Integrating Logical Rules into LLM Reasoning for Document-level Relation Extraction

Hye-Yoon Baek (백혜윤)

Graph & Language Intelligence Laboratory
Department of Computer Science and Engineering
Konkuk University

2026.01.08

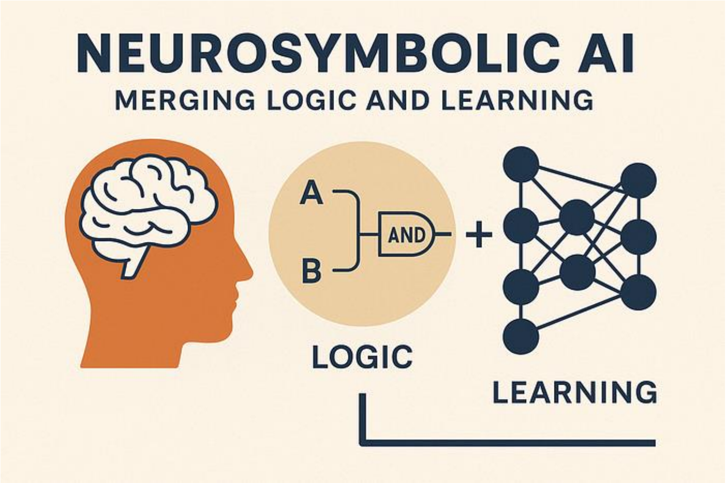
CONTENTS

1. Background
   - Conceptual Background
   - Task Background
2. Related Works
3. Research in Progress
4. Conceptual Insights

Conceptual Background

- "Stacked neural layers is all we need"
  - ✓ 현재 AI 연구의 주된 방향성
  - ✓ 대다수의 모델은 더 깊게, 더 크게 만드는 방식에 집중
  - ✓ 근본적 한계
    - ▪ explainability - 왜 이런 판단을 했는가?
    - ▪ trustworthiness – 결과를 얼마나 신뢰할 수 있는가?
    - ▪ data efficiency – 적은 데이터에서도 잘 동작하는가?
    - ▪ curse of dimensionality – 모델 크기 확장의 한계

- Neuro-symbolic AI emerges as an alternative direction
  - ✓ by combining
    - ▪ data-driven learning (Neural)
    - ▪ explainable, logic-based representations (Symbolic)
  - ✓ leads to
    - ▪ higher generalization capabilities
    - ▪ more structured and controllable reasoning



이미지 출처: https://medium.com/@jainultrivedi55555/neurosymbolic-ai-merging-logic-and-learning-for-smarter-machines-1eea0c5b0806

Conceptual Background

Why Neuro-Symbolic?

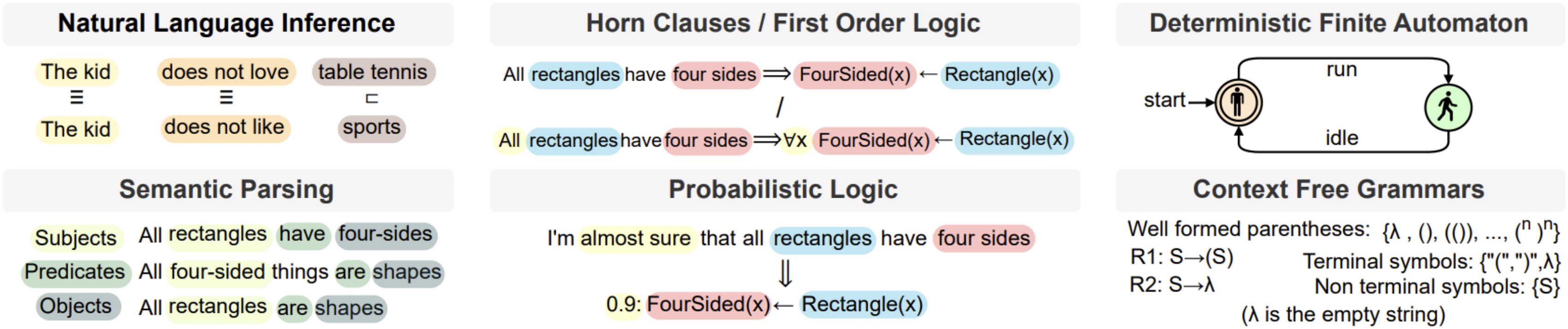
- Symbolic AI
  - ✓ 기호, 논리, 규칙을 사용해서 지식을 명시적으로 표현
  - ✓ 모델의 결정에 대해 명확하고 해석 가능한 설명 가능
  - ✓ 적은 데이터에서도 작동
  - ✓ 유연하지 못하며, 대량의 데이터를 잘 처리하지 못함

- Connectionist AI (Neural Network)
  - ✓ 대량의 데이터로부터 패턴과 관계를 학습하는 데 탁월
  - ✓ NLP, CV 등에서 강력한 성능
  - ✓ 불투명함, 많은 데이터 필요, 논리적 추론에 약함

Conceptual Background

Why Neuro-Symbolic?

- Neuro-Symbolic AI
  - ✓ 두 가지 방식의 강점을 결합 _ 신경망의 학습능력 + 심볼릭 AI의 추론 능력 및 해석가능성
  - ✓ 목표:
    - ▪ 경험으로부터 학습함과 동시에,
    - ▪ 학습한 것에 대해 논리적으로 추론할 수 있는
    - ▪ AI 시스템 구축


Natural Language Inference
The kid ≡ The kid
does not love ≡ does not like
table tennis ⊏ sports
Semantic Parsing
Subjects
Predicates
Objects
All rectangles have four-sides
All four-sided things are shapes
All rectangles are shapes
Horn Clauses / First Order Logic
All rectangles have four sides ⟹ FourSided(x) ← Rectangle(x)
/
All rectangles have four sides ⟹ ∀x FourSided(x) ← Rectangle(x)
Probabilistic Logic
I'm almost sure that all rectangles have four sides
⇓
0.9: FourSided(x) ← Rectangle(x)
Deterministic Finite Automaton
start
run
idle
Context Free Grammars
Well formed parentheses: {λ , (), (()), ..., (ⁿ)ⁿ}
R1: S→(S)
R2: S→λ
Terminal symbols: {"(",")",λ}
Non terminal symbols: {S}
(λ is the empty string)

이미지 출처: Neuro-Symbolic Artificial Intelligence: A Task-Directed Survey in the Black-Box Models Era (IJCAI'25)

Conceptual Background

Reasoning & Explainability

- Two different approaches:
  - ✓ Rule-guided reasoning
    - ▪ rules constrain the inference process
    - ▪ outputs are forced to meet specific criteria
  - ✓ Post-inference forensic
    - ▪ Interpretable reasoning pathways are reconstructed after prediction

Task Background

Sentence-level Relation Extraction

- Identify entity and relations within a single sentence

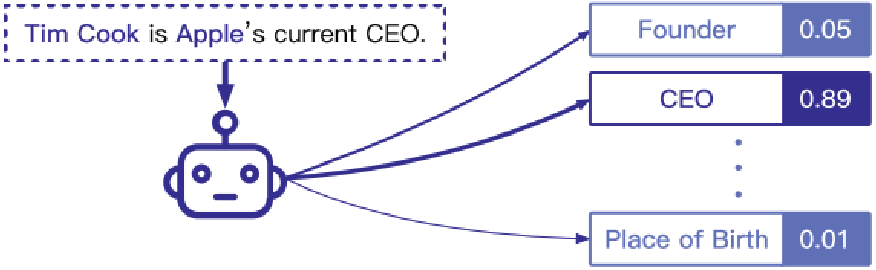


Fig. 4.6 An example of sentence-level relation extraction

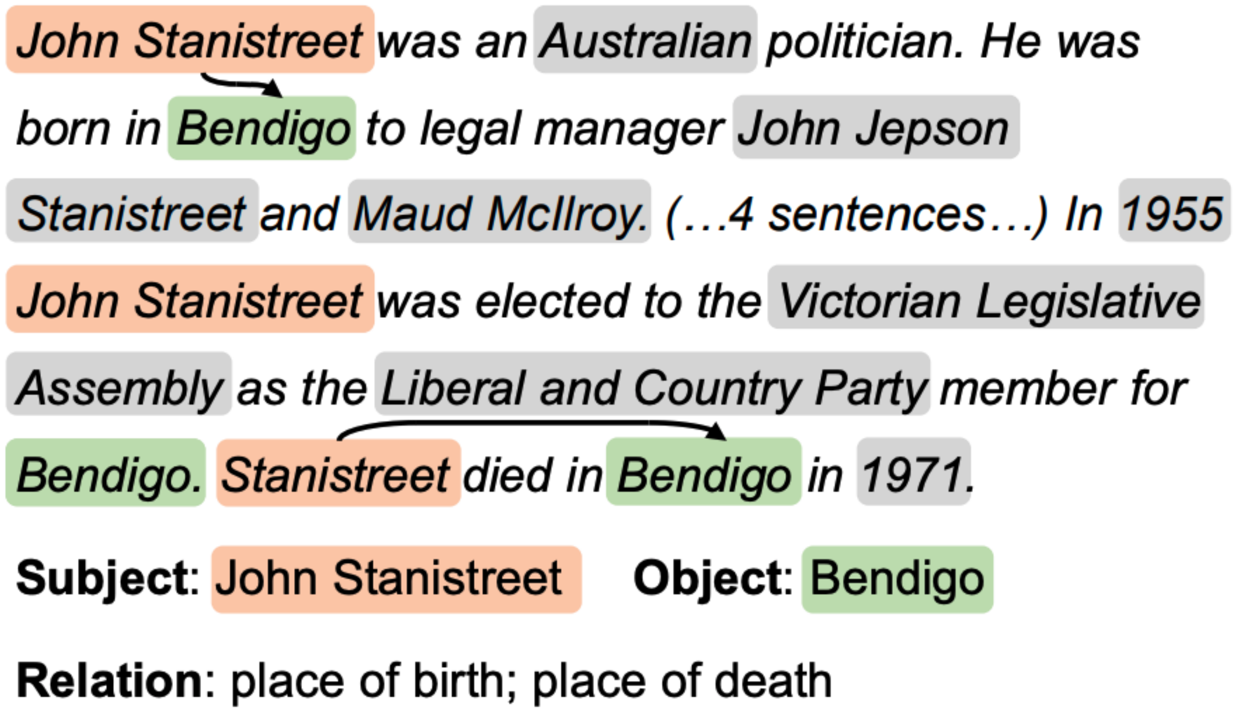
- Limitations
  - ✓ real world application 취약함 _ 문장 내부 정보만 사용
  - ✓ 문장 간에 걸쳐 나타나는 관계는 추출 불가

이미지 출처: https://www.researchgate.net/publication/342684048_Sentence_Representation

Task Background

Document-level Relation Extraction (DocRE)

- Identify all relations between each entity pair
- across an entire document



John Stanistreet was an Australian politician. He was born in Bendigo to legal manager John Jepson Stanistreet and Maud McIlroy. (…4 sentences…) In 1955 John Stanistreet was elected to the Victorian Legislative Assembly as the Liberal and Country Party member for Bendigo. Stanistreet died in Bendigo in 1971.

Subject: John Stanistreet **Object**: Bendigo

Relation: place of birth; place of death

이미지 출처: Document-Level Relation Extraction with Adaptive Thresholding and Localized Context Pooling (AAAI'21)

Task Background

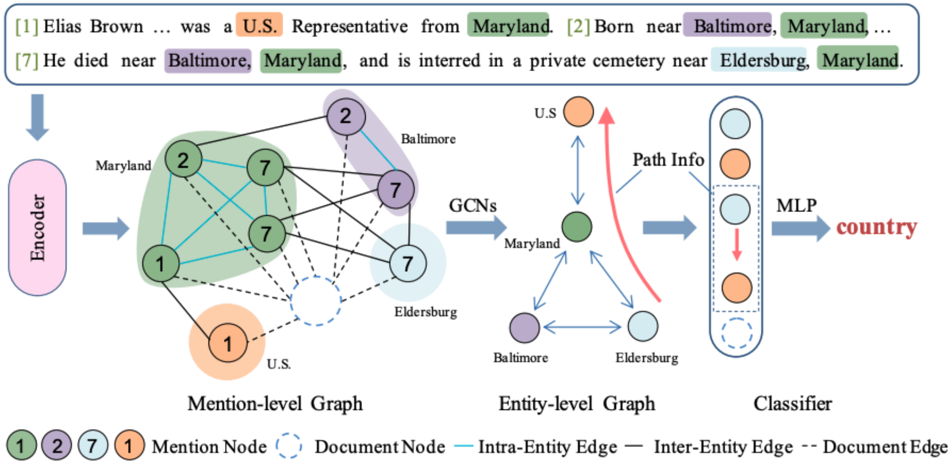
Why is DocRE challenging?

- **Long-context & Cross-sentence Modeling**
  - ✓ 단일 문장을 넘어서는 긴 문맥을 모델링해야 함
  - ✓ 관계에 등장하는 주어, 목적어 엔티티가 서로 다른 문장에 존재할 수 있음
  - ✓ 관계에 대한 단서가 여러 문장에 분산되어 나타나는 경우가 일반적임
- **Entity Representation**
  - ✓ 하나의 엔티티가 문서 전반에 걸쳐 여러 번 반복 등장함
  - ✓ 하나의 엔티티가 다양한 멘션(mentions)이나 별칭(aliases)으로 표현될 수 있음
  - ✓ 문서 수준에서 엔티티를 추적하고 통합적으로 표현해야함
- **Multi-entity & Multi-label structure**
  - ✓ 하나의 문서에는 다수의 엔티티와 엔티티 쌍이 존재함
  - ✓ 동일한 엔티티 쌍이 여러 개의 관계(label)를 동시에 가질 수 있음
- **Implicit & Latent Reasoning**
  - ✓ 관계가 특정 문장에 명시적으로 드러나지 않는 경우가 많음
  - ✓ 문장 수준에서 관찰되지 않는 잠재적, 암묵적 정보를 활용해야 함

Task Background

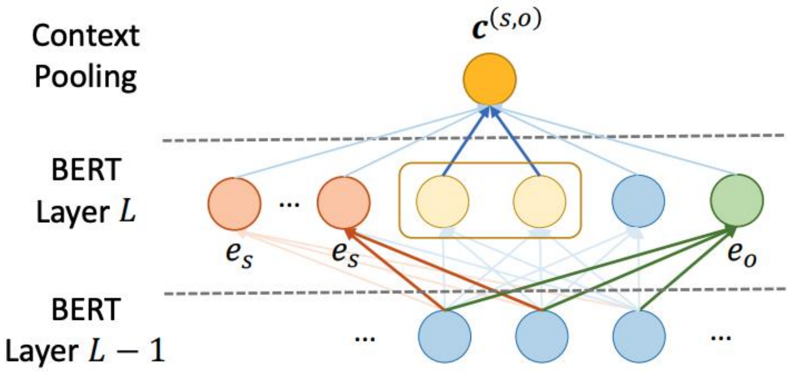
Plain DocRE models

- **Focus on strengthening entity-pair representations**
  - ✓ capture interdependencies between entity pairs
  - ✓ by learning powerful representations through neural models



[1] Elias Brown ... was a U.S. Representative from Maryland. [2] Born near Baltimore, Maryland, ...
[7] He died near Baltimore, Maryland, and is interred in a private cemetery near Eldersburg, Maryland.

[1]



[2]

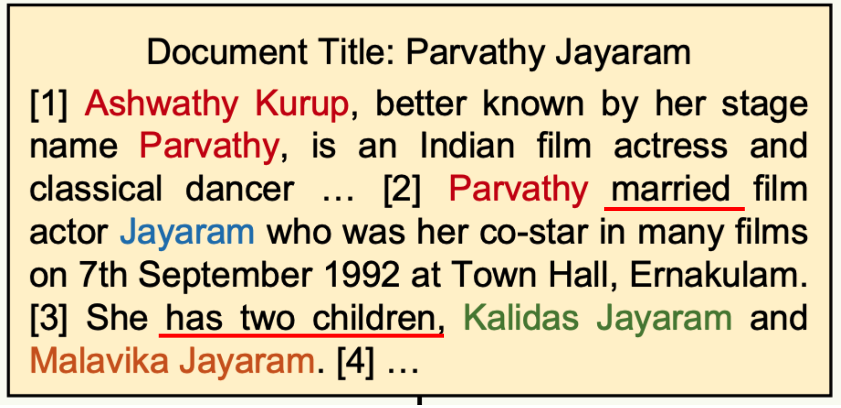
[1] [2020][EMNLP][GAIN] _ graph-based
Double Graph Based Reasoning for Document-level Relation Extraction

[2] [2021][AAAI][ATLOP] _ transformer-based
Document-Level Relation Extraction with Adaptive Thresholding and Localized Context Pooling

Task Background

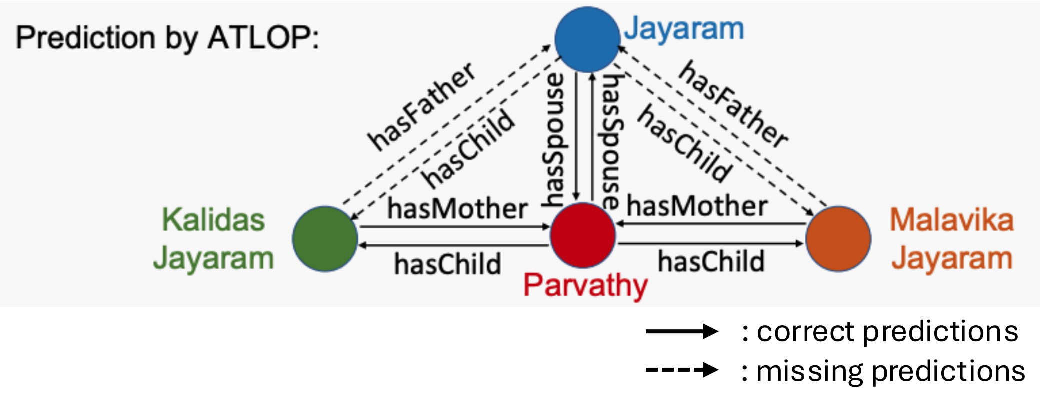
Plain DocRE models

- Limitations)
  - ✓ 강력한 문맥 표현을 통해 명시적으로 드러난 관계(apparent facts) 추출에는 효과적이나,
  - ✓ 텍스트에 직접 나타나지 않은 잠재적 관계(potential/implicit facts)를 도출하는 데 한계



Document Title: Parvathy Jayaram

[1] Ashwathy Kurup, better known by her stage name Parvathy, is an Indian film actress and classical dancer … [2] Parvathy married film actor Jayaram who was her co-star in many films on 7th September 1992 at Town Hall, Ernakulam. [3] She has two children, Kalidas Jayaram and Malavika Jayaram. [4] …

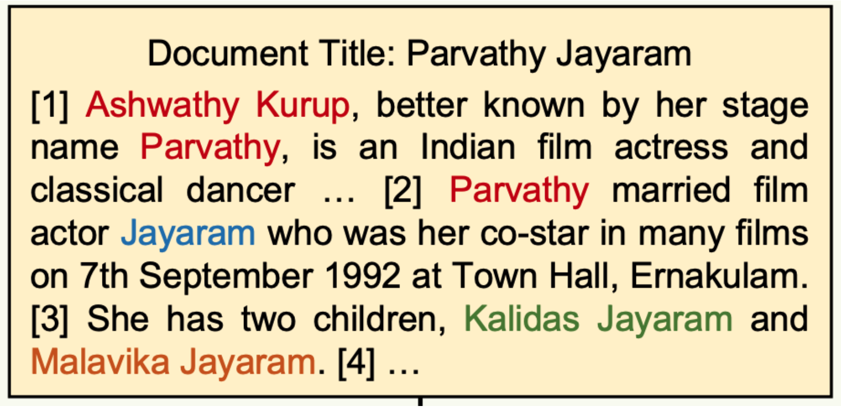


이미지 출처: End-to-end Learning of Logical Rules for Enhancing Document-level Relation Extraction (ACL'24)

Task Background

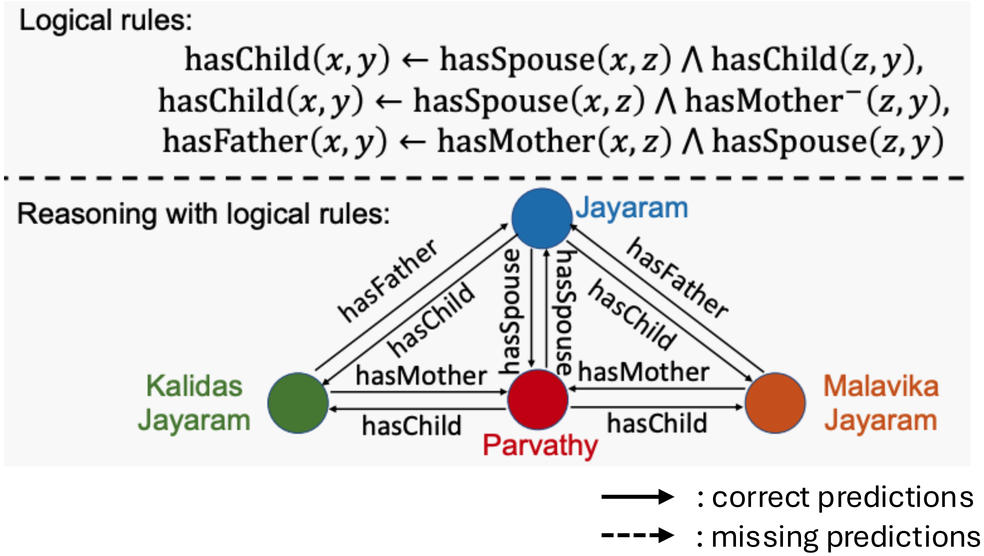
DocRE with Logical Rules

- Motivation
  - ✓ Logical rules를 이용해 기존에 추출된 관계로부터 누락된 사실(missing facts)을 추론함으로써 DocRE 성능을 향상시킬 수 있음



Document Title: Parvathy Jayaram

[1] Ashwathy Kurup, better known by her stage name Parvathy, is an Indian film actress and classical dancer … [2] Parvathy married film actor Jayaram who was her co-star in many films on 7th September 1992 at Town Hall, Ernakulam. [3] She has two children, Kalidas Jayaram and Malavika Jayaram. [4] …



Logical rules:

$$\text{hasChild}(x,y) \leftarrow \text{hasSpouse}(x,z) \wedge \text{hasChild}(z,y),$$
$$\text{hasChild}(x,y) \leftarrow \text{hasSpouse}(x,z) \wedge \text{hasMother}^{-}(z,y),$$
$$\text{hasFather}(x,y) \leftarrow \text{hasMother}(x,z) \wedge \text{hasSpouse}(z,y)$$

Reasoning with logical rules:

이미지 출처: End-to-end Learning of Logical Rules for Enhancing Document-level Relation Extraction (ACL'24)

Task Background

DocRE with Logical Rules

- Why Logical Rules?
  - ✓ 관계 간의 논리적 결합을 신경망이 암묵적으로 학습해야하는 부담을 완화함
  - ✓ 관계 간 내재적 상관관계(intrinsic correlations)를 명시적으로 표현할 수 있어 결과를 보다 해석 가능하게 설명할 수 있음

  => 장점: 잠재적 관계에 대한 추론 근거 제시, blackbox 모델 대비 추론 과정의 투명성 향상

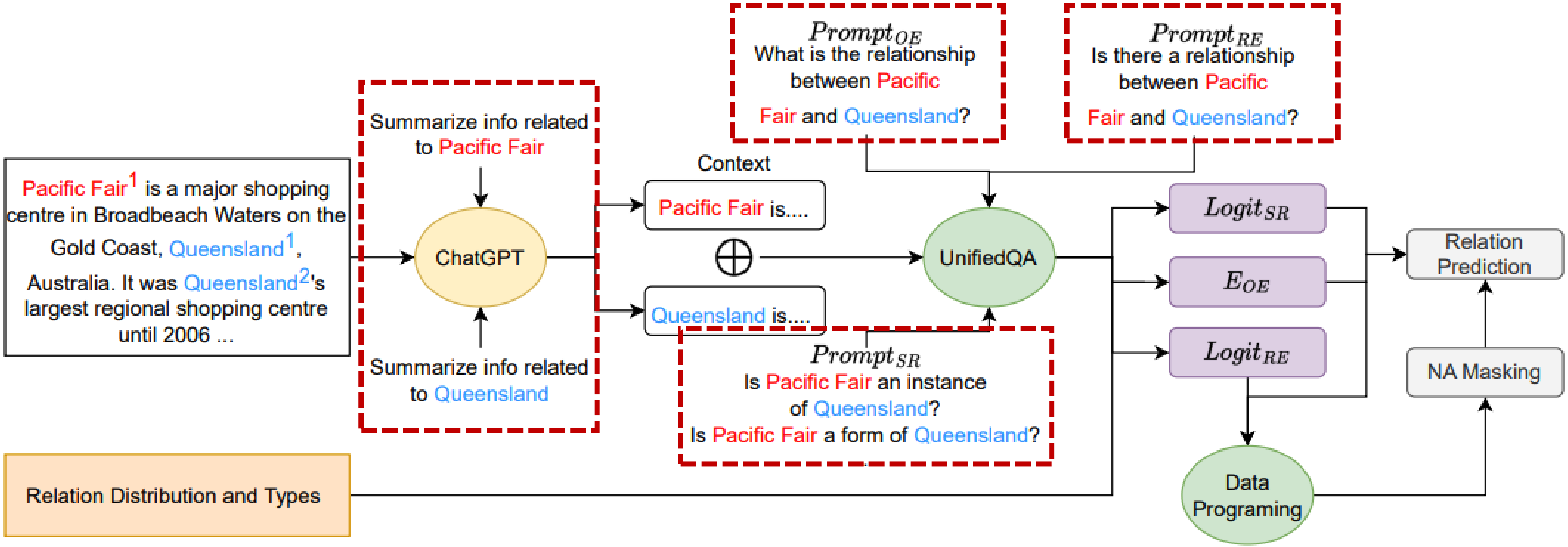
- 대표 논문
  - ✓ [2024][ACL][JMRL] _ neural module
    End-to-end Learning of Logical Rules for Enhancing Document-level Relation Extraction
  - ✓ [2025][ICLR][RuAG] _ LLM
    Learned-Rule-Augmented Generation for Large Language Models

Task Background

DocRE with LLMs

- PromptRE (arXiv'23)
  - ✓ 여러 종류의 프롬프트를 통해 약한 판단 신호들을 만들고
  - ✓ 이를 data programming으로 결합해 pseudo-label을 생성함

=> LLM은 프롬프팅에 기반해 판단 신호 생성 을 위해 사용됨



Pacific Fair[1] is a major shopping centre in Broadbeach Waters on the Gold Coast, Queensland[1], Australia. It was Queensland[2]'s largest regional shopping centre until 2006 ...
Summarize info related to Pacific Fair
ChatGPT
Summarize info related to Queensland
Context
Pacific Fair is....
Queensland is....
$\oplus$
$Prompt_{OE}$
What is the relationship between Pacific Fair and Queensland?
$Prompt_{RE}$
Is there a relationship between Pacific Fair and Queensland?
UnifiedQA
$Prompt_{SR}$
Is Pacific Fair an instance of Queensland?
Is Pacific Fair a form of Queensland?
$Logit_{SR}$
E_{OE}
$Logit_{RE}$
Relation Prediction
NA Masking
Data Programing
Relation Distribution and Types

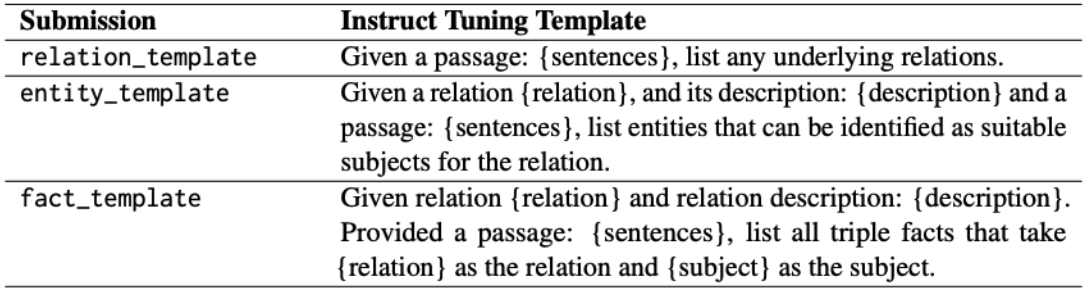
[1] PromptRE: Weakly-Supervised Document-Level Relation Extraction via Prompting-Based Data Programming (arXiv'23)

Task Background

DocRE with LLMs

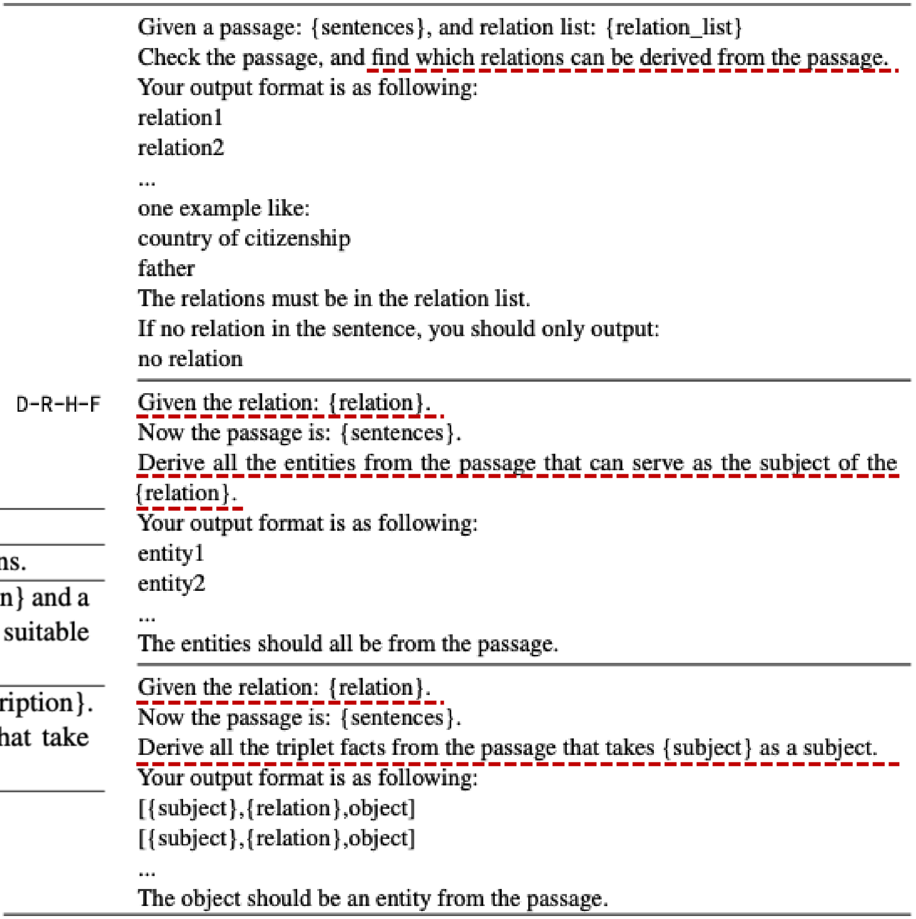
- AutoRE (ACL'24)
  - ✓ RHF(Relation-Head-Facts) 패러다임 제안
    -> DocRE를 단계적 subtask로 분해
  - ✓ Instruction Tuning (QLoRA)

  => LLM을 DocRE 추출 파이프라인으로 fine-tuning



Submission	Instruct Tuning Template
`relation_template`	Given a passage: {sentences}, list any underlying relations.
`entity_template`	Given a relation {relation}, and its description: {description} and a passage: {sentences}, list entities that can be identified as suitable subjects for the relation.
`fact_template`	Given relation {relation} and relation description: {description}. Provided a passage: {sentences}, list all triple facts that take {relation} as the relation and {subject} as the subject.

instruction tuning template for RHF



Given a passage: {sentences}, and relation list: {relation_list}
Check the passage, and find which relations can be derived from the passage.
Your output format is as following:
relation1
relation2
...
one example like:
country of citizenship
father
The relations must be in the relation list.
If no relation in the sentence, you should only output:
no relation

D-R-H-F

Given the relation: {relation}.
Now the passage is: {sentences}.
Derive all the entities from the passage that can serve as the subject of the {relation}.
Your output format is as following:
entity1
entity2
...
The entities should all be from the passage.

Given the relation: {relation}.
Now the passage is: {sentences}.
Derive all the triplet facts from the passage that takes {subject} as a subject.
Your output format is as following:
[{subject},{relation},object]
[{subject},{relation},object]
...
The object should be an entity from the passage.

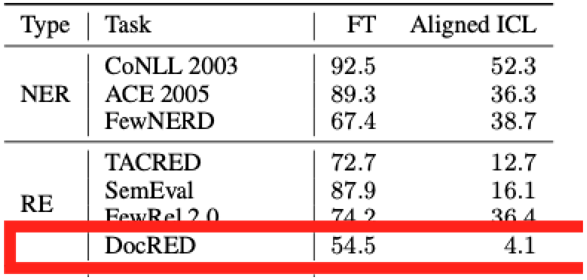
prompt template for RE

[1] AutoRE: Document-Level Relation Extraction with Large Language Models (ACL'24)

Task Background

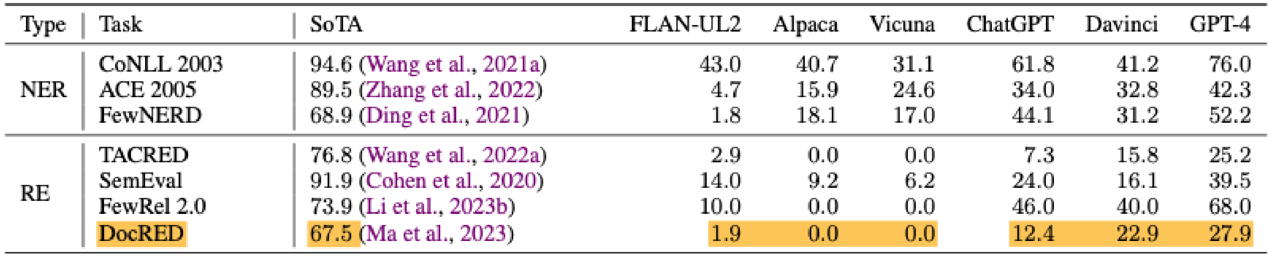
DocRE with LLMs

- When Does In-Context Learning Fall Short and Why? A Study on Specification-Heavy Tasks (arXiv'23)
  - ✓ LLM의 ICL이 specification-heavy tasks(DocRE 포함)에서 부족한 이유
    - ▪ 구체적인 컨텍스트 이해 능력 부족
    - ▪ 인간과의 태스크 스키마 이해 불일치 (인풋 크기 제한으로 인해 underspecified 됨)
    - ▪ Long-context 이해 능력 부족
      - – DocRE와 같은 긴 컨텍스트를 특징으로 하는 태스크에서 LLM의 성능이 특히 저조
  - ✓ DocRE와 같은 specification-heavy task에서는 <u>in-context learning 만으로는 효과적인 성능을 내기 어려움</u>



Type	Task	FT	Aligned ICL
NER	CoNLL 2003	92.5	52.3
	ACE 2005	89.3	36.3
	FewNERD	67.4	38.7
RE	TACRED	72.7	12.7
	SemEval	87.9	16.1
	FewRel 2.0	74.2	36.4
	DocRED	54.5	4.1

fine-tuning and zero-shot ICL after alignment



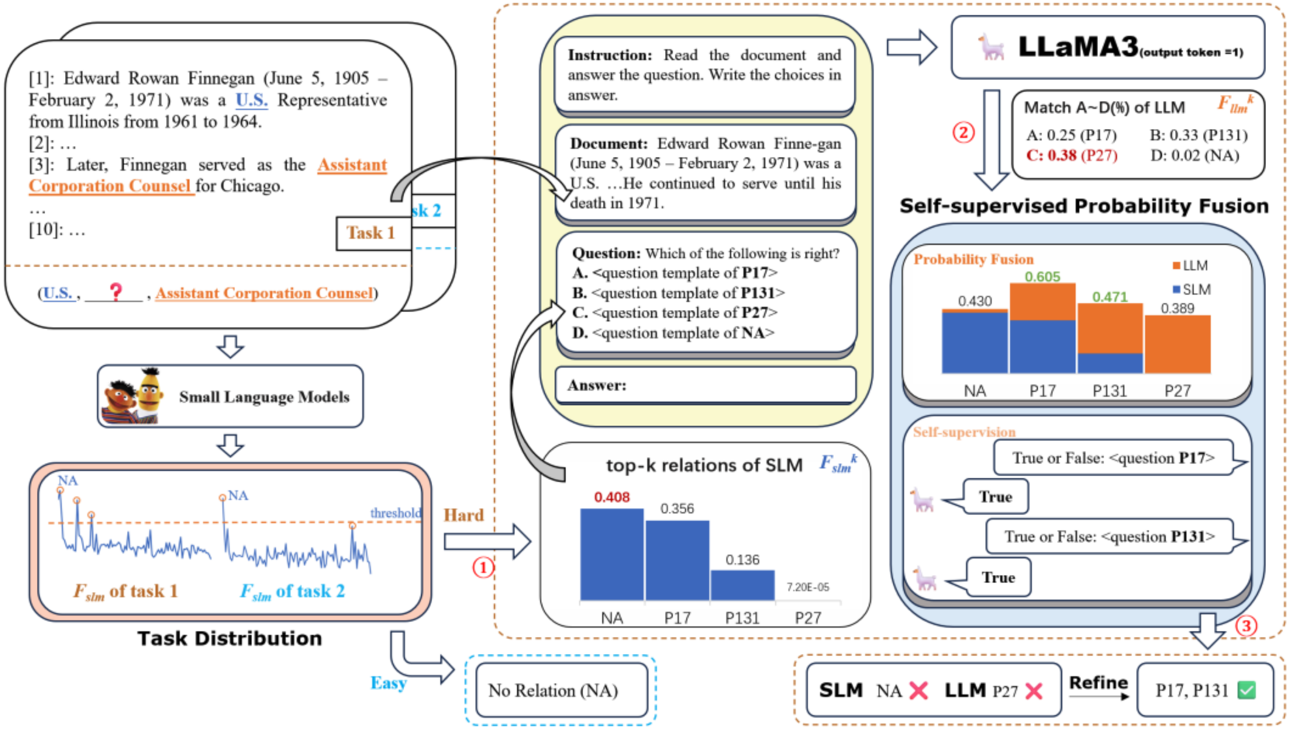
Type	Task	SoTA	FLAN-UL2	Alpaca	Vicuna	ChatGPT	Davinci	GPT-4
NER	CoNLL 2003	94.6 (Wang et al., 2021a)	43.0	40.7	31.1	61.8	41.2	76.0
	ACE 2005	89.5 (Zhang et al., 2022)	4.7	15.9	24.6	34.0	32.8	42.3
	FewNERD	68.9 (Ding et al., 2021)	1.8	18.1	17.0	44.1	31.2	52.2
RE	TACRED	76.8 (Wang et al., 2022a)	2.9	0.0	0.0	7.3	15.8	25.2
	SemEval	91.9 (Cohen et al., 2020)	14.0	9.2	6.2	24.0	16.1	39.5
	FewRel 2.0	73.9 (Li et al., 2023b)	10.0	0.0	0.0	46.0	40.0	68.0
	DocRED	67.5 (Ma et al., 2023)	1.9	0.0	0.0	12.4	22.9	27.9

ICL performance

Task Background

DocRE with LLMs

- Drell (NAACL'25)
  - ✓ LLM을 refiner로 사용하여,
  - ✓ 먼저 SLM(기존 모델)으로 문서 내 엔티티 쌍에 대해 예측을 수행하고, easy/hard 케이스를 분류한 뒤 hard 케이스에 대해 LLM이 refinement를 수행


[1]: Edward Rowan Finnegan (June 5, 1905 – February 2, 1971) was a U.S. Representative from Illinois from 1961 to 1964.
[2]: …
[3]: Later, Finnegan served as the Assistant Corporation Counsel for Chicago.
…
[10]: …

(U.S. , ? , Assistant Corporation Counsel)

Instruction: Read the document and answer the question. Write the choices in answer.

Document: Edward Rowan Finne-gan (June 5, 1905 – February 2, 1971) was a U.S. …He continued to serve until his death in 1971.

Question: Which of the following is right?
A. <question template of P17>
B. <question template of P131>
C. <question template of P27>
D. <question template of NA>

Answer:

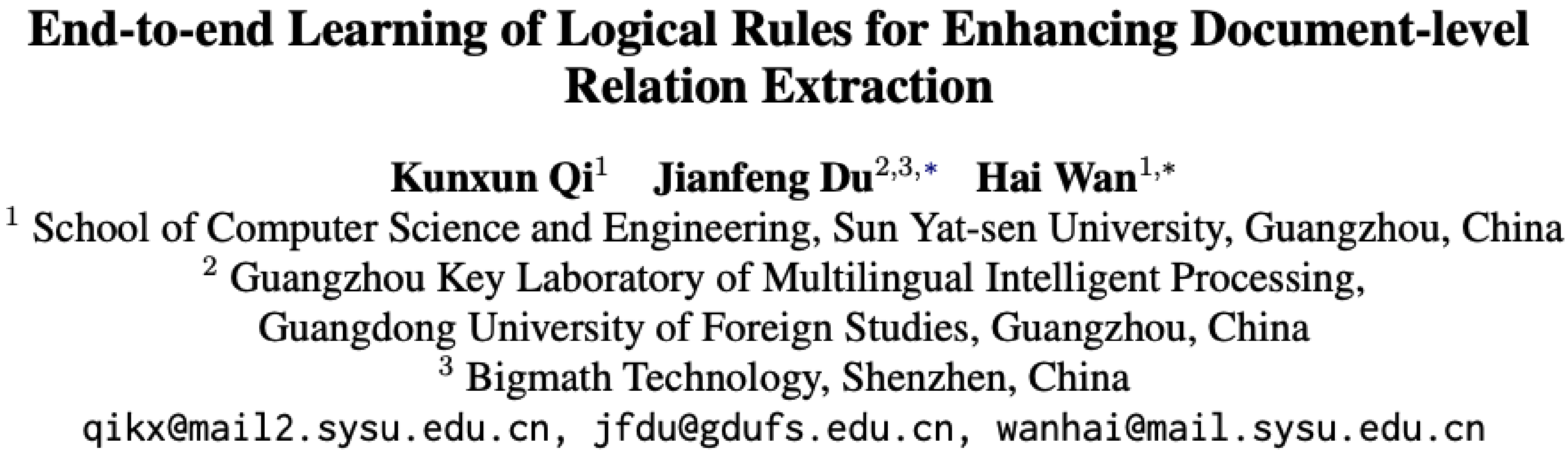
True or False: <question P17>
True
True or False: <question P131>
True

[1] Rethinking the Role of LLMs for Document-level Relation Extraction: a Refiner with Task Distribution and Probability Fusion (NAACL'25)

CONTENTS

1. Background
2. Related Works
    - JMRL (ACL'24)
    - RuAG (ICLR'25)
3. Research in Progress
4. Conceptual Insights

[2024][ACL][JMRL]



End-to-end Learning of Logical Rules for Enhancing Document-level Relation Extraction

Kunxun Qi[1] Jianfeng Du[2,3,*] Hai Wan[1,*]
[1] School of Computer Science and Engineering, Sun Yat-sen University, Guangzhou, China
[2] Guangzhou Key Laboratory of Multilingual Intelligent Processing,
Guangdong University of Foreign Studies, Guangzhou, China
[3] Bigmath Technology, Shenzhen, China
qikx@mail2.sysu.edu.cn, jfdu@gdufs.edu.cn, wanhai@mail.sysu.edu.cn

[2024][ACL][JMRL]

Key Idea

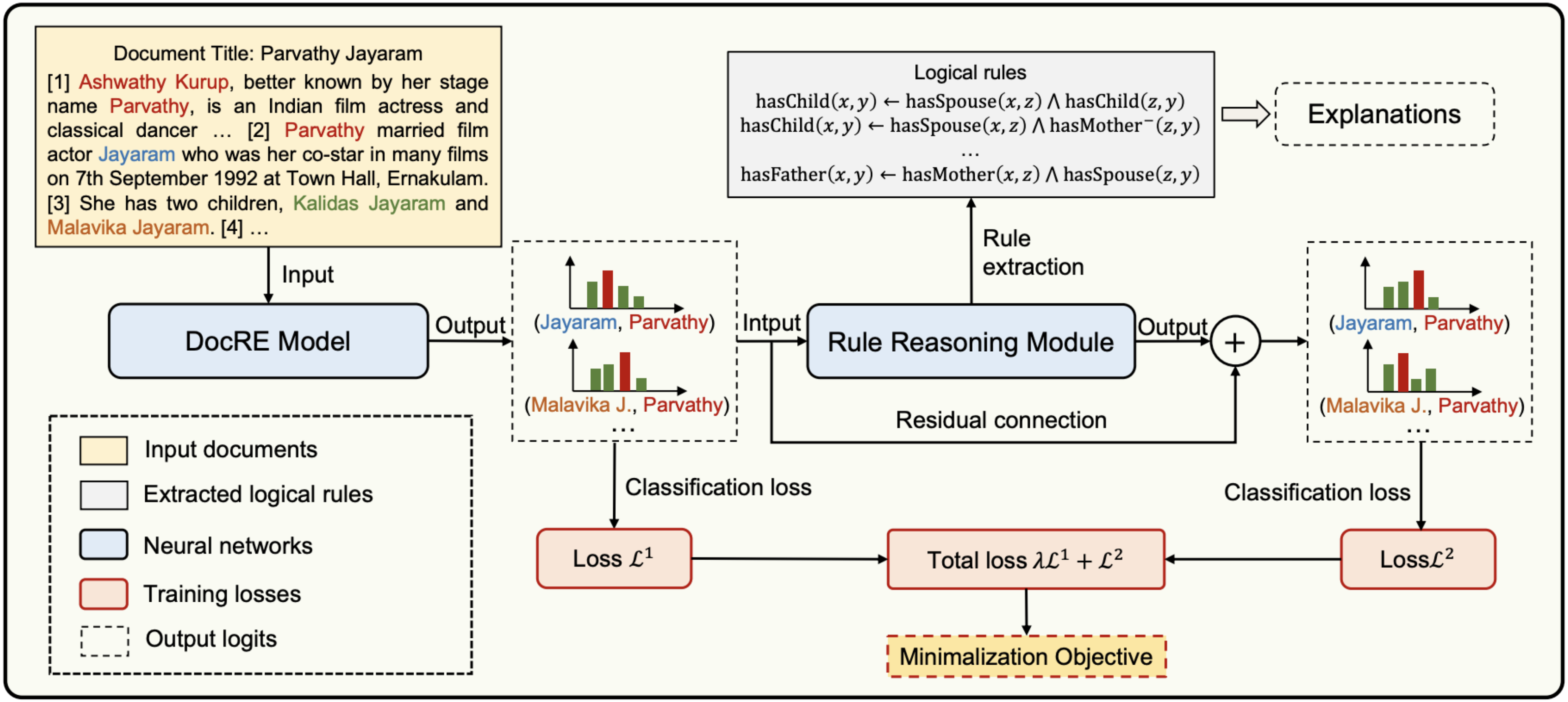
_ End-to-end Learning of Logical Rules for Enhancing Document-level Relation Extraction

- Joint Modeling Relation extraction and Logical rules
  - ✓ jointly learning
    - ▪ a neural model for DocRE
    - ▪ and a neural model for approximating logical rules
  - ✓ in an end-to-end fashion
- Motivation
  - ✓ 기존의 logical rule based DocRE 모델[1, 2] 은 pipeline 구조로, error propagation 한계를 가짐
- Contribution
  - ✓ first end-to-end approach for imposing logical rules upon DocRE models

[1] Learning Logic Rules for Document-level Relation Extraction (EMNLP'21)
[2] Boosting Document-Level Relation Extraction by Mining and Injecting Logical Rules (EMNLP'22)

[2024][ACL][JMRL]

Overall Architecture



Document Title: Parvathy Jayaram

[1] Ashwathy Kurup, better known by her stage name Parvathy, is an Indian film actress and classical dancer … [2] Parvathy married film actor Jayaram who was her co-star in many films on 7th September 1992 at Town Hall, Ernakulam. [3] She has two children, Kalidas Jayaram and Malavika Jayaram. [4] …

Logical rules

$\mathrm{hasChild}(x, y) \leftarrow \mathrm{hasSpouse}(x, z) \wedge \mathrm{hasChild}(z, y)$

$\mathrm{hasChild}(x, y) \leftarrow \mathrm{hasSpouse}(x, z) \wedge \mathrm{hasMother}^{-}(z, y)$

…

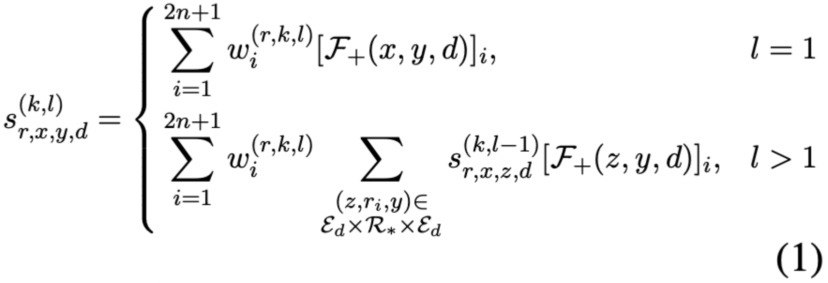
$\mathrm{hasFather}(x, y) \leftarrow \mathrm{hasMother}(x, z) \wedge \mathrm{hasSpouse}(z, y)$

- 🟨 Input documents
- ⬜ Extracted logical rules
- 🟦 Neural networks
- 🟥 Training losses
- ⬚ Output logits

[2024][ACL][JMRL]

Rule reasoning module

- 각 관계 r에 대해 최대 N개의 규칙, 규칙당 길이 L만큼의 body를 chain 형태로 학습
- 목표: 각 문서 d의 모든 팩트 (x,r,y)에 대해, 길이 L의 chain rule N개로 그 사실이 추론될 수 있는 정도를 나타내는 truth degree $s_{r,x,y,d}^{(N,L)}$ 를 추정함
- Intermediate estimated truth degree



$$s_{r,x,y,d}^{(k,l)} = \begin{cases} \sum_{i=1}^{2n+1} w_i^{(r,k,l)} [\mathcal{F}_+(x,y,d)]_i, & l = 1 \\ \sum_{i=1}^{2n+1} w_i^{(r,k,l)} \sum_{\substack{(z,r_i,y)\in \\ \mathcal{E}_d \times \mathcal{R}_* \times \mathcal{E}_d}} s_{r,x,z,d}^{(k,l-1)} [\mathcal{F}_+(z,y,d)]_i, & l > 1 \end{cases} \quad (1)$$

- 각 규칙의 body atom은 trainable weight $w_{r,k,l}$로 표현
  - ✓ weight on predicate selection for [head atom이 r인 k번째 rule의, l번째 body atom]
  - ✓ confined to [0,1] by a softmax layer
  - ✓ $w_{r,k,l}$가 1이면, i번째 relation r_i가 selected as the predicate of the l번째 body atom

[2024][ACL][JMRL]

Rule reasoning module

- $\alpha_r^{(k)}$: r의 k번째 규칙에 대한 confidence score _ trainable weight
- 규칙 집계 _ 최종 truth degree of N개의 규칙

$$s_{r,x,y,d}^{(N,L)} = \sum_{k=1}^{N} \alpha_r^{(k)} s_{r,x,y,d}^{(k,L)} \qquad (2)$$

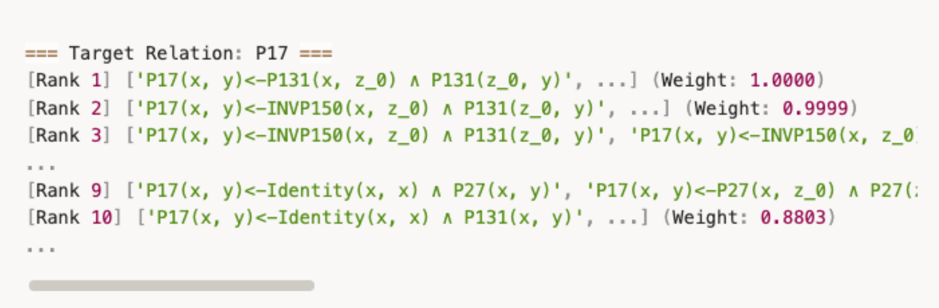
- Residual connection for final truth degree

$$\phi_r^{(x,y,d)} = [\mathcal{F}(x,y,d)]_r + s_{r,x,y,d}^{(N,L)} \qquad (3)$$

[2024][ACL][JMRL]

Rule Extraction

- 학습된 soft rule 파라미터를 사람이 읽을 수 있는 symbolic rule로 변환 및 정렬
- 유동적으로 유효한 길이만큼의 룰을 뽑는 게 아니라 T 값을 지정하면 모든 타겟 관계에 대해 T 길이만큼의 룰을 추출해냄. 즉 노이즈가 많아짐
- 고정된 길이로 출력하기 때문에, 모든 target relation에 대해 T 길이만큼의 body chain을 강제로 생성함
  - ✓ 짧은 길이만으로 충분한 규칙도 불필요한 relation들이 붙음
  - ✓ 유효하지 않은 relation hop까지 포함되어 노이즈가 많은 규칙이 만들어짐

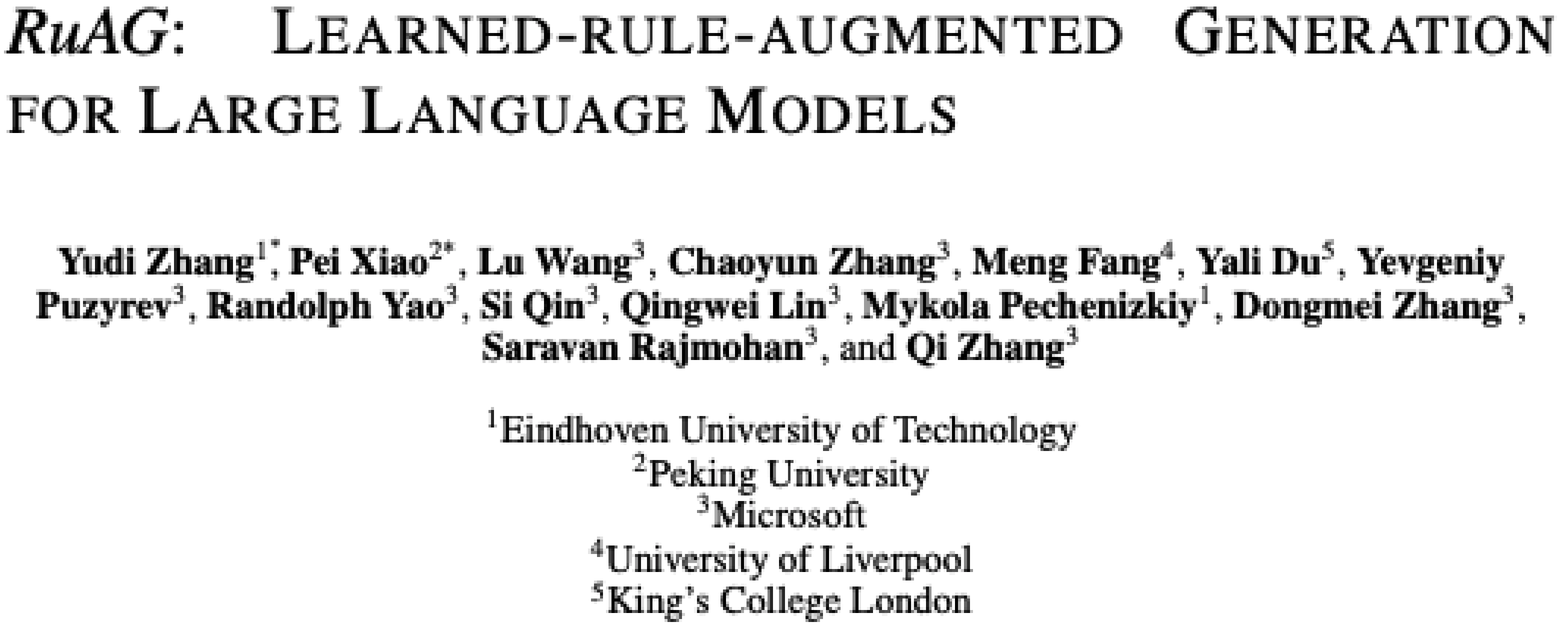


```
=== Target Relation: P17 ===
[Rank 1] ['P17(x, y)<-P131(x, z_0) ∧ P131(z_0, y)', ...] (Weight: 1.0000)
[Rank 2] ['P17(x, y)<-INVP150(x, z_0) ∧ P131(z_0, y)', ...] (Weight: 0.9999)
[Rank 3] ['P17(x, y)<-INVP150(x, z_0) ∧ P131(z_0, y)', 'P17(x, y)<-INVP150(x, z_0
...
[Rank 9] ['P17(x, y)<-Identity(x, x) ∧ P27(x, y)', 'P17(x, y)<-P27(x, z_0) ∧ P27(:
[Rank 10] ['P17(x, y)<-Identity(x, x) ∧ P131(x, y)', ...] (Weight: 0.8803)
...
```

[Rank 1] [규칙 길이 2
'P17(x, y)<-P131(x, z_0) ∧ P131(z_0, y)',
'P17(x, y)<-P131(x, z_0) ∧ P1376(z_0, y)',
'P17(x, y)<-P131(x, z_0) ∧ INVP36(z_0, y)',
'P17(x, y)<-P131(x, z_0) ∧ P27(z_0, y)',
'P17(x, y)<-P131(x, z_0) ∧ P37(z_0, y)'
] (Weight: 1.0000)

[Rank 1] [규칙 길이 3
'P17(x, y)<-P131(x, z_0) ∧ INVP1376(z_0, z_1) ∧ P131(z_1, y)',
'P17(x, y)<-P131(x, z_0) ∧ P36(z_0, z_1) ∧ P131(z_1, y)',
'P17(x, y)<-P131(x, z_0) ∧ INVP27(z_0, z_1) ∧ P131(z_1, y)',
'P17(x, y)<-P131(x, z_0) ∧ INVP495(z_0, z_1) ∧ P131(z_1, y)',
'P17(x, y)<-P131(x, z_0) ∧ P527(z_0, z_1) ∧ P131(z_1, y)'
] (Weight: 1.0000)

[2025][ICLR][RuAG]



RuAG: LEARNED-RULE-AUGMENTED GENERATION FOR LARGE LANGUAGE MODELS

Yudi Zhang[1*], **Pei Xiao**[2*], **Lu Wang**[3], **Chaoyun Zhang**[3], **Meng Fang**[4], **Yali Du**[5], **Yevgeniy Puzyrev**[3], **Randolph Yao**[3], **Si Qin**[3], **Qingwei Lin**[3], **Mykola Pechenizkiy**[1], **Dongmei Zhang**[3], **Saravan Rajmohan**[3], and **Qi Zhang**[3]

[1]Eindhoven University of Technology
[2]Peking University
[3]Microsoft
[4]University of Liverpool
[5]King's College London

[2025][ICLR][RuAG]

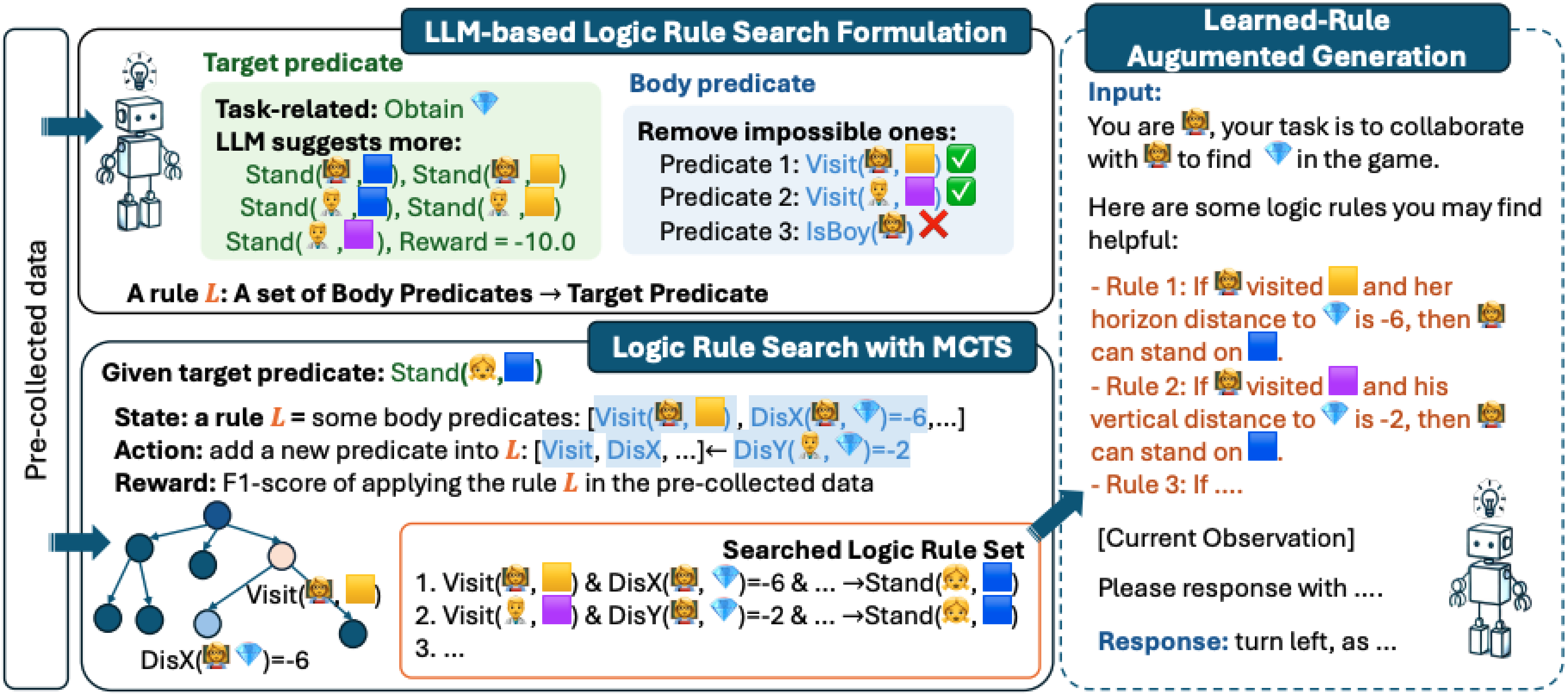
Motivation

- LLM의 도메인 지식/외부 정보 기반한 추론을 위한 외부 지식 주입 기법들(SFT, ICL, RAG, KG) 모두 한계를 가짐
- LLM의 이해력과 추론 능력 향상을 위한 새로운 Knowledge transformation으로 logic rules를 제안함
    - ✓ 기존 방법에 비해 복잡한 관계를 간결하고 구조화된 문장으로 표현할 수 있고 자연어로 쉽게 번역되어 LLM의 프롬프트에 삽입 가능하며 결과적으로 LLM의 이해와 추론 효율을 향상시킴
    - ✓ 즉, logic rule은 자연어로 표현 가능하고 LLM의 학습방식과 잘 정렬되며 복잡한 지식을 간결하게 전달하는 수단이 됨
- Contribution
    - ✓ SFT, ICL, RAG, KG-based 접근을 대체할 수 있는 방법 제안, 외부 지식을 해석 가능한 논리 규칙으로 압축하여 사용함
    - ✓ MCTS 적용을 통한 탐색 효율화
    - ✓ 다양한 태스크(DocRE, log anomaly detection, decision-making)에서의 실험 검증

[2025][ICLR][RuAG]

Framework

- LLM formulates the MCTS search by defining the target and body predicates
- Apply MCTS to generate structured first-order logic rules
- Learned-rule augmented generation



Pre-collected data

LLM-based Logic Rule Search Formulation

Target predicate

Task-related: Obtain 💎
LLM suggests more:
Stand(👩, 🟦), Stand(👩, 🟨)
Stand(👨, 🟦), Stand(👨, 🟨)
Stand(👨, 🟪), Reward = -10.0

Body predicate

Remove impossible ones:
Predicate 1: Visit(👩, 🟨) ✅
Predicate 2: Visit(👨, 🟪) ✅
Predicate 3: IsBoy(👩) ❌

A rule *L*: A set of Body Predicates → Target Predicate

Logic Rule Search with MCTS

Given target predicate: Stand(👧, 🟦)

State: a rule *L* = some body predicates: [Visit(👩, 🟨), DisX(👩, 💎)=-6,...]
Action: add a new predicate into *L*: [Visit, DisX, ...]← DisY(👨, 💎)=-2
Reward: F1-score of applying the rule *L* in the pre-collected data

Visit(👩, 🟨)

DisX(👩 💎)=-6

Searched Logic Rule Set
1. Visit(👩, 🟨) & DisX(👩, 💎)=-6 & ... →Stand(👧, 🟦)
2. Visit(👨, 🟪) & DisY(👩, 💎)=-2 & ... →Stand(👧, 🟦)
3. ...

Learned-Rule Augmented Generation

Input:
You are 👩, your task is to collaborate with 👩 to find 💎 in the game.

Here are some logic rules you may find helpful:

- Rule 1: If 👩 visited 🟨 and her horizon distance to 💎 is -6, then 👩 can stand on 🟦.
- Rule 2: If 👩 visited 🟪 and his vertical distance to 💎 is -2, then 👩 can stand on 🟦.
- Rule 3: If ....

[Current Observation]

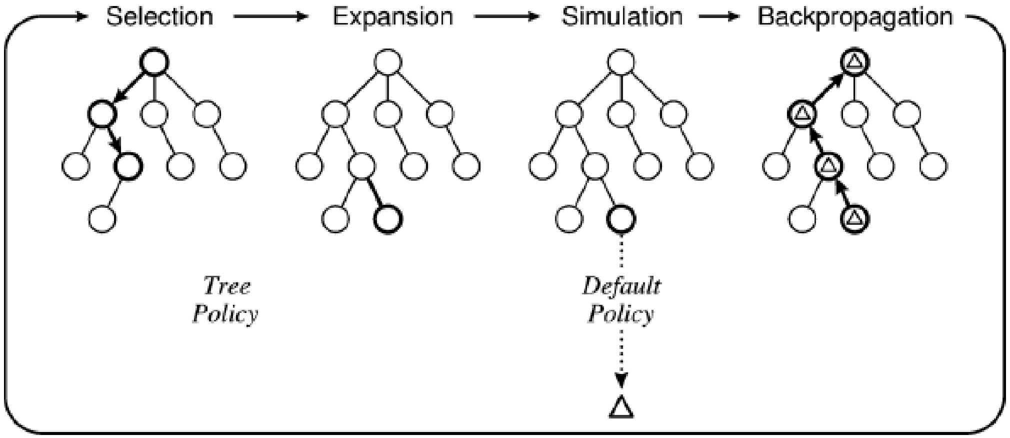
Please response with ....

Response: turn left, as ...

[2025][ICLR][RuAG]

Rule search process

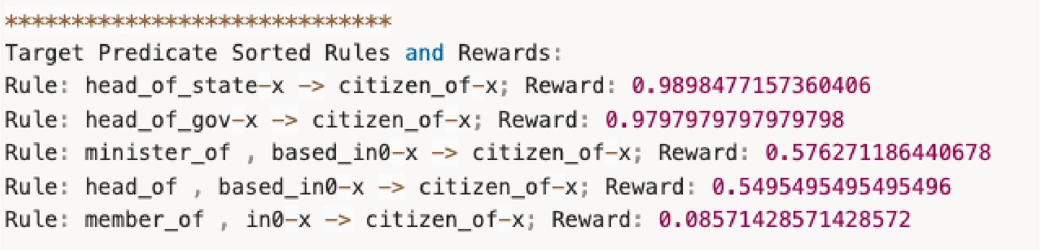
- Monte Carlo Tree Search
  - ✓ Node: 하나의 rule _ child로 내려갈 수록 body predicate을 하나씩 추가하면서 규칙이 길어짐
    - ▪ Root Node: [ ] (규칙 없음)
    - ▪ Child: [born_in]
    - ▪ Child: [born_in, located_in]
  - ✓ Leaf node_ is_terminal() 조건: 규칙 길이가 2 이상이 되면 종료
    - ▪ [simulation] Leaf에 도달하면 규칙을 평가 -> reward(precision) 반환
  - ✓ MCTS가 rollout을 충분히 반복한 뒤, 가장 보상이 큰 rule path를 선택



[2025][ICLR][RuAG]

Rule search process

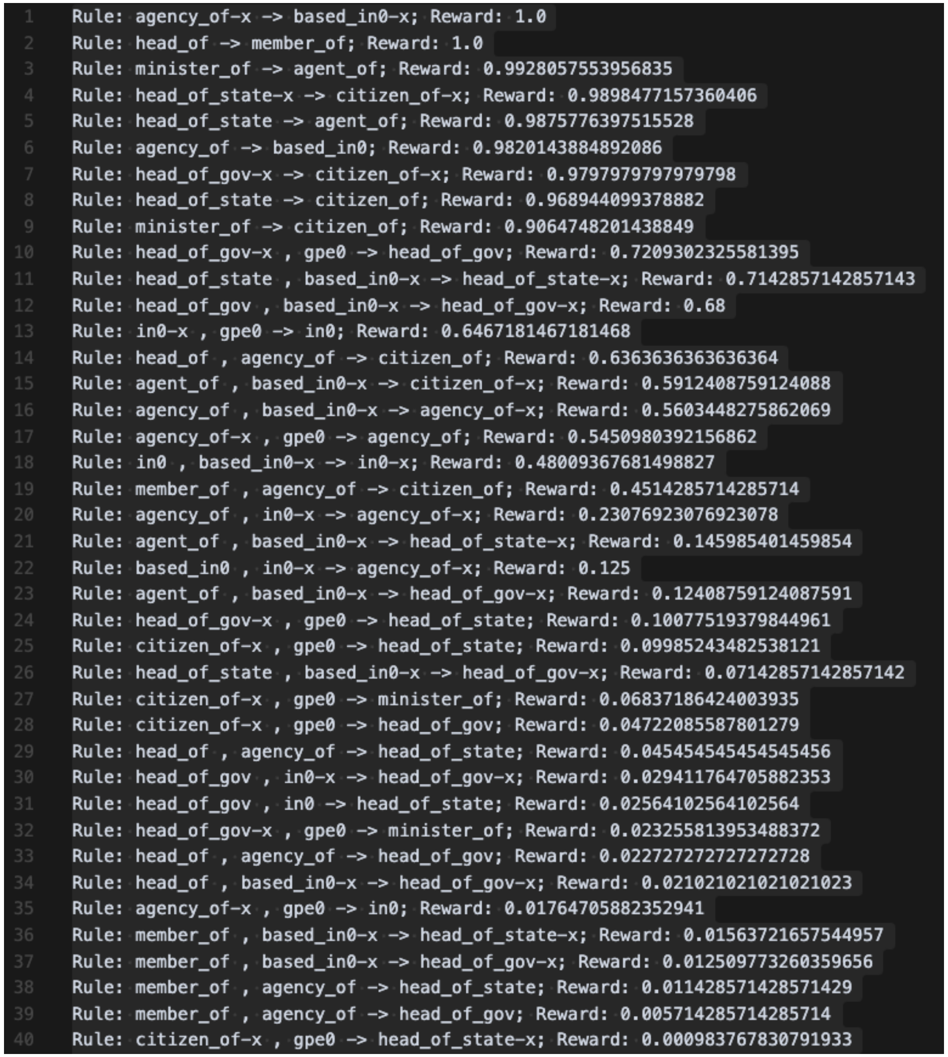
- MCTS를 통해 하나의 target relation에 대해 유효한 규칙들을 찾음
  - ✓ Reward(precision) : precision = correct_predictions / total_predictions
    - ▪ total_predictions: 규칙에 따라 예측된 트리플 수
    - ▪ correct_predictions: 규칙에 따라 예측된 트리플 중 실제 정답인 트리플 수
- Ex) target relation: citizen_of-x



```
*****************************
Target Predicate Sorted Rules and Rewards:
Rule: head_of_state-x -> citizen_of-x; Reward: 0.9898477157360406
Rule: head_of_gov-x -> citizen_of-x; Reward: 0.9797979797979798
Rule: minister_of , based_in0-x -> citizen_of-x; Reward: 0.576271186440678
Rule: head_of , based_in0-x -> citizen_of-x; Reward: 0.5495495495495496
Rule: member_of , in0-x -> citizen_of-x; Reward: 0.08571428571428572
```

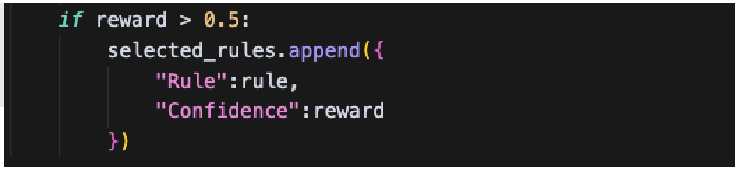
Learned Rule

전체 규칙

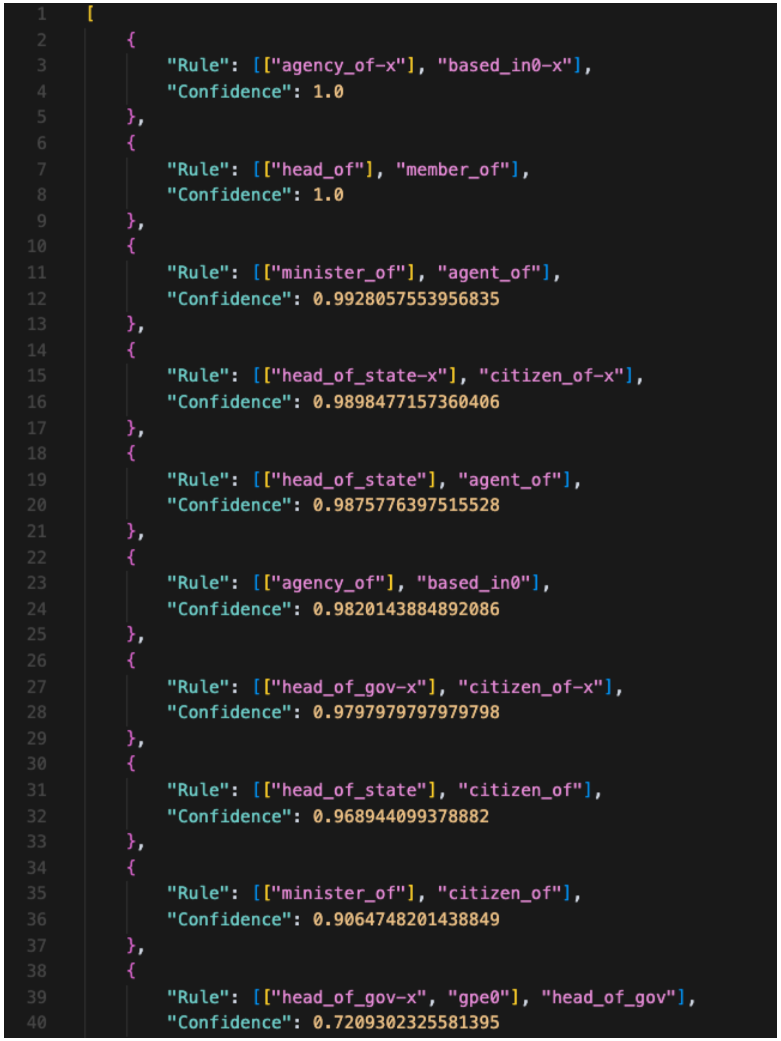


```
Rule: agency_of-x -> based_in0-x; Reward: 1.0
Rule: head_of -> member_of; Reward: 1.0
Rule: minister_of -> agent_of; Reward: 0.9928057553956835
Rule: head_of_state-x -> citizen_of-x; Reward: 0.9898477157360406
Rule: head_of_state -> agent_of; Reward: 0.9875776397515528
Rule: agency_of -> based_in0; Reward: 0.9820143884892086
Rule: head_of_gov-x -> citizen_of-x; Reward: 0.9797979797979798
Rule: head_of_state -> citizen_of; Reward: 0.968944099378882
Rule: minister_of -> citizen_of; Reward: 0.9064748201438849
Rule: head_of_gov-x , gpe0 -> head_of_gov; Reward: 0.7209302325581395
Rule: head_of_state , based_in0-x -> head_of_state-x; Reward: 0.7142857142857143
Rule: head_of_gov , based_in0-x -> head_of_gov-x; Reward: 0.68
Rule: in0-x , gpe0 -> in0; Reward: 0.6467181467181468
Rule: head_of , agency_of -> citizen_of; Reward: 0.6363636363636364
Rule: agent_of , based_in0-x -> citizen_of-x; Reward: 0.5912408759124088
Rule: agency_of , based_in0-x -> agency_of-x; Reward: 0.5603448275862069
Rule: agency_of-x , gpe0 -> agency_of; Reward: 0.5450980392156862
Rule: in0 , based_in0-x -> in0-x; Reward: 0.48009367681498827
Rule: member_of , agency_of -> citizen_of; Reward: 0.4514285714285714
Rule: agency_of , in0-x -> agency_of-x; Reward: 0.23076923076923078
Rule: agent_of , based_in0-x -> head_of_state-x; Reward: 0.145985401459854
Rule: based_in0 , in0-x -> agency_of-x; Reward: 0.125
Rule: agent_of , based_in0-x -> head_of_gov-x; Reward: 0.12408759124087591
Rule: head_of_gov-x , gpe0 -> head_of_state; Reward: 0.10077519379844961
Rule: citizen_of-x , gpe0 -> head_of_state; Reward: 0.09985243482538121
Rule: head_of_state , based_in0-x -> head_of_gov-x; Reward: 0.07142857142857142
Rule: citizen_of-x , gpe0 -> minister_of; Reward: 0.06837186424003935
Rule: citizen_of-x , gpe0 -> head_of_gov; Reward: 0.04722085587801279
Rule: head_of , agency_of -> head_of_state; Reward: 0.045454545454545456
Rule: head_of_gov , in0-x -> head_of_gov-x; Reward: 0.029411764705882353
Rule: head_of_gov , in0 -> head_of_state; Reward: 0.02564102564102564
Rule: head_of_gov-x , gpe0 -> minister_of; Reward: 0.023255813953488372
Rule: head_of , agency_of -> head_of_gov; Reward: 0.022727272727272728
Rule: head_of , based_in0-x -> head_of_gov-x; Reward: 0.021021021021021023
Rule: agency_of-x , gpe0 -> in0; Reward: 0.01764705882352941
Rule: member_of , based_in0-x -> head_of_state-x; Reward: 0.01563721657544957
Rule: member_of , based_in0-x -> head_of_gov-x; Reward: 0.012509773260359656
Rule: member_of , agency_of -> head_of_state; Reward: 0.011428571428571429
Rule: member_of , agency_of -> head_of_gov; Reward: 0.005714285714285714
Rule: citizen_of-x , gpe0 -> head_of_state-x; Reward: 0.000983767830791933
```

selected 규칙



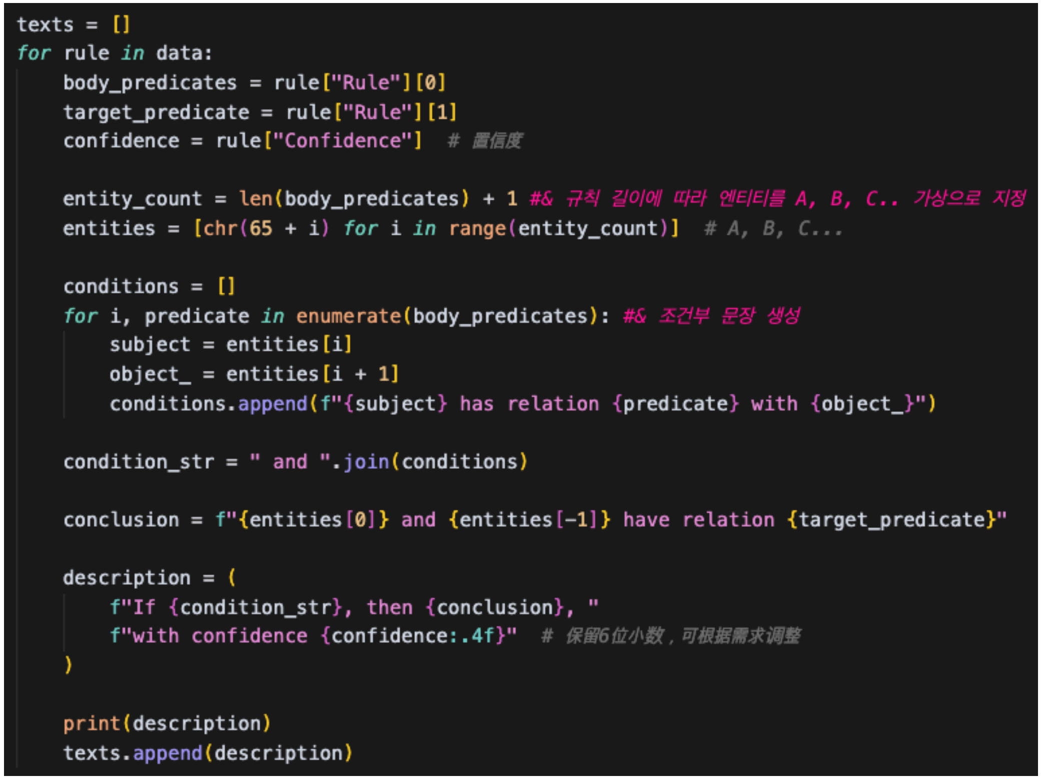
```
if reward > 0.5:
    selected_rules.append({
        "Rule":rule,
        "Confidence":reward
    })
```



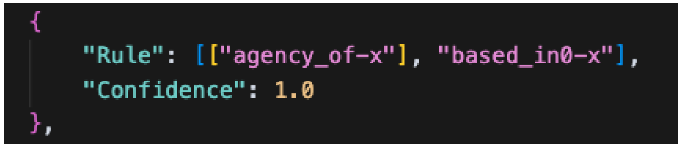
```
[
    {
        "Rule": [["agency_of-x"], "based_in0-x"],
        "Confidence": 1.0
    },
    {
        "Rule": [["head_of"], "member_of"],
        "Confidence": 1.0
    },
    {
        "Rule": [["minister_of"], "agent_of"],
        "Confidence": 0.9928057553956835
    },
    {
        "Rule": [["head_of_state-x"], "citizen_of-x"],
        "Confidence": 0.9898477157360406
    },
    {
        "Rule": [["head_of_state"], "agent_of"],
        "Confidence": 0.9875776397515528
    },
    {
        "Rule": [["agency_of"], "based_in0"],
        "Confidence": 0.9820143884892086
    },
    {
        "Rule": [["head_of_gov-x"], "citizen_of-x"],
        "Confidence": 0.9797979797979798
    },
    {
        "Rule": [["head_of_state"], "citizen_of"],
        "Confidence": 0.968944099378882
    },
    {
        "Rule": [["minister_of"], "citizen_of"],
        "Confidence": 0.9064748201438849
    },
    {
        "Rule": [["head_of_gov-x", "gpe0"], "head_of_gov"],
        "Confidence": 0.7209302325581395
```

rules to text _ rule description.txt 생성

- selected_rules 를 텍스트 설명으로 전환

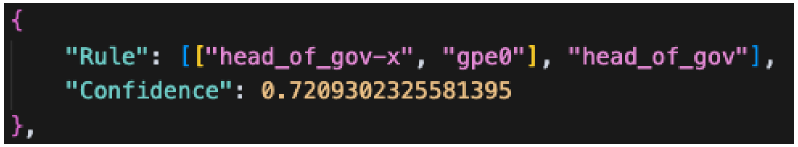


```python
texts = []
for rule in data:
    body_predicates = rule["Rule"][0]
    target_predicate = rule["Rule"][1]
    confidence = rule["Confidence"]  # 置信度

    entity_count = len(body_predicates) + 1 #& 규칙 길이에 따라 엔티티를 A, B, C.. 가상으로 지정
    entities = [chr(65 + i) for i in range(entity_count)]  # A, B, C...

    conditions = []
    for i, predicate in enumerate(body_predicates): #& 조건부 문장 생성
        subject = entities[i]
        object_ = entities[i + 1]
        conditions.append(f"{subject} has relation {predicate} with {object_}")

    condition_str = " and ".join(conditions)

    conclusion = f"{entities[0]} and {entities[-1]} have relation {target_predicate}"

    description = (
        f"If {condition_str}, then {conclusion}, "
        f"with confidence {confidence:.4f}"  # 保留6位小数，可根据需求调整
    )

    print(description)
    texts.append(description)
```



```json
{
    "Rule": [["agency_of-x"], "based_in0-x"],
    "Confidence": 1.0
},
```

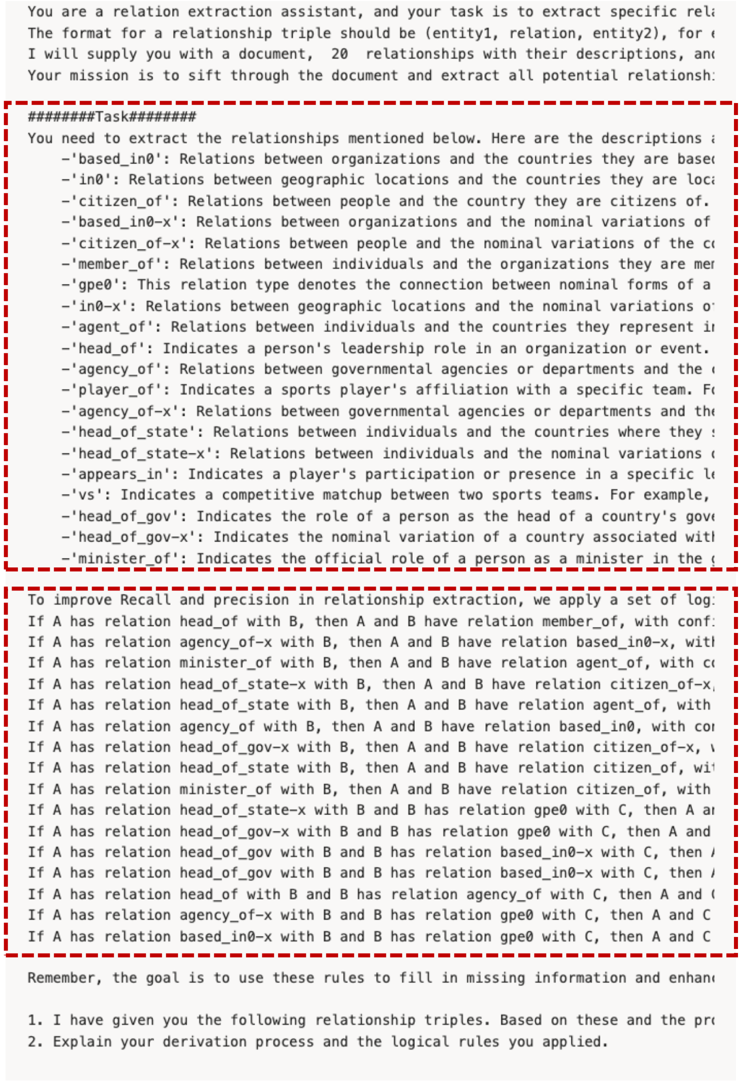
If A has relation agency_of-x with B,
then A and B have relation based_in0-x,
with confidence 1.0000



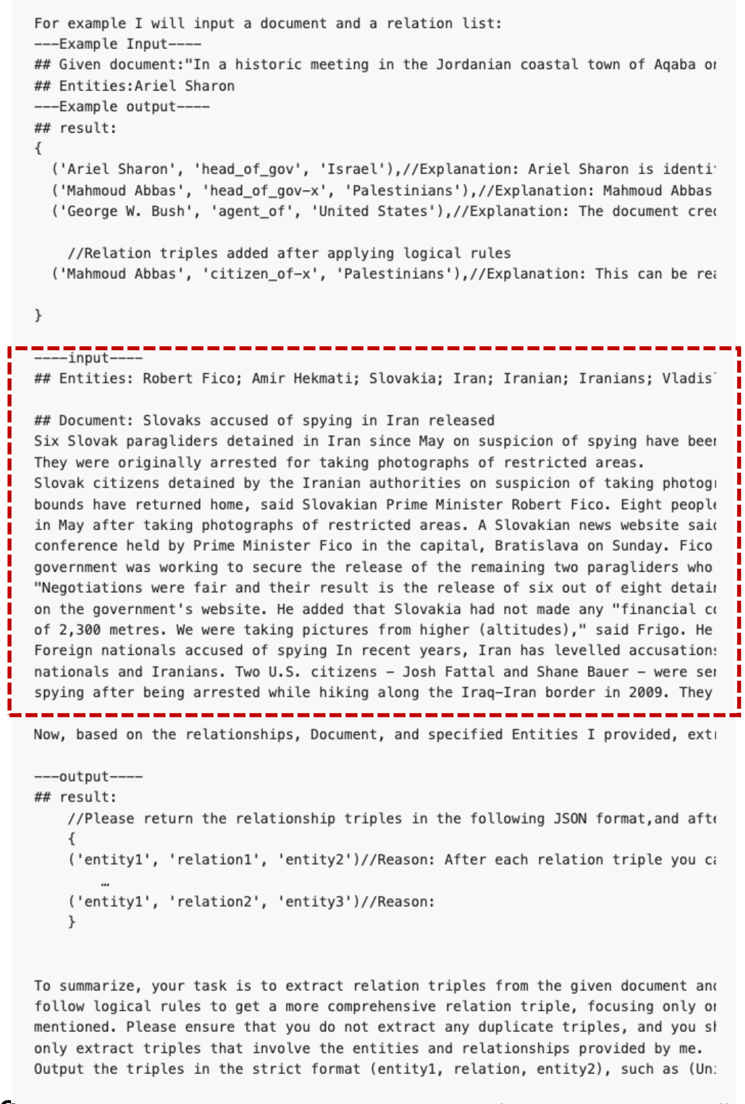
```json
{
    "Rule": [["head_of_gov-x", "gpe0"], "head_of_gov"],
    "Confidence": 0.7209302325581395
},
```

If A has relation head_of_gov-x with B
and B has relation gpe0 with C,
then A and C have relation head_of_gov,
with confidence 0.7209

Relation Extraction _ LLM prompt



```
You are a relation extraction assistant, and your task is to extract specific rel
The format for a relationship triple should be (entity1, relation, entity2), for
I will supply you with a document,  20  relationships with their descriptions, an
Your mission is to sift through the document and extract all potential relationsh

########Task########
You need to extract the relationships mentioned below. Here are the descriptions
    -'based_in0': Relations between organizations and the countries they are base
    -'in0': Relations between geographic locations and the countries they are loc
    -'citizen_of': Relations between people and the country they are citizens of.
    -'based_in0-x': Relations between organizations and the nominal variations of
    -'citizen_of-x': Relations between people and the nominal variations of the c
    -'member_of': Relations between individuals and the organizations they are me
    -'gpe0': This relation type denotes the connection between nominal forms of a
    -'in0-x': Relations between geographic locations and the nominal variations o
    -'agent_of': Relations between individuals and the countries they represent i
    -'head_of': Indicates a person's leadership role in an organization or event.
    -'agency_of': Relations between governmental agencies or departments and the
    -'player_of': Indicates a sports player's affiliation with a specific team. F
    -'agency_of-x': Relations between governmental agencies or departments and th
    -'head_of_state': Relations between individuals and the countries where they
    -'head_of_state-x': Relations between individuals and the nominal variations
    -'appears_in': Indicates a player's participation or presence in a specific l
    -'vs': Indicates a competitive matchup between two sports teams. For example,
    -'head_of_gov': Indicates the role of a person as the head of a country's gov
    -'head_of_gov-x': Indicates the nominal variation of a country associated wit
    -'minister_of': Indicates the official role of a person as a minister in the

To improve Recall and precision in relationship extraction, we apply a set of log
If A has relation head_of with B, then A and B have relation member_of, with conf
If A has relation agency_of-x with B, then A and B have relation based_in0-x, wit
If A has relation minister_of with B, then A and B have relation agent_of, with c
If A has relation head_of_state-x with B, then A and B have relation citizen_of-x,
If A has relation head_of_state with B, then A and B have relation agent_of, with
If A has relation agency_of with B, then A and B have relation based_in0, with co
If A has relation head_of_gov-x with B, then A and B have relation citizen_of-x,
If A has relation head_of_state with B, then A and B have relation citizen_of, wi
If A has relation minister_of with B, then A and B have relation citizen_of, with
If A has relation head_of_state-x with B and B has relation gpe0 with C, then A a
If A has relation head_of_gov-x with B and B has relation gpe0 with C, then A and
If A has relation head_of_gov with B and B has relation based_in0-x with C, then
If A has relation head_of_gov with B and B has relation based_in0-x with C, then
If A has relation head_of with B and B has relation agency_of with C, then A and
If A has relation agency_of-x with B and B has relation gpe0 with C, then A and C
If A has relation based_in0-x with B and B has relation gpe0 with C, then A and C

Remember, the goal is to use these rules to fill in missing information and enhan

1. I have given you the following relationship triples. Based on these and the pr
2. Explain your derivation process and the logical rules you applied.
```



```
For example I will input a document and a relation list:
---Example Input----
## Given document:"In a historic meeting in the Jordanian coastal town of Aqaba o
## Entities:Ariel Sharon
---Example output----
## result:
{
  ('Ariel Sharon', 'head_of_gov', 'Israel'),//Explanation: Ariel Sharon is identi
  ('Mahmoud Abbas', 'head_of_gov-x', 'Palestinians'),//Explanation: Mahmoud Abbas
  ('George W. Bush', 'agent_of', 'United States'),//Explanation: The document cre

    //Relation triples added after applying logical rules
  ('Mahmoud Abbas', 'citizen_of-x', 'Palestinians'),//Explanation: This can be re

}
----input----
## Entities: Robert Fico; Amir Hekmati; Slovakia; Iran; Iranian; Iranians; Vladis

## Document: Slovaks accused of spying in Iran released
Six Slovak paragliders detained in Iran since May on suspicion of spying have bee
They were originally arrested for taking photographs of restricted areas.
Slovak citizens detained by the Iranian authorities on suspicion of taking photog
bounds have returned home, said Slovakian Prime Minister Robert Fico. Eight peopl
in May after taking photographs of restricted areas. A Slovakian news website sai
conference held by Prime Minister Fico in the capital, Bratislava on Sunday. Fico
government was working to secure the release of the remaining two paragliders who
"Negotiations were fair and their result is the release of six out of eight detai
on the government's website. He added that Slovakia had not made any "financial c
of 2,300 metres. We were taking pictures from higher (altitudes)," said Frigo. He
Foreign nationals accused of spying In recent years, Iran has levelled accusation
nationals and Iranians. Two U.S. citizens - Josh Fattal and Shane Bauer - were se
spying after being arrested while hiking along the Iraq-Iran border in 2009. They

Now, based on the relationships, Document, and specified Entities I provided, ext

---output----
## result:
    //Please return the relationship triples in the following JSON format,and aft
    {
    ('entity1', 'relation1', 'entity2')//Reason: After each relation triple you c
        …
    ('entity1', 'relation2', 'entity3')//Reason:
    }


To summarize, your task is to extract relation triples from the given document an
follow logical rules to get a more comprehensive relation triple, focusing only o
mentioned. Please ensure that you do not extract any duplicate triples, and you s
only extract triples that involve the entities and relationships provided by me.
Output the triples in the strict format (entity1, relation, entity2), such as (Un
```

CONTENTS

1. Background
2. Related Works
3. Research in Progress
4. Conceptual Insights

Goal

- DocRE with LLMs is inherently challenging
- Recent studies attempt to integrate logical rules into DocRE
- Existing LLM-based approach (RuAG) relies on prompt-based rule injection
  - ✓ RuAG incorporates rules by explicitly providing rules in prompts

- **Limitations**
  - ✓ Rules are externally injected rather than internalized
  - ✓ Rule usage can be inconsistent and brittle, and requires repeated prompting

- **My Goal**
  - ✓ Enable LLMs to internalize logical rules without fine-tuning
  - ✓ Allow LLMs to utilize rules implicitly during DocRE inference

Motivation

왜 LLM이 규칙을 '이해'해야할까?

- Post-filtering
  - ✓ 잘못된 출력을 사후적으로 제거
  - ✓ 단순히 오류 감소에 한정됨

- Constraint decoding
  - ✓ 생성 과정에서 모델의 선택 공간을 제약하는 방식
  - ✓ 규칙을 위반하는 후보를 사전에 배제하여 오류 억제
  - ✓ Precision 향상, syntax correctness 보장의 효과
  - ✓ 그러나, semantic correctness나 의미적 정답성은 충분히 보장하지 못함

Motivation

왜 LLM이 규칙을 '이해'해야할까?

- Prompting
  - ✓ 규칙을 설명하여 모델이 규칙을 지키도록 유도함
  - ✓ 모델의 내재적 규칙 이해 수준은 제한적이며, 모델 자체의 추론 능력에 많이 의존함
  - ✓ 복잡한 규칙, 형식, 논리적 제약에는 취약함

- Fine-tuning
  - ✓ 고비용 및 재학습 이슈
  - ✓ 규칙 변화에 유연하지 않음

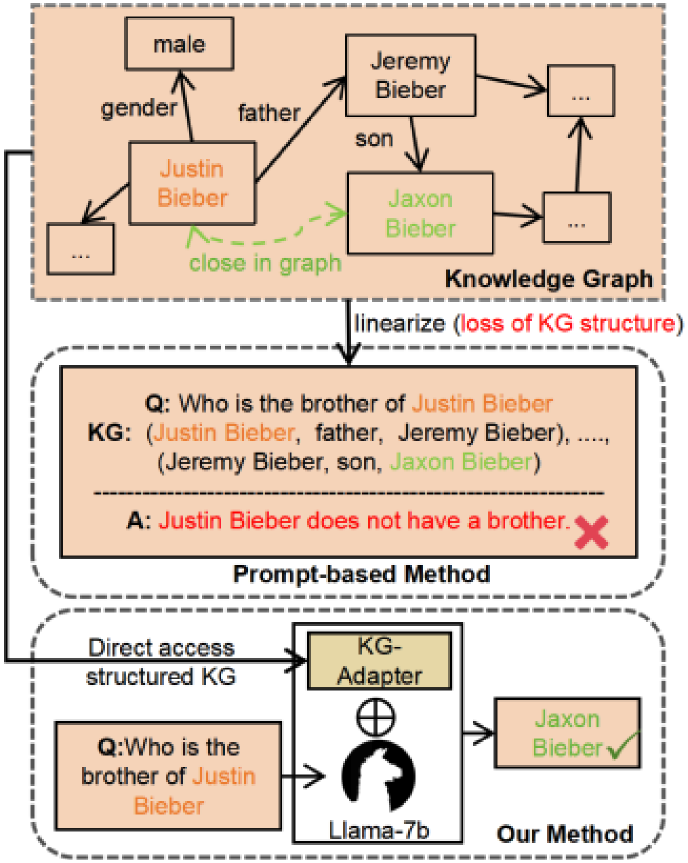
Motivation

왜 LLM이 규칙을 ‘이해’해야할까?

- 규칙을 LLM 내부에 직접적으로 이해시키면,
  - ✓ 확률적 생성 구조 자체가 개선(학습)됨
    - ▪ 단순 오류 억제가 아닌, 모델이 스스로 규칙을 우선시하도록 됨
  - ✓ 단순 오류 억제가 아닌 정답 생성 확률 자체를 높이는 방향
  - ✓ semantic, logical level의 correctness를 다룰 수 있음

Approach Overview

- Explicit logical rules are given
  - ✓ pre-defined logical rules (e.g. Horn clauses, first-order logic, tree-like ..)
- Inject rules into LLM via an adapter-style mechanism
  - ✓ Plug and play integration
  - ✓ No full-model fine-tuning
  - ✓ Enables rule internalization within the model
- Flexible rule management
  - ✓ Rules can be added, removed, or modified
  - ✓ No re-training required

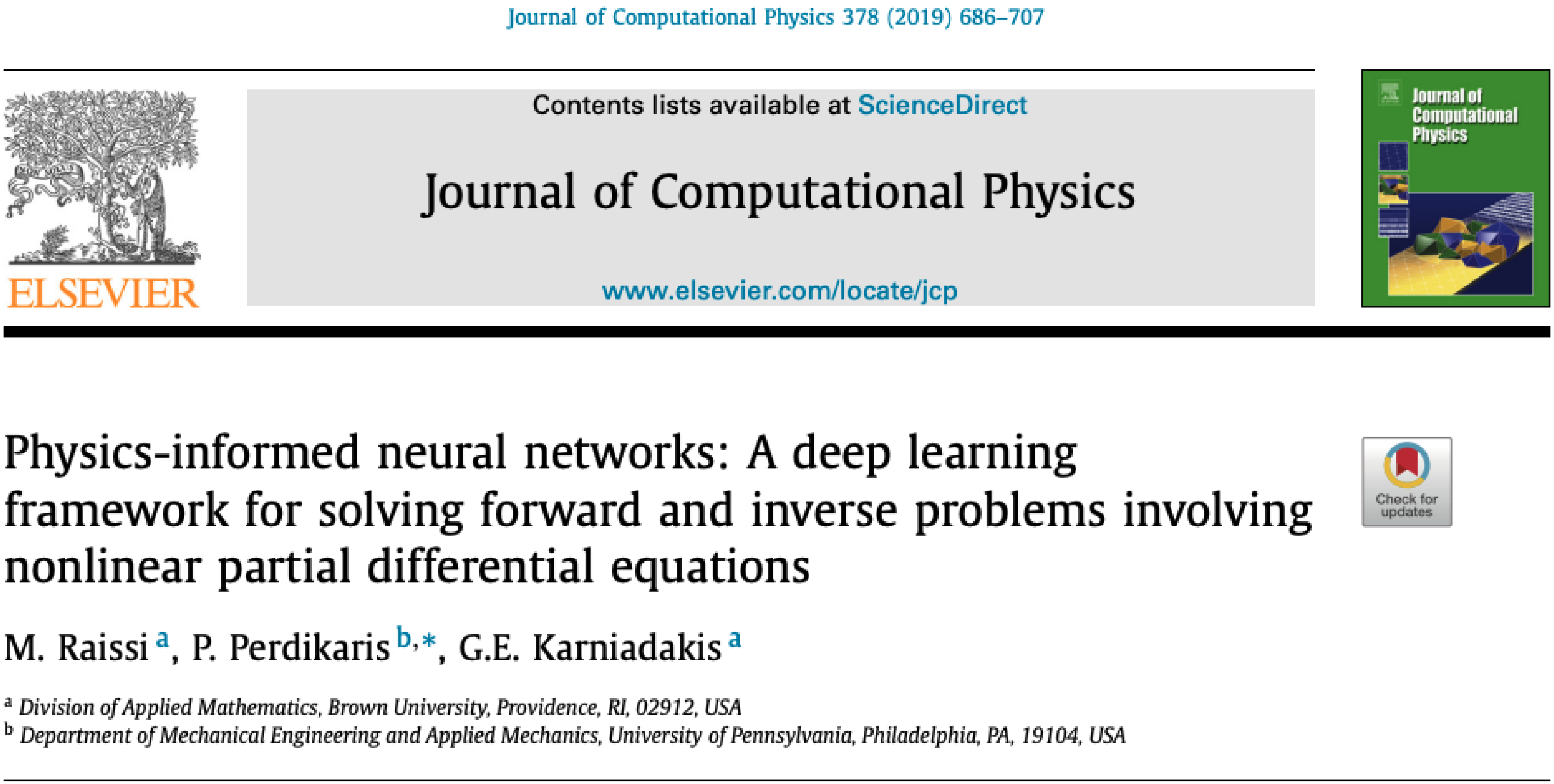

male
gender
father
Jeremy Bieber
son
...
Justin Bieber
Jaxon Bieber
...
...
close in graph
Knowledge Graph
linearize (loss of KG structure)
Q: Who is the brother of Justin Bieber
KG: (Justin Bieber, father, Jeremy Bieber), ...,
(Jeremy Bieber, son, Jaxon Bieber)
A: Justin Bieber does not have a brother.
Prompt-based Method
Direct access structured KG
KG-Adapter
Q:Who is the brother of Justin Bieber
Llama-7b
Jaxon Bieber
Our Method

이미지 출처: KG-Adapter: Enabling Knowledge Graph Integration in Large Language Models through Parameter-Efficient Fine-Tuning (ACL'24)

CONTENTS

1. Background
2. Related Works
3. Research in Progress
4. Conceptual Insights

PINN: Physics-Informed Neural Network



Journal of Computational Physics 378 (2019) 686–707

Contents lists available at ScienceDirect

Journal of Computational Physics

www.elsevier.com/locate/jcp

Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations

M. Raissi [a], P. Perdikaris [b,*], G.E. Karniadakis [a]

[a] Division of Applied Mathematics, Brown University, Providence, RI, 02912, USA
[b] Department of Mechanical Engineering and Applied Mechanics, University of Pennsylvania, Philadelphia, PA, 19104, USA

# of citations 19,686

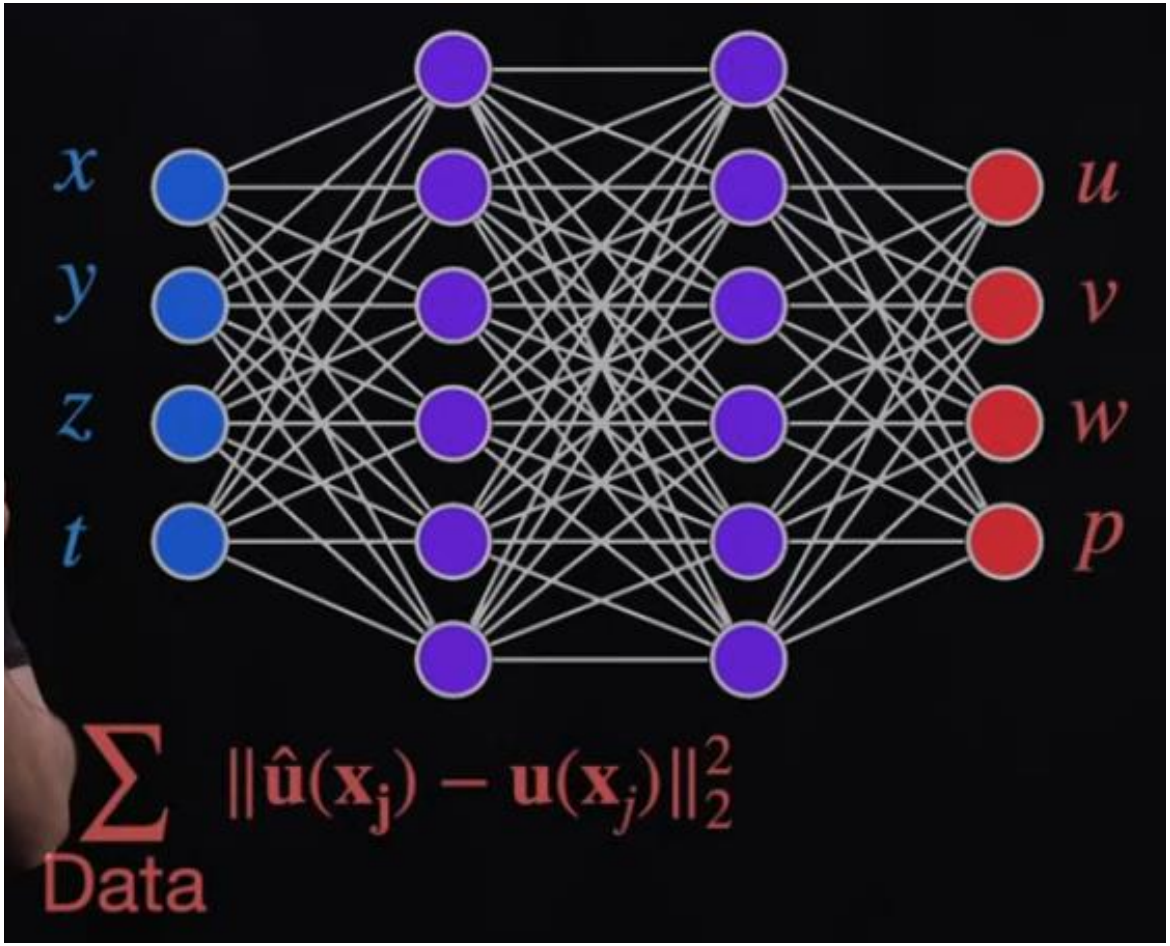
PINN _ Background

Limitations of Deep Learning

- small data regime _ 소규모 데이터 환경
  - ✓ 물리·공학 등의 분야에서는 데이터 수집 비용이 매우 높음
    -> partial information 하에서 결론을 도출해야 하는 경우가 많음
  - ✓ 이러한 환경에서 기존 딥러닝 모델은
    - ▪ robustness 부족
    - ▪ 일반화 능력이 부족 -> 오버피팅 위험이 높음
    - ▪ 수렴(convergence)에 대한 보장 없음
    - ▪ 즉, 데이터가 적을수록 purely data-driven learning 모델은 불안정해짐
- 사전 지식(prior knowledge)의 미반영
  - ✓ 물리적·공학적 시스템에는 이미 물리법칙, 경험적 규칙, 도메인 지식 등의 방대한 사전 지식이 존재함
  - ✓ 그러나 현대 기계학습 실무에서는 이러한 사전 지식이 거의 활용되지 않음. 순수 data-driven learning 모델은 현상의 물리적 기작을 활용하지 못하며, 물리적으로 말이 안되는 예측 결과를 낼 수 있음
  - ✓ prior knowledge as regularization
    : prior information은 허용 가능한 해의 공간을 관리 가능한 크기로 제한하는 정규화 요소로 작용할 수 있음 -> 이를 학습 과정에 인코딩하면 모델이 접하는 데이터의 정보량이 증폭되어,
    소규모 데이터 환경에서도 빠른 수렴과 일반화가 가능해짐

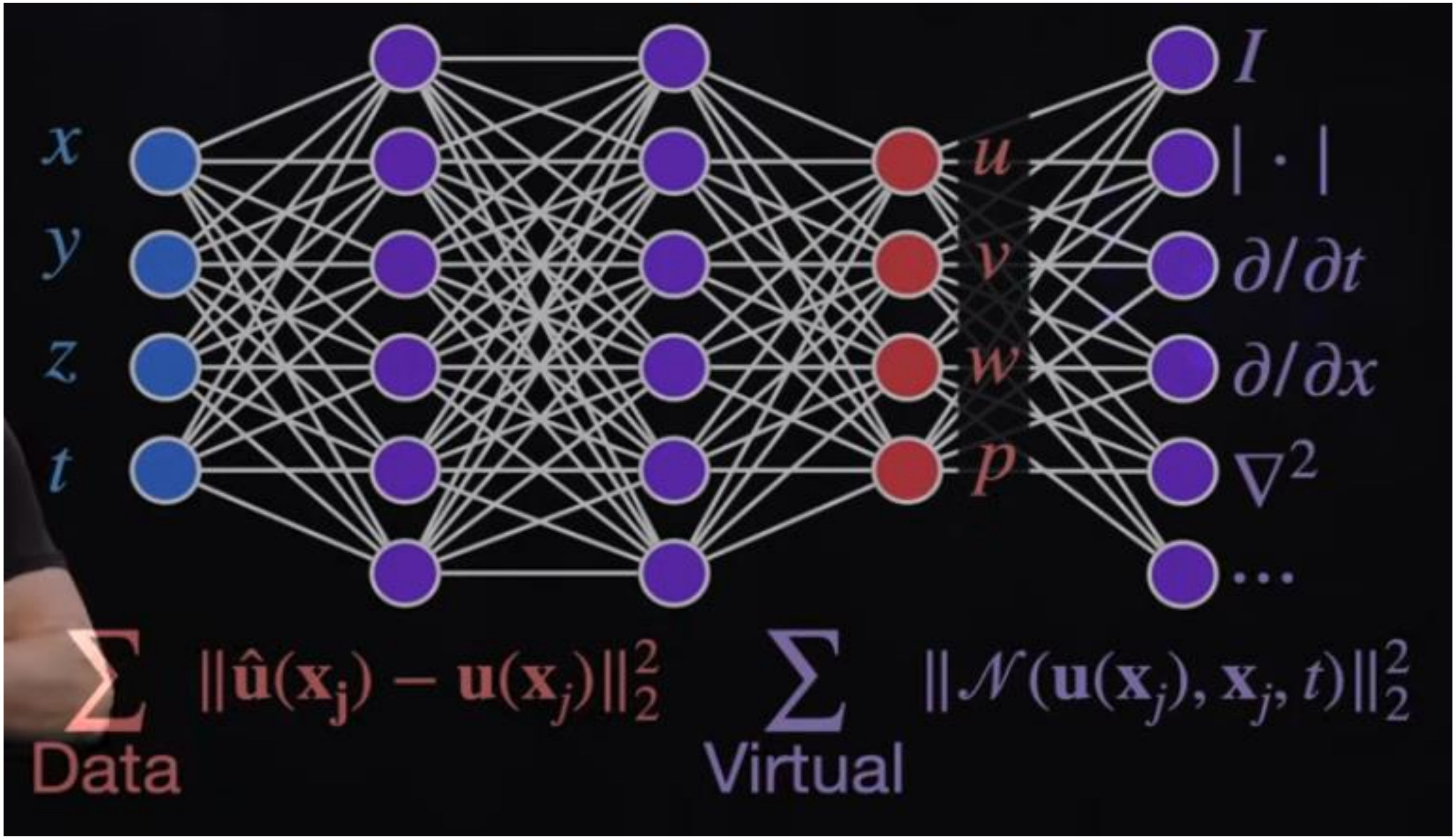
PINN: Physics-Informed Neural Network

- 물리적 관계 (편미분 방정식)를 loss에 반영하여 학습하는 뉴럴 네트워크


x
y
z
t
u
v
w
p
$\sum_{\text{Data}} \|\hat{\mathbf{u}}(\mathbf{x}_j) - \mathbf{u}(\mathbf{x}_j)\|_2^2$

PINN: Physics-Informed Neural Network

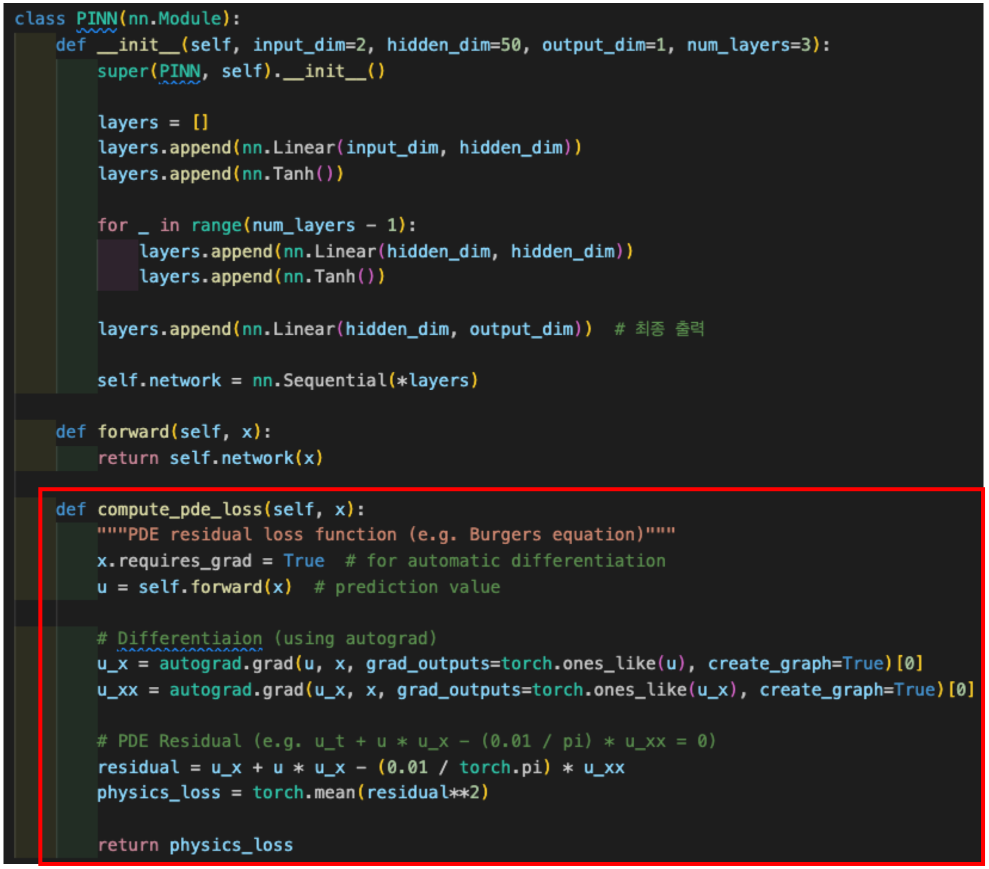
- 물리적 관계 (편미분 방정식)를 loss에 반영하여 학습하는 뉴럴 네트워크



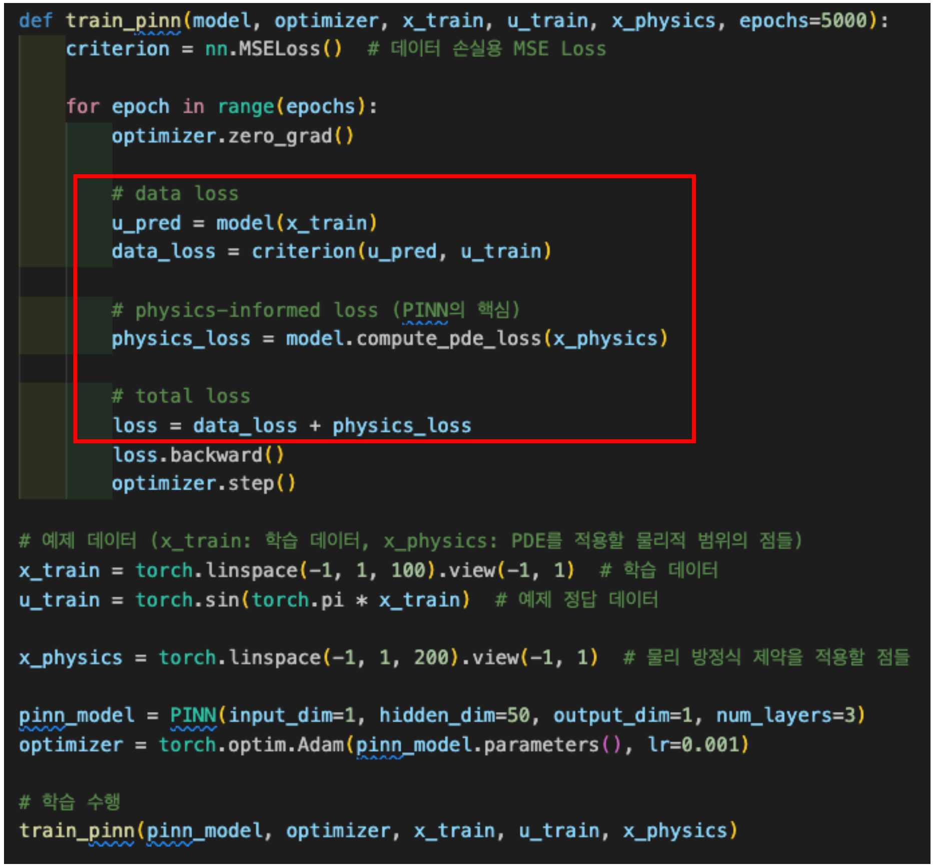
x
y
z
t
u
v
w
p
I
$|\cdot|$
$\partial/\partial t$
$\partial/\partial x$
∇^2
...
$\sum \|\hat{\mathbf{u}}(\mathbf{x}_\mathbf{j}) - \mathbf{u}(\mathbf{x}_j)\|_2^2$
Data
$\sum \|\mathcal{N}(\mathbf{u}(\mathbf{x}_j), \mathbf{x}_j, t)\|_2^2$
Virtual

PINN: Physics-Informed Neural Network

Implementation



```python
class PINN(nn.Module):
    def __init__(self, input_dim=2, hidden_dim=50, output_dim=1, num_layers=3):
        super(PINN, self).__init__()

        layers = []
        layers.append(nn.Linear(input_dim, hidden_dim))
        layers.append(nn.Tanh())

        for _ in range(num_layers - 1):
            layers.append(nn.Linear(hidden_dim, hidden_dim))
            layers.append(nn.Tanh())

        layers.append(nn.Linear(hidden_dim, output_dim))  # 최종 출력

        self.network = nn.Sequential(*layers)

    def forward(self, x):
        return self.network(x)

    def compute_pde_loss(self, x):
        """PDE residual loss function (e.g. Burgers equation)"""
        x.requires_grad = True  # for automatic differentiation
        u = self.forward(x)  # prediction value

        # Differentiaion (using autograd)
        u_x = autograd.grad(u, x, grad_outputs=torch.ones_like(u), create_graph=True)[0]
        u_xx = autograd.grad(u_x, x, grad_outputs=torch.ones_like(u_x), create_graph=True)[0]

        # PDE Residual (e.g. u_t + u * u_x - (0.01 / pi) * u_xx = 0)
        residual = u_x + u * u_x - (0.01 / torch.pi) * u_xx
        physics_loss = torch.mean(residual**2)

        return physics_loss
```



```python
def train_pinn(model, optimizer, x_train, u_train, x_physics, epochs=5000):
    criterion = nn.MSELoss()  # 데이터 손실용 MSE Loss

    for epoch in range(epochs):
        optimizer.zero_grad()

        # data loss
        u_pred = model(x_train)
        data_loss = criterion(u_pred, u_train)

        # physics-informed loss (PINN의 핵심)
        physics_loss = model.compute_pde_loss(x_physics)

        # total loss
        loss = data_loss + physics_loss
        loss.backward()
        optimizer.step()

# 예제 데이터 (x_train: 학습 데이터, x_physics: PDE를 적용할 물리적 범위의 점들)
x_train = torch.linspace(-1, 1, 100).view(-1, 1)  # 학습 데이터
u_train = torch.sin(torch.pi * x_train)  # 예제 정답 데이터

x_physics = torch.linspace(-1, 1, 200).view(-1, 1)  # 물리 방정식 제약을 적용할 점들

pinn_model = PINN(input_dim=1, hidden_dim=50, output_dim=1, num_layers=3)
optimizer = torch.optim.Adam(pinn_model.parameters(), lr=0.001)

# 학습 수행
train_pinn(pinn_model, optimizer, x_train, u_train, x_physics)
```

코드 출처: https://ffighting.net/deep-learning-paper-review/deep-simulation/pinn/

PINN: Physics-Informed Neural Network

Features

- very simple, but powerful and intuitive
  - ✓ 기존 신경망 구조에 물리 법칙을 loss function 형태로 추가
- allows you to work with relatively small datasets
  - ✓ 실제 측정 데이터가 많지 않아도 됨
  - ✓ 데이터가 없는 위치에서도 물리 방정식 만족 여부 평가 가능 (virtual points)
- physics is informed, not enforced (only suggesting)
  - ✓ 물리 법칙을 강제하는 것이 아닌, 만족하도록 유도하는 방식
  - ✓ data loss, physics loss 간의 trade-off _ 학습과정에서 항상 경쟁
  - ✓ physics loss가 정확히 0이 되는 경우는 거의 없음
    - ▪ 질량, 운동량 보존이 근사적으로만 만족됨
    - ▪ 응용에 따라
      - – small but non-zero physics loss가 충분히 허용되는 경우도 있지만,
      - – exact physics enforcement가 필요한 경우도 있음

        -> Architectural design, actual constraint optimization 등이 필요

AGI: Artificial General Intelligence

: 인간과 유사한 수준의 이해·학습·추론 능력을 발휘할 수 있는 인공지능

- Narrow AI (현재의 인공지능)
  - ✓ 특정 목적 (task-specific)을 수행하기 위해 개발됨
  - ✓ 한정된 작업 영역에서는 인간 수준 또는 그 이상의 성능을 달성
  - ✓ 사전 정의된 데이터와 파라미터 범위 내에서만 동작

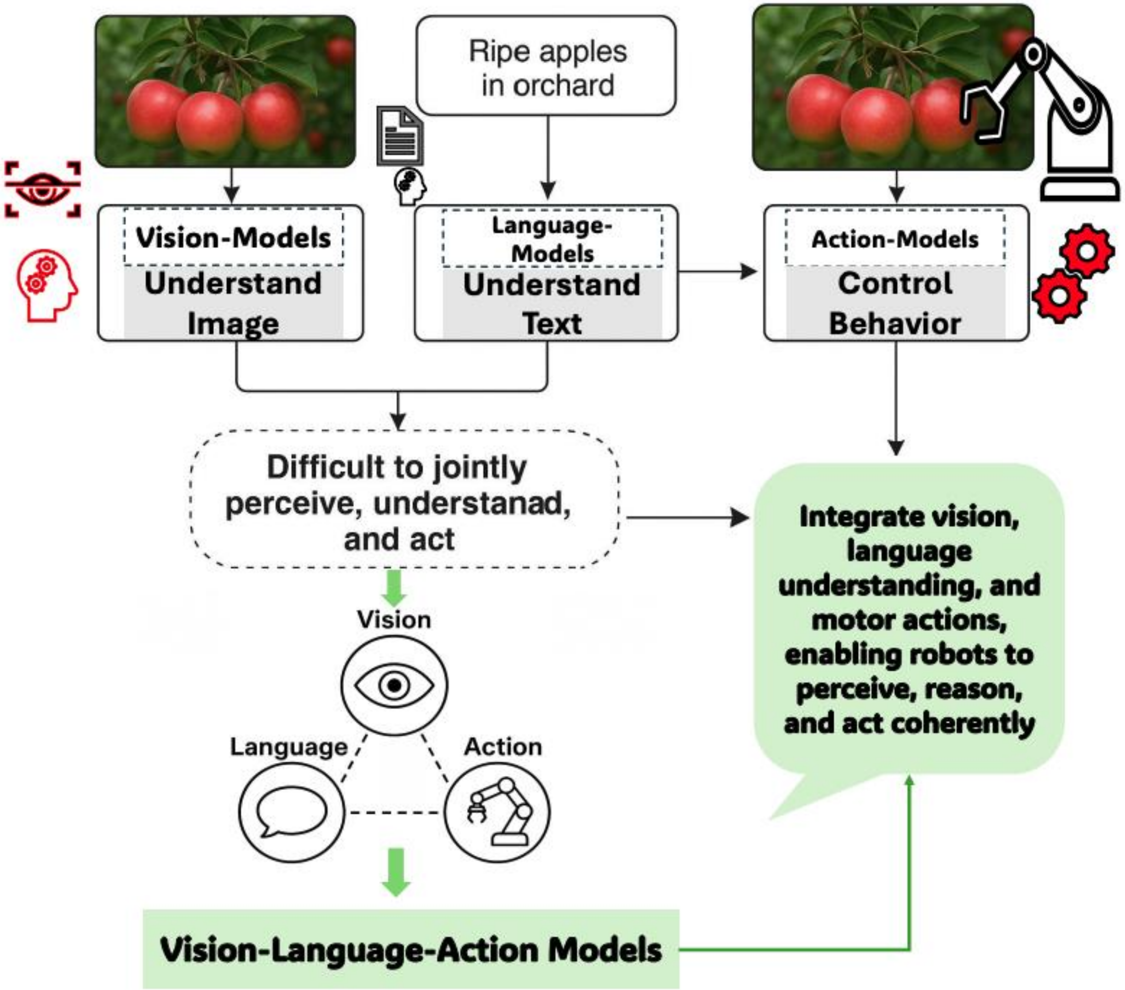
    e.g. 이미지 인식 및 생성에 대해 학습한 AI 모델은 웹사이트를 설계하거나 구축할 수 없음

- AGI (인공 일반 지능)
  - ✓ 다양한 분야에서 인간과 유사한 학습, 이해, 추론 능력을 지닌 범용 지능을 지향
  - ✓ 인간은 언어 학습, 문제 해결, 예술 창작 등 다양한 분야에서 일정 수준의 능력을 발휘할 수 있는데, 이러한 인간의 일반적 지능을 컴퓨터에서 재현하는 것이 목표
  - ✓ 새로운 환경과 문제에 대해 별도의 재학습 없이 개념을 전이하여 유연하게 문제 해결

[1] https://aws.amazon.com/ko/what-is/artificial-general-intelligence/

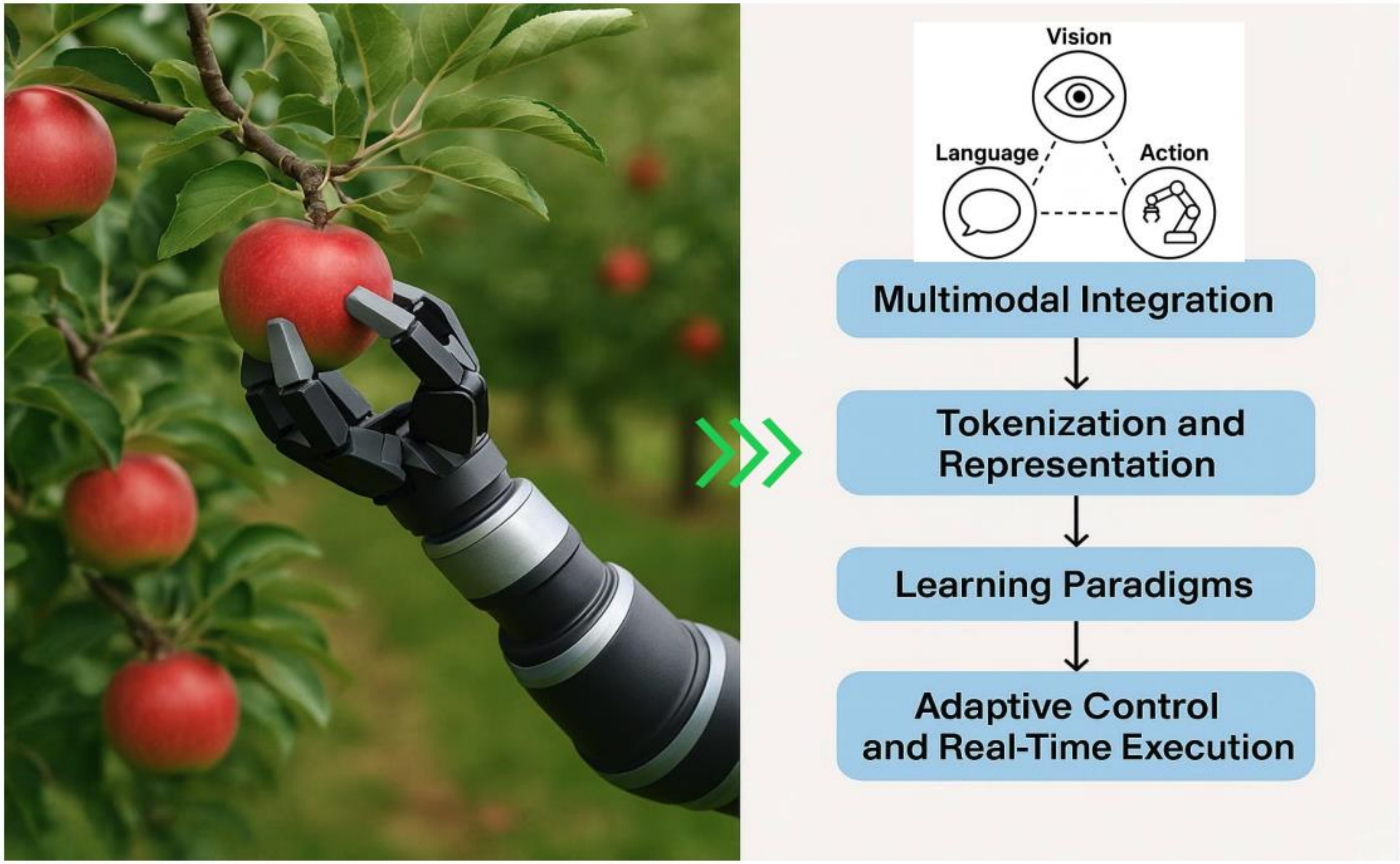
[2] https://www.samsungsds.com/kr/insights/artificial_general_intelligence_20240417.html

VLA: Vision-Language-Action Models


Ripe apples in orchard
Vision-Models
Understand Image
Language-Models
Understand Text
Action-Models
Control Behavior
Difficult to jointly perceive, understanad, and act
Integrate vision, language understanding, and motor actions, enabling robots to perceive, reason, and act coherently
Vision
Language
Action
Vision-Language-Action Models

이미지 출처: Vision-Language-Action Models: Concepts, Progress, Applications and Challenges (arXiv'25)

VLA : Vision-Language-Action Models


Vision
Language
Action
Multimodal Integration
Tokenization and Representation
Learning Paradigms
Adaptive Control and Real-Time Execution

이미지 출처: Vision-Language-Action Models: Concepts, Progress, Applications and Challenges (arXiv’25)

VLA: Vision-Language-Action Models

- 등장 배경
  - ✓ 2021-2022 개념화
  - ✓ Google Deepmind 의 RT-2 (Robotic Transformer 2)가 선구적 역할
    _ # of citations 2,392
- 핵심 개념
  - ✓ Vision, Language, Action 을 단일 통합 아키텍처로 결합
  - ✓ 인식(perception)–추론(reasoning)–제어(control) 를 하나의 프레임워크 안에서 공동학습
- 구조적 특징
  - ✓ 기존 VLM(vision-language model)을 확장하여 action token을 포함
  - ✓ 시각·언어·행동을 공통 표현 공간에서 정렬
  - ✓ 텍스트처럼 행동을 생성·예측하는 트랜스포머 기반 end-to-end 구조
- 의의
  - ✓ 단순한 멀티모달 결합 모델이 아니라, 맥락 이해 기반 추론을 통해 실제 행동을 생성하는 통합 지능 모델
  - ✓ 기존 파이프라인 형태의 로봇 시스템의 한계를 넘어, adaptive, generalizable, 실제 환경 배치가 가능한 embodied AI로 발전 중

[1] [2025][arXiv][-] Vision-Language-Action Models: Concepts, Progress, Applications and Challenges _ # of citations 63

VLA: Vision-Language-Action Models

VLA for AGI

- VLA는 embodied AI의 핵심 하위 분야이자, AGI를 향한 중요한 연구 영역
- AGI의 논의가 디지털 AI -> 물리 세계의 embodied AI 로 이동
  - ✓ 물리세계는 본질적으로 비정형적이며 예측 불가능하기 때문에 디지털 AI보다 훨씬 어려움
  - ✓ 일상 환경 자체가 OOD(out-of-distribution)의 연속
    - ▪ 자유로운 인간의 대화
    - ▪ 넘어져 있는 의자, 어질러진 방
    - ▪ 물체 가림, 조명 변화 등
- VLA 연구 분야가 빠르게 확장되고 있으며, 아직 초기 단계이지만 제한적인 수준의 지능과 일반화 능력을 보여주고 있음

[1] [2025][arXiv][-] A Survey on Vision-Language-Action Models: An Action Tokenization Perspective _ # of citations 37

Thank you for your attention!
Any questions?

Hye-Yoon Baek (hannah100@konkuk.ac.kr / hyeyoonbaek@gmail.com)

Konkuk University, Seoul, Republic of Korea

2026.01.08