Novelty Assessment Report

Paper: RECAST: Expanding the Boundaries of LLMs' Complex Instruction Following with Multi-Constraint Data
PDF URL: https://openreview.net/pdf?id=90tCp2KszA
Venue: ICLR 2026 Conference Submission
Year: 2026
Report Generated: 2025-12-29

Abstract

Large language models (LLMs) are increasingly expected to tackle complex tasks, driven by their expanding applications and users' growing proficiency in crafting sophisticated prompts. However, as the number of explicitly stated requirements increases (particularly more than 10 constraints), LLMs often struggle to accurately follow such complex instructions, which limits their applicability in complex real-world scenarios. To the best of our knowledge, existing datasets do not exceed 10 constraints per instance. To address this challenge, we propose RECAST, an efficient and scalable framework for synthesizing datasets where each example incorporates far more constraints than those in existing benchmarks, aiming to challenge and extend the boundaries of models' ability to follow complex instructions. These constraints are extracted from real-world prompt-response pairs to ensure practical relevance. Using this framework, we construct RECAST-30K, a large-scale, high-quality dataset comprising 30k instances spanning 19 constraint types. Experimental results demonstrate that models fine-tuned on RECAST-30K substantially improve in following complex instructions while maintaining their general capabilities without degradation. Moreover, RECAST enables automatic verification of constraint satisfaction via rule-based validators for quantitative constraints and LLM-based validators for qualitative ones, the verifiability provided by RECAST enables the design of reward functions for reinforcement learning, which further boosts model performance on complex and challenging tasks.

Core Task Landscape

This paper addresses: **Complex Instruction Following with Multiple Constraints**
A total of **50 papers** were analyzed and organized into a taxonomy with **22 categories**.

Taxonomy Overview

The research landscape has been organized into the following main categories:
- **Training Data Construction and Synthesis**
- **Training Methods and Optimization**
- **Evaluation Benchmarks and Metrics**
- **Inference-Time Strategies and Self-Correction**
- **Mechanistic Analysis and Model Internals**
- **Generalization and Transfer**
- **Domain-Specific Applications**
- **Multi-Task Learning Architectures**
- **Optimization and Search in Constrained Spaces**
- **Unrelated Technical Papers**

Complete Taxonomy Tree

- Complex Instruction Following with Multiple Constraints Survey Taxonomy
- Training Data Construction and Synthesis
  - Constraint-Based Data Generation ★ (4 papers)
  - [0] RECAST: Expanding the Boundaries of LLMs' Complex Instruction Following with Multi-Constraint Data (Anon et al., 2026) View paper
  - [3] Constraint back-translation improves complex instruction following of large language models (Yunjia Qi, 2025) View paper
  - [6] RECAST: Strengthening LLMs' Complex Instruction Following with Constraint-Verifiable Data (W Liu, 2025) View paper
  - [7] Conifer: Improving complex constrained instruction-following ability of large language models (Sun Hao-Ran, 2024) View paper
  - Iterative and Refinement-Based Generation (2 papers)
  - [18] Complex logical instruction generation (Zhang Mian, 2025) View paper
  - [24] Air: Complex instruction generation via automatic iterative refinement (Liu Wei, 2025) View paper
  - Long-Form and Multi-Constraint Datasets (2 papers)
  - [2] Suri: Multi-constraint instruction following for long-form text generation (Pham, 2024) View paper
  - [28] Suri: Multi-constraint Instruction Following in Long-form Text Generation (Chau Minh Pham, 2024) View paper
- Training Methods and Optimization
  - Contrastive and Preference Learning (3 papers)
  - [1] Musc: Improving complex instruction following with multi-granularity self-contrastive training (Huang Hui, 2025) View paper
  - [11] Step-by-step mastery: Enhancing soft constraint following ability of large language models (Ren Qingyu, 2025) View paper
  - [21] Iopo: Empowering llms with complex instruction following via input-output preference optimization (Zhang, 2025) View paper
  - Multi-Task and Context-Aware Learning (2 papers)
  - [8] Enhancing Complex Instruction Following for Large Language Models with Mixture-of-Contexts Fine-tuning (Lu Yuheng, 2025) View paper

○ [49] Multi-Task Learning Enhanced Physics-Informed Neural Network for Solving Fluid-Structure Interaction Equations (Yihe Zhao, 2024) View paper

Narrative

Core task: complex instruction following with multiple constraints. This field addresses how language models can satisfy diverse, simultaneous requirements—ranging from format and length restrictions to content and stylistic rules—within a single response. The taxonomy organizes research into several major branches: Training Data Construction and Synthesis focuses on generating high-quality constraint-rich examples, often through automated pipelines or back-translation techniques (e.g., Constraint Back-translation[3], Conifer[7]); Training Methods and Optimization explores curriculum strategies and specialized loss functions; Evaluation Benchmarks and Metrics develops testbeds that measure adherence to multi-dimensional constraints (e.g., FollowBench[17], CFBench[23]); Inference-Time Strategies and Self-Correction examines runtime verification and iterative refinement; and additional branches cover mechanistic analysis, generalization, domain-specific applications, and multi-task architectures. Together, these branches reflect a maturing effort to move beyond single-objective instruction tuning toward systems that juggle competing or layered requirements.

Within the data construction landscape, a handful of works emphasize constraint-based generation to produce training corpora that stress-test models on verifiable rules. RECAST[0] sits squarely in this cluster, proposing methods to synthesize instruction–response pairs where constraints are explicit and checkable, much like RECAST Verifiable[6] and Conifer[7]. Compared to Constraint Back-translation[3], which reverses the generation process to ensure coverage, RECAST[0] prioritizes scalable synthesis with built-in verification hooks. Meanwhile, evaluation-focused efforts such as LIFEBench[4] and Multi-Dimensional Constraint[5] highlight the need for rigorous benchmarks that go beyond surface-level metrics. Open questions remain around balancing constraint diversity with data efficiency, and whether models trained on synthetic constraint-heavy data generalize to real-world multi-constraint scenarios. RECAST[0] contributes to this ongoing dialogue by offering a framework that bridges data synthesis and verifiability, positioning itself as a practical tool for scaling constraint-aware instruction tuning.

Related Works in Same Category

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

1. Constraint back-translation improves complex instruction following of large language models

Authors: Yunjia Qi, Hao Peng, Xiaozhi Wang, Bin Xu, Lei Hou, et al. (6 authors total) | **Year/Venue**: 2025 | **URL**: View paper

Abstract

Large language models (LLMs) struggle to follow instructions with complex constraints in format, length, etc. Following the conventional instruction-tuning practice, previous works conduct post-training on complex instruction-response pairs generated by feeding complex instructions to advanced LLMs. However, even advanced LLMs cannot follow complex instructions well, thus limiting the quality of generated data. In this work, we find that existing datasets inherently contain implicit complex cons...

Relationship Analysis

Both papers belong to the Constraint-Based Data Generation category, focusing on extracting or synthesizing constraints from existing data to create instruction-response pairs for improving LLMs' complex instruction-following abilities. While RECAST extracts constraints from responses in seed data and then augments instructions by adding these constraints, the candidate paper (Crab) uses constraint back-translation to identify implicit constraints already satisfied by existing high-quality responses and adds them to instructions. The key difference is that RECAST builds a comprehensive constraint pool through rule-based and model-based extraction to create datasets with unprecedented constraint complexity (30+ constraints), whereas Crab focuses on back-translating implicit constraints from existing pairs to reduce data noise and generation costs.

2. RECAST: Strengthening LLMs' Complex Instruction Following with Constraint-Verifiable Data

Authors: W Liu, Z Guo, M Xie, J Xu, Z Huang, et al. (6 authors total) | **Year/Venue**: 2025 | **URL**: View paper

Abstract

â␣ We present RECAST, a comprehensive framework for improving large language modelsâ␣␣ ability to follow complex, multi-constraint instructions. As shown in Figure 2, the RECAST â␣

⚠ Similarity Notice

These papers appear to be the same work or very closely related variants. Both share the identical title 'RECAST: Expanding the Boundaries of LLMs' Complex Instruction Following with Multi-Constraint Data' (with minor subtitle variations), describe the same RECAST framework for constraint extraction and synthesis, present the same RECAST-30K dataset with 19 constraint types, and detail the same RLVC reinforcement learning approach. The core technical contributions, methodology, and experimental setup are essentially identical across both papers.

3. Conifer: Improving complex constrained instruction-following ability of large language models

Authors: Sun Hao-Ran, Liu Li-xin, Li Junjie, Lixin Liu, Wang Feng-Yu, et al. (12 authors total) | **Year/Venue**: 2024 | **URL**: View paper

Abstract

The ability of large language models (LLMs) to follow instructions is crucial to real-world applications. Despite recent advances, several studies have highlighted that LLMs struggle when faced with challenging instructions, especially those that include complex constraints, hindering their effectiveness in various tasks. To address this challenge, we introduce Conifer, a novel instruction tuning dataset, designed to enhance LLMs to follow multi-level instructions with complex constraints. Utili...

Relationship Analysis

Both papers belong to the Constraint-Based Data Generation category, focusing on extracting or synthesizing constraints from existing data to create instruction-response pairs for training LLMs. They overlap in their approach of mining constraints from seed datasets (RECAST uses Tülu 3 Persona IF, Conifer uses ShareGPT) and incorporating multiple constraints into instructions to improve complex instruction-following capabilities. However, RECAST emphasizes scaling to unprecedented constraint complexity (far exceeding 10 constraints per instance) with dual rule-based and model-based verification mechanisms, while Conifer focuses on a progressive learning scheme with easy-to-hard progression and process feedback, organizing constraints into multi-turn conversations with difficulty levels 1-5.

Contributions Analysis

Overall novelty summary. The paper introduces RECAST, a framework for synthesizing instruction-following datasets with far more constraints per instance than existing benchmarks (exceeding 10 constraints), targeting practical scenarios with 19 constraint types). It resides in the 'Constraint-Based Data Generation' leaf under 'Training Data Construction and Synthesis', alongside three sibling papers: Constraint Back-translation, RECAST Verifiable, and Conifer. This leaf represents a focused but not overcrowded research direction within the broader taxonomy of 50 papers across 22 leaf nodes, indicating moderate activity in constraint-based data synthesis methods.

The taxonomy reveals neighboring leaves such as 'Iterative and Refinement-Based Generation' (automated complexity expansion) and 'Long-Form and Multi-Constraint Datasets' (specialized corpora for extended text). RECAST's approach diverges from iterative refinement methods by emphasizing static constraint extraction from real-world prompt-response pairs, aligning more closely with verification-oriented synthesis. The broader 'Training Data Construction and Synthesis' branch contains six papers total, suggesting this is an active but not saturated area. Related branches like 'Evaluation Benchmarks and Metrics' (14 papers) and 'Training Methods and Optimization' (7 papers) indicate the field prioritizes assessment and optimization alongside data generation.

Among 30 candidates examined, the RECAST framework (Contribution 1) shows one refutable candidate out of 10 examined, suggesting some overlap with prior constraint-based synthesis work. The RECAST-30K dataset (Contribution 2) and RLVC training method (Contribution 3) each examined 10 candidates with zero refutations, indicating these contributions appear more distinct within the limited search scope. The framework's emphasis on scaling beyond 10 constraints per instance and extracting constraints from real-world data may differentiate it from existing synthesis pipelines, though the single refutable match warrants attention to prior constraint extraction techniques.

Based on the top-30 semantic matches examined, the work appears to occupy a moderately explored niche within constraint-based data generation. The taxonomy structure suggests the field is actively developing evaluation benchmarks and training methods, but data synthesis approaches remain less saturated. The analysis does not cover exhaustive literature review or domain-specific constraint generation methods outside the examined candidates, leaving open questions about overlap with specialized constraint frameworks or industry-scale synthesis pipelines not captured in this search scope.

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

Contribution 1: RECAST framework for synthesizing multi-constraint instruction-following datasets

Description: The authors introduce RECAST, a data-synthesis framework that systematically mines and verifies both rule-based and model-based constraints from existing prompt-response pairs. This framework enables the construction of instruction-following datasets with unprecedented constraint complexity, addressing limitations in existing benchmarks that typically contain fewer than 10 constraints per instance.

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. UltraIF: Advancing Instruction Following from the Wild
URL: View paper

Prior Art Analysis

UltraIF[54] demonstrates that prior work exists for synthesizing instruction-following datasets with multiple verifiable constraints. Both papers decompose real-world instructions into constraints and use verification mechanisms to ensure response quality. UltraIF[54] decomposes user instructions into simplified queries and constraints with evaluation questions, then trains a composer model to synthesize complex instructions. Similarly, the original paper mines constraints from existing data and uses both rule-based and model-based validators. The key overlap is that both frameworks systematically extract constraints from real-world data and incorporate verification mechanisms, challenging the novelty claim that RECAST is the first to enable construction of datasets with unprecedented constraint complexity through systematic mining and verification.

Evidence

Evidence 1 - **Rationale**: Both papers claim to propose scalable frameworks for synthesizing complex instruction-following datasets by leveraging real-world instructions, demonstrating prior work exists in this area. - **Original**: we propose recast, an efficient and scalable framework for synthesizing datasets where each example incorporates far more constraints than those in existing benchmarks, aiming to challenge and extend the boundaries of models' ability to follow complex instructions. these constraints are extracted fr... - **Candidate**: we propose ultraif, a simple and scalable approach that leverages real-world user instructions to train a composer model,ultracomposer, enabling the synthesis of complex and diverse instructions with correct responses.

Evidence 2 - **Rationale**: Both frameworks decompose instructions into components with verification mechanisms. UltraIF[54] uses evaluation questions while the original uses rule-based and model-based validators, showing similar approaches to constraint verification existed prior. - **Original**: to effectively model the multifaceted nature of real-world instructions, we develop a comprehensive constraint pool that captures both objective and subjective requirements. we categorize constraints into two complementary types based on their verification mechanisms, enabling precise assessment of ... - **Candidate**: ultraif focuses on generating diverse, complex instructions with correct responses. to achieve this, we propose theultracomposer, a specialized model to synthesize diverse instructions and generate corresponding evaluation questions. building this composer model involves three key steps: instruction...

Evidence 3 - **Rationale**: Both papers systematically extract and categorize constraints from instructions, with verification mechanisms to ensure constraint satisfaction, demonstrating that the approach of mining verifiable constraints from real-world data was not novel to RECAST. - **Original**: rule-based constraints represent objective requirements that can be verified through deterministic methods, including structural elements (paragraph count), lexical specifications (keyword inclusion/exclusion), and quantitative parameters (word limits). these constraints are particularly valuable as... - **Candidate**: instruction decompositionthe decomposition process leverages llms to decompose complex instructions into different components. these components consist of a set of simplified instructions paired with constraints that represent the underlying requirements of the original instruction.

Evidence 4 - **Rationale**: Both frameworks address subjective constraints requiring semantic judgment. UltraIF[54] uses LLM-generated evaluation questions while the original uses model-based validators, showing similar dual-verification approaches existed before RECAST. - **Original**: model-based constraints encompass subjective requirements that necessitate semantic understanding or qualitative judgment, such as stylistic elements (formality level), tonal qualities (politeness), and content characteristics (persuasiveness). these constraints are crucial for capturing the nuanced... - **Candidate**: evaluation question generationwhile xu et al. (2023); wang et al. (2023) focus on improving the complexity of instructions, omitting the quality of generated responses often leads to lowquality samples. inspired by qin et al. (2024), we utilize llm to generate evaluation questions for each constrain...

---

2. Mm-ifengine: Towards multimodal instruction following
URL: View paper

Brief Assessment

Mm-ifengine[57] focuses on multimodal instruction following with vision inputs, generating image-instruction pairs through a three-stage pipeline (image filtering, task generation, constraint integration). In contrast, RECAST operates on text-only prompt-response pairs and systematically mines constraints from existing responses. The two frameworks address different modalities and employ distinct data synthesis approaches.

---

3. Evaluating the Instruction-following Abilities of Language Models using Knowledge Tasks
URL: View paper

Brief Assessment

Knowledge Tasks Evaluation[58] focuses on instruction-following evaluation using knowledge QA tasks with verifiable constraints, not on synthesizing multi-constraint datasets from prompt-response pairs. The candidate evaluates existing models rather than proposing a data synthesis framework.

4. DecIF: Improving Instruction-Following through Meta-Decomposition

URL: View paper

Brief Assessment

DecIF[53] focuses on generating instruction-following data through meta-decomposition (domains, requests, scenarios) without relying on external documents, whereas RECAST systematically mines constraints from existing prompt-response pairs. The two approaches differ fundamentally in their data sources and constraint extraction methods.

5. Towards Better Instruction Following Retrieval Models

URL: View paper

Brief Assessment

Instruction Following Retrieval[56] focuses on information retrieval tasks with instruction-query-passage triplets for retrieval models, not general instruction-following datasets with multi-constraint synthesis for LLMs.

6. SynthRL: Scaling Visual Reasoning with Verifiable Data Synthesis

URL: View paper

Brief Assessment

SynthRL[55] focuses on synthesizing visual reasoning data for vision-language models using verifiable rewards in RL, not on multi-constraint instruction-following datasets. The technical domains and objectives differ fundamentally.

7. Codeif-bench: Evaluating instruction-following capabilities of large language models in interactive code generation

URL: View paper

Brief Assessment

CodeIF-bench[52] focuses on interactive code generation with verifiable instructions in multi-turn programming scenarios, not on general instruction-following dataset synthesis with multiple constraints across diverse domains.

8. Suri: Multi-constraint instruction following for long-form text generation

URL: View paper

Brief Assessment

Suri[2] focuses on long-form text generation (2-5k words) with backtranslated instructions from existing human-written texts, whereas RECAST systematically mines constraints from prompt-response pairs to create datasets with unprecedented constraint complexity (>10 constraints). The approaches and domains differ substantially.

9. RECAST: Strengthening LLMs' Complex Instruction Following with Constraint-Verifiable Data

URL: View paper

Brief Assessment

RECAST Verifiable[6] presents the same RECAST framework with identical methodology. This is the same work, not prior work that refutes novelty.

10. Infobench: Evaluating instruction following ability in large language models

URL: View paper

Brief Assessment

InfoBench[51] focuses on evaluating instruction-following through decomposed requirements, not on synthesizing training datasets with verifiable constraints like RECAST.

Contribution 2: RECAST-30K dataset with 30k instances spanning 19 constraint types

Description: The authors release RECAST-30K, a large-scale dataset constructed using the RECAST framework. This dataset contains 30,000 training instances with diverse, verifiable constraints across 19 types, designed to benchmark and improve complex instruction-following performance in language models.

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. Conifer: Improving complex constrained instruction-following ability of large language models

URL: View paper

Brief Assessment

Conifer[7] focuses on multi-level constraint instructions with a different construction methodology (query reframing, constraint generation, recombination) and uses 13k samples, not 30k. The constraint taxonomy and data generation approach differ substantially from RECAST-30K's extraction-based framework.

2. EIFBENCH: Extremely Complex Instruction Following Benchmark for Large Language Models

URL: View paper

Brief Assessment

EIFBENCH[10] focuses on multi-instruction multi-constraint scenarios with 1,000 instances across different task types, while RECAST-30K addresses single-instruction scenarios with 30,000 instances spanning 19 constraint types. The datasets serve different evaluation purposes and complexity dimensions.

3. Cif-bench: A chinese instruction-following benchmark for evaluating the generalizability of large language models

URL: View paper

Brief Assessment

CIF-bench[69] focuses on evaluating Chinese instruction-following capabilities across 150 tasks with 45,000 instances, not on constraint-type diversity for complex instruction-following. The datasets serve different purposes: RECAST-30K targets multi-constraint satisfaction training, while CIF-bench evaluates zero-shot generalizability in Chinese language tasks.

4. Constraint back-translation improves complex instruction following of large language models

URL: View paper

Brief Assessment

Constraint Back-translation[3] focuses on back-translating constraints from existing responses rather than extracting and synthesizing constraints from scratch. The CRAB dataset uses a different methodology (constraint back-translation) compared to RECAST's forward constraint extraction and augmentation pipeline.

5. Benchmarking Complex Instruction-Following with Multiple Constraints Composition

URL: View paper

Brief Assessment

Multiple Constraints Composition[16] focuses on evaluating complex instruction-following through composition types (and, chain, selection, nested) rather than constructing large-scale training datasets. Their benchmark contains 1,150 instances for evaluation, not 30k training instances like RECAST-30K.

6. Followbench: A multi-level fine-grained constraints following benchmark for large language models

URL: View paper

Brief Assessment

FollowBench[17] focuses on evaluating instruction-following through five constraint categories (content, situation, style, format, example) with a multi-level mechanism, rather than providing a large-scale training dataset like RECAST-30K. The two works serve complementary purposes: evaluation benchmark versus training resource.

7. Unnatural instructions: Tuning language models with (almost) no human labor

URL: View paper

Brief Assessment

Unnatural Instructions[68] focuses on generating diverse instruction-following tasks through LLM prompting, not on constructing datasets with explicit, verifiable constraints across multiple constraint types as RECAST does.

8. WizardLM: Empowering large pre-trained language models to follow complex instructions

URL: View paper

Brief Assessment

WizardLM[70] focuses on generating complex instructions through evolutionary methods (Evol-Instruct) rather than constraint-based instruction following. While both create large-scale instruction datasets, WizardLM[70] does not address verifiable constraint types or multi-constraint satisfaction as RECAST does.

9. RECAST: Strengthening LLMs' Complex Instruction Following with Constraint-Verifiable Data

URL: View paper

Brief Assessment

RECAST Verifiable[6] describes the identical RECAST-30K dataset. This is the same contribution, not a refutation by prior work.

10. Infobench: Evaluating instruction following ability in large language models

URL: View paper

Brief Assessment

InfoBench[51] provides a 500-instruction evaluation benchmark with decomposed questions, not a large-scale training dataset like RECAST-30K.

Contribution 3: RLVC reinforcement learning method using constraint-specific rewards

Description: The authors propose RLVC (Reinforcement Learning with Verifiable Constraints), which exploits the verifiable nature of constraints in RECAST-30K to provide fine-grained, per-constraint reward signals during policy optimization. This method treats each constraint as a separate optimization target, enabling more effective learning for complex multi-constraint scenarios.

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. A hierarchical compliance-based contextual policy search for robotic manipulation tasks with multiple objectives

URL: View paper

Brief Assessment

Hierarchical Compliance Policy[64] addresses robotic manipulation with hierarchical policy search for impedance control and contact-rich tasks, not multi-constraint instruction following with verifiable rewards for language models.

2. Light-IF: Endowing LLMs with Generalizable Reasoning via Preview and Self-Checking for Complex Instruction Following

URL: View paper

Brief Assessment

Light-IF[67] focuses on eliciting generalizable reasoning patterns (preview and self-checking) for instruction following through entropy-preserving training and token-wise entropy-adaptive RL. It does not employ constraint-specific reward signals for multi-objective optimization as RLVC does.

3. The perfect blend: Redefining rlhf with mixture of judges

URL: View paper

Brief Assessment

Mixture of Judges[61] focuses on multi-task RLHF with mixture of judges for constraint evaluation, not specifically on per-constraint reward signals for multi-constraint instruction following as in RLVC.

4. Omni-thinker: Scaling cross-domain generalization in llms via multi-task rl with hybrid rewards

URL: View paper

Brief Assessment

Omni-thinker[60] focuses on multi-task RL across diverse domains (coding, math, QA, creative writing) using hybrid rewards combining rule-based verifiers and LLM-as-a-judge evaluations. The original paper's RLVC specifically targets multi-constraint instruction following with per-constraint reward channels for simultaneous optimization of multiple constraints within single instructions, which is a different technical focus than Omni-thinker's cross-domain generalization approach.

5. GR (1)-guided deep reinforcement learning for multi-task motion planning under a stochastic environment
URL: View paper

Brief Assessment

GR Motion Planning[62] focuses on motion planning for robotics using temporal logic (GR(1)) to guide movement strategies in stochastic environments, not on multi-constraint instruction following with per-constraint reward signals for language models.

6. UCP: a unified framework for code generation with pseudocode-based multi-task learning and reinforcement alignment: Y. Wen et al.
URL: View paper

Brief Assessment

UCP Pseudocode[63] applies reinforcement learning for pseudocode-based code generation tasks with multi-task optimization, not for multi-constraint instruction following with per-constraint reward signals as in RLVC.

7. Multi-Level Aware Preference Learning: Enhancing RLHF for Complex Multi-Instruction Tasks
URL: View paper

Brief Assessment

Multi-Level Aware Preference[66] focuses on preference learning at intra-sample and inter-sample levels for multi-instruction tasks, not on constraint-specific reward signals for individual constraints as in RLVC.

8. Efficient multi-task reinforcement learning via task-specific action correction
URL: View paper

Brief Assessment

Task-specific Action Correction[59] addresses multi-task robotic manipulation with task-specific action corrections and goal-oriented sparse rewards, not multi-constraint instruction following with per-constraint verification signals.

9. RECAST: Strengthening LLMs' Complex Instruction Following with Constraint-Verifiable Data
URL: View paper

Brief Assessment

RECAST Verifiable[6] presents the same RLVC method with identical constraint-specific reward design. This is the same work, not prior work.

10. Multi Task Inverse Reinforcement Learning for Common Sense Reward
URL: View paper

Brief Assessment

Common Sense Reward[65] focuses on learning task-agnostic common-sense rewards through multi-task inverse reinforcement learning, not on constraint-specific rewards for multi-constraint instruction following as in RLVC.

Appendix: Text Similarity Detection

Textual similarity detection checked 27 papers and found 4 similarity segment(s) across 2 paper(s).

The following **2 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. RECAST: Strengthening LLMs' Complex Instruction Following with Constraint-Verifiable Data
Detected in: Core Task (sibling), Contribution: contribution_1, Contribution: contribution_2, Contribution: contribution_3

⚠ **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.

2. Followbench: A multi-level fine-grained constraints following benchmark for large language models
Detected in: Contribution: contribution_2

⚠ **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] RECAST: Expanding the Boundaries of LLMs' Complex Instruction Following with Multi-Constraint Data View paper
- [1] Musc: Improving complex instruction following with multi-granularity self-contrastive training View paper
- [2] Suri: Multi-constraint instruction following for long-form text generation View paper
- [3] Constraint back-translation improves complex instruction following of large language models View paper
- [4] LIFEBench: Evaluating Length Instruction Following in Large Language Models View paper
- [5] A Multi-Dimensional Constraint Framework for Evaluating and Improving Instruction Following in Large Language Models View paper
- [6] RECAST: Strengthening LLMs' Complex Instruction Following with Constraint-Verifiable Data View paper
- [7] Conifer: Improving complex constrained instruction-following ability of large language models View paper
- [8] Enhancing Complex Instruction Following for Large Language Models with Mixture-of-Contexts Fine-tuning View paper
- [9] Attention satisfies: A constraint-satisfaction lens on factual errors of language models View paper
- [10] EIFBENCH: Extremely Complex Instruction Following Benchmark for Large Language Models View paper
- [11] Step-by-step mastery: Enhancing soft constraint following ability of large language models View paper
- [12] Spar: Self-play with tree-search refinement to improve instruction-following in large language models View paper
- [13] Understanding information storage and transfer in multi-modal large language models View paper
- [14] A Hierarchical and Evolvable Benchmark for Fine-Grained Code Instruction Following with Multi-Turn Feedback View paper

- [15] Mpcc: A novel benchmark for multimodal planning with complex constraints in multimodal large language models View paper
- [16] Benchmarking Complex Instruction-Following with Multiple Constraints Composition View paper
- [17] Followbench: A multi-level fine-grained constraints following benchmark for large language models View paper
- [18] Complex logical instruction generation View paper
- [19] Incentivizing Reasoning for Advanced Instruction-Following of Large Language Models View paper
- [20] From instructions to constraints: Language model alignment with automatic constraint verification View paper
- [21] Iopo: Empowering llms with complex instruction following via input-output preference optimization View paper
- [22] Marco-Bench-MIF: On Multilingual Instruction-Following Capability of Large Language View paper
- [23] Cfbench: A comprehensive constraints-following benchmark for llms View paper
- [24] Air: Complex instruction generation via automatic iterative refinement View paper
- [25] Constraint satisfaction in large language models View paper
- [26] When thinking fails: The pitfalls of reasoning for instruction-following in llms View paper
- [27] Concurrent Prediction of Timing and wire Length Using A Multi-Task Graph Neural Network View paper
- [28] Suri: Multi-constraint Instruction Following in Long-form Text Generation View paper
- [29] Advancing mental health detection in texts via multi-task learning with soft-parameter sharing transformers View paper
- [30] Learning to branch for multi-task learning View paper
- [31] Generalizing Verifiable Instruction Following View paper
- [32] Modeling Task Relationships in Multi-task Learning with Multi-gate Mixture-of-Experts View paper
- [33] Generalizing from short to long: Effective data synthesis for long-context instruction tuning View paper
- [34] Orthographic Constraint Satisfaction and Human Difficulty Alignment in Large Language Models View paper
- [35] Large language models can solve real-world planning rigorously with formal verification tools View paper
- [36] Real-time equilibrium reconstruction by multi-task learning neural network based on HL-3 tokamak View paper
- [37] Coordinate structure constraint in the linguistic competence of large language models View paper
- [38] Task-Interaction-Free multi-task learning with efficient hierarchical feature representation View paper
- [39] Cs4: Measuring the creativity of large language models automatically by controlling the number of story-writing constraints View paper
- [40] Multi-task learning for dense prediction tasks: A survey View paper
- [41] Constraint Decomposition for Multi-Objective Instruction-Following in Large Language Models View paper
- [42] Multi-dimensional Constraint-based Test Case Generation and Evaluation Framework for Large Language Models View paper
- [43] Divide-Verify-Refine: Can LLMs Self-Align with Complex Instructions? View paper
- [44] Adashare: Learning what to share for efficient deep multi-task learning View paper
- [45] Enhancing Demand Prediction: A Multi-Task Learning Approach for Taxis and TNCs View paper
- [46] WildIFEval: Instruction Following in the Wild View paper
- [47] Large language model-aided evolutionary search for constrained multiobjective optimization View paper
- [48] TOD-ProcBench: Benchmarking Complex Instruction-Following in Task-Oriented Dialogues View paper
- [49] Multi-Task Learning Enhanced Physics-Informed Neural Network for Solving Fluid-Structure Interaction Equations View paper
- [50] 3DBench: A scalable benchmark for object and scene-level instruction-tuning of 3D large language models. View paper
- [51] Infobench: Evaluating instruction following ability in large language models View paper
- [52] Codeif-bench: Evaluating instruction-following capabilities of large language models in interactive code generation View paper
- [53] DecIF: Improving Instruction-Following through Meta-Decomposition View paper
- [54] UltraIF: Advancing Instruction Following from the Wild View paper
- [55] SynthRL: Scaling Visual Reasoning with Verifiable Data Synthesis View paper
- [56] Towards Better Instruction Following Retrieval Models View paper
- [57] Mm-ifengine: Towards multimodal instruction following View paper
- [58] Evaluating the Instruction-following Abilities of Language Models using Knowledge Tasks View paper
- [59] Efficient multi-task reinforcement learning via task-specific action correction View paper
- [60] Omni-thinker: Scaling cross-domain generalization in llms via multi-task rl with hybrid rewards View paper
- [61] The perfect blend: Redefining rlhf with mixture of judges View paper
- [62] GR (1)-guided deep reinforcement learning for multi-task motion planning under a stochastic environment View paper
- [63] UCP: a unified framework for code generation with pseudocode-based multi-task learning and reinforcement alignment: Y. Wen et al. View paper
- [64] A hierarchical compliance-based contextual policy search for robotic manipulation tasks with multiple objectives View paper
- [65] Multi Task Inverse Reinforcement Learning for Common Sense Reward View paper
- [66] Multi-Level Aware Preference Learning: Enhancing RLHF for Complex Multi-Instruction Tasks View paper
- [67] Light-IF: Endowing LLMs with Generalizable Reasoning via Preview and Self-Checking for Complex Instruction Following View paper
- [68] Unnatural instructions: Tuning language models with (almost) no human labor View paper
- [69] Cif-bench: A chinese instruction-following benchmark for evaluating the generalizability of large language models View paper
- [70] WizardLM: Empowering large pre-trained language models to follow complex instructions View paper