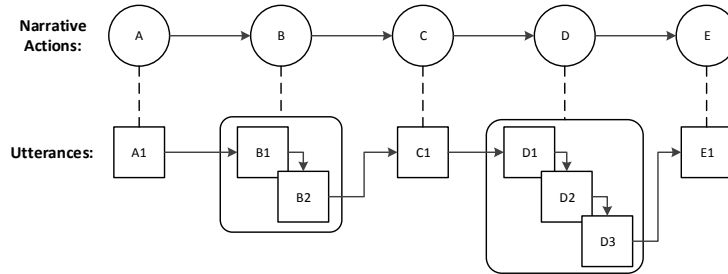
Fig. 1. The dialogue generation model.

representation which is followed by a conversion from this generated event representation to a natural language sentence. Inspired by this approach of story representation, we present our new approach for narrative dialogue generation.

3.1   Dialogue Generation Model

In Figure 1, we present a dialogue generation model based on a sequence of narrative actions, which are generated by a plan-based narrative system (similar to [20,21]). These pre-generated narrative actions are displayed as circles in Figure 1. Each narrative action can be represented as one or several utterances (displayed as small squares in Figure 1). For example, utterances *A1*, *C1*, and *E1* represent narrative actions *A*, *C*, and *E* respectively. An utterance pair, *B1* and *B2* represent the narrative action *B*. Also, *D1*, *D2*, and *D3* represent narrative action *D* jointly.

Each narrative action is supposed to be translated into one or multiple utterances as illustrated in Figure 1 shown through a pre-trained Seq2Seq based neural network. Along with narrative actions, other necessary elements of the dialogue such as who is the "doer" (subject) and who is the "doee" (object), often referred to as characters. Besides, where the dialogue happens (location) is also considered. All these elements can be extracted from story plot (e.g. film descriptions) according to [27] and are able to provide dialogue styles, described in the next section.

In this generation process, the context coherence between each turn within the dialogue is an important issue. Some open-domain dialogue systems [13,31] tackle coherence issues by discovering more relations between turn-pairs. In our model, the generated utterances should follow the direction of the narrative actions because the content of the story is represented by these utterances. Therefore, the coherence we consider here is between generated utterances and their corresponding narrative actions. An evaluation process must be undertaken after every utterance generation. Martin et al. [16] has shown that it is feasible to evaluate the correlation between the story events and the matching sentences using two metrics: Perplexity and BLEU [17].

Table 1. An example of how narrative actions can be translated into utterances.

Narrative Actions	Utterance turn pairs
(*Offer-To-Repay-Loan, A D C*)	D: Antonio, Do you acknowledge the contract? A: Yes, I do. D: Then you have to pay the money according to the contract and law.
(*Show-Contempt-For, S D C*)	S: Oh, wise young judge, I honour you! D: Shylock, they're offering you three times the money you lent. S: No, I made an oath in heaven. I urge you to deliver your verdict. I'm sticking to the contract.
(*Call-For-Judgement, S D C*)	D: A pound of this merchant's flesh is yours. The court awards it and the law authorizes it. S: What a righteous judge! D: But this contract doesn't give you any blood.
(*Release-From-Forfeit, S D C*)	S: So give me my money and let me go. D: The verdict is your penalty has been released.
(*Receive-Verdict-Of-Court, S D C*)	D: Are you satisfied, Shylock? S: I'm satisfied.

3.2 Example

Here we describe our dialogue generation model with an example displayed as Table 1. These five sequential narrative actions shown here are collected from [19], adapted from a well-known play of William Shakespeare, *Merchant of Venice*. The acronyms in narrative actions represent the names of the main characters and the locations of the scene. For the characters' name, *A*, *D*, and *S* indicates Antonio, the Duke of Venice, and Shylock respectively. The selected plots all occur at the court, which is indicated by *C*. As we can see in the table, a narrative action usually tends to be represented with more than one utterance, but normally no more than three utterances. If so, that narrative action should be split into some sub-narrative-actions, assuming the representation can be more precise with smaller granularity.

4   Stylistic Dialogue Modelling

In this section, we introduce our thoughts on adding story-level styles in dialogues during the dialogue generation process.

Many other works that address the issues of stylistic dialogues focus more prominently on linguistic features (e.g. sentence length, number of nouns/verbs) [30] or speaker profile features (e.g. age, gender, and occupation) [12]. We aim to rely on the potential that the story-level features are able to affect dialogue style.

Table 2. Dialogue styles in different genres of film scripts.

Genre	Dialogue Script Pieces
Romance	Ilsa: Hello, Sam. Sam: Hello, Miss Ilsa. I never expected to see you again.
	Violet: Good afternoon, Mr. Bailey. George: Hello, Violet. Hey, you look good. That's some dress you got on there.
Western	Ringo: Well hello, Sherm. Mcmasters: You wanted to talk? Ringo: Yeah, kinda, wanted to see if you'd join back up with us.
	Woman: Hello? Chigurh: Is Llewelyn there? Woman: Llewelyn?! No, he ain't.
Crime	Senator: Hi, Ace. Ace: Hello, Senator. Senator: Hey, I need a room. Need a room.
	Nick: Hello... Send him up. Nora: Who's that? Nick: Macaulay.

4.1 Genre Diversity

Table 2 shows film script pieces grouped in different genres, which are collected from Cornell Movie-Dialogs Corpus[1]. We select 6 pieces of script representing *greeting* in 3 different genres: romance, western, and crime. It is easy to observe that within each genre, dialogues are expressed through differing styles, often reflected in their linguistic aspects. For example, in the western genre, there exist informal spoken expressions (e.g. kinda, ain't). However, in our work, we are looking for a relation between story-level style and the genres. For example, we can see in the table that in western and crime genres, characters tend to come straight to the point for propelling the storyline when they start a dialogue with greetings (fast-paced). However, for the romance genre, characters are more likely to use witticisms to exaggerate the atmosphere of the scenes or make the audience more engaged (slow-paced). We aim to focus on the relations between genre and story-level style in dialogues, and analyse and quantify these relations.

4.2 Characters

For the purpose of characters, we put emphasis on the relationship between characters who participate in the dialogue exchange, which is different from previous literature that focus on the profile of interlocutors. In a story, a character normally plays a similar role across the narrative though will see an evolution in the relationship with other the characters. Porteous et al. [21] have presented a mechanism for narrative generation, which demonstrated that changes in the

---

[1] https://www.cs.cornell.edu/c̃ristian/Cornell_Movie-Dialogs_Corpus.html
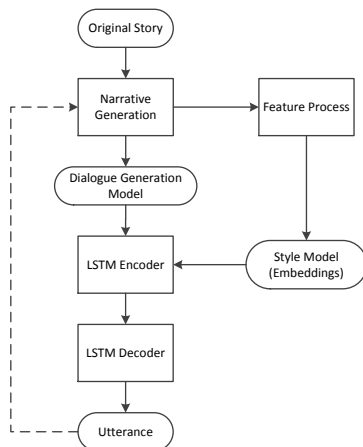
Fig. 2. Overview of the framework for our proposed approach.

relationships between characters in a social network can yield significant qualitative difference in narratives of some genres of drama (e.g. medical drama). This approach can also be used for providing styles in dialogue generation. We are supporting our generation of different utterances on the basis of the different character relationships defined as part of a same narrative action.

5   Neural Network Framework Design

Here, we describe our neural network framework as illustrated in Figure 2. We designed our dialogue generation framework based on Seq2Seq [28] with Attention mechanism [1], and on LSTM [11] for the encoder and decoder.

First, in the left principal stream, the original story will be preprocessed using a plan-based narrative engine, producing the sequence of relevant narrative actions through narrative generation. The generated narrative actions is used as the input into a LSTM encoder which is pre-trained using parallel narrative action/utterance corpora. After this stage, the learnt style embeddings will be incorporated with the LSTM encoder and the corresponding narrative actions, as the right branch stream shown. The LSTM decoder will receive the result from the encoder and generate the final utterance as the output.

Dialogue Generation There are two potential ways to generate dialogue with narrative actions. The first is dialogue generation within each narrative action. This way, each narrative action is the input to be used for the neural network, and the output is one or multiple utterances. The other is to consider the dialogue generation process over the whole pre-generated story. An initial utterance

is generated by the first narrative action, and the following utterances are iteratively generated based on the last generated utterance through evaluation. Either way, the story coherence is managed through the narrative generation process itself. If using the second one, the evaluation of whether the generated utterance is matching the narrative action is essential. Therefore, we choose the first mechanism to generate utterance through translating the narrative action.

Style Embeddings Style embeddings is an independent model which is incorporated as part of the main pipeline. Several prior dialogue works use embedding to improve the word-level embeddings using multiple dimensions vectors to represent word distribution and cluster semantically similar words into near spaces, which can be visualised as a two-dimensional space projected from the high-dimension vectors.

Here, it is achievable to create sentence-level embeddings for representing each style aspect we mentioned before (e.g. genre, character). Lin et al [15] use a two-dimensional matrix to represent the embeddings instead of a vector, with each row of the matrix corresponding to a different part of the sentence. For instance, the genre embedding can be created by promoting a supervised learning problem. Given a film script corpus, of which each utterance can be categorised into single or multiple genres, the genre embedding is being created by encoding utterances into matrices, since utterances with same genre(s) have similar encoding. This embedding can be viewed as a map containing the relations between genres and parts of the utterances. With this kind of embedding, we can add similar expressions through matching genres to our dialogue generation process.

6   Conclusion

In this paper, we have investigated the potential to integrate a data-driven approach within dialogue generation in narratives and have reasoned its feasibility theoretically and technically. We described a new model for dialogue generation based on the relations between utterances and narrative actions. We also introduced our style model in terms of story-level features, such as genre and character.

Our approach uses LSTM-based recurrent neural networks to generate dialogues, which will be derived from the given storyline. Although we have been working under the assumption of the complete knowledge of the narrative, within an *interactive* narrative paradigm, the narrative generation process must follows authorial control, which will also have to be managed for the purpose of dialogue generation. We may have to evaluate any potential remediation which will need to occur between the dialogue and the narrative generated interactively.

Our approach could be used in interactive computer games or interactive narratives. Normally, authors will create the necessary content components (e.g. characters, locations), where our approach can generate dialogues based upon those storylines and provide different styles learned from pre-existing corpora for providing more compelling and as such engaging dialogues.

References

1. Bahdanau, D., Cho, K., Bengio, Y.: Neural machine translation by jointly learning to align and translate. arXiv preprint arXiv:1409.0473 (2014)
2. Bohus, D., Rudnicky, A.I.: Ravenclaw: Dialog management using hierarchical task decomposition and an expectation agenda. In: Eighth European Conference on Speech Communication and Technology (2003)
3. Bowden, K.K., Lin, G.I., Reed, L.I., Tree, J.E.F., Walker, M.A.: M2d: Monolog to dialog generation for conversational story telling. In: International Conference on Interactive Digital Storytelling. pp. 12–24. Springer (2016)
4. Burgan, D.: Dialogue systems and dialogue management. Tech. rep., DST Group Edinburgh Edinburgh SA Australia (2016)
5. Cavazza, M., Charles, F.: Dialogue generation in character-based interactive storytelling. In: Proceedings of the First AAAI Conference on Artificial Intelligence and Interactive Digital Entertainment. pp. 21–26. AAAI Press (2005)
6. Chambers, N., Jurafsky, D.: Unsupervised learning of narrative event chains. Proceedings of ACL-08: HLT pp. 789–797 (2008)
7. Chen, H., Liu, X., Yin, D., Tang, J.: A survey on dialogue systems: Recent advances and new frontiers. ACM SIGKDD Explorations Newsletter **19**(2), 25–35 (2017)
8. Elson, D.K., McKeown, K.R.: A tool for deep semantic encoding of narrative texts. In: Proceedings of the ACL-IJCNLP 2009 Software Demonstrations. pp. 9–12. Association for Computational Linguistics (2009)
9. Gatt, A., Krahmer, E.: Survey of the state of the art in natural language generation: Core tasks, applications and evaluation. Journal of Artificial Intelligence Research **61**, 65–170 (2018)
10. Gervás, P., Díaz-Agudo, B., Peinado, F., Hervás, R.: Story plot generation based on CBR. In: Applications and Innovations in Intelligent Systems XII, pp. 33–46. Springer (2005)
11. Hochreiter, S., Schmidhuber, J.: Long short-term memory. Neural computation **9**(8), 1735–1780 (1997)
12. Li, J., Galley, M., Brockett, C., Spithourakis, G., Gao, J., Dolan, B.: A persona-based neural conversation model. In: Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers). vol. 1, pp. 994–1003 (2016)
13. Li, J., Monroe, W., Ritter, A., Jurafsky, D., Galley, M., Gao, J.: Deep reinforcement learning for dialogue generation. In: Proceedings of the 2016 Conference on Empirical Methods in Natural Language Processing. pp. 1192–1202 (2016)
14. Lin, G., Walker, M.: Stylistic variation in television dialogue for natural language generation. In: Proceedings of the Workshop on Stylistic Variation. pp. 85–93 (2017)
15. Lin, Z., Feng, M., Santos, C.N.D., Yu, M., Xiang, B., Zhou, B., Bengio, Y.: A structured self-attentive sentence embedding (2017)
16. Martin, L., Ammanabrolu, P., Wang, X., Hancock, W., Singh, S., Harrison, B., Riedl, M.: Event representations for automated story generation with deep neural nets. In: AAAI Conference on Artificial Intelligence (2018)
17. Papineni, K., Roukos, S., Ward, T., Zhu, W.J.: Bleu: a method for automatic evaluation of machine translation. In: Proceedings of the 40th annual meeting on association for computational linguistics. pp. 311–318. Association for Computational Linguistics (2002)

18. Pichotta, K., Mooney, R.J.: Learning statistical scripts with lstm recurrent neural networks. In: AAAI. pp. 2800–2806 (2016)
19. Porteous, J., Cavazza, M., Charles, F.: Applying planning to interactive storytelling: Narrative control using state constraints. ACM Transactions on Intelligent Systems and Technology (TIST) **1**(2), 10 (2010)
20. Porteous, J., Cavazza, M., Charles, F.: Narrative generation through characters' point of view. In: Proceedings of the 9th International Conference on Autonomous Agents and Multiagent Systems: volume 1-Volume 1. pp. 1297–1304. International Foundation for Autonomous Agents and Multiagent Systems (2010)
21. Porteous, J., Charles, F., Cavazza, M.: Networking: using character relationships for interactive narrative generation. In: Proceedings of the 2013 international conference on Autonomous agents and multi-agent systems. pp. 595–602. International Foundation for Autonomous Agents and Multiagent Systems (2013)
22. Riedl, M.O., Young, R.M.: Narrative planning: Balancing plot and character. Journal of Artificial Intelligence Research **39**, 217–268 (2010)
23. Rishes, E., Lukin, S.M., Elson, D.K., Walker, M.A.: Generating different story tellings from semantic representations of narrative. In: International Conference on Interactive Digital Storytelling. pp. 192–204. Springer (2013)
24. Rowe, J.P., Ha, E.Y., Lester, J.C.: Archetype-driven character dialogue generation for interactive narrative. In: Proceedings of the 8th international conference on Intelligent Virtual Agents. pp. 45–58. Springer-Verlag (2008)
25. Schank, R.C., Abelson, R.P.: Scripts, plans, goals, and understanding: an inquiry into human knowledge structures (1977)
26. Serban, I.V., Lowe, R., Henderson, P., Charlin, L., Pineau, J.: A survey of available corpora for building data-driven dialogue systems. arXiv preprint arXiv:1512.05742 (2015)
27. Tozzo, A., Jovanovic, D., Amer, M.: Neural event extraction from movies description. In: Proceedings of the First Workshop on Storytelling. pp. 60–66 (2018)
28. Vinyals, O., Le, Q.: A neural conversational model. arXiv preprint arXiv:1506.05869 (2015)
29. Walker, M., Lin, G., Sawyer, J.: An annotated corpus of film dialogue for learning and characterizing character style. In: Proceedings of the Eighth International Conference on Language Resources and Evaluation (LREC-2012) (2012)
30. Walker, M.A., Grant, R., Sawyer, J., Lin, G.I., Wardrip-Fruin, N., Buell, M.: Perceived or not perceived: Film character models for expressive nlg. In: International Conference on Interactive Digital Storytelling. pp. 109–121. Springer (2011)
31. Wen, T., Vandyke, D., Mrkšíc, N., Gašíc, M., Rojas-Barahona, L., Su, P., Ultes, S., Young, S.: A network-based end-to-end trainable task-oriented dialogue system. In: 15th Conference of the European Chapter of the Association for Computational Linguistics, EACL 2017-Proceedings of Conference. vol. 1, pp. 438–449 (2017)
32. Wen, T.H., Gasic, M., Mrkšić, N., Su, P.H., Vandyke, D., Young, S.: Semantically Conditioned LSTM-based Natural Language Generation for Spoken Dialogue Systems. In: Proceedings of the 2015 Conference on Empirical Methods in Natural Language Processing. pp. 1711–1721 (2015)