Controllable Neural Plot Generation via Reward Shaping

PRADYUMNA TAMBWEKAR*, MURTAZA DHULIAWALA*, **LARA J. MARTIN**, ANIMESH MEHTA, BRENT HARRISON, AND MARK O. RIEDL

*Equal contribution

Why Storytelling?

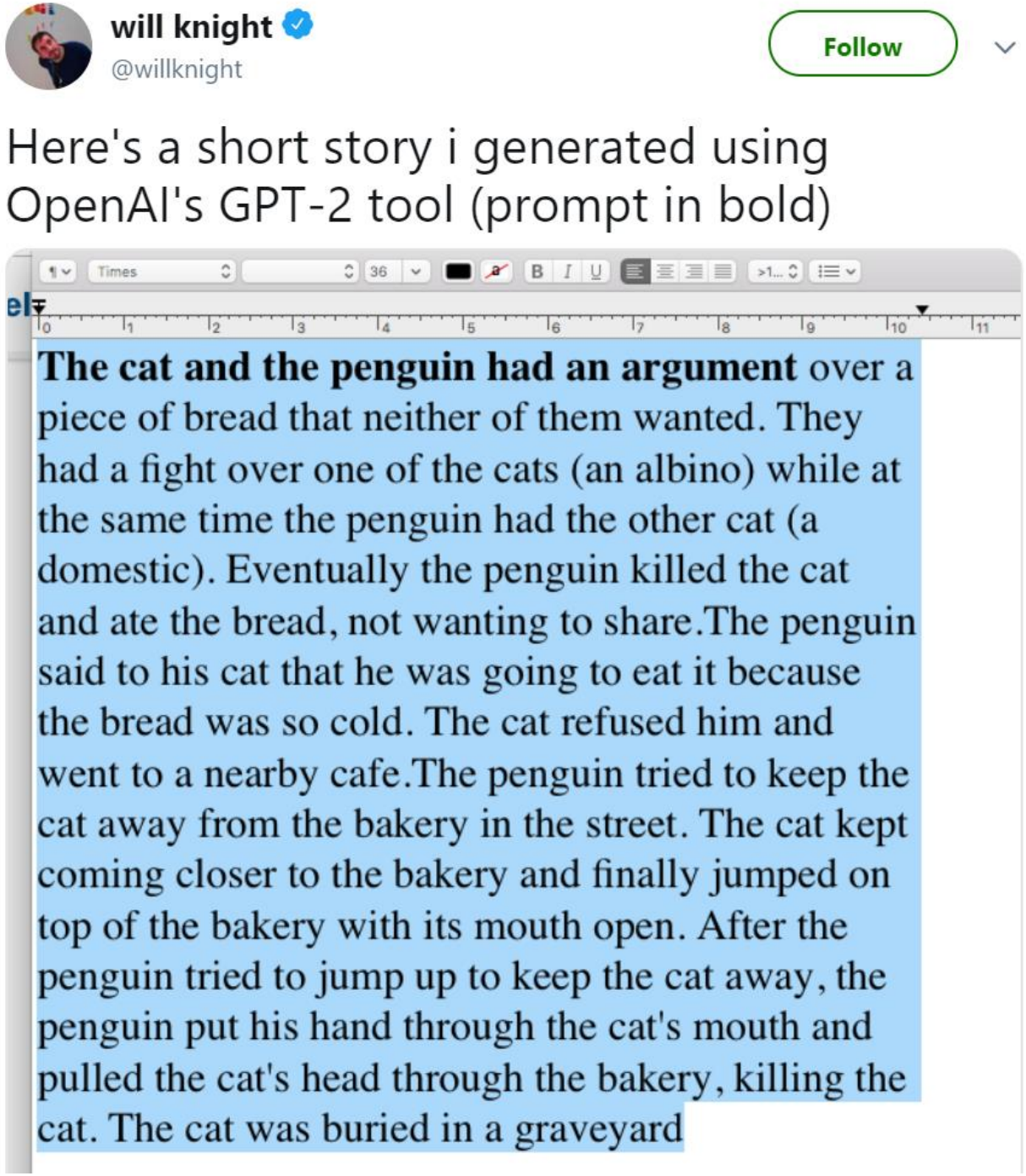
Image from: https://www.nowplayingutah.com/event/2018-vernalutah-storytelling-festival/

Automated Storytelling

Icon from flaticon.com by Freepik

Stories can…

- Help us plan
- Teach us
- Train us for hypothetical scenarios
- Do anything else that requires long-term context and commonsense information!


will knight
@willknight
Follow
Here's a short story i generated using OpenAI's GPT-2 tool (prompt in bold)
Times
36
The cat and the penguin had an argument over a piece of bread that neither of them wanted. They had a fight over one of the cats (an albino) while at the same time the penguin had the other cat (a domestic). Eventually the penguin killed the cat and ate the bread, not wanting to share.The penguin said to his cat that he was going to eat it because the bread was so cold. The cat refused him and went to a nearby cafe.The penguin tried to keep the cat away from the bakery in the street. The cat kept coming closer to the bakery and finally jumped on top of the bakery with its mouth open. After the penguin tried to jump up to keep the cat away, the penguin put his hand through the cat's mouth and pulled the cat's head through the bakery, killing the cat. The cat was buried in a graveyard

Plot Generation

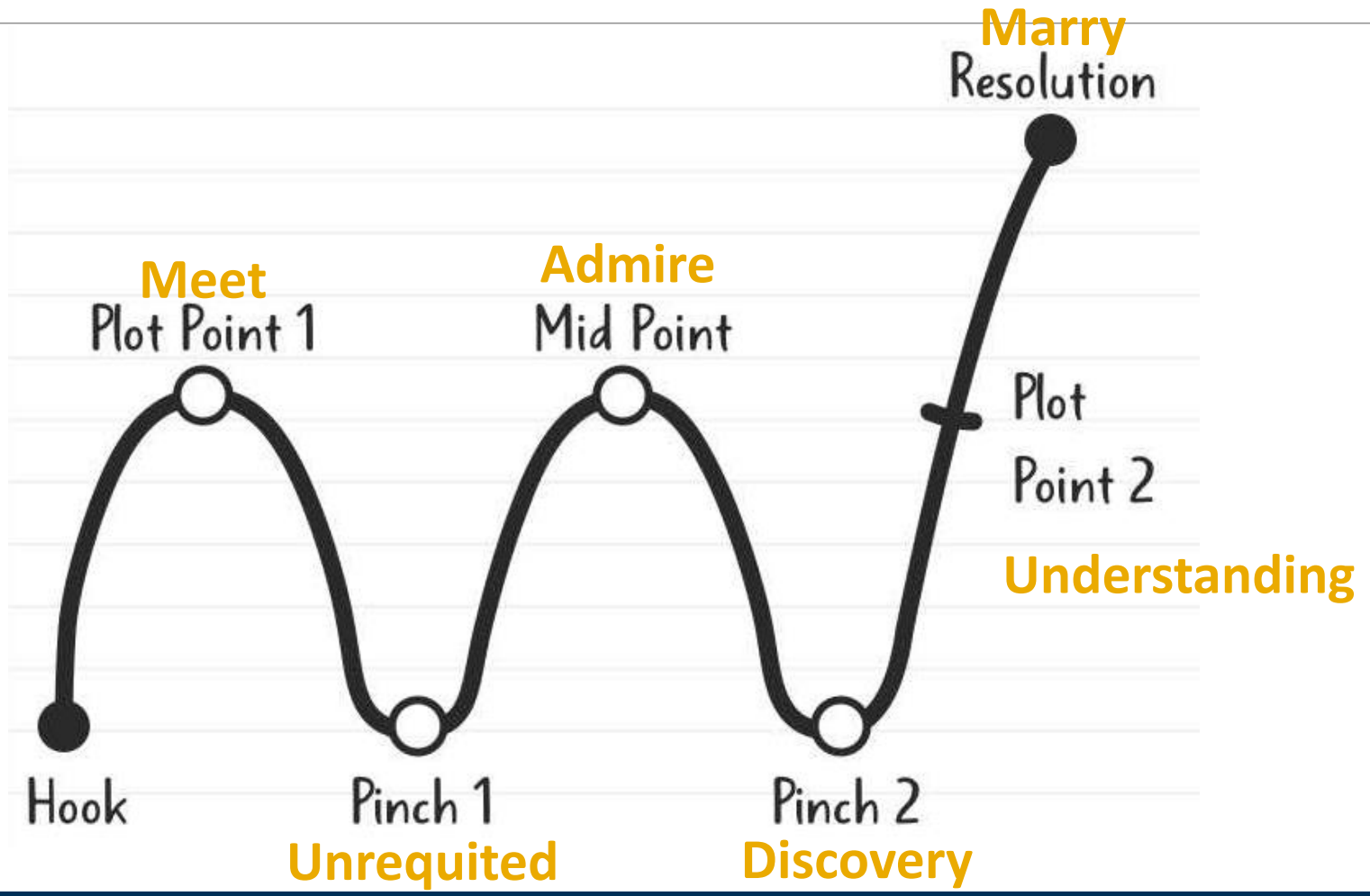


Image source: https://blog.reedsy.com/plot-point/

How can we make *controllable* neural storytellers?

Controllable Story Generation

We need a criteria for success → Reach a “goal verb”

- Given any start of the story, we want it to end a certain way
- E.g. “I want a story where…”
  - The bad guys *lose.*
  - The couple *marries*.

What we did:

We use reinforcement learning with **reward shaping**

to create a storytelling system

that can **incrementally** head toward a plot goal

Outline

1. The problem: generating a sequence of plot points
2. Reinforcement learning storytelling
3. Our reward shaping technique
4. Automated evaluation
5. Human evaluation

Event/Sentence Generation

Simonetta learns of Tito's affections for her.

She loved Tito before she loved Luigi.

Sentence Sparsity

Simonetta learns of Tito’s affections for her.

Problem: Sentences like this only appear once in the dataset

Solution: Fixing sparsity by separating semantics (meaning) from syntax (grammar)

Event Representation

⟨subject, verb, direct object, modifier⟩

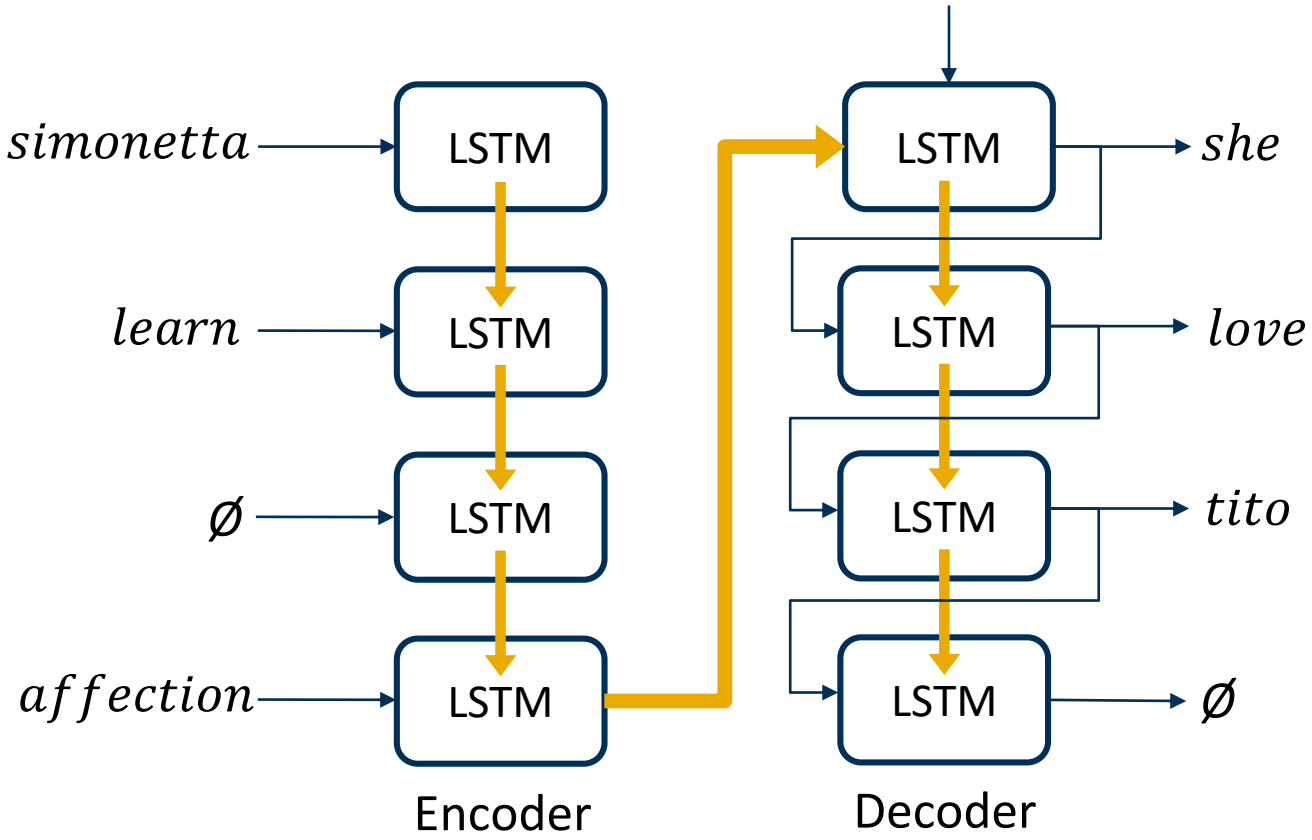
Original sentence: simonetta learns of tito s affections for her

Event: ⟨simonetta, learn, Ø, affection⟩

Generalized Event: ⟨<PERSON>0, learn-14-1, Ø, state.n.02⟩

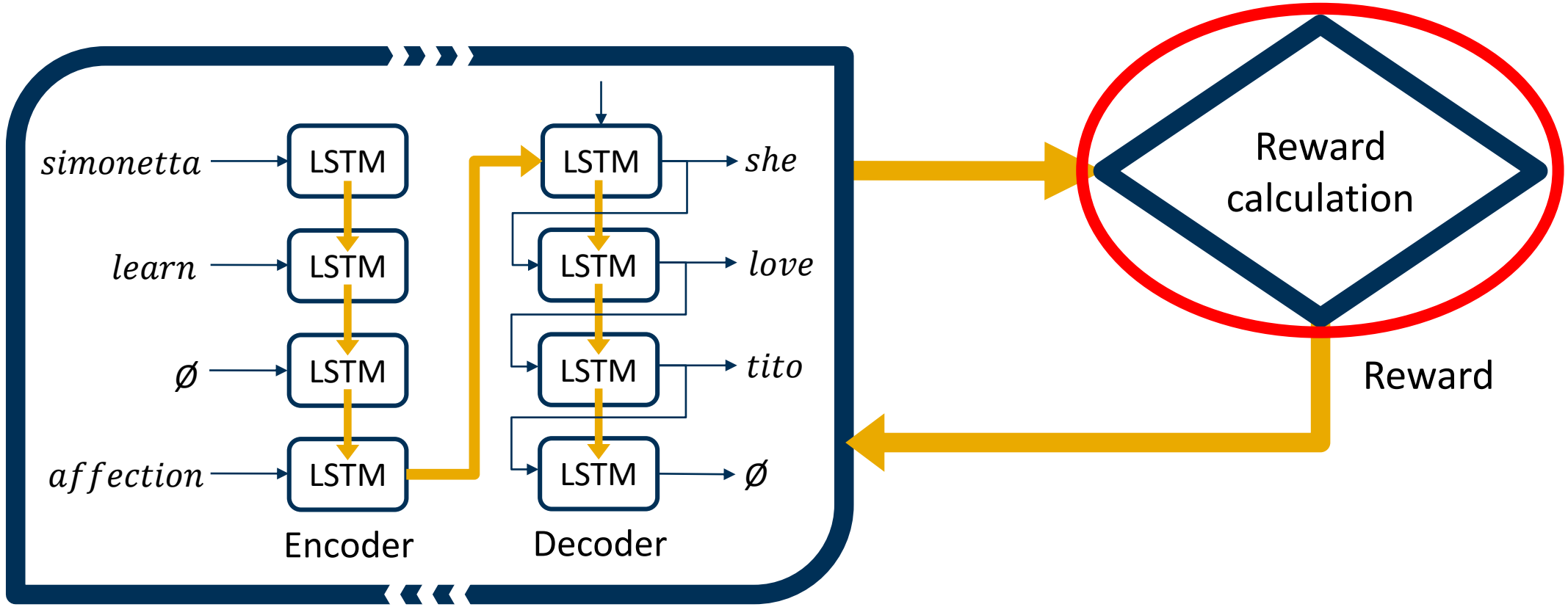
Martin, L. J., Ammanabrolu, P., Wang, X., Hancock, W., Singh, S., Harrison, B., & Riedl, M. O. (2018). Event Representations for Automated Story Generation with Deep Neural Nets. In *AAAI* (pp. 868–875).

Sequence-to-Sequence Refresher



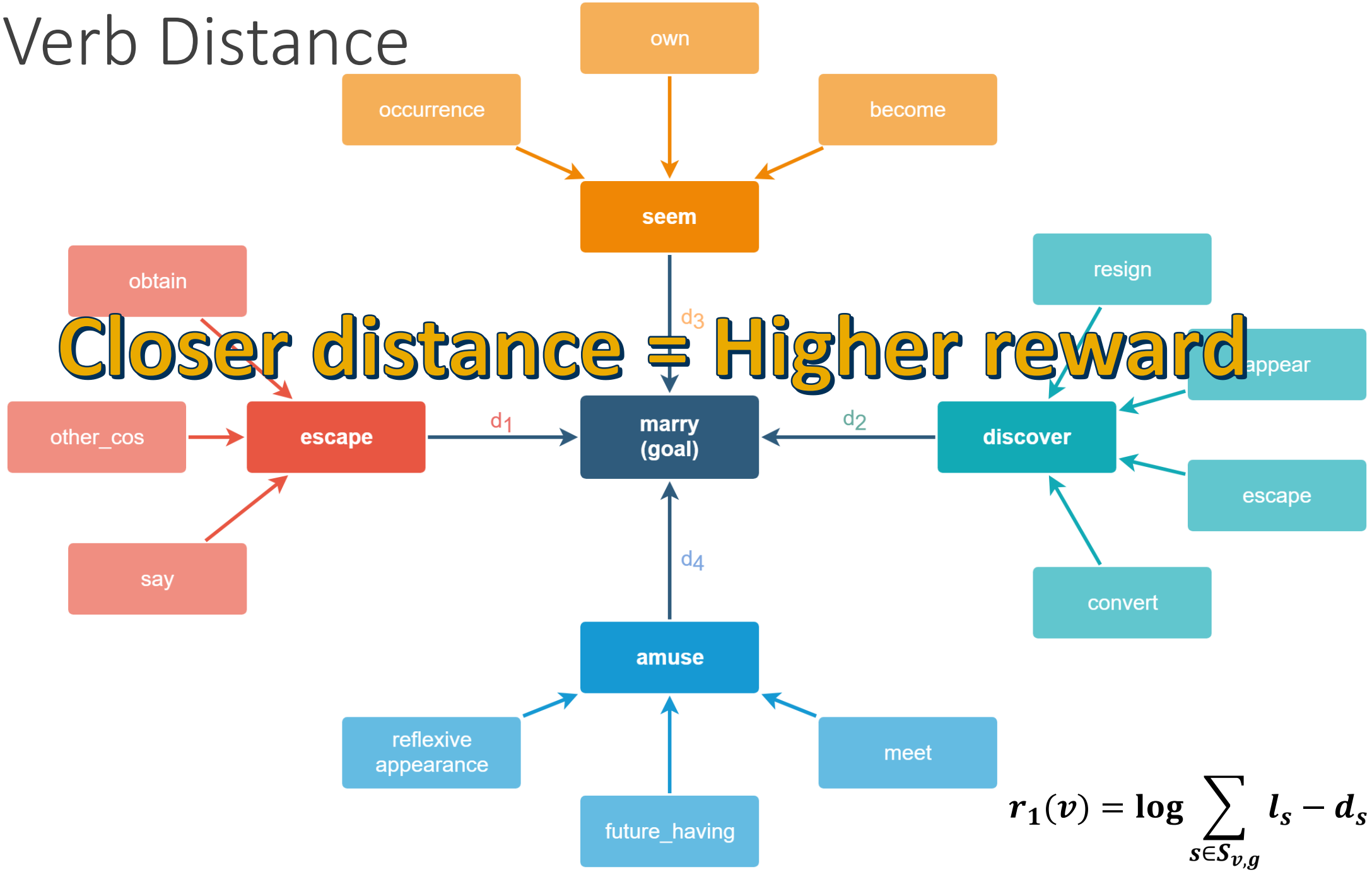
Sutskever, I., Vinyals, O., & Le, Q. V. (2014). Sequence to Sequence Learning with Neural Networks. In *Advances in Neural Information Processing Systems* (pp. 3104–3112).

REINFORCE (Seq2Seq++)



Williams, R. J. (1992). Simple Statistical Gradient-Following Algorithms for Connectionist Reinforcement Learning. *Machine Learning*, *8*(3), 229–256.

#1 Verb Distance

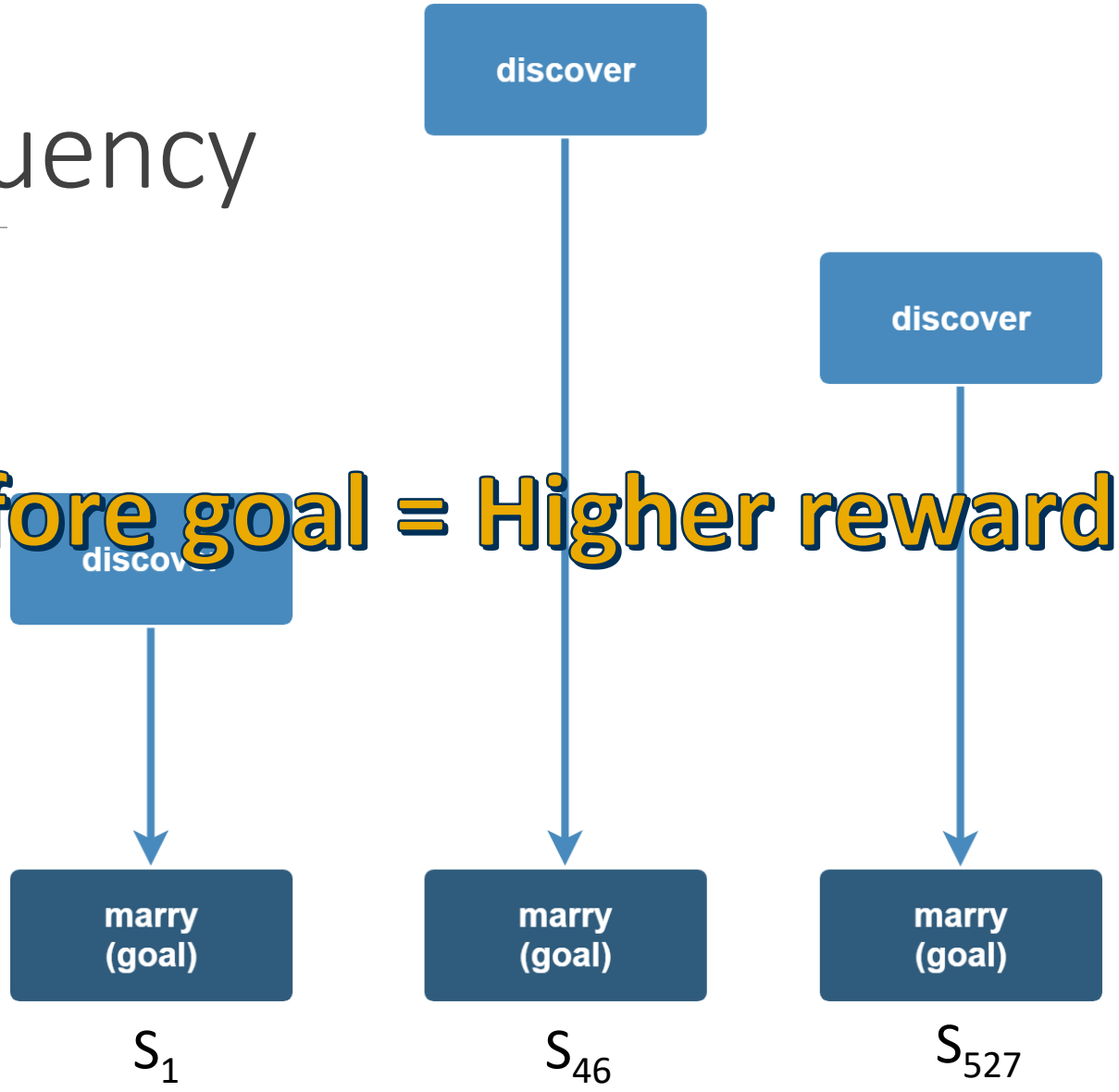


$$r_1(v) = \log \sum_{s \in S_{v,g}} l_s - d_s(v, g)$$

#2 Story-Verb Frequency

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

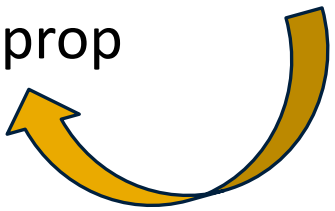
Appear frequently before goal = Higher reward



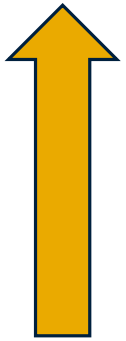
Final Reward Equation

Affects step size for backprop

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



Verb Distance to Goal



Story-Verb Frequency

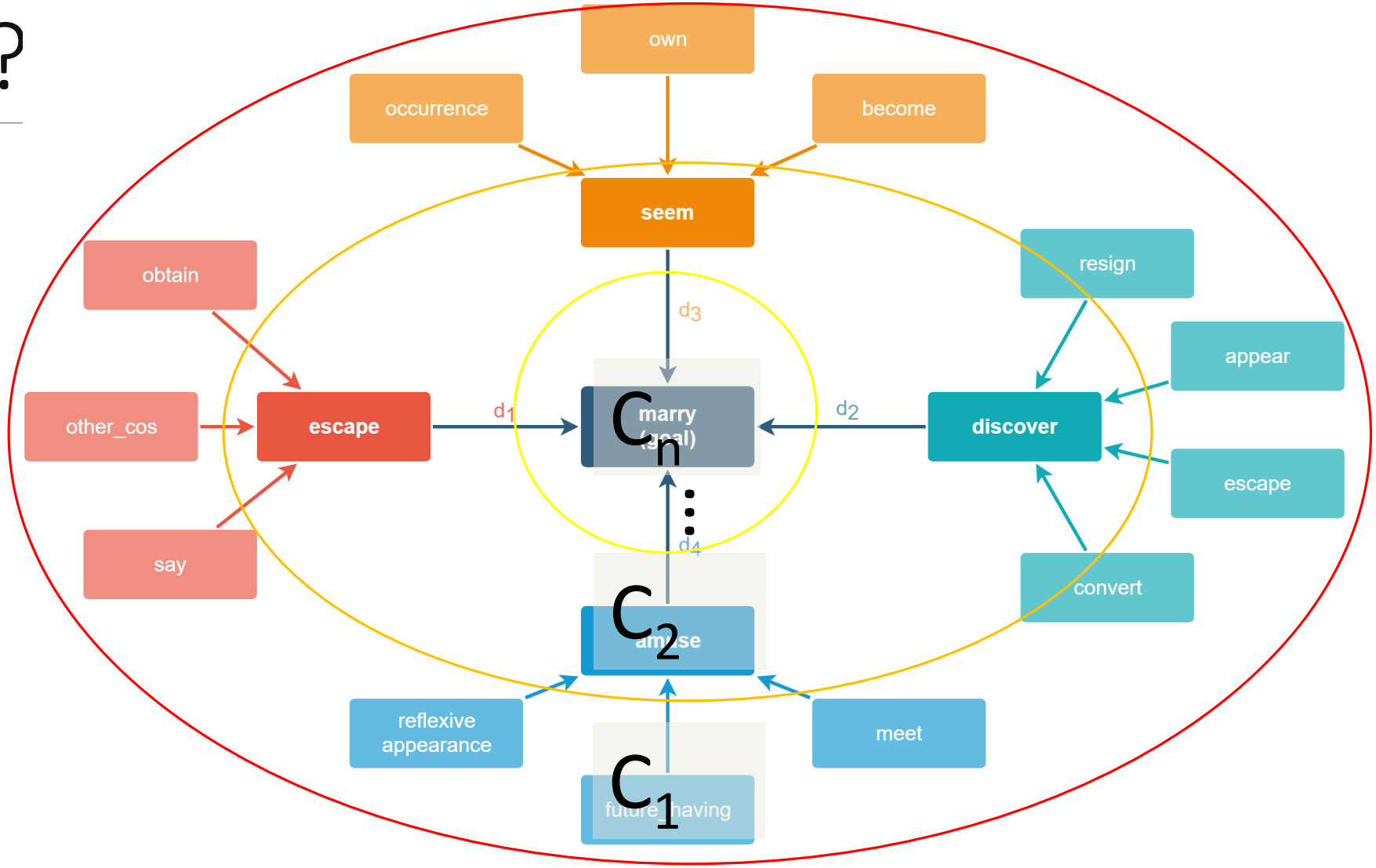
Results

Goal	Model	Average Story Length	Average Perplexity	Goal Achievement Rate
admire	Seq2Seq	7.11	48.06	35.52%
	REINFORCE	7.32	**5.73**	15.82%
marry	Seq2Seq	6.94	48.06	39.92%
	REINFORCE	7.38	**9.78**	24.05%

What now?

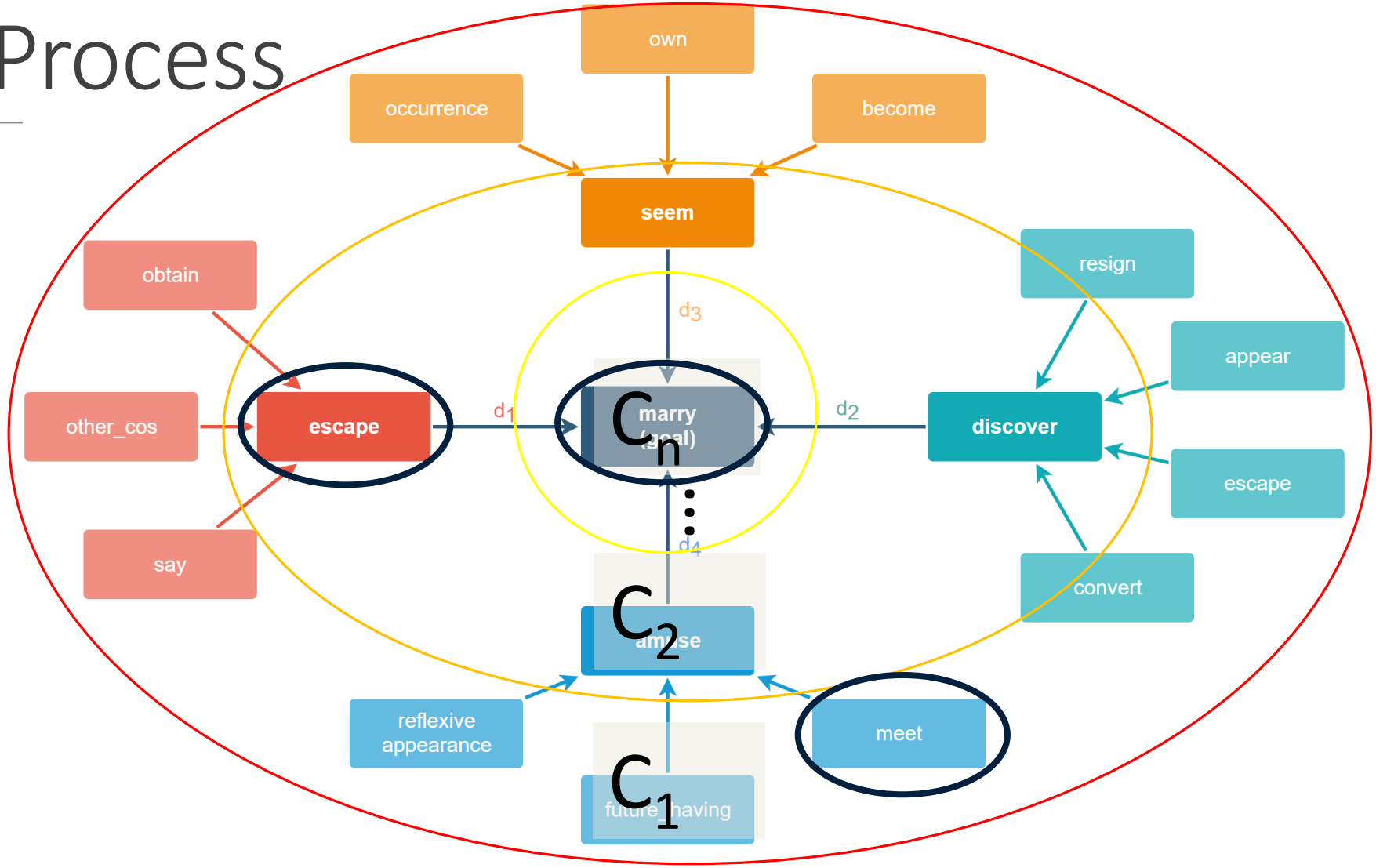
Cluster based on reward score

Constrain system to sample from next cluster



Clustering Process

1. Jenks Natural Breaks
2. Sample event
3. Replace verb if needed

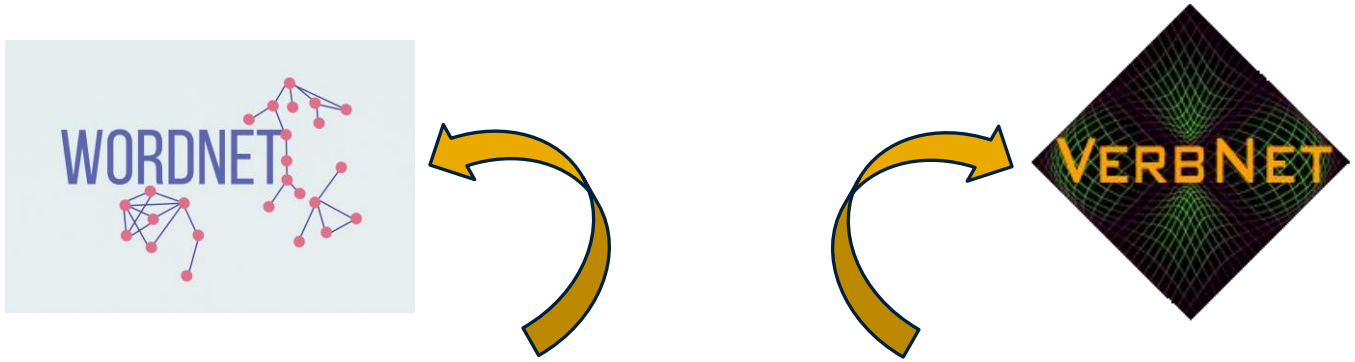


Results

Goal	Model	Average Story Length	Average Perplexity	Goal Achievement Rate
admire	Seq2Seq	7.11	48.06	35.52%
	REINFORCE	7.32	**5.73**	15.82%
	REINFORCE + Clustering	4.90	7.61	**94.29%**
marry	Seq2Seq	6.94	48.06	39.92%
	REINFORCE	7.38	9.78	24.05%
	REINFORCE + Clustering	5.76	**7.05**	**93.35%**

But are the stories actually any *good*?

Event Translation via Humans



⟨ relative.n.01, disappearance-48.2, Ø, Ø ⟩

My cousin died.

http://www.cs.princeton.edu/courses/archive/spring19/cos226/images/assignment-logos/600-by-400/wordnet.png
https://verbs.colorado.edu/verbnet/images/verbnet.gif

Example (Goal: hate/admire)

DRL Event Output ⟨ subject, verb, object, modifier ⟩	Translated Sentence
⟨ relative.n.01, disappearance-48.2, Ø, Ø ⟩	My cousin died.
⟨ NE1, say-37.7-1, visit, Ø ⟩	Alexander insisted on a visit.
⟨ NE1, meet-36.3-1, female.n.02, Ø ⟩	Alexander met her.
⟨ NE0, correspond-36.1, Ø, NE1 ⟩	Barbara commiserated with Alexander.
⟨ physical_entity.n.01, marry-36.2, Ø, Ø ⟩	They hugged.
⟨ group.n.01, contribute-13.2-2, Ø, LOCATION ⟩	The gathering dispersed to Hawaii.
⟨ gathering.n.01, characterize-29.2-1-1, time_interval.n.01, Ø ⟩	The community remembered their trip.
⟨ physical_entity.n.01, cheat-10.6, pack, Ø ⟩	They robbed the pack.
⟨ physical_entity.n.01, admire-31.2, social_gathering.n.01, Ø ⟩	They adored the party.

Human Evaluation Methods

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

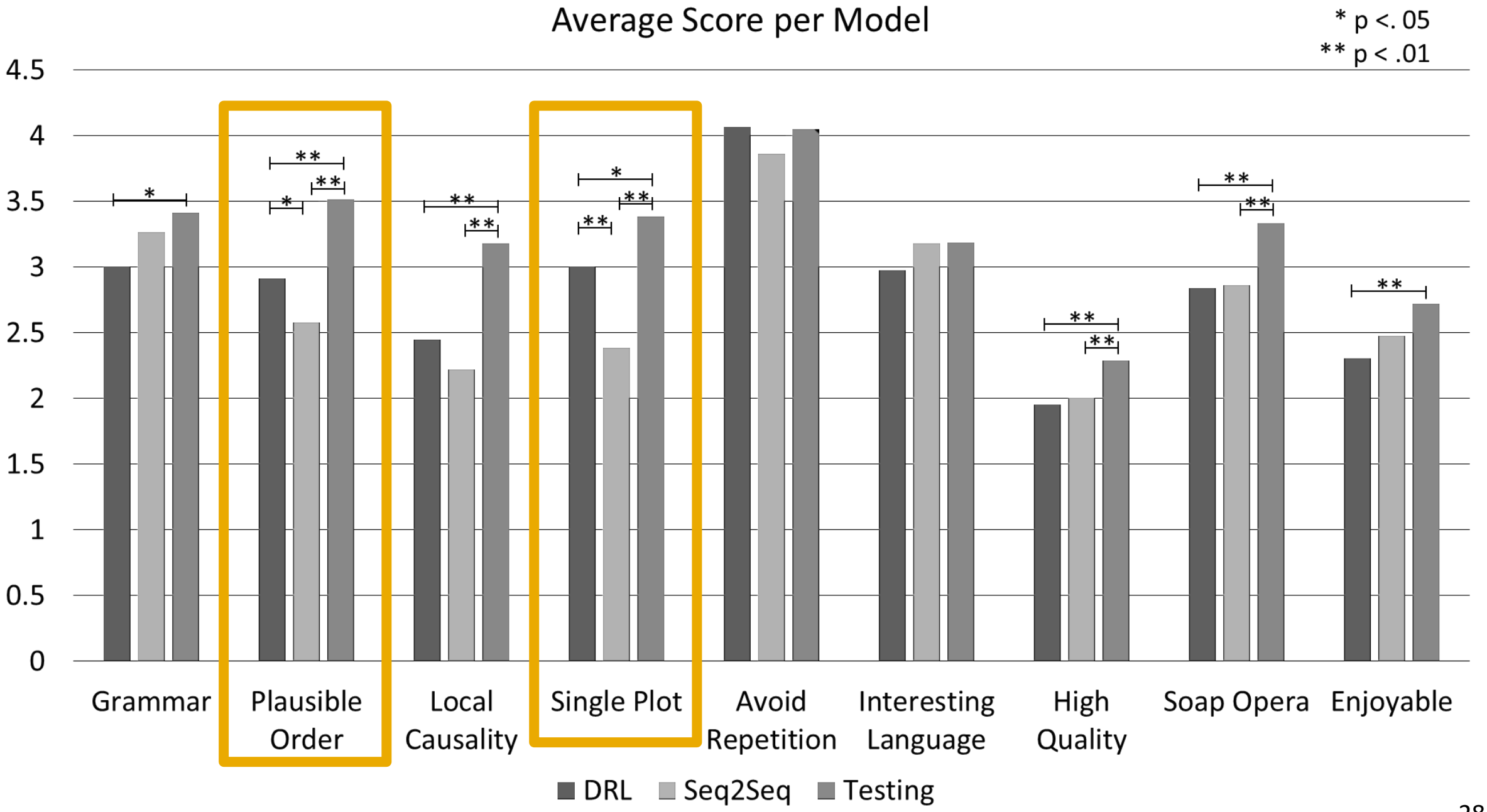
For each of 3 conditions:

- REINFORCE + Clustering (Ours)
- Baseline Seq2Seq
- Testing Set Stories (Translated Events; Gold Standard)

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.

Purdy, C., Wang, X., He, L., & Riedl, M. (2018). Towards Predicting Generated Story Quality with Quantitative Metrics. In *14th AAAI Conference on Artificial Intelligence and Interactive Digital Entertainment (AIIDE '18)*.


Average Score per Model
* p <. 05
** p < .01
4.5
4
3.5
3
2.5
2
1.5
1
0.5
0
Grammar
Plausible Order
Local Causality
Single Plot
Avoid Repetition
Interesting Language
High Quality
Soap Opera
Enjoyable
DRL
Seq2Seq
Testing

In Conclusion…

- Most neural storytelling methods lack “controllability”
- We used reinforcement learning to guide the story toward a goal (verb)
- Reward shaping and clustering → logical plot progression
- RL plots resulted in stories with more of a “single plot” and “plausible ordering” than Seq2Seq baseline

Thank you!

Read the paper on arXiv!
https://arxiv.org/abs/1809.10736

QUESTIONS?

LJMARTIN@GATECH.EDU / TWITTER: @LADOGNOME

entertainment intelligence lab