Neurosymbolic Automated Story Generation

CIS-700 INTERACTIVE FICTION AND TEXT GENERATION

MODULE 6 – 4/7/2022

DR. LARA J. MARTIN

Why is storytelling important?

Most natural way of communicating

What if computers could tell stories?

Entertain us

Photo by Mike Erskine on Unsplash

Plan with us

Bixby, what makes for a good party?

Most people seem to like parties where…

Photo by Glenn Carstens-Peters on Unsplash

Teach us

Siri, how do I bake cookies?

First, you'll want to gather the ingredients…

https://www.bettycrocker.com/how-to/tipslibrary/baking-tips/baking-with-kids

Prepare us

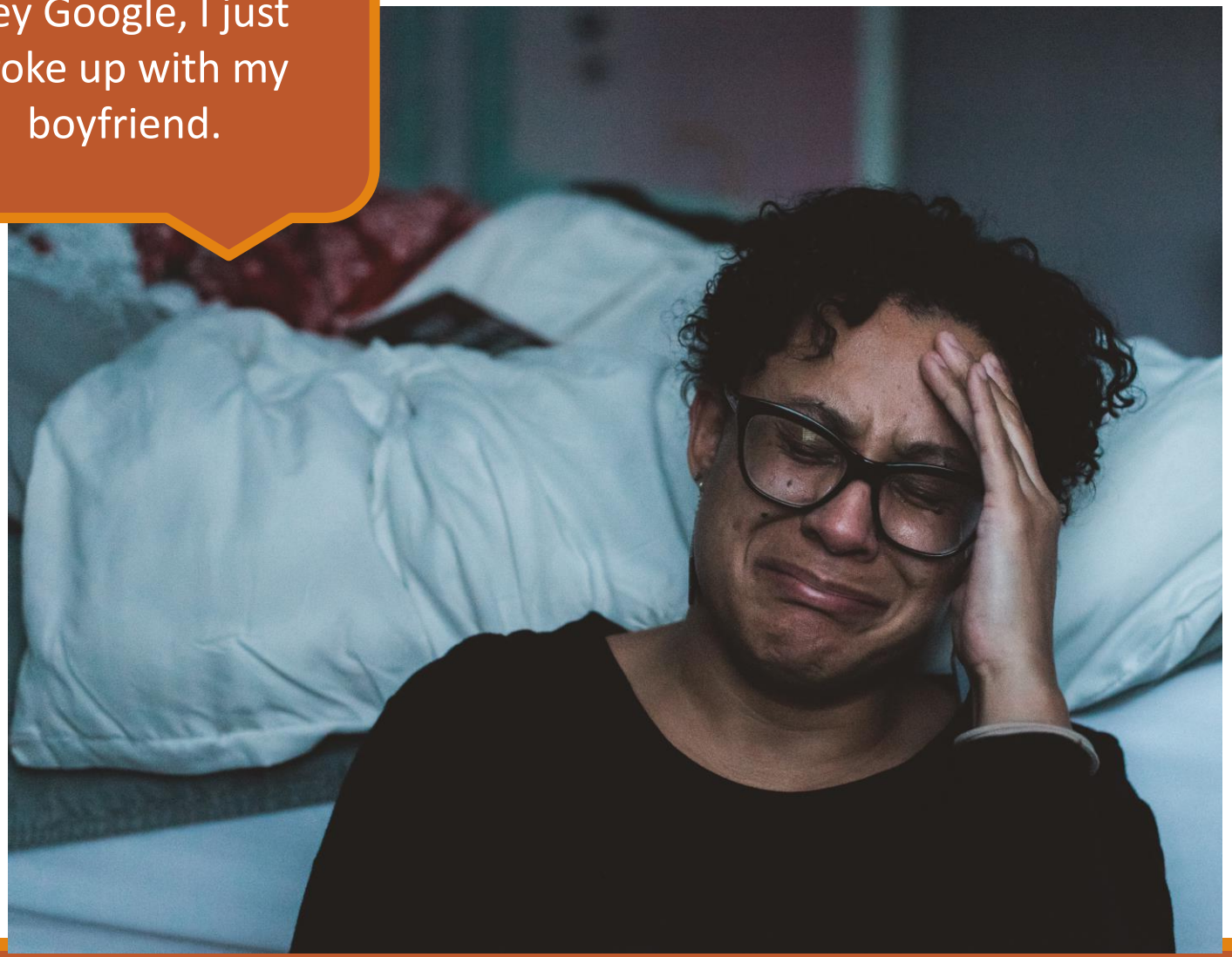
Cortana, how do I ask this guy out?

The key to a man’s heart…

Photo by Verena Yunita Yapi on Unsplash

Support us

Hey Google, I just broke up with my boyfriend.



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

Photo by Claudia Wolff on Unsplash

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):

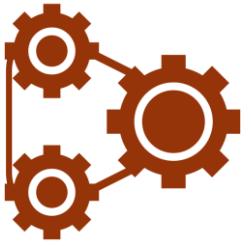
You travel to the city.
You ask a knight to kill the sorcerer.
The knight buys a sharp sword at the market.
The knight travels to the tower.
The knight challenges the sorcerer to a fight to the death.
The sorcerer reveals that he is your father.
The knight defeats the sorcerer.
The prince travels to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

Meehan, J.R. "TALE-SPIN, an Interactive Program that Writes Stories." *IJCAI* 1977.

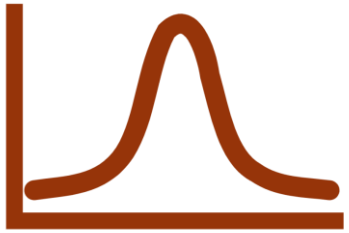
Ware, S. et al. "Computational Model of Plan-Based Narrative Conflict at the Fabula Level." *IEEE Transactions on Computational Intelligence and AI in Games* 2014..

Symbolic Systems

Schemas



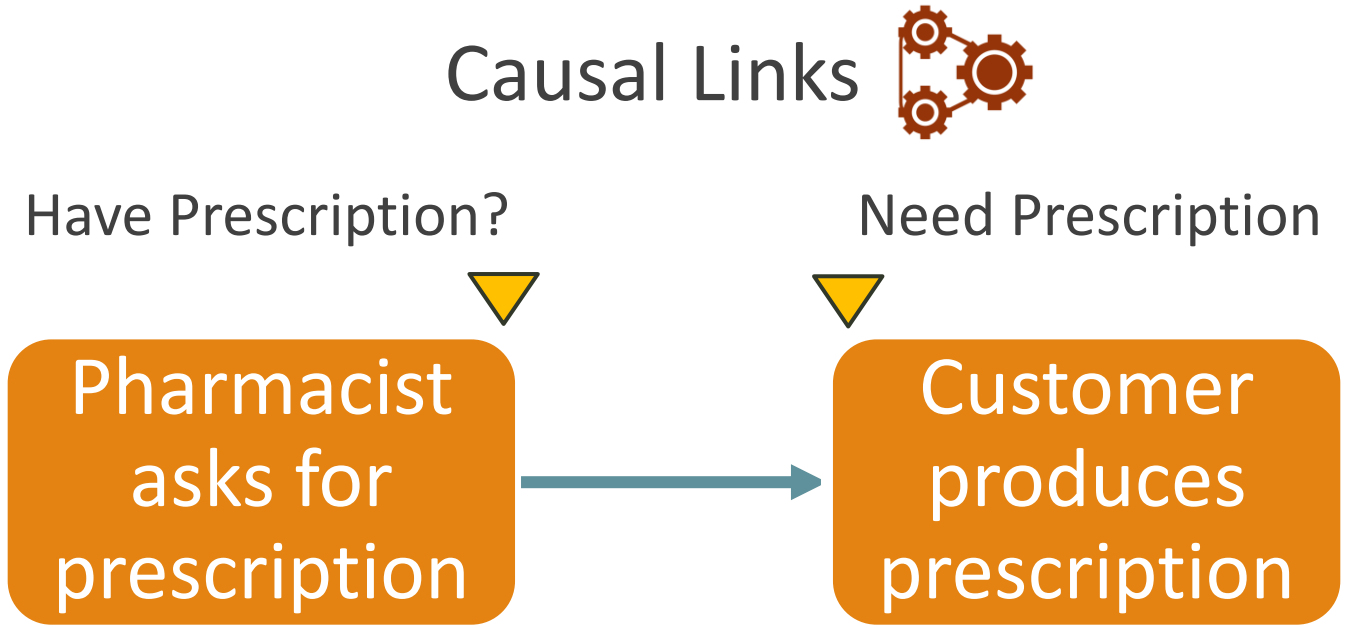
Causality



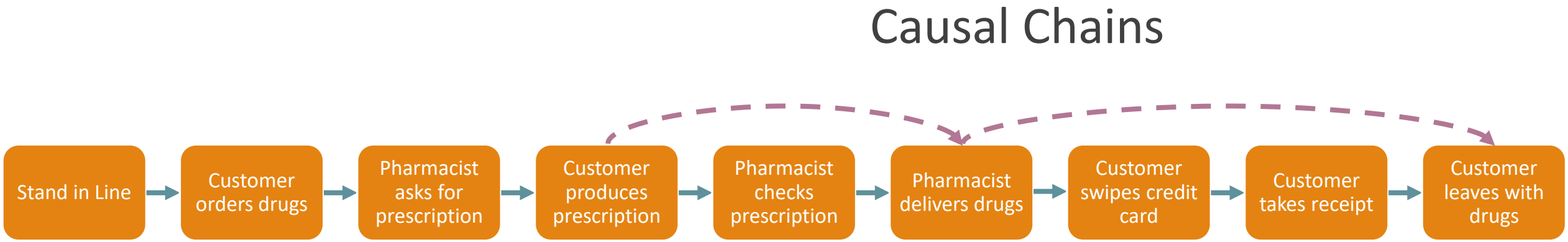
Planning

Symbolic Systems

Symbolic Systems

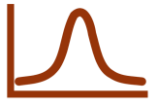


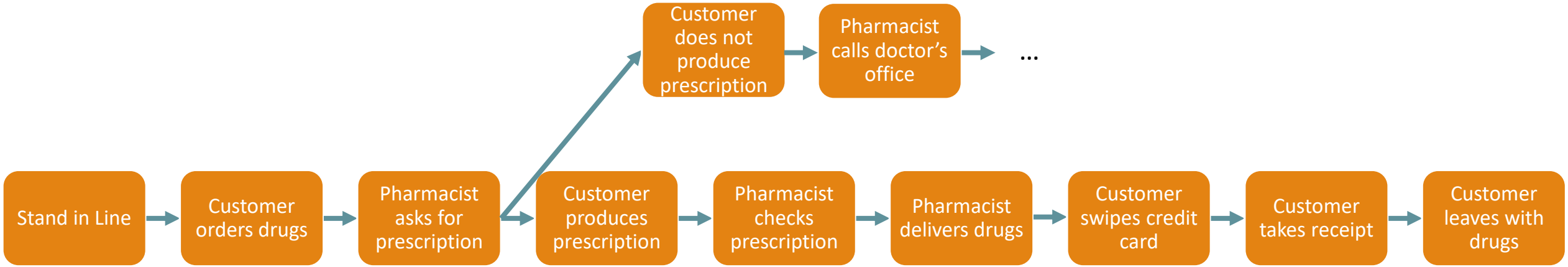
Symbolic Systems



Symbolic Systems

Branching → Planning





Pros & Cons of Symbolic Models

+ Really coherent

- Lots of knowledge engineering
- Possibly lacking diversity & novelty

} Limited Domain

Closed vs Open World

All possible events that can occur in a story

Closed vs Open World

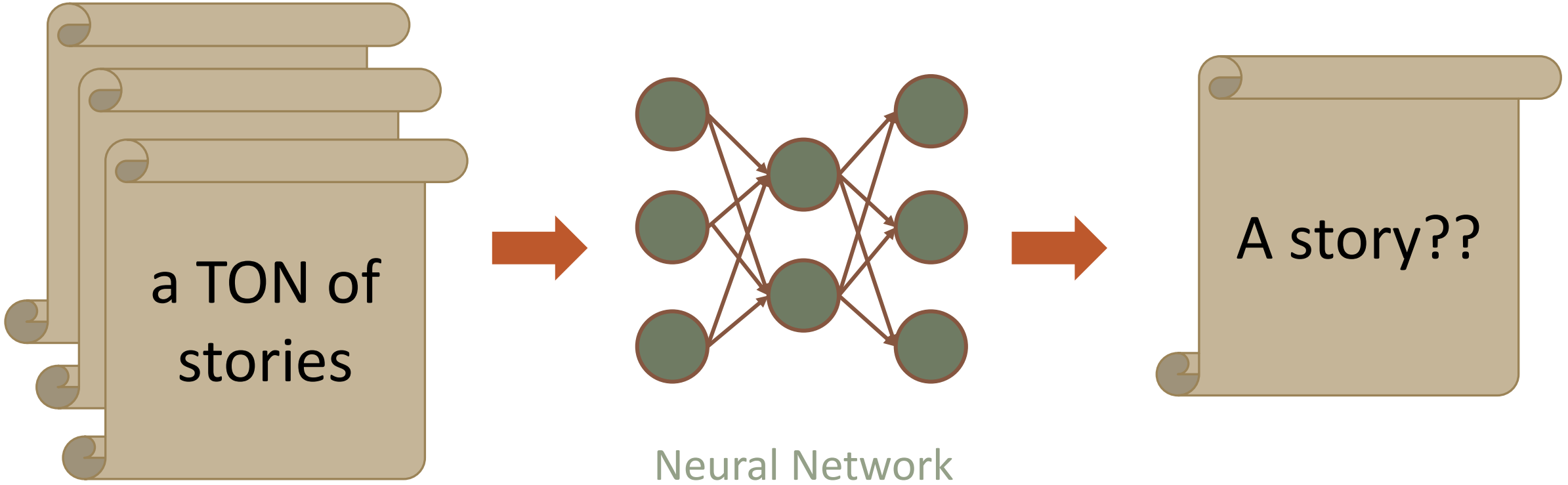
All possible thoughts a human can think of and express through language

Story Prompt
(First Sentence)

Rest of the Story

(about anything)

Neural Storytellers




SUNSPRING

Seq2Seq Neural Networks

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:

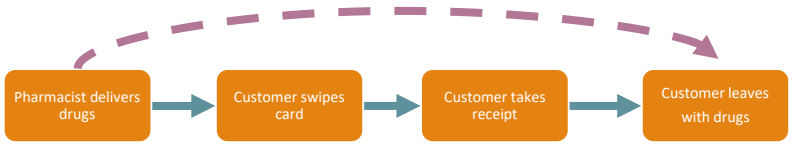
can this block gives him the advantage to personally run around with a large stick of cheese

Comparison

SYMBOLIC SYSTEMS

+Coherent stories

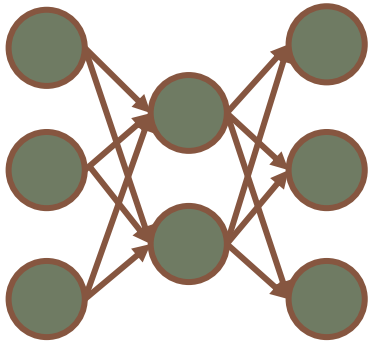
–Limited domain



NEURAL SYSTEMS

+Open world stories

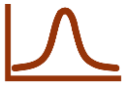
–Coherence is terrible



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 

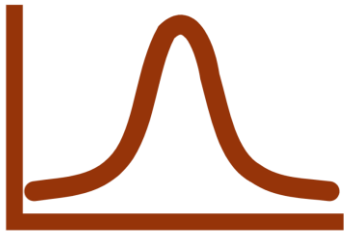
causal reasoning. 

Outline

1

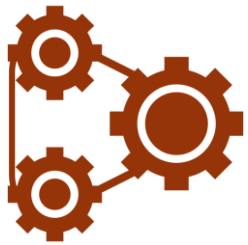
Events

2



Plot Progression

3



Reasoning

Seq2Seq Neural Networks

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.

Huge Language Model (GPT-2)

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

Too unique!

? **Problem** Sparsity!

! **Solution** Separating semantics (meaning) from syntax (grammar)

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

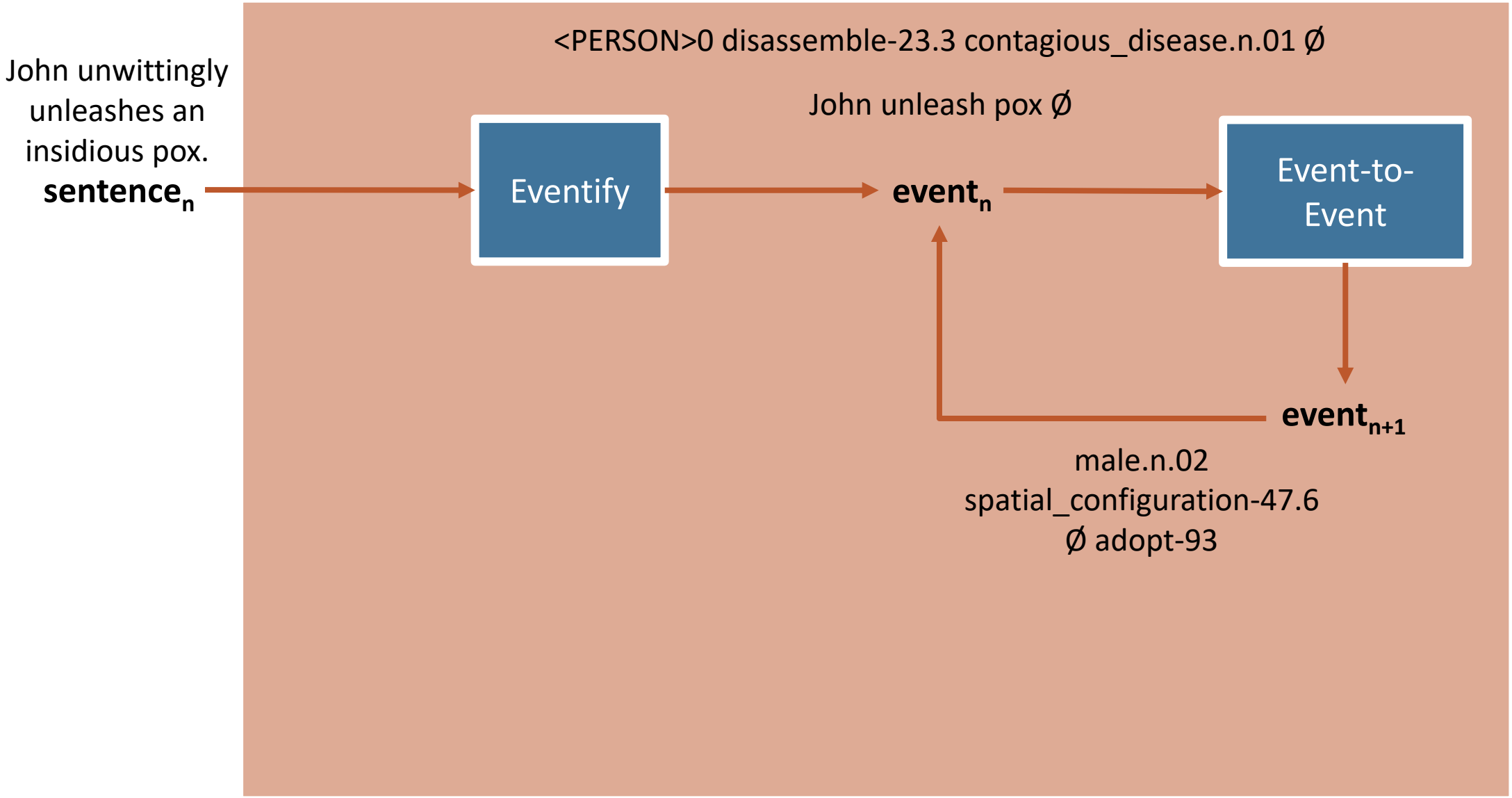
Generalized Event: <PERSON>0 fit-54.3 power.n.01 Ø

Martin, Lara J., et al. "Event representations for automated story generation with deep neural nets." *Thirty-Second AAAI Conference on Artificial Intelligence*. 2018.

Introducing ASTER

AUTOMATED STORY-TELLING WITH EVENT REPRESENTATIONS

ASTER Pipeline



Martin, Lara J., et al. "Event representations for automated story generation with deep neural nets." *Thirty-Second AAAI Conference on Artificial Intelligence*. 2018.

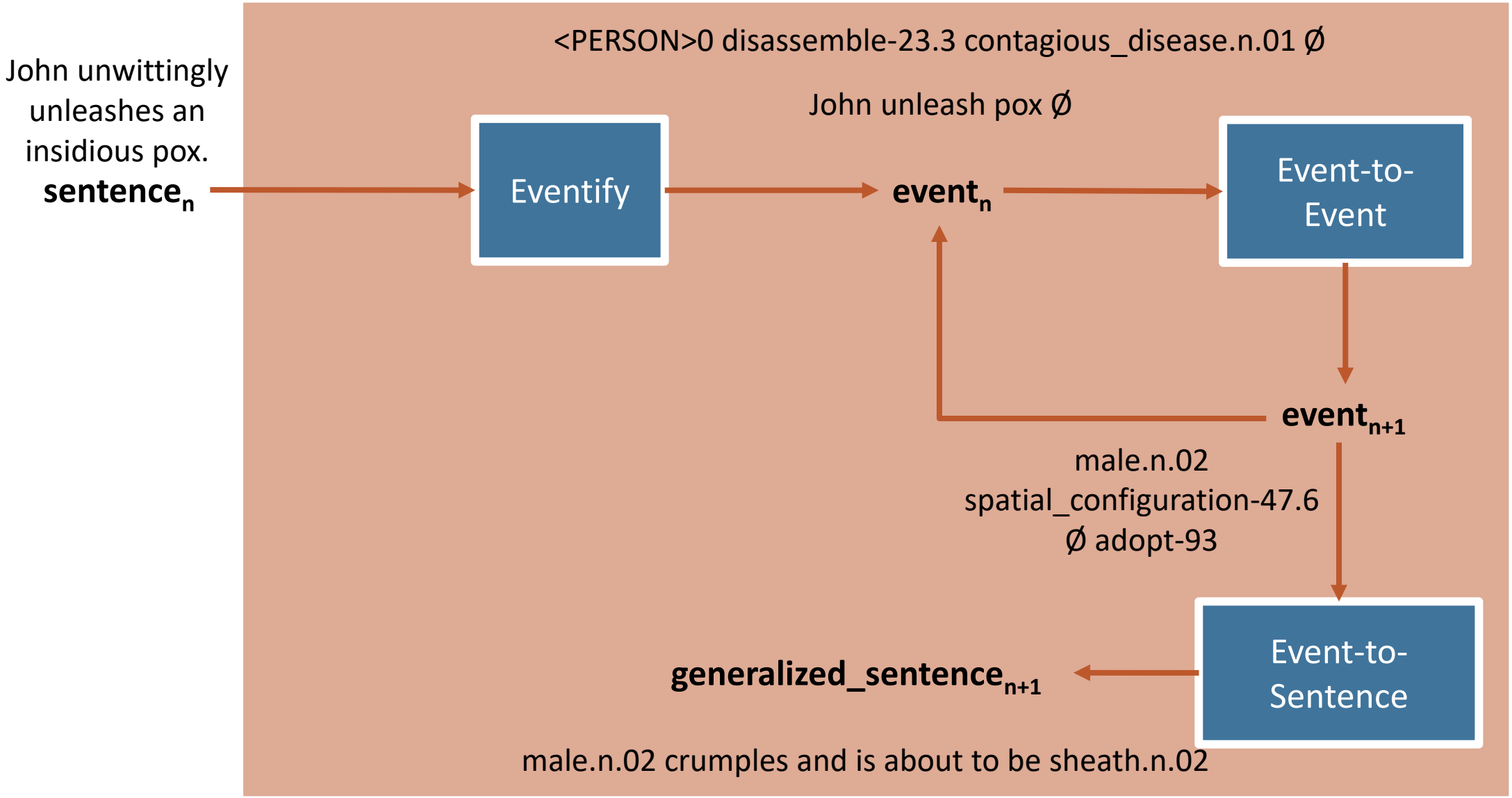
Event2Event (Seq2Seq Baseline)

Experiment	Perplexity
(0) Original Sentences	704.815
(1) Original Words Baseline	748.914
(2) Original Words with PERSONs	166.646
(3) Generalized Baseline	**54.231**

$$\boldsymbol{Perplexity} = 2^{-\sum_x \boldsymbol{p}(x) \log_2 \boldsymbol{p}(x)} \qquad \boldsymbol{p}(x) = \frac{\boldsymbol{count}(x)}{\sum_{y \epsilon Y} \boldsymbol{count}(y)}$$

Martin, Lara J., et al. "Event representations for automated story generation with deep neural nets." *Thirty-Second AAAI Conference on Artificial Intelligence*. 2018.

ASTER Pipeline



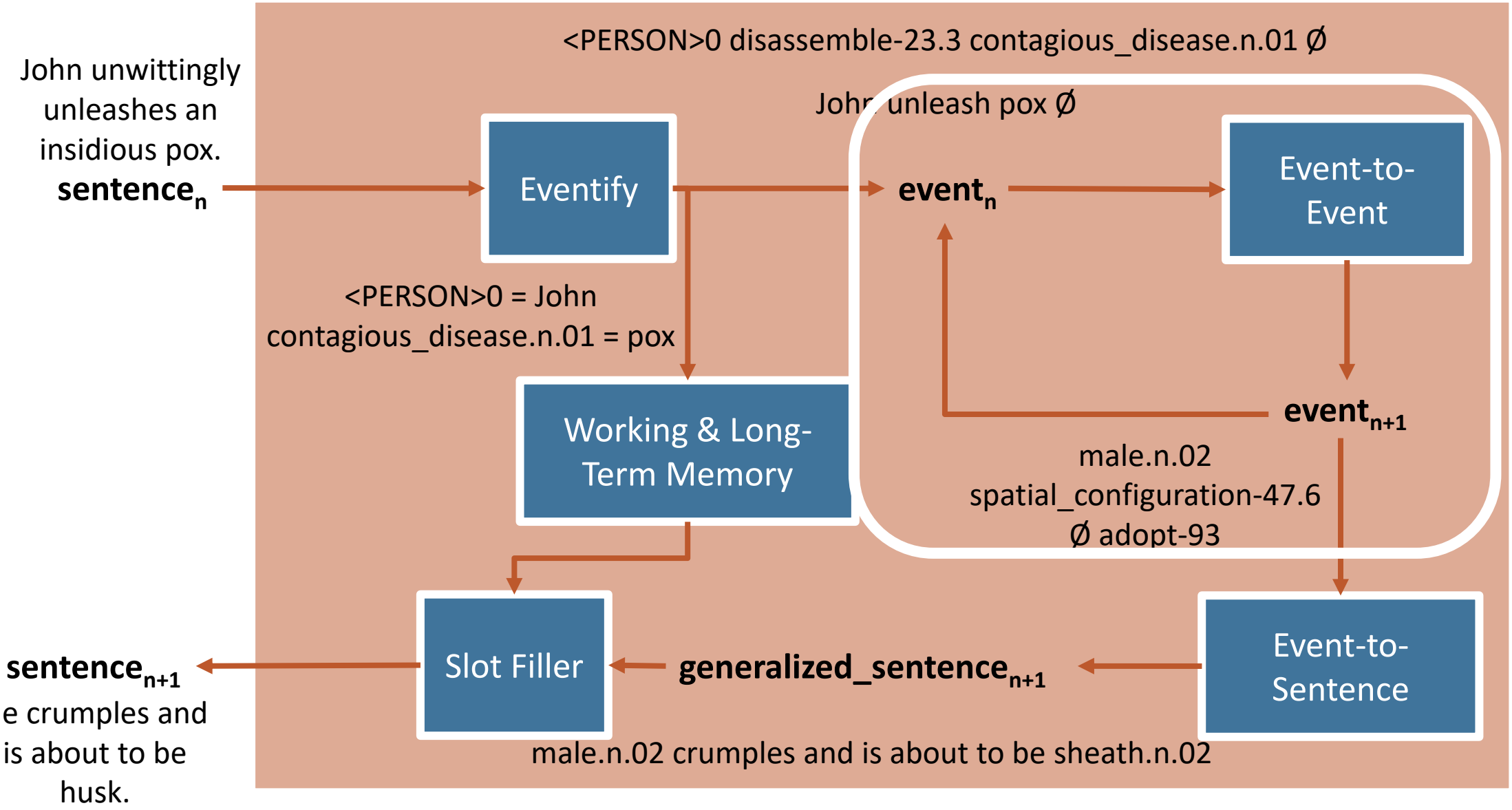
Martin, Lara J., et al. "Event representations for automated story generation with deep neural nets." *Thirty-Second AAAI Conference on Artificial Intelligence*. 2018.

Event2Sentence

Experiment	Perplexity	BLEU
Original Words Event → Original Sentence	1585.46	0.0016
Generalized Event → Generalized Sentence	**56.516**	0.0331
All Generalized Events → Generalized Sentence	59.106	**0.0366**

Martin, Lara J., et al. "Event representations for automated story generation with deep neural nets." *Thirty-Second AAAI Conference on Artificial Intelligence*. 2018.

ASTER Pipeline



Martin, Lara J., et al. "Event representations for automated story generation with deep neural nets." *Thirty-Second AAAI Conference on Artificial Intelligence*. 2018.

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

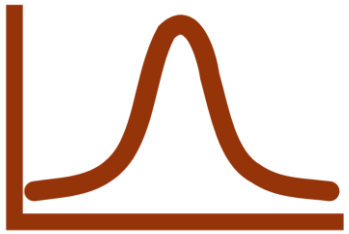
This created a giant system that still rambles

Outline

1

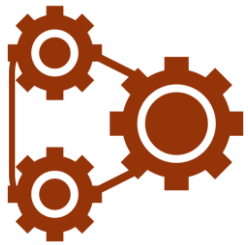
Events

2



Plot Progression

3



Reasoning

Global Coherence

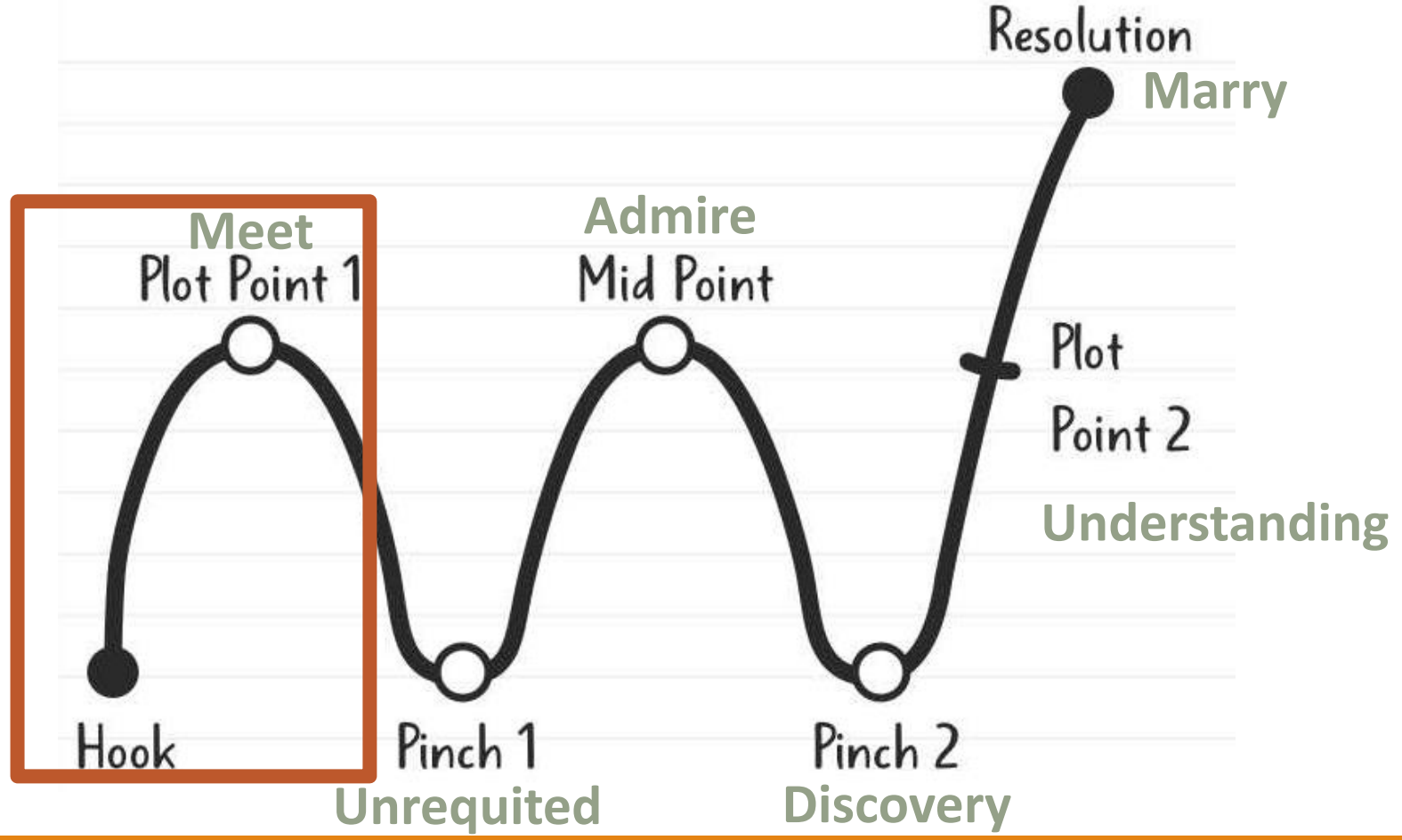
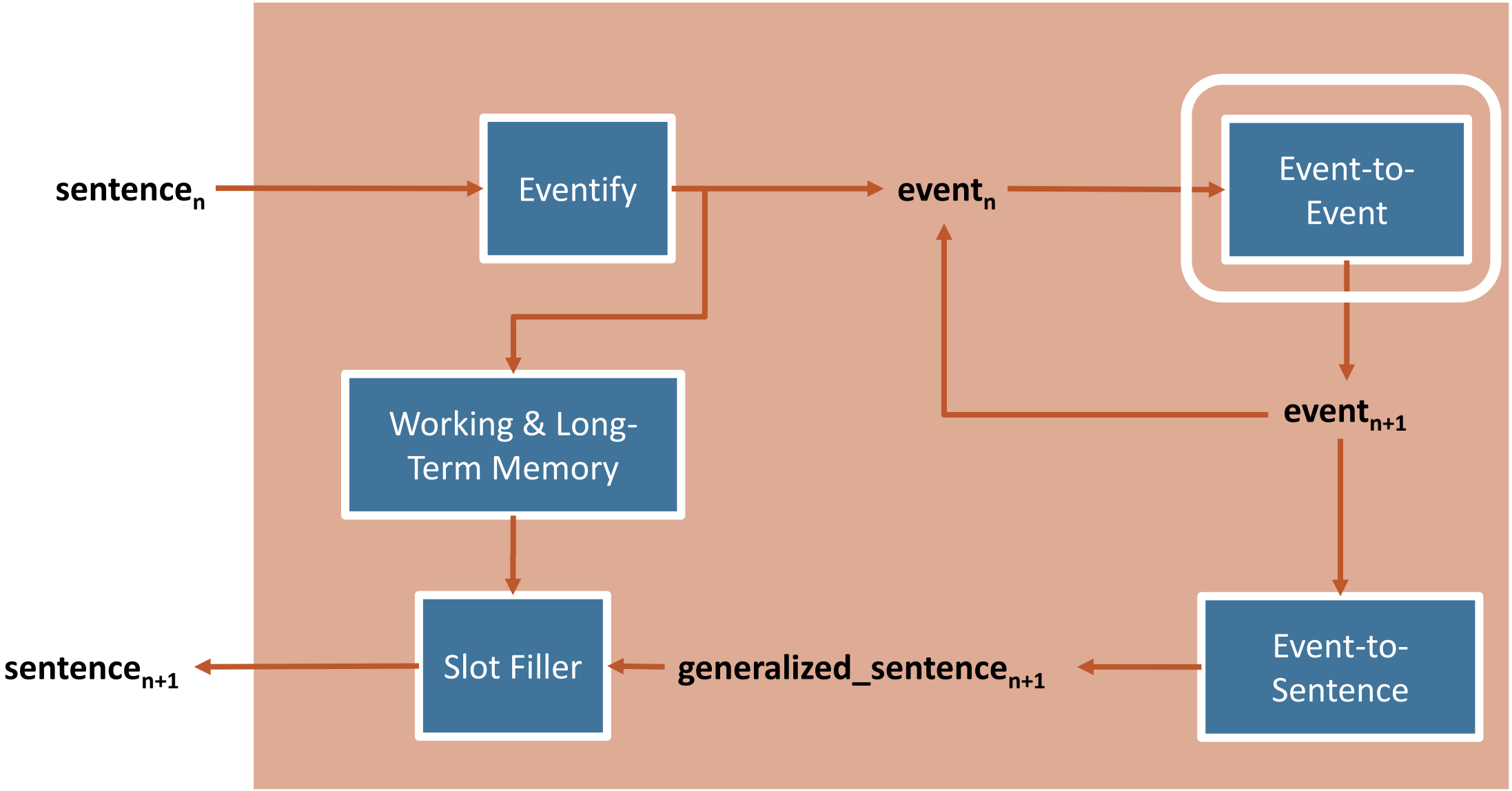
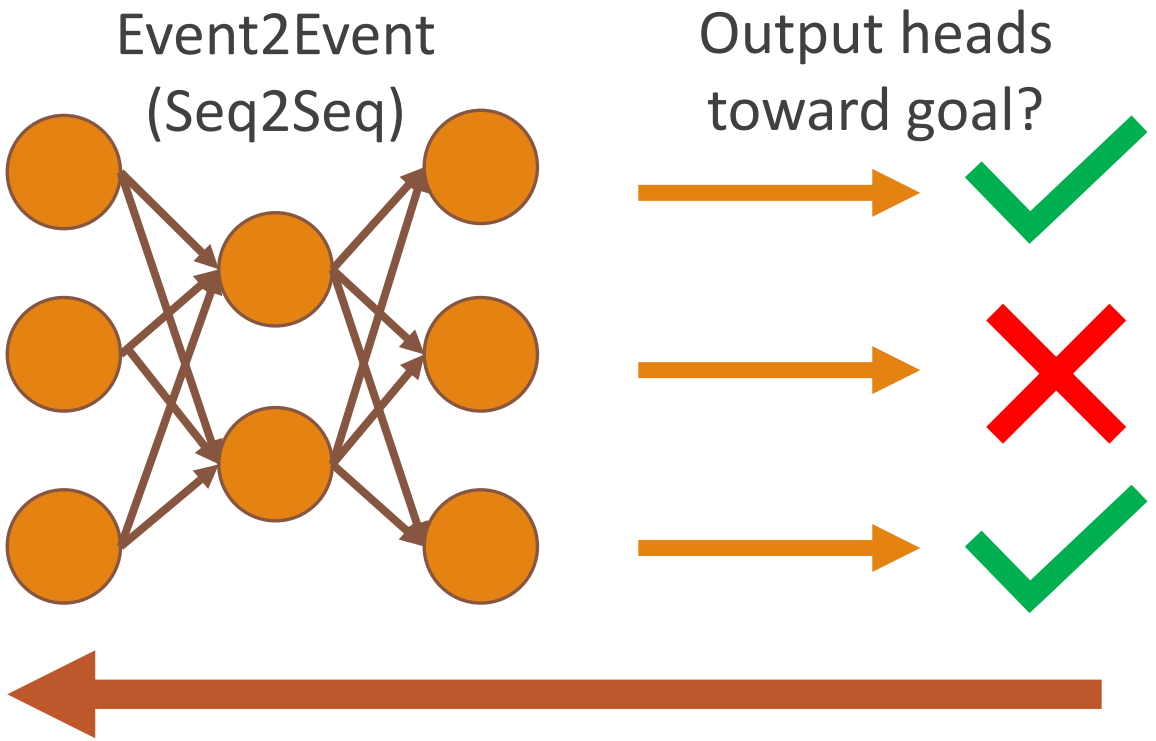


Image source: https://blog.reedsy.com/plot-point/

ASTER Pipeline

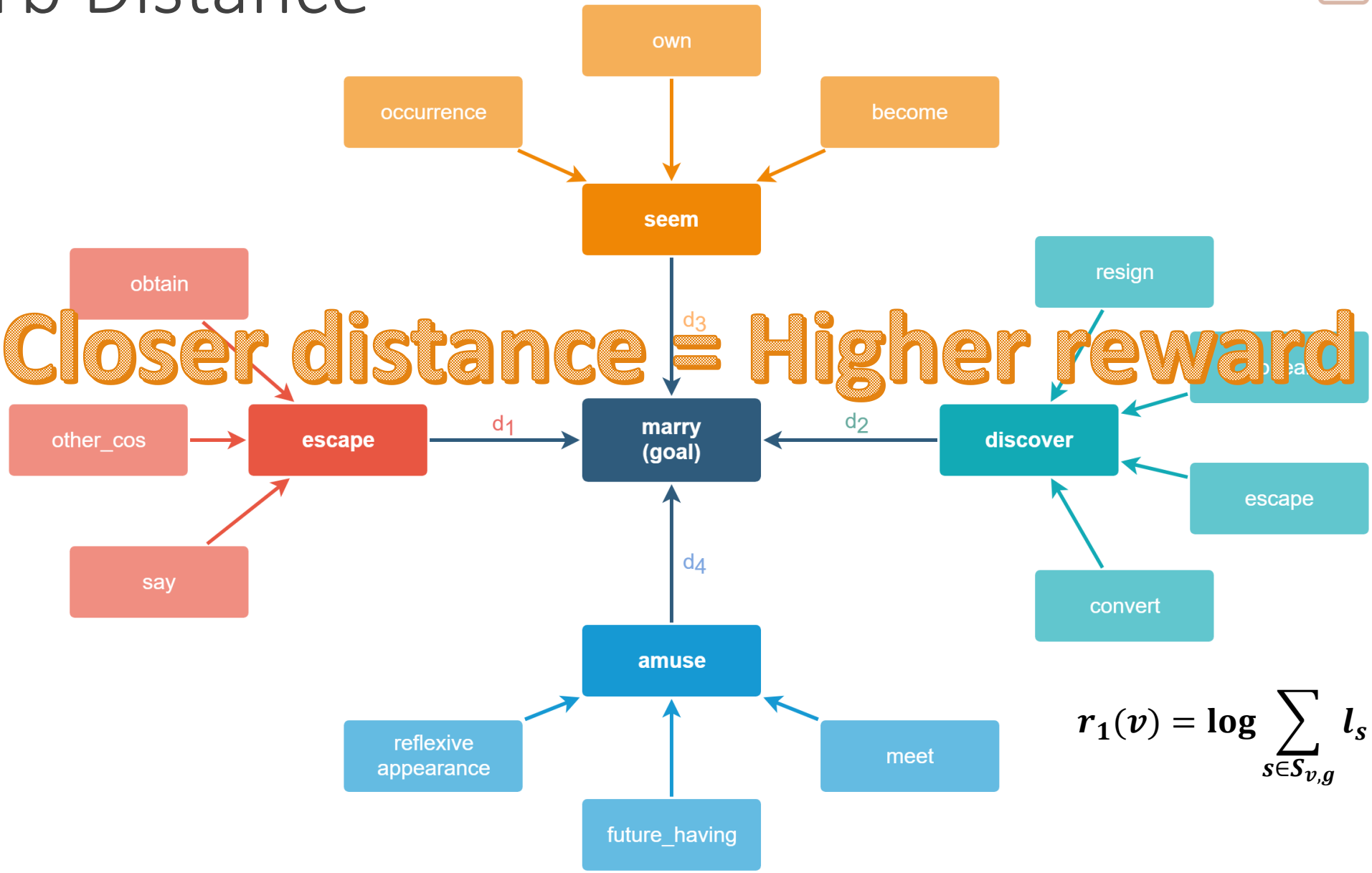


Policy Gradient DRL (REINFORCE)



Tambwekar, P.*, Dhuliawala, M.*, **Martin, L. J.,** Mehta, A., Harrison, B., & Riedl, M. O. "Controllable Neural Story Plot Generation via Reinforcement Learning." *IJCAI 2019.*

#1 Verb Distance



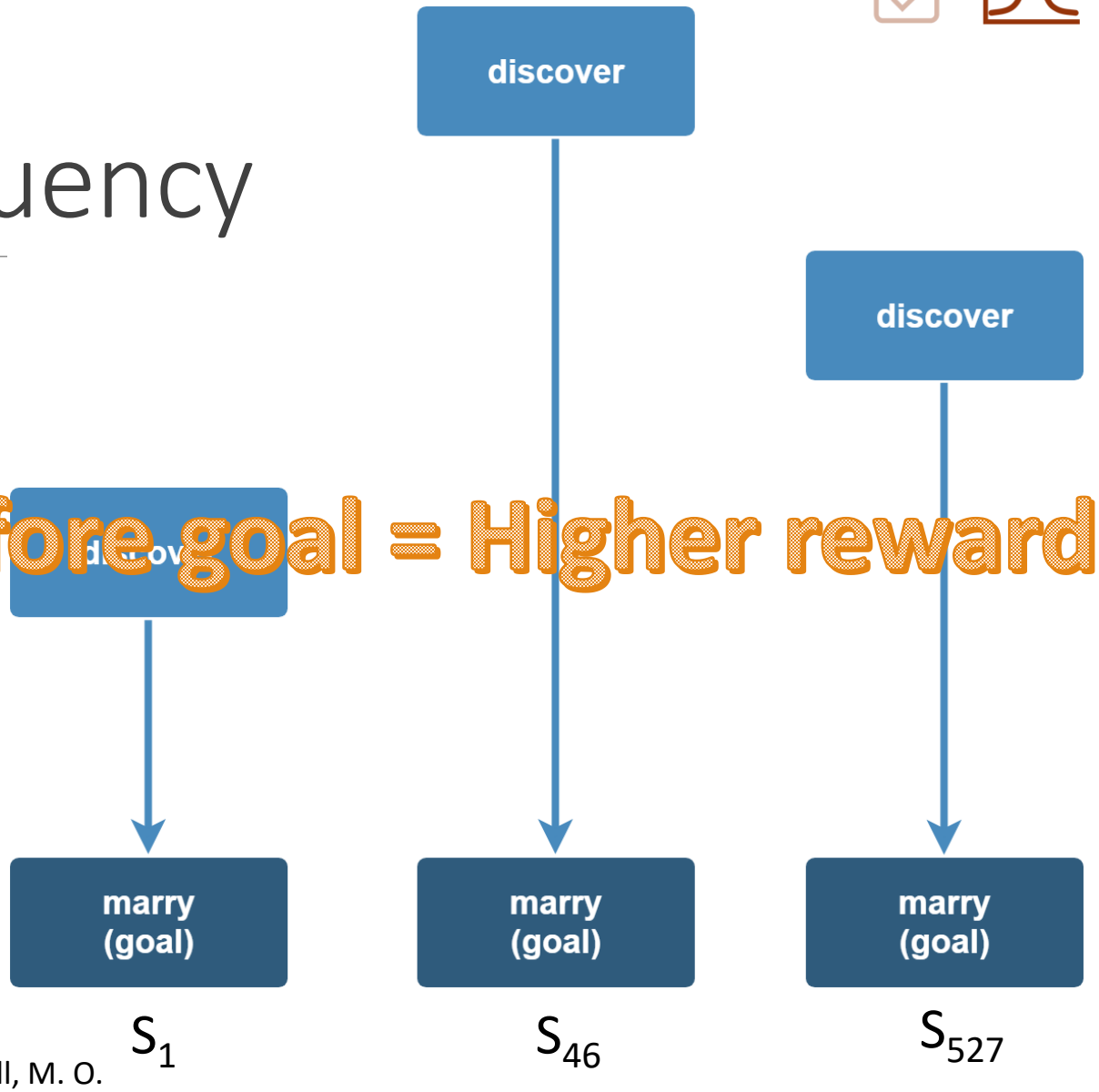
$$r_1(v) = \log \sum_{s \in S_{v,g}} l_s - d_s(v, g)$$

Tambwekar, P.*, Dhuliawala, M.*, **Martin, L. J.,** Mehta, A., Harrison, B., & Riedl, M. O. "Controllable Neural Story Plot Generation via Reinforcement Learning." *IJCAI 2019.*

#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. "Controllable Neural Story Plot Generation via Reinforcement Learning." *IJCAI 2019.*

Final Reward Equation

$$\boldsymbol{R(v) = \alpha \times r_1(v) \times r_2(v)}$$

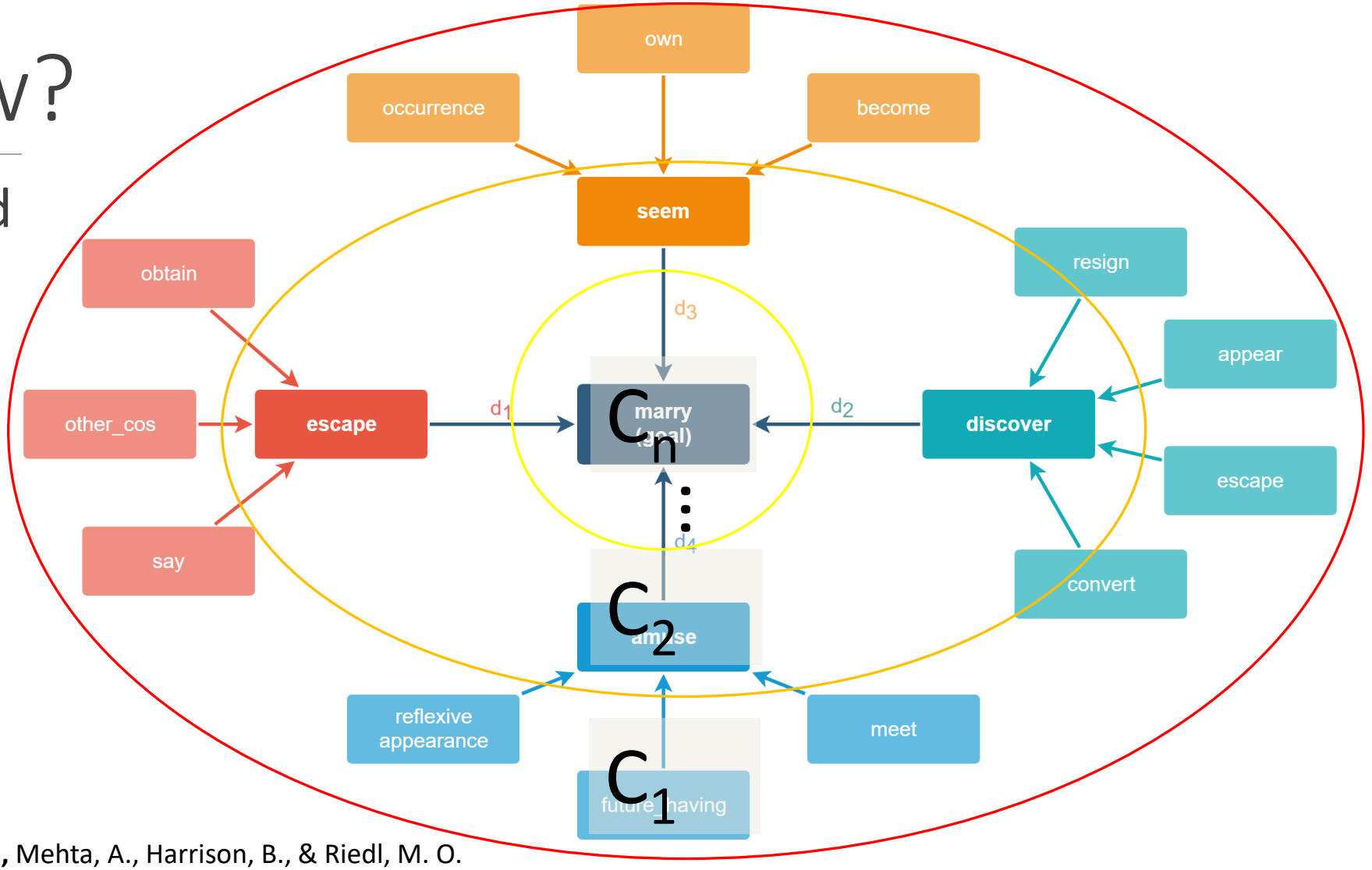
Verb Distance to Goal

Story-Verb Frequency

Tambwekar, P.*, Dhuliawala, M.*, **Martin, L. J.,** Mehta, A., Harrison, B., & Riedl, M. O. "Controllable Neural Story Plot Generation via Reinforcement Learning." *IJCAI 2019.*

What now?

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



Tambwekar, P.*, Dhuliawala, M.*, **Martin, L. J.,** Mehta, A., Harrison, B., & Riedl, M. O. "Controllable Neural Story Plot Generation via Reinforcement Learning." *IJCAI 2019.*

Policy Gradient Results

Goal	Model	Goal Achievement Rate	Average Perplexity	Average Story Length
admire	Test Corpus	20.30%	n/a	7.59
	Seq2Seq	35.52%	48.06	7.11
	Policy Gradient DRL	**94.29%**	**7.61**	4.90
marry	Test Corpus	24.64%	n/a	7.37
	Seq2Seq	39.92%	48.06	6.94
	Policy Gradient DRL	**93.35%**	**7.05**	5.76

Tambwekar, P.*, Dhuliawala, M.*, **Martin, L. J.,** Mehta, A., Harrison, B., & Riedl, M. O. "Controllable Neural Story Plot Generation via Reinforcement Learning." *IJCAI 2019.*

But are the stories actually any *good*?

Methods

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

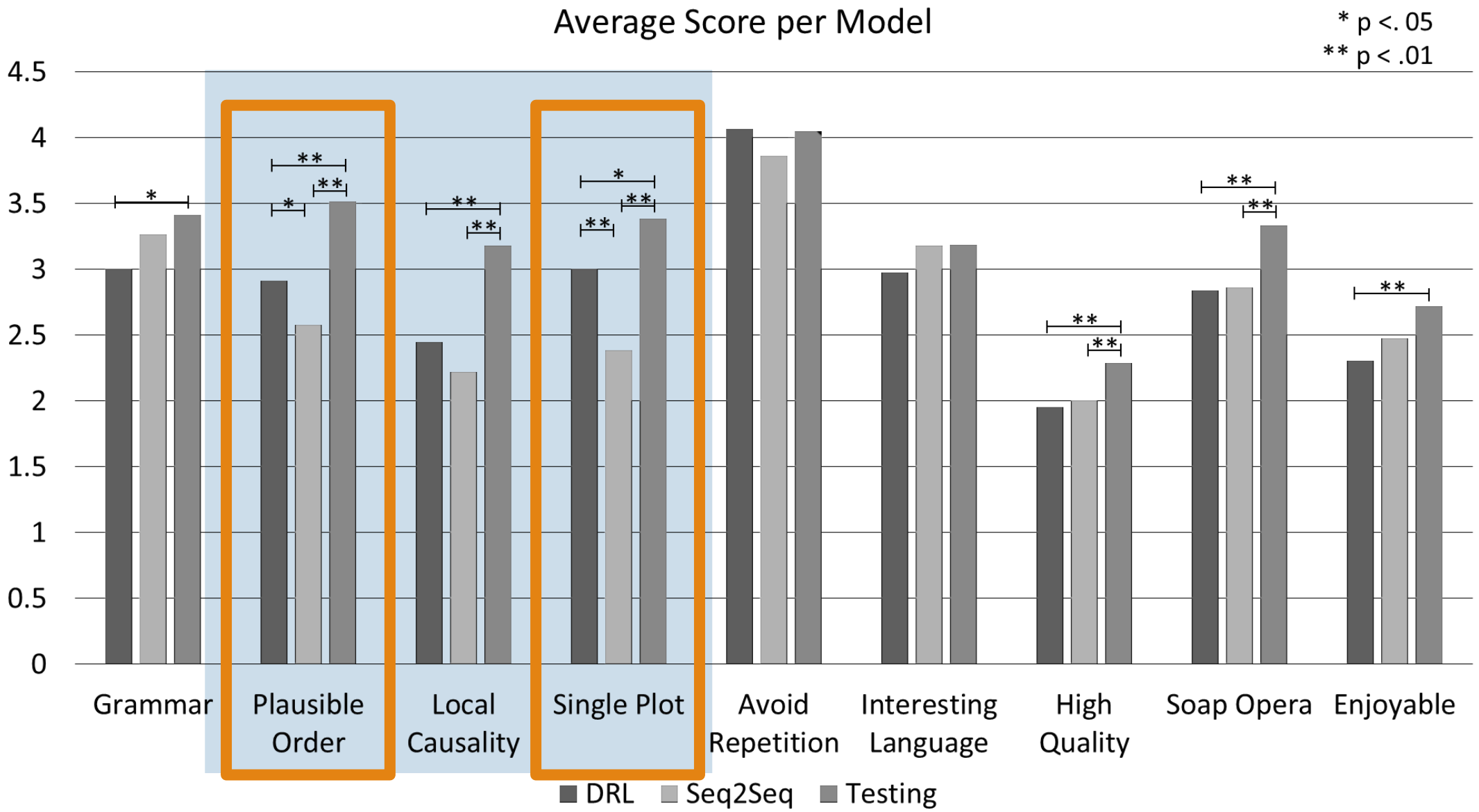
For each of 3 conditions:

- Policy Gradient DRL
- Baseline Seq2Seq
- Testing Set Stories (Gold Standard)

Tambwekar, P.*, Dhuliawala, M.*, **Martin, L. J.,** Mehta, A., Harrison, B., & Riedl, M. O. "Controllable Neural Story Plot Generation via Reinforcement Learning." *IJCAI 2019.*

Questionnaire

1. This story exhibits CORRECT GRAMMAR.
2. This story's events occur in a PLAUSIBLE ORDER. Coherence
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.



Tambwekar, P.*, Dhuliawala, M.*, **Martin, L. J.,** Mehta, A., Harrison, B., & Riedl, M. O. "Controllable Neural Story Plot Generation via Reinforcement Learning." *IJCAI 2019.*

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.

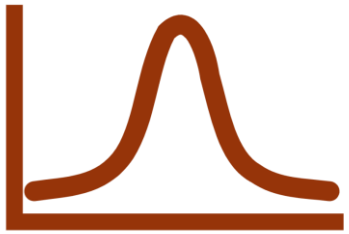
Penelope detested the jungle gym.

Outline

1

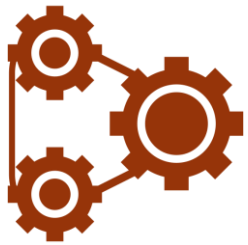
Events

2



Plot Progression

3



Reasoning

Remember Causal Chains?

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

Initial_Location : location

Theme : concrete

Agent : animate or organization

SELECTIONAL RESTRICTIONS

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)~~

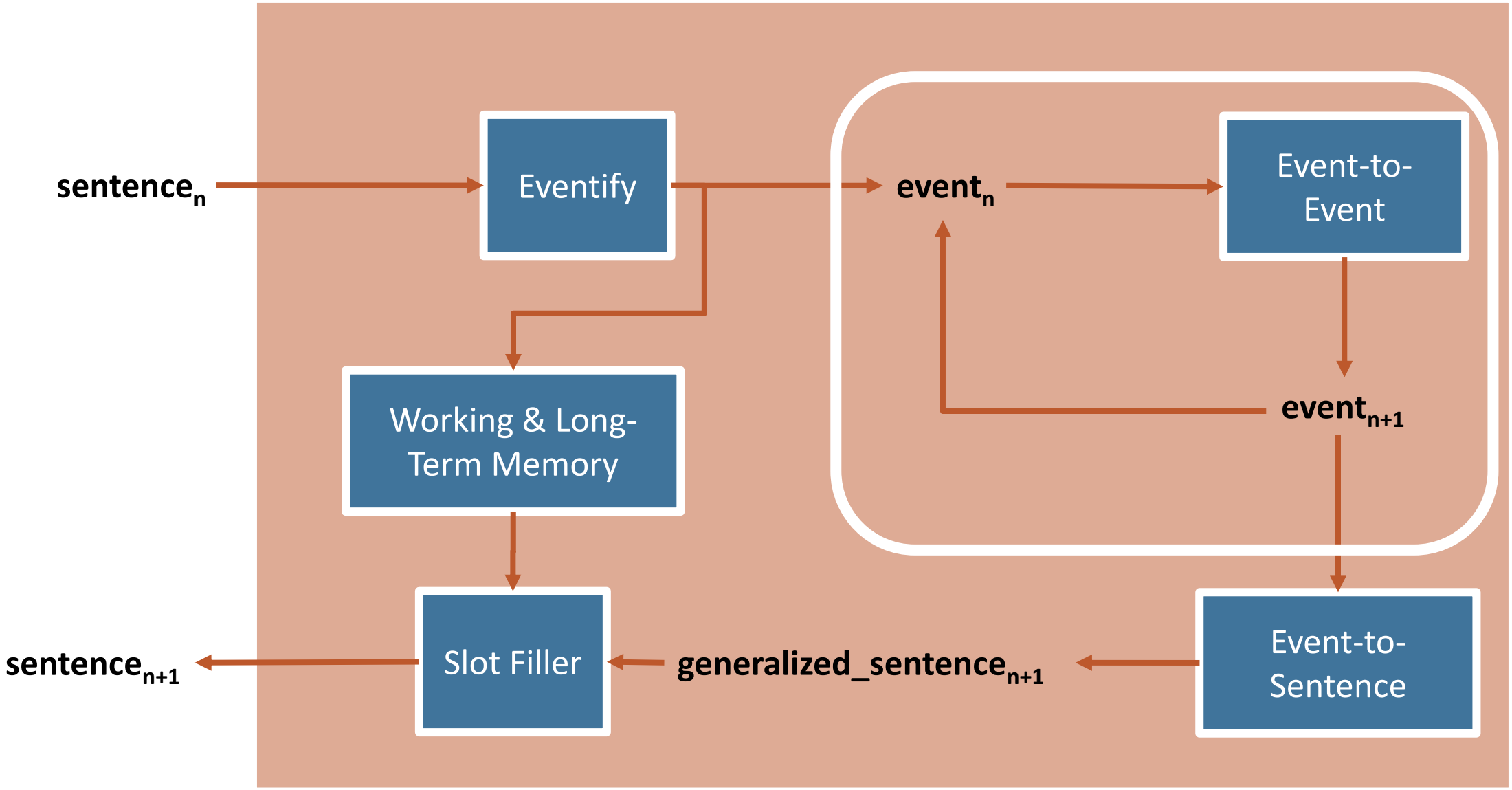
~~motion(e3, book)~~

!has_location(e3, book, Atlanta)

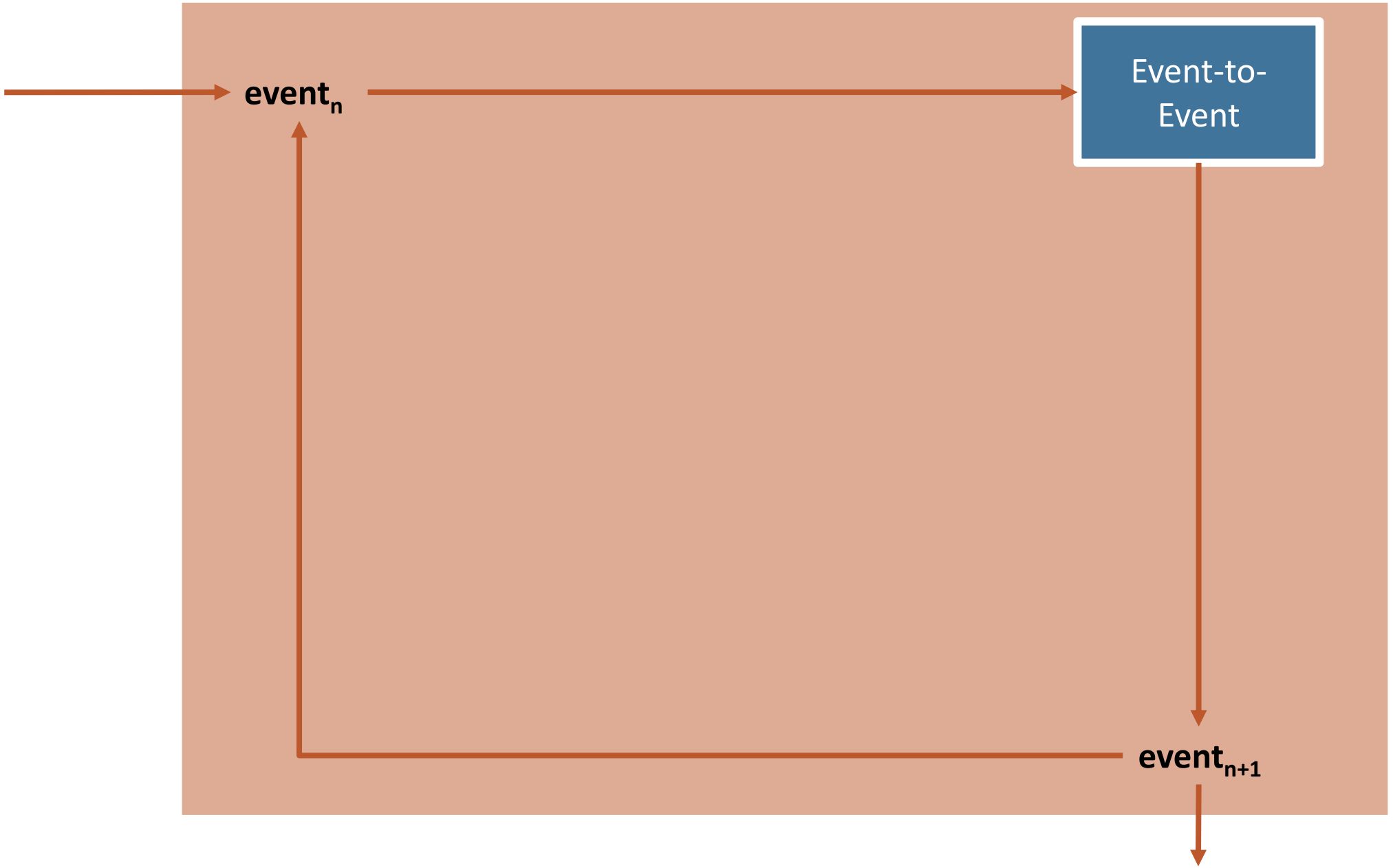
has_location(e4, book, Remy)

Then VerbNet can model the story world!

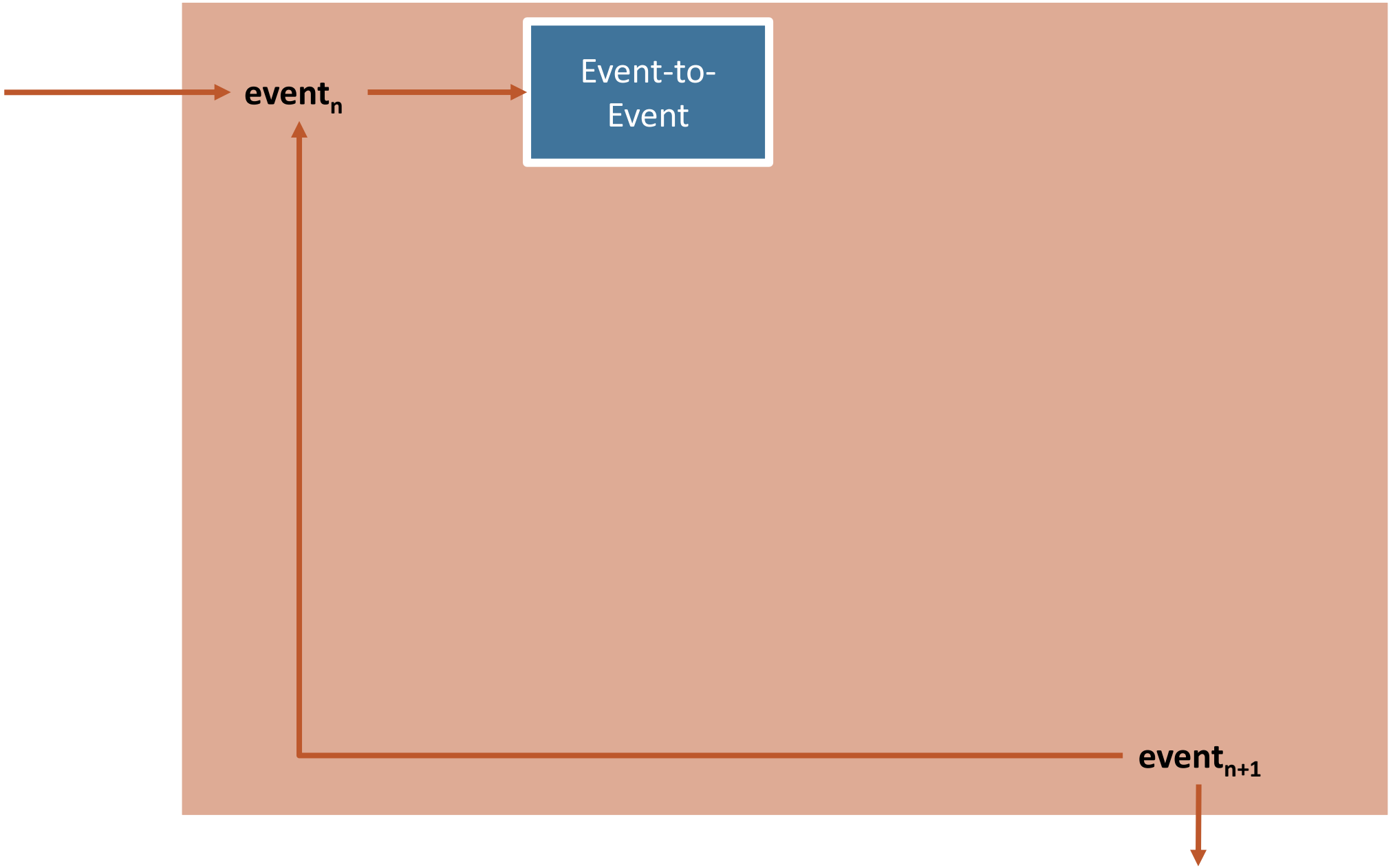
ASTER Pipeline



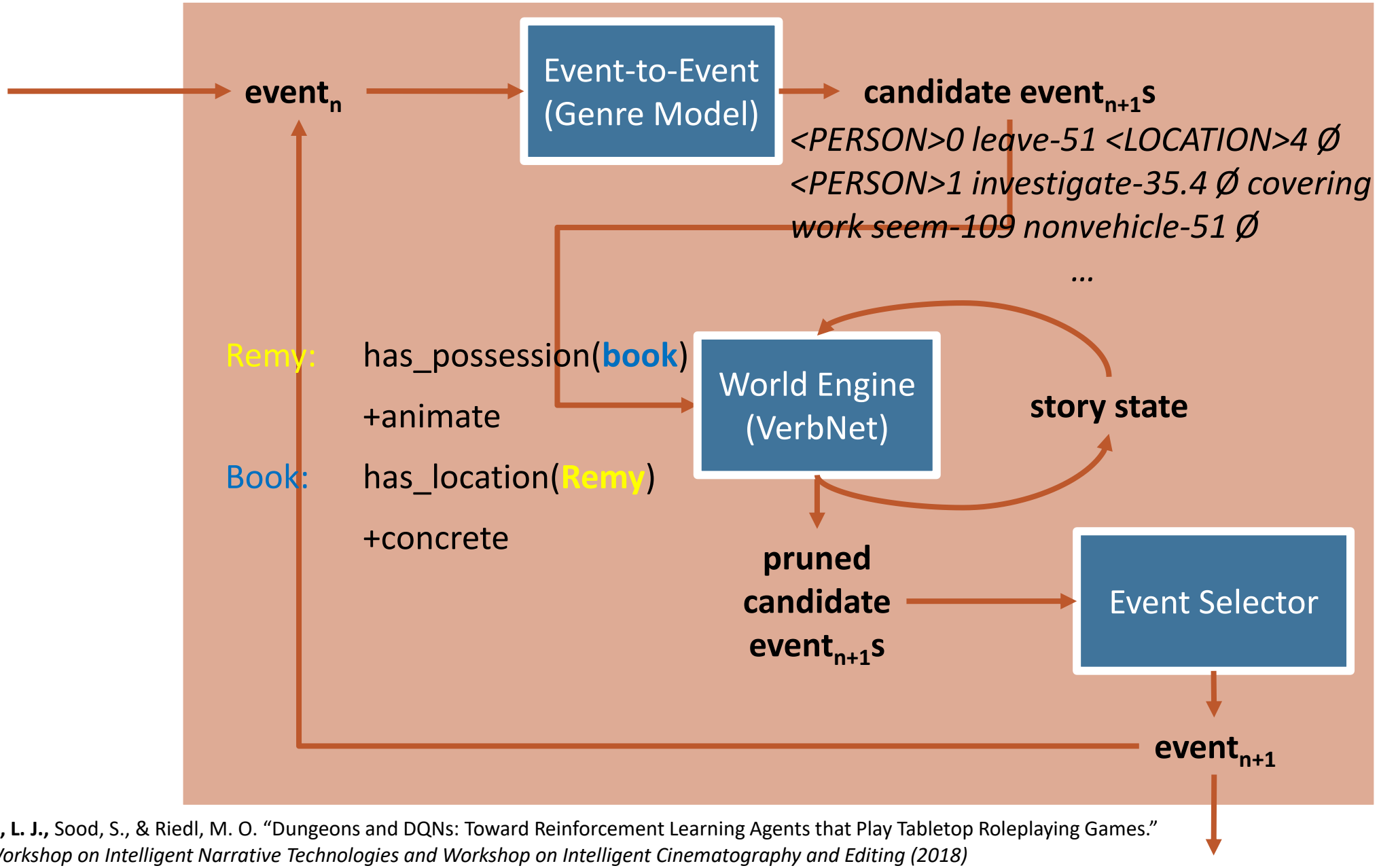
Event-to-Event



Event-to-Event



ASTER-X: ASTER (Symbolic) eXtension



Martin, L. J., Sood, S., & Riedl, M. O. "Dungeons and DQNs: Toward Reinforcement Learning Agents that Play Tabletop Roleplaying Games." *Joint Workshop on Intelligent Narrative Technologies and Workshop on Intelligent Cinematography and Editing (2018)*

ASTER-X Example (Hand-Translated)

ASTER-X (NEUROSYMBOLIC)

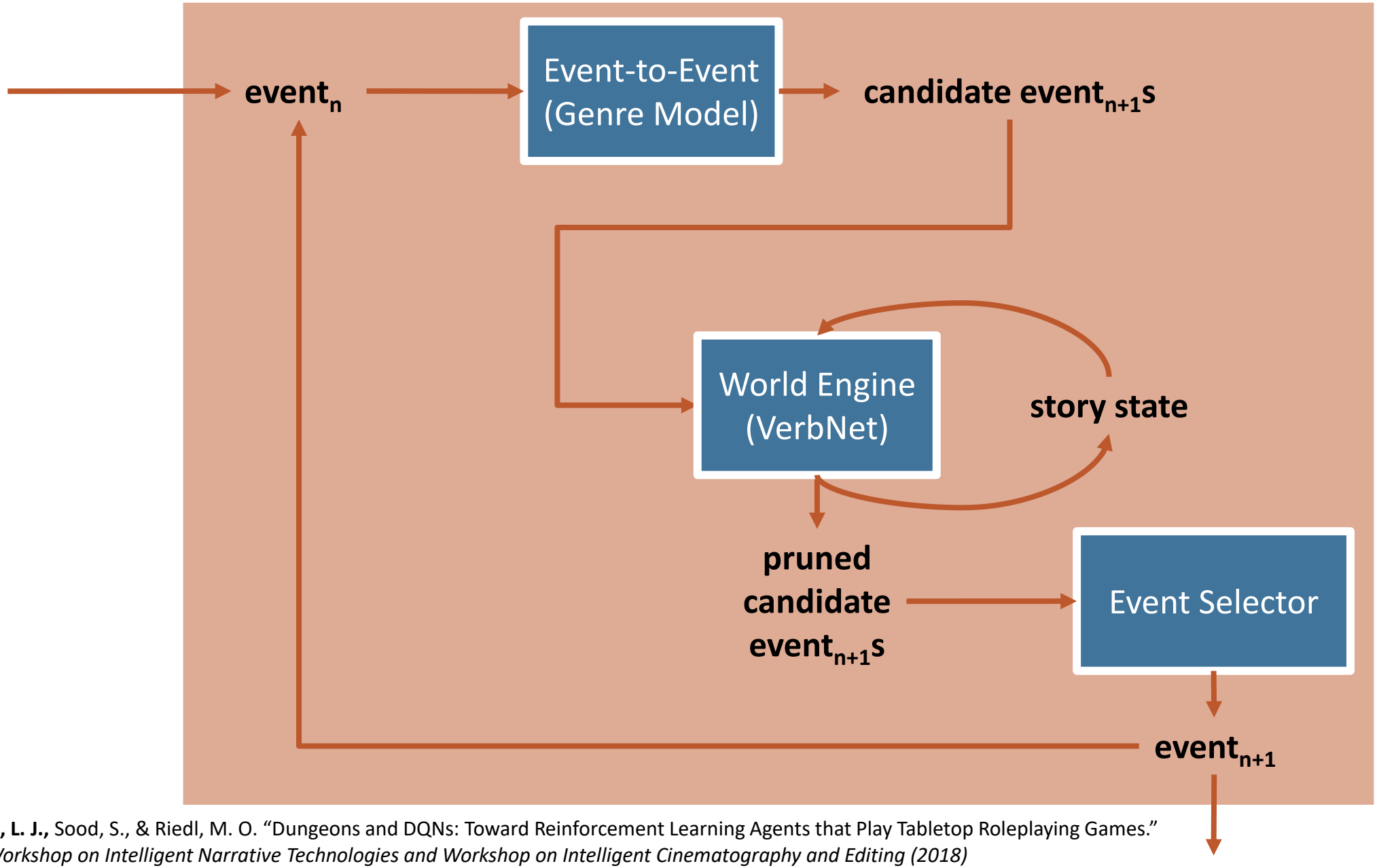
The traveler succeeded.
The traveler materialized the Voyager.
The traveler vaporized the Voyager.
Evelyn sought the Voyager to Paul.
What found the farewell order to the Voyager?
Wendy demanded to judge for the vote.
Wendy asked the vote up Kevin.
Ruby consulted the draft to the Voyager.
The Voyager 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.

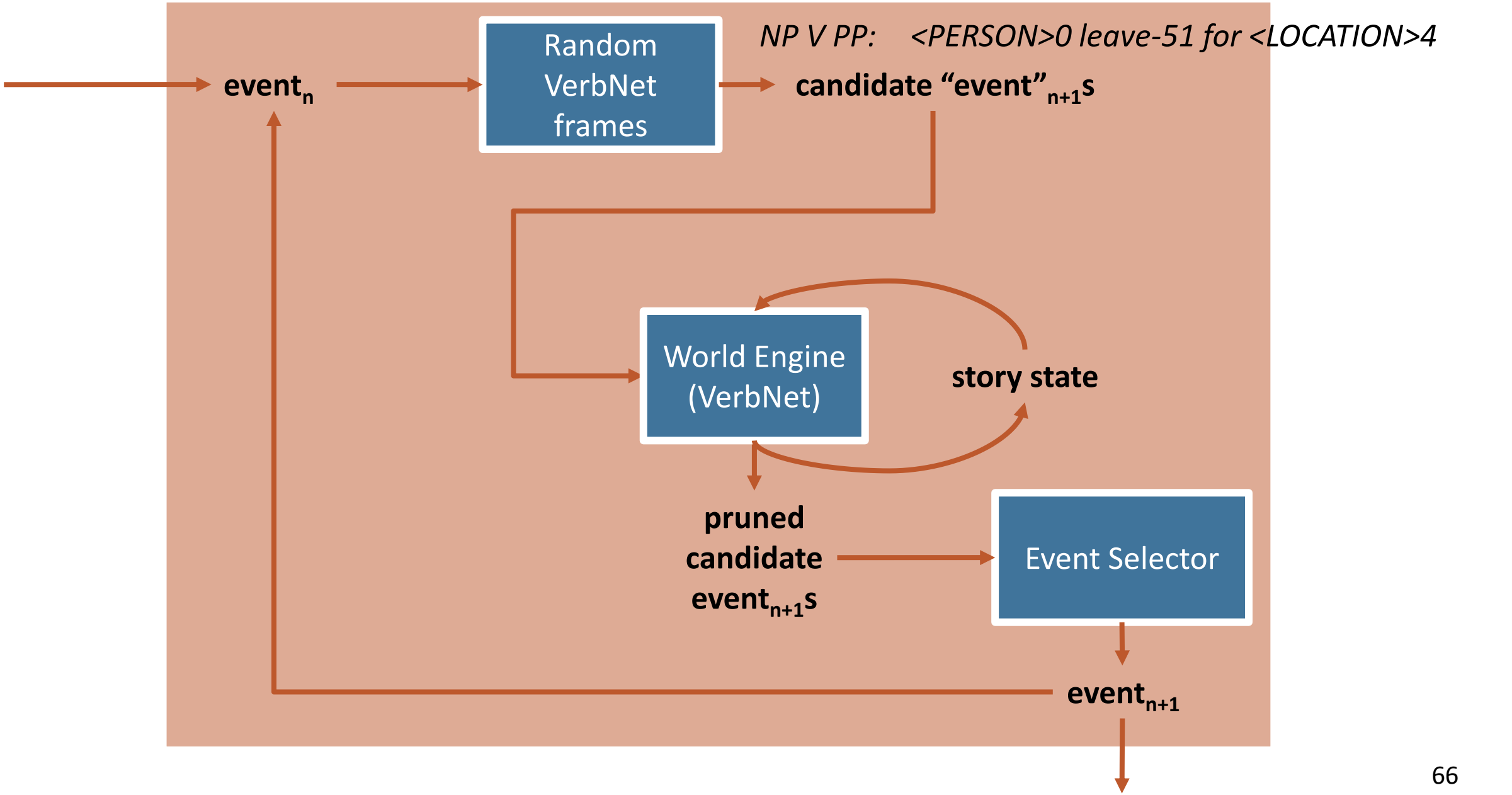
How does this compare to symbolic-only systems?

ASTER-X: ASTER (Symbolic) eXtension



Martin, L. J., Sood, S., & Riedl, M. O. "Dungeons and DQNs: Toward Reinforcement Learning Agents that Play Tabletop Roleplaying Games." *Joint Workshop on Intelligent Narrative Technologies and Workshop on Intelligent Cinematography and Editing (2018)*

Symbolic-Only System Baseline

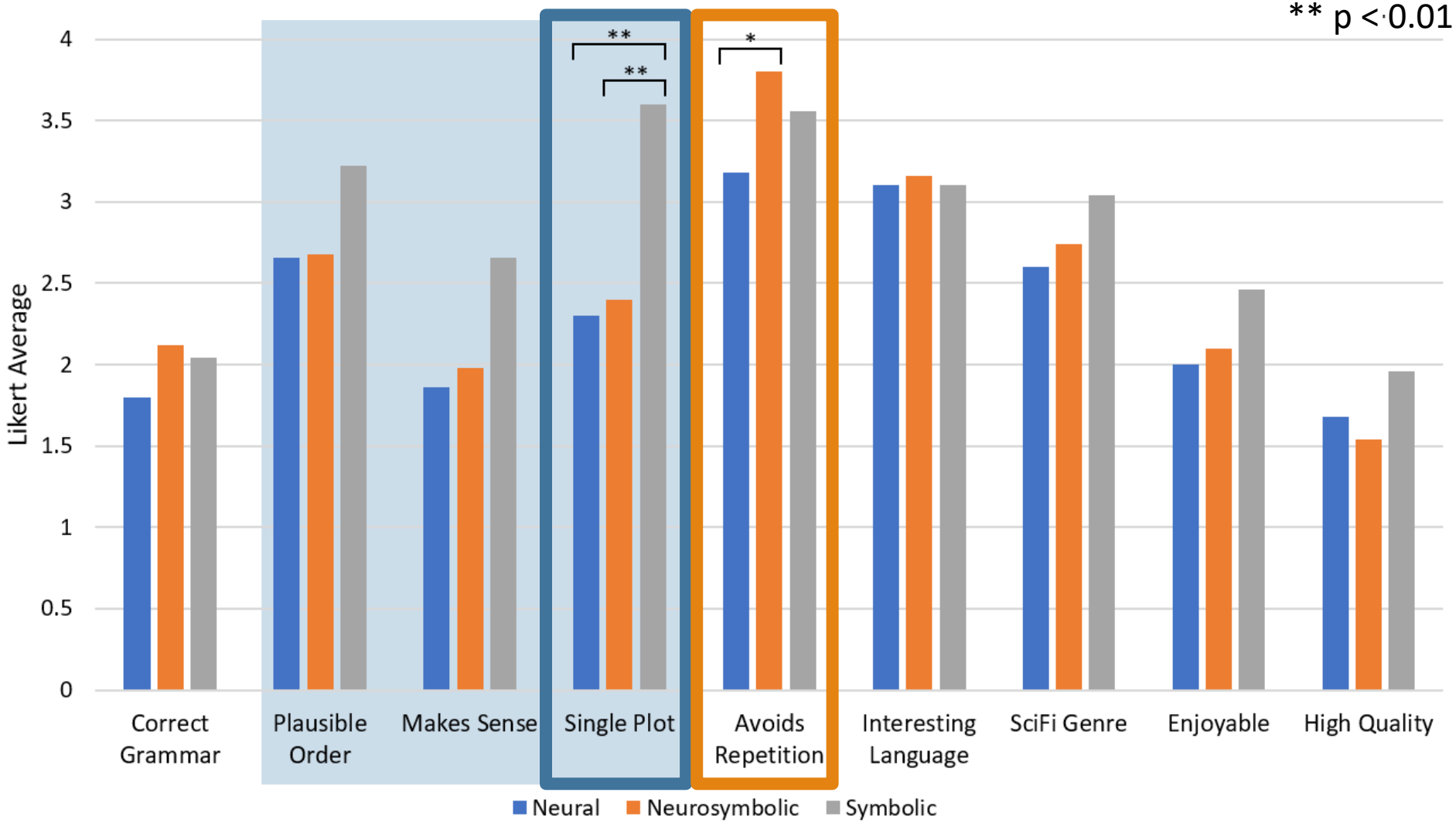


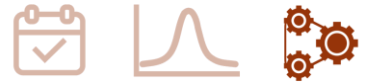
Questionnaire

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

* $p < 0.05$
** $p < 0.01$

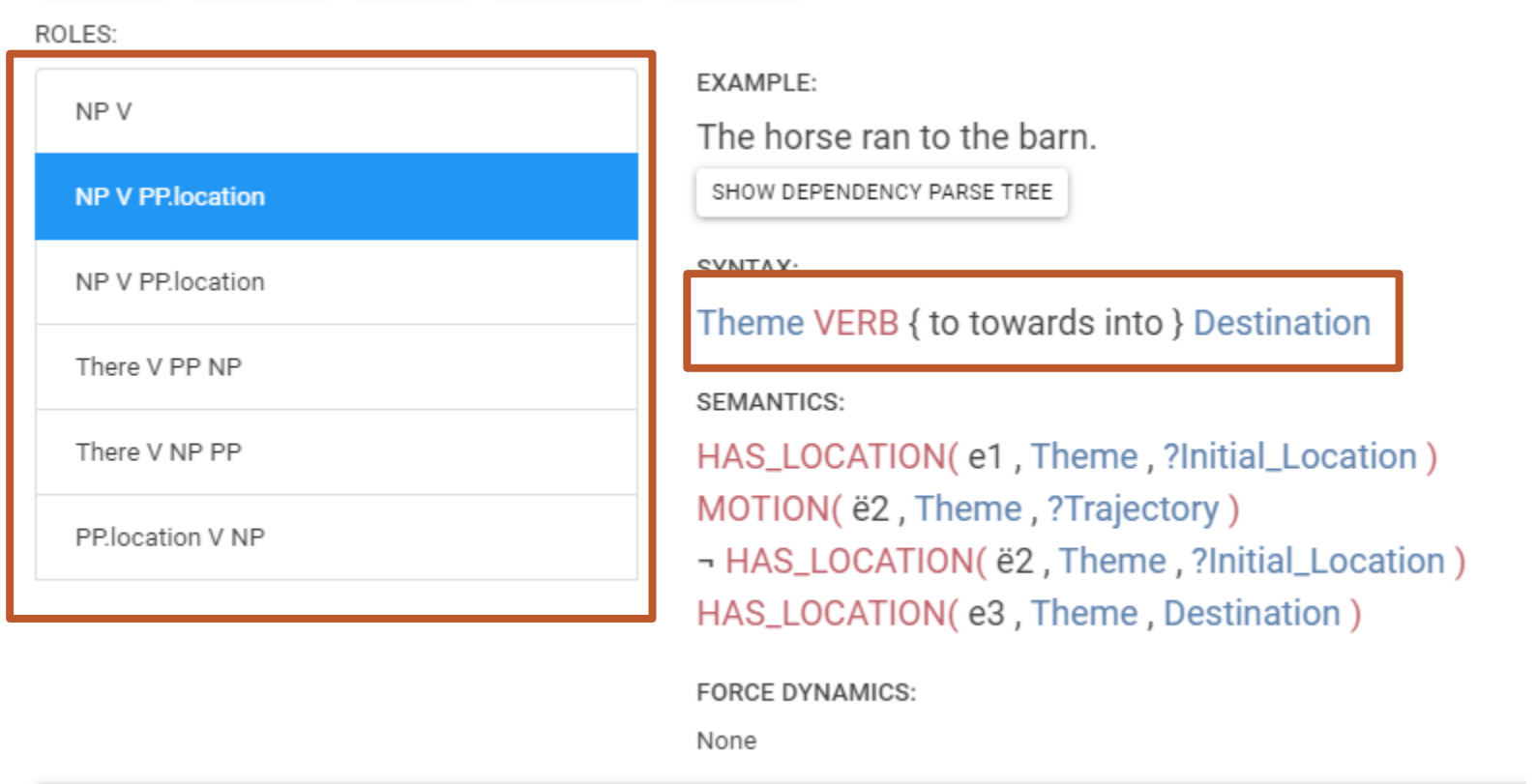




“While difficult for me to understand, the story does seem to follow in plausible order.”

–A Turker, about the Symbolic-only system

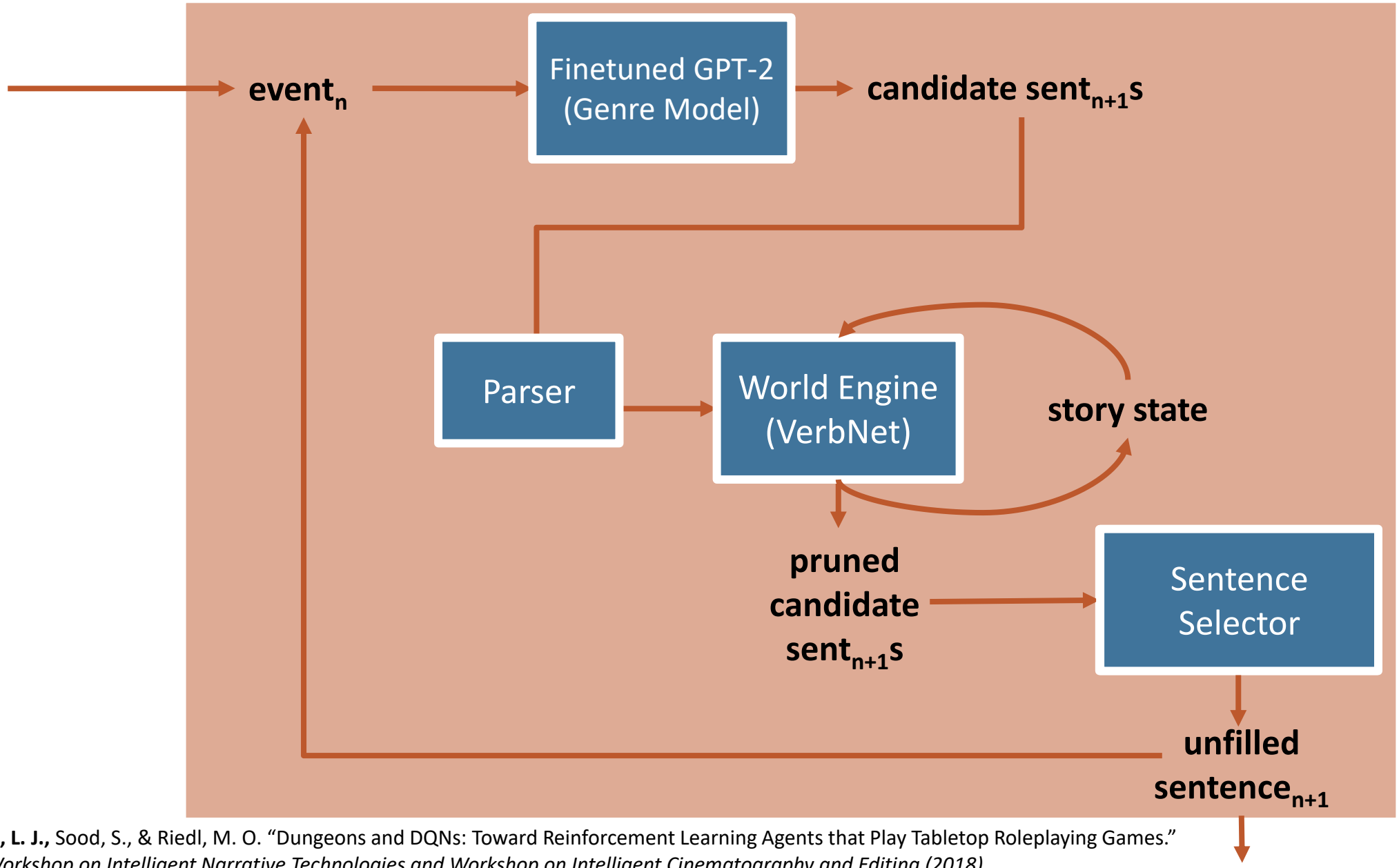
Using VerbNet for Syntax Improved Readability



Screenshot from https://uvi.colorado.edu/uvi_search

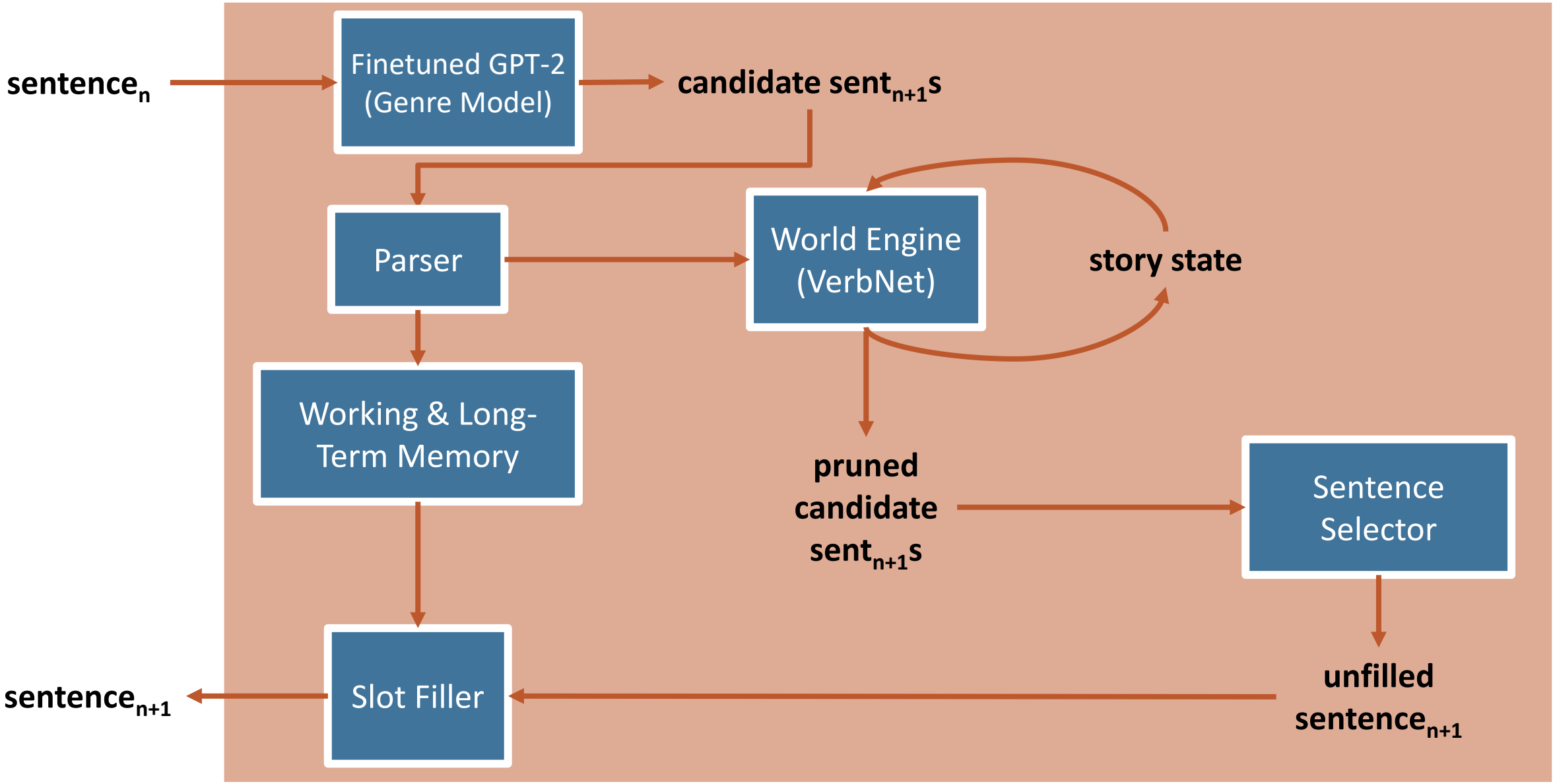
To help the neurosymbolic system with **readability**, what if my event representation was a full sentence?

ASTER-XT: ASTER-X with Transformers



Martin, L. J., Sood, S., & Riedl, M. O. "Dungeons and DQNs: Toward Reinforcement Learning Agents that Play Tabletop Roleplaying Games." *Joint Workshop on Intelligent Narrative Technologies and Workshop on Intelligent Cinematography and Editing (2018)*

ASTER-XT: ASTER-X with Transformers



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!

(Coherence)

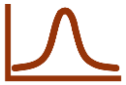
Ability to solve long-term plans! (Also, coherence)

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 Turker

Thank you!