Literature Review for Final Project Milestone Two

Base Description Generation

COMET-ATOMIC: On Symbolic and Neural Commonsense Knowledge Graphs

- **Year: 2021**

- **Summary:** This article introduces the COMET training framework that enables language models to learn common sense knowledge from pre-constructed knowledge graphs.
- **Thoughts:** We pay particular attention to the factual knowledge encoding, i.e., how to use a COMET language model as a knowledge base for item description.

Language Models as Knowledge Bases?

- **Year: 2019**

- **Summary:** This work explores the possibility of using large pre-trained language models as a knowledge base. The authors experiment with querying the language models with natural language.
- **Thoughts:** It confirms that we can use large language models as simple knowledge bases to help generate base descriptions of various objects, which are primarily factual.

Style Transfer

Deep Learning for Text Style Transfer: A Survey

- **Year: 2021**

- **Summary:** This article systematically reviews the Text Style Transfer (TST) task. The survey includes commonly used datasets, SOTA models, evaluation metrics, and benchmarks.
- **Thoughts:** The article helps us formalize our game-description-rendering problem as a TST task based on a (pseudo) parallel dataset and guides us to search for sequence-to-sequence and generative models for implementations. It also provides us with ideas about evaluating the quality of style transferring.

Dear Sir or Madam, May I Introduce the GYAFC Dataset: Corpus, Benchmarks and Metrics for Formality Style Transfer

- **Year: 2018**

- **Summary:** This work introduces a collection of (formality) style transfer datasets and shows that we can adopt standard sequence-to-sequence models for TST tasks.
- **Thoughts:** This gives us a preliminary research result that Seq2Seq models like BART can be used for our purpose.

Dataset

Justifying Recommendations using Distantly-Labeled Reviews and Fine-Grained Aspects

- **Year: 2018**

- **Summary:** This dataset includes a comprehensive set of items' descriptions posted on the Amazon platform.
- **Thoughts:** We use the metadata of the items to guide a language model to generate realistic item descriptions.

Learning to Speak and Act in a Fantasy Text Adventure Game

- **Year: 2019**

- **Summary:** The LIGHT dataset contains a set of fantasy adventure game items metadata.
- **Thoughts:** We use the LIGHT dataset to develop baseline models that render base item descriptions to game objects.