



# Enhanced Story Representation by ConceptNet for Predicting Story Endings

Paper Authors:  
Shanshan Huang  
Kenny Q. Zhu  
Qianzi Liao



# Introduction

Story-ending prediction is a significant challenge in AI, requiring commonsense reasoning and understanding of story structure.

Pre-trained models like BERT have shown progress but often rely on patterns rather than true comprehension.

This paper proposes an improved method by simplifying sentences into key concepts using **ConceptNet** and modeling relationships between those concepts.

The approach enhances story representation and improves prediction accuracy on the **Story Cloze Test (SCT)**.



# Dataset Used

Source Dataset Used: Story Cloze Test (SCT) dataset

An example from the story-cloze task:

## **Context:**

1. Tiffany was getting overwhelmed at work.
2. While she liked her job, she longed for a break.
3. One day, she tripped outside on uneven pavement.
4. She broke her ankle and had to be off work for a couple of months.

## **Ending:**

5. Tiffany went in to work the next day.
6. She was in pain but happy to have a rest.



## Summary

The paper presents a method for improving story comprehension and prediction of story endings using commonsense knowledge from ConceptNet.

The authors propose a system that **simplifies story sentences** by extracting key concepts from ConceptNet and **incorporating the relationships** between these concepts to better represent the story.

This improved representation enhances the performance of pre-trained language models like DSSM, SKBC and BERT for predicting story endings. The method outperforms several state-of-the-art techniques on the SCT dataset.



# Strengths of the Paper

**Innovative Sentence Simplification:** The authors streamline story sentences by focusing on key concepts rather than all the words, reducing noise and improving semantic understanding.

**Use of ConceptNet:** Integrating structured commonsense knowledge from ConceptNet(Simplification + Concept Encoding) provides a deeper understanding of story logic, helping models make better predictions.

**Improved Accuracy:** The method demonstrates substantial improvement over existing methods in predicting story endings, achieving better accuracy on the SCT dataset.

**Versatile Approach:** The proposed system is flexible and can be applied to various pre-trained models (e.g., BERT, DSSM, SKBC), showing its adaptability.



# Weaknesses of the Paper

## Limited Dataset:

The results are primarily evaluated on the SCT dataset, which has been criticized for biases, such as human-authorship bias in validation data.

The paper addresses this issue by creating a new training set, but **it still relies heavily on SCT.**

**Simplification Challenges:** While sentence simplification improves prediction accuracy, **there's a risk of losing critical context**, which may affect understanding of more complex narratives.



# Relation to Story Generation and Interactive Fiction

## Story Generation:

- The concept of simplification and concept encoding (SI + CE) can be used to create logical & reasonable storylines, by using commonsense knowledge from ConceptNet, we can generate plausible story continuations, making it a useful tool for interactive storytelling systems.

## Interactive Fiction:

- The proposed idea (SI + CE approach) can be used in interactive fiction games, where the player choices influences the narration in a meaningful way with more believable and engaging story progression, by understanding the key elements of the player's choices and how they connect to future events, the system can generate appropriate narrative responses.



# Conclusion

Key concepts and events are sufficient for story comprehension when **informative words** are preserved.

**Capturing relationships** between concepts across sentences is crucial for logical inference in the story.

It is observed that SKBC model with our methods achieves 69.7% accuracy(+5% improvement), which is the best among our experiments. It performs better than any other commonly-used models we tested.

Results show improved performance using **unbiased, automatically generated datasets**.





THANK YOU !!!  
ANY QUESTIONS