Cue Me In: Content Inducing Approaches to Interactive Story Generation

Faeze Brahman

Alexandru Petrusca

Snigdha Chaturvedi
Introduction

What?

- Local and global coherence
- Thematic consistency
- Creativity

Why?

- Addressing literary learning needs
- Explore creative writing at an early age
- Collaborative Authorship

Faeze Brahman, Alexandru Petrusca and Snigdha Chaturvedi. *Cue Me In: Content-Inducing Approaches to Interactive Story Generation*. AACL 2020
Motivation

Existing SOTA models for story generation generate a story in one go with limited initial input from a user.

When humans write, they incrementally edit and refine the text they produce.

We explore the problem of interactive story generation where a user provides cue phrases to guide the model generation.
Overview

**Task**: Interactive Story Generation

**Models**: Content-Inducing Decoders

**Experimental Setup**: Dataset, Training Details, Evaluation Measures

**Empirical Results**: Automatic / Human
Problem Statement

Prompt: John walked into the kitchen.

Cue phrase: dark

It was **dark** and cold.

He turned on the **light switch**.

**But** then John smelled something burning.

He **checked the oven** to find the smell.
Problem Formulation

Context: \( x = \{x_1, x_2, \ldots, x_T\} \)

Cue phrase: \( c = \{c_1, c_2, \ldots, c_k\} \)

Next sentence: \( y = \{y_1, y_2, \ldots, y_M\} \)

When generating the n-th sentence, the model takes all n-1 sentence as context along with the cue phrase.
Overview

**Task**: Interactive Story Generation

**Models**: Content-Inducing Decoders

**Experimental Setup**: Dataset, Training Details, Evaluation Measures

**Empirical Results**: Automatic / Human
Overview

**Task**: Interactive Story Generation

**Models**: Content-Inducing Decoders

**Experimental Setup**: Dataset, Training Details, Evaluation Measures

**Empirical Results**: Automatic / Human
Model Architecture

- Use the same encoding mechanism and differ only in their decoders.
- Dual encoding

Faeze Brahman, Alexandru Petrusca and Snigdha Chaturvedi. *Cue Me In: Content-Inducing Approaches to Interactive Story Generation*. AACL 2020
Dual Encoders

\[ X^{t+1} = \text{ENCBLOCK}(X^t, X^t, X^t) \]

\[ c^{t+1} = \text{ENCBLOCK}(c^t, c^t, c^t) \]
Content-Inducing Decoders — Cued Writer

Agreement between context & output

Agreement between output & cue phrase

Intra-sentence sentence agreement

\[ Y_{self}^{l+1} = \text{MULTIH}(Y_{self}^{l}, Y^{l}, Y^{l}) \]

\[ Y_{dec}^{l+1} = \text{MULTIH}(Y_{self}^{l+1}, X^{L}, X^{L}) \]

\[ Y_{cued}^{l+1} = \text{MULTIH}(Y_{self}^{l+1}, c^{L}, c^{L}) \]
Content-Inducing Decoders — Cued Writer

Feed-forward with ReLU

Integrate semantic representations

$$Y^{t+1} = \text{FFN}(Y_{int}^{t+1})$$

$$g^{t+1} = \sigma(W_1[Y_{dec}^{t+1}; Y_{cued}^{t+1}])$$

$$Y_{int}^{t+1} = W_2(g^{t+1} \circ [Y_{dec}^{t+1}; Y_{cued}^{t+1}])$$
Content-Inducing Decoders — Relevance Cued Writer

Captured relevance provided to decoder

Characterize the relevance btw context & cue phrase
Decoding and Objective

Likelihood of predicting $y_i$ given the preceding text $y_{<i}$

Cross-entropy loss:

$$- \sum_{i=1}^{M} \log \mathbb{P}(y_i | y_{<i}, X, c, \theta)$$
Overview

Task: Interactive Story Generation

Models: Content-Inducing Decoders

Experimental Setup: Dataset, Training Details, Evaluation Measures

Empirical Results: Automatic / Human
Overview

**Task:** Interactive Story Generation

**Models:** Content-Inducing Decoders

**Experimental Setup:** Dataset, Training Details, Evaluation Measures

**Empirical Results:** Automatic / Human
Training Details

- **Dataset**: ROCStories 98,161 five-sentence long stories

- **Cue phrases**: To scale model training, we use RAKE algorithm to automatically extract cue phrases from target sentence.
Evaluation

- **Baseline:**
  - SEQ2SEQ
  - Dynamic (Yao et al. 2019):
  - Static (Yao et al. 2019):
  - Vanilla:
Evaluation

• Automatic Eval.
  ✦ Perplexity
  ✦ BLEU (1-2-3)
  ✦ Greedy-Match (GM)
  ✦ Repetition-4

• Human Eval.
  ✦ Story-level
  ✦ Sentence-level
Overview

**Task:** Interactive Story Generation

**Models:** Content-Inducing Decoders

**Experimental Setup:** Dataset, Training Details, Evaluation Measures

**Empirical Results:** Automatic / Human
Overview

Task: Interactive Story Generation

Models: Content-Inducing Decoders

Experimental Setup: Dataset, Training Details, Evaluation Measures

Empirical Results: Automatic / Human
Results — Automatic Eval.

Both models outperform Dynamic and Static

Additive gain from modeling relevance

Stories more related to cue phrases

Proposed models significantly outperform all baselines

<table>
<thead>
<tr>
<th>Models</th>
<th>PPL (↓)</th>
<th>BLEU-1 (↑)</th>
<th>BLEU-2 (↑)</th>
<th>BLEU-3 (↑)</th>
<th>GM (↑)</th>
<th>Repetition-4 (↓)</th>
</tr>
</thead>
<tbody>
<tr>
<td>DYNAMIC [35]</td>
<td>29.49</td>
<td>30.05</td>
<td>9.16</td>
<td>4.59</td>
<td>0.73</td>
<td>44.36</td>
</tr>
<tr>
<td>Static [35]</td>
<td>20.81</td>
<td>33.25</td>
<td>9.64</td>
<td>4.77</td>
<td>0.75</td>
<td>26.26</td>
</tr>
<tr>
<td>SEQ2SEQ</td>
<td>20.97</td>
<td>33.91</td>
<td>10.01</td>
<td>3.09</td>
<td>0.82</td>
<td>33.23</td>
</tr>
<tr>
<td>VANILLA</td>
<td>15.78</td>
<td>40.30</td>
<td>16.09</td>
<td>7.19</td>
<td>0.89</td>
<td>20.87</td>
</tr>
<tr>
<td>Cued Writer</td>
<td>14.80</td>
<td>41.50</td>
<td>16.72</td>
<td>7.25</td>
<td>0.92</td>
<td>15.08</td>
</tr>
<tr>
<td>Rel. Cued Writer</td>
<td><strong>14.66</strong></td>
<td><strong>42.65</strong></td>
<td><strong>17.33</strong></td>
<td><strong>7.59</strong></td>
<td><strong>0.94</strong></td>
<td>16.23</td>
</tr>
</tbody>
</table>

Our models outperform all baselines.
The two proposed models are almost comparable but show a general superior performance.
Results — Human Eval.

Percentage of preference for our stronger model, Rel. Cued Writer, over other baselines

Story-level

Sentence-level

Our stronger model, Rel. Cued Writer, was preferred over other baselines in both experiments.
Qualitative Results

Prompt: Harry really wanted to see the new Star Wars movie.
Cue phrases: opening day - ticket - entire movie - really

Static

One day, she was in a hurry to get to the theater. She wasn't sure if she would like to go. She was excited to try the new movie. She was so excited to see the movie. [opening day] [ticket]

Vanilla

On opening day, Harry was very nervous. He bought a ticket to the theater. He bought Harry ticket tickets to the theater. He really didn't like the movie. [entire movie]

Rel. Cued Writer

He decided to watch it on opening day. He got to the theater and got a ticket. He watched the entire movie. He was really excited about it.

Our novel Content-Inducing approach generates stories that suffer less from listed issues.
Summary

🌟 We explored the problem of interactive storytelling, which leverages human and computer collaboration for creative language generation.

🌟 We propose two content-inducing approaches to incorporate additional information into the generation phase.

🌟 We evaluate the utility of our approaches through several experiments and demonstrate that our methods outperformed competitive baselines.

🌟 We suggest other significant aspects of story generation, such modeling of discourse relations, and representation of key narrative elements, which could be possible future directions.
Thank you!
Any question?