Novelty Assessment Report

Paper: Mixing Mechanisms: How Language Models Retrieve Bound Entities In-Context

PDF URL: https://openreview.net/pdf?id=UJ2UUjT2ko

Venue: ICLR 2026 Conference Submission

Year: 2026

Report Generated: 2025-12-27

Abstract

A key component of in-context reasoning is the ability of language models (LMs) to bind entities for later retrieval. For example, an LM might represent Ann loves pie by binding Ann to pie, allowing it to later retrieve Ann when asked Who loves pie? Prior research on short lists of bound entities found strong evidence that LMs implement such retrieval via a **positional mechanism**, where Ann is retrieved based on its position in context. In this work, we find that this mechanism generalizes poorly to more complex settings; as the number of bound entities in context increases, the positional mechanism becomes noisy and unreliable in middle positions. To compensate for this, we find that LMs supplement the positional mechanism with a **lexical mechanism** (retrieving Ann using its bound counterpart pie) and a **reflexive mechanism** (retrieving Ann through a direct pointer). Through extensive experiments on nine models and ten binding tasks, we uncover a consistent pattern in how LMs mix these mechanisms to drive model behavior. We leverage these insights to develop a causal model combining all three mechanisms that estimates next token distributions with 95\% agreement. Finally, we show that our model generalizes to substantially longer inputs of open-ended text interleaved with entity groups, further demonstrating the robustness of our findings in more natural settings. Overall, our study establishes a more complete picture of how LMs bind and retrieve entities in-context.

Core Task Landscape

This paper addresses: **Entity Binding and Retrieval in Language Models**

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:

- **Entity Binding Mechanisms and Representations**
- **Entity Linking Systems**
- **Named Entity Recognition**
- **Retrieval-Augmented Entity Processing**
- **In-Context Learning and Entity Reasoning**
- **Specialized Entity Applications**

Complete Taxonomy Tree

- Entity Binding and Retrieval in Language Models Survey Taxonomy
- Entity Binding Mechanisms and Representations
  - In-Context Entity Binding and Tracking ★ (2 papers)
  - [0] Mixing Mechanisms: How Language Models Retrieve Bound Entities In-Context (Anon et al., 2026) View paper
  - [4] Representational analysis of binding in language models (Qin Dai, 2024) View paper
  - Entity Knowledge Representation in Pretrained Models (5 papers)
  - [17] Language models as knowledge bases: On entity representations, storage capacity, and paraphrased queries (Benjamin Heinzerling, 2021) View paper
  - [26] Locating and Extracting Relational Concepts in Large Language Models (Wang Zi-jian, 2024) View paper
  - [27] How can we know what language models know? (Zhengbao Jiang, 2020) View paper
  - [40] On Entity Identification in Language Models (Heinzerling, 2025) View paper
  - [44] Vector Space Transformations to Uncover Knowledge Graphs in Neural Language Models (Giacomo Munda, 2025) View paper
  - Compositional Binding in Vision-Language Models (2 papers)
  - [24] Understanding the limits of vision language models through the lens of the binding problem (Declan Campbell, 2024) View paper
  - [28] ComAlign: Compositional Alignment in Vision-Language Models (Abdollah, 2024) View paper
- Entity Linking Systems
  - Neural Entity Linking Architectures (4 papers)
  - [8] Neural entity linking: A survey of models based on deep learning (Sevgili, 2022) View paper
  - [9] Scalable zero-shot entity linking with dense entity retrieval (Ledell Wu, 2020) View paper
  - [15] Rel: An entity linker standing on the shoulders of giants (J, 2020) View paper
  - [45] End-to-end neural entity linking (Kolitsas, 2018) View paper
  - LLM-Based Entity Linking (3 papers)
  - [6] Llmael: Large language models are good context augmenters for entity linking (Hou Lei, 2025) View paper
  - [11] EntGPT: Entity Linking with Generative Large Language Models (Ding, 2024) View paper
  - [29] Instructed language models with retrievers are powerful entity linkers (Zilin Xiao, 2023) View paper
  - Multimodal and Domain-Specific Entity Linking (5 papers)

    - [13] Clinical entity recognition and linking in greek discharge letters using multilingual-llm-based multi-stage system (Huang, 2025) View paper
    - [19] Unimel: A unified framework for multimodal entity linking with large language models (Qi Liu, 2024) View paper
    - [34] Cross-lingual Unified Medical Language System entity linking in online health communities (Yonatan Bitton, 2020) View paper
    - [36] BioLinkerAI: Capturing Knowledge Using LLMs to Enhance Biomedical Entity Linking (Ahmad Sakor, 2024) View paper
    - [46] Open-set entity alignment using large language models with retrieval augmentation (Linyao Yang, 2024) View paper
    - End-to-End Entity Linking with BERT (1 papers)
    - [30] Investigating entity knowledge in BERT with simple neural end-to-end entity linking (Broscheit, 2019) View paper
- Named Entity Recognition
    - LLM-Based Named Entity Recognition (4 papers)
    - [1] Gpt-ner: Named entity recognition via large language models (Shuhe Wang, 2025) View paper
    - [5] Universalner: Targeted distillation from large language models for open named entity recognition (Zhou, 2023) View paper
    - [10] Linkner: Linking local named entity recognition models to large language models using uncertainty (Zhen Zhang, 2024) View paper
    - [50] Large Language Models Struggle in Token-Level Clinical Named Entity Recognition. (Lu Qiuhao, 2025) View paper
    - Neural and Feature-Based NER (3 papers)
    - [35] Robust lexical features for improved neural network named-entity recognition (Abbas Ghaddar, 2018) View paper
    - [37] End-to-end named entity recognition and relation extraction using pre-trained language models (Giorgi, 2019) View paper
    - [38] Can BERT dig it? Named entity recognition for information retrieval in the archaeology domain (Brandsen, 2022) View paper
    - Next-Generation NER and Retrieval (1 papers)
    - [33] NERetrieve: Dataset for next generation named entity recognition and retrieval (Cohen, 2023) View paper
- Retrieval-Augmented Entity Processing
    - General Retrieval-Augmented Language Models (4 papers)
    - [2] Retrieve anything to augment large language models (Zhang Pei-tian, 2023) View paper
    - [12] Retrieval augmented language model pre-training (Kelvin Guu, 2020) View paper
    - [18] Self-retrieval: End-to-end information retrieval with one large language model (Jiawei Chen, 2024) View paper
    - [22] Reimagining retrieval augmented language models for answering queries (Halevy, 2023) View paper
    - Entity-Centric Dense Retrieval (3 papers)
    - [20] Information retrieval with entity linking (Shehata, 2024) View paper
    - [48] Dense retrieval with entity views (Hai-Dang Tran, 2022) View paper
    - [49] Entity-aware transformers for entity search (Gerritse, 2022) View paper
    - Entity Retrieval for Question Answering (3 papers)
    - [14] Entity retrieval for answering entity-centric questions (Sarkar, 2025) View paper
    - [32] Entity retrieval using fine-grained entity aspects (Shubham Chatterjee, 2021) View paper
    - [42] Exploiting entity linking in queries for entity retrieval (Faegheh Hasibi, 2016) View paper
    - Entity-Based Language Models for Retrieval (2 papers)
    - [21] Barack's wife hillary: Using knowledge-graphs for fact-aware language modeling (Robert Logan, 2019) View paper
    - [23] Document retrieval using entity-based language models (Hadas Raviv, 2016) View paper
- In-Context Learning and Entity Reasoning (2 papers)
    - [16] Explore the in-context learning capability of large language models (Tianle, 2024) View paper
    - [41] Recall, retrieve and reason: Towards better in-context relation extraction (Li Guozheng, 2024) View paper
- Specialized Entity Applications
    - Biomedical and Scientific Entity Processing (3 papers)
    - [7] Rethinking Text-based Protein Understanding: Retrieval or LLM? (Liu Zi-jing, 2025) View paper
    - [39] Phenomics assistant: An interface for LLM-based biomedical knowledge graph exploration (Shawn T. Oâ□□Neil, 2024) View paper
    - [47] Real: A retrieval-augmented entity linking approach for biomedical concept recognition (Darya Shlyk, 2024) View paper
    - Visual Entity Recognition and Grounding (1 papers)
    - [3] Grounding language models for visual entity recognition (Zilin Xiao, 2024) View paper
    - Structured Data and Linked Data Retrieval (1 papers)
    - [31] Towards Enhancing Linked Data Retrieval in Conversational UIs Using Large Language Models (Omar Mussa, 2024) View paper
    - Domain-Specific Neural Production Systems (1 papers)
    - [25] Twitter stance detection via neural production systems (Shenzhen, 2023) View paper
    - Diffusion Models and Entity Manipulation (1 papers)
    - [43] Magnet: We never know how text-to-image diffusion models work, until we learn how vision-language models function (Pan Gao, 2024) View paper

Narrative

Core task: entity binding and retrieval in language models. This field examines how models associate linguistic mentions with structured entity representations and retrieve relevant entity information during inference. The taxonomy organizes research into several major branches: Entity Binding Mechanisms and Representations explores how models internally encode and track entities, including in-context binding strategies; Entity Linking Systems focuses on mapping text spans to knowledge base entries, often through dense retrieval methods like Dense Entity Retrieval[9] or neural architectures surveyed in Neural Entity Linking Survey[8]; Named Entity Recognition addresses the identification and classification of entity mentions, with approaches ranging from traditional methods to large language model adaptations such as GPT NER[1] and UniversalNER[5]; Retrieval-Augmented Entity Processing integrates external knowledge sources to enhance entity understanding, exemplified by Retrieve Anything[2] and Retrieval Augmented Pretraining[12]; In-Context Learning and Entity Reasoning investigates how models leverage contextual examples to perform entity-related tasks, as seen in In Context Learning[16]; and Specialized Entity Applications targets domain-specific challenges in areas like biomedicine, with works such as Greek Clinical Entity[13] and BioLinkerAI[36].

Across these branches, a central tension emerges between end-to-end neural approaches that learn entity representations implicitly and modular systems that explicitly retrieve and bind entities to external knowledge. Many studies explore how retrieval mechanisms can be tightly integrated with language model architectures, balancing computational efficiency with representational richness. Mixing Mechanisms[0] sits within the Entity Binding Mechanisms branch, specifically addressing in-context entity binding and tracking—a line of work concerned with how models maintain entity references across discourse without explicit symbolic grounding. This emphasis aligns

closely with Binding Representational Analysis[4], which examines the internal representational structures that support entity tracking, and contrasts with retrieval-heavy approaches like Dense Entity Retrieval[9] that rely on external knowledge bases. While Grounding Visual Entity[3] extends binding to multimodal settings, Mixing Mechanisms[0] focuses on the linguistic mechanisms that enable coherent entity reference resolution within textual contexts, highlighting ongoing questions about the sufficiency of implicit versus explicit entity representations.

Related Works in Same Category

The following **1 sibling papers** share the same taxonomy leaf node with the original paper:

1. Representational analysis of binding in language models

Authors: Qin Dai, Benjamin Heinzerling, Kentaro Inui | **Year/Venue**: 2024 | **URL**: View paper

Abstract

â¦ Entity tracking is essential for complex reasoning. To perform in-context entity tracking, language models (LMs) must bind an entity to its attribute (eg, bind a container to its content) to â¦

Relationship Analysis

Both papers belong to the same taxonomy category studying in-context entity binding and tracking mechanisms in language models. They share overlapping focus on how models bind entities to attributes and retrieve them, with both employing causal intervention methods (interchange interventions/activation patching) to probe internal mechanisms. The key difference is that the original paper identifies three distinct retrieval mechanisms (positional, lexical, and reflexive) and their mixture across context positions, while the candidate paper focuses specifically on discovering and analyzing an "Ordering ID" (OI) subspace through dimensionality reduction techniques like PCA, providing a representational analysis of how positional information is encoded in model activations.

Contributions Analysis

Overall novelty summary. The paper investigates how language models retrieve bound entities in context, identifying three distinct mechanisms: positional (retrieving based on context position), lexical (using bound counterparts), and reflexive (direct pointers). It resides in the 'In-Context Entity Binding and Tracking' leaf, which contains only two papers total, indicating a relatively sparse research direction. This leaf focuses specifically on internal binding mechanisms during inference, distinguishing it from the broader 'Entity Knowledge Representation in Pretrained Models' branch (five papers) that examines parametric entity storage rather than dynamic in-context tracking.

The taxonomy reveals that entity binding research divides into several neighboring areas: entity linking systems (four leaves, ~16 papers) focus on mapping mentions to external knowledge bases, while retrieval-augmented approaches (four leaves, ~13 papers) integrate external knowledge sources. The paper's leaf explicitly excludes these external-knowledge methods, positioning the work within a narrower investigation of purely internal mechanisms. The 'In-Context Learning and Entity Reasoning' branch (two papers) addresses related phenomena but emphasizes task performance over mechanistic analysis, whereas this work dissects the underlying retrieval strategies.

Among 20 candidates examined across three contributions, no clearly refuting prior work was identified. The 'Discovery of three mechanisms' contribution examined one candidate with no refutation; the 'Causal model combining mechanisms' examined nine candidates, none refuting; and the 'Counterfactual intervention methodology' examined ten candidates, also without refutation. This suggests that within the limited search scope, the specific combination of positional, lexical, and reflexive mechanisms—and their integration into a unified causal model—appears not to have direct precedent in the examined literature.

The analysis reflects a constrained literature search (top-20 semantic matches), not an exhaustive survey. The sparse population of the target leaf (two papers) and absence of refuting candidates among examined works suggest the mechanistic decomposition may be novel within this scope. However, the limited search scale means potentially relevant work in adjacent areas—such as attention mechanism studies or broader interpretability research—may not have been captured, leaving open questions about the contribution's novelty relative to the full field.

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

Contribution 1: Discovery of three mechanisms for entity retrieval

Description: The authors identify that language models use not just a positional mechanism but also lexical and reflexive mechanisms to retrieve bound entities in context. The lexical mechanism retrieves entities using their bound counterparts, while the reflexive mechanism uses direct self-referential pointers.

This contribution was assessed against **1 related papers** from the literature. Papers with potential prior art are analyzed in detail with textual evidence; others receive brief assessments.

1. Named Entity Recognition in Persian Language based on Self-attention Mechanism with Weighted Relational Position Encoding

URL: View paper

Brief Assessment

Persian NER Self Attention[61] focuses on named entity recognition using weighted relative positional encoding for Persian language processing, not on entity retrieval mechanisms in language models' in-context reasoning.

Contribution 2: Causal model combining three mechanisms

Description: The authors develop a formal causal model that combines positional, lexical, and reflexive mechanisms as a position-weighted mixture to predict next token distributions. This model achieves 95% agreement with actual language model behavior.

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. A Mechanistic Interpretation of Arithmetic Reasoning in Language Models using Causal Mediation Analysis

URL: View paper

Brief Assessment

Arithmetic Reasoning Mediation[67] focuses on arithmetic reasoning in language models using causal mediation analysis to identify which model components (MLPs, attention) process numerical information. The ORIGINAL paper addresses entity binding and retrieval in context using positional, lexical, and reflexive mechanisms—a fundamentally different task domain.

---

2. Causal Head Gating: A Framework for Interpreting Roles of Attention Heads in Transformers

URL: View paper

Brief Assessment

Causal Head Gating[64] focuses on interpreting attention head roles in transformers through learned soft gates and causal taxonomy, not on developing causal models that combine positional, lexical, and reflexive mechanisms for next token prediction in entity binding tasks.

---

3. Using deep autoregressive models as causal inference engines

URL: View paper

Brief Assessment

Autoregressive Causal Inference[70] focuses on using deep autoregressive models for causal inference with high-dimensional confounders and sequential actions, not on developing causal models that combine positional, lexical, and reflexive mechanisms for next token prediction in language models.

---

4. CAMEF: Causal-augmented multi-modality event-driven financial forecasting by integrating time series patterns and salient macroeconomic announcements

URL: View paper

Brief Assessment

CAMEF[69] focuses on causal learning for financial forecasting by integrating textual and time-series data, not on causal models for next token prediction in language models. The domains and technical approaches are fundamentally different.

---

5. Fine-Grained Pavement Performance Prediction Based on Causal-Temporal Graph Convolution Networks

URL: View paper

Brief Assessment

Pavement Performance Prediction[66] focuses on pavement deterioration prediction using causal discovery for infrastructure management, not on language model mechanisms for next token prediction. The domains and technical approaches are entirely different.

---

6. Token-Level Uncertainty-Aware Objective for Language Model Post-Training

URL: View paper

Brief Assessment

Token Level Uncertainty[71] focuses on token-level uncertainty estimation and training objectives (masked MLE, self-distillation) for language model post-training. It does not address causal models for next token prediction that combine positional, lexical, and reflexive mechanisms for entity binding and retrieval.

---

7. Non-markovian discrete diffusion with causal language models

URL: View paper

Brief Assessment

Non Markovian Diffusion[68] focuses on discrete diffusion models for sequence generation, not on causal models for next token prediction in language models. The paper addresses non-Markovian diffusion processes rather than mechanisms for entity binding and retrieval.

---

8. Causal Intervention Is What Large Language Models Need for Spatio-Temporal Forecasting

URL: View paper

Brief Assessment

Causal Intervention Forecasting[63] focuses on spatio-temporal forecasting with causal intervention encoders for adaptive graphs, not on developing causal models for next token prediction in language models.

---

9. Diffusion Forcing: Next-token Prediction Meets Full-Sequence Diffusion

URL: View paper

Brief Assessment

Diffusion Forcing[65] focuses on a training paradigm for diffusion models with per-token noise levels for sequence generation, not on developing causal models that combine positional, lexical, and reflexive mechanisms for next token prediction in language models.

Contribution 3: Counterfactual intervention methodology

Description: The authors design a novel counterfactual dataset construction method where interchange interventions on paired inputs cause each of the three proposed mechanisms to predict different entities, enabling systematic separation and validation of the mechanisms.

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. The Non-Linear Representation Dilemma: Is Causal Abstraction Enough for Mechanistic Interpretability?

URL: View paper

Brief Assessment

Non Linear Representation Dilemma[53] focuses on theoretical limitations of causal abstraction with arbitrary alignment maps, not on designing counterfactual datasets to distinguish specific mechanisms. The original paper's contribution is a novel dataset construction method for separating positional, lexical, and reflexive mechanisms in entity binding tasks.

---

2. Counterfactual explanations for face forgery detection via adversarial removal of artifacts

URL: View paper

Brief Assessment

Counterfactual Forgery Detection[58] focuses on generating counterfactual explanations for face forgery detection by adversarially removing artifacts in deepfake images, not on validating distinct mechanism predictions through counterfactual interventions on paired inputs as in the original paper's entity binding task.

---

3. Gender slopes: Counterfactual fairness for computer vision models by attribute manipulation

URL: View paper

Brief Assessment

Gender Slopes[59] focuses on fairness evaluation in computer vision models using image attribute manipulation to measure demographic bias, not on validating distinct mechanism predictions in language models through counterfactual dataset construction.

---

4. An exact and robust conformal inference method for counterfactual and synthetic controls

URL: View paper

Brief Assessment

Conformal Inference Counterfactual[54] focuses on statistical inference methods for policy evaluation using counterfactual predictions in aggregate time series settings, not on designing counterfactual datasets to validate distinct mechanism predictions in language models.

---

5. Causal inference and counterfactual prediction in machine learning for actionable healthcare

URL: View paper

Brief Assessment

Causal Inference Healthcare[51] focuses on causal inference and counterfactual prediction in healthcare applications, not on mechanistic interpretability of language models or entity binding validation through counterfactual dataset construction.

---

6. Causal Abstraction: A Theoretical Foundation for Mechanistic Interpretability

URL: View paper

Brief Assessment

Causal Abstraction Foundation[57] provides a theoretical framework for mechanistic interpretability and unifies various intervention methods, but does not present a novel counterfactual dataset construction method for systematically separating and validating distinct mechanisms through interchange interventions as described in the original paper.

---

7. Spatiotemporal Causal Inference With Mechanistic Ecological Models: Evaluating Targeted Culling on Chronic Wasting Disease Dynamics in Cervids

URL: View paper

Brief Assessment

Spatiotemporal Causal Inference[60] focuses on causal inference for wildlife disease interventions using inverse-probability-of-treatment-weighted techniques with PDEs, not on counterfactual dataset construction for validating distinct mechanism predictions in language models.

---

8. Causal discovery in heterogeneous environments under the sparse mechanism shift hypothesis

URL: View paper

Brief Assessment

Sparse Mechanism Shift[55] focuses on causal discovery in heterogeneous environments using mechanism shift scores, not on counterfactual dataset construction for validating distinct mechanism predictions in language models.

---

9. Large-scale hypothesis testing for causal mediation effects with applications in genome-wide epigenetic studies

URL: View paper

Brief Assessment

Causal Mediation Hypothesis Testing[56] focuses on statistical hypothesis testing for mediation effects in genome-wide epigenetic studies, not on counterfactual dataset construction methods for validating mechanism predictions in language models.

---

10. Strategic branch management in commercial banks: Driving profitability through operational efficiency

URL: View paper

Brief Assessment

Branch Management Profitability[52] focuses on operational efficiency in commercial banking using matched branch pools for counterfactual trajectories, not on validating distinct mechanism predictions in language models through interchange interventions on paired inputs.

Appendix: Text Similarity Detection

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

References

- [0] Mixing Mechanisms: How Language Models Retrieve Bound Entities In-Context View paper
- [1] Gpt-ner: Named entity recognition via large language models View paper
- [2] Retrieve anything to augment large language models View paper
- [3] Grounding language models for visual entity recognition View paper
- [4] Representational analysis of binding in language models View paper
- [5] Universalner: Targeted distillation from large language models for open named entity recognition View paper
- [6] Llmael: Large language models are good context augmenters for entity linking View paper
- [7] Rethinking Text-based Protein Understanding: Retrieval or LLM? View paper
- [8] Neural entity linking: A survey of models based on deep learning View paper
- [9] Scalable zero-shot entity linking with dense entity retrieval View paper
- [10] Linkner: Linking local named entity recognition models to large language models using uncertainty View paper
- [11] EntGPT: Entity Linking with Generative Large Language Models View paper
- [12] Retrieval augmented language model pre-training View paper
- [13] Clinical entity recognition and linking in greek discharge letters using multilingual-llm-based multi-stage system View paper
- [14] Entity retrieval for answering entity-centric questions View paper
- [15] Rel: An entity linker standing on the shoulders of giants View paper
- [16] Explore the in-context learning capability of large language models View paper
- [17] Language models as knowledge bases: On entity representations, storage capacity, and paraphrased queries View paper
- [18] Self-retrieval: End-to-end information retrieval with one large language model View paper
- [19] Unimel: A unified framework for multimodal entity linking with large language models View paper
- [20] Information retrieval with entity linking View paper
- [21] Barack's wife hillary: Using knowledge-graphs for fact-aware language modeling View paper
- [22] Reimagining retrieval augmented language models for answering queries View paper

- [23] Document retrieval using entity-based language models View paper
- [24] Understanding the limits of vision language models through the lens of the binding problem View paper
- [25] Twitter stance detection via neural production systems View paper
- [26] Locating and Extracting Relational Concepts in Large Language Models View paper
- [27] How can we know what language models know? View paper
- [28] ComAlign: Compositional Alignment in Vision-Language Models View paper
- [29] Instructed language models with retrievers are powerful entity linkers View paper
- [30] Investigating entity knowledge in BERT with simple neural end-to-end entity linking View paper
- [31] Towards Enhancing Linked Data Retrieval in Conversational UIs Using Large Language Models View paper
- [32] Entity retrieval using fine-grained entity aspects View paper
- [33] NERetrieve: Dataset for next generation named entity recognition and retrieval View paper
- [34] Cross-lingual Unified Medical Language System entity linking in online health communities View paper
- [35] Robust lexical features for improved neural network named-entity recognition View paper
- [36] BioLinkerAI: Capturing Knowledge Using LLMs to Enhance Biomedical Entity Linking View paper
- [37] End-to-end named entity recognition and relation extraction using pre-trained language models View paper
- [38] Can BERT dig it? Named entity recognition for information retrieval in the archaeology domain View paper
- [39] Phenomics assistant: An interface for LLM-based biomedical knowledge graph exploration View paper
- [40] On Entity Identification in Language Models View paper
- [41] Recall, retrieve and reason: Towards better in-context relation extraction View paper
- [42] Exploiting entity linking in queries for entity retrieval View paper
- [43] Magnet: We never know how text-to-image diffusion models work, until we learn how vision-language models function View paper
- [44] Vector Space Transformations to Uncover Knowledge Graphs in Neural Language Models View paper
- [45] End-to-end neural entity linking View paper
- [46] Open-set entity alignment using large language models with retrieval augmentation View paper
- [47] Real: A retrieval-augmented entity linking approach for biomedical concept recognition View paper
- [48] Dense retrieval with entity views View paper
- [49] Entity-aware transformers for entity search View paper
- [50] Large Language Models Struggle in Token-Level Clinical Named Entity Recognition. View paper
- [51] Causal inference and counterfactual prediction in machine learning for actionable healthcare View paper
- [52] Strategic branch management in commercial banks: Driving profitability through operational efficiency View paper
- [53] The Non-Linear Representation Dilemma: Is Causal Abstraction Enough for Mechanistic Interpretability? View paper
- [54] An exact and robust conformal inference method for counterfactual and synthetic controls View paper
- [55] Causal discovery in heterogeneous environments under the sparse mechanism shift hypothesis View paper
- [56] Large-scale hypothesis testing for causal mediation effects with applications in genome-wide epigenetic studies View paper
- [57] Causal Abstraction: A Theoretical Foundation for Mechanistic Interpretability View paper
- [58] Counterfactual explanations for face forgery detection via adversarial removal of artifacts View paper
- [59] Gender slopes: Counterfactual fairness for computer vision models by attribute manipulation View paper
- [60] Spatiotemporal Causal Inference With Mechanistic Ecological Models: Evaluating Targeted Culling on Chronic Wasting Disease Dynamics in Cervids View paper
- [61] Named Entity Recognition in Persian Language based on Self-attention Mechanism with Weighted Relational Position Encoding View paper
- [62] Dag: A dual causal network for time series forecasting with exogenous variables View paper
- [63] Causal Intervention Is What Large Language Models Need for Spatio-Temporal Forecasting View paper
- [64] Causal Head Gating: A Framework for Interpreting Roles of Attention Heads in Transformers View paper
- [65] Diffusion Forcing: Next-token Prediction Meets Full-Sequence Diffusion View paper
- [66] Fine-Grained Pavement Performance Prediction Based on Causal-Temporal Graph Convolution Networks View paper
- [67] A Mechanistic Interpretation of Arithmetic Reasoning in Language Models using Causal Mediation Analysis View paper
- [68] Non-markovian discrete diffusion with causal language models View paper
- [69] CAMEF: Causal-augmented multi-modality event-driven financial forecasting by integrating time series patterns and salient macroeconomic announcements View paper
- [70] Using deep autoregressive models as causal inference engines View paper
- [71] Token-Level Uncertainty-Aware Objective for Language Model Post-Training View paper