Novelty Assessment Report

Paper: StoryAlign: Evaluating and Training Reward Models for Story Generation

PDF URL: https://openreview.net/pdf?id=a3JmkJtTDV

Venue: ICLR 2026 Conference Submission

Year: 2026

Report Generated: 2026-01-01

Abstract

Story generation aims to automatically produce coherent, structured, and engaging narratives. Although large language models (LLMs) have significantly advanced text generation, stories generated by LLMs still diverge from human-authored works regarding complex narrative structure and human-aligned preferences. A key reason is the absence of effective modeling of human story preferences, which are inherently subjective and under-explored. In this work, we systematically evaluate the modeling of human story preferences and introduce StoryRMB, the first benchmark for assessing reward models on story preferences. StoryRMB contains $1,133$ high-quality, human-verified instances, each consisting of a prompt, one chosen story, and three rejected stories. We find existing reward models struggle to select human-preferred stories, with the best model achieving only 66.3% accuracy. To address this limitation, we construct roughly $100,000$ high-quality story preference pairs across diverse domains and develop StoryReward, an advanced reward model for story preference trained on this dataset. StoryReward achieves state-of-the-art (SoTA) performance on StoryRMB, outperforming much larger models. We also adopt StoryReward in downstream test-time scaling applications for best-of-n (BoN) story selection and find that it generally chooses stories better aligned with human preferences. We will release our dataset, model, and code to facilitate future research.

Core Task Landscape

This paper addresses: **Evaluating and Training Reward Models for Story Generation**

A total of **50 papers** were analyzed and organized into a taxonomy with **20 categories**.

Taxonomy Overview

The research landscape has been organized into the following main categories:

- **Reward Model Architecture and Training Methodology**
- **Reward-Guided Generation and Decoding Strategies**
- **Reinforcement Learning for Story and Creative Text Generation**
- **Reward Model Evaluation and Benchmarking**
- **Intermediate Rewards and Critique-Based Methods**
- **Reward Models for Controlled Generation and Safety**
- **Surveys and Theoretical Perspectives**
- **Visual Storytelling and Multimodal Narrative Generation**
- **Specialized Applications and Domains**

Complete Taxonomy Tree

- Evaluating and Training Reward Models for Story Generation Survey Taxonomy
- Reward Model Architecture and Training Methodology
  - Reward Model Design and Parametrization (3 papers)
  - [4] On the Low-Rank Parametrization of Reward Models for Controlled Language Generation (Troshin, 2024) View paper
  - [26] GRAM: A Generative Foundation Reward Model for Reward Generalization (Wang Cheng-long, 2025) View paper
  - [30] Efficient controlled language generation with low-rank autoregressive reward models (S Troshin, 2024) View paper
  - Reward Model Training and Learning Frameworks (3 papers)
  - [15] Helpsteer 2: Open-source dataset for training top-performing reward models (Olivier Delalleau, 2024) View paper
  - [49] Self-Generated Critiques Boost Reward Modeling for Language Models (Yue Yu, 2024) View paper
  - [50] Probabilistic Uncertain Reward Model (Cheng Xiang, 2025) View paper
  - Multimodal and Domain-Specific Reward Models (4 papers)
  - [1] Imagereward: Learning and evaluating human preferences for text-to-image generation (XU Jiazheng, 2023) View paper
  - [6] Visionreward: Fine-grained multi-dimensional human preference learning for image and video generation (XU Jiazheng, 2024) View paper
  - [23] Multimodal LLMs as Customized Reward Models for Text-to-Image Generation (Zhou Shi-jie, 2025) View paper
  - [45] Unified Reward Model for Multimodal Understanding and Generation (Wang Yi-bin, 2025) View paper
  - Context-Aware and Personalized Reward Modeling (3 papers)
  - [3] RAGferee: Building contextual reward models for retrieval-augmented generation (Andrei Catalin Coman, 2025) View paper
  - [29] Personalized Reward Modeling for Text-to-Image Generation (Jeongeun Lee, 2025) View paper
  - [47] RoleRMBench & RoleRM: Towards Reward Modeling for Profile-Based Role Play in Dialogue Systems (Hang Ding, 2025) View paper
- Reward-Guided Generation and Decoding Strategies
  - Test-Time Alignment and Reward-Guided Decoding (4 papers)
  - [7] A critical look at tokenwise reward-guided text generation (Rashid Ahmad, 2024) View paper

- [9] Args: Alignment as reward-guided search (Burapacheep, 2024) View paper
- [12] GenARM: Reward Guided Generation with Autoregressive Reward Model for Test-time Alignment (Xu, 2024) View paper
- [16] Reward-augmented decoding: Efficient controlled text generation with a unidirectional reward model (Deng Hai-kang, 2023) View paper
- Search-Based and Sampling Methods with Rewards (1 papers)
- [13] Cos (m+ o) s: Curiosity and rl-enhanced mcts for exploring story space via language models (Materzok, 2025) View paper

- Reinforcement Learning for Story and Creative Text Generation
  - RL with Human Feedback for Story Generation ★ (4 papers)
  - [0] StoryAlign: Evaluating and Training Reward Models for Story Generation (Anon et al., 2026) View paper
  - [2] Learning to reason for long-form story generation (Gurung, 2025) View paper
  - [24] Recursively summarizing books with human feedback (Wu, 2021) View paper
  - [32] BabyStories: Can Reinforcement Learning Teach Baby Language Models to Write Better Stories? (Zhao, 2023) View paper
  - Adversarial and Automated Reward Learning for Stories (3 papers)
  - [20] No metrics are perfect: Adversarial reward learning for visual storytelling (Xin Wang, 2018) View paper
  - [40] Show, reward and tell: Automatic generation of narrative paragraph from photo stream by adversarial training (Jing Wang, 2018) View paper
  - [43] Show, reward, and tell: Adversarial visual story generation (J Tang, 2019) View paper
  - Reward Shaping and Goal-Directed Story Generation (3 papers)
  - [35] Goal-directed story generation: Augmenting generative language models with reinforcement learning (Alabdulkarim, 2021) View paper
  - [38] From Plots to Endings: A Reinforced Pointer Generator for Story Ending Generation (Zhao Yan, 2018) View paper
  - [44] Controllable neural story plot generation via reward shaping (Pradyumna Tambwekar, 2018) View paper
  - Hierarchical and Structured RL for Narrative Coherence (2 papers)
  - [19] Hierarchically structured reinforcement learning for topically coherent visual story generation (Qiuyuan Huang, 2019) View paper
  - [34] Narrative order aware story generation via bidirectional pretraining model with optimal transport reward (Lu, 2023) View paper
  - Multi-Objective and Diversity-Aware RL (3 papers)
  - [8] Jointly reinforcing diversity and quality in language model generations (Li Tianjian, 2025) View paper
  - [17] Rlmr: Reinforcement learning with mixed rewards for creative writing (Liao Jian-xing, 2025) View paper
  - [28] What makes a good story? designing composite rewards for visual storytelling (Junjie Hu, 2020) View paper
  - Preference Learning and Contrastive RL for Stories (2 papers)
  - [10] Cpo: Addressing reward ambiguity in role-playing dialogue via comparative policy optimization (Jing Ye, 2025) View paper
  - [41] Robust preference learning for storytelling via contrastive reinforcement learning (Castricato, 2022) View paper
  - Style-Conditioned and Authorial RL for Stories (1 papers)
  - [42] Capturing Classic Authorial Style in Long-Form Story Generation with GRPO Fine-Tuning (Jinlong Liu, 2025) View paper
- Reward Model Evaluation and Benchmarking
  - Reward Model Benchmarks and Evaluation Frameworks (1 papers)
  - [37] EvolvR: Self-Evolving Pairwise Reasoning for Story Evaluation to Enhance Generation (Wang Xinda, 2025) View paper
  - Analysis of Reward Model Behavior and Biases (2 papers)
  - [14] A long way to go: Investigating length correlations in rlhf (Singhal, 2023) View paper
  - [18] ReMoDetect: Reward Models Recognize Aligned LLM's Generations (Hyunseok Lee, 2024) View paper
- Intermediate Rewards and Critique-Based Methods (2 papers)
  - [11] Beyond sparse rewards: Enhancing reinforcement learning with language model critique in text generation (Cao Meng, 2024) View paper
  - [46] Teacher Forcing Recovers Reward Functions for Text Generation (Hao, 2022) View paper
- Reward Models for Controlled Generation and Safety (2 papers)
  - [5] Pcgrllm: Large language model-driven reward design for procedural content generation reinforcement learning (Baek, 2025) View paper
  - [25] Reward modeling for mitigating toxicity in transformer-based language models (Faal, 2023) View paper
- Surveys and Theoretical Perspectives (4 papers)
  - [21] Deep reinforcement learning and creativity (Franceschelli, 2025) View paper
  - [31] A survey of deep learning applied to story generation (Chenglong Hou, 2019) View paper
  - [33] Automatic story generation: A survey of approaches (AI Alhussain, 2021) View paper
  - [39] Transforming human interactions with AI via reinforcement learning with human feedback (RLHF) (Liu, 2023) View paper
- Visual Storytelling and Multimodal Narrative Generation (1 papers)
  - [22] StoryLLaVA: enhancing visual storytelling with multi-modal large language models (L Yang, 2025) View paper
- Specialized Applications and Domains (3 papers)
  - [27] Curiosity-Driven Reinforcement Learning from Human Feedback (Haoran Sun, 2025) View paper
  - [36] Automated scenario generation to support competency-based experiential learning in GIFT (A Smith, 2024) View paper
  - [48] Formalizing Adaptive Team Feedback in Synthetic Training Environments with Reinforcement Learning (A Smith, 2022) View paper

Narrative

Core task: evaluating and training reward models for story generation. The field organizes around several complementary branches that address different facets of guiding narrative systems toward human preferences. One branch focuses on reward model architecture and training methodology, exploring how to build robust scoring functions from human feedback or other signals (e.g., ImageReward[1], VisionReward[6]). A second branch examines reward-guided generation and decoding strategies, investigating how to integrate learned rewards into sampling or search procedures (e.g., Reward-Augmented Decoding[16]). A third major area is reinforcement learning for story and creative text generation, where policy optimization techniques adapt language models to produce coherent, engaging narratives (e.g., Hierarchical Story Generation[19], Recursively Summarizing Books[24]). Additional branches cover evaluation and benchmarking of reward models (e.g., RoleRMBench[47]), intermediate rewards and critique-based methods that provide step-by-step guidance (e.g., Self-Generated Critiques[49]), controlled generation and safety (e.g., Mitigating Toxicity[25]), and specialized applications ranging from visual storytelling (e.g., StoryLLaVA[22]) to domain-specific scenarios (e.g., Automated Scenario Generation[36]).

Within the reinforcement learning branch, a particularly active line of work applies RL with human feedback to story generation, balancing creativity with alignment to user preferences. StoryAlign[0] sits squarely in this cluster, emphasizing the challenge of training reward models that capture nuanced narrative quality while avoiding common pitfalls such as length biases (cf. Length Correlations[14]) or reward hacking. Nearby efforts like Learning to Reason[2] and RAGferee[3] explore how to incorporate reasoning or retrieval-augmented signals into the reward landscape, whereas BabyStories[32] investigates simpler narrative domains to isolate core alignment questions. A central tension across these works is whether to rely on end-to-end learned rewards, composite hand-crafted signals (e.g., Composite Rewards[28]), or hybrid critique-based approaches. StoryAlign[0] contributes to this conversation by proposing methods that directly address reward model reliability in open-ended creative settings, positioning itself among studies that seek principled ways to scale human oversight for long-form, stylistically diverse narratives.

Related Works in Same Category

The following **3 sibling papers** share the same taxonomy leaf node with the original paper:

1. Learning to reason for long-form story generation

Authors: Gurung, Alexander, Lapata, Mirella, Alexander Gurung, et al. (6 authors total) | **Year/Venue**: 2025 | **URL**: View paper

Abstract

Generating high-quality stories spanning thousands of tokens requires competency across a variety of skills, from tracking plot and character arcs to keeping a consistent and engaging style. Due to the difficulty of sourcing labeled datasets and precise quality measurements, most work using large language models (LLMs) for long-form story generation uses combinations of hand-designed prompting techniques to elicit author-like behavior. This is a manual process that is highly dependent on the spe...

Relationship Analysis

Both papers belong to the RL with Human Feedback for Story Generation category, focusing on using reinforcement learning to improve story generation quality. While StoryAlign focuses on evaluating and training reward models specifically for story preferences using human-verified benchmarks and preference pairs, this candidate paper focuses on learning reasoning capabilities for long-form story generation through RL with verifiable rewards based on completion likelihood improvement. The key difference is that StoryAlign develops reward models to capture human story preferences, whereas the candidate develops reasoning traces to plan and generate better story chapters.

---

2. Recursively summarizing books with human feedback

Authors: Wu, Jeff, Ouyang Long, Jeff Wu, Ziegler, et al. (19 authors total) | **Year/Venue**: 2021 | **URL**: View paper

Abstract

A major challenge for scaling machine learning is training models to perform tasks that are very difficult or time-consuming for humans to evaluate. We present progress on this problem on the task of abstractive summarization of entire fiction novels. Our method combines learning from human feedback with recursive task decomposition: we use models trained on smaller parts of the task to assist humans in giving feedback on the broader task. We collect a large volume of demonstrations and comparis...

Relationship Analysis

Both papers belong to the RL with Human Feedback for Story Generation category, using human preference data to train models for creative text generation. They overlap in employing RLHF techniques to align language models with human preferences for narrative quality, with both collecting human feedback on generated text and training reward models. The key difference is that the original paper (StoryAlign) focuses specifically on evaluating and training reward models for story preference with a dedicated benchmark, while the candidate paper applies recursive task decomposition to book summarization, using RLHF to train models that summarize entire novels through hierarchical subtask completion.

---

3. BabyStories: Can Reinforcement Learning Teach Baby Language Models to Write Better Stories?

Authors: Zhao, Xingmeng, Xingmeng Zhao, Tongnian Wang, Sheri Osborn, et al. (9 authors total) | **Year/Venue**: 2023 | **URL**: View paper

Abstract

Language models have seen significant growth in the size of their corpus, leading to notable performance improvements. Yet, there has been limited progress in developing models that handle smaller, more human-like datasets. As part of the BabyLM shared task, this study explores the impact of reinforcement learning from human feedback (RLHF) on language models pretrained from scratch with a limited training corpus. Comparing two GPT-2 variants, the larger model performs better in storytelling tas...

Relationship Analysis

Both papers belong to the RL with Human Feedback for Story Generation category, using RLHF techniques to improve story generation models. The original paper (StoryAlign) focuses on evaluating and training reward models specifically for story preferences, introducing a benchmark (StoryRMB) and a specialized reward model (StoryReward) trained on 100K preference pairs from human-written stories. The candidate paper (BabyStories) investigates whether small language models pretrained on limited child-like data can benefit from RLHF for storytelling, comparing GPT-2 base and large models trained on the BabyLM dataset with only 500 preference pairs, emphasizing model size effects rather than reward model development.

Contributions Analysis

This paper presents **3 main contributions**, each analyzed against relevant prior work:

Contribution 1: StoryRMB benchmark for story preference evaluation

Description: The authors present StoryRMB, a benchmark containing 1,133 high-quality, human-verified instances for evaluating how well reward models capture human story preferences. Each instance includes a prompt, one chosen story, and three rejected stories across five evaluation dimensions: coherence, creativity, characterization, fluency, and relevance.

This contribution was assessed against **9 related papers** from the literature. Papers with potential prior art are analyzed in detail with textual evidence; others receive brief assessments.

---

1. OpenMEVA: A benchmark for evaluating open-ended story generation metrics

URL: View paper

Brief Assessment

OpenMEVA[52] focuses on evaluating automatic metrics for story generation (e.g., BLEU, perplexity) rather than reward models for story preferences. It contains manually annotated stories to assess metric correlation with human judgments, not preference pairs for training reward models.

---

2. The authenticity gap in human evaluation

URL: View paper

Brief Assessment

Authenticity Gap[60] focuses on theoretical limitations of human evaluation protocols (Likert scales, utility theory) rather than creating story preference benchmarks. It does not present a competing benchmark for story evaluation.

3. LitBench: A Benchmark and Dataset for Reliable Evaluation of Creative Writing

URL: View paper

Prior Art Analysis

LitBench[54] demonstrates that a similar benchmark for evaluating story preferences with human annotations existed prior to the original paper's submission. LitBench[54] presents a standardized benchmark comprising 2,480 human-labeled story comparisons for creative writing evaluation, along with a 43,827-pair training corpus of human preference labels. Both benchmarks serve the same fundamental purpose: providing human-verified instances for evaluating how well models capture human story preferences, though they differ in scale and construction methodology.

Evidence

Evidence 1 - **Rationale**: Both papers claim to introduce 'the first benchmark' for evaluating story preferences with human annotations. LitBench[54] provides a benchmark with 2,480 human-labeled story comparisons, while the original paper presents StoryRMB with 1,133 human-verified instances. This directly challenges the novelty claim of being 'the first benchmark' for this purpose. - **Original**: we introduce story rmb , the first benchmark for assessing reward models on story preferences. story rmb contains 1, 133 high-quality, human-verified instances, each consisting of a prompt, one chosen story, and three rejected stories. - **Candidate**: we introduce litbench, the first standardized benchmark and paired dataset for creative writing verification, comprising a held-out test set of 2,480 debiased, human-labeled story comparisons drawn from reddit and a 43,827-pair training corpus of human preference labels.

Evidence 2 - **Rationale**: The original paper claims StoryRMB is 'the first benchmark for evaluating reward models of story preferences,' but LitBench[54] presents a similar benchmark for creative writing evaluation with human preference labels, which serves the same fundamental purpose of evaluating models on story preferences. - **Original**: to address this gap, we systematically evaluate the modeling of human story preferences and develop an advanced reward model for story generation. specifically, we introduce story rmb , the first benchmark for evaluating reward models of story preferences. - **Candidate**: in pursuit of robust evaluation for creative writing, we introduce litbench, the first standardized benchmark and paired dataset for creative writing verification, comprising a held-out test set of 2,480 debiased, human-labeled story comparisons drawn from reddit and a 43,827-pair training corpus of...

4. Meta-evaluation methodology and benchmark for automatic story generation

URL: View paper

Brief Assessment

Meta-Evaluation Methodology[58] focuses on meta-evaluation methodology for automatic story generation systems, not on benchmarking reward models for story preference. The candidate paper evaluates story generation systems using human criteria (relevance, coherence, empathy, surprise, engagement, complexity) rather than preference pairs with chosen/rejected stories for reward model training.

5. Hybrid preferences: Learning to route instances for human vs. AI feedback

URL: View paper

Brief Assessment

Hybrid Preferences[56] focuses on routing preference instances between human and AI annotators for general language model alignment, not on story-specific preference evaluation benchmarks. The candidate addresses a different problem domain (hybrid annotation routing) rather than story generation evaluation.

6. Are large language models capable of generating human-level narratives?

URL: View paper

Brief Assessment

Human-Level Narratives[57] focuses on narrative discourse analysis (story arcs, turning points, arousal/valence) in movie synopses, not on reward model evaluation or story preference benchmarking with chosen/rejected pairs.

7. OpenGenAlign: A Preference Dataset and Benchmark for Trustworthy Reward Modeling in Open-Ended, Long-Context Generation

URL: View paper

Brief Assessment

OpenGenAlign[53] focuses on open-ended long-context generation across QA, data-to-text, and summarization tasks, not specifically on story preference evaluation with narrative-specific dimensions like creativity and characterization.

8. Ares: An automated evaluation framework for retrieval-augmented generation systems

URL: View paper

Brief Assessment

ARES[51] focuses on evaluating retrieval-augmented generation (RAG) systems for question answering and fact-checking tasks, not story generation or story preference modeling. The benchmark and evaluation dimensions are fundamentally different from StoryRMB.

9. Storybench: A multifaceted benchmark for continuous story visualization

URL: View paper

Brief Assessment

StoryBench[59] focuses on continuous story visualization (text-to-video generation) with human annotations for video generation tasks, not on evaluating reward models for story preference as in StoryRMB.

Contribution 2: Automated method for collecting story preference pairs

Description: The authors develop an automated pipeline for constructing approximately 100,000 story preference pairs from human-written stories using three methods: premise back-generation, prompt-guided rewriting, and human-guided continuation. This dataset captures real-world human preferences from online literary platforms.

This contribution was assessed against **10 related papers** from the literature. Papers with potential prior art are analyzed in detail with textual evidence; others receive brief assessments.

1. StoryER: Automatic story evaluation via ranking, rating and reasoning

URL: View paper

Prior Art Analysis

StoryER[68] demonstrates prior work in automated collection of story preference pairs from human-written stories. The candidate paper describes collecting 100k story pairs by pairing highly-upvoted stories (upvotes ≥50) with lowly-upvoted stories (upvotes ≤0) from the WritingPrompt dataset, using upvote counts as preference signals. This automated pipeline predates the original paper's contribution and captures real-world human preferences from online literary platforms through a similar methodology of leveraging existing metadata (upvote counts) to construct preference pairs without manual annotation.

Evidence

Evidence 1 - **Rationale**: Both papers describe automated collection of story preference pairs from online platforms using upvote counts as preference signals. StoryER[68] collected 116,971 story pairs by pairing highly-upvoted with lowly-upvoted stories from WritingPrompt, demonstrating the same core methodology of leveraging existing metadata to construct preference pairs. - **Original**: we collect stories and associated metadata, such as category and upvote counts, from chinese and english online literary platforms, and construct preference pairs using three methods - **Candidate**: we first collect 193,842 stories prior to 03/2020 from wp along with their prompt, the number of upvotes and uncategorized comments. we remove the stories updated from 12/2019 to 03/2020, since newly-updated stories usually have few upvotes regardless of whether they are good or bad. then, we exclus...

Evidence 2 - **Rationale**: Both papers use upvote counts from online platforms to automatically determine preference labels. StoryER[68]'s method of using upvote counts to label chosen/rejected stories is conceptually identical to the original paper's premise back-generation approach, though StoryER[68] operates at a larger scale (100k pairs). - **Original**: premise back-generation. we randomly sample story pairs from the same category with available upvote information and use an llm to back-translate corresponding premises (qi et al., 2024). the story with more upvotes is denoted as chosen, and the other as rejected, yielding about 6, 000 preference pa... - **Candidate**: as we mentioned above, ranking method is more flexible and better than discrimination when evaluating the story (we also experimentally compare them in sec. f.1). we thus prepare 100k pairwise ranking data for training the model. to this end, we first collect 193,842 stories prior to 03/2020 from wp...

---

2. Eliciting human preferences with language models

URL: View paper

Brief Assessment

Eliciting Human Preferences[69] focuses on interactive elicitation of user preferences through language model-generated questions for personalization tasks, not on automated collection of story preference pairs from human-written stories using back-generation, rewriting, or continuation methods.

---

3. Folktale Story Generation and Automatic Evaluation of Generated Text

URL: View paper

Brief Assessment

Folktale Story Generation[72] focuses on generating folktale stories and evaluating generated text quality, but does not describe an automated pipeline for constructing story preference pairs from human-written stories using methods like premise back-generation, prompt-guided rewriting, or human-guided continuation as described in the original paper.

---

4. LitBench: A Benchmark and Dataset for Reliable Evaluation of Creative Writing

URL: View paper

Prior Art Analysis

LitBench[54] demonstrates prior work in automated collection of story preference pairs from human-written stories. The candidate paper describes collecting 43,827 human preference labels from Reddit, providing a large-scale dataset of story preferences from an online platform. While the original paper's methods (premise back-generation, prompt-guided rewriting, human-guided continuation) differ in specific techniques, LitBench[54] shows that automated collection of story preference pairs from human-written content on online platforms was already established.

Evidence

Evidence 1 - **Rationale**: Both papers describe automated methods for collecting story preference pairs from online platforms with human-written content. LitBench[54] collected 43,827 human preference labels from Reddit, while the original paper collected from Douban and WritingPrompts. This shows that automated collection of story preference pairs from online literary platforms was already established. - **Original**: we collect stories and associated metadata, such as category and upvote counts, from chinese and english online literary platforms, and construct preference pairs using three methods: (1) premise back-generation. we randomly sample story pairs from the same category with available upvote information - **Candidate**: we introduce litbench, the first standardized benchmark and paired dataset for creative writing verification, comprising a held-out test set of 2,480 debiased, human-labeled story comparisons drawn from reddit and a 43,827-pair training corpus of human preference labels.

---

5. Hierarchical neural story generation

URL: View paper

Brief Assessment

Hierarchical Neural Story[70] focuses on hierarchical story generation from prompts using a two-stage generation process (premise then story), not on automated collection of story preference pairs from human-written stories for reward model training.

---

6. Robust preference learning for storytelling via contrastive reinforcement learning

URL: View paper

Brief Assessment

Contrastive RL[41] focuses on using contrastive learning (CARP model) to align stories with human critiques for preference modeling, not on automated collection methods for story preference pairs from human-written stories using techniques like premise back-generation or prompt-guided rewriting.

---

7. Tailored Tales: Enhancing Children's Reading Comprehension with Preference-Tuned Automatic Story Generation

URL: View paper

Brief Assessment

Tailored Tales[73] focuses on children's reading comprehension with age-appropriate stories, not on automated collection of story preference pairs from human-written stories using methods like premise back-generation or prompt-guided rewriting.

---

8. Automatic story generation: A survey of approaches

URL: View paper

Brief Assessment

Automatic Story Survey[33] is a survey paper that reviews existing approaches to story generation. The provided context fragments mention 'resources, corpora, and evaluation methods' and 'reward function' but contain no evidence of automated methods for collecting story preference pairs from human-written stories using techniques like premise back-generation or prompt-guided rewriting.

9. GRAVITY: A Framework for Personalized Text Generation via Profile-Grounded Synthetic Preferences

URL: View paper

Brief Assessment

GRAVITY[71] focuses on personalized text generation using synthetic preference data derived from user profiles (demographics, personality, values), not on collecting story preference pairs from human-written stories. The candidate addresses a different task domain (product description personalization) with different data sources (user profiles and psychological frameworks rather than literary platforms).

10. Aligning Text-to-Music Evaluation with Human Preferences

URL: View paper

Brief Assessment

Text-to-Music Alignment[74] focuses on evaluating text-to-music generation systems through divergence metrics and human preference data collection for music, not story generation or automated story preference pair construction.

Contribution 3: StoryReward advanced reward model

Description: The authors introduce StoryReward, a reward model trained on their large-scale preference dataset that achieves state-of-the-art performance on StoryRMB. The model outperforms much larger models and demonstrates effectiveness in test-time scaling applications such as best-of-n story selection.

This contribution was assessed against **10 related papers** from the literature. Papers with potential prior art are analyzed in detail with textual evidence; others receive brief assessments.

1. User-centric Subjective Leaderboard by Customizable Reward Modeling

URL: View paper

Brief Assessment

User-Centric Leaderboard[63] focuses on customizable reward models for general LLM evaluation across diverse user preferences, not specifically on story generation or narrative quality assessment.

2. StoryER: Automatic story evaluation via ranking, rating and reasoning

URL: View paper

Brief Assessment

StoryER[68] focuses on story evaluation through ranking, rating, and reasoning using LED architecture, not on training a reward model for RLHF or test-time scaling applications like best-of-n selection. The technical approaches and applications differ substantially.

3. Swag: Storytelling with action guidance

URL: View paper

Brief Assessment

SWAG[65] focuses on iterative story generation through action guidance using a feedback loop between two LLMs, not on reward model training for story preference evaluation. The candidate paper does not address reward modeling or preference datasets for story evaluation.

4. LitBench: A Benchmark and Dataset for Reliable Evaluation of Creative Writing

URL: View paper

Prior Art Analysis

LitBench[54] demonstrates prior work in training reward models for story preference evaluation. The candidate paper describes training both Bradley-Terry and generative reward models on their human preference dataset, achieving 78% accuracy. These models were specifically designed for evaluating creative writing and story preferences, demonstrating that advanced reward models for story evaluation existed before the original paper's submission.

Evidence

Evidence 1 - **Rationale**: LitBench[54] describes training reward models (Bradley-Terry and generative reward models) specifically for story preference evaluation, which challenges the novelty of StoryReward as an 'advanced reward model for story preference.' - **Original**: we propose story reward , an advanced reward model for story preference that achieves sota performance on story rmb. we further validate its effectiveness in real-world applications. - **Candidate**: using litbench, we (i) benchmark zero-shot llm judges, (ii) train bradley terry and generative reward models, and (iii) conduct an online human study to validate reward model rankings on newly llm-generated stories.

Evidence 2 - **Rationale**: LitBench[54] demonstrates that reward models trained on story preferences were already achieving strong performance (78% accuracy) and were validated through human studies, showing that advanced reward models for story evaluation existed prior to StoryReward. - **Original**: story reward achieves state-of-the-art performance on story rmb, even surpassing much larger models. we further apply story reward in test-time scaling, conducting best-of-n (bon) experiments where reward models select a better story. - **Candidate**: among trained reward models, bradley-terry and generative reward models both attain an accuracy of 78%, outperforming all off-the-shelf judges. an online human study further confirms that our trained reward models consistently align with human preferences in novel llm-generated stories.

5. ChARM: Character-based Act-adaptive Reward Modeling for Advanced Role-Playing Language Agents

URL: View paper

Brief Assessment

ChARM[61] focuses on role-playing dialogue agents with character-based preferences, not story generation or narrative evaluation. The technical domains are distinct.

6. Robust preference learning for storytelling via contrastive reinforcement learning

URL: View paper

Brief Assessment

Contrastive RL[41] presents CARP as a contrastive bi-encoder for story-critique alignment, not a reward model specifically trained on large-scale story preference datasets for test-time scaling applications like best-of-n selection as described for StoryReward.

7. Larger or Smaller Reward Margins to Select Preferences for Alignment?

URL: View paper

Brief Assessment

Reward Margins[67] focuses on evaluating preference data quality through alignment potential metrics for general LLM alignment, not on developing reward models specifically for story generation tasks or story preference evaluation.

---

8. Reward collapse in aligning large language models: A prompt-aware approach to preference rankings

URL: View paper

Brief Assessment

Reward Collapse[66] focuses on the phenomenon of reward collapse in RLHF training and proposes prompt-aware utility functions to address it. This is fundamentally different from StoryReward, which is a domain-specific reward model trained on story preference data for story generation tasks.

---

9. Beyond Scalar Reward Model: Learning Generative Judge from Preference Data

URL: View paper

Brief Assessment

Generative Judge[64] focuses on training generative reward models using self-generated contrastive judgments for general preference alignment, not specifically for story generation or story preference evaluation as in StoryReward.

---

10. Ai-slop to ai-polish? aligning language models through edit-based writing rewards and test-time computation

URL: View paper

Brief Assessment

AI-Slop to Polish[62] focuses on general writing quality assessment across fiction, non-fiction, and marketing domains, training WQRM models on edit-based preferences. The original paper specifically targets story generation with StoryReward trained on story-specific preference pairs using methods like premise back-generation and human-guided continuation for narrative evaluation.

Appendix: Text Similarity Detection

Textual similarity detection checked 29 papers and found 2 similarity segment(s) across 1 paper(s).

The following **1 paper(s)** were detected to have high textual similarity with the original paper. These may represent different versions of the same work, duplicate submissions, or papers with substantial textual overlap. Readers are advised to verify these relationships independently.

1. Are large language models capable of generating human-level narratives?

Detected in: Contribution: contribution_1

⚠ **Note**: This paper shows substantial textual similarity with the original paper. It may be a different version, a duplicate submission, or contain significant overlapping content. Please review carefully to determine the nature of the relationship.

References

- [0] StoryAlign: Evaluating and Training Reward Models for Story Generation View paper
- [1] Imagereward: Learning and evaluating human preferences for text-to-image generation View paper
- [2] Learning to reason for long-form story generation View paper
- [3] RAGferee: Building contextual reward models for retrieval-augmented generation View paper
- [4] On the Low-Rank Parametrization of Reward Models for Controlled Language Generation View paper
- [5] Pcgrllm: Large language model-driven reward design for procedural content generation reinforcement learning View paper
- [6] Visionreward: Fine-grained multi-dimensional human preference learning for image and video generation View paper
- [7] A critical look at tokenwise reward-guided text generation View paper
- [8] Jointly reinforcing diversity and quality in language model generations View paper
- [9] Args: Alignment as reward-guided search View paper
- [10] Cpo: Addressing reward ambiguity in role-playing dialogue via comparative policy optimization View paper
- [11] Beyond sparse rewards: Enhancing reinforcement learning with language model critique in text generation View paper
- [12] GenARM: Reward Guided Generation with Autoregressive Reward Model for Test-time Alignment View paper
- [13] Cos (m+ o) s: Curiosity and rl-enhanced mcts for exploring story space via language models View paper
- [14] A long way to go: Investigating length correlations in rlhf View paper
- [15] Helpsteer 2: Open-source dataset for training top-performing reward models View paper
- [16] Reward-augmented decoding: Efficient controlled text generation with a unidirectional reward model View paper
- [17] Rlmr: Reinforcement learning with mixed rewards for creative writing View paper
- [18] ReMoDetect: Reward Models Recognize Aligned LLM's Generations View paper
- [19] Hierarchically structured reinforcement learning for topically coherent visual story generation View paper
- [20] No metrics are perfect: Adversarial reward learning for visual storytelling View paper
- [21] Deep reinforcement learning and creativity View paper
- [22] StoryLLaVA: enhancing visual storytelling with multi-modal large language models View paper
- [23] Multimodal LLMs as Customized Reward Models for Text-to-Image Generation View paper
- [24] Recursively summarizing books with human feedback View paper
- [25] Reward modeling for mitigating toxicity in transformer-based language models View paper
- [26] GRAM: A Generative Foundation Reward Model for Reward Generalization View paper
- [27] Curiosity-Driven Reinforcement Learning from Human Feedback View paper
- [28] What makes a good story? designing composite rewards for visual storytelling View paper
- [29] Personalized Reward Modeling for Text-to-Image Generation View paper
- [30] Efficient controlled language generation with low-rank autoregressive reward models View paper
- [31] A survey of deep learning applied to story generation View paper
- [32] BabyStories: Can Reinforcement Learning Teach Baby Language Models to Write Better Stories? View paper
- [33] Automatic story generation: A survey of approaches View paper
- [34] Narrative order aware story generation via bidirectional pretraining model with optimal transport reward View paper

- [35] Goal-directed story generation: Augmenting generative language models with reinforcement learning View paper
- [36] Automated scenario generation to support competency-based experiential learning in GIFT View paper
- [37] EvolvR: Self-Evolving Pairwise Reasoning for Story Evaluation to Enhance Generation View paper
- [38] From Plots to Endings: A Reinforced Pointer Generator for Story Ending Generation View paper
- [39] Transforming human interactions with AI via reinforcement learning with human feedback (RLHF) View paper
- [40] Show, reward and tell: Automatic generation of narrative paragraph from photo stream by adversarial training View paper
- [41] Robust preference learning for storytelling via contrastive reinforcement learning View paper
- [42] Capturing Classic Authorial Style in Long-Form Story Generation with GRPO Fine-Tuning View paper
- [43] Show, reward, and tell: Adversarial visual story generation View paper
- [44] Controllable neural story plot generation via reward shaping View paper
- [45] Unified Reward Model for Multimodal Understanding and Generation View paper
- [46] Teacher Forcing Recovers Reward Functions for Text Generation View paper
- [47] RoleRMBench & RoleRM: Towards Reward Modeling for Profile-Based Role Play in Dialogue Systems View paper
- [48] Formalizing Adaptive Team Feedback in Synthetic Training Environments with Reinforcement Learning View paper
- [49] Self-Generated Critiques Boost Reward Modeling for Language Models View paper
- [50] Probabilistic Uncertain Reward Model View paper
- [51] Ares: An automated evaluation framework for retrieval-augmented generation systems View paper
- [52] OpenMEVA: A benchmark for evaluating open-ended story generation metrics View paper
- [53] OpenGenAlign: A Preference Dataset and Benchmark for Trustworthy Reward Modeling in Open-Ended, Long-Context Generation View paper
- [54] LitBench: A Benchmark and Dataset for Reliable Evaluation of Creative Writing View paper
- [55] Baton: Aligning text-to-audio model with human preference feedback View paper
- [56] Hybrid preferences: Learning to route instances for human vs. AI feedback View paper
- [57] Are large language models capable of generating human-level narratives? View paper
- [58] Meta-evaluation methodology and benchmark for automatic story generation View paper
- [59] Storybench: A multifaceted benchmark for continuous story visualization View paper
- [60] The authenticity gap in human evaluation View paper
- [61] ChARM: Character-based Act-adaptive Reward Modeling for Advanced Role-Playing Language Agents View paper
- [62] Ai-slop to ai-polish? aligning language models through edit-based writing rewards and test-time computation View paper
- [63] User-centric Subjective Leaderboard by Customizable Reward Modeling View paper
- [64] Beyond Scalar Reward Model: Learning Generative Judge from Preference Data View paper
- [65] Swag: Storytelling with action guidance View paper
- [66] Reward collapse in aligning large language models: A prompt-aware approach to preference rankings View paper
- [67] Larger or Smaller Reward Margins to Select Preferences for Alignment? View paper
- [68] StoryER: Automatic story evaluation via ranking, rating and reasoning View paper
- [69] Eliciting human preferences with language models View paper
- [70] Hierarchical neural story generation View paper
- [71] GRAVITY: A Framework for Personalized Text Generation via Profile-Grounded Synthetic Preferences View paper
- [72] Folktale Story Generation and Automatic Evaluation of Generated Text View paper
- [73] Tailored Tales: Enhancing Children's Reading Comprehension with Preference-Tuned Automatic Story Generation View paper
- [74] Aligning Text-to-Music Evaluation with Human Preferences View paper