



# Statistical Language Modelling for Automatic Story Generation

Marta Vicente, Cristina Barros and Elena Lloret

University of Alicante

{mvicente,cbarros,elloret}@dlsi.ua.es

**Abstract**—This abstract summarises the research work published in the *Journal of Intelligent & Fuzzy Systems*, vol. 34 (5), 2018. In this publication, we proposed and analysed an end-to-end Natural Language Generation approach to automatically create fiction stories using statistical language models, both in the macroplanning and surface realisation stages. The results obtained from the evaluation conducted showed that the combination of the aforementioned stages in an end-to-end approach was appropriate and had positive effects in the resulting generated text. Specifically, the use of language models allows the generation task to be more flexible, as far as the adaptation of the approach to different languages, domains and textual genres is concerned.

## I. INTRODUCTION

The generation of stories is a difficult task in which complex cognitive processes are involved. The process of creating a story requires both knowledge of the world and ability to perform tasks such as planning events or resolving situations [8]. Natural Language Generation (NLG), as a subarea of Natural Language Processing, can provide, in turn, useful mechanisms to create a story automatically. In order to generate a text, the responsibilities of NLG include, to determine “*what to say*” and “*how to say it*”. This dual purpose can be addressed as a pipeline of two stages: macroplanning and surface realisation. Macroplanning is in charge of determining the structure of the text to be produced, as well as deciding the content it will have. Surface realisation is focused on taking that content as input, and transforming it into human language, using the appropriate vocabulary, syntactic structures, etc.

In this research work, we studied the problem of whether a story could be automatically generated from scratch without having any clue about the characters, topics, or facts to describe, just having some previous information from other existing stories that could serve as an inspiration source.

To achieve this, a statistical perspective was adopted. Specifically, macroplanning was addressed using Positional Language Models (PLM) while, in the surface realisation stage, Factored Language Models (FLM) were employed. As output, the approach produced a set of sentences in the form of a story, inspired by the structure and linguistic elements of an existing one, but with a different realisation (different events, actions and vocabulary).

The preliminary results showed that statistical language models can definitely contribute to the development of more adaptive and flexible story generation systems, thus providing

mechanisms that can be extended to other kind of domains, languages and textual genres.

## II. BACKGROUND

Recent work in narrative and storytelling has been focused on regenerating stories from graphs of intentions [4] or approaching the task working with discourse and story planning simultaneously to differentiate levels of narrative [9]. Regarding the statistical models used in this research, on the one hand, PLMs have been successfully employed in some language-related areas such as summarisation [2] or information retrieval [5]. However, they have not been directly implemented within the generation framework. Also, in recent years, FLM have been employed for NLG, such in BAGEL [6] or in [7]. The main novelty of our approach is that human intervention in the NLG process is not necessary at any stage, so we avoid the hand-coding of the story constraints, and we increase the automation of the generation process.

## III. END-TO-END STORY GENERATION APPROACH

The NLG approach consists of a macroplanning and a surface realisation stage that are sequentially executed.

### A. Macroplanning.

In the NLG process, macroplanning is responsible for both selecting the content and providing the structure that articulates the output.

Different from the common bag-of-words perspective, PLMs are able to take into account the positions of words together with the number of their occurrences. On this basis, for each position  $i$  of a text, we computed a model  $P(w | i)$  by which every word  $w$  of the vocabulary gets a value for that location  $i$  conditioned by the distance to other appearances of that word within the text.

Using this model, we could obtain the distribution of the elements along the text, considering those with higher values the most relevant ones. We translated this information into a document plan.

### B. Surface realisation.

From the document plan provided by the macroplanning stage, the realisation of the final output was performed relying on FLMs. For this type of models, a word  $w$  is seen as a collection of  $K$  parallel factors, they can be any linguistic feature related to the word, from whose combination a statistical model is built.

We used over-generation and ranking, where a set of candidate sentences is generated and ranked according to the probabilities given by the model. We took into account a simple grammar and prioritised the information specified in the document plan to select the words belonging to each candidate.

#### IV. EXPERIMENTAL SCENARIO

Some experiments were conducted to analyse the performance of the approach. Each stage was empirically adjusted to achieve an optimal configuration and afterwards, a subset of the corpora was used to produce the document plans. The surface realisation module generated from them the final stories. In this research work, from each document plan only one story was generated.

Regarding the macroplanning, some decisions were made about the vocabulary and the design of the document plan itself. The vocabulary was formed by synsets in order to incorporate grammatical and semantic features to the process and, to conform the document plan, we produced one line per each sentence of the source text, containing synsets of verbs, nouns, adjectives and adverbs.

Once the document plans were created, the surface realisation was in charge of generating the final stories following the guidelines the plans had established while being also consistent with the grammar.

We worked with a collection of English children stories, including the Lobo and Matos corpus [3] and automatically gathered stories<sup>1</sup> from *Bedtime stories* and *Hans Christian Andersen: Fairy Tales and Stories*.

#### V. EVALUATION, RESULTS AND DISCUSSION

In order to evaluate how the document plan impacted in the generated text as well as the quality of the generated fiction story, we performed two tests.

On the one hand, we evaluated how the synsets and their distribution in the document plan were reflected in the resulting stories. To estimate the influence of the document plan in the generation of the new tales, we analysed the relation between both documents, measuring how the elements were shared.

On the other hand, a user evaluation was performed to analyse the general problems and errors of the approach. We used a point score scale, and three users manually read and analysed a set of stories in order to assess the system performance and provide feedback based on the detection of problems. They shed light on the limitations of our approach.

At a word level, the users highlighted the appropriate variety and richness of the vocabulary. Even though, for some examples they remarked that it would be adequate to use more synonyms, in order to prevent finding consecutive sentences repeating exactly the same terms.

At a sentence level, the evaluation revealed that, independently of the relation with their neighbours, sentences, in general, would become more meaningful once inflected. To

understand why some excerpts present an oddly shape, we should attend to the grammar beneath, on the one side, and to the semantic aspect, on the other.

Thirdly, at the level of discourse, adapted to our defined framework, the users were able to identify themes and guiding threads. They also noted the presence of characters as being involved in possible actions. They stated that the repetition of elements was essential to detect those underlying features. This reinforces our initial consideration regarding the relevance of a proper distribution of the elements through the text. Along with these comments, the users indicated that in some of the examples the absence of enough information impeded the assumption of any of those elements (themes, characters), but mostly this difficulty was related to the semantics of the generated tales.

#### VI. CONCLUSION

This paper proposed and evaluated a novel end-to-end NLG approach for generating fiction stories that integrates the macroplanning and the surface realisation stages within a statistical framework.

Evaluation showed that all the sentences generated by the approach contained elements from the document plan, 81% of the sentences having at least one of the highest rated elements. Furthermore, as there were not only one type of element within the document plan, on average, 83% of the content of each generated tale came from the document plan. Concerning the results of the user evaluation, we received positive feedback from 47% of the generated stories. We consider these results to be an indicator of the positive effects of combining the two stages (i.e. macroplanning and surface realisation).

Although the results are still preliminary, there are some issues which need to be improved and some interesting research lines opening new directions for future work.

#### ACKNOWLEDGMENT

Funded by the Generalitat Valenciana (ACIF/2016/501) and Spanish Government (TIN2015-65100-R, TIN2015-65136-C2-2-R).

#### REFERENCES

- [1] C. Baker, C.J. Fillmore, and J.B. Lowe, *The Berkeley FrameNet Project*, in: Procs. of the 36th Meeting of the ACL – Volume 1, 1998, 86–90.
- [2] S. Liu et al, *Positional Language Modelling for Extractive Broadcast News Speech Summarization*, in: INTERSPEECH, 2015, 2729–33.
- [3] P.V. Lobo, and D.M. De Matos, *Fairy Tale Corpus Organization Using Latent Semantic Mapping and an Item-to-Item Top-N Recommendation Algorithm*, in: LREC, 2010.
- [4] S.M. Lukin, L.I. Reed and M.A. Walker, *Generating Sentence Planning Variations for Story Telling*, in: 16th SIGdial, 2015, 188.
- [5] Y. Lv, Yuanhua, and C. Zhai, *Positional Language Models for Information Retrieval*, in: Procs. of the 32nd ACM SIGIR, New York, 2009, 299–306.
- [6] F. Mairesse and S. Young, *Stochastic Language Generation in Dialogue Using Factored Language Models*, Computational Linguistics, 40(4), 2014, 763–99.
- [7] E.M. Novais, and I. Paraboni, *Portuguese Text Generation Using Factored Language Models*, Journal of the Brazilian Computer Society 19(2), 2012, 135–46.
- [8] S.R. Turner, *Minstrel: A Computer Model of Creativity and Storytelling*, Ph.D. Dissertation. University of California at Los Angeles, 1993.
- [9] D.R. Winer and R.M. Young, *Discourse-Driven Narrative Generation with Bipartite Planning*, in: The 9th INLG, 2016, 11–20.

<sup>1</sup><https://freestoriesforkids.com/>, <http://hca.gilead.org.il/>