Neurosymbolic Automated Story Generation
Why is storytelling important?

Most natural way of communicating

What if computers could tell stories?
Alexa, tell us a spooky story about the Jersey Devil.
Plan with us

Bixby, what makes for a good party?

Most people seem to like parties where...
Siri, how do I bake cookies?

First, you’ll want to gather the ingredients...

Prepare us

Cortana, how do I ask this guy out?

The key to a man’s heart...
Hey Google, I just broke up with my boyfriend.

I’m so sorry to hear that. Let me tell you about the time...
Automated Story Generation

TEACHING COMPUTERS TO TELL STORIES
Examples

TALE-SPIN (1977):
One day,
JOE WAS THIRSTY.
JOE WANTED NOT TO BE THIRSTY.
JOE WANTED TO BE NEAR THE WATER.

CPOCL Planning Algorithm (2014):
You travel to the city.
You ask a knight to kill the sorcerer.
The knight buys a sharp sword at the market.
The knight travels to the tower.
The knight challenges the sorcerer to a fight to the death.
The sorcerer reveals that he is your father.
The knight defeats the sorcerer.
The prince travels to the city.
The king gives you a bag of gold.
The king makes you a knight.


Ware, S. et al. “Computational Model of Plan-Based Narrative Conflict at the Fabula Level.” IEEE Transactions on Computational Intelligence and AI in Games 2014.
Symbolic Systems

- Schemas
- Causality
- Planning
Symbolic Systems

Pharmacist_asks: prescription

- Stand in Line
- Customer orders drugs
- Pharmacist asks for prescription
- Customer produces prescription
- Pharmacist checks prescription
- Pharmacist delivers drugs
- Customer swipes credit card
- Customer takes receipt
- Customer leaves with drugs
Symbolic Systems

Causal Links

Have Prescription?

Pharmacist asks for prescription

Need Prescription

Customer produces prescription
Causal Chains

Stand in Line → Customer orders drugs → Pharmacist asks for prescription → Customer produces prescription → Pharmacist checks prescription → Pharmacist delivers drugs → Customer swipes credit card → Customer takes receipt → Customer leaves with drugs
Symbolic Systems

Stand in Line → Customer orders drugs → Pharmacist asks for prescription → Customer produces prescription → Pharmacist checks prescription → Pharmacist delivers drugs → Customer swipes credit card → Customer takes receipt → Customer leaves with drugs

Branching → Planning

Customer does not produce prescription → Pharmacist calls doctor’s office → ...

Symbolic Systems
Pros & Cons of **Symbolic Models**

+ Really coherent

- Lots of knowledge engineering
- Possibly lacking diversity & novelty
Closed vs Open World

All possible events that can occur in a story
Closed vs Open World

All possible events that can occur in a story
All possible thoughts a human can think of and express through language
Story Prompt
(First Sentence)

Rest of the Story
(about anything)
Neural Storytellers

a TON of stories

Neural Network

A story??
SUNSPRING
r 2 d 2 carrying some drinks on a tray strapped to his back passes yoda who uses his force powers to hog the drinks

Expected:

obi wan and anakin are drinking happily when chewbacca takes a polaroid picture of anakin and obi wan

Predicted:

can this block gives him the advantage to personally run around with a large stick of cheese
Comparison

**SYMBOLIC SYSTEMS**

+ Coherent stories
– Limited domain

**NEURAL SYSTEMS**

+ Open world stories
– Coherence is terrible

Pharmacist delivers drugs → Customer swipes card → Customer takes receipt → Customer leaves with drugs
I posit that...

The *perceived coherence* of stories produced by neural-based automated story generation systems can be improved by incorporating symbolic approaches—such as schemas, goals, and causal reasoning.
Outline

1. Events
2. Plot Progression
3. Reasoning
r 2 d 2 carrying some drinks on a tray strapped to his back passes yoda who uses his force powers to hog the drinks

**Expected:**

obi wan and anakin are drinking happily when chewbacca takes a polaroid picture of anakin and obi wan

**Predicted:**

can this block gives him the advantage to personally run around with a large stick of cheese
But what if the model was just bigger?
Huge Language Model (GPT-2)

The hungry dog licked her lips as she watched her owner eat.
The hungry dog licked her lips as she watched her owner eat.
She could see the white teeth in his jaw, then moving behind his right ear.

The hunter picked up his sword, directing the dogs in a cloud of mutterings. "Did you just feed a dog?" she demanded, suddenly even more agitated than before.

"Yes." Jack's voice was rough, and he hated that word. His face was tight and angry. "She has killed something important to us, and now the house is filled with blood. It makes me sad. She is one of the most dangerous wolves I have ever met, and if she bites me she will lose."

John shrugged. "Well, it's not your fault," he said.
Why is this happening?!

r2d2 carrying some drinks on a tray strapped to his back passes yoda who uses his force powers to hog the drinks can this block gives him the advantage to personally run around with a large stick of cheese

Problem: Sparsity!

Solution: Separating semantics (meaning) from syntax (grammar)

Too unique!

More overlap!
Event Representations

Use linguistic knowledge to bootstrap the neural network
From sentence, extract event representation
(subject, verb, direct object, modifier)

Original sentence: yoda uses the force to take apart the platform
Event: yoda use force Ø
Generalized Event: <PERSON>0 fit-54.3 power.n.01 Ø

Introducing ASTER

AUTOMATED STORY-TELLING WITH EVENT REPRESENTATIONS
John unwittingly unleashes an insidious pox.

John unleash pox Ø

ASTER Pipeline

<PERSON>0 disassemble-23.3 contagious_disease.n.01 Ø

John unleash pox Ø

<PERSON>0 disassemble-23.3 contagious_disease.n.01 Ø

male.n.02

spatial_configuration-47.6

Ø adopt-93

# Event2Event (Seq2Seq Baseline)

<table>
<thead>
<tr>
<th>Experiment</th>
<th>Perplexity</th>
</tr>
</thead>
<tbody>
<tr>
<td>(0) Original Sentences</td>
<td>704.815</td>
</tr>
<tr>
<td>(1) Original Words Baseline</td>
<td>748.914</td>
</tr>
<tr>
<td>(2) Original Words with PERSONs</td>
<td>166.646</td>
</tr>
<tr>
<td>(3) Generalized Baseline</td>
<td>54.231</td>
</tr>
</tbody>
</table>

Perplexity = \(2^{-\sum_x p(x) \log_2 p(x)}\)

\[ p(x) = \frac{\text{count}(x)}{\sum_{y \in Y} \text{count}(y)} \]

---

John unwittingly unleashes an insidious pox.

John unleash pox Ø

<PERS0N>0 disassemble-23.3 contagious_disease.n.01 Ø

male.n.02 spatial_configuration-47.6 Ø adopt-93

<PERSON>0 crumples and is about to be sheath.n.02

---

## Event2Sentence

<table>
<thead>
<tr>
<th>Experiment</th>
<th>Perplexity</th>
<th>BLEU</th>
</tr>
</thead>
<tbody>
<tr>
<td>Original Words Event → Original Sentence</td>
<td>1585.46</td>
<td>0.0016</td>
</tr>
<tr>
<td>Generalized Event → Generalized Sentence</td>
<td>56.516</td>
<td>0.0331</td>
</tr>
<tr>
<td>All Generalized Events → Generalized Sentence</td>
<td>59.106</td>
<td>0.0366</td>
</tr>
</tbody>
</table>

Precision using n-grams

---

John unwittingly unleashes an insidious pox.

He crumples and is about to be husk.

John unleash pox

He crumples and is about to be sheath.

---

Summary

Writers like to be unique with their writings
This makes it hard for ML to see patterns
I created a schema to aid in the pattern-matching
This created a giant system that still rambles
Outline

1. Events
2. Plot Progression
3. Reasoning
Global Coherence

Meet
Plot Point 1

Admire
Mid Point

Unrequited

Pinch 1

Resolution
Marry

Pinch 2
Discovery

Hook

Plot
Point 2

Understanding

image source: https://blog.reedsy.com/plot-point/
ASTER Pipeline

- `sentence_n` → **Eventify** → `event_n` → **Event-to-Event** → `event_{n+1}`
- `sentence_{n+1}` → **Working & Long-Term Memory** → **Slot Filler** → `generalized_sentence_{n+1}` → **Event-to-Sentence**
Policy Gradient DRL (REINFORCE)

Tambwekar, P.*, Dhuliawala, M.*, Martin, L. J., Mehta, A., Harrison, B., & Riedl, M. O.
#1 Verb Distance

\[ r_1(v) = \log \sum_{s \in S_{v,g}} l_s - d_s(v, g) \]

#2 Story-Verb Frequency

\[ r_2(v) = \log \frac{k_{v,g}}{N_v} \]

Appear frequently before goal = Higher reward

Tambwekar, P.*, Dhuliawala, M.*, Martin, L. J., Mehta, A., Harrison, B., & Riedl, M. O.
Final Reward Equation

\[ R(v) = \alpha \times r_1(v) \times r_2(v) \]

Verb Distance to Goal

Story-Verb Frequency

What now?

1. Cluster based on reward score
2. Constrain system to sample from next cluster

## Policy Gradient Results

<table>
<thead>
<tr>
<th>Goal</th>
<th>Model</th>
<th>Goal Achievement Rate</th>
<th>Average Perplexity</th>
<th>Average Story Length</th>
</tr>
</thead>
<tbody>
<tr>
<td>admire</td>
<td>Test Corpus</td>
<td>20.30%</td>
<td>n/a</td>
<td>7.59</td>
</tr>
<tr>
<td></td>
<td>Seq2Seq</td>
<td>35.52%</td>
<td>48.06</td>
<td>7.11</td>
</tr>
<tr>
<td></td>
<td>Policy Gradient DRL</td>
<td><strong>94.29%</strong></td>
<td><strong>7.61</strong></td>
<td><strong>4.90</strong></td>
</tr>
<tr>
<td>marry</td>
<td>Test Corpus</td>
<td>24.64%</td>
<td>n/a</td>
<td>7.37</td>
</tr>
<tr>
<td></td>
<td>Seq2Seq</td>
<td>39.92%</td>
<td>48.06</td>
<td>6.94</td>
</tr>
<tr>
<td></td>
<td>Policy Gradient DRL</td>
<td><strong>93.35%</strong></td>
<td><strong>7.05</strong></td>
<td><strong>5.76</strong></td>
</tr>
</tbody>
</table>

But are the stories actually any good?
Methods

175 Mechanical Turkers rated statements on a 5-point Likert scale

For each of 3 conditions:
- Policy Gradient DRL
- Baseline Seq2Seq
- Testing Set Stories (Gold Standard)

Tambwekar, P.*, Dhuliawala, M.*, Martin, L. J., Mehta, A., Harrison, B., & Riedl, M. O.
Questionnaire

1. This story exhibits CORRECT GRAMMAR.
2. This story's events occur in a PLAUSIBLE ORDER.
3. This story's sentences MAKE SENSE given sentences before and after them.
4. This story FOLLOWS A SINGLE PLOT.
5. This story AVOIDS REPETITION.
6. This story uses INTERESTING LANGUAGE.
7. This story is of HIGH QUALITY.
8. This story REMINDS ME OF A SOAP OPERA.
9. This story is ENJOYABLE.

Coherence
Tambwekar, P.*, Dhuliawala, M.*, Martin, L. J., Mehta, A., Harrison, B., & Riedl, M. O.
So far...

I have created a schema so that semantic events can be generated separately from syntax.

This created more semantically-accurate plot generation but it would still lose coherence over time.

So I created a way to finetune the event generator to behave as RL and created artificial states through reward clustering.

This made for more consistent, plot-driven generation.
But the stories still aren’t causally coherent...
Example (Goal: hate/admire)

Our sister died.
Greggory executed during the visit.
Greggory adopted the girl.
The girl looked like her mom.
She was appalled.
Penelope detested the jungle gym.
Outline

1. Events
2. Plot Progression
3. Reasoning
Remember Causal Chains?

1. Stand in Line
2. Customer orders drugs
3. Pharmacist asks for prescription
4. Customer produces prescription
5. Pharmacist checks prescription
6. Pharmacist delivers drugs
7. Customer swipes credit card
8. Customer takes receipt
9. Customer leaves with drugs
Using VerbNet

**Jen** sent the **book** to **Remy** from **Atlanta**.

**ROLES**  
Agent  
Theme  
Destination  
Initial_Location

- has_location(e1, book, Atlanta)
- do(e2, Jen)
- cause(e2, e3)
- motion(e3, book)
- !has_location(e3, book, Atlanta)
- has_location(e4, book, Remy)

**PREDICATES**

**SELECTIONAL RESTRICTIONS**

Initial_Location : location  
Theme : concrete  
Agent : animate or organization
Using VerbNet

**Jen** sent the **book** to **Remy** from **Atlanta**.

**Pre-Conditions**
- `has_location(e1, book, Atlanta)`
- `Atlanta : location`
- `book : concrete`
- `Jen : animate or organization`

**Post-Conditions**
- `do(e2, Jen)`
- `cause(e2, e3)`
- `motion(e3, book)`
- `!has_location(e3, book, Atlanta)`
- `has_location(e4, book, Remy)`
Then VerbNet can model the story world!
ASTER Pipeline

- **Sentence$_n$** → **Eventify** → **Event$_n$** → **Event-to-Event** → **Event$_{n+1}$** → **Event-to-Sentence**
- **Sentence$_{n+1}$** → **Slot Filler** → **generalized_sentence$_{n+1}$** → **Sentence$_{n+1}$**
Event-to-Event

$\text{event}_n \rightarrow \text{Event-to-Event} \rightarrow \text{event}_{n+1}$
Event-to-Event

$\text{event}_n \rightarrow \text{Event-to-Event} \rightarrow \text{event}_n$

$\text{event}_n \rightarrow \text{Event-to-Event} \rightarrow \text{event}_{n+1}$
Event - to - Event (Genre Model)
ASTER-X Example (Hand-Translated)

ASTER-X (NEUROSYMBOLIC)
The traveler succeeded.
The traveler materialized the Voyager.
The traveler vaporized the Voyager.
Evelyn sought the Voyager to Paul.
What found the farewell order to the Voyager?
Wendy demanded to judge for the vote.
Wendy asked the vote up Kevin.
Ruby consulted the draft to the Voyager.
The Voyager inquired on the refusal on the draft.
Evelyn aided the draft into The Pacific Command.
The traveler waited.
The traveler knelt inside the order.
The traveler plotted to return.
Kevin moved to escape its deck.

ASTER (NEURAL)
The traveler succeeded.
The tape died.
The tape repeated.
The effect authorized the tape.
The tape burned.
The tape expelled the starboard.
The starboard continued.
The starboard confessed.
The starboard depicted the builder.
The builder condense the Voyager to a particle.
The mutation was accomplished.
Richard pained to achieve the irreversible process.
Richard revealed the radio beam.
How does this compare to symbolic-only systems?
ASTER-X: ASTER (Symbolic) eXtension

**Event**

- event\(_n\) → Event-to-Event (Genre Model) → candidate event\(_{n+1}\)s

- World Engine (VerbNet)
  - pruned candidate event\(_{n+1}\)s

- Event Selector
  - event\(_{n+1}\)

- story state

---

**References**

Symbolic-Only System Baseline

NP V PP: <PERSON>0 leave-51 for <LOCATION>4

World Engine (VerbNet)

Random VerbNet frames

candidate "event"_{n+1}s

Event Selector

pruned candidate event_{n+1}s

story state

event_{n}
event_{n+1}
Questionnaire

1. This story exhibits CORRECT GRAMMAR.
2. This story's events occur in a PLAUSIBLE ORDER.
3. This story's sentences MAKE SENSE given sentences before and after them.
4. This story AVOIDS REPETITION.
5. This story uses INTERESTING LANGUAGE.
6. This story is of HIGH QUALITY.
7. This story is ENJOYABLE.
8. This story REMINDS ME OF A SPACE OPERA.
9. This story FOLLOWS A SINGLE PLOT.
ASTER vs ASTER-X vs Symbolic

- * p < 0.05
- ** p < 0.01
While difficult for me to understand, the story does seem to follow in plausible order.

—A Turk, about the Symbolic-only system
Using VerbNet for Syntax

Improved Readability

**Example:**
The horse ran to the barn.

**Syntax:**

- **Theme** **VERB** {to, towards, into} **Destination**

**Semantics:**

- `HAS_LOCATION(e1, Theme, ?Initial_Location)`
- `MOTION(e2, Theme, ?Trajectory)`
- `-HAS_LOCATION(e2, Theme, ?Initial_Location)`
- `HAS_LOCATION(e3, Theme, Destination)`

**Force Dynamics:**
None
To help the neurosymbolic system with **readability**, what if my event representation was a full sentence?
ASTER-XT: ASTER-X with Transformers

- event \(_n\)
- Finetuned GPT-2 (Genre Model)
- candidate sent \(_{n+1}\)
- Parser
- World Engine (VerbNet)
- story state
- pruned candidate sent \(_{n+1}\)
- Sentence Selector
- unfilled sentence \(_{n+1}\)

ASTER-XT: ASTER-X with Transformers

- Sentence $n$: Input to Finetuned GPT-2 (Genre Model)
  - Output: Candidate sentences $n+1$ (candidate sent$_{n+1}$s)
- Sentence $n+1$: Input to Sentence Selector
  - Output: Unfilled sentence $n+1$ (unfilled sent$_{n+1}$)
- Parser
- World Engine (VerbNet)
  - Output: Pruned candidate sentences $n+1$ (pruned sent$_{n+1}$s)
- Working & Long-Term Memory
- Slot Filler
ASTER-XT Example

A little boy lies awake in the dark as he hears scratching noises on the window.

GPT-2
He is Chekov.
He is lying awake.
The force of his own voice is heard.
He is awake in the Melllvar Epran.
He is in Resh's body.

ASTER-XT
Koroth and Nona find him.
The man explains he and Nona were supposed to be watching a recording of the Breen, and that they were supposed to go to the local Germantown.
Suddenly, Nona hears the basement door open and is forced to call out to the woman.
The woman explains that they found nothing.
Nona's holo were missing.
Recap

Storytelling systems are important!

Separate semantics & syntax

Controllability

Hybrid system for reasoning & maintaining state

Language generation can be longer & can concentrate on the meaning throughout!

Ability to solve long-term plans! (Also, coherence)

Generative systems that are consistent! (COHERENCE)
The perceived coherence of stories produced by neural-based automated story generation systems can be improved by incorporating symbolic approaches—such as schemas, goals, and causal reasoning.
making sense is most important because this does not make any sense. I have nothing more to say here. Have your next story make sense.

“–A frustrated Turk
Thank you!