Novelty Assessment Report

Paper: Attend to the Active: Structure-Aware Dynamic Attention in LLMs for Compositional Instruction Following
PDF URL: https://openreview.net/pdf?id=wwA0X3UfAn
Venue: ICLR 2026 Conference Submission
Year: 2026
Report Generated: 2026-01-05

Abstract

Large language models (LLMs) have exhibited strong instruction-following capabilities; however, they often struggle with compositional instructions involving multiple interleaved yet logically independent sub-tasks. These sub-tasks are typically organized in mutually exclusive structures, such as branching, chaining, or paralleling, where only one sub-task should be active at each generation step, while the others remain dormant. Despite their inactivity, dormant sub-tasks can inadvertently attract the model's attention due to structural entanglement within the input context or intermediate representations, leading to interference that compromises output fidelity. To address this challenge, we propose ATA, a structure-aware dynamic attention mechanism grounded in compositional structures, which dynamically identifies the active sub-task during generation while suppressing attention to inactive ones. By precisely steering the model's focus, ATA mitigates interference and explicitly enhances model adherence to the active sub-task. Importantly, ATA operates within a single forward pass without requiring parameter updates. Extensive experiments show that ATA consistently enhances LLMs' instruction-following ability across various compositional structures, effectively mitigating attention distraction and demonstrating a strong generalization ability.

> **Disclaimer**
>
> This report is **AI-GENERATED** using Large Language Models and WisPaper (a scholar search engine). It analyzes academic papers' tasks and contributions against retrieved prior work. While this system identifies **POTENTIAL** overlaps and novel directions, **ITS COVERAGE IS NOT EXHAUSTIVE AND JUDGMENTS ARE APPROXIMATE**. These results are intended to assist human reviewers and **SHOULD NOT** be relied upon as a definitive verdict on novelty.
>
> Note that some papers exist in multiple, slightly different versions (e.g., with different titles or URLs). The system may retrieve several versions of the same underlying work. The current automated pipeline does not reliably align or distinguish these cases, so human reviewers will need to disambiguate them manually.
>
> If you have any questions, please contact: mingzhang23@m.fudan.edu.cn

Core Task Landscape

This paper addresses: **core task of the paper (unspecified)**
A total of **50 papers** were analyzed and organized into a taxonomy with **30 categories**.

Taxonomy Overview

The research landscape has been organized into the following main categories:
- **Attention Mechanisms and Model Architecture**
- **Multi-Task and Auxiliary Learning Frameworks**
- **Training Objectives and Optimization**
- **Task Decomposition and Planning**
- **Benchmark Datasets and Evaluation**
- **Task-Specific Applications**
- **Modular System Design and Software Engineering**
- **Research Methodology and Problem Formulation**
- **Organizational and Policy Studies**
- **Theoretical and Foundational Studies**
- ... and 2 more categories

Complete Taxonomy Tree

- core task of the paper (unspecified) Survey Taxonomy
- Attention Mechanisms and Model Architecture
  - Dynamic and Structure-Aware Attention ★ (1 papers)
  - [0] Attend to the Active: Structure-Aware Dynamic Attention in LLMs for Compositional Instruction Following (Anon et al., 2026) View paper
  - Hyperbolic and Non-Euclidean Representations (1 papers)
  - [47] Hyperbolic Deep Learning in Computer Vision: A Survey (Pascal Mettes, 2024) View paper
- Multi-Task and Auxiliary Learning Frameworks
  - Auxiliary Task Design and Knowledge Transfer (2 papers)
  - [29] Dynamic Auxiliary Task-Based Evolutionary Multitasking for Constrained Multiobjective Optimization (Kangjia Qiao, 2023) View paper
  - [34] Prior knowledge-informed multi-task dynamic learning for few-shot machinery fault diagnosis (Tianci Zhang, 2025) View paper
  - Multi-Objective Optimization Strategies (2 papers)
  - [13] Spatial multi-objective optimization of primary healthcare facilities: A case study in Singapore (Zhong Wang, 2024) View paper
  - [14] PANDA study: subtyping of Parkinson's disease – cohort study protocol (Rui Yang, 2025) View paper
  - Unified Multi-Task Architectures (1 papers)
  - [49] A Simple Language Model for Task-Oriented Dialogue (Ehsan Hosseini-Asl, 2020) View paper
- Training Objectives and Optimization
  - Preference Modeling and Alignment (1 papers)
  - [8] Interpretable Preferences via Multi-Objective Reward Modeling and Mixture-of-Experts (Haoxiang Wang, 2024) View paper
  - Complementary and Auxiliary Objectives (1 papers)
  - [31] Complement objective training (Chen, 2019) View paper

- ◦ [36] Research problems (I. Daubechies, 2025) View paper
- Medical and Biological Research Studies
  - ◦ Oncology and Cancer Research (1 papers)
  - ◦ [5] Implications of Adnexal Invasions in Primary Extramammary Paget's Disease: A Systematic Review (Sabita Aryal, 2024) View paper
  - ◦ Biological Research Methodology and Bias (1 papers)
  - ◦ [27] Are hormones a "female problem" for animal research? (Rebecca M. Shansky, 2019) View paper
- Educational and Training Studies (1 papers)
  - ◦ [24] [Competency-based learning objectives of primary qualifying nursing studies and intended task profiles]. (Ingrid Darmann-Finck, 2024) View paper

Narrative

Core task: The paper addresses an unspecified core task, yet the taxonomy reveals a rich landscape spanning diverse methodological and application domains. At the highest level, the field divides into twelve major branches that range from technical machine learning concerns—such as Attention Mechanisms and Model Architecture, Multi-Task and Auxiliary Learning Frameworks, and Training Objectives and Optimization—to more applied and interdisciplinary areas including Task-Specific Applications, Benchmark Datasets and Evaluation, and even Organizational and Policy Studies. Within the technical branches, researchers explore how models can be designed with flexible attention schemes, how multiple objectives can be balanced during training (as seen in works like Multi-Objective Reward Modeling[8]), and how tasks can be decomposed or planned (Interactive Planning LLMs[12]). Meanwhile, branches devoted to benchmarks (MATH Dataset[6]) and task-specific applications (Precision Agriculture Datasets[43]) provide the empirical testbeds and real-world contexts that ground these methodological advances. The taxonomy also includes branches on Research Methodology and Problem Formulation, as well as studies in medical, biological, educational, and policy domains, reflecting the broad interdisciplinary reach of the field.

Within this landscape, a particularly active line of work focuses on dynamic and structure-aware attention mechanisms, where models adapt their focus based on input structure or task demands. Active Structure Attention[0] sits squarely in this branch, emphasizing how attention can be conditioned on structural properties rather than relying solely on static or content-based weighting. This contrasts with more general architectural innovations and with multi-task frameworks that prioritize shared representations across objectives. Nearby efforts in modular system design (Object Oriented Modularization[28]) and task decomposition (Interactive Planning LLMs[12]) share a concern for flexible, interpretable computation, yet they typically address modularity at the level of entire subsystems rather than within the attention mechanism itself. The open questions in this area revolve around how to balance expressiveness and efficiency when attention must respect complex structural constraints, and how such mechanisms generalize across diverse tasks and data modalities.

Related Works in Same Category

No sibling papers were found in the same taxonomy leaf. A taxonomy-subtopic-level comparison will be produced instead.

Taxonomy-Level Summary

The original leaf focuses on adaptive attention mechanisms that dynamically adjust based on task structure or compositional requirements during inference, emphasizing runtime modulation. The sibling subtopic explores non-Euclidean geometric spaces (particularly hyperbolic geometry) for representing hierarchical or complex visual structures. Both address structural awareness in representation learning, but through fundamentally different mechanisms: dynamic attention modulation versus geometric embedding spaces.

Similarities: - Both categories address structure-aware representation learning in computer vision - Both aim to capture compositional or hierarchical relationships in visual data - Both go beyond standard Euclidean or static approaches to better model complex visual structures

Differences: - Dynamic and Structure-Aware Attention operates through adaptive modulation mechanisms at inference time, while Hyperbolic representations use fixed geometric spaces with inherent hierarchical properties - The original leaf focuses on attention mechanisms as the primary tool, whereas the sibling uses alternative geometric frameworks (hyperbolic, non-Euclidean spaces) - Dynamic attention methods adapt to task-specific requirements during inference, while hyperbolic methods encode structure through the choice of embedding space itself - The original leaf excludes static architectural designs, while hyperbolic methods typically involve static geometric space selection

Suggested Search Directions: - Investigate whether dynamic attention mechanisms could be combined with hyperbolic embeddings for adaptive hierarchical reasoning - Explore how structure-aware attention patterns might emerge naturally in non-Euclidean spaces - Compare the expressiveness of dynamic attention modulation versus geometric space selection for capturing compositional structures

Sibling Subtopics

- **Hyperbolic and Non-Euclidean Representations** (leaves: 1, papers: 1)
- Scope: Approaches using hyperbolic or other non-Euclidean geometric spaces for learning hierarchical or complex visual representations.
- Exclude: Excludes Euclidean-space methods and text-focused models; see Language Model Architectures or Task-Specific Applications.

Contributions Analysis

Overall novelty summary. The paper proposes ATA, a structure-aware dynamic attention mechanism that adaptively identifies active sub-tasks in compositional instructions while suppressing attention to dormant ones. Within the taxonomy, it resides in the 'Dynamic and Structure-Aware Attention' leaf under 'Attention Mechanisms and Model Architecture'. Notably, this leaf contains only the original paper itself—no sibling papers are present—indicating a relatively sparse research direction. The broader parent branch includes one other leaf on hyperbolic representations, suggesting that dynamic attention conditioned on compositional structure represents an emerging rather than crowded area.

The taxonomy reveals that neighboring branches address related but distinct concerns. 'Multi-Task and Auxiliary Learning Frameworks' explores training paradigms with multiple objectives, while 'Task Decomposition and Planning' focuses on breaking complex tasks into executable sub-tasks. ATA diverges by operating within a single forward pass at the attention level, rather than through multi-objective training or explicit task decomposition. The scope note for the original leaf emphasizes 'adaptively modulating attention based on task structure during inference', distinguishing it from static architectural designs and general multi-task methods found in adjacent branches.

Across three identified contributions, the literature search examined 30 candidates total, with 10 candidates per contribution. None of the contributions were clearly refuted by prior work among these candidates. Contribution A (the ATA mechanism itself) examined 10 papers with zero refutable matches, as did Contribution B (identifying three prototypical composition structures) and Contribution C (mutual attention masking). This suggests that within the limited search scope—top-K semantic matches plus citation expansion—no directly overlapping prior work was identified, though the search was not exhaustive.

Given the sparse taxonomy position and absence of refutable prior work among 30 examined candidates, the paper appears to occupy a relatively novel niche. However, the limited search scope means that related work in broader attention mechanism literature or compositional reasoning may exist beyond the candidates examined. The analysis captures novelty within the surveyed subset but does not constitute a comprehensive field-wide assessment.

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

Contribution 1: ATA: Structure-aware dynamic attention mechanism for compositional instructions

Description: The authors introduce ATA, a novel attention mechanism that analyzes compositional instruction structures (chain, branch, parallel) to dynamically identify which sub-task is active at each generation step and suppresses attention to structurally exclusive inactive sub-tasks. This mechanism operates within a single forward pass without parameter updates.

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 multilevel attention network with sub-instructions for continuous vision-and-language navigation
URL: View paper

Brief Assessment

Multilevel Attention Navigation[72] focuses on vision-and-language navigation tasks with sub-instruction processing for robot navigation, not on compositional instruction following in LLMs with mutually exclusive sub-task structures (chain, branch, parallel).

2. Episodic transformer for vision-and-language navigation
URL: View paper

Brief Assessment

Episodic Transformer Navigation[71] focuses on vision-and-language navigation with transformer-based episode history encoding, not on structure-aware dynamic attention for compositional instruction sub-task identification and suppression as proposed in the original paper.

3. Scale Your Instructions: Enhance the Instruction-Following Fidelity of Unified Image Generation Model by Self-Adaptive Attention Scaling
URL: View paper

Brief Assessment

Instruction Attention Scaling[52] focuses on unified image generation models handling multimodal inputs with text instruction neglect issues, not on LLMs processing compositional text instructions with mutually exclusive sub-task structures.

4. Generate subgoal images before act: Unlocking the chain-of-thought reasoning in diffusion model for robot manipulation with multimodal prompts
URL: View paper

Brief Assessment

Subgoal Images Diffusion[73] focuses on generating visual subgoal images for robot manipulation tasks using diffusion models, not on dynamic attention mechanisms for compositional instruction following in LLMs.

5. Augmented hierarchical scene prior learning with context-based scene completion network for visual semantic navigation
URL: View paper

Brief Assessment

Hierarchical Scene Prior[70] focuses on visual semantic navigation with scene representation learning, not on compositional instruction following or dynamic attention mechanisms for sub-task identification in language models.

6. HAPFI: History-Aware Planning based on Fused Information
URL: View paper

Brief Assessment

History-Aware Planning HAPFI[78] focuses on embodied instruction following with multi-modal fusion for robotic action planning, not on attention mechanisms for compositional instruction structures in language models.

7. MLANet: Multi-Level Attention Network with Sub-instruction for Continuous Vision-and-Language Navigation
URL: View paper

Brief Assessment

MLANet Navigation[77] focuses on vision-and-language navigation with sub-instructions in continuous 3D environments, not on compositional instruction following with mutually exclusive sub-task structures (chain/branch/parallel). The attention mechanisms serve different purposes: MLANet uses multi-level attention to fuse vision and language features for navigation actions, while ATA dynamically suppresses attention to structurally inactive sub-tasks during text generation.

8. Hierarchical spatial proximity reasoning for vision and language navigation
URL: View paper

Brief Assessment

Hierarchical Spatial Proximity[74] focuses on vision-and-language navigation tasks using spatial proximity reasoning for path planning in physical environments, not on compositional instruction following with sub-task identification in language models.

9. Neuro-Symbolic Robotics
URL: View paper

Brief Assessment

Neuro-Symbolic Robotics[75] focuses on task decomposition into sub-goal descriptions for robotics applications, not on dynamic attention mechanisms for compositional instruction following in LLMs.

10. Describe, explain, plan and select: interactive planning with llms enables open-world multi-task agents
URL: View paper

Brief Assessment

Interactive Planning LLMs[76] focuses on task planning and goal decomposition in open-world environments using LLMs, not on attention mechanisms for compositional instruction following. The candidate addresses planning challenges through describe-explain-plan-select cycles, while the original contribution proposes a novel attention mechanism (ATA) that dynamically identifies active sub-tasks and suppresses attention to inactive ones within the model's forward pass.

Contribution 2: Systematic identification of three prototypical composition structures

Description: The authors systematically identify and formalize three fundamental composition structures in compositional instructions: chaining (sequential execution), branching (conditional selection), and paralleling (parallel independent tasks). They are the first to introduce the parallel structure in this research area.

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. Superfast Multi-Robot-Arm Motion Planning and Execution
URL: View paper

Brief Assessment

Multi-Robot Motion Planning[69] focuses on multi-robot-arm motion planning in physical manipulation tasks, not on compositional instruction structures in natural language processing. The candidate addresses collision checking and trajectory planning for robotic systems, which is fundamentally different from identifying composition structures (chaining, branching, paralleling) in language model instructions.

2. Pel, a programming language for orchestrating ai agents
URL: View paper

Brief Assessment

PEL Programming Language[65] focuses on designing a programming language for orchestrating AI agents with features like piping mechanisms and closures, not on identifying composition structures in multi-step instructions like chaining, branching, and paralleling.

3. Ai chains: Transparent and controllable human-ai interaction by chaining large language model prompts
URL: View paper

Brief Assessment

AI Chains Transparency[61] focuses on chaining LLM steps for transparency and controllability in complex tasks, not on systematically identifying composition structures (chaining, branching, paralleling) in compositional instructions as the original paper does.

4. A Preliminary Exploration of Evolving Agent Societies through Simple Local Rules
URL: View paper

Brief Assessment

Evolving Agent Societies[68] focuses on multi-agent coordination through local interaction rules (separation, alignment, cohesion) for tool-building ecosystems, not on compositional instruction structures. The candidate does not address chaining, branching, or parallel composition structures in instructions.

5. Multi-level compositional reasoning for interactive instruction following
URL: View paper

Brief Assessment

Compositional Instruction Following[54] focuses on multi-level compositional reasoning for interactive instruction following in robotic navigation tasks, not on systematically identifying composition structures in general compositional instructions. The candidate does not address the parallel structure introduced by the original paper.

6. Demystifying chains, trees, and graphs of thoughts
URL: View paper

Brief Assessment

Chains Trees Graphs[63] focuses on prompting topologies (chains, trees, graphs) for LLM reasoning structures, not on compositional instruction structures (chaining, branching, paralleling) for multi-step task execution as in the original paper.

7. Incentivizing Reasoning for Advanced Instruction-Following of Large Language Models
URL: View paper

Brief Assessment

Incentivizing Reasoning Instructions[64] acknowledges the same three composition structures (paralleling, chaining, and branching) but does not claim to be the first to identify them. The candidate focuses on reinforcement learning methods for handling these structures rather than claiming novelty in their identification.

8. Exposing Limitations of Language Model Agents in Sequential-Task Compositions on the Web
URL: View paper

Brief Assessment

Sequential Task Limitations[67] focuses on web automation tasks with sequential compositions and reverse-order instructions, not on systematically identifying and formalizing composition structures (chaining, branching, paralleling) as fundamental paradigms across general compositional instructions.

9. Scaling long-horizon llm agent via context-folding
URL: View paper

Brief Assessment

Context Folding LLM[62] focuses on context management through branching and folding mechanisms for long-horizon tasks, not on identifying composition structures (chaining, branching, paralleling) in compositional instructions as the original paper does.

10. An LLM-Tool Compiler for Fused Parallel Function Calling
URL: View paper

Brief Assessment

LLM Tool Compiler[66] focuses on fusing parallel function calls in API execution for efficiency optimization, not on identifying or formalizing composition structures (chaining, branching, paralleling) in compositional instructions as the original paper does.

Contribution 3: Mutual attention masking between exclusive sub-tasks during encoding

Description: The authors propose a mutual attention masking technique that prevents attention flow between structurally exclusive sub-task pairs during the encoding phase. This prevents blending comprehension of multiple mutually exclusive sub-tasks and ensures their representations remain independent.

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. Masked Diffusion with Task-awareness for Procedure Planning in Instructional Videos
URL: View paper

Brief Assessment

Masked Diffusion Planning[56] applies masking to action dimensions in procedure planning for instructional videos, not to prevent attention flow between exclusive sub-tasks during instruction encoding in LLMs. The technical contexts are fundamentally different: video-based action planning versus compositional instruction following in language models.

2. Focus on your instruction: Fine-grained and multi-instruction image editing by attention modulation
URL: View paper

Brief Assessment

Instruction Attention Modulation[51] focuses on image editing with attention modulation between image regions and instruction keywords, not on preventing attention flow between mutually exclusive sub-tasks in compositional instruction following for LLMs.

3. Internal Chain-of-Thought: Empirical Evidence for Layer‑wise Subtask Scheduling in LLMs
URL: View paper

Brief Assessment

Internal Chain of Thought[55] focuses on layer-wise subtask scheduling in composite tasks (e.g., antonym then uppercase), not on preventing attention flow between mutually exclusive sub-tasks during encoding as proposed in the original paper's ATA framework.

4. Attention Hijackers: Detect and Disentangle Attention Hijacking in LVLMs for Hallucination Mitigation
URL: View paper

Brief Assessment

Attention Hijacking Detection[59] focuses on masking attention from disruptive instruction tokens to image regions in vision-language models to reduce hallucinations. The ORIGINAL paper addresses masking attention between mutually exclusive sub-tasks in text-only compositional instructions. These are fundamentally different technical problems in different domains (vision-language vs. text-only LLMs).

5. Plan, eliminate, and track--language models are good teachers for embodied agents
URL: View paper

Brief Assessment

Plan Eliminate Track[57] focuses on using LLMs to simplify control tasks for embodied agents through planning, object elimination, and progress tracking in interactive environments. It does not address attention masking techniques for disentangling exclusive sub-tasks during instruction encoding in LLMs.

6. Scale Your Instructions: Enhance the Instruction-Following Fidelity of Unified Image Generation Model by Self-Adaptive Attention Scaling
URL: View paper

Brief Assessment

Instruction Attention Scaling[52] addresses cross-attention conflicts between neglected sub-instructions and input image activations in image generation, not mutual attention masking between structurally exclusive text sub-tasks during encoding.

7. Multi-level compositional reasoning for interactive instruction following
URL: View paper

Brief Assessment

Compositional Instruction Following[54] proposes a multi-level hierarchical framework with separate policies for navigation and interaction, but does not describe mutual attention masking techniques for disentangling exclusive sub-tasks during the encoding phase as proposed in the original paper.

8. Dual Attention Alignment for Clause-Grounded Instruction-based Legal Question Answering
URL: View paper

Brief Assessment

Clause Grounded Legal[58] focuses on dual attention alignment for legal QA with clause retrieval and answer generation, not on masking attention between mutually exclusive sub-tasks in compositional instruction following.

9. Long-vla: Unleashing long-horizon capability of vision language action model for robot manipulation
URL: View paper

Brief Assessment

Long Horizon VLA[53] focuses on phase-aware masking for robotic manipulation tasks (moving vs. interaction phases), not on preventing attention flow between mutually exclusive sub-tasks in compositional instruction following.

10. Disentangling Instruction Influence in Diffusion Transformers for Parallel Multi-Instruction-Guided Image Editing
URL: View paper

Brief Assessment

Disentangling Instruction Influence[60] focuses on image editing with diffusion transformers, not LLM instruction-following. The masking operates on visual attention between image regions and editing instructions, not on textual sub-task comprehension in compositional instructions.

Appendix: Text Similarity Detection

No high-similarity text segments were detected across any compared papers.

References

- [0] Attend to the Active: Structure-Aware Dynamic Attention in LLMs for Compositional Instruction Following View paper
- [1] Objectives of the Study View paper
- [2] The primary objectives of the show View paper
- [3] Social equity â□ primary objective of the Parliament of the Republic of Moldova View paper
- [4] Modern research problems of modern sociology View paper
- [5] Implications of Adnexal Invasions in Primary Extramammary Pagetâ□s Disease: A Systematic Review View paper
- [6] Measuring Mathematical Problem Solving With the MATH Dataset View paper
- [7] Research problems in function theory View paper
- [8] Interpretable Preferences via Multi-Objective Reward Modeling and Mixture-of-Experts View paper
- [9] Pragmatism as a research paradigm and its implications for social work research View paper
- [10] The main tasks of ensuring the cybersecurity of ENERGY View paper
- [11] Terahertz Band Communication: An Old Problem Revisited and Research Directions for the Next Decade View paper
- [12] Describe, Explain, Plan and Select: Interactive Planning with Large Language Models Enables Open-World Multi-Task Agents View paper
- [13] Spatial multiâ□objective optimization of primary healthcare facilities: A case study in Singapore View paper
- [14] PANDA study: subtyping of Parkinsonâ□s disease â□ cohort study protocol View paper
- [15] Carcinoma ex Pleomorphic Adenoma: Multiâ□Institutional Retrospective Cohort Study View paper
- [16] Biomedical Research: The Research Problem Matters View paper
- [17] Objectives, design and main findings until 2020 from the Rotterdam Study View paper
- [18] Text algorithms in economics View paper
- [19] Management and outcome of primary CNS lymphoma in the modern era: An LOC network study. View paper
- [20] Endogeneity bias in marketing research: Problem, causes and remedies View paper
- [21] A guide to research methodology: An overview of research problems, tasks and methods View paper
- [22] Screening for Colorectal Cancer: Updated Evidence Report and Systematic Review for the US Preventive Services Task Force. View paper
- [23] Evaluation of gastroschisis feeding protocol: A retrospective cohort study View paper
- [24] [Competency-based learning objectives of primary qualifying nursing studies and intended task profiles]. View paper
- [25] SemEval-2023 Task 10: Explainable Detection of Online Sexism View paper
- [26] The Problem of Abortion and the Doctrine of the Double Effect View paper
- [27] Are hormones a â□female problemâ□ for animal research? View paper
- [28] Modularization in Object Oriented Software: A Comparative Study View paper
- [29] Dynamic Auxiliary Task-Based Evolutionary Multitasking for Constrained Multiobjective Optimization View paper
- [30] Is Quantum Advantage the Right Goal for Quantum Machine Learning? View paper
- [31] Complement objective training View paper
- [32] Minimum information for studies of extracellular vesicles (MISEV) as toolbox for rigorous, reproducible and homogeneous studies on extracellular vesicles. View paper
- [33] An efficient grasping shared control architecture for unpredictable and unspecified tasks View paper
- [34] Prior knowledge-informed multi-task dynamic learning for few-shot machinery fault diagnosis View paper
- [35] SemEval-2022 Task 5: Multimedia Automatic Misogyny Identification View paper
- [36] Research problems View paper
- [37] Towards the IMO's GHG goals: A critical overview of the perspectives and challenges of the main options for decarbonizing international shipping View paper
- [38] Practice Standards for Addressing Social Justice in Social Work Research View paper
- [39] Formulating a good research question: Pearls and pitfalls View paper
- [40] Impact of the COVID-19 pandemic on the core functions of primary care: will the cure be worse than the disease? A qualitative interview study in Flemish GPs View paper
- [41] The Research Problem View paper
- [42] Efficacy of Pembrolizumab in Patients With Noncolorectal High Microsatellite Instability/Mismatch Repair-Deficient Cancer: Results From the Phase II KEYNOTE-158 Study. View paper
- [43] A survey of public datasets for computer vision tasks in precision agriculture View paper
- [44] Examining the Double-Edged Sword Effect of AI Usage on Work Engagement: The Moderating Role of Core Task Characteristics Substitution View paper
- [45] What is a researchable question in design? View paper
- [46] Detecting primary aldosteronism in Australian primary care: a prospective study View paper
- [47] Hyperbolic Deep Learning in Computer Vision: A Survey View paper
- [48] Identifying key policy objectives for strong primary care: a cross-sectional study View paper
- [49] A Simple Language Model for Task-Oriented Dialogue View paper
- [50] Task Partitioning and Offloading in DNN-Task Enabled Mobile Edge Computing Networks View paper
- [51] Focus on your instruction: Fine-grained and multi-instruction image editing by attention modulation View paper
- [52] Scale Your Instructions: Enhance the Instruction-Following Fidelity of Unified Image Generation Model by Self-Adaptive Attention Scaling View paper
- [53] Long-vla: Unleashing long-horizon capability of vision language action model for robot manipulation View paper
- [54] Multi-level compositional reasoning for interactive instruction following View paper
- [55] Internal Chain-of-Thought: Empirical Evidence for Layerâ□wise Subtask Scheduling in LLMs View paper
- [56] Masked Diffusion with Task-awareness for Procedure Planning in Instructional Videos View paper
- [57] Plan, eliminate, and track--language models are good teachers for embodied agents View paper
- [58] Dual Attention Alignment for Clause-Grounded Instruction-based Legal Question Answering View paper
- [59] Attention Hijackers: Detect and Disentangle Attention Hijacking in LVLMs for Hallucination Mitigation View paper

- [60] Disentangling Instruction Influence in Diffusion Transformers for Parallel Multi-Instruction-Guided Image Editing View paper
- [61] Ai chains: Transparent and controllable human-ai interaction by chaining large language model prompts View paper
- [62] Scaling long-horizon llm agent via context-folding View paper
- [63] Demystifying chains, trees, and graphs of thoughts View paper
- [64] Incentivizing Reasoning for Advanced Instruction-Following of Large Language Models View paper
- [65] Pel, a programming language for orchestrating ai agents View paper
- [66] An LLM-Tool Compiler for Fused Parallel Function Calling View paper
- [67] Exposing Limitations of Language Model Agents in Sequential-Task Compositions on the Web View paper
- [68] A Preliminary Exploration of Evolving Agent Societies through Simple Local Rules View paper
- [69] Superfast Multi-Robot-Arm Motion Planning and Execution View paper
- [70] Augmented hierarchical scene prior learning with context-based scene completion network for visual semantic navigation View paper
- [71] Episodic transformer for vision-and-language navigation View paper
- [72] A multilevel attention network with sub-instructions for continuous vision-and-language navigation View paper
- [73] Generate subgoal images before act: Unlocking the chain-of-thought reasoning in diffusion model for robot manipulation with multimodal prompts View paper
- [74] Hierarchical spatial proximity reasoning for vision and-language navigation View paper
- [75] Neuro-Symbolic Robotics View paper
- [76] Describe, explain, plan and select: interactive planning with llms enables open-world multi-task agents View paper
- [77] MLANet: Multi-Level Attention Network with Sub-instruction for Continuous Vision-and-Language Navigation View paper
- [78] HAPFI: History-Aware Planning based on Fused Information View paper