Text Rendering for Automatic Story Generation

Lin de Huybrecht, Nick Harley, and Geraint Wiggins

Advisor: Dr. Nick Harley
Promotor: Prof. Dr. Dr. Geraint Wiggins
Vrije Universiteit Brussel, Belgium
lin.de.huybrecht@vub.be

Most research in automatic story generation does not make a clear distinction between fabula generation (what happens in a story), sjuzhet generation (how the story is told) and text rendering (generating the natural language). Van der Sande [5] and Huybrecht [2] argued that these tasks should be separated and designed a three-tier (fabula/sjuzhet/text rendering) architecture for automatic story generation. Each subsystem fulfils a different task of automatic story generation. First, the Fabula Generating System [5] creates the content of the story (the *Fabula*) in the form of a Directed Acyclic Graph. Next, the Sjuzhet Generating System transforms the Fabula of the previous system into a linear sequence, the *Sjuzhet*. This structure specifies what needs to be told to the reader in the order it should be told, which can be different from the temporal and logical order in the Fabula. Finally, the Text Rendering System [2] creates the final Text Rendering from the Sjuzhet. This is the end-product of the three-tier system in human-readable format. All three subsystems share a Knowledge Base consisting of a Story Ontology, which restricts the story logic that should be used by the systems. The ontology can be replaced for the systems to create different types of stories.

This paper presents a text rendering system, called the Story Text Rendering System (STRS). The STRS generates natural language from a Sjuzhet that it gets as input from a Sjuzhet Generating System. For this, it uses a modular approach to natural language generation. First, the Microplanner gets as input a Knowledge Base consisting of the Sjuzhet and the Story Ontology. These are used to generate Sentence Plans, which make up the Text Plan. The Sentence Planner generates each Sentence Plan by first selecting a grammatical structure to convey the meaning of a Story Node of the Sjuzhet (Grammaticalisation Phase). Next, function words are inserted according to the chosen structure. Depending on the Lexical History, words are selected to be added to the Sentence Plan. During this process, linguistic variation is achieved by providing synonyms for content words (verbs and nouns), and by choosing between multiple grammatical structures for each Verb Class. These alternations are based on "English Verb Classes and Alternations: a Preliminary Investigation" by Levin [3]. When the Sentence Planner outputs a Sentence Plan, other mechanisms of linguistic variation are executed by the Microplanner, namely aggregation, ellipsis and referring expression generation (REG). However, REG also takes place during grammaticalisation, i.e. when choosing between definite ("the") or indefinite articles ("a") for Noun Phrases. After that, the Microplanner passes on the Text Plan to the SimpleNLG Surface Realiser. The Surface Realiser is responsible

for choosing the appropriate syntax (e.g. constituent order), agreement between noun and verb, inflection (e.g. of words), and orthography (e.g. punctuation). Finally, the SimpleNLG Surface Realiser [1] generates the Text Rendering from the Text Plans.

The output of the system was evaluated by an audience through a questionnaire, which was used to study the effect of sentence aggregation, verb synonyms, ellipsis and referring expression generation on enjoyability, semantics and grammaticality. The participants were asked to rate two versions of two stories on a five-point Likert scale ranging from "Strongly disagree" (1) to "Strongly agree" (5). More specifically, the participants received the following statements for each story version: (1) *I think the story is fun to read*, (2) *I think the story is understandable*, (3) *I think the story is grammatically correct*. These statements were inspired by the evaluation of the WeirdAnalogyMatic system [4]. After having read the two versions of a story, the participants were asked to rank the versions based on their overall preference.

The two stories are adapted versions of "The Hare and the Tortoise" and "The Little Red Hen". The system was designed and implemented using the former as an example. As a result, the second story was adapted more to suit the functionalities of the system and illustrate them as good as possible. The plot was altered significantly and the story was renamed to "The Hen and the Seed". For both stories, two versions were generated with the following settings: (1) all settings for linguistic variation enabled (i.e. verb synonyms, sentence aggregation, ellipsis and referring expression generation), (2) all settings disabled. The order of the stories and their versions was randomised, to avoid bias relating to the order in which the participants get to see the different stories and their versions. This experimental setup was inspired by the evaluation in the work of Viaene [6].

Statistical tests did not find any significant difference in the scores between the versions with the settings disabled and those with all the settings enabled, for all three statements. There were, however slight trends towards a significant difference for the story of The Hare and the Tortoise with regard to grammaticality and for both stories with regard to understandability. More people agreed with the grammatical correctness of the version of The Hare and the Tortoise with all settings disabled, than for the version with all settings enabled. This might due to the system's current functionalities being tailored to that story. Therefore, the sentences of that story are likely of a higher quality than those of the other story. Moreover, this could cause participants to notice that, for example the sentence aggregation is done rather naively, by sometimes combining two sentences with "and". It could the the case that, participants found the usage of "and" where a more suitable word could have been chosen, more disturbing, in a story that is in general of a higher quality. For the first story there seems to be a consensus that both versions are understandable, this is less so for the second story. This could be due to the large cuts that were made in the story lines of the original plot of The Little Red Hen. However, for both stories, the majority of the participants agreed that the stories are understandable.

References

1. Gatt, A., Reiter, E.: SimpleNLG: A realisation engine for practical applications. In: Proceedings of the 12th European Workshop on Natural Language Generation - ENLG '09. pp. 90–93. Association for Computational Linguistics. https://doi.org/10.3115/1610195.1610208
2. Huybrecht, L.d.: Text Rendering for Automatic Story Generation. Master's thesis, Vrije Universiteit Brussel (2022)
3. Levin, B.: English Verb Classes and Alternations: A Preliminary Investigation. University of Chicago Press
4. Oliveira, H.G.: WeirdAnalogyMatic: Experimenting with Analogy for Lyrics Transformation. In: Proceedings of the Eleventh International Conference on Computational Creativity. pp. 228–235. Association for Computational Creativity, http://computationalcreativity.net/iccc20/papers/047-iccc20.pdf
5. Van der Sande, E.: Fabula Generation with BCDI Characters and Narrative Tension. Master's thesis, Vrije Universiteit Brussel (2022)
6. Viaene, M.: Motivating Character Design in Automatically Generated Narratives. Master's thesis, Vrije Universiteit Brussel (2021)