THEaiTRE: Artificial Intelligence to Write a Theatre Play

Rudolf Rosa[μ], **Ondřej Dušek**[μ], **Tom Kocmi**[μ], **David Mareček**[μ], **Tomáš Musil**[μ], **Patrícia Schmidtová**[μ], **Dominik Jurko**[μ], **Ondřej Bojar**[μ], **Daniel Hrbek**[σδ], **David Košťák**[σ], **Martina Kinská**[σ], **Josef Doležal**[δ] and **Klára Vosecká**[δ]

[μ]Charles University, Faculty of Mathematics and Physics, Institute of Formal and Applied Linguistics

[σ]The Švanda Theatre in Smíchov, Prague

[δ]The Academy of Performing Arts in Prague, Theatre Faculty (DAMU)

uru@ufal.mff.cuni.cz, hrbek@svandovodivadlo.cz

Abstract

We present THEaiTRE, a starting research project aimed at automatic generation of theatre play scripts. This paper reviews related work and drafts an approach we intend to follow. We plan to adopt generative neural language models and hierarchical generation approaches, supported by summarization and machine translation methods, and complemented with a human-in-the-loop approach.

1 Introduction

We introduce the THEaiTRE project,[1] which aims to produce and stage the first computer-generated theatre play. This play will be presented on the occasion of the 100th anniversary of Karel Čapek's play *R.U.R.* [Čapek, 1920], for which the word "robot" was invented by Čapek.

The project, currently in its early stages, is at the intersection of artificial intelligence research and theatre studies. The core of our approach is to use state-of-the-art deep neural models trained and fine-tuned on theatre play data. However, our team includes both experts on natural language processing and theatre experts, and our solution will be based on research and exchange of experience from both fields.

In this paper, we first review related previous works (Section 2) and data resources available to us (Section 3). We then draft the approaches we are following and intending to follow in the project (Section 4) and present the project timeline (Section 5).

2 Related Work

2.1 Narrative Natural Language Generation

While we are not aware of any generation systems specifically aimed at theatre play generation, research in story/narrative generation has been quite active in the past years, with computer-aided systems allowing various degrees of automation and different abilities in learning from data [Kybartas and Bidarra, 2017; Riedl, 2018]. Since recurrent neural networks (RNN) were applied for text generation [Bahdanau *et al.*, 2015; Sutskever *et al.*, 2014], research in story generation has mostly focused on fully data-driven, fully automated approaches. As plain RNNs were found unsuitable for producing longer, coherent texts [Wiseman *et al.*, 2017], multiple improvements have been proposed.

The first line of work focuses on providing a higher-level semantic representation to the networks and conditioning the generation on it. Martin *et al.* [2018] and Ammanabrolu *et al.* [2019; 2020] use an event-based representation, where an event roughly represents a clause (predicate, subject, direct and indirect object). The model generates the story at the event level and subsequently realizes the individual events to surface sentences. Tu *et al.* [2019] take a similar approach, using frame semantics and also conditioning sentence generation on other information, such as sentiment.

Other works focus on explicit entity modelling across the generated story, e.g., Clark *et al.* [2018]. Here, each entity has its own distributed representation (embedding), which is updated on each mention of the entity in the story.

Multiple authors attempt to increase long-term coherence by hierarchical story generation. Fan *et al.* [2018] generate first a short prompt/tagline, then use it to condition the full story generation. Yao *et al.* [2019] take a similar approach, using a "storyline" – a list of entities and items to be introduced in the story in the given order. Fan *et al.* [2019] then combine the hierarchical generation with explicit entity modelling. Their system generates outputs using anonymized but tracked entities, which are subsequently lexicalized in the context of the story by generating referring expressions.

Several works experiment with altering the base RNN architecture: Wang and Wan [2019] use a modified Transformer architecture [Vaswani *et al.*, 2017], which is trained as a conditional variational autoencoder. Tambwekar *et al.* [2019] utilize reinforcement learning with automatically induced rewards to train their event-based model. Ammanabrolu *et al.* [2019; 2020] extend this work by experimenting with various sentence realization techniques, including retrieval from database and post-editing.

Latest works use massive pretrained language models based on the Transformer architecture, such as GPT-2 [Rad-

In: A. Jorge, R. Campos, A. Jatowt, A. Aizawa (eds.): Proceedings of the first AI4Narratives Workshop, Yokohama, Japan, January 2021, published at http://ceur-ws.org

[1] https://www.theaitre.com/

ford *et al.*, 2019], for generation. See *et al.* [2019] use GPT-2 directly and show that it is superior to plain RNNs. Mao *et al.* [2019] apply GPT-2 fine-tuned for both story generation and common-sense reasoning to improve coherence.

While research in this area has progressed considerably, most experiments have been performed on rather short and simple stories, such as the ROCStories corpus [Mostafazadeh *et al.*, 2016]. Many works focus on limited tasks, such as single-sentence continuation generation [Tu *et al.*, 2019]. The state-of-the-art results still cannot match human performance, producing repetitive and dull outputs [See *et al.*, 2019].

2.2 Dramatic Analysis

For our needs, we are mostly interested in classifications and abstractions over theatre play scripts or their parts. In the field of theatre studies, there is a vast amount of research on the structure and interpretation of theatre plays. Unfortunately, the results of such research are not made available in forms and formats that would easily allow us to use these as data and annotations in machine learning approaches.

The Thirty-Six Dramatic Situations by Polti [1921][2] is a classic work, in which the author presented a supposedly ultimate list of all categories of possible dramatic situations that can occur in a theatre play (e.g. "adultery" or "conflict with a god"), further subclassified into 323 situational possibilities.

Although not directly related to theatre plays, the work of [Propp, 1968] is also essential. Propp analyzed Russian folk tales and identified 31 *functions*, similar to Polti's situations but somewhat more down-to-earth (e.g. "villainy" or "wedding"), as well as 7 abstract character types (e.g. "villain" or "hero") and other abstractions.

Polti's and Propp's categorizations are sometimes used in analyzing and generating narratives, although typically not in drama. The works closest to our focus is probably that of [Gervás *et al.*, 2016] or Lombardo *et al.* [2018], who devised an ontologies of abstractions for annotating scripts, based on both of the mentioned works, as well as on more recent plot categorization studies [Booker, 2004; Tobias, 2011].

There are also works producing drama analyses in the form of networks, capturing various relations between the characters in the play [Moretti, 2014; Horstmann, 2019; Fischer *et al.*, 2019].

2.3 Computer-Generated Art

There already is a range of partially or fully artificially generated works of art – e.g. a short sci-fi movie with an LSTM-generated and human-post-edited script [Benjamin *et al.*, 2016], a musical based on suggestions from several automated tools [Colton *et al.*, 2016], a human-picked collection of computer generated poems [Materna, 2016], or a theatre play written with the help of a next word suggestion tool [Helper, 2018]. While this demonstrates the technical possibility of such an approach, the mixed reception of the outcomes shows that the employed technologies are not (yet?) on par with humans [See *et al.*, 2019]. We thus believe a more specialized and complex approach is needed here.

[2] https://en.wikipedia.org/wiki/The_Thirty-Six_Dramatic_Situations

3 Data Resources

Theatre play scripts are not easily available for our purposes. As no reasonable corpus is available, we have to create one ourselves. The corpus will contain Czech and English theatre play scripts and synopses (plot summaries), and will be used to train and fine-tune our systems, described in following sections. We are also collecting film and TV series scripts, which are easier to obtain in large quantities, although they are not a perfect match for our setting. Unfortunately, due to copyright reasons, we will not be able to release the full corpus.

In most cases, scripts cannot be downloaded for free, and for most scripts it seems that they are only available in print or scanned. Even electronically available scripts come in various formats and there seems to be no technical standards in this respect. For our project, we need to devise a common representation format, and automatically or semi-automatically convert and normalize the data into the format, marking character names, lines, scenic notes, scene settings, etc. [Croce *et al.*, 2019]. Also the scripts and synopses need to be paired together. At the moment, we only have collected and partially converted several hundreds of documents.

4 Planned Approach

As a theatre play script is a highly structured and complex piece of text, we plan to take a hierarchical approach composed of several steps to generating the full script, also employing human inputs in the process. The overall idea is to start from a brief description of the play, gradually expanding it into more detailed act and scene synopses, and finally generating the individual scene dialogues. We currently envision using generative neural models for the final step (Section 4.1), conditioned by prompts generated by hierarchical generation approaches (Section 4.2).

4.1 Applying Neural Language Models

Large neural *language models* (LMs), such as GPT-2 [Radford *et al.*, 2019; see Section 2.1], are able to generate believable texts in certain domains (e.g. news articles). This is not the case for the domain of theatre plays. The original GPT-2 must have had a number of plays (or movie scripts) in the training data, which is evident when it is presented with a suitable starting prompt. It can produce a text that follows the formal structure and has some level of content coherence. However, the basic attributes of a dramatic situation are missing: there is no plot, and the scene is not moving towards a conclusion. Other problems include having new characters appear randomly in the middle of the scene or falling into a state of repeating the same sentence forever.

Our basic workflow would be to seed an LM with a prompt which is the beginning of a dramatic situation. The LM would generate the rest of the whole dialogue. We plan to finetune the LM to theatre plays to see how far this approach can go. Then we plan to restrict the generation by enforcing that only certain predetermined characters speak, possibly in a pregenerated order. This can be achieved by stopping the generation at the end of a character's line, adding the name of the next desired character and then resuming the generation process.

To make the characters more internally consistent and different from each other at the same time, we plan to devise individual LMs specialized to specific character types, based on a clustering of the characters across plays. The part of each character would then be generated by a different LM; i.e., the script would consist of several LMs "talking" to each other.

4.2 Hierarchical Generation

We also plan to extend our experiments with hierarchical generation from large pretrained LMs. We will use an approach similar to Fan *et al.* [2018] and Yao *et al.* [2019] (see Section 2.1): starting with generating a title or a prompt for the story, then generating a textual synopsis. The generation of the play from the synopsis will follow as a novel step, not present in previous works. We are considering multiple options of what to choose as the synopsis representation: the play background/setting from play databases, more detailed synopses from fan websites, or scenic remarks extracted from texts of plays themselves. Ultimately, the choice will be made based on data availability. The setup will also include generating "play metadata", such as the main theme, list of characters, narrative type, etc.

The final step will use a similar approach as the base LM generation (see Section 4.1). We also plan on using explicit embeddings for individual characters in the play and using explicit entity tracking/coreference [Clark *et al.*, 2018; Fan *et al.*, 2019]. Since the available automatic coreference tools [e.g., Clark and Manning, 2016; Lee *et al.*, 2017] are typically not trained for processing dialogic texts, they may require adaptation.

4.3 Data Synthesis through Summarization

The hierarchical generation approach relies on data that contain information of various granularity, as described in Section 4.2. However, most of the available data contain only the title and the script of the play, missing other invaluable information. In our project, we intend to synthesize the missing data; synthetic data are frequently used in various tasks, such as machine translation [Bojar and Tamchyna, 2011; Sennrich *et al.*, 2016].

We can generate synthetic data with the use of the classical task of text summarization; abstractive summarization in particular [Rush *et al.*, 2015]. The main idea is to take a long document and summarize it into a few sentences, then take these synthetic data and use them for training the generative models in the hierarchical approach. With various summarizing models, we can first abstract the whole script of a theatre play into a detailed synopsis, then the detailed synopsis into a short plot synopsis, and eventually the short synopsis into the play title. With these summarizing models, we can fill the gaps in our datasets, so that the hierarchical generation models can be trained on all theatre scripts available to us, even if they lack some or all higher-level summaries.

We plan to train the Transformer model [Vaswani *et al.*, 2017] for the summarization tasks. As we expect the amount of available training play-summary pairs to be scarce, we will pretrain our models on other summarization tasks, such as news abstract generation for which plenty of parallel data is available [Straka *et al.*, 2018], followed by fine-tuning the pretrained models on our in-domain theatre data.

Due to the specific nature of the genre, where a lot of what is meant is not explicitly said by any of the characters, we know that the summarization may be difficult or impossible to do, and this component thus cannot be entirely relied on.

4.4 Machine Translation

We plan on using machine translation (MT) for two purposes: (1) Since we have limited amounts of training data scattered across both English and Czech, we need the generation to take advantage of data in both languages. Therefore, we plan to generate new training data by translating either Czech texts to English or vice versa. (2) We would like the same resulting generated play to be available instantly in both languages. Therefore, we plan to generate it in one of the languages and use MT to bring the result over to the other language.

For both applications, we are going to use our in-house state-of-the-art Czech-English model [Popel, 2018]. However, theatre play scripts are a specific domain of data for which our MT models were not specifically trained. To tackle this problem, we will finetune [Miceli Barone *et al.*, 2017] the general MT models on theatre parallel data, possibly also applying automated heuristical pre-processing and/or post-editing [Rosa *et al.*, 2012].

4.5 Human in the Loop

To ensure a satisfactory result, we intend to complement the automated generation with interventions from theatre professionals, using a *human-in-the-loop* approach.

We currently envision using the automated system to generate texts and the human to choose parts of the output to use in the play. This could be done e.g. in an iterative interactive way, where the system generates several options for a line of the script, the human picks one of the options to add to the script, the system generates continuation options, etc.

Moreover, only the dialogues of the characters will be fully automatically generated. The subsequent realization and performance of the play will be in the hands of theatre professionals, who will analyze and interpret the script, devise stage directions, rehearse the play, design the scene, and finally perform the play for a live audience, all of which will further shape the perception of the play by the spectators.

5 Conclusion and Future Work

After some preliminary work, the project started in April 2020. The first automatically generated THEaiTRE play will be premiered in January 2021, at the occasion of the 100th anniversary of the premiere of the play *R.U.R.* [Čapek, 1920]. A premiere of a second play, generated from an improved version of our system, is planned for 2022.

The project can be tracked at https://theaitre.com

Acknowledgments

The THEaiTRE project is supported by the Technology Agency of the Czech Republic grant TL03000348. and partially supported by SVV project number 260 575.

References

[Ammanabrolu *et al.*, 2019] Prithviraj Ammanabrolu, Ethan Tien, Wesley Cheung, Zhaochen Luo, William Ma, Lara Martin, and Mark Riedl. Guided Neural Language Generation for Automated Storytelling. In *Proceedings of the Second Workshop on Storytelling*, pages 46–55, Florence, Italy, August 2019. Association for Computational Linguistics.

[Ammanabrolu *et al.*, 2020] Prithviraj Ammanabrolu, Ethan Tien, Wesley Cheung, Zhaochen Luo, William Ma, Lara J. Martin, and Mark O. Riedl. Story Realization: Expanding Plot Events into Sentences. In *AAAI*, New York, NY, USA, February 2020. arXiv: 1909.03480.

[Bahdanau *et al.*, 2015] Dzmitry Bahdanau, Kyunghyun Cho, and Yoshua Bengio. Neural Machine Translation by Jointly Learning to Align and Translate. In *3rd International Conference on Learning Representations (ICLR2015)*, San Diego, CA, USA, May 2015. arXiv: 1409.0473.

[Benjamin *et al.*, 2016] AI Benjamin, Oscar Sharp, and Ross Goodwin. Sunspring, a sci-fi short film starring Thomas Middleditch, 2016. https://www.youtube.com/watch?v=LY7x2Ihqjmc.

[Bojar and Tamchyna, 2011] Ondřej Bojar and Aleš Tamchyna. Improving Translation Model by Monolingual Data. In *Proceedings of WMT*, pages 330–336, Edinburgh, Scotland, 2011. ACL.

[Booker, 2004] Christopher Booker. *The seven basic plots: Why we tell stories*. A&C Black, 2004.

[Čapek, 1920] Karel Čapek. *R.U.R. (Rossum's Universal Robots)*. Aventinum, Ot. Štorch-Marien, Praha, 1920.

[Clark and Manning, 2016] Kevin Clark and Christopher D. Manning. Deep Reinforcement Learning for Mention-Ranking Coreference Models. In *Proceedings of EMNLP*, pages 2256–2262, Austin, Texas, November 2016. Association for Computational Linguistics.

[Clark *et al.*, 2018] Elizabeth Clark, Yangfeng Ji, and Noah A. Smith. Neural Text Generation in Stories Using Entity Representations as Context. In *Proceedings of NAACL-HLT*, pages 2250–2260, New Orleans, Louisiana, June 2018. Association for Computational Linguistics.

[Colton *et al.*, 2016] Simon Colton, Maria Teresa Llano, Rose Hepworth, John Charnley, Catherine V. Gale, Archie Baron, François Pachet, Pierre Roy, Pablo Gervás, Nick Collins, Bob Sturm, Tillman Weyde, Daniel Wolff, and James Robert Lloyd. The Beyond the Fence musical and Computer Says Show documentary. In *Proceedings of the Seventh International Conference on Computational Creativity*, 2016.

[Croce *et al.*, 2019] Danilo Croce, Roberto Basili, Vincenzo Lombardo, and Eleonora Ceccaldi. Automatic recognition of narrative drama units: A structured learning approach. In *Text2Story@ECIR*, 2019.

[Fan *et al.*, 2018] Angela Fan, Mike Lewis, and Yann Dauphin. Hierarchical Neural Story Generation. In *Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, New Orleans, LA, USA, June 2018. arXiv: 1805.04833.

[Fan *et al.*, 2019] Angela Fan, Mike Lewis, and Yann Dauphin. Strategies for Structuring Story Generation. In *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics*, pages 2650–2660, Florence, Italy, July 2019. Association for Computational Linguistics.

[Fischer *et al.*, 2019] Frank Fischer, Ingo Börner, Mathias Göbel, Angelika Hechtl, Christopher Kittel, Carsten Milling, and Peer Trilcke. Programmable corpora. die digitale literaturwissenschaft zwischen forschung und infrastruktur am beispiel von dracor. In *DHd 2019 Digital Humanities: multimedial & multimodal. Konferenzabstracts*, pages 194–197, Frankfurt am Main, March 2019. Zenodo. https://github.com/dracor-org/gerdracor.

[Gervás *et al.*, 2016] Pablo Gervás, Raquel Hervás, Carlos León, and Catherine V Gale. Annotating musical theatre plots on narrative structure and emotional content. In *7th Workshop on Computational Models of Narrative (CMN 2016)*. Schloss Dagstuhl-Leibniz-Zentrum fuer Informatik, 2016.

[Helper, 2018] Roslyn Helper. Lifestyle of the Richard and family, 2018. https://www.roslynhelper.com/lifestyle-of-the-richard-and-family.

[Horstmann, 2019] Jan Horstmann. DraCor: Drama corpora project. In *forTEXT. Literatur digital erforschen*, 2019. https://dracor.org/.

[Kybartas and Bidarra, 2017] B. Kybartas and R. Bidarra. A Survey on Story Generation Techniques for Authoring Computational Narratives. *IEEE Transactions on Computational Intelligence and AI in Games*, 9(3):239–253, September 2017.

[Lee *et al.*, 2017] Kenton Lee, Luheng He, Mike Lewis, and Luke Zettlemoyer. End-to-end Neural Coreference Resolution. In *Proceedings of the 2017 Conference on Empirical Methods in Natural Language Processing*, pages 188–197, Copenhagen, Denmark, September 2017. Association for Computational Linguistics.

[Lombardo *et al.*, 2018] Vincenzo Lombardo, Rossana Damiano, and Antonio Pizzo. Drammar: A comprehensive ontological resource on drama. In *International Semantic Web Conference*, pages 103–118. Springer, 2018.

[Mao *et al.*, 2019] Huanru Henry Mao, Bodhisattwa Prasad Majumder, Julian McAuley, and Garrison Cottrell. Improving Neural Story Generation by Targeted Common Sense Grounding. In *Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing (EMNLP-IJCNLP)*, pages 5987–5992, Hong Kong, China, November 2019. Association for Computational Linguistics.

[Martin *et al.*, 2018] Lara J Martin, Prithviraj Ammanabrolu, Xinyu Wang, William Hancock, Shruti Singh, Brent Harrison, and Mark O Riedl. Event Representations for Automated Story Generation with Deep Neural Nets. In *AAAI*, New Orleans, LA, USA, 2018.

[Materna, 2016] Jiří Materna. *Poezie umělého světa*. Backstage Books, 2016.

[Miceli Barone *et al.*, 2017] Antonio Valerio Miceli Barone, Barry Haddow, Ulrich Germann, and Rico Sennrich. Regularization techniques for fine-tuning in neural machine translation. In *Proceedings of the 2017 Conference on Empirical Methods in Natural Language Processing*, pages 1489–1494, Copenhagen, Denmark, September 2017. Association for Computational Linguistics.

[Moretti, 2014] Franco Moretti. "Operationalizing": or, the function of measurement in modern literary theory. Pamphlet 6, Stanford Literary Lab, 2014.

[Mostafazadeh *et al.*, 2016] Nasrin Mostafazadeh, Lucy Vanderwende, Wen-tau Yih, Pushmeet Kohli, and James Allen. Story Cloze Evaluator: Vector Space Representation Evaluation by Predicting What Happens Next. In *Proceedings of the 1st Workshop on Evaluating Vector-Space Representations for NLP*, pages 24–29, Berlin, Germany, August 2016. Association for Computational Linguistics.

[Polti, 1921] Georges Polti. *The thirty-six dramatic situations*. JK Reeve, 1921.

[Popel, 2018] Martin Popel. Cuni transformer neural mt system for wmt18. In *Proceedings of the Third Conference on Machine Translation*, pages 486–491, Belgium, Brussels, October 2018. Association for Computational Linguistics.

[Propp, 1968] Vladimir Propp. Morphology of the folktale, trans. *Louis Wagner, 2d. ed.(1928*, 1968.

[Radford *et al.*, 2019] Alec Radford, Jeffrey Wu, Rewon Child, David Luan, Dario Amodei, and Ilya Sutskever. Language Models are Unsupervised Multitask Learners. Technical report, OpenAI, February 2019.

[Riedl, 2018] Mark Riedl. Computational Narrative Intelligence: Past, Present, and Future. *Medium*, February 2018.

[Rosa *et al.*, 2012] Rudolf Rosa, David Mareček, and Ondřej Dušek. Depfix: A system for automatic correction of Czech MT outputs. In *Proceedings of WMT*, pages 362–368, 2012.

[Rush *et al.*, 2015] Alexander M. Rush, Sumit Chopra, and Jason Weston. A neural attention model for abstractive sentence summarization. In *Proceedings of the 2015 Conference on Empirical Methods in Natural Language Processing*, pages 379–389, Lisbon, Portugal, September 2015. Association for Computational Linguistics.

[See *et al.*, 2019] Abigail See, Aneesh Pappu, Rohun Saxena, Akhila Yerukola, and Christopher D. Manning. Do Massively Pretrained Language Models Make Better Storytellers? In *CoNLL*, Hong Kong, November 2019. arXiv: 1909.10705.

[Sennrich *et al.*, 2016] Rico Sennrich, Barry Haddow, and Alexandra Birch. Improving neural machine translation models with monolingual data. In *Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 86–96, Berlin, Germany, August 2016. Association for Computational Linguistics.

[Straka *et al.*, 2018] Milan Straka, Nikita Mediankin, Tom Kocmi, Zdeněk Žabokrtský, Vojtěch Hudeček, and Jan Hajic. SumeCzech: Large Czech News-Based Summarization Dataset. In *Proceedings of the Eleventh International Conference on Language Resources and Evaluation (LREC 2018)*, Miyazaki, Japan, May 7-12, 2018 2018. European Language Resources Association (ELRA).

[Sutskever *et al.*, 2014] Ilya Sutskever, Oriol Vinyals, and Quoc VV Le. Sequence to sequence learning with neural networks. In *Advances in Neural Information Processing Systems*, pages 3104–3112, 2014. arXiv:1409.3215.

[Tambwekar *et al.*, 2019] Pradyumna Tambwekar, Murtaza Dhuliawala, Animesh Mehta, Lara J. Martin, Brent Harrison, and Mark O. Riedl. Controllable Neural Story Plot Generation via Reinforcement Learning. In *2019 International Joint Conference on Artificial Intelligence*, Macau, August 2019. arXiv: 1809.10736.

[Tobias, 2011] Ronald B Tobias. *20 MASTER Plots: and how to build them*. Penguin, 2011.

[Tu *et al.*, 2019] Lifu Tu, Xiaoan Ding, Dong Yu, and Kevin Gimpel. Generating Diverse Story Continuations with Controllable Semantics. In *3rd Workshop on Neural Generation and Translation (WNGT 2019)*, Hong Kong, November 2019. arXiv: 1909.13434.

[Vaswani *et al.*, 2017] Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N. Gomez, Lukasz Kaiser, and Illia Polosukhin. Attention Is All You Need. In *31st Conference on Neural Information Processing Systems (NIPS)*, Long Beach, CA, USA, December 2017. arXiv: 1706.03762.

[Wang and Wan, 2019] Tianming Wang and Xiaojun Wan. T-CVAE: Transformer-Based Conditioned Variational Autoencoder for Story Completion. In *Proceedings of the Twenty-Eighth International Joint Conference on Artificial Intelligence*, pages 5233–5239, Macao, China, August 2019. International Joint Conferences on Artificial Intelligence Organization.

[Wiseman *et al.*, 2017] Sam Wiseman, Stuart M. Shieber, and Alexander M. Rush. Challenges in Data-to-Document Generation. In *Proceedings of the 2017 Conference on Empirical Methods in Natural Language Processing*, pages 2243–2253, Copenhagen, Denmark, September 2017. arXiv: 1707.08052.

[Yao *et al.*, 2019] Lili Yao, Nanyun Peng, Ralph Weischedel, Kevin Knight, Dongyan Zhao, and Rui Yan. Plan-And-Write: Towards Better Automatic Storytelling. In *AAAI*, Honolulu, HI, USA, January 2019. arXiv: 1811.05701.