

This thesis explores different ways of improving theatre play script generation. To generate a theatre play script, we fine-tune the GPT-2 medium language model on a mixture of theatre plays, movies, and TV show scripts, training it to continue a human-written script start. As plain language-model generation is not sufficient for generating a coherent full-length theatre play, we propose a two step hierarchical generation approach: first generating a plot summary and then conditioning the script generation on the plot summary. We train the hierarchical model on a dataset which we created by aligning plot summaries to script scenes using dynamic programming. We also propose additional approaches to make the generated text more consistent (with respect to characters and contradictions). We compare our models to a strong vanilla GPT-2 XL baseline and achieve comparable performance with our models being smaller and faster. Moreover, our hierarchical generation approach allows generating a coherent full-length theatre play script under limited human supervision, which has also been verified in practice by theatre professionals within the THEaiTRE project.