Stochastic Interactive Storytelling via Commonsense-Guided Reinforcement Learning

Team Members

Ninad Deo (UIN: 535005965)
Dhruvraj Singh Rathore (UIN: 735007897)

November 2025

---

1 Literature Survey

Project Summary

Our project models interactive storytelling as a *stochastic branching Markov Decision Process (MDP)*, in which an agent selects one of several candidate narrative continuations at each decision point. The state representation includes a sentence embedding of the current line plus character emotion/motivation features (from a commonsense narrative dataset). The action space consists of discrete candidate next sentences (one true continuation plus alternate generated options). The transition function is probabilistic, simulating narrative uncertainty (i.e., the same action may lead to different next states). The reward is designed to promote coherence, emotional consistency and satisfying story endings. We train the agent using a Deep Q-Network (DQN) with experience replay and a target network.

Review of Related Work

Interactive Narrative RL Wang et al. (2017) in "Interactive Narrative Personalization with Deep Reinforcement Learning" propose a Q-network–based framework that personalizes story events in the educational narrative CRYSTAL ISLAND. They model the Adaptive Event Sequence (AES) selection process as a sequential decision problem and train multiple Deep Q-Networks (DQN) to optimize narrative planning actions using synthetic interaction data from an LSTM-based bipartite player simulation model. The state space comprises 25 handcrafted features (player action history, questionnaire data, and event encodings), and rewards are computed from normalized learning gain (NLG). Stabilization techniques such as experience replay, target networks, and asynchronous gradient descent are incorporated to ensure convergence.

Our work is similar in using reinforcement learning for narrative control and deep Q-networks for decision optimization. However, it differs by representing states with sentence embeddings and emotion–motivation features instead of handcrafted player metrics, employing discrete sentence-level continuation actions rather than event-detail responses, and introducing stochastic branching transitions to capture narrative uncertainty rather than deterministic educational event flows.

Goal-Directed Generation with RL Alabdulkarim et al. (2021) in "Goal-Directed Story Generation: Augmenting Generative Language Models with Reinforcement Learning" fine-tune a transformer-based GPT-2 model using Proximal Policy Optimization (PPO) and reward shaping to make story generation goal-oriented. Their system constructs a knowledge-graph world model from story text using OpenIE triples ⟨subject, relation, object⟩, which defines the state space, and trains a policy network with graph attention to select candidate continuations generated by GPT-2 that best advance the story toward a defined goal verb class. The reward function is computed from VerbNet-based verb clusters and co-occurrence distances to the goal event, achieving high goal-completion rates (98%).

Our project is similar in modelling storytelling as a sequential decision process with reinforcement learning and discrete continuation selection. However, unlike their continuous token-level policy optimization, we frame storytelling as a stochastic MDP with discrete candidate actions, sentence-level state embeddings (not full knowledge graphs), and branching transitions to capture narrative uncertainty rather than deterministic goal achievement.

Coherence via Skeleton Models Xu et al. (2018) propose a skeleton-based reinforcement learning model to improve sentence-to-sentence coherence in narrative story generation. Their system decomposes story generation into two coupled modules: a skeleton extraction module, which automatically identifies key phrases representing the semantic core of each sentence, and a skeleton-based generative module, consisting of an input-to-skeleton encoder-decoder and a skeleton-to-sentence decoder, both implemented using LSTM-based Seq2Seq architectures with attention. Because skeleton extraction involves discrete word selection, they employ a policy-gradient reinforcement learning framework that jointly optimizes extraction and generation through feedback rewards based on reconstruction loss. This dual optimization yields a 20% improvement in human-rated coherence (G-score) compared to Seq2Seq baselines.

Our work is similar in using reinforcement learning to enhance narrative coherence and in treating sentence-level structure as a learning signal. However, unlike their deterministic sequential model, we model storytelling as a stochastic branching MDP with discrete continuation actions, using sentence embeddings and emotional/motivational features rather than learned skeletons as state representations, thereby capturing uncertainty in story progression rather than relying on single linear continuation paths.

Visual Storytelling with Hierarchical RL Huang et al. (2018) propose a Hierarchically Structured Reinforcement Learning (HSRL) framework to generate topically coherent multi-sentence stories from image sequences in the VIST dataset. Their model employs a two-level decoder: a Manager LSTM that plans high-level semantic subgoals (topics) for each image, and a Worker Semantic Compositional Network (SCN) that generates sentences conditioned on those topics. The two modules are trained jointly via a mixed maximum-likelihood and self-critical RL objective, where rewards are based on sequence-level metrics such as CIDEr. This hierarchical setup enables the model to capture both global narrative structure and local fluency, outperforming flat RL baselines by up to 30% on coherence metrics.

Our work shares the use of reinforcement learning for hierarchical narrative planning but differs by targeting text-only storytelling rather than visual inputs, representing states via sentence embeddings and emotional/motivational features instead of image or topic vectors, and introducing stochastic branching transitions across candidate continuations rather than deterministic topic-to-sentence mappings.

Uncertainty Branching in RL Weber et al. (2017) propose the Imagination-Augmented Agent (I2A), a hybrid architecture that integrates model-based and model-free reinforcement learning by combining real environment observations with simulated rollouts from a learned environment model. The model uses an imagination core that predicts future observations and rewards, an LSTM-based rollout encoder to interpret imagined trajectories, and a policy network trained with Asynchronous Advantage Actor-Critic (A3C). The approach achieves superior data efficiency and robustness on domains like Sokoban and MiniPacman, even with imperfect environment models.

Our work shares the concept of leveraging model-based imagination to enhance policy learning but differs by applying it to stochastic narrative environments rather than visual games. We use sentence-level state embeddings and discrete branching transitions to capture narrative uncertainty, instead of frame-based predictions and continuous control in deterministic environments.

RL for Text Generation with Sparse Reward Guo et al. (2022) present a Soft Q-Learning (SQL) framework for text generation that combines the strengths of on-policy and off-policy reinforcement learning through Path Consistency Learning (PCL). Their model reinterprets the generation logits as Q-values, enabling stable updates across all candidate actions simultaneously, even in large vocabulary spaces. The SQL objective integrates entropy-regularized rewards and supports training from noisy or negative examples, adversarial generation, and controllable prompt learning without maximum-likelihood pretraining.

Our work shares the goal of improving RL efficiency for text-based generation but applies these ideas to stochastic narrative branching rather than token-level continuation. While Guo et al. address sparse reward and stability via soft Q-value propagation, we operate in a discrete, story-level MDP, using sentence embeddings and narrative rewards (coherence and emotion alignment) instead of token-level entropy optimization.

Knowledge-Enhanced Story Gen Guan et al. (2020) propose a knowledge-enhanced pretraining framework built on GPT-2, designed to improve long-range coherence and logical consistency in commonsense story generation. Their approach integrates external knowledge bases—ConceptNet and ATOMIC— by transforming if-then triples into natural-language sentences for post-training, allowing the model to capture causal and temporal dependencies between events. They further introduce multi-task learning, combining a language modeling objective with a classification loss that distinguishes true stories from synthetically corrupted ones (via sentence shuffling or repetition). This dual objective helps the model learn causal and temporal reasoning implicitly, improving BLEU, distinct-n, and coherence metrics on ROC-Stories.

Our project shares their focus on commonsense-driven narrative coherence but differs fundamentally in framing storytelling as a stochastic MDP trained via reinforcement learning rather than supervised fine-tuning. Instead of implicit knowledge learning through pretraining, we explicitly model state transitions, rewards, and actions at the sentence level, using emotion- and motivation-aware embeddings and branching continuations to represent narrative uncertainty.

Branching RL Formalism Du et al. (2022) formalize Branching Reinforcement Learning (Branching RL) as a generalization of standard RL in which the agent observes only a single trajectory from a branching stochastic process but must infer an underlying tree-structured transition model. They derive sample-complexity bounds for both tabular and function-approximation settings, showing that the optimal policy can be efficiently learned with polynomial samples under certain realizability assumptions. The paper introduces the Branching Contextual Decision Process (BCDP) framework, which handles non-sequential dependencies among latent branches using Bellman-consistent value estimation.

Our work aligns conceptually with this formulation by viewing narrative generation as a branching decision process where a single continuation may probabilistically lead to multiple next states. However, we apply this idea to a text-based storytelling environment, using sentence embeddings as states and Deep Q-Networks for learning, rather than focusing on theoretical guarantees or value-function estimation bounds.

LLM-Based Branching Narratives Leandro et al. (2023) introduce GENEVA, a two-stage system that uses GPT-4 to generate and visualize branching narrative graphs for storytelling and game design. The model first produces multiple interconnected storylines—each composed of "narrative beats," defined as minimal story units—under designer-specified constraints (e.g., number of starts, endings, and storylines). It then translates these storylines into a Directed Acyclic Graph (DAG) representation using structured prompting for node-edge generation compatible with a D3.js visualization tool. By iteratively prompting GPT-4 to create one storyline at a time, GENEVA ensures branching and reconverging structures across narratives grounded in user-defined contexts (e.g., Frankenstein in the 21st century).

Our project aligns with GENEVA in its goal of modeling branching narrative structures, but differs fundamentally in formulation: GENEVA relies on prompt-based generation and visualization without a reinforcement-learning component, whereas our work frames story progression as a stochastic MDP trained via Deep Q-Learning, where each continuation represents an action and transitions are governed by probabilistic rewards tied to coherence and emotional consistency.

Multimodal and Storyline-Guided Generation Kim et al. (2023) present a multi-modal story generation framework that combines BERT-based storyline guidance, GPT-2-based paragraph generation, and a diffusion-based visualization model. The system first predicts the next storyline entities—characters, events, and places—via a multiple-choice question answering (MCQA) task using fine-tuned BERT embeddings. These predicted entities guide a GPT-2 decoder to generate coherent paragraphs, while a latent diffusion model visualizes the resulting scenes. The framework is trained and evaluated on the Storium dataset, achieving higher BLEU, Recall@k, and BERTScore-coherence scores compared to baseline models.

Our work is related through its use of multi-stage planning for narrative coherence, yet differs in formulation and learning objectives. While their model operates through supervised sequence prediction and multi-modal grounding, ours frames storytelling as a stochastic MDP trained with Deep Q-Learning, emphasizing reinforcement-based coherence rewards and probabilistic branching rather than entity-driven deterministic planning.

Summary

The reviewed literature demonstrates the feasibility of using reinforcement learning for narrative generation, interactive storytelling and branching decision-making. Nevertheless, three gaps remain: (1) explicit modelling of discrete candidate continuation actions at sentence-level branching points, (2) incorporation of emotion and motivation features into state representations, and (3) explicit stochastic transition modelling in narrative decision trajectories. Our project addresses these gaps by leveraging a commonsense narrative dataset with emotion/motivation annotations, designing discrete candidate selection actions, representing states via sentence embeddings plus character attributes, and modelling branching transitions stochastically within a DQN training framework.

References

1. Wang, P., Rowe, J., Min, W., Mott, B., & Lester, J. (2017). Interactive Narrative Personalization with Deep Reinforcement Learning. *Proceedings of the 26th International Joint Conference on Artificial Intelligence (IJCAI-17)*, 3852-3858.

2. Alabdulkarim, A., Li, W., Martin, L. J., & Riedl, M. O. (2021). Goal-Directed Story Generation: Augmenting Generative Language Models with Reinforcement Learning. arXiv preprint arXiv:2112.08593.

3. Xu, J., Ren, X., Zhang, Y., Zeng, Q., Cai, X., & Sun, X. (2018). A Skeleton-Based Model for Promoting Coherence Among Sentences in Narrative Story Generation. arXiv preprint arXiv:1808.06945.

4. Huang, Q., Gan, Z., Celikyilmaz, A., Wu, D., Wang, J., & He, X. (2018). Hierarchically Structured Reinforcement Learning for Topically Coherent Visual Story Generation. arXiv preprint arXiv:1805.08191.

5. Weber, T., Racanière, S., Reichert, D. P., Buesing, L., Guez, A., Jimenez Rezende, D., & Silver, D. (2017). Imagination-Augmented Agents for Deep Reinforcement Learning. arXiv preprint arXiv:1707.06203.

6. Guo, H., Tan, B., Liu, Z., Xing, E. P., & Hu, Z. (2021). Efficient (Soft) Q-Learning for Text Generation with Limited Good Data. arXiv preprint arXiv:2106.07704.

7. Guan, J., Huang, F., Zhao, Z., Zhu, X., & Huang, M. (2020). A Knowledge-Enhanced Pretraining Model for Commonsense Story Generation. arXiv preprint arXiv:2001.05139.

8. Du, Y. (2022). Branching Reinforcement Learning (Branching RL). arXiv preprint arXiv:2202.07995.

9. Leandro, J. (2023). Generating and Visualizing Branching Narratives using Large Language Models. arXiv preprint arXiv:2311.09213.

10. Kim, J., Heo, Y., Yu, H., & Nang, J. (2022). A Multi-Modal Story Generation Framework with AI-Driven Storyline Guidance. Electronics, 12(6), 1289.