

A Experiment Details

Computational Budget. Long Story Short uses GPT-3 (175B parameters) via OpenAI API as the backbone. An average prompt to summarize a video segment processes ~ 3000 tokens, while a QA prompt usually takes ~ 4000 tokens. For CLIPCheck, we extract CLIP features and compute the cosine similarity using a single NVIDIA A6000 GPU: it takes 0.5 hours to process video frames for the MovieQA validation split.

Hyperparameters. All hyperparameters are pre-defined by analyzing a single training sample. For narrative search, we use sentence similarity threshold $\alpha \geq 0.5$ to find plot pieces when GPT-3 does not output a single index. We use the binary entropy threshold $E' \geq 0.4$ in CLIPCheck. We run each experiment only once, as our method is deterministic and is not susceptible to randomness in initialization.

Video Segmentation Scheme. There are predefined segment boundary annotations for all datasets we utilize in this paper. Also, all plot pieces have aligned clip segments in turn since we perform summarization on each clip segmented with the predefined boundaries. Also, before applying LSS we filter out clip segments that 1. are too short, 2. have no aligned image frame, or 3. have no text context to make sure that we can retrieve the clip segments using plot summaries.

External Libraries. We use OpenAI API to access GPT-3 language model. The CLIP features are computed with the Huggingface implementations (https://huggingface.co/docs/transformers/main/en/model_doc/clip).

B Prompt Samples

We use the following prompts for each stage of Long Story Short. We break lines for visibility and instead denote the actual linebreaks with `\n`. Also, listed items within the prompts are abbreviated using ellipses (...).

Screenplay to Plot.

```
I am a highly intelligent storytelling bot.
If you give me a script, I will give
you the short synopsis in detail.\n\n
[Generated Screenplay]\n\n
Synopsis:
```

Plot Index Lookup.

```
Plot:\n
(1) [Plot1]\n
(2) [Plot2]\n
...\n
(N) [PlotN]\n\n
I am a highly intelligent question answering bot.
If you provide me with a question, I will give you
an index of the plot you should lookup to solve it.\n
Q: [Question]\n
Top 1 Plot Index: (
```

Answering Question.

Plot:\n
(1) [Plot1]\n
(2) [Plot2]\n
...\n
(N) [PlotN]\n\n[Generated Screenplay]\n\nI am a highly intelligent plot question answering bot.
If you ask me a question and candidates, I will give you
the index of answer.\n
Q: [Question]\n
Candidates:\n
(1): [Answer1]\n
...\n
(5): [Answer5]\n
A: (