Guided Story Generation

Module 2 – 10/1/2024

CMSC 491/691 – Interactive Fiction and Text Generation

Lara J. Martin

Learning Objectives

- Appraise the different ways people have used script/plot-like structures to guide neural networks
- Consider how a guided system would work with transformers
- Compare and contrast previously-made guided systems
- Define the Story Cloze Test and determine its place in guided story generation

Plan-and-Write

The Thirty-Third AAAI Conference on Artificial Intelligence (AAAI-19)

Plan-and-Write: Towards Better Automatic Storytelling

Lili Yao,[1,3*] **Nanyun Peng,**[2*] **Ralph Weischedel,**[2] **Kevin Knight,**[2] **Dongyan Zhao,**[1] **Rui Yan**[1†]
liliyao@tencent.com, {npeng,weisched,knight}@isi.edu
{zhaodongyan,ruiyan}@pku.edu.cn
[1]Institute of Computer Science and Technology, Peking University
[2]Information Sciences Institute, University of Southern California, [3]Tencent AI Lab

Abstract

Automatic storytelling is challenging since it requires generating long, coherent natural language to describes a sensible sequence of events. Despite considerable efforts on automatic story generation in the past, prior work either is restricted in plot planning, or can only generate stories in a narrow domain. In this paper, we explore open-domain story generation that writes stories given a title (topic) as input. We propose a *plan-and-write* hierarchical generation framework that first plans a storyline, and then generates a story based on the storyline. We compare two planning strategies. The *dynamic* schema interweaves story planning and its surface realization in text, while the *static* schema plans out the entire storyline before generating stories. Experiments show that with explicit storyline planning, the generated stories are more diverse, coherent, and on topic than those generated without creating a full plan, according to both automatic and human evaluations.

Introduction

A narrative or story is anything which is told in

Title (Given)	The Bike Accident
Storyline (Extracted)	Carrie → bike → sneak → nervous → leg
Story (Human Written)	Carrie had just learned how to ride a bike. She didn't have a bike of her own. Carrie would sneak rides on her sister's bike. She got nervous on a hill and crashed into a wall. The bike frame bent and Carrie got a deep gash on her leg.

Table 1: An example of title, storyline and story in our system. A storyline is represented by an ordered list of words.

and Young 2010), we propose to decompose story generation into two steps: 1) story planning which generates plots, and 2) surface realization which composes natural language text based on the plots. We propose a *plan-and-write* hierarchical generation framework that combines plot planning and surface realization to generate stories fr

Yao, L., Peng, N., Weischedel, R., Knight, K., Zhao, D., & Yan, R. (2019). Plan-And-Write: Towards Better Automatic Storytelling. *AAAI Conference on Artificial Intelligence (AAAI)*, *33*(1), 7378–7385. https://aaai.org/ojs/index.php/AAAI/article/view/4726

Extracting Plots

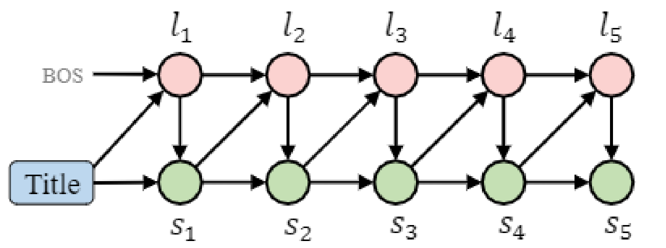
Carrie had just learned how to ride a bike. She didn't have a bike of her own. Carrie would sneak rides on her sister's bike. She got nervous on a hill and crashed into a wall. The bike frame bent and Carrie got a deep gash on her leg.

Carrie→bike→sneak→nervous→leg

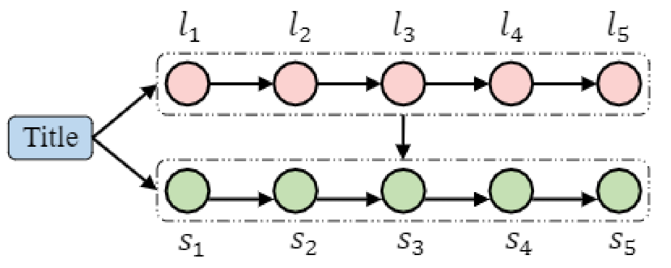
Yao, L., Peng, N., Weischedel, R., Knight, K., Zhao, D., & Yan, R. (2019). Plan-And-Write: Towards Better Automatic Storytelling. *AAAI Conference on Artificial Intelligence (AAAI)*, *33*(1), 7378–7385. https://aaai.org/ojs/index.php/AAAI/article/view/4726

Plan-and-Write Overview

- Extracted most important word from each sentence using RAKE algorithm (keyword extraction) to create a storyline (aka plot)
- Used storyline as input to plan out stories
- Dynamic generation → using storyline and sentences to inform each other
- Static generation → plan ahead and then generate



(a) Dynamic schema work-flow.



(b) Static schema work-flow.

Yao, L., Peng, N., Weischedel, R., Knight, K., Zhao, D., & Yan, R. (2019). Plan-And-Write: Towards Better Automatic Storytelling. *AAAI Conference on Artificial Intelligence (AAAI)*, *33*(1), 7378–7385. https://aaai.org/ojs/index.php/AAAI/article/view/4726

System Diagram

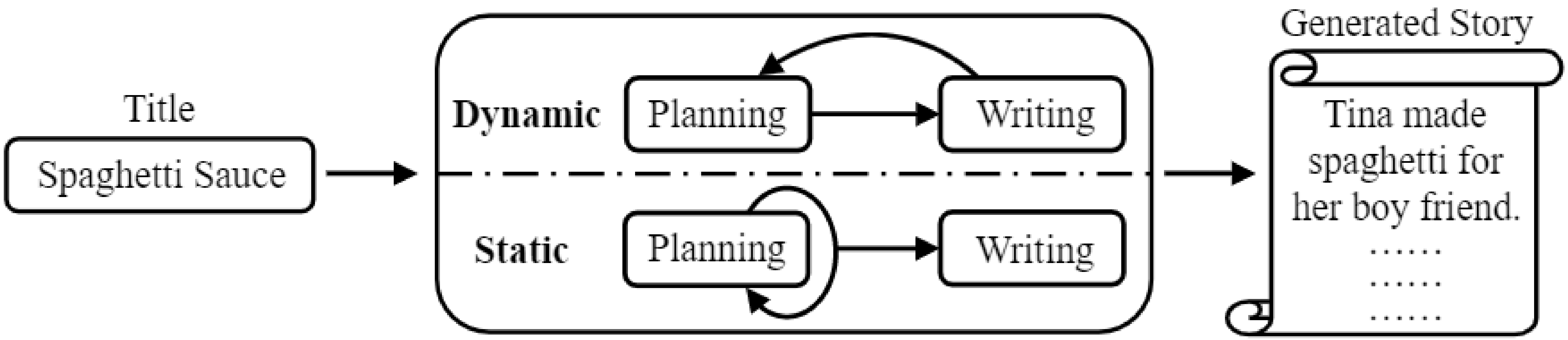


Figure 1: An overview of our system.

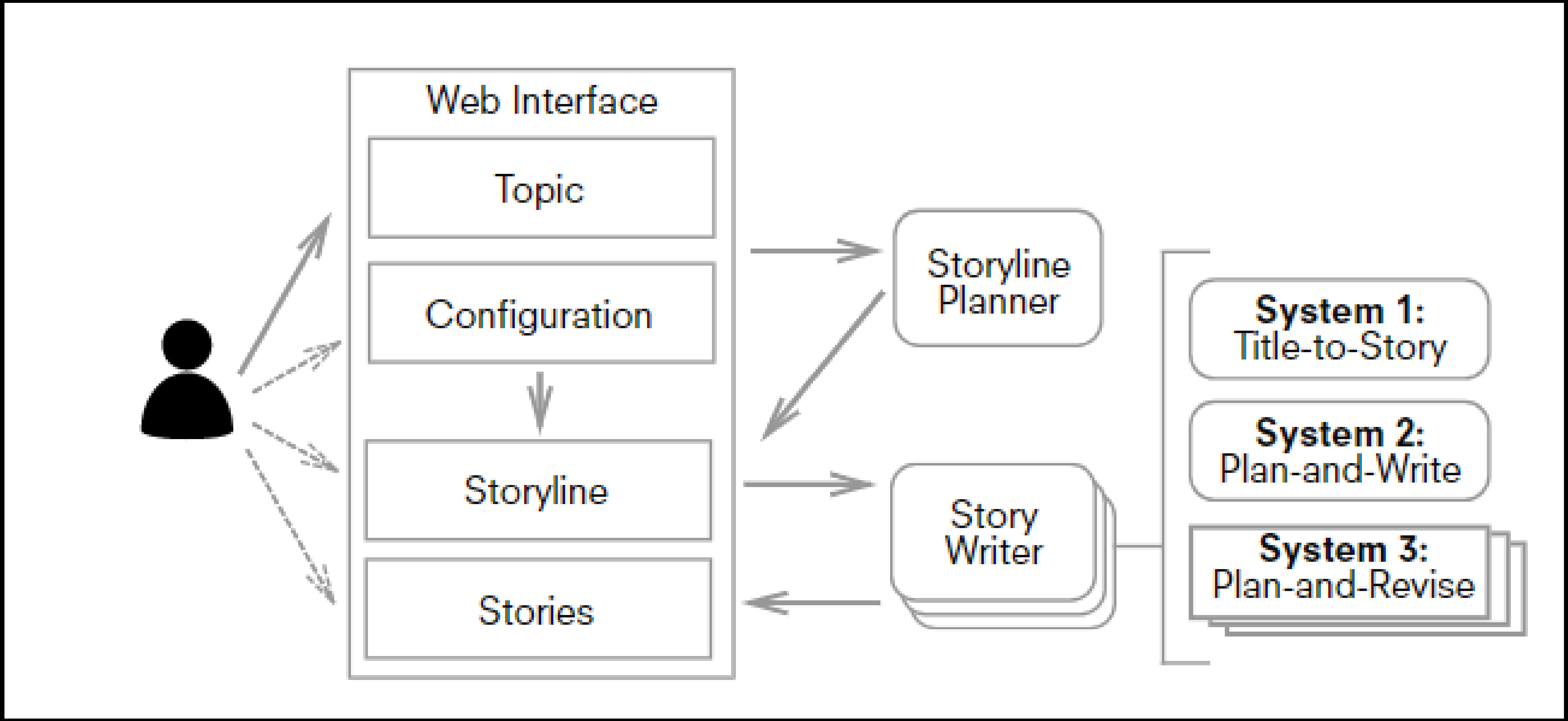
Yao, L., Peng, N., Weischedel, R., Knight, K., Zhao, D., & Yan, R. (2019). Plan-And-Write: Towards Better Automatic Storytelling. *AAAI Conference on Artificial Intelligence (AAAI)*, *33*(1), 7378–7385. https://aaai.org/ojs/index.php/AAAI/article/view/4726

Examples

Title: Computer		
Baselines	Inc-S2S	Tom's computer broke down. He needed to buy a new computer. He decided to buy a new computer. Tom bought a new computer. Tom was able to buy a new computer.
	Cond-LM	The man bought a new computer. He went to the store. He bought a new computer. He bought the computer. He installed the computer.
Dynamic	Storyline	needed $\rightarrow$ money $\rightarrow$ computer $\rightarrow$ bought $\rightarrow$ happy
	Story	John needed a computer for his birthday. He worked hard to earn money. John was able to buy his computer. He went to the store and bought a computer. John was happy with his new computer.
Static	Storyline	computer $\rightarrow$ slow $\rightarrow$ work $\rightarrow$ day $\rightarrow$ buy
	Story	I have an old computer. It was very slow. I tried to work on it but it wouldn't work. One day, I decided to buy a new one. I bought a new computer .

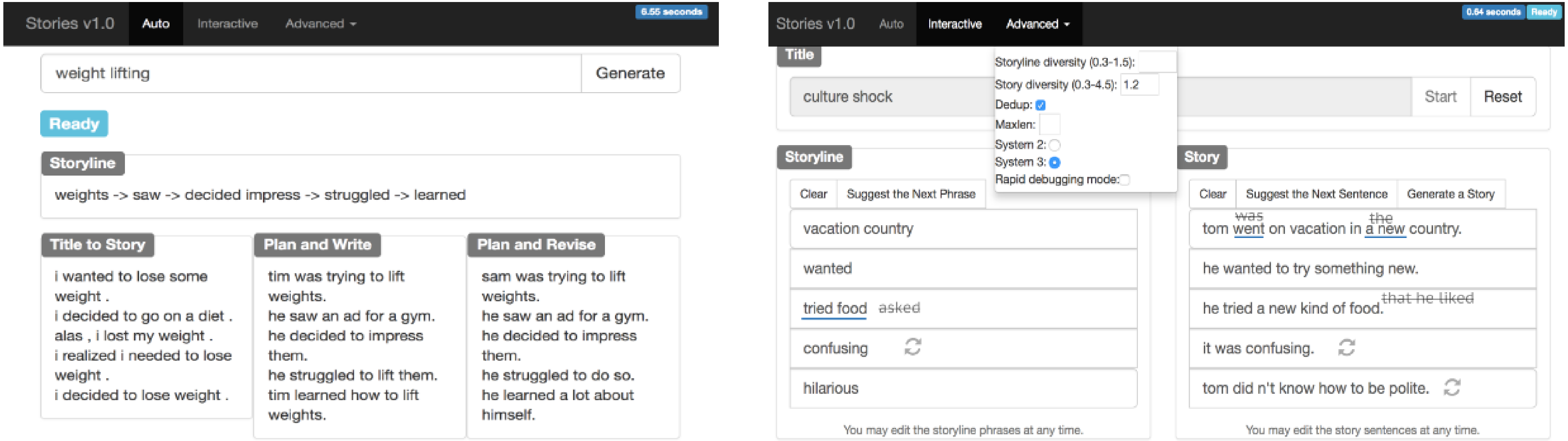
Yao, L., Peng, N., Weischedel, R., Knight, K., Zhao, D., & Yan, R. (2019). Plan-And-Write: Towards Better Automatic Storytelling. *AAAI Conference on Artificial Intelligence (AAAI)*, *33*(1), 7378–7385. https://aaai.org/ojs/index.php/AAAI/article/view/4726

Plan, Write, *and Revise*



Goldfarb-Tarrant, S., Feng, H., & Peng, N. (2019). Plan, Write, and Revise: an Interactive System for Open-Domain Story Generation. *Conference of the North American Chapter of the Association for Computational Linguistics (NAACL) Demo Track*.

Plan, Write, and Revise



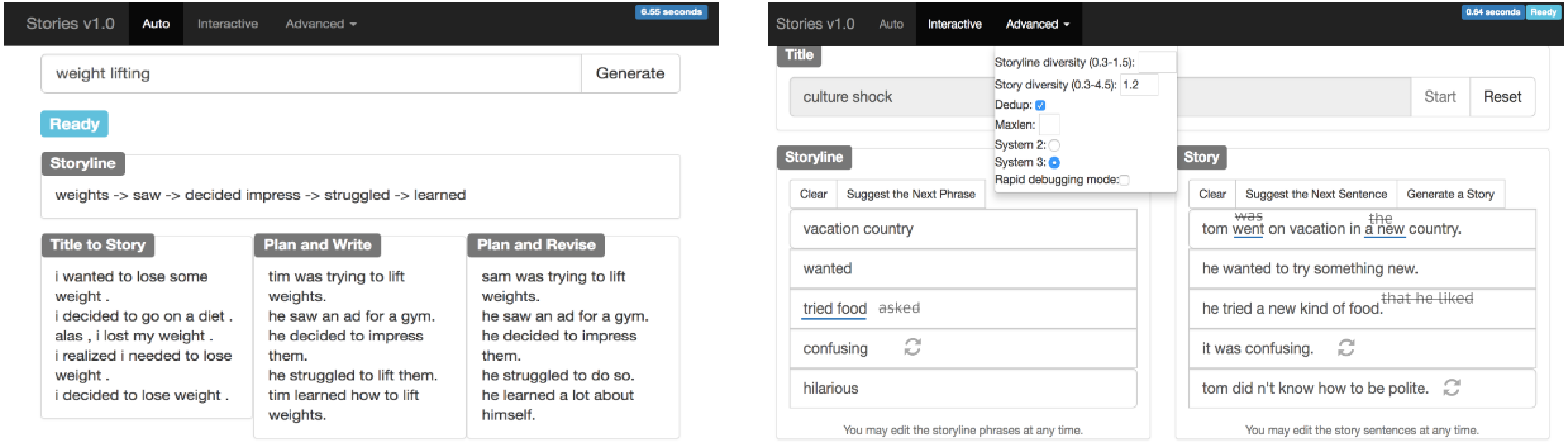
(a) cross-model interaction, comparing three models with advanced options to alter the storyline and story diversities.

(b) intra-model interaction, showing advanced options and annotated with user interactions from an example study.

Goldfarb-Tarrant, S., Feng, H., & Peng, N. (2019). Plan, Write, and Revise: an Interactive System for Open-Domain Story Generation. *Conference of the North American Chapter of the Association for Computational Linguistics (NAACL) Demo Track*.

Think-Pair-Share

- Plan, Write, and Revise was written in 2019. How might you “update” this work? Does it need to be updated?



(a) cross-model interaction, comparing three models with advanced options to alter the storyline and story diversities.

(b) intra-model interaction, showing advanced options and annotated with user interactions from an example study.

Re^3

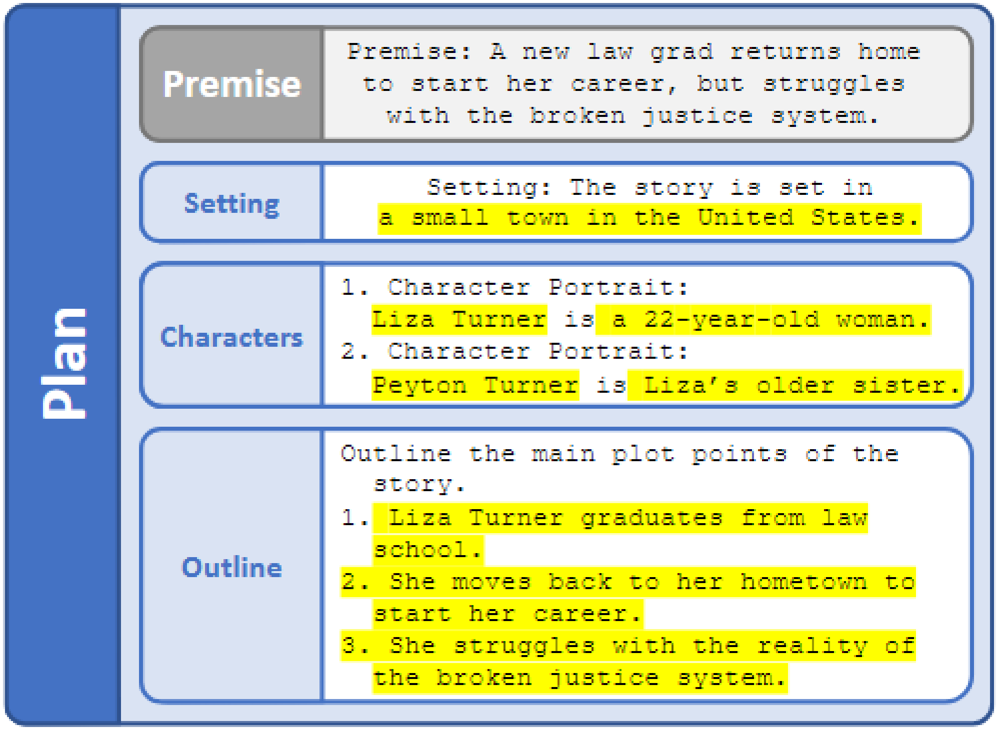


Figure 2: Illustration of Re^3's Plan module, which prompts a language model to generate a setting, characters, and outline based on the premise. Highlighting indicates generated text.

Re^3: Generating Longer Stories With Recursive Reprompting and Revision

Kevin Yang[1] **Yuandong Tian**[2] **Nanyun Peng**[3] **Dan Klein**[1]
[1]UC Berkeley, [2]Meta AI, [3]UCLA
{yangk,klein}@berkeley.edu,yuandong@meta.com,violetpeng@cs.ucla.edu

Abstract

We consider the problem of automatically generating longer stories of over two thousand words. Compared to prior work on shorter stories, long-range plot coherence and relevance are more central challenges here. We propose the Recursive Reprompting and Revision framework (Re^3) to address these challenges by (a) prompting a general-purpose language model to construct a structured overarching plan, and (b) generating story passages by repeatedly injecting contextual information from both the plan and current story state into a language model prompt. We then revise by (c) reranking different continuations for plot coherence and premise relevance, and finally (d) editing the best continuation for factual consistency. Compared to similar-length stories generated directly from the same base model, human evaluators judged substantially more of Re^3's stories as having a coherent overarching plot (by 14% absolute increase), and relevant to the given initial premise (by 20%).

1 Introduction

Generating long-term coherent stories is a long-standing challenge for artificial intelligence, requir-

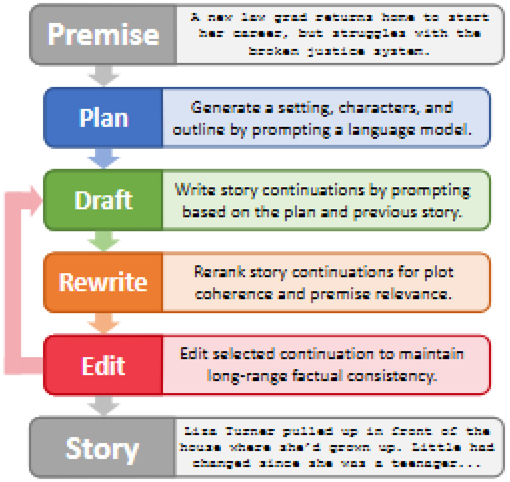


Figure 1: High-level overview of Re^3.

length increases limited primarily by evaluation rather than technical issues.[1] Generating stories of such length faces qualitatively new challenges compared to prior work on shorter stories. First, the system must maintain a coherent overarching plot over thousands of words. Given an initial premise, it should maintain relevance to this premise over thousands of words as well. Additional challenges include preservation of narration style and avoiding

K. Yang, Y. Tian, N. Peng, and D. Klein, "Re^3: Generating Longer Stories With Recursive Reprompting and Revision," in *Conference on Empirical Methods in Natural Language Processing (EMNLP)*, Abu Dhabi, United Arab Emirates: Association for Computational Linguistics, Dec. 2022, pp. 4393–4479. doi: 10.18653/v1/2022.emnlp-main.296.

Guided Open Story Generation Using Probabilistic Graphical Models

Guided Open Story Generation Using Probabilistic Graphical Models

Sagar Gandhi*
University of Kentucky
Lexington, Kentucky
sga267@uky.edu

Brent Harrison
University of Kentucky
Lexington, Kentucky
harrison@cs.uky.edu

ABSTRACT

In this work, we present an approach for performing computational storytelling in open domain based on Author Goals. Author Goals are constraints placed on a story event directed by the author of the system. There are two challenges present in this type of story generation: (1) automatically acquiring a model of story progression, and (2) guiding the progress of story progression in light of different goals. We propose a novel approach to story generation based on probabilistic graphical models and Loopy Belief Propagation (LBP) that addresses both of these problems. We show the applicability of our technique through a case study on the Visual Storytelling (VIST) 2017 dataset. We use image descriptions as author goals. This empirical analysis suggests that our approach is able to utilize goals information to better automatically generate stories.

CCS CONCEPTS

• **Computing methodologies** → **Probabilistic reasoning**; *Discourse, dialogue and pragmatics*; *Natural language generation*; *Learning in probabilistic graphical models*.

problem *domain model* [3, 5, 25]. This type of story generation is sometimes called *Closed Story Generation*. While stories generated in this way are often coherent, the space of possible stories that can be generated is tightly coupled with the domain model. In order to generate different types of stories, a new domain model must be authored, which is often a time consuming task.

Open Story Generation seeks to address this limitation by enabling stories to be told in any conceivable domain through the use of machine learning. The goal is to use these systems to tell a variety of stories without the need for re-learning specific domain models.

One limitation of open story generation systems is that, due to the complexity of the models used for training, it can be difficult to encode specific author goals into the story generation process. To address this limitation of open story generation systems, we propose a new framework that can reason about the overall structure of a story as well as how to incorporate author goals into the story generation process. This system enables many different types of stories to be generated based on these author goals without needing to retrain or re-learn a domain model.

In this paper, we use a novel approach for o[illegible] generation

Gandhi, S., & Harrison, B. (2019). Guided open story generation using probabilistic graphical models. *International Conference on the Foundations of Digital Games (FDG)*, 1–7. https://doi.org/10.1145/3337722.3341871

Overview

- Use discourse representation structure (DRS) parser to get semantic relationships

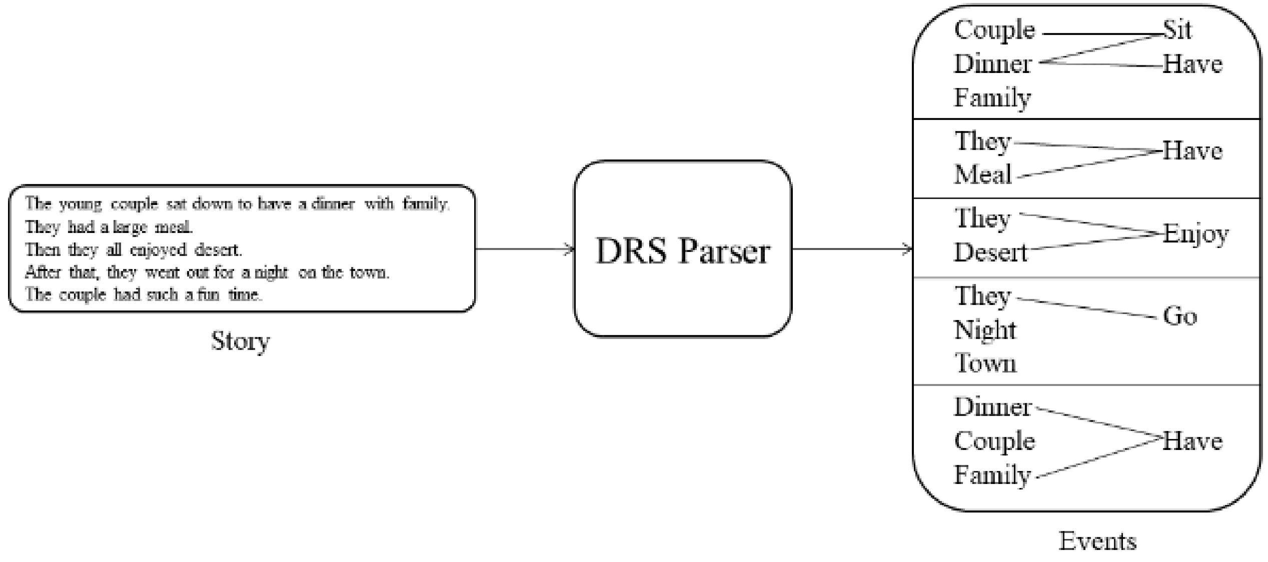


Figure 1: Event Representation using DRS Parser. Note that the words without edges are removed while forming the graphs.

Gandhi, S., & Harrison, B. (2019). Guided open story generation using probabilistic graphical models. *International Conference on the Foundations of Digital Games (FDG)*, 1–7. https://doi.org/10.1145/3337722.3341871

Overview

- Use discourse representation structure (DRS) parser to get semantic relationships
- Story graph: “edges […] between story verb nodes and story subject-object nodes.”

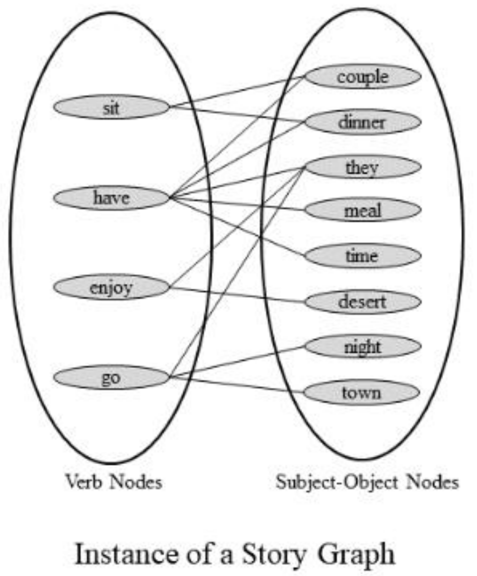


Figure 2: Story graph example. By adding the events generated in Figure 1, we get this instance of a graph.

Gandhi, S., & Harrison, B. (2019). Guided open story generation using probabilistic graphical models. *International Conference on the Foundations of Digital Games (FDG)*, 1–7. https://doi.org/10.1145/3337722.3341871

Overview

- Use discourse representation structure (DRS) parser to get semantic relationships
- Story graph: “edges […] between story verb nodes and story subject-object nodes.”
- Goal Graph: “relationship between author goals and story events”

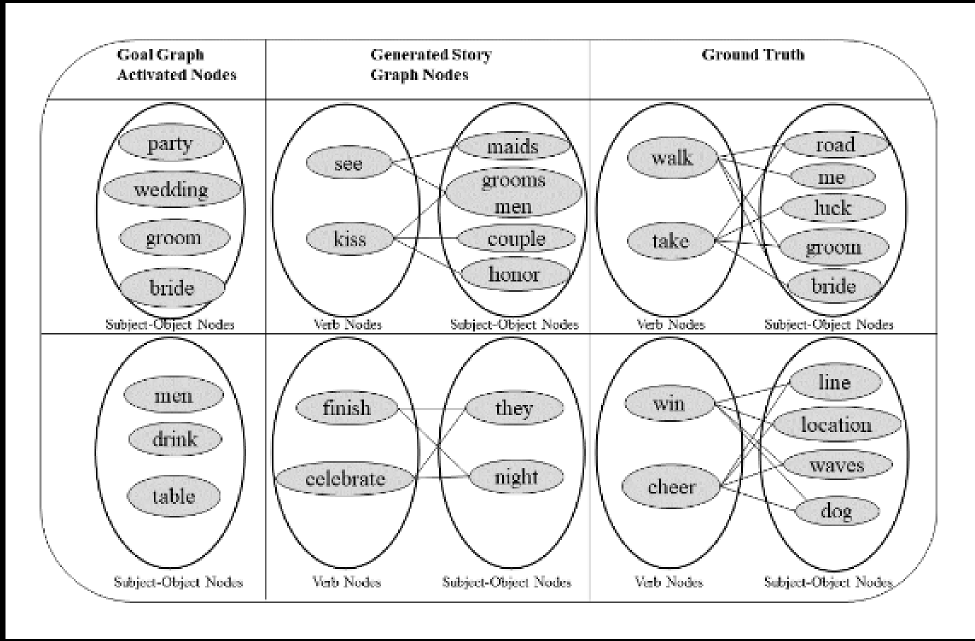


Figure 4: Evaluation Intricacies. Shown are two examples of generated action verbs and subject-action words from the given author goal (Description).

Gandhi, S., & Harrison, B. (2019). Guided open story generation using probabilistic graphical models. *International Conference on the Foundations of Digital Games (FDG)*, 1–7. https://doi.org/10.1145/3337722.3341871

Overview

- Use discourse representation structure (DRS) parser to get semantic relationships
- Story graph: “edges […] between story verb nodes and story subject-object nodes.”
- Goal Graph: “relationship between author goals and story events”
- Use VIST Storytelling dataset

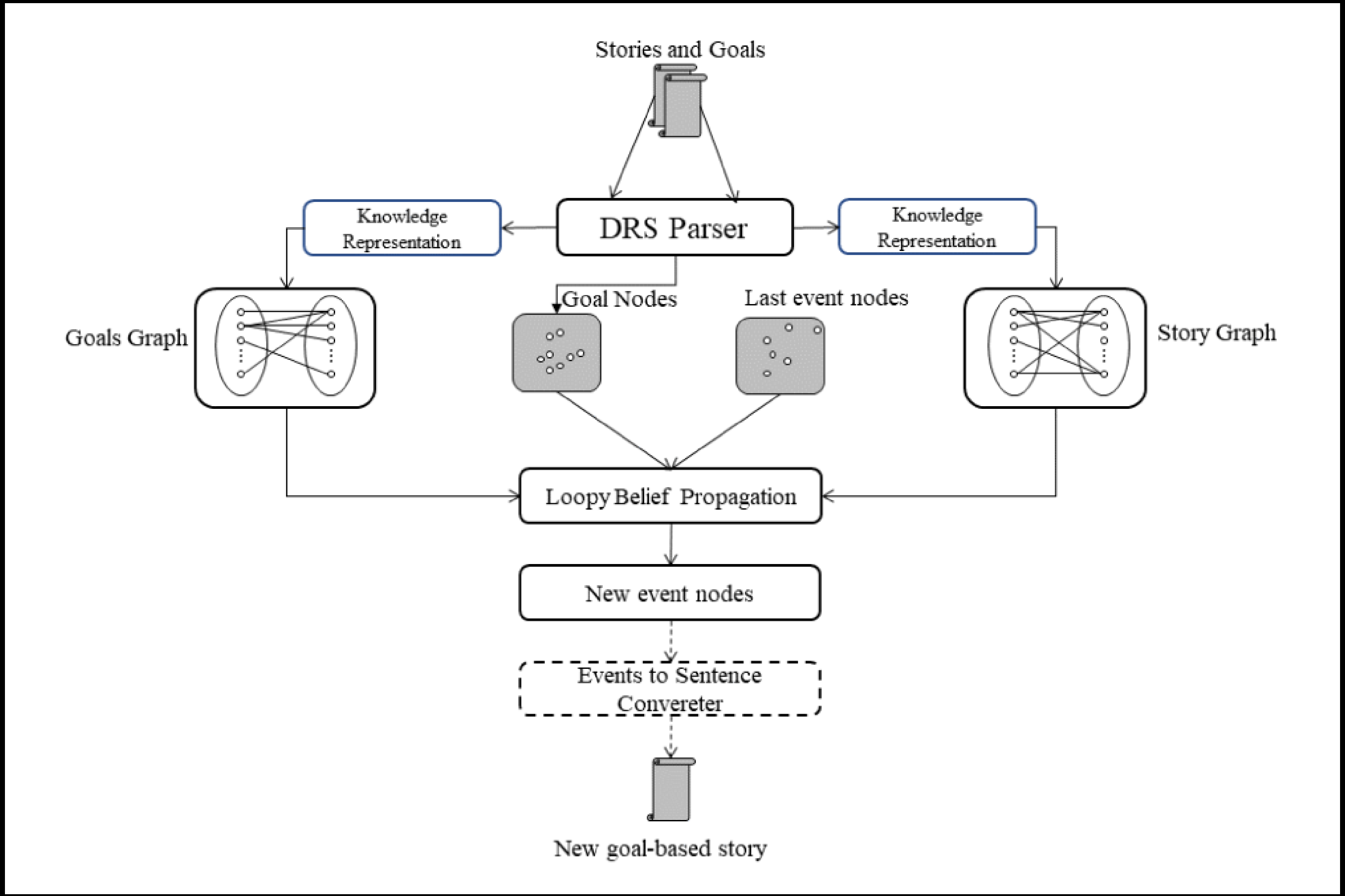
Gandhi, S., & Harrison, B. (2019). Guided open story generation using probabilistic graphical models. *International Conference on the Foundations of Digital Games (FDG)*, 1–7. https://doi.org/10.1145/3337722.3341871

VIST

From [VIST: Visual Storytelling Dataset (visionandlanguage.net)]

T.-H. Huang *et al.*, “Visual Storytelling,” in *Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies (NAACL)*, San Diego, California: Association for Computational Linguistics, Jun. 2016, pp. 1233–1239. Available: https://aclanthology.org/N16-1147/

System Diagram



Gandhi, S., & Harrison, B. (2019). Guided open story generation using probabilistic graphical models. *International Conference on the Foundations of Digital Games (FDG)*, 1–7. https://doi.org/10.1145/3337722.3341871

Story Generation

ALGORITHM 1: Story generation algorithm

```
Data: Story Graph, Goal Graph
Result: Story S
for event_i ← 1 : n do
    if CG != ParseCurrentGoal() then
        CG = ParseCurrentGoal()
    end
    Initial SVN = LBPInfer(StoryGraph, SSON_{i-1})
    SVN_i = LBPInfer(GoalGraph, Initial SVN, CG)
    SSON_i = LBPInfer(StoryGraph, SVN_i)
    S+ = GetEvent(SVN_i, SSON_i)
end
```

Subject-Verb

Subject-Object

Gandhi, S., & Harrison, B. (2019). Guided open story generation using probabilistic graphical models. *International Conference on the Foundations of Digital Games (FDG)*, 1–7. https://doi.org/10.1145/3337722.3341871

Story Realization: Expanding Plot Events into Sentences

The Thirty-Fourth AAAI Conference on Artificial Intelligence (AAAI-20)

Story Realization: Expanding Plot Events into Sentences

Prithviraj Ammanabrolu, Ethan Tien, Wesley Cheung, Zhaochen Luo, William Ma, Lara J. Martin, Mark O. Riedl
School of Interactive Computing
Georgia Institute of Technology
{raj.ammanabrolu, etien, wcheung8, zluo, wma61, ljmartin, riedl}@gatech.edu

Abstract

Neural network based approaches to automated story plot generation attempt to learn how to generate novel plots from a corpus of natural language plot summaries. Prior work has shown that a semantic abstraction of sentences called *events* improves neural plot generation and and allows one to decompose the problem into: (1) the generation of a sequence of events (event-to-event) and (2) the transformation of these events into natural language sentences (event-to-sentence). However, typical neural language generation approaches to event-to-sentence can ignore the event details and produce grammatically-correct but semantically-unrelated sentences. We present an ensemble-based model that generates natural language guided by events. We provide results—including a human subjects study—for a full end-to-end automated story generation system showing that our method generates more coherent and plausible stories than baseline approaches [1].

1 Introduction

Automated story plot generation is the problem of creating a sequence of main plot points for a story in a given domain. Generated plots must remain consistent across th entire story, preserve long-term dependencies, an

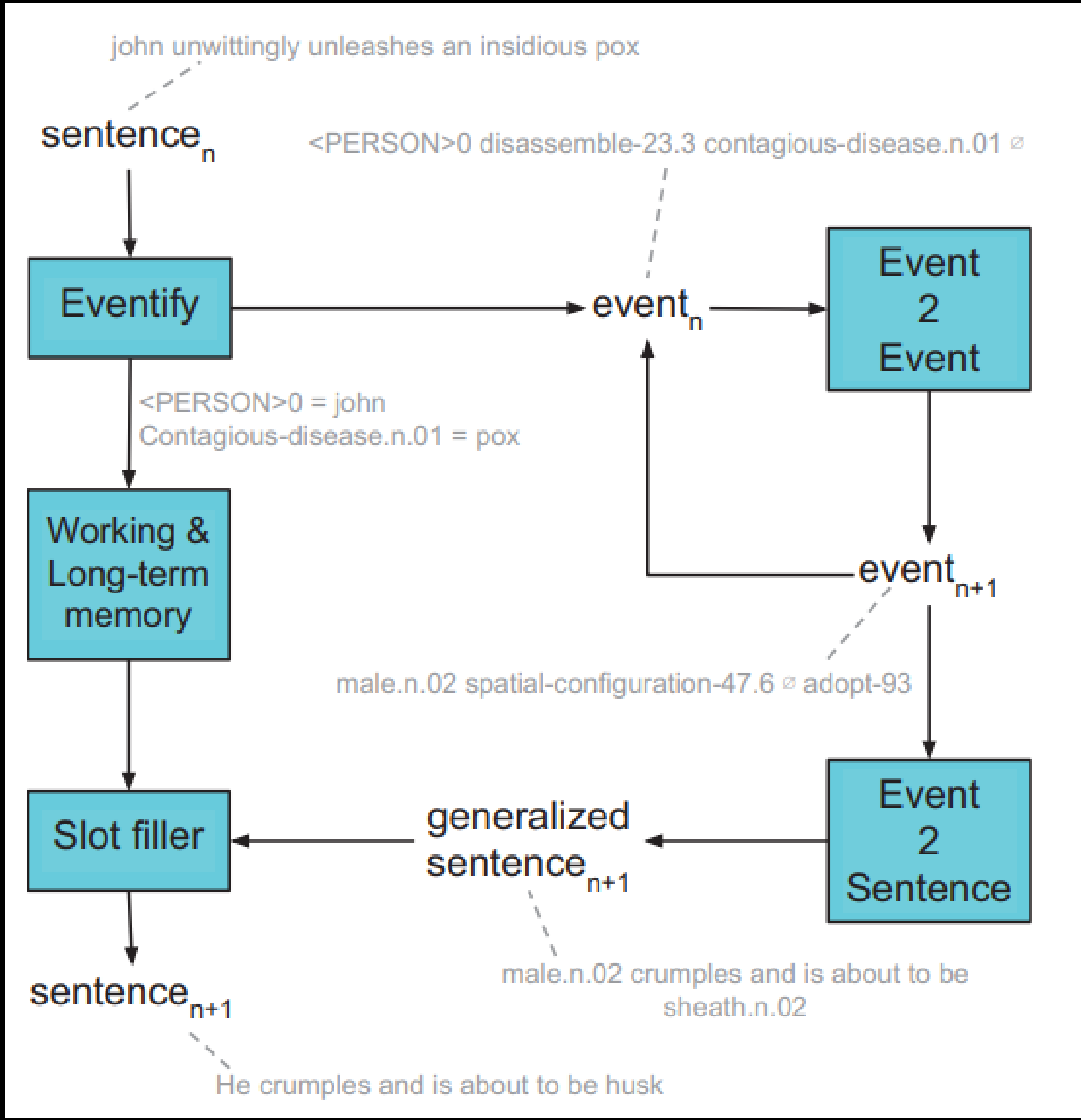
for explicit domain modeling beyond providing a corpus of example stories. The primary pitfall of neural language model approaches for story generation is that the space of stories that can be generated is huge, which in turn, implies that, in a textual story corpora, any given sentence will likely only be seen once.

Martin et al. (2018) propose the use of a semantic abstraction called an *event*, reducing the sparsity in a dataset that comes from an abundance of unique sentences. They define an event to be a unit of a story that creates a change in the story world's state. Technically, an event is a tuple containing a subject, verb, direct object, and some additional disambiguation token(s).

The event representation enables the decomposition of the plot generation task into two sub-problems: *event-to-event* and *event-to-sentence*. Event-to-event is broadly the problem of generating the sequence of events that together comprise a plot. Models used to address this problem are also responsible for maintaining plot coherence and consistency. Once new events are generated, however, they are still not human-readable. Thus the second sub-problem, event-to-sentence, focuses on transforming these events into natural language sentences.

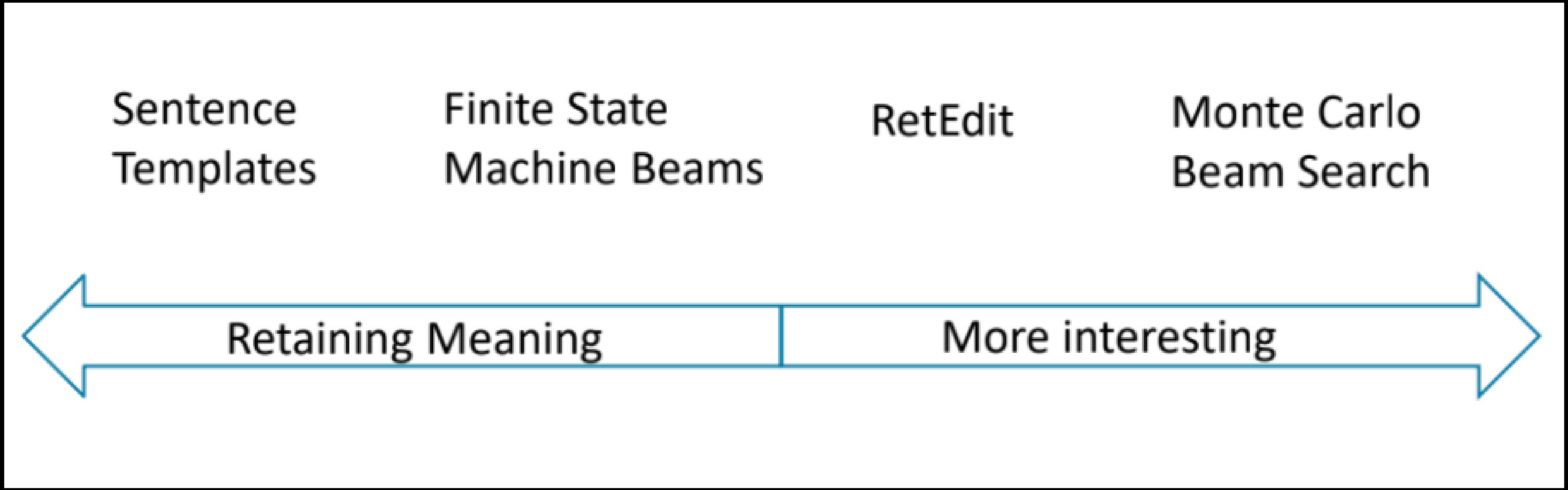
Overview

- Extract events from stories
- Generate the plot using a seq2seq network
- Use an ensemble of methods to find the best sentence given an event
- Get a confidence score from each model, and accept the sentence if it's above a threshold



Ammanabrolu, P., Tien, E., Cheung, W., Luo, Z., Ma, W., Martin, L. J., & Riedl, M. O. (2020). Story Realization: Expanding Plot Events into Sentences. *AAAI Conference on Artificial Intelligence (AAAI)*, *34*(5), 7375–7382. https://ojs.aaai.org//index.php/AAAI/article/view/6232

Balance



Ammanabrolu, P., Tien, E., Cheung, W., Luo, Z., Ma, W., Martin, L. J., & Riedl, M. O. (2020). Story Realization: Expanding Plot Events into Sentences. *AAAI Conference on Artificial Intelligence (AAAI)*, *34*(5), 7375–7382. https://ojs.aaai.org//index.php/AAAI/article/view/6232

Cascading Ensemble

- RetEdit
- Sentence Templates
- Monte Carlo Beam Search
- Finite State Machine Constrained Beams
- Seq2Seq

Ammanabrolu, P., Tien, E., Cheung, W., Luo, Z., Ma, W., Martin, L. J., & Riedl, M. O. (2020). Story Realization: Expanding Plot Events into Sentences. *AAAI Conference on Artificial Intelligence (AAAI)*, *34*(5), 7375–7382. https://ojs.aaai.org//index.php/AAAI/article/view/6232

RetEdit

- Retrieve: Map event to its closest sentence and retrieve that
- Edit:
  Seq2seq with attention and copying (Gu et al., 2016)
  Takes the retrieved sentence and the original input event, then edits
- Confidence score: proportional to 1 – retrieval distance

Ammanabrolu, P., Tien, E., Cheung, W., Luo, Z., Ma, W., Martin, L. J., & Riedl, M. O. (2020). Story Realization: Expanding Plot Events into Sentences. *AAAI Conference on Artificial Intelligence (AAAI)*, *34*(5), 7375–7382. https://ojs.aaai.org//index.php/AAAI/article/view/6232

Sentence Templates

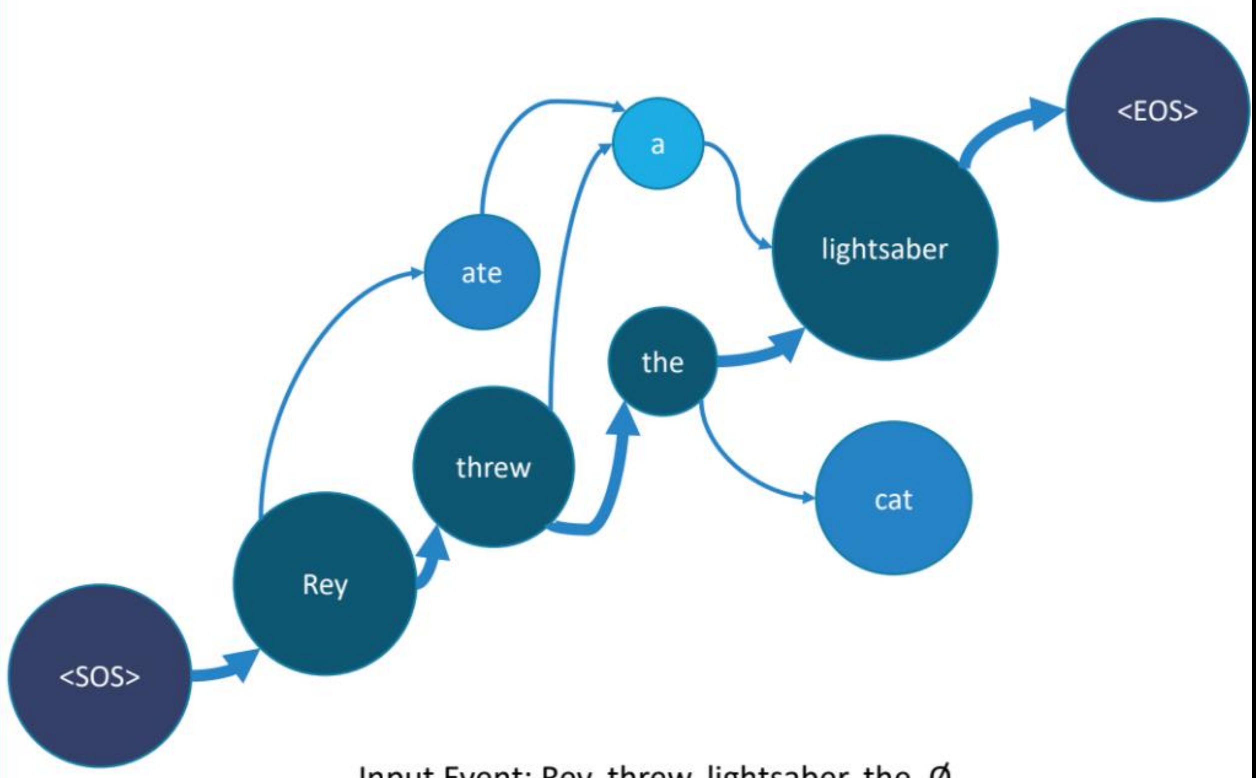
$$S \rightarrow NP \quad v \quad (NP) \quad (PP)$$
$$NP \rightarrow d \quad n$$
$$PP \rightarrow p \quad NP$$

$$[\_\_ \, s]\{v \, [\_\_ \, o] \, [p \, \_\_ \, m]\}$$

Confidence score: 1 - $\frac{\sum loss}{sentence\ length}$

Ammanabrolu, P., Tien, E., Cheung, W., Luo, Z., Ma, W., Martin, L. J., & Riedl, M. O. (2020). Story Realization: Expanding Plot Events into Sentences. *AAAI Conference on Artificial Intelligence (AAAI)*, *34*(5), 7375–7382. https://ojs.aaai.org//index.php/AAAI/article/view/6232

Monte Carlo Beam Search



Input Event: Rey, threw, lightsaber, the, Ø

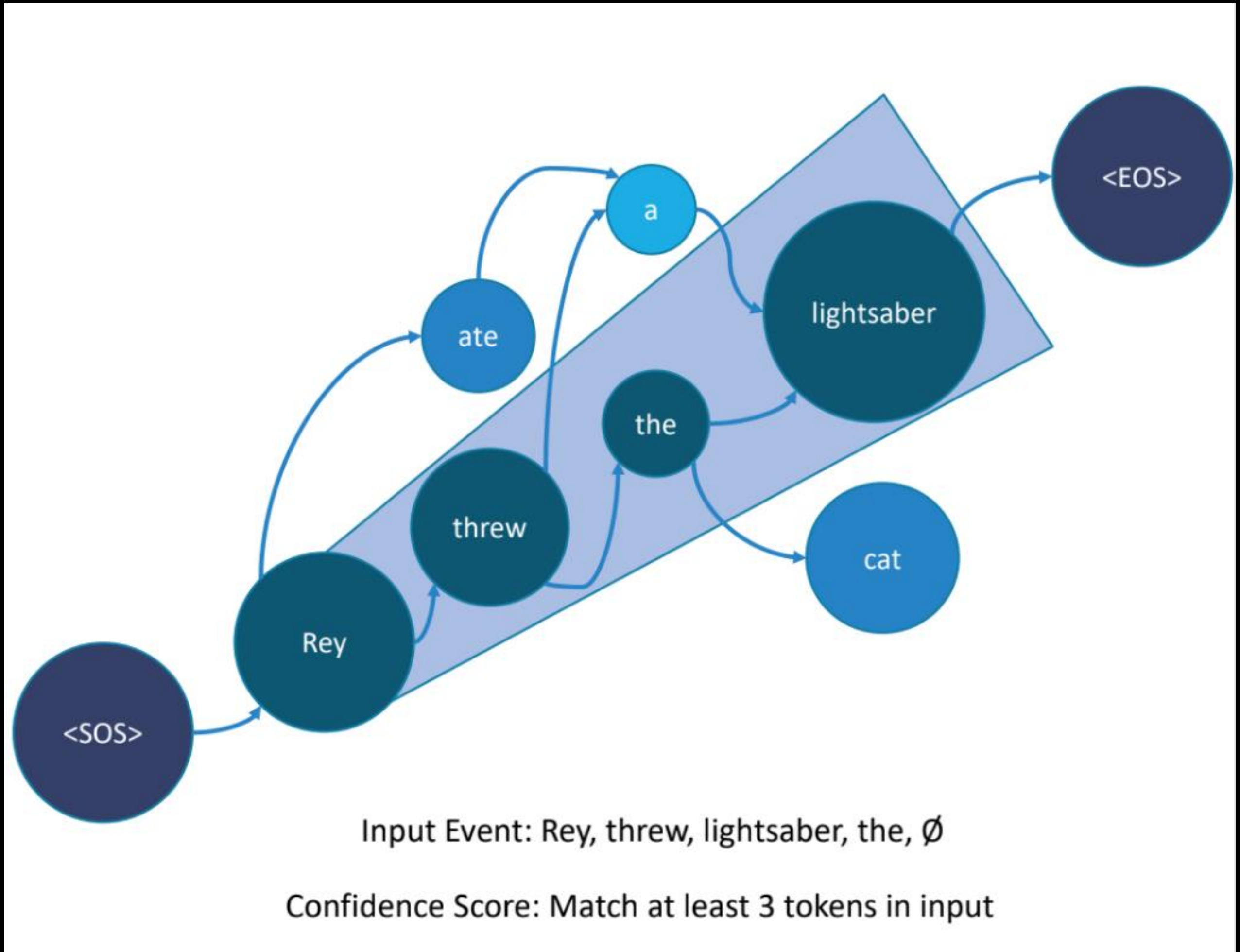
Confidence Score: $s_t = \alpha * s_{t-1} + (1 - \alpha) * AVG(playout_t)$

Beams weighted to favor tokens in input

Weights learned at validation

Ammanabrolu, P., Tien, E., Cheung, W., Luo, Z., Ma, W., Martin, L. J., & Riedl, M. O. (2020). Story Realization: Expanding Plot Events into Sentences. *AAAI Conference on Artificial Intelligence (AAAI)*, *34*(5), 7375–7382. https://ojs.aaai.org//index.php/AAAI/article/view/6232

Finite State Machine Constrained Beams



Ammanabrolu, P., Tien, E., Cheung, W., Luo, Z., Ma, W., Martin, L. J., & Riedl, M. O. (2020). Story Realization: Expanding Plot Events into Sentences. *AAAI Conference on Artificial Intelligence (AAAI)*, *34*(5), 7375–7382. https://ojs.aaai.org//index.php/AAAI/article/view/6232

Examples

Table 1: Event-to-sentence examples for each model. ∅ represents an empty parameter; <PRP> is a pronoun.

Input Event	**RetEdit**	**Templates**	**Monte Carlo**	**FSM**	**Gold Standard**
⟨<PRP>, act-114-1-1, to, ∅, event.n.01⟩	<PRP> and <PERSON>0 move to the event.n.01 of the natural_object.n.01.	<PRP> act-114-1-1 to event.n.01.	<PRP> moves to the nearest natural_object.n.01.	physical_entity.n.01 move back to the phenomenon.n.01 of the craft.n.02...	<PRP> move to the event.n.01.
⟨<PERSON>2, send-11.1, through, <PERSON>6, <LOCATION>1⟩	<PERSON>2 sends <PERSON>6 through the <LOCATION>1.	The <PERSON>2 send-11.1 the <PERSON>6 through <LOCATION>1.	<PERSON>2 passes this undercover in the body_part.n.01 and collapses.	∅	In activity.n.01 to avoid <PRP> out.n.01 <PERSON>2 would transport <PERSON>6 through the <LOCATION>1.

Ammanabrolu, P., Tien, E., Cheung, W., Luo, Z., Ma, W., Martin, L. J., & Riedl, M. O. (2020). Story Realization: Expanding Plot Events into Sentences. *AAAI Conference on Artificial Intelligence (AAAI)*, *34*(5), 7375–7382. https://ojs.aaai.org//index.php/AAAI/article/view/6232

End-to-End Examples

Table 2: End-to-end pipeline examples on previously-unseen input data. The Event-to-Sentence model used is the full ensemble. Sentences are generated using both the extracted and generated events.

Input Sent.	**Extracted event**	**Generated Events (Event-to-Event)**	**Generated Sentences (Event-to-Sentence)**	**Slot-filled Sentences**
On Tatooine, Jabba the Hutt inspects the drone barge recently delivered to him.	⟨<ORG>0, assessment-34.1, ∅, vessel.n.02, ∅⟩	⟨<PERSON>1, settle-36.1.2, ∅, indicator.n.03, indicator.n.03 ⟩ ; ⟨music.n.01, escape-51.1-1, from, ∅, ∅⟩; ⟨<PRP>, discover-84, to, run-51.3.2, progenitor.n.01⟩	The <ORG>0 can not scan the vessel.n.02 of the <VESSEL>0. <PERSON>1 decides to be a little person.n.01 at the structure.n.01. the music.n.01 arrives. <PRP> finds a lonely person.n.01 on the upper one of the craft.n.02 which is not a personal_letter.n.01 but does not respond to hails .	The Jabba the Hutt can not scan the bareboat of the Uss Lakota. O Yani decides to be a little mailer at the airdock. The Music arrives. She finds a lonely mailer on the upper one of the bareboat which is not a love letter but does not respond to hails.
Boba Fett has just chased down another bounty, a Rodian art dealer who sold fake works to Gebbu the Hutt.	⟨<PERSON>0, chase-51.6, ∅, bounty.n.04, ∅⟩	⟨<PERSON>0, chase-51.6, to, magnitude.n.01, ∅⟩; ⟨magnitude.n.01, comprehend-87.2, off, craft.n.02, magnitude.n.01⟩; ⟨<PERSON>2, amuse-31.1, off, ∅, ∅⟩; ⟨<PERSON>2, discover-84, off, change_of_integrity.n.01, ∅⟩	<PERSON>0 enters the bounty.n.04 and tells <PRP>. <PERSON>0 attaches the explosive.a.01 to the person.n.01 who is trying to fix the device.n.01. the magnitude.n.01 doesn't know the craft.n.02 off the craft.n.02. <PERSON>2 is surprised when <PRP> learns that the person.n.01 is actually <PERSON>7. <PERSON>2 sees the change_of_integrity.n.01 and tells <PRP>.	Boba Fett enters the bounty and tells it. Boba Fett attaches the explosive to the peer who is trying to fix the toy. The multiplicity doesn't know the bounty off the bounty. Dark Jedi Lomi Plo is surprised when it learns that the peer is actually Mrs Conners. Dark Jedi Lomi Plo sees the combination off the Orbs and tells them.

Ammanabrolu, P., Tien, E., Cheung, W., Luo, Z., Ma, W., Martin, L. J., & Riedl, M. O. (2020). Story Realization: Expanding Plot Events into Sentences. *AAAI Conference on Artificial Intelligence (AAAI)*, *34*(5), 7375–7382. https://ojs.aaai.org//index.php/AAAI/article/view/6232

Proceedings of the Twenty-Eighth International Joint Conference on Artificial Intelligence (IJCAI-19)

Controllable Neural Story Plot Generation via Reward Shaping

Pradyumna Tambwekar[1*], **Murtaza Dhuliawala**[1*], **Lara J. Martin**[1], **Animesh Mehta**[1],
Brent Harrison[2] and **Mark O. Riedl**[1]
[1]School of Interactive Computing, Georgia Institute of Technology
[2]Department of Computer Science, University of Kentucky
{ptambwekar3, murtaza.d.210, ljma
harrison@cs.uky.edu,

Abstract

Language-modeling–based approaches to story plot generation attempt to construct a plot by sampling from a language model (LM) to predict the next character, word, or sentence to add to the story. LM techniques lack the ability to receive guidance from the user to achieve a specific goal, resulting in stories that don't have a clear sense of progression and lack coherence. We present a reward-shaping technique that analyzes a story corpus and produces intermediate rewards that are backpropagated into a pre-trained LM in order to guide the model towards a given goal. Automated evaluations show our technique can create a model that generates story plots which consistently achieve a specified goal. Human-subject studies show that the generated stories have more plausible event ordering than baseline plot generation techniques.

1 Introduction

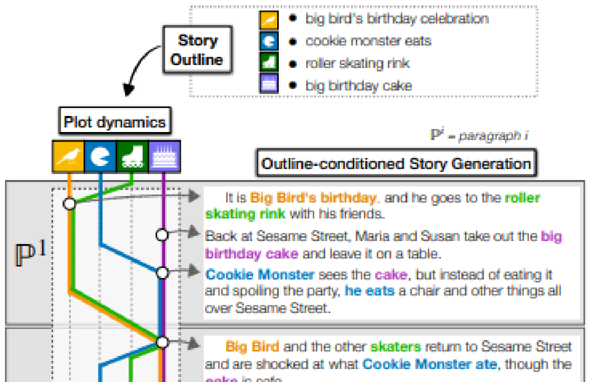
Automated plot generation is the problem of creating a sequence of main plot points for a story in a given domain and

PlotMachines: Outline-Conditioned Generation with Dynamic Plot State Tracking

Hannah Rashkin[1], **Asli Celikyilmaz**[2], **Yejin Choi**[1,3], **Jianfeng Gao**[2]
[1] Paul G. Allen School of Computer Science & Engineering, University of Washington
[2] Microsoft Research, Redmond, WA, USA
[3] Allen Institute for Artificial Intelligence, Seattle, WA, USA
{hrashkin,yejin}@cs.washington.edu, {aslicel,jfgao}@microsoft.com

Abstract

We propose the task of *outline-conditioned story generation*: given an outline as a set of phrases that describe key characters and events to appear in a story, the task is to generate a coherent narrative that is consistent with the provided outline. This task is challenging as the input only provides a rough sketch of the plot, and thus, models need to generate a story by interweaving the key points provided in the outline. This requires the model to keep track



How do all of these guided systems differ?

When poll is active respond at **PollEv.com/laramartin527**

Send **laramartin527** and your message to **22333**

How are all they similar?

Think-Pair-Share

We saw how Plan & Write has evolved.
How might some of these other techniques differ when using transformers (if at all)?

Guided Open Story Generation Using Probabilistic Graphical Models

Sagar Gandhi*
University of Kentucky
Lexington, Kentucky
sga267@uky.edu

Brent Harrison
University of Kentucky
Lexington, Kentucky
harrison@cs.uky.edu

ABSTRACT

In this work, we present an approach for performing computational storytelling in open domain based on Author Goals. Author Goals are constraints placed on a story event directed by the author of the system. There are two challenges present in this type of story generation: (1) automatically acquiring a model of story progression, and (2) guiding the progress of story progression in light of different goals. We propose a novel approach to story generation based on probabilistic graphical models and Loopy Belief Propagation (LBP) that addresses both of these problems. We show the applicability of our technique through a case study on the Visual Storytelling (VIST) 2017 dataset. We use image descriptions as author goals. This empirical analysis suggests that our approach is able to utilize goals information to better automatically generate stories.

CCS CONCEPTS

• **Computing methodologies → Probabilistic reasoning**; *Discourse, dialogue and pragmatics*; *Natural language generation*; *Learning in probabilistic graphical models.*

problem *domain model* [3, 5, 25]. This type of story generation is sometimes called *Closed Story Generation*. While stories generated in th…

The Thirty-Fourth AAAI Conference on Artificial Intelligence (AAAI-20)

Story Realization: Expanding Plot Events into Sentences

Prithviraj Ammanabrolu, Ethan Tien, Wesley Cheung,
Zhaochen Luo, William Ma, Lara J. Martin, Mark O. Riedl
School of Interactive Computing
Georgia Institute of Technology
{raj.ammanabrolu, etien, wcheung8, zluo, wma61, ljmartin, riedl}@gatech.edu

Abstract

Neural network based approaches to automated story plot

for explicit domain modeling beyond providing a corpus of example stories. The primary pitfall of neural language model approaches for story generation is that the space of

PLOTMACHINES: Outline-Conditioned Generation with Dynamic Plot State Tracking

Hannah Rashkin[1], Asli Celikyilmaz[2], Yejin Choi[1,3], Jianfeng Gao[2]
[1] Paul G. Allen School of Computer Science & Engineering, University of Washington
[2] Microsoft Research, Redmond, WA, USA
[3] Allen Institute for Artificial Intelligence, Seattle, WA, USA
{hrashkin,yejin}@cs.washington.edu, {aslicel,jfgao}@microsoft.com

Abstract

We propose the task of *outline-conditioned story generation*: given an outline as a set of phrases that describe key characters and events to appear in a story, the task is to generate a coherent narrative that is consistent with the provided outline. This task is challenging as the input only provides a rough sketch of the plot, and thus, models need to generate a story by interweaving the key points provided in the outline. This requires the model to keep track of the dynamic states of the latent plot, conditioning on the input outline while generating the full story. We present PLOTMACHINES, a neural narrative model that learns to transform an outline into a coherent story by tracking the dynamic plot states. In addition, we enrich PLOTMACHINES with high-level discourse structure so that the model can learn different writing styles corresponding to different parts of the narrative. Comprehensive experiments over three fiction and non-fiction

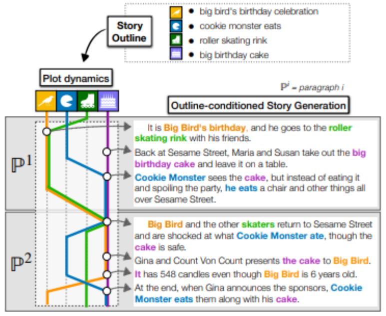


Figure 1: An outline (input) paired with a story (output) from the Wikiplots training set. Plot elements from the outline can appear and reappear non-linearly throughout the plot, as shown in plot dynamics graph. Composing stories from an outline requires keeping track of how outline phrases have been used while writing.

The Story Cloze Test

What is a Cloze Test?

- Something is removed from a text; try to guess what's missing
- Used for reading comprehension, grammar, etc. (with humans)

Unsupervised Learning of Narrative Event Chains

Unsupervised Learning of Narrative Event Chains

Nathanael Chambers and **Dan Jurafsky**
Department of Computer Science
Stanford University
Stanford, CA 94305
{natec,jurafsky}@stanford.edu

Abstract

Hand-coded *scripts* were used in the 1970-80s as knowledge backbones that enabled inference and other NLP tasks requiring deep semantic knowledge. We propose unsupervised induction of similar schemata called *narrative event chains* from raw newswire text.

A narrative event chain is a partially ordered set of events related by a common protagonist. We describe a three step process to learning narrative event chains. The first uses unsupervised distributional methods to learn narrative relations between events sharing corefe

tate learning, and thus this paper addresses the three tasks of chain induction: *narrative event induction*, *temporal ordering of events* and *structured selection* (pruning the event space into discrete sets).

Learning these prototypical schematic sequences of events is important for rich understanding of text. Scripts were central to natural language understanding research in the 1970s and 1980s for proposed tasks such as summarization, coreference resolution and question answering. For example, Schank and Abelson (1977) proposed that understanding text about restaurants required knowledge about the ...staurant Script, including the partici... (Cus-

Narrative Cloze Test

- Evaluate “event relatedness”
- Find which events could be missing from a narrative chain
- Uses verbs only

N. Chambers and D. Jurafsky, “Unsupervised Learning of Narrative Event Chains,” in *Annual Meeting of the Association for Computational Linguistics: Human Language Technologies (ACL-HLT)*, 2008, pp. 789–797, doi: 10.1.1.143.1555.

Narrative Cloze Test

Known events:

(pleaded subj), (admits subj), (convicted obj)

Likely Events:

sentenced obj	0.89	indicted obj	0.74
paroled obj	0.76	fined obj	0.73
fired obj	0.75	denied subj	0.73

Figure 1: Three narrative events and the six most likely events to include in the same chain.

X pleaded _

X admits _

_ convicted X

N. Chambers and D. Jurafsky, "Unsupervised Learning of Narrative Event Chains," in *Annual Meeting of the Association for Computational Linguistics: Human Language Technologies (ACL-HLT)*, 2008, pp. 789–797, doi: 10.1.1.143.1555.

A Corpus and Cloze Evaluation for Deeper Understanding of Commonsense Stories

A Corpus and Cloze Evaluation for Deeper Understanding of Commonsense Stories

Nasrin Mostafazadeh[1], **Nathanael Chambers**[2], **Xiaodong He**[3], **Devi Parikh**[4], **Dhruv Batra**[4], **Lucy Vanderwende**[3], **Pushmeet Kohli**[3], **James Allen**[1,5]

1 University of Rochester, 2 United States Naval Academy, 3 Microsoft Research, 4 Virginia Tech, 5 The Institute for Human & Machine Cognition

{nasrinm, james}@cs.rochester.edu, nchamber@usna.edu, {parikh, dbatra}@vt.edu, {xiaohe, lucyv, pkohli}@microsoft.com

Abstract

Representation and learning of commonsense knowledge is one of the foundational problems in the quest to enable deep language understanding. This issue is particularly challenging for understanding casual and correlational relationships between events. While this topic has received a lot of interest in the NLP community, research has been hindered by the lack of a proper evaluation framework. This paper attempts to address this problem with a new framework for evaluating st

Recently, there has been a renewed interest in story and narrative understanding based on progress made in core NLP tasks. This ranges from generic story telling models to building systems which can compose meaningful stories in collaboration with humans (Swanson and Gordon, 2008). Perhaps the biggest challenge of story understanding (and story generation) is having commonsense knowledge for the interpretation of narrative events. The question is how to provide commonsense knowledge regarding daily events to machines.

Finish the story

Gina was worried the cookie dough in the tube would be gross.

She was very happy to find she was wrong.

The cookies from the tube were as good as from scratch.

Gina intended to only eat 2 cookies and save the rest.

A. Gina liked the cookies so much she ate them all in one sitting. ✓
B. Gina gave the cookies away at her church.

Mostafazadeh, N., Chambers, N., He, X., Parikh, D., Batra, D., Vanderwende, L., Kohli, P., & Allen, J. (2016). A Corpus and Cloze Evaluation for Deeper Understanding of Commonsense Stories. *Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies (NAACL-HLT)*, 839–849. http://www.aclweb.org/anthology/N16-1098

Story Cloze Test

- Predict/select the most likely story *ending*
  Given the first 4 sentences of the story
- Full sentences
- Multiple choice evaluation

Mostafazadeh, N., Chambers, N., He, X., Parikh, D., Batra, D., Vanderwende, L., Kohli, P., & Allen, J. (2016). A Corpus and Cloze Evaluation for Deeper Understanding of Commonsense Stories. *Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies (NAACL-HLT)*, 839–849. http://www.aclweb.org/anthology/N16-1098

An RNN-based Binary Classifier for the Story Cloze Test

Melissa Roemmele
Institute for Creative Technologies
University of Southern California
roemmele@ict.usc.edu

Sosuke Kobayashi*
Preferred Networks, Inc.
sosk@preferred.jp

Naoya Inoue
Tohoku University
naoya-i@ecei.tohoku.ac.jp

Andrew M. Gordon
Institute for Creative Technologies
University of Southern California
gordon@ict.usc.edu

Enhanced Story Representation by ConceptNet for Predicting Story Endings

Shanshan Huang
huangss_33@sjtu.edu.cn
Shanghai Jiao Tong University

Kenny Q. Zhu*
kzhu@cs.sjtu.edu.cn
Shanghai Jiao Tong University

Qianzi Liao
liaoqz@sjtu.edu.cn
Shanghai Jiao Tong University

Libin Shen
libin@leyantech.com
Leyan Tech

Yinggong Zhao
ygzhao@leyantech.com
Leyan Tech

ABSTRACT

Predicting endings f… machine commonsen… resentation of the sto… Pre-trained language… in this task by explo… dataset, instead of "u… we propose to improv… fying the sentences to… latent relationship be… enhanced sentence re… guage models, makes… the popular Story Clo… data.

CCS CONCEPTS

Toward Better Storylines with Sentence-Level Language Models

Daphne Ippolito*
daphnei@seas.upenn.edu

David Grangier
grangier@google.com

Douglas Eck
deck@google.com

Chris Callison-Burch
ccb@seas.upenn.edu

Abstract

We propose a sentence-level language model which selects the next sentence in a story from a finite set of fluent alternatives. Since it does not need to model fluency, the sentence-level language model can focus on longer range dependencies, which are crucial for multi-sentence coherence. Rather than dealing with individual words, our method treats the story so far as a list of pre-trained sentence embeddings and predicts an embedding for the next sentence, which is more efficient than predicting word embeddings. Notably this allows us to consider a large number of candidates for the next sentence during training. We demonstrate the effectiveness of our approach with state-of-the-art accuracy on the unsupervised Story Cloze task and with promising results on larger-scale next sentence prediction tasks.

quence of imag… roles (Liu et a… Our work is… than consideri… pose a model v… of context and… a large set of f… age pre-traine… 2019) to build… Given the em… of the story, c… embedding of… This task is… dependencies… words, which… our model onl… candidate sen… tinuation to th… and time to lea…

IEEE/ACM TRANSACTIONS ON AUDIO, SPEECH, AND LANGUAGE PROCESSING, VOL. 27, NO. 4, APRIL 2019 719

Story Ending Selection by Finding Hints From Pairwise Candidate Endings

Mantong Zhou, Minlie Huang, and Xiaoyan Zhu

…strong indica… …ory Cloze Test …ding compre… …andidate end… …sting methods …d that operate …text, therefore …te endings can …hich misleads …ress this issue, …sion by utiliz… …two candidate …feature vector …d then refines …the difference …se feature vec… …is regarded as …approach can …mprehension. …story compre…

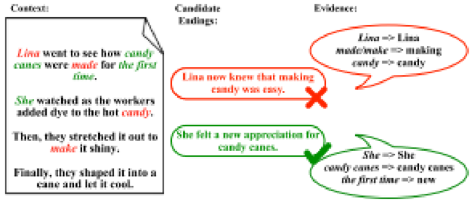


Fig. 1. **Evidence bias issue**: both a wrong ending (in red) and a correct ending (in green) can obtain sufficient evidence from the story context.

important linkages between a story context and a candidate ending. They suffer from the issue of **evidence bias**: both the wrong and correct endings can obtain sufficient support from the story context. As illustrated in Fig. 1, the wrong ending (in red) and the correct ending (in green) can be supported by the red-colored evidence and the green-colored evidence in the story context, respectively. Thus, it is difficult for matching-based models to distinguish such cases. The situation is not rare because both correct and wrong endings are written to fit the world of a story

Tackling the Story Ending Biases in The Story Cloze Test

Rishi Sharma[1], **James F. Allen**[1,2], **Omid Bakhshandeh**[3], **Nasrin Mostafazadeh**[4*]
1 University of Rochester, 2 Institute for Human and Machine Cognition, 3 Verneek.ai 4 Elemental Cognition
rishi.sharma@rochester.edu, nasrinm@cs.rochester.edu

Abstract

The Story Cloze Test (SCT) is a recent framework for evaluating story comprehension and script learning. There have been a variety of models tackling the SCT so far. Although the original goal behind the SCT was to require systems to perform deep language understanding and commonsense reasoning for successful narrative understanding, some recent models could perform significantly better than the initial baselines by leveraging human-authorship biases discovered in the SCT dataset. In order to shed some

this issue. This test evaluates a story comprehension system where the system is given a four-sentence short story as the 'context' and two alternative endings and to the story, labeled 'right ending' and 'wrong ending.' Then, the system's task is to choose the right ending. In order to support this task, Mostafazadeh et al. also provide the ROC Stories dataset, which is a collection of crowd-sourced complete five sentence stories through Amazon Mechanical Turk (MTurk). Each story follows a character through a fairly simple series of events to a conclusion.

Several shallow and neural models, including the state-of-the-art script learning approaches, were presented as baselines (Mostafazadeh et al.,

How could you use the Story Cloze Task for generation?

Class Stuff

Thursday’s Class

- Keerthi Sree Kopparaju - Reasoning about Goals, Steps, and Temporal Ordering with WikiHow
- Hanuma Sashank Samudrala - Enhanced Story Representation by ConceptNet for Predicting Story Endings
- Patty Delafuente - Controllable Neural Story Plot Generation via Reward Shaping
- Arya Honraopatil - Implicit Representations of Meaning in Neural Language Models
- Pooja Guttal - What do Large Language Models Learn about Scripts?

Slides are due 5pm tomorrow.
I will be using those slides on my computer. We will not be switching computers.
I will be keeping the presentations to a strict 8 minutes, so please practice ahead of time!!

Deadlines

- Project proposals are due **tonight at 11:59pm**
  https://laramartin.net/interactive-fiction-class/homeworks/project/project-proposal.html
- HW 2 due 10/7 at 11:59pm
  https://laramartin.net/interactive-fiction-class/homeworks/generating-descriptions/generating-descriptions.html