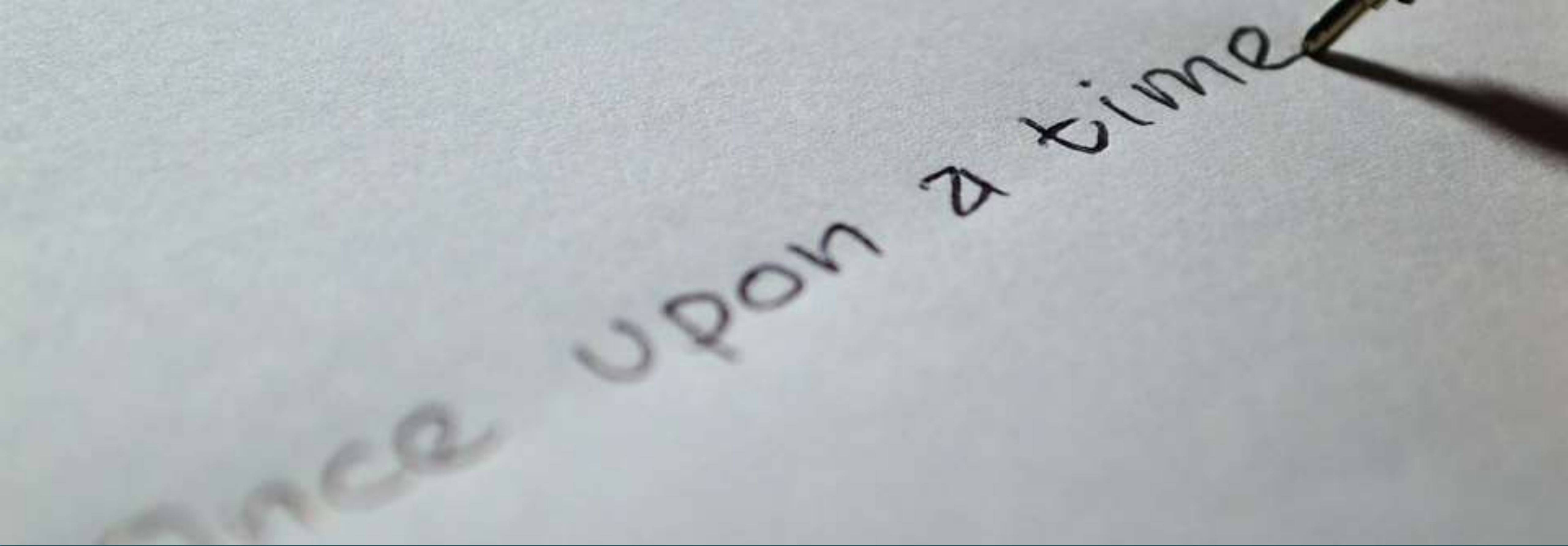

Controllable Neural Story Plot Generation via Reward Shaping

Tambwekar, Pradyumna, Murtaza Dhuliawala, Lara J. Martin, Animesh Mehta, Brent Harrison, and Mark O. Riedl. “Controllable Neural Story Plot Generation via Reward Shaping.” In *Proceedings of the Twenty-Eighth International Joint Conference on Artificial Intelligence*, 5982–88. Macao, China: International Joint Conferences on Artificial Intelligence Organization, 2019. https://doi.org/10.24963/ijcai.2019/829.

Agenda

Introduction

Traditional Language-modeling approaches struggle with story plot generation

Using a reward-shaping technique that analyzes a story corpus and provides intermediate rewards is shown to successfully guide a baseline model towards more plausible event ordering and consistent achievement of goals, as demonstrated by both automated and human evaluations.

CHALLENGES

Challenges in Story and Plot Generation

- Large neural language models work well with short-term tasks
- Prone to generating stories with limited direction and coherence
- RNNs tend to lose coherence in long-term context
- Most systems do not provide way for user to provide guidance on goal

Related Work: Early Story and Plot Generation Systems

- Early Story and Plot Generation Systems
  - Relied on symbolic planning and case-based reasoning
  - Generated stories for predetermined, well-defined domains
  - Provided long-term causal coherence
- Machine Learning Story Generation Techniques
  - Textual case-based reasoning trained on blogs
  - Probabilistic graphical models learned from crowdsourced example stories
  - Recurrent neural networks (RNNs) trained on large dataset of stories
- Recent Research Efforts
  - RNNs used to generate stories and plots
  - Solving the Story Cloze Test
  - Focus on generating entire stories to fit specified endings

Approach

- Using Reinforcement Learning (RL) for Plot Generation
  - RL agent learns a policy to maximize future expected reward
  - System generates plots similar to training corpus and moves plot towards a goal
- Primary Contribution: Reward-Shaping Technique
  - Reinforces weights in a neural language model
  - Guides generation of plot points towards a given goal
- Evaluation of Technique
  - Compared to standard language modeling technique
  - Human subject study for subjective ratings

Initial Language Model

Language Model

- Encoder-decoder network used as starting language model and baseline
- Trained to generate sequences of text for dialogue or story generation

Event Representation

- Improved predictive accuracy by using events derived from sentences
- Event is a tuple consisting of verb, subject, object, and additional significant noun
- Parameters can take special value of empty if no object or additional information

$$P(e_{i+1}|e_i;\theta) \text{ where } e_i = \langle s_i, v_i, o_i, m_i \rangle$$

Reinforcement Learning for Plot Generation

Sequence of events transitions the state of the world
- Goal: specific verb occurs in the final event

Reward Shaping
- Replace sparse rewards with dense reward signals
- Provide rewards at intermediate states rather than just only when goal is achieved
- Final rewards advantage significant verbs that appear close to target

Verb Clustering
- Jenks Natural Breaks Optimization to cluster verbs & prevent fast jump to goal

Experiments

- Data
  - CMU move summary corpus
  - Clusters into genre and subset of ones with soap-opera like plots to derive event sequence
- Model training
  - Encoder-decoder network
  - Hidden layer size of 1024
  - Pre-trained for 200 epochs
  - Mini-batch gradient descent with batch size of 64
  - Seq2Seq Model - baseline
  - DRL-clustered Model – continued training with RL with clustering and vocab restriction
  - DRL-unrestricted Model – training with RL without vocabulary restriction

Results and Discussion

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
	Unrestricted	15.82%	5.73	7.32
	Clustered	94.29%	7.61	4.90
marry	Test Corpus	24.64%	n/a	7.37
	Seq2Seq	39.92%	48.06	6.94
	Unrestricted	24.05%	9.78	7.38
	Clustered	93.35%	7.05	5.76

- Goal Achievement Rate
  - 22.47% of stories in the testing set end in desired goals
  - DRL-clustered model achieves goals 93.82% of the time
  - Baseline Seq2Seq achieves goals 37.72% of the time
  - DRL-unrestricted model achieves goals 19.95% -negatively affected by not using verb-clustering
- Average Story Length
  - Baseline Seq2Seq model produces stories similar in length to testing corpus

Human Evaluation: Experimental Setup

Leveraged human translators to translate LLM output events into short sentences

Questionnaire

- 175 participants recruited on Amazon Mechanical Turk
- Each participant compensated $10
- Translated plots rated on a 5-point Likert scale
- Rated included grammar, plausibility, coherence, repetition, language, quality, enjoyment, soap opera resemblance, and plot consistency

DRL stories perceived to have more plausible orderings than baseline

Conclusions

Traditional language models have challenges with stories with cohesive plots

Using reward shaping to provide rewards at intermediates steps can guide a models to generate stories that are cohesive and follow a goal