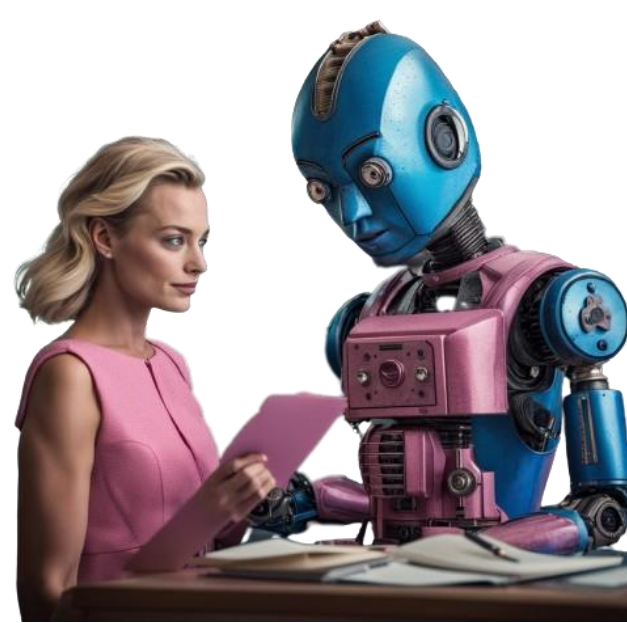


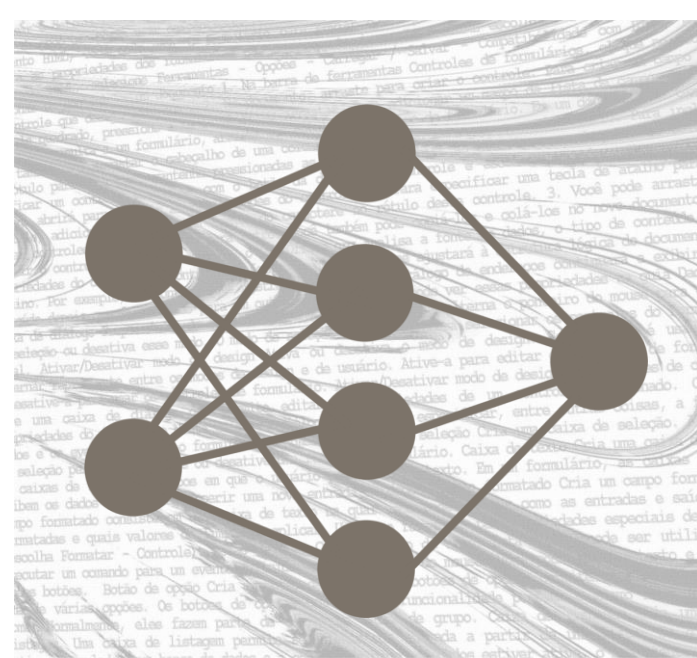
A Tale of Two LLMs: Long-form Story Generation with Large Language Models

Karan Taneja, Richard Segal, Richard Goodwin @ IBM Computational Creativity Group

Introduction



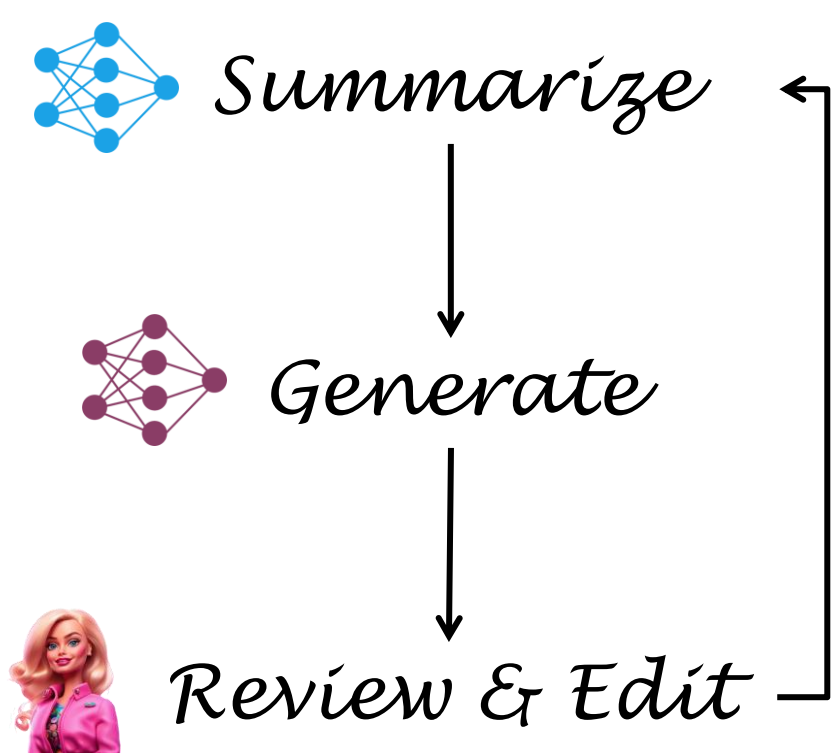
Human-AI
Collaborative Writing



Large Language Models
for Text Generation



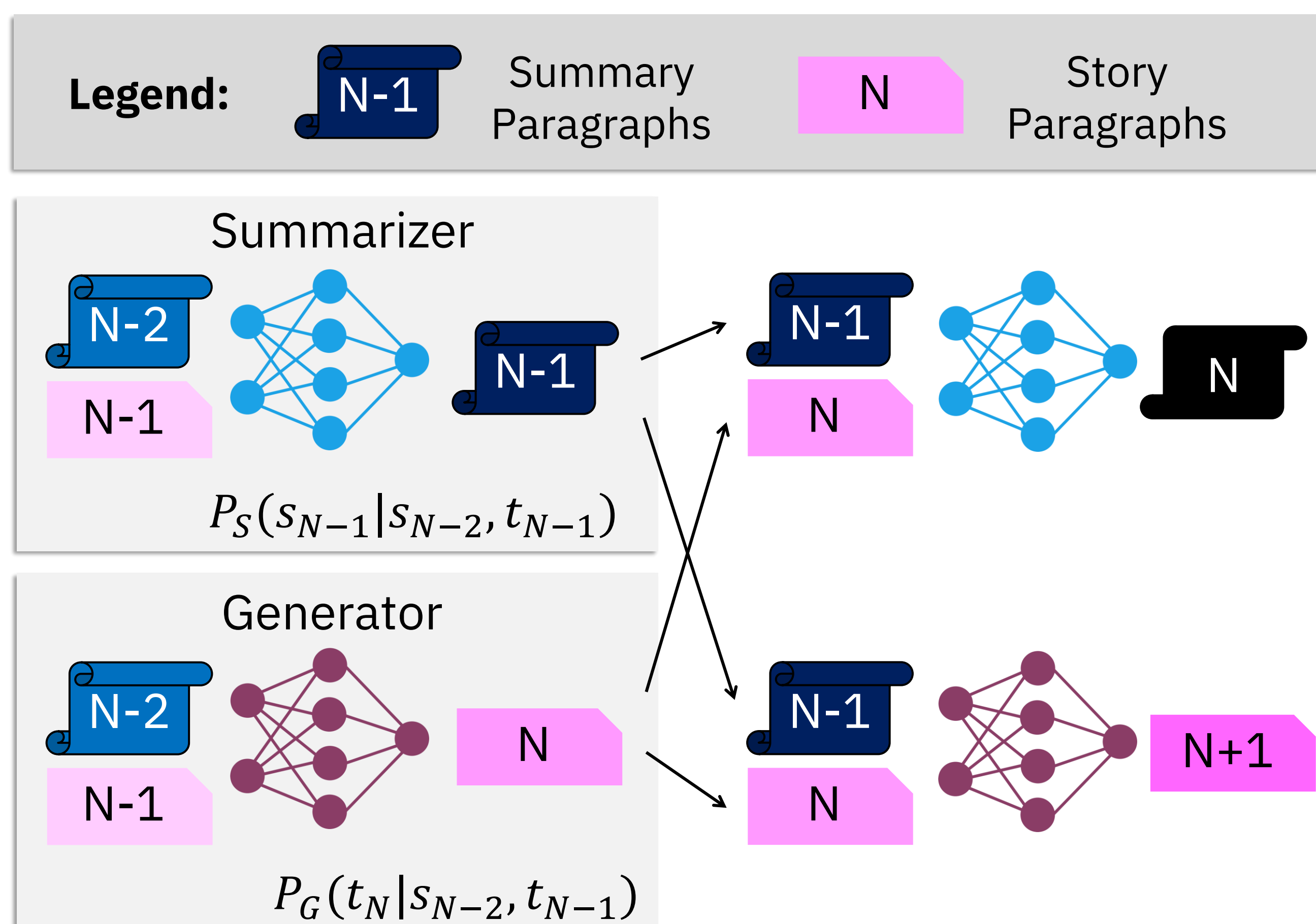
Limited Input and
Output Length



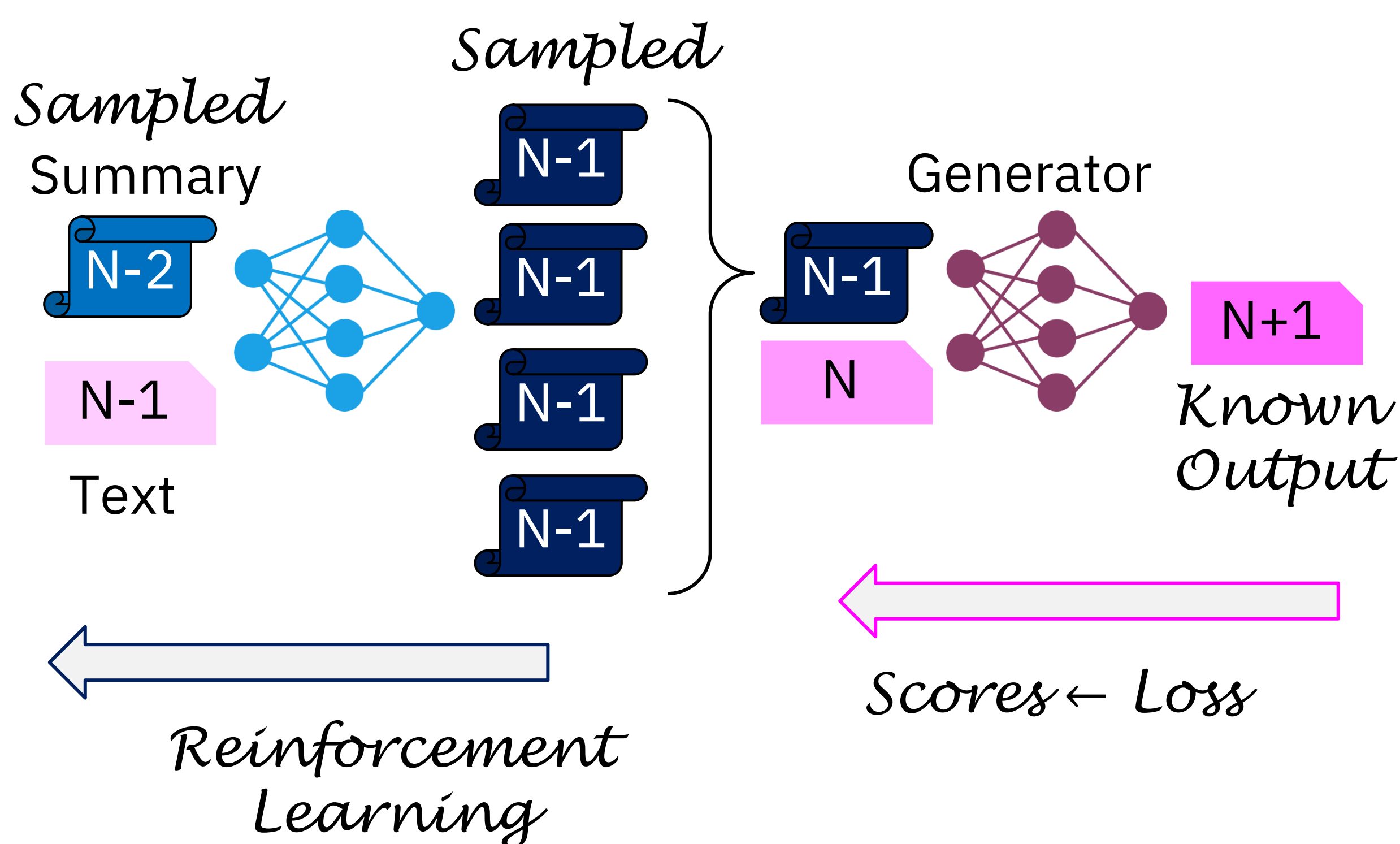
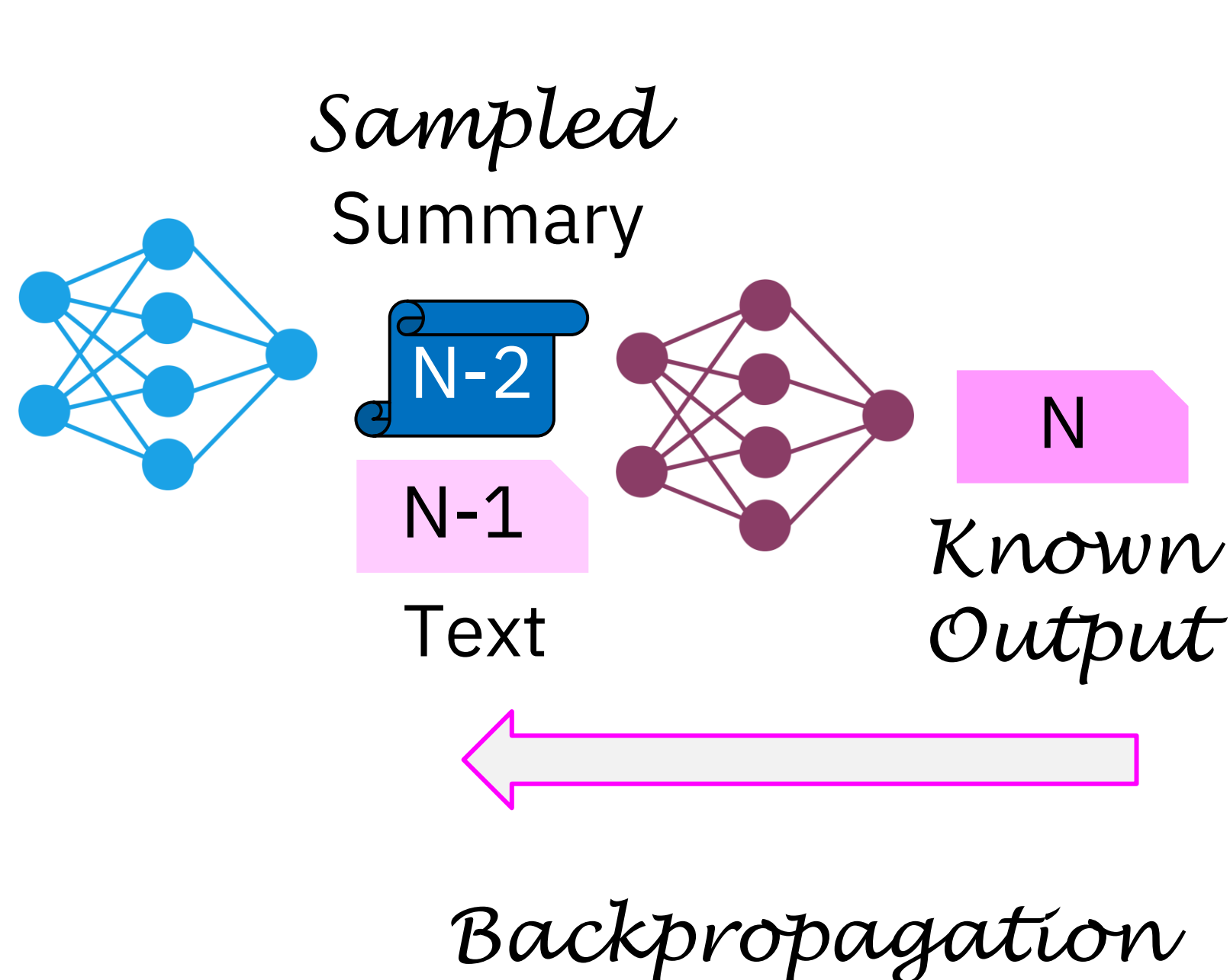
A Tale of Two LLMs
(and a human)

Generating Long-form Stories

- Summary Input:
 - Succinctly summarizes the story up to a given point in time.
 - Provides information about events and objects.
- Text Input:
 - Text from last paragraph.
 - Stylistic features such as writing style.
- Generator output is the next paragraph.
- Summarizer output is the updated summary.



Training Generator and Summarizer



Experiments

**Project
Gutenberg**

> 10k Fictional Stories



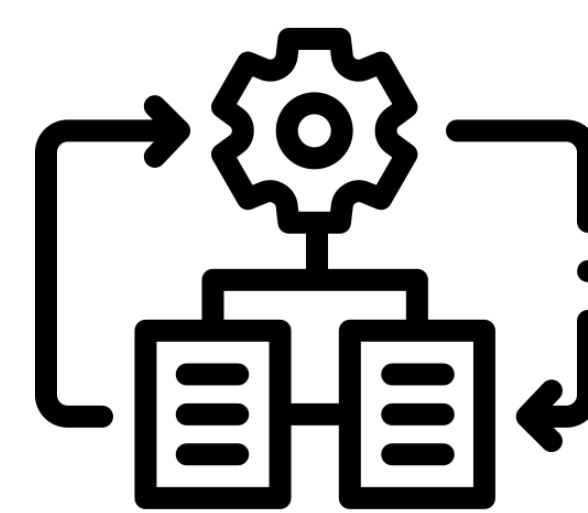
Flan-T5-XL
Flan-T5-XXL



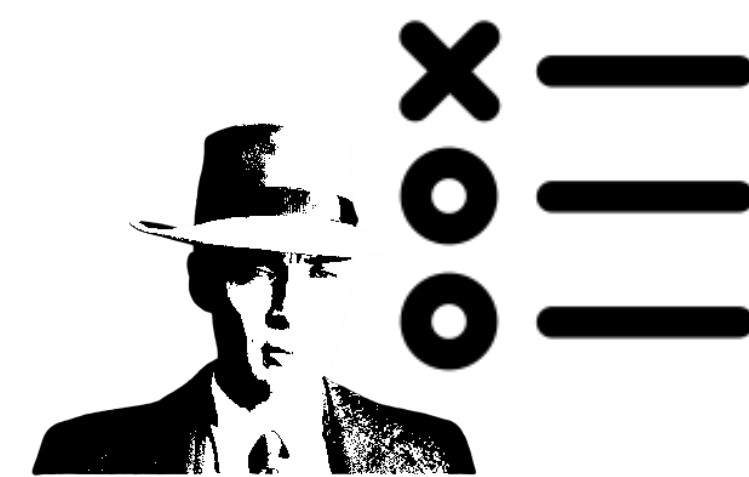
deepspeed

Distributed Training

Evaluation



Perplexity, ROUGE



Human Preferences