Lecture#4: Representation Learning on Hyper-relational Knowledge Graphs

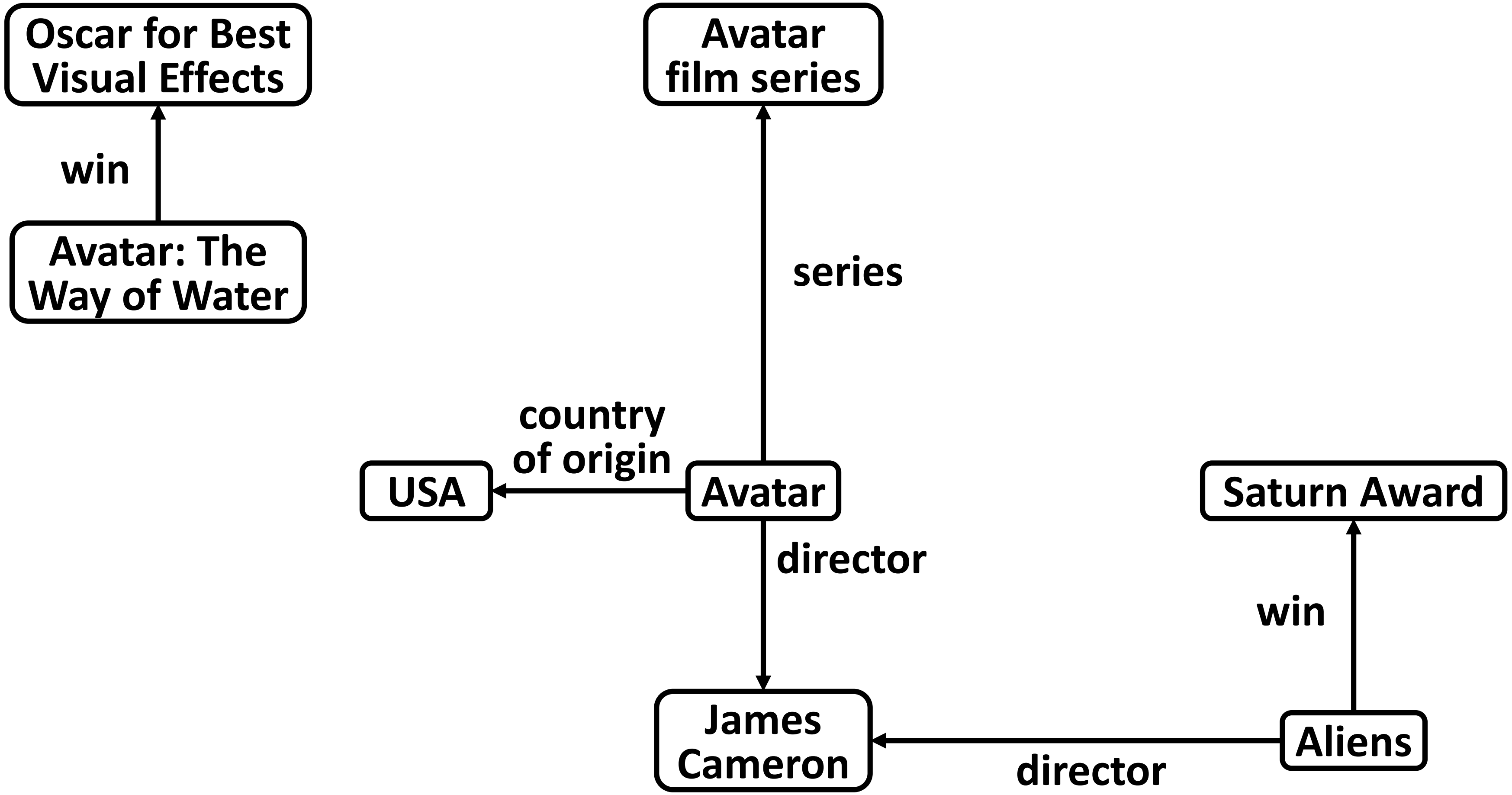
Joyce Jiyoung Whang
School of Computing, KAIST

Key Facets in Modern Knowledge Graph Representation Learning (KeyKGRL), ISWC 2025 Tutorial

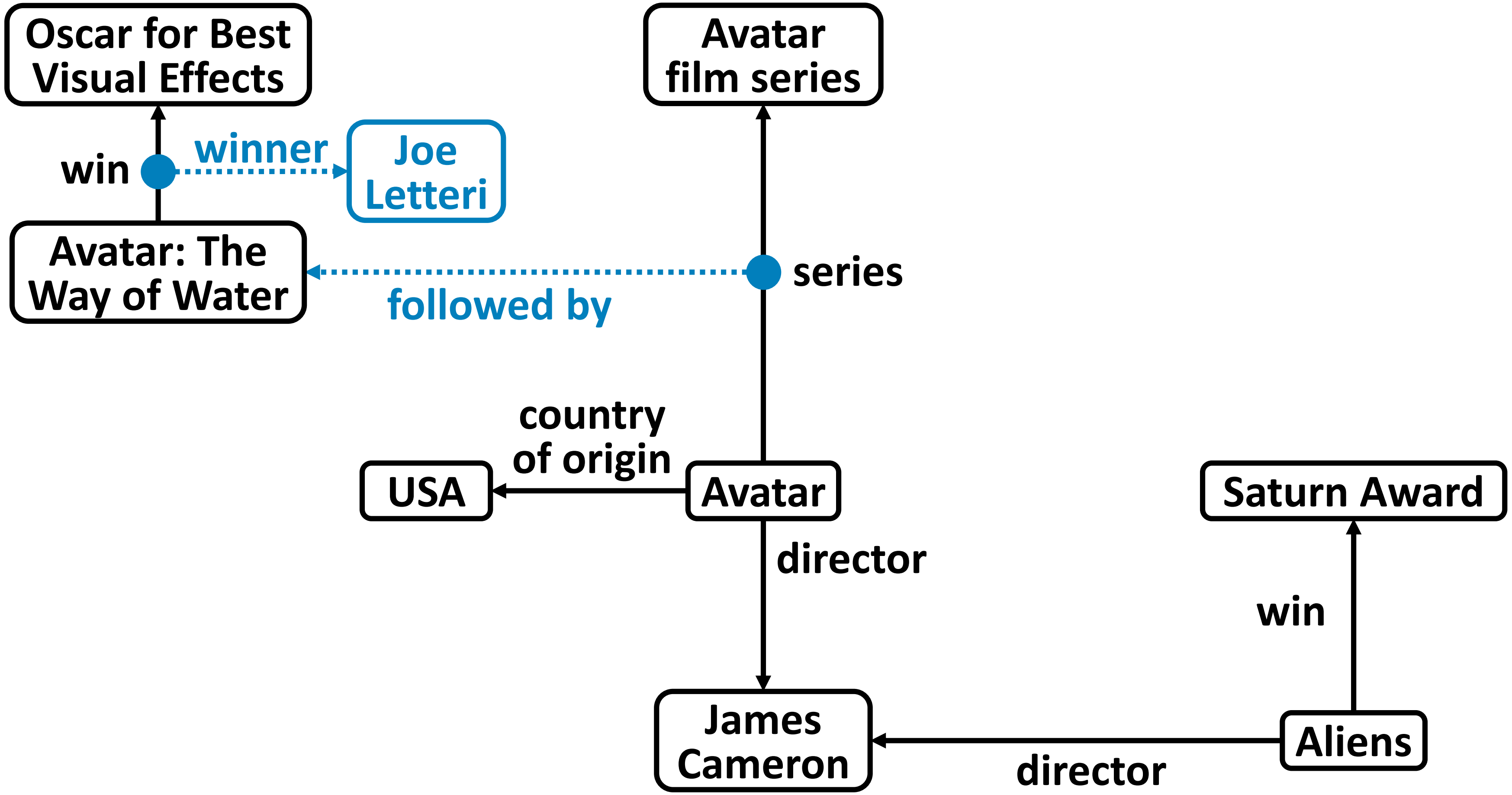
https://bdi-lab.kaist.ac.kr

KAIST BDILab
BIG DATA INTELLIGENCE

01 Knowledge Graphs



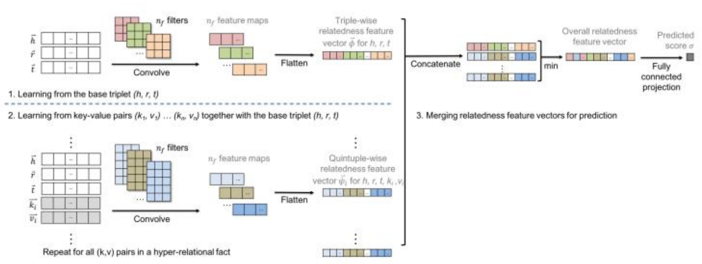
01 Hyper-relational Knowledge Graphs (HKGs)



01 Representation Learning on HKGs

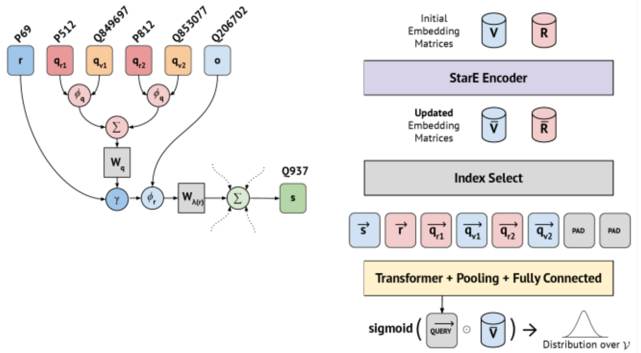
HINGE (TheWebConf 2020)

- A hyper-relational KG embedding model that directly learns from hyper-relational facts in a KG
- Captures the structural information of the KG encoded in the triplets
  - Also captures the correlation between each triplet and its associated qualifiers



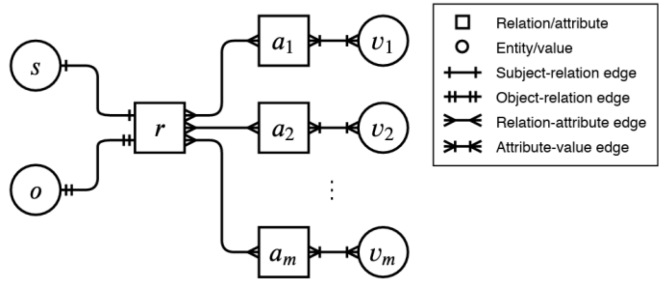
StarE (EMNLP 2020)

- A message passing based graph encoder that is capable of modeling hyper-relational KGs
- Can encode qualifiers along with the main triplet
- Demonstrates that existing benchmarks for evaluating LP on HKGs suffer from flaws



GRAN (ACL 2021 Findings)

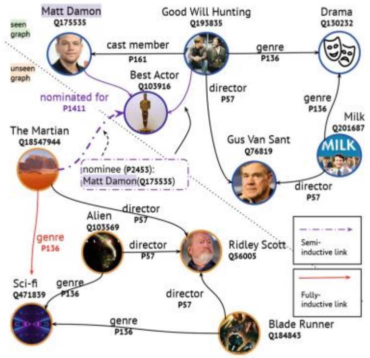
- Represents the structure of a hyper-relational fact as a small heterogeneous graph
  - Models the heterogeneous graph with edge-biased fully-connected attention
- Can fully model global and local dependencies in each fact



01 Representation Learning on HKGs

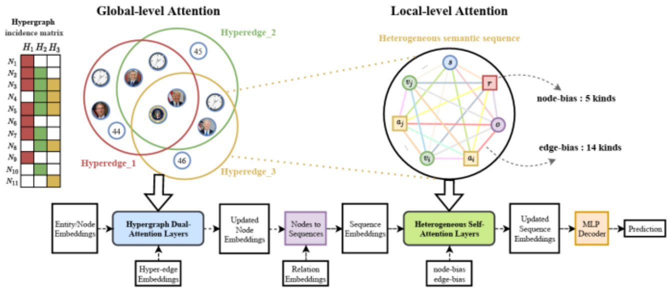
QBLP (ISWC 2021)

- Proposes a classification of inductive LP scenarios that describes the settings formally
- Adapts two existing baseline models for the inductive LP tasks and probes them into HKGs
- Shows that using hyper-relational facts can improve inductive KGC



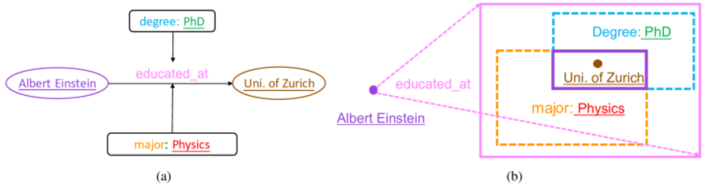
HAHE (ACL 2023)

- Represents the global structure of HKG as a hypergraph and the local structure as a semantic sequence
- Separately models the graphical structure of HKGs and sequential structure inside facts
- Performs multi-position prediction in hyper-relational KGs



ShrinkE (ACL 2023)

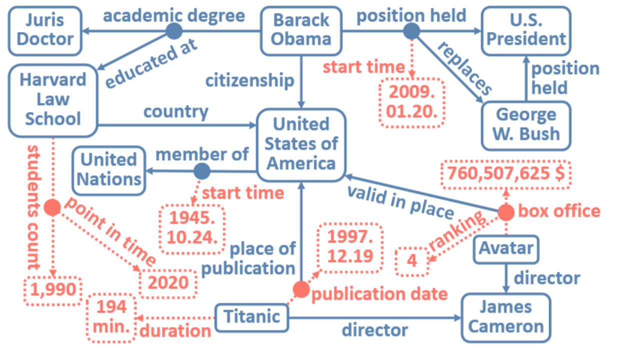
- A geometric HKG embedding method aiming to explicitly model essential inference patterns of facts
- Models a primary triplet as a spatial-functional transformation from the head into a relation-specific box
  - Each qualifier shrinks the box to narrow down the answer set



01 Representation Learning on HKGs

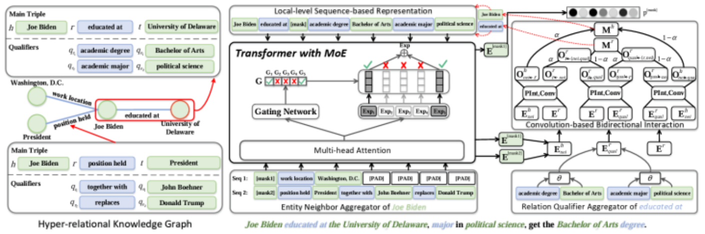
HyNT (KDD 2023)

- A unified framework that learns representations of a HKG containing numeric literals in triplets/qualifiers
- Reduces the cost of transformers by learning compact representations of triplets and qualifiers
- Predicts missing entities, relations, and numeric values in a HKG



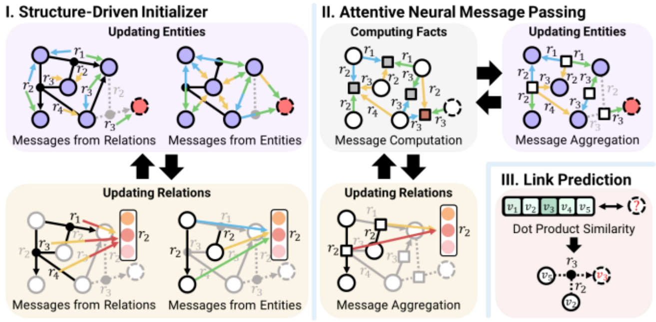
HyperFormer (CIKM 2023)

- A model that considers local-level sequential information
  - Encodes the content of the entities, relations, and qualifiers of a triplet
- Introduces a Mixture-of-Experts strategy to strengthen the representation capabilities



MAYPL (ICML 2025)

- Demonstrates that thoroughly leveraging the structure of an HKG is crucial for reasoning on HKGs
- The first structural representation learning method for HKGs that can be applied in both transductive and inductive learning settings

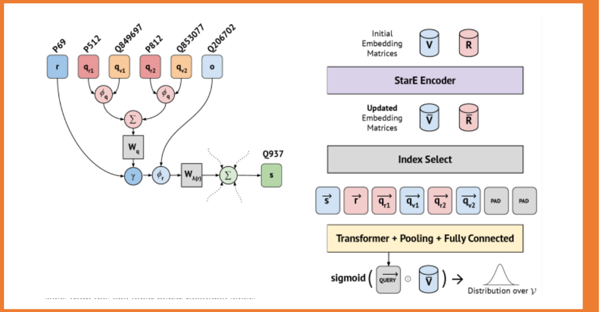


I

Message Passing for Hyper-Relational Knowledge Graphs

Mikhail Galkin, Priyansh Trivedi, Gaurav Maheshwari, Ricardo Usbeck, and Jens Lehmann

EMNLP 2020

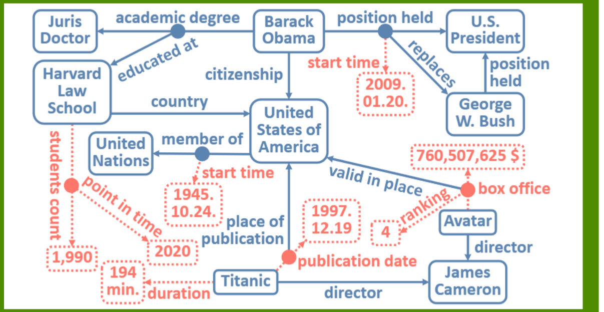


II

Representation Learning on Hyper-Relational and Numeric Knowledge Graphs with Transformers

Chanyoung Chung‡, Jaejun Lee‡, and Joyce Jiyoung Whang*

KDD 2023

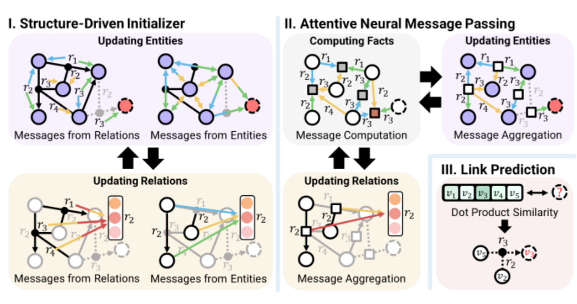


III

Structure Is All You Need: Structural Representation Learning on Hyper-Relational Knowledge Graphs

Jaejun Lee and Joyce Jiyoung Whang*

ICML 2025

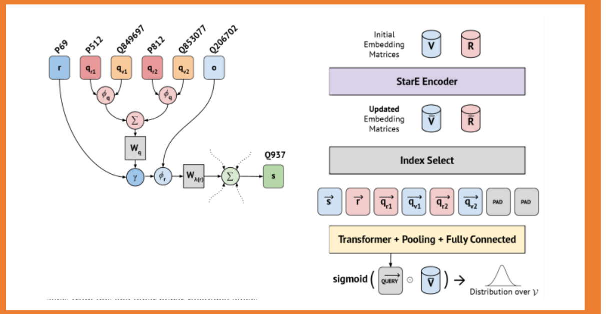


I

Message Passing for Hyper-Relational Knowledge Graphs

Mikhail Galkin, Priyansh Trivedi, Gaurav Maheshwari, Ricardo Usbeck, and Jens Lehmann

EMNLP 2020

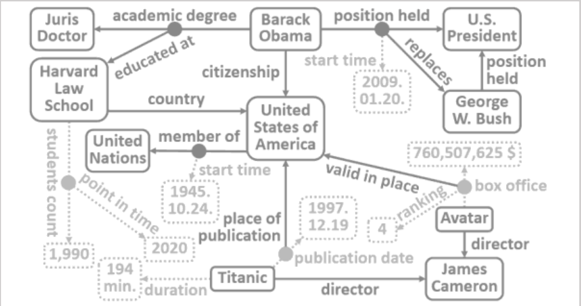


II

Representation Learning on Hyper-Relational and Numeric Knowledge Graphs with Transformers

Chanyoung Chung‡, Jaejun Lee‡, and Joyce Jiyoung Whang*

KDD 2023

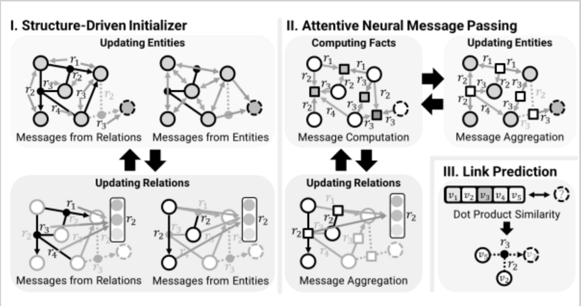


III

Structure Is All You Need: Structural Representation Learning on Hyper-Relational Knowledge Graphs

Jaejun Lee and Joyce Jiyoung Whang*

ICML 2025



02 Motivation

- Existing representation learning approaches for HKGs largely treat a hyper-relational fact as a **n-ary composed relation**

- This may **lose entity-relation attribution** or **ignore the semantic difference** between a triplet relation and qualifier relation

- Some others **decompose a hyper-relational fact** into multiple quintuples comprised of a triplet and one qualifier key-value pair

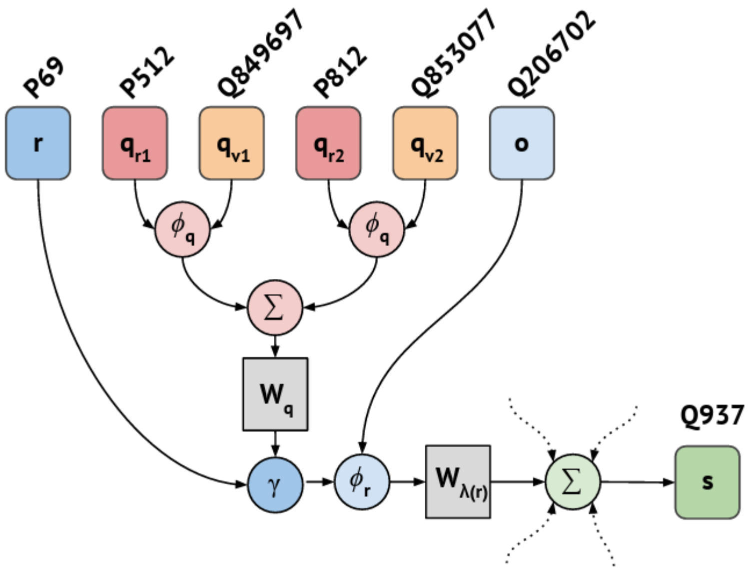
02 Contributions

- Propose an alternate graph representation learning mechanism capable of encoding HKGs with an **arbitrary number of qualifiers**, while keeping the **semantic roles of qualifiers and triplets intact**

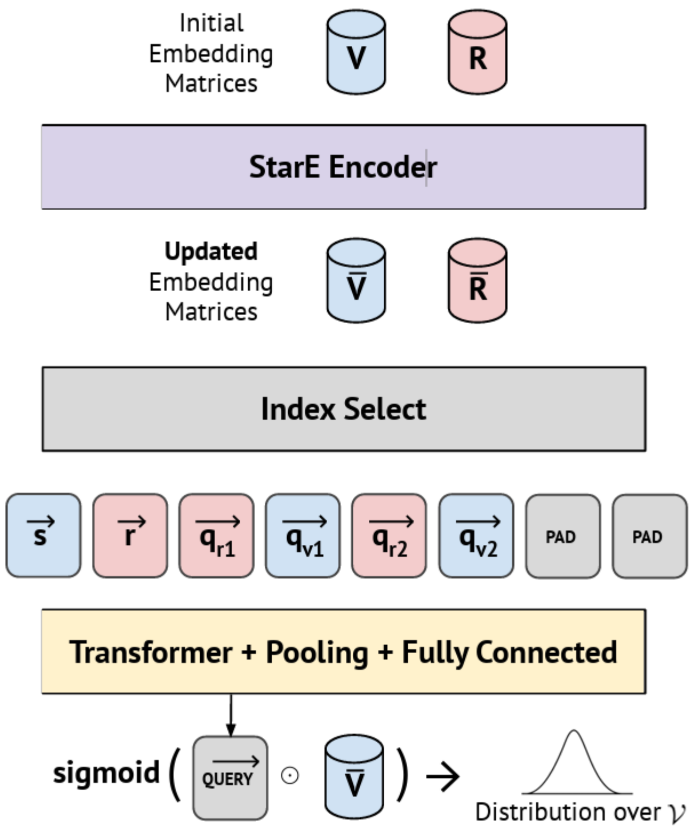
- Propose **StarE**, a GNN-based approach capable of handling HKGs

- Show that existing benchmarks for link prediction over hyper-relational KGs **exhibit some design flaws**
  - Propose a new hyper-relational link prediction dataset that contains facts with varying amounts of qualifiers

Overview of StarE



StarE Encoder

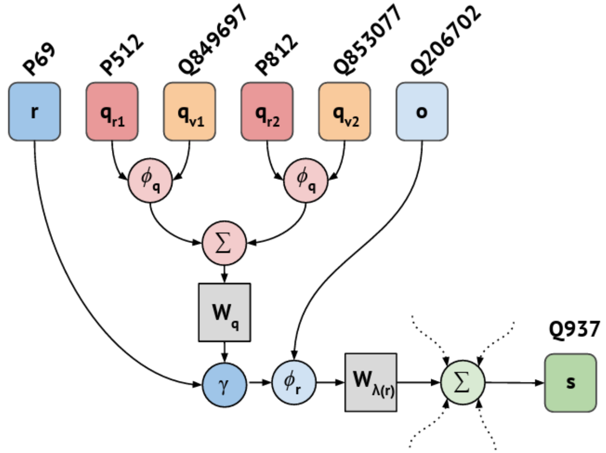


StarE-based Link Prediction Model

02 StarE Encoder

- StarE **incorporates the qualifiers and the primary triplet** into a message passing process
  - Combines the relation embedding $\boldsymbol{h}_r$ with a vector $\boldsymbol{h}_q$ that represents all qualifiers in the fact

- The qualifier vector $\boldsymbol{h}_q$ is obtained by **aggregating representations of the qualifiers**
  - A composition function ϕ_q is used to compute a qualifier representation using the representations of the qualifier entity q_v and the qualifier relation q_r



02 WD50K Dataset

- WikiPeople and JF17K are benchmarks for learning representations on HKGs
  - 13% of facts contain a literal in WikiPeople
    - After removing literals, less than 3% of the facts contain qualifier
  - 44.5% of the facts in the test set share the same primary triplet as the facts in the training set in JF17K

- Propose a new dataset, WD50K, extracted from WikiData
  - 14% of facts have at least one qualifier

Dataset	Statements	w/ Quals (%)	Entities	Relations	E in quals	R in quals	Train	Valid	Test
WD50K	236,507	32,167 (13.6%)	47,156	532	5460	45	166,435	23,913	46,159
WikiPeople	369,866	9,482 (2.6%)	34,839	375	416	35	294,439	37,715	37,712
JF17K	100,947	46,320 (45.9%)	28,645	322	3652	180	76,379	-	24,568

02 Experiments

Exp #	Method	WikiPeople				JF17K			
		MRR	H@1	H@5	H@10	MRR	H@1	H@5	H@10
1	m-TransH	0.063	0.063	-	0.300	0.206	0.206	-	0.463
1	RAE	0.059	0.059	-	0.306	0.215	0.215	-	0.469
1	NaLP-Fix	0.420	0.343	-	0.556	0.245	0.185	-	0.358
1	HINGE	0.476	**0.415**	-	0.585	0.449	0.361	-	0.624
1,4	Transformer (H)	0.469	0.403	0.538	0.586	0.512	0.434	0.593	0.665
1,4	StarE (H) + Transformer(H)	**0.491**	0.398	**0.592**	**0.648**	**0.574**	**0.496**	**0.658**	**0.725**
4	Transformer (T)	0.474	0.419	0.532	0.575	0.537	0.473	0.606	0.663
4	StarE (T) + Transformer (T)	0.493	0.400	0.592	0.648	0.562	0.493	0.637	0.702

Exp #	Dataset → Method ↓	WD50K		
		MRR	H@1	H@10
4	Baseline (Transformer (T))	0.275	0.207	0.404
4	StarE (T) + Transformer(T)	0.308	0.228	0.465
4	NaLP-Fix	0.177	0.131	0.264
4	HINGE	0.243	0.176	0.377
1,2,4	Baseline (Transformer (H))	0.286	0.222	0.406
1,2,4	StarE (H) + Transformer(H)	**0.349**	**0.271**	**0.496**

02 Conclusion

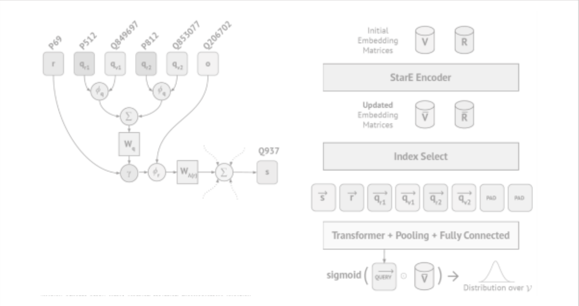
- Present **StarE**, an instance of message passing framework for representation learning over hyper-relational KGs
- StarE performs competitively on link prediction tasks over existing hyper-relational approaches and greatly outperforms triplet-only baselines
- Identify significant **flaws in existing link prediction datasets** and propose WD50K
  - WD50K: Wikidata-based hyper-relational dataset that is closer to real-world graphs and better captures the complexity of the link prediction task on HKGs

I

Message Passing for Hyper-Relational Knowledge Graphs

Mikhail Galkin, Priyansh Trivedi, Gaurav Maheshwari, Ricardo Usbeck, and Jens Lehmann

EMNLP 2020

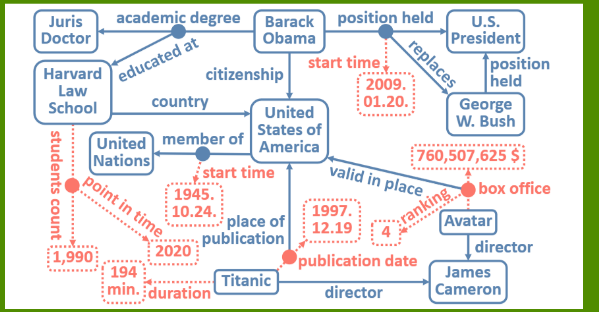


II

Representation Learning on Hyper-Relational and Numeric Knowledge Graphs with Transformers

Chanyoung Chung‡, Jaejun Lee‡, and Joyce Jiyoung Whang*

KDD 2023

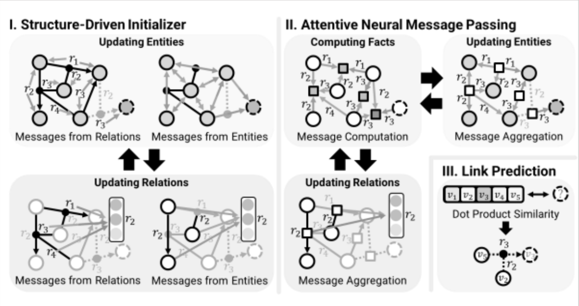


III

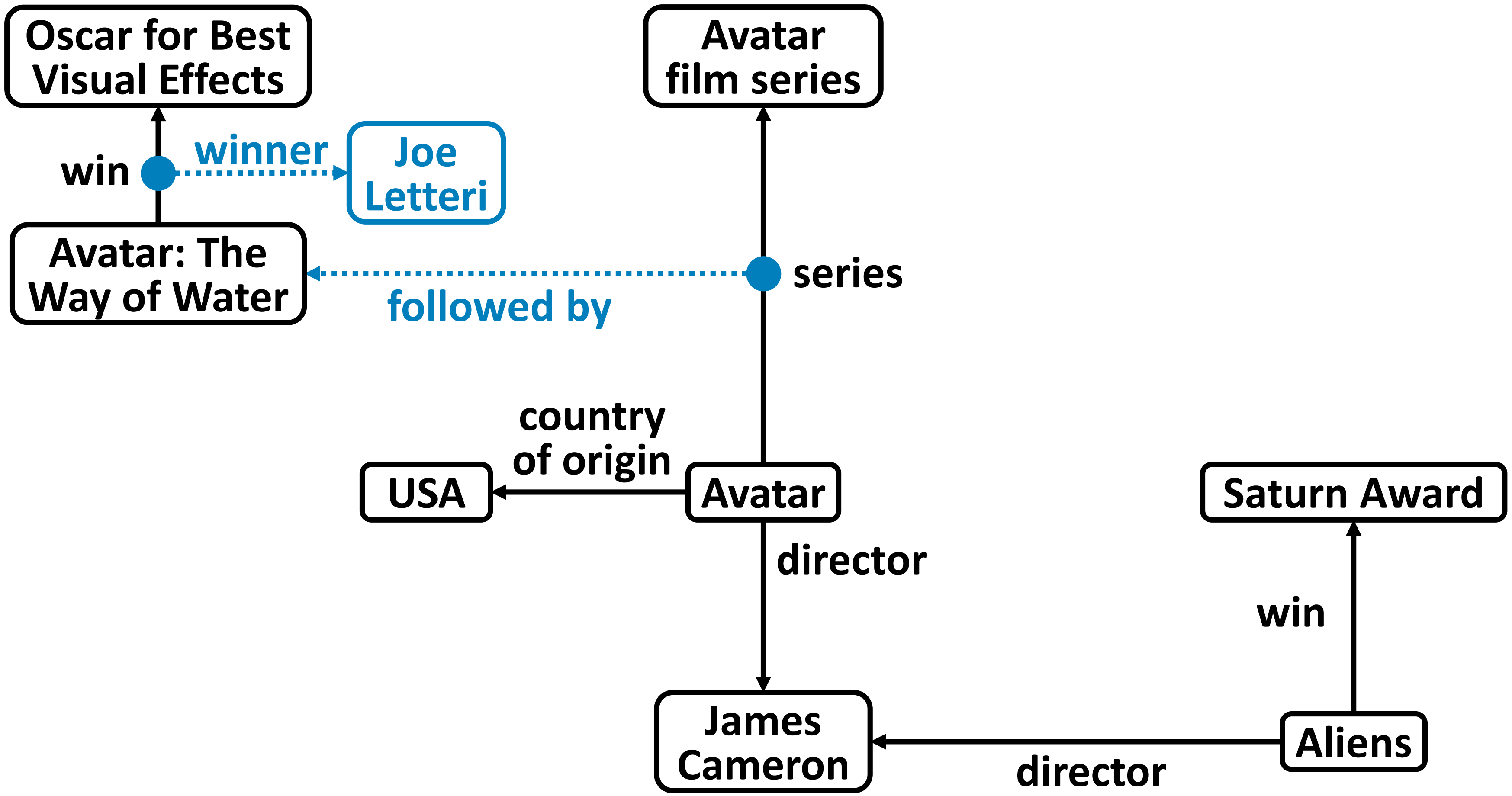
Structure Is All You Need: Structural Representation Learning on Hyper-Relational Knowledge Graphs

Jaejun Lee and Joyce Jiyoung Whang*

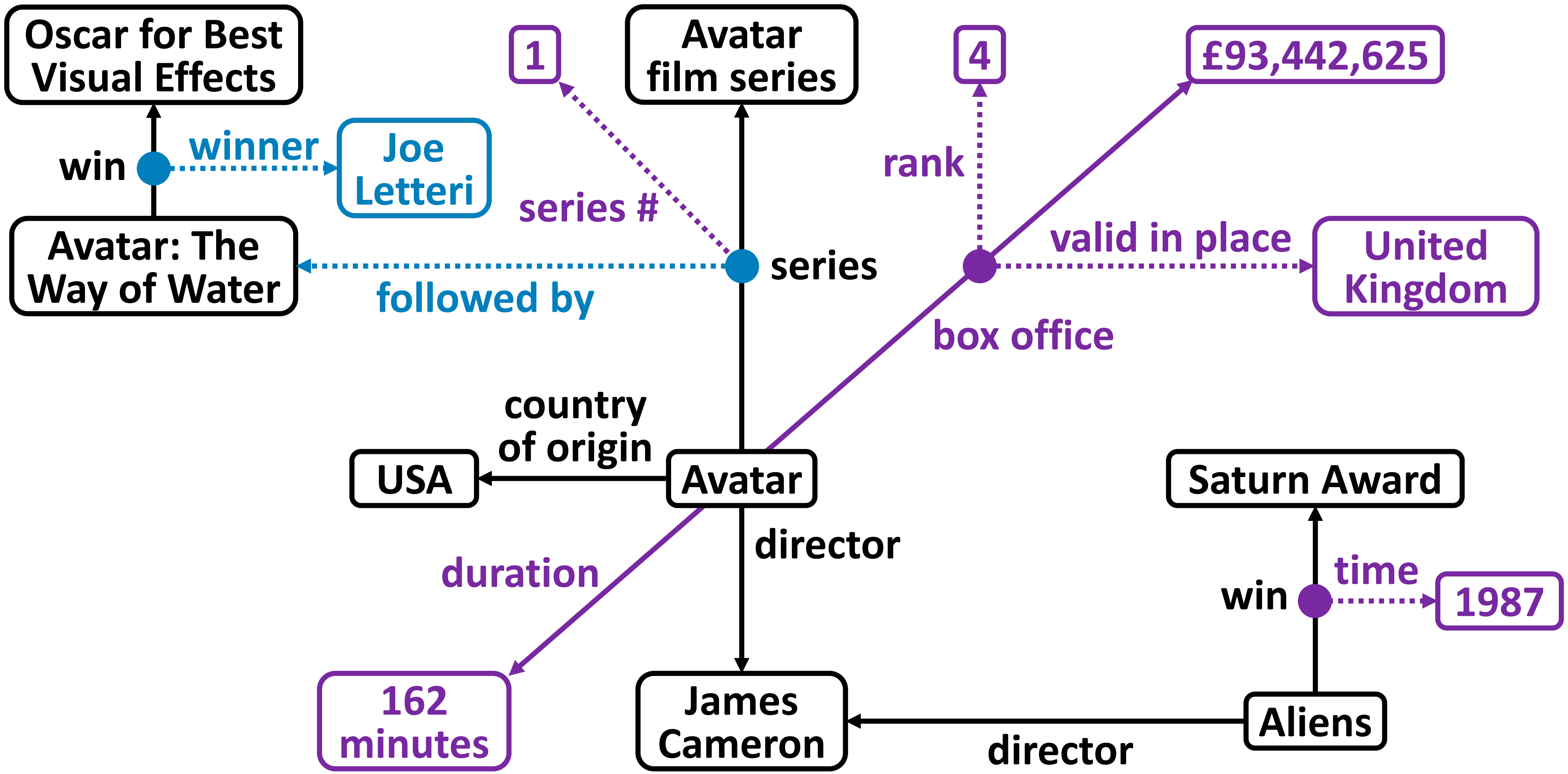
ICML 2025



03 Hyper-relational Knowledge Graphs



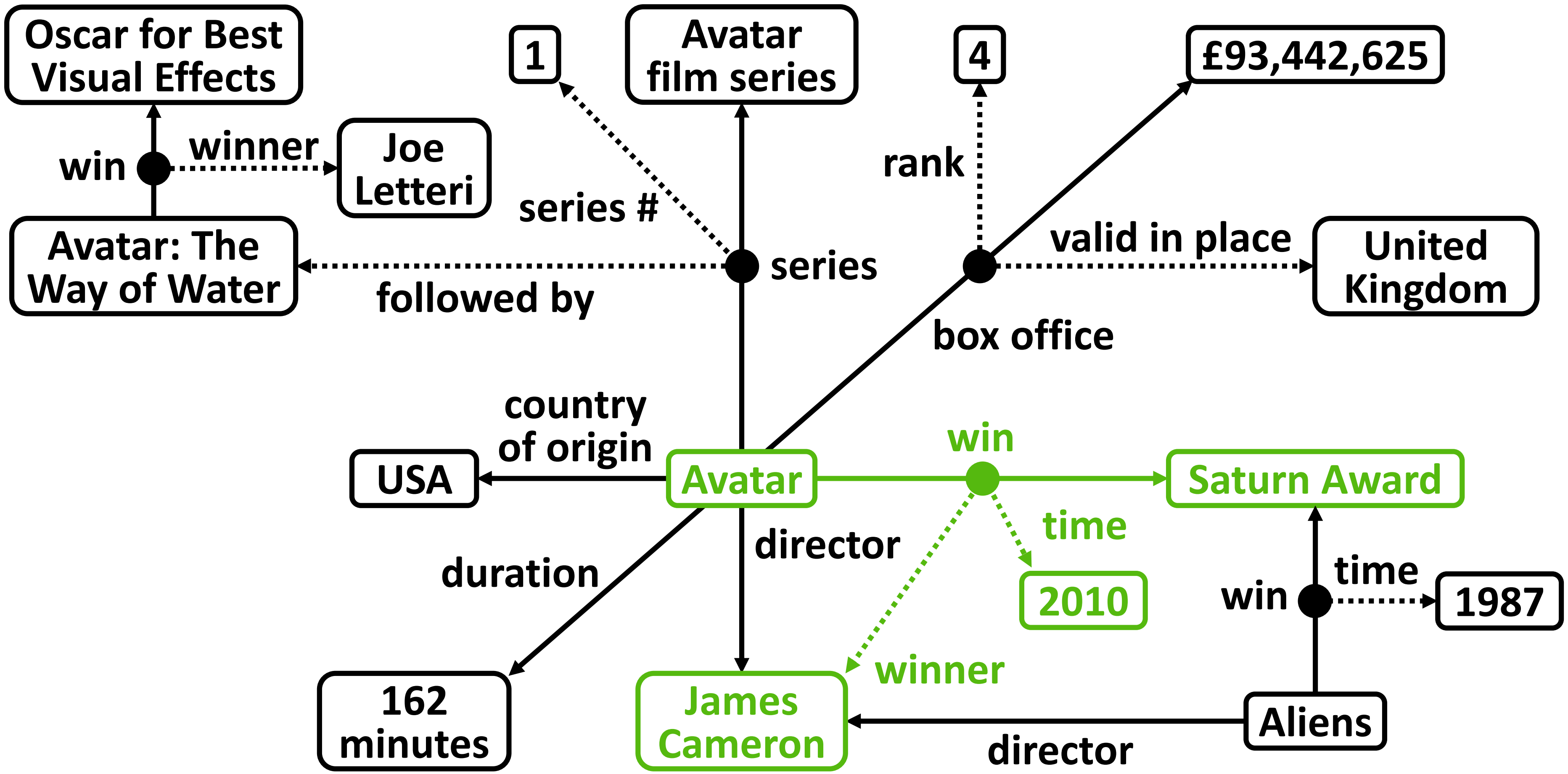
03 Hyper-relational and Numeric Knowledge Graphs



03 Hyper-relational and Numeric Knowledge Graphs

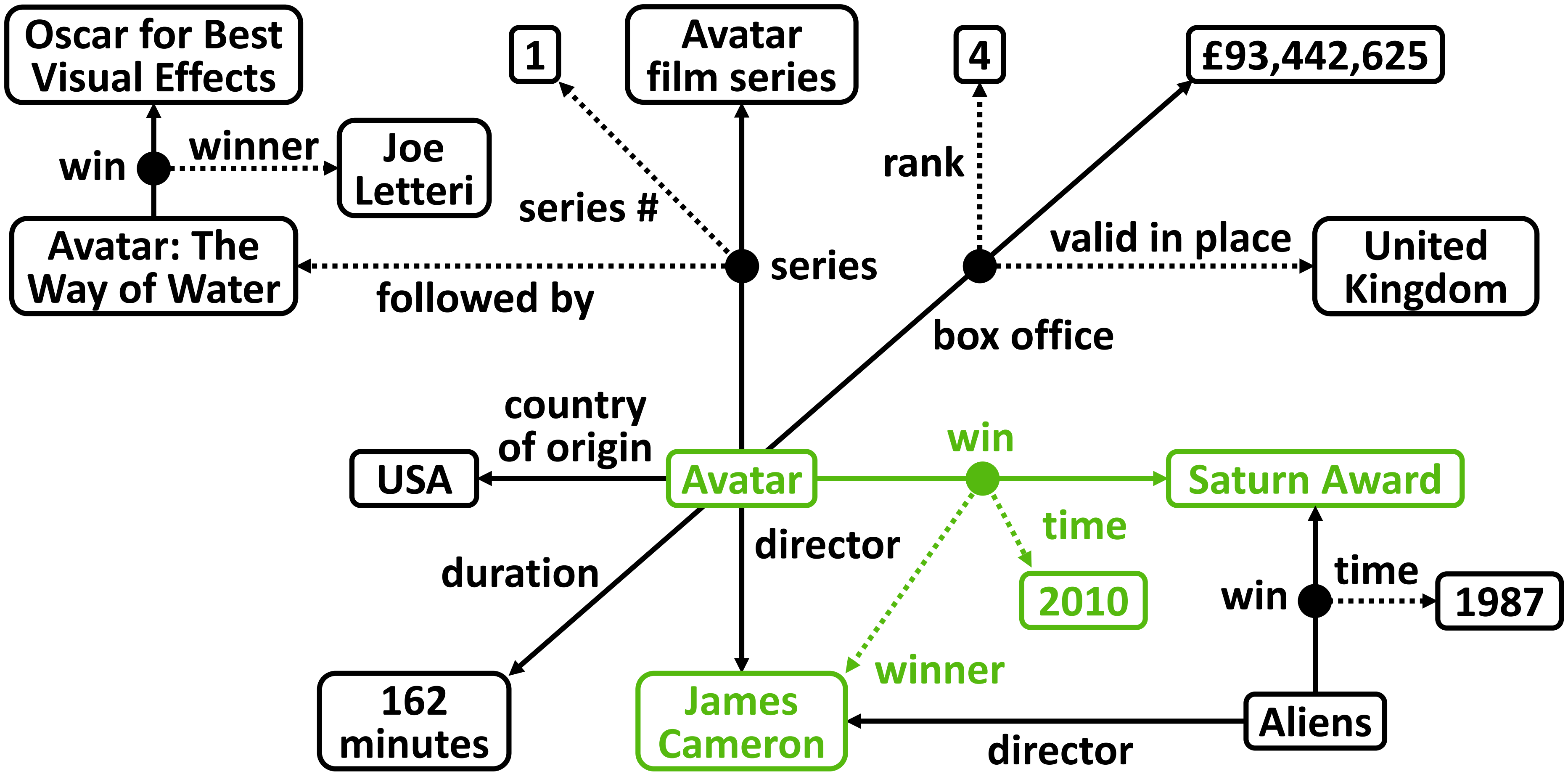
((Avatar, win, Saturn_Award), {(winner, James_Cameron), (time, 2010)})

Primary Triplet — Qualifier 1 — Qualifier 2



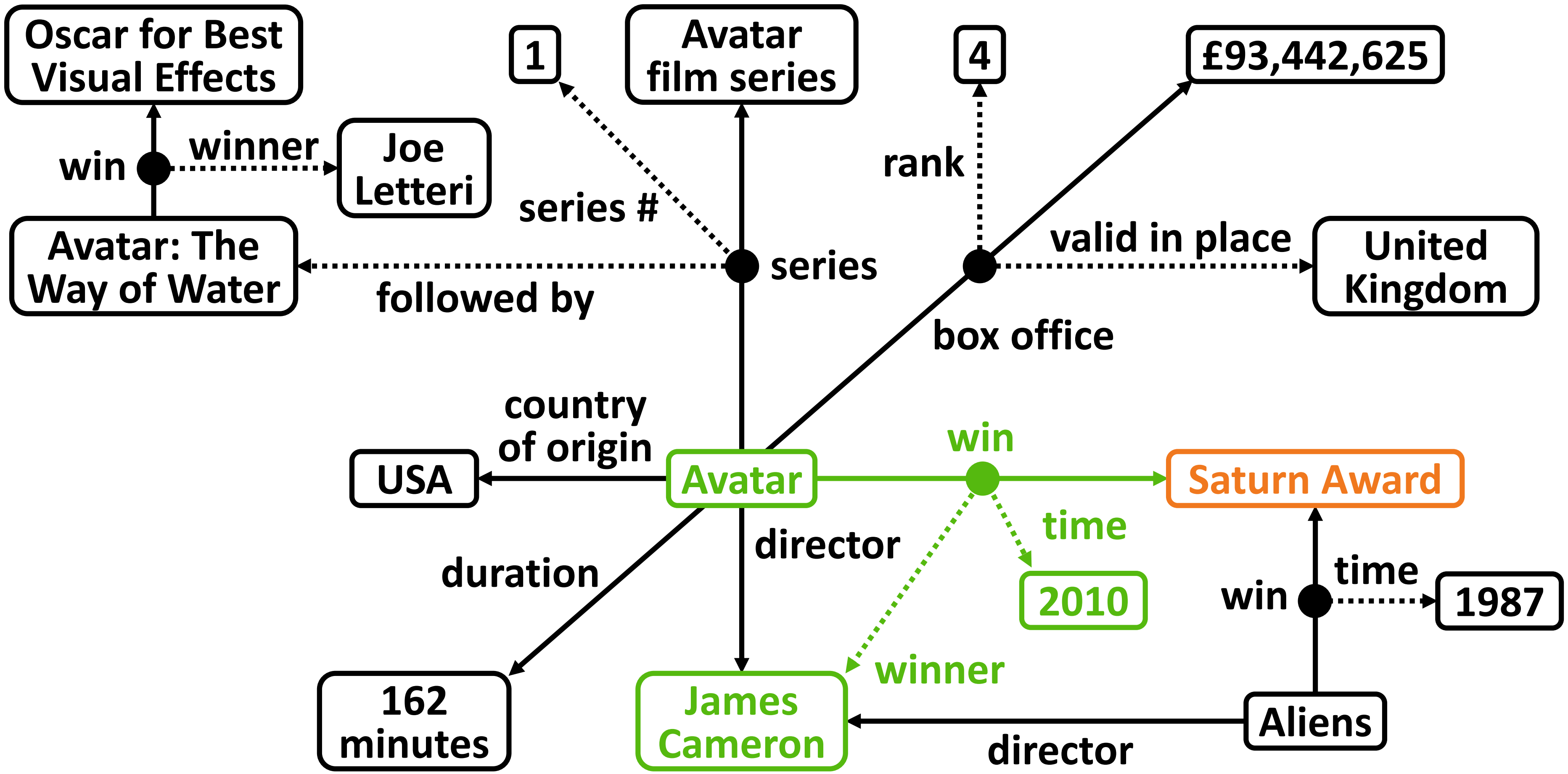
03 Link Prediction on HN-KGs

((Avatar, win, Saturn_Award), {(winner, James_Cameron), (time, 2010)})



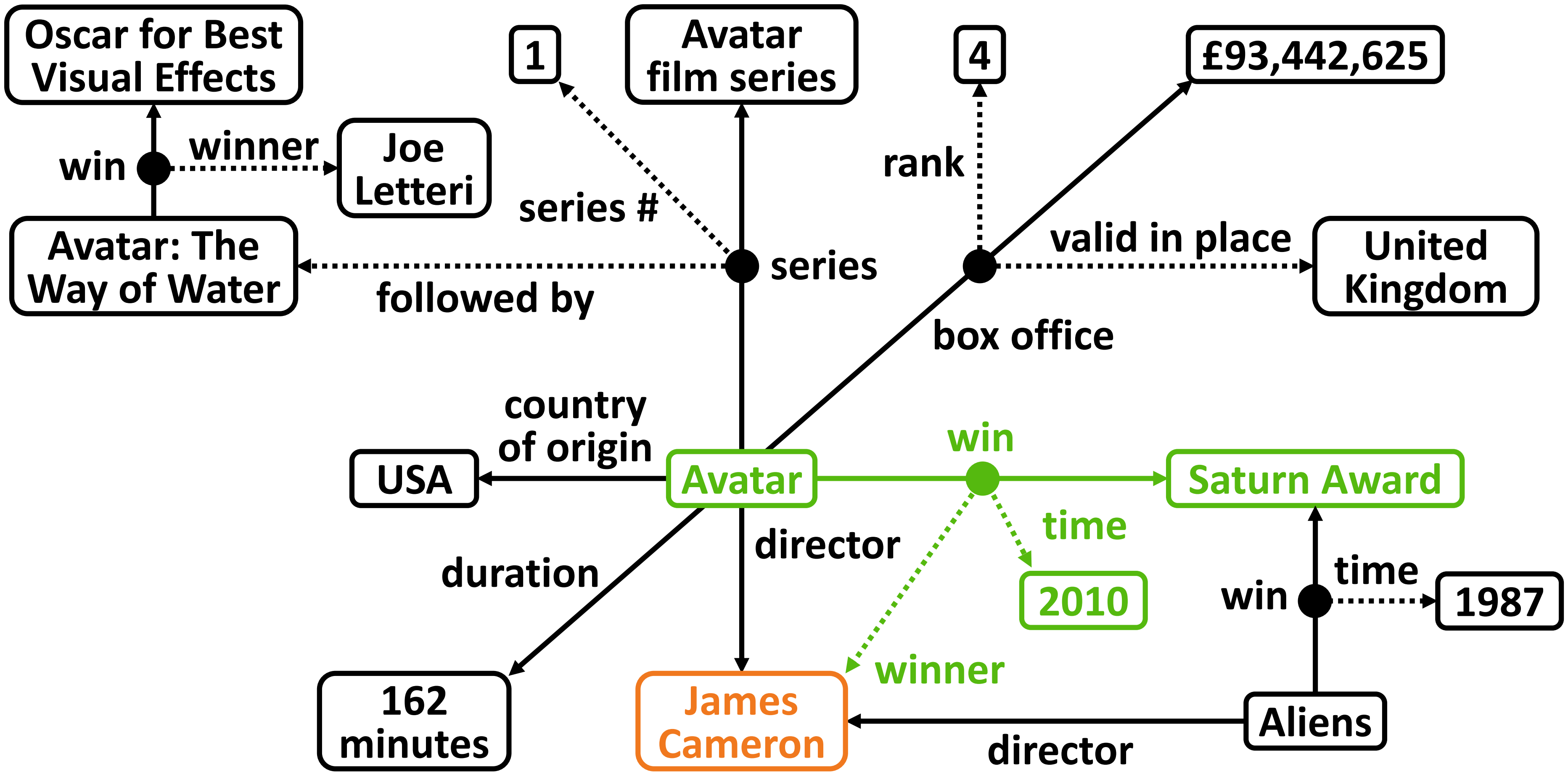
03 Link Prediction on HN-KGs

((Avatar, win, ?), {(winner, James_Cameron), (time, 2010)})



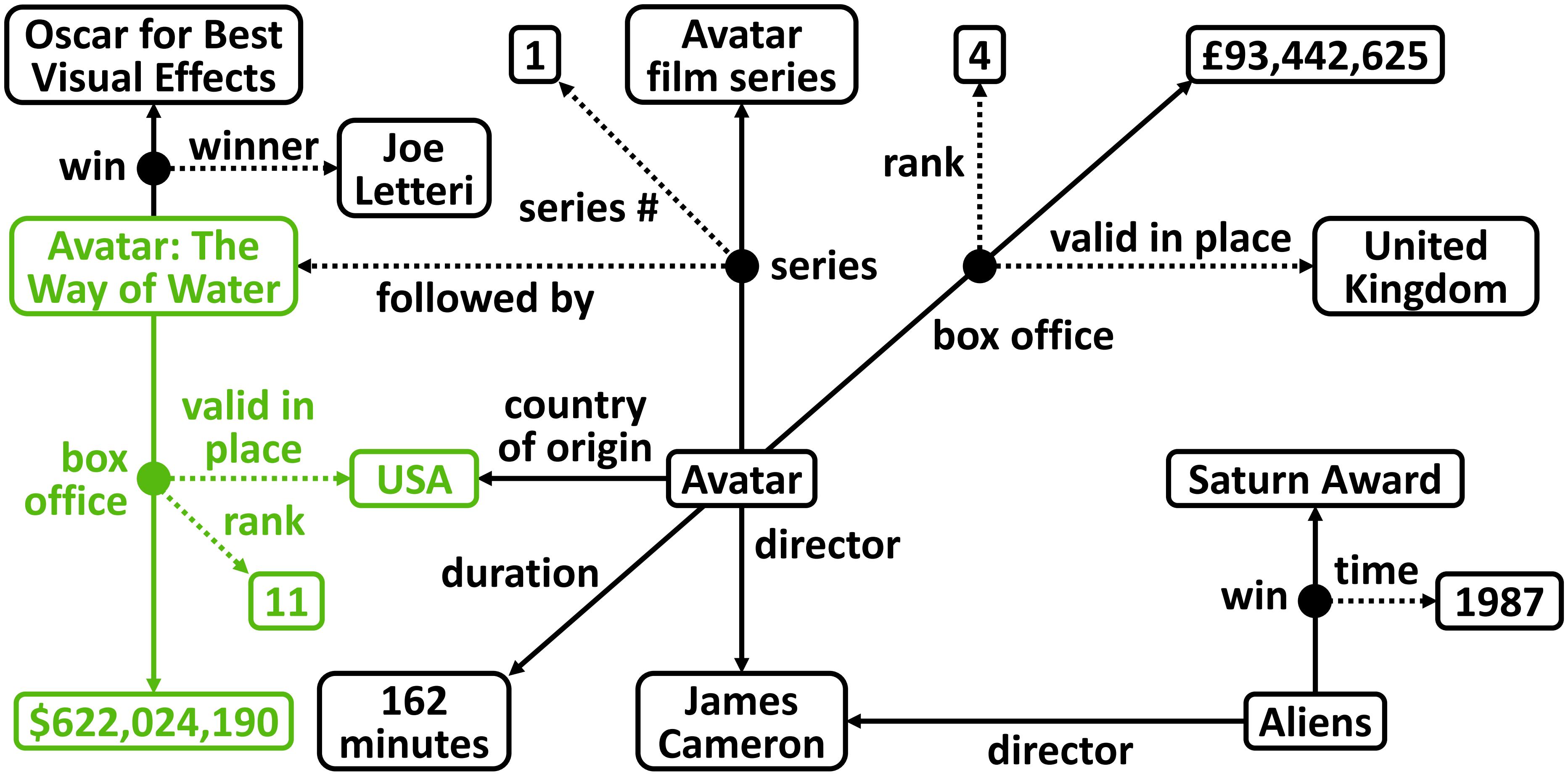
03 Link Prediction on HN-KGs

((Avatar, win, Saturn_Award), {(winner, ?), (time, 2010)})



03 Numeric Value Prediction on HN-KGs

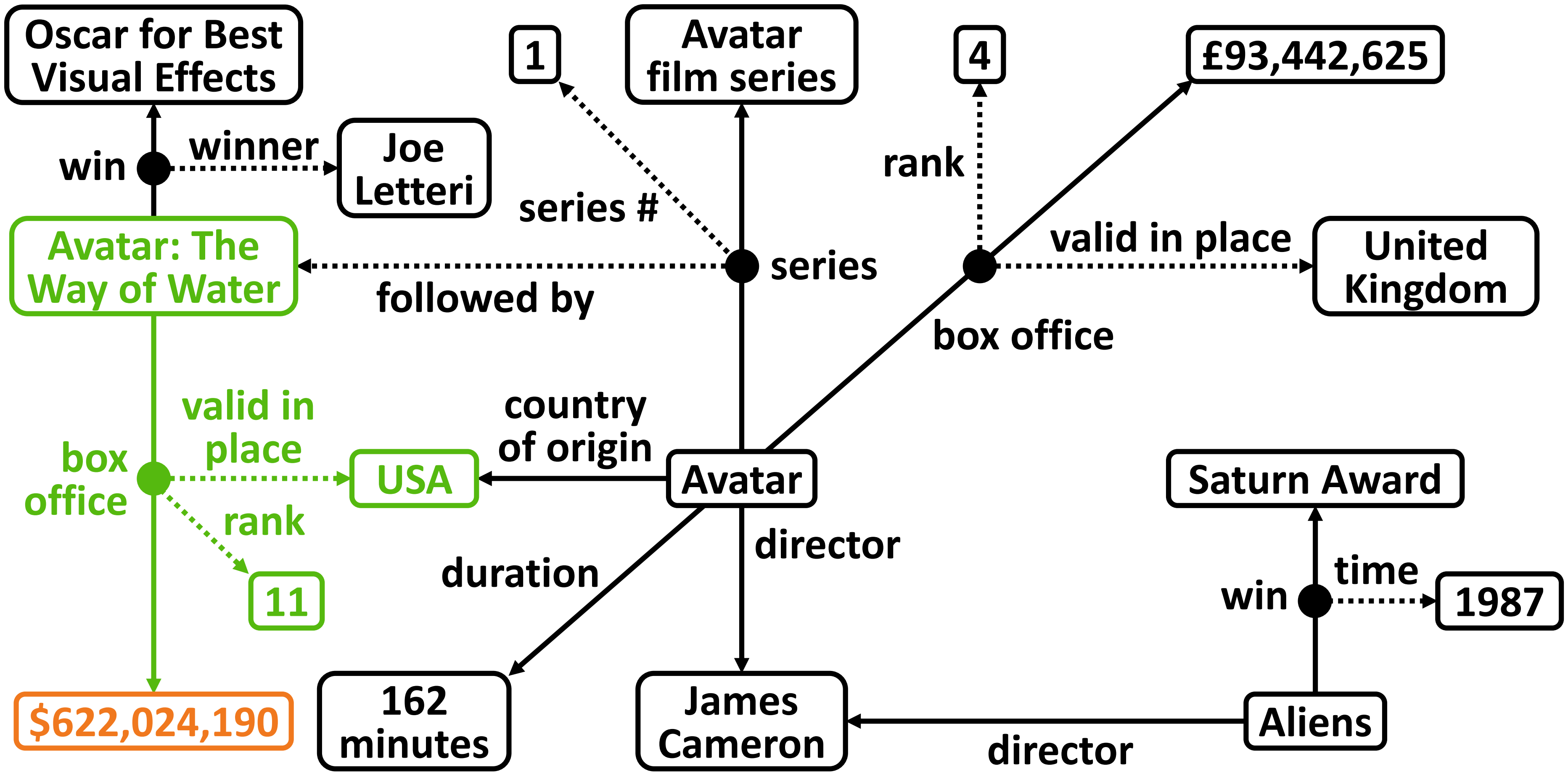
((Avatar:The_Way_of_Water, box_office, $622,024,190), {(rank, 11), (valid_in_place, USA)})



03

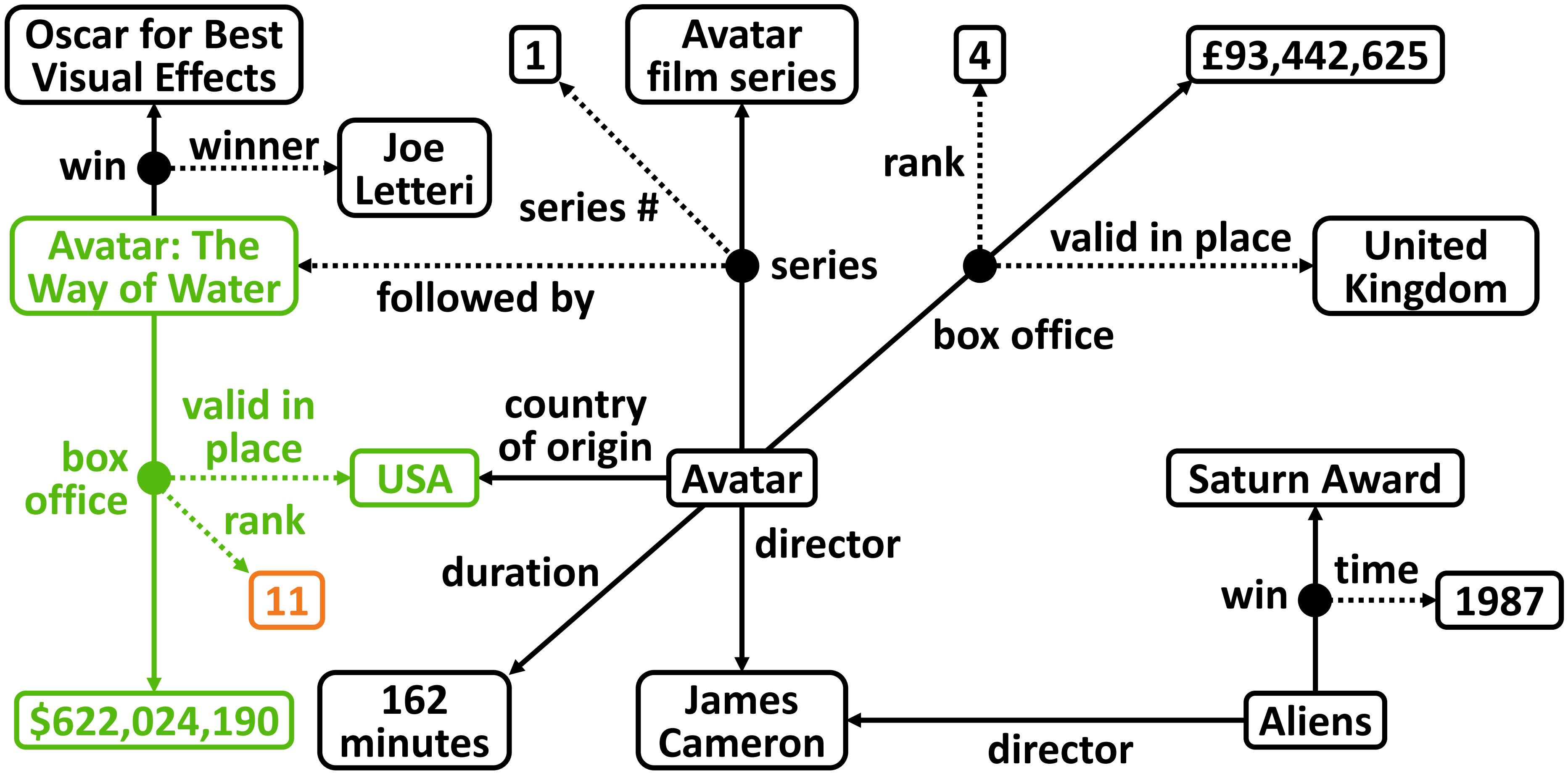
Numeric Value Prediction on HN-KGs

((Avatar:The_Way_of_Water, box_office, ?), {(rank, 11), (valid_in_place, USA)})



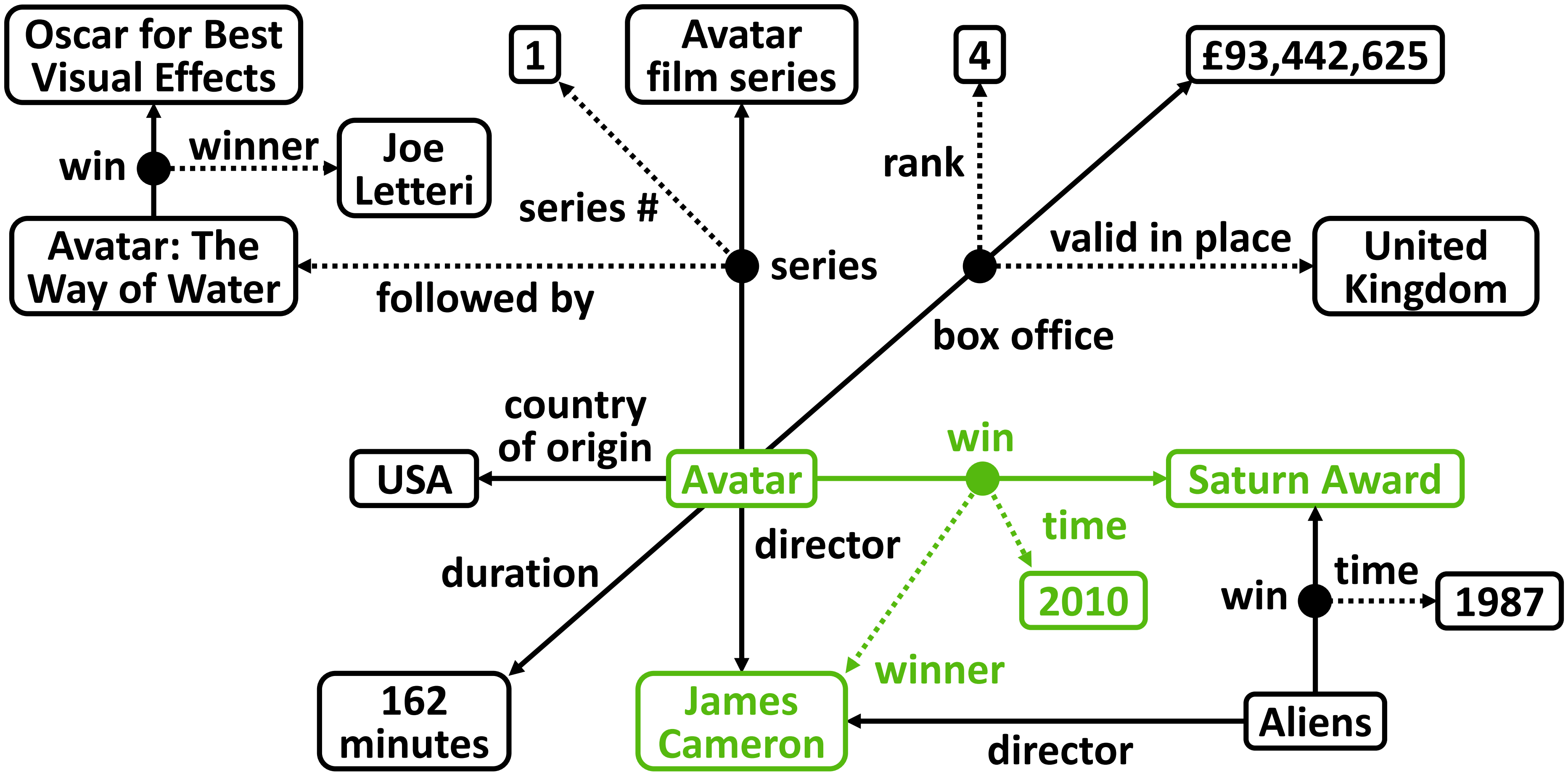
03 Numeric Value Prediction on HN-KGs

((Avatar:The_Way_of_Water, box_office, $622,024,190), {(rank, ?), (valid_in_place, USA)})



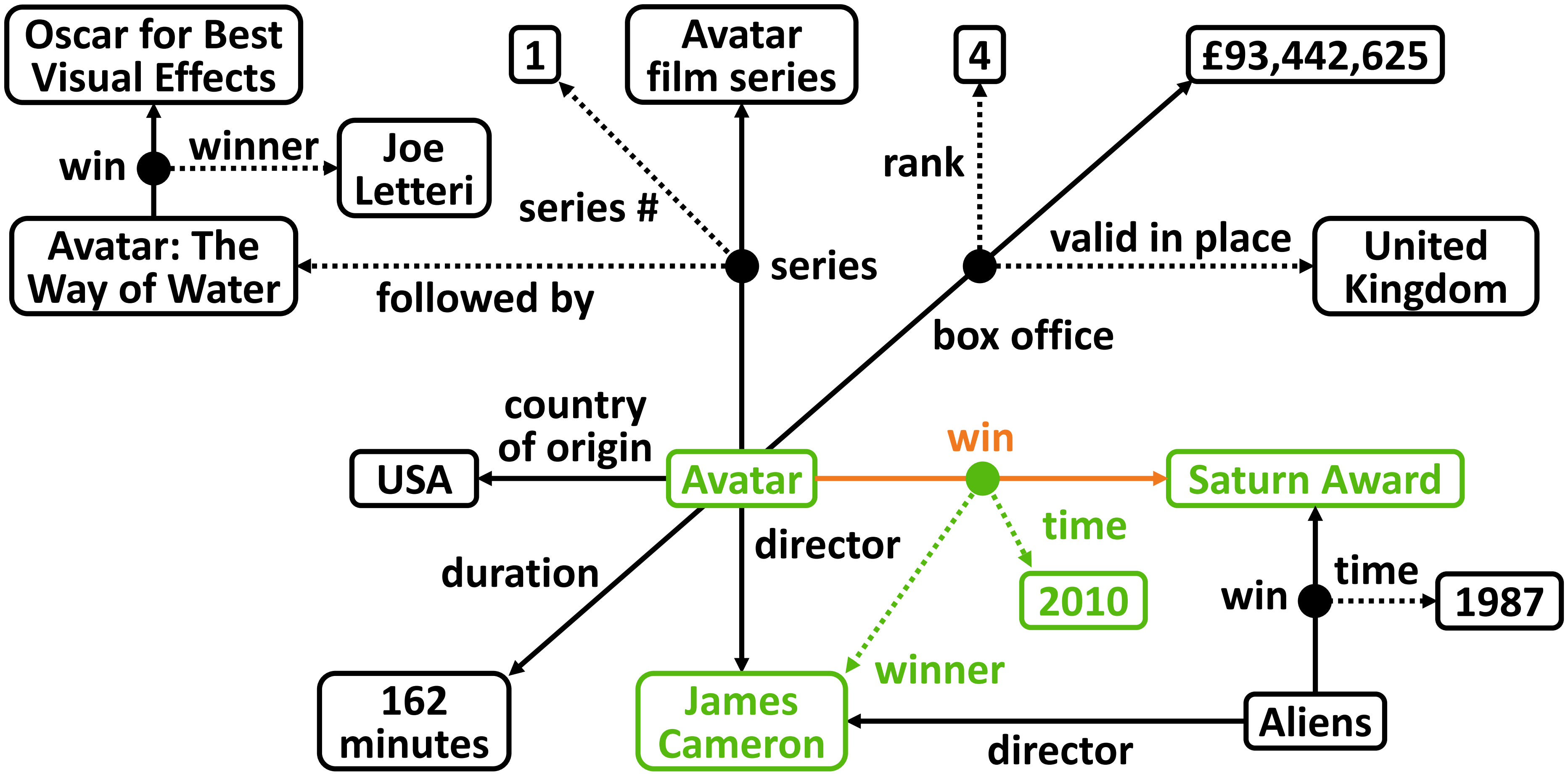
03 Relation Prediction on HN-KGs

((Avatar, win, Saturn_Award), {(winner, James_Cameron), (time, 2010)})



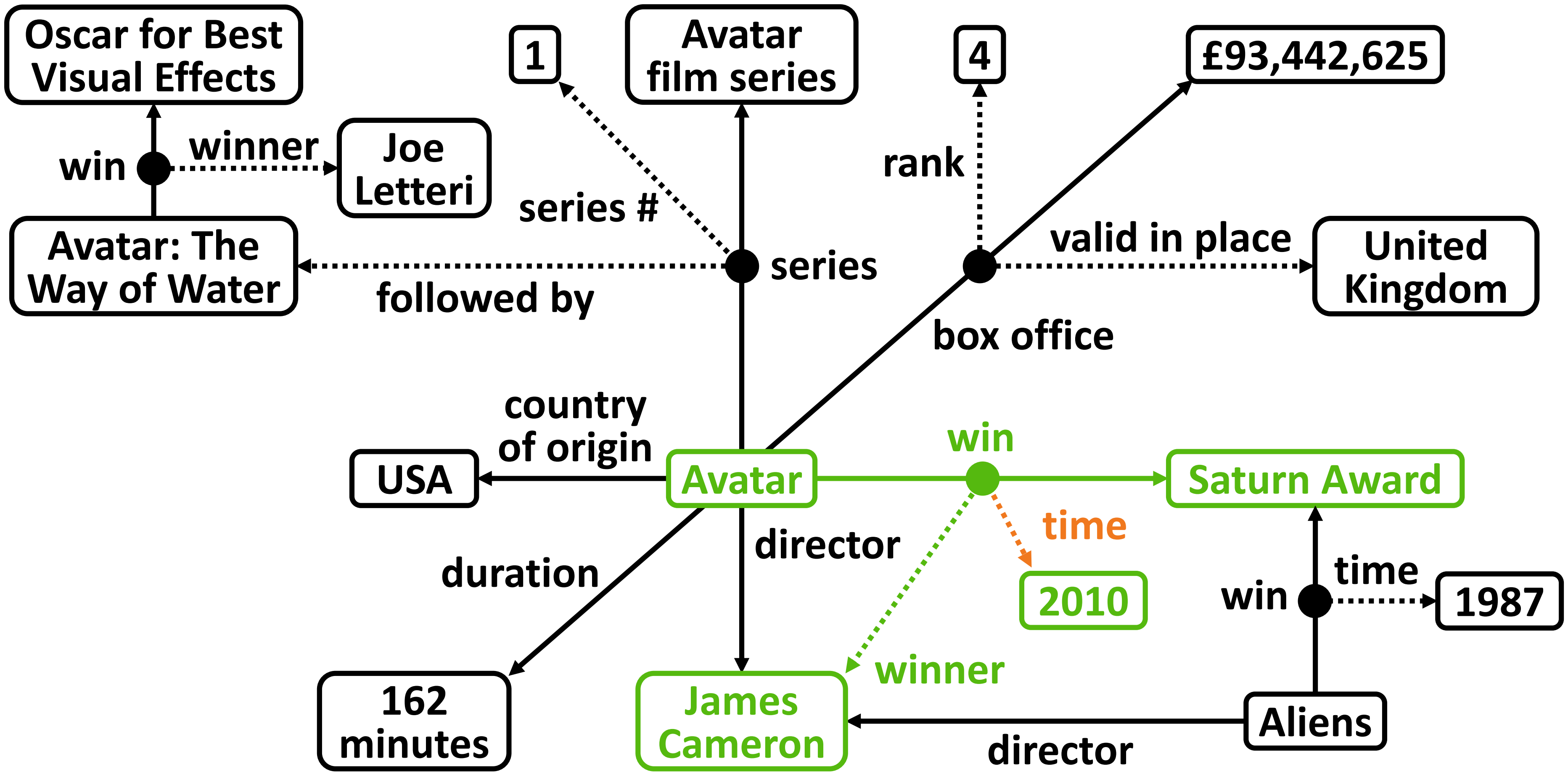
03 Relation Prediction on HN-KGs

((Avatar, ? , Saturn_Award), {(winner, James_Cameron), (time, 2010)})



03 Relation Prediction on HN-KGs

((Avatar, win, Saturn_Award), {(winner, James_Cameron), (? , 2010)})



03 Contributions

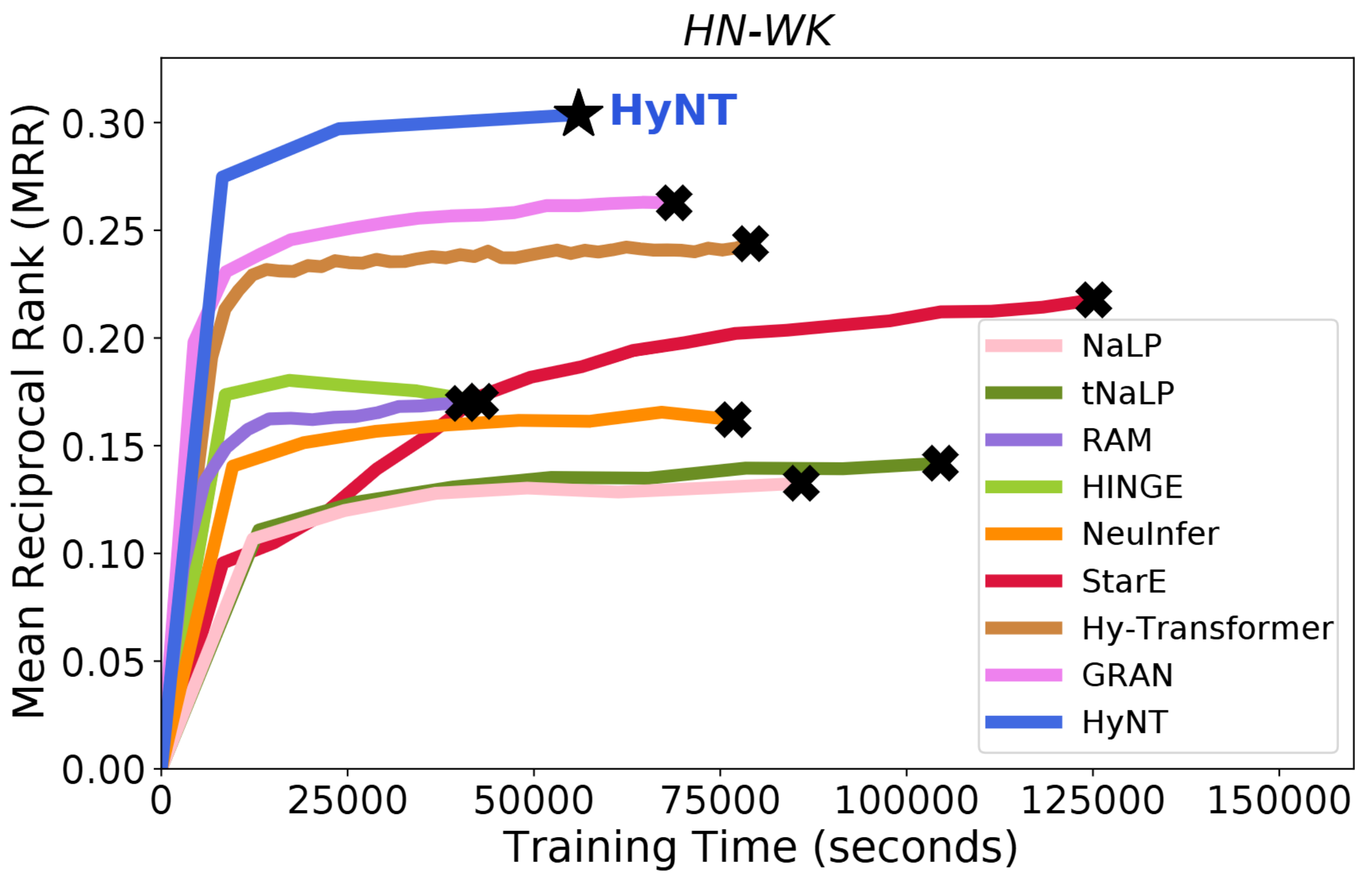
- Define **Hyper-relational and Numeric Knowledge Graphs**
  - Create 4 real-world HN-KG datasets

- Propose **HyNT**, **Hy**per-relational knowledge graph embedding with **N**umeric literals using **T**ransformers
  - Define a context transformer and a prediction transformer
  - Reduce the cost by learning compact representations of triplets and qualifiers

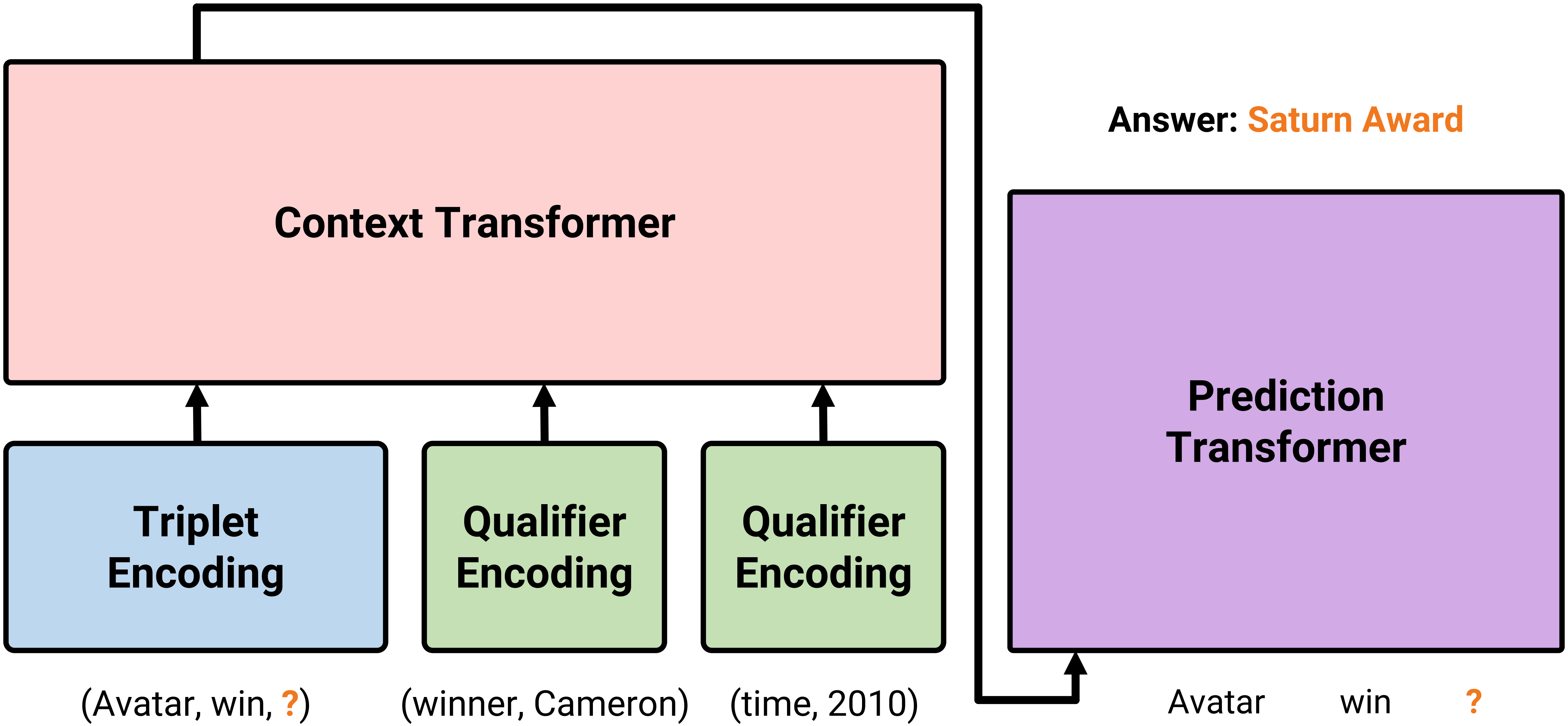
- HyNT significantly outperforms 12 different state-of-the-art methods for **link prediction**, **numeric value prediction**, and **relation prediction**

03 Contributions

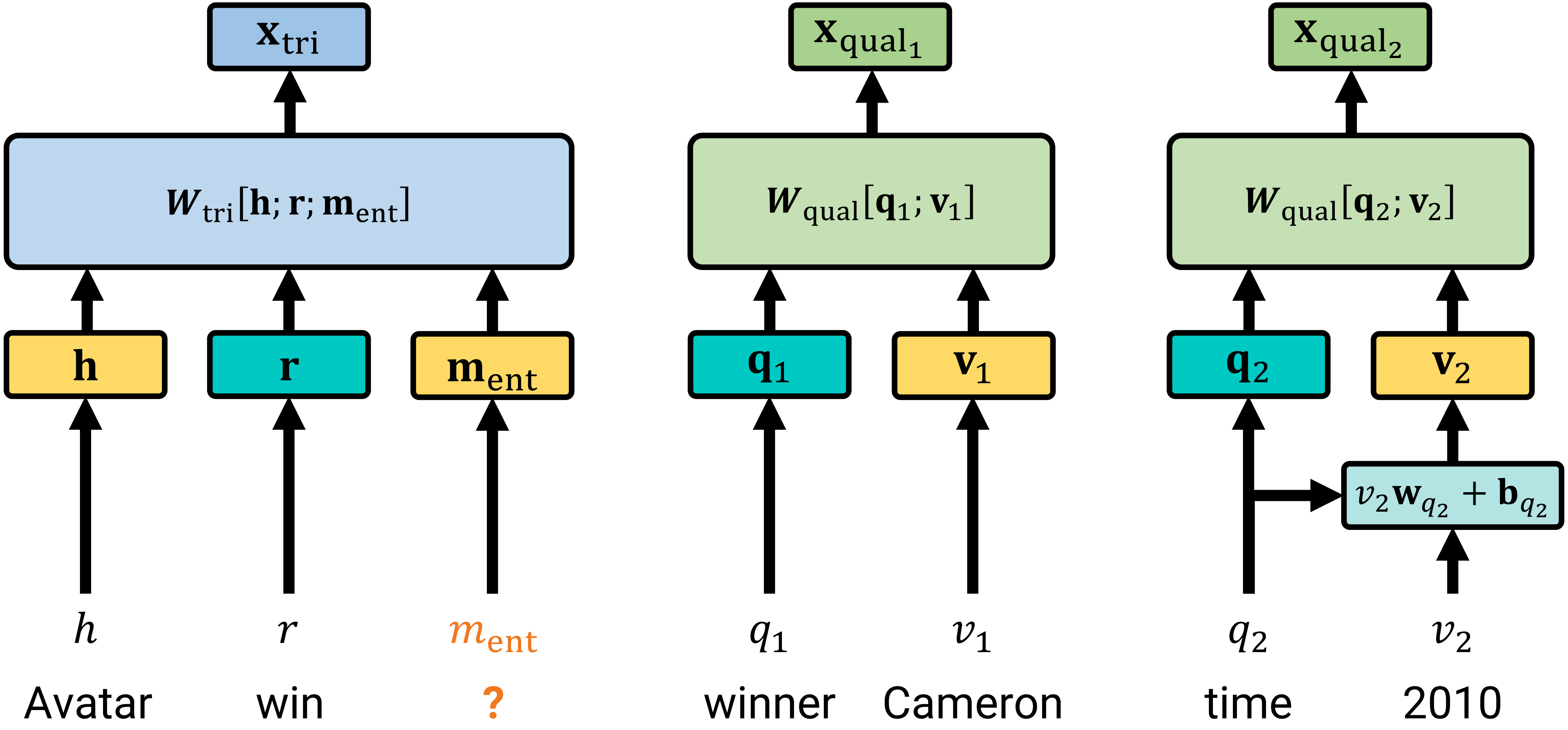
- Link Prediction Performance vs. Training Time



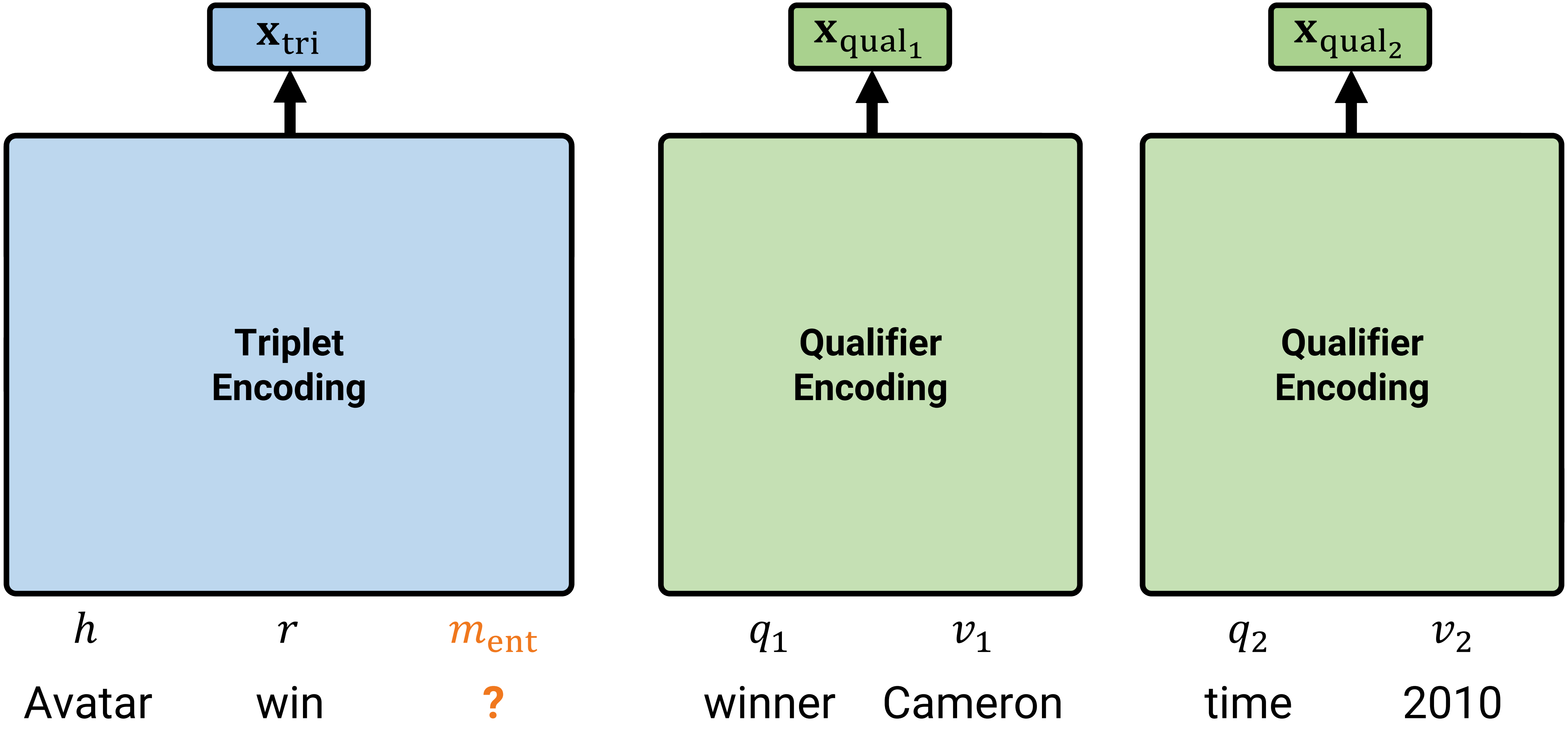
03 Overview of HyNT



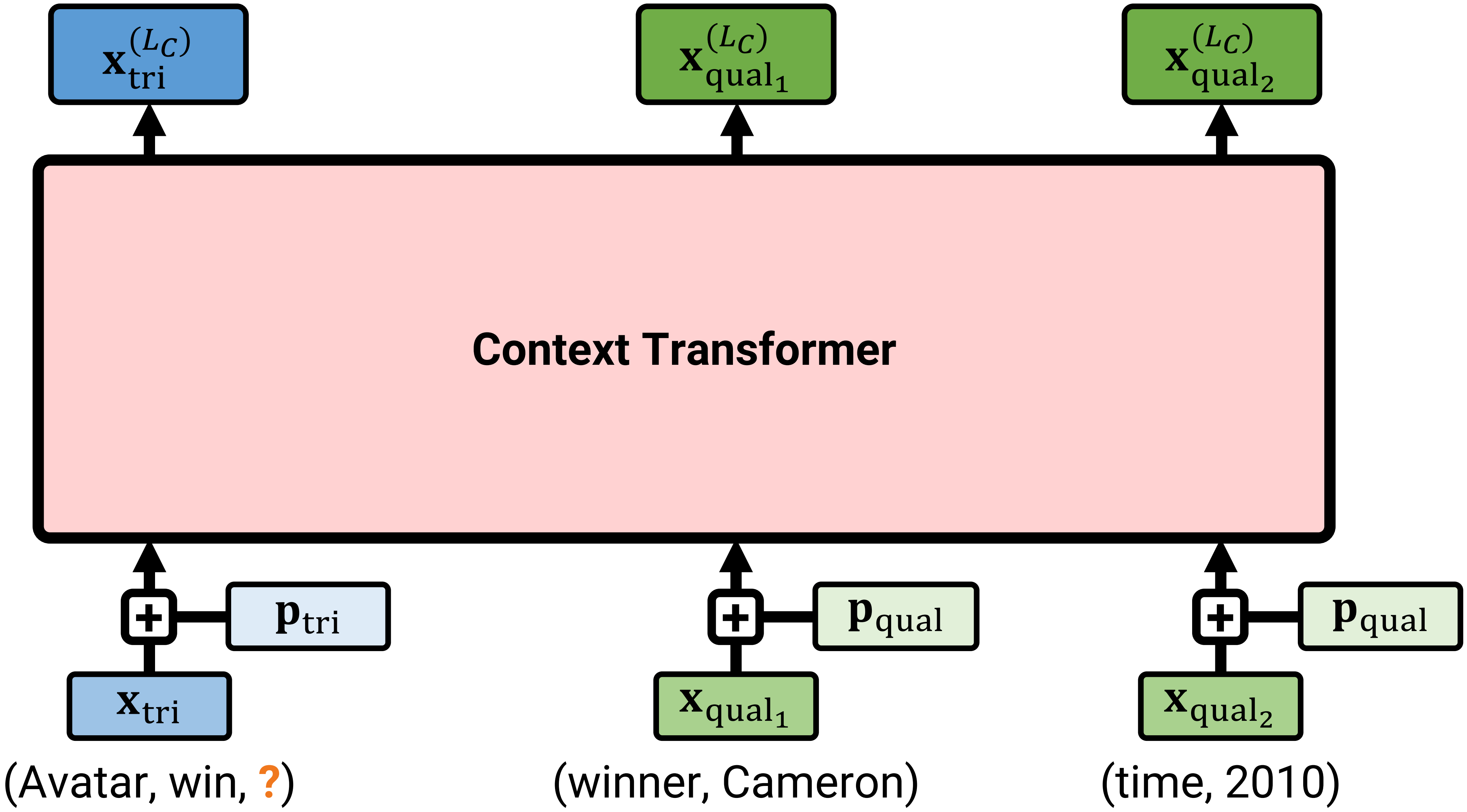
03 Triplet/Qualifier Encoding



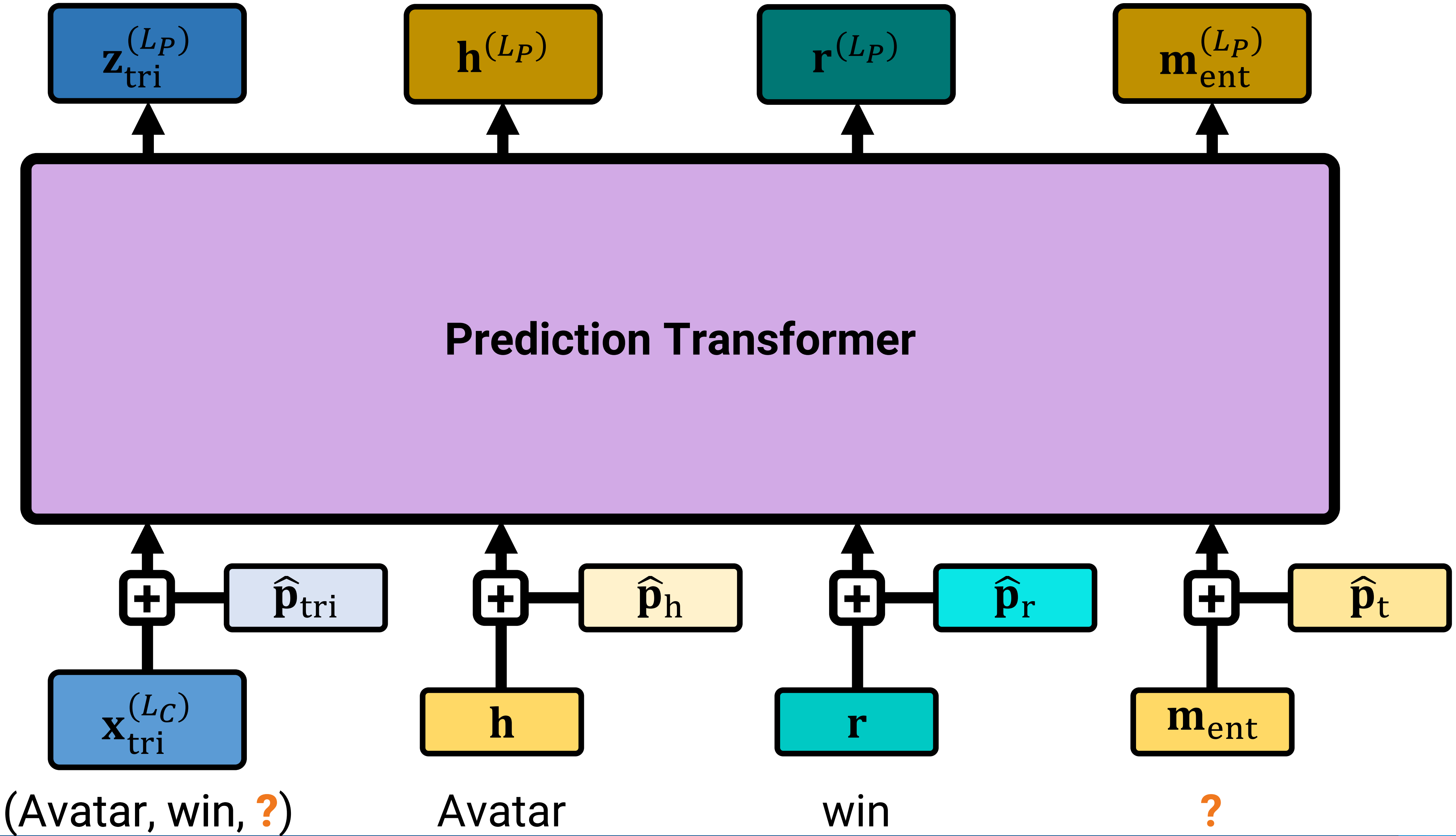
Triplet/Qualifier Encoding



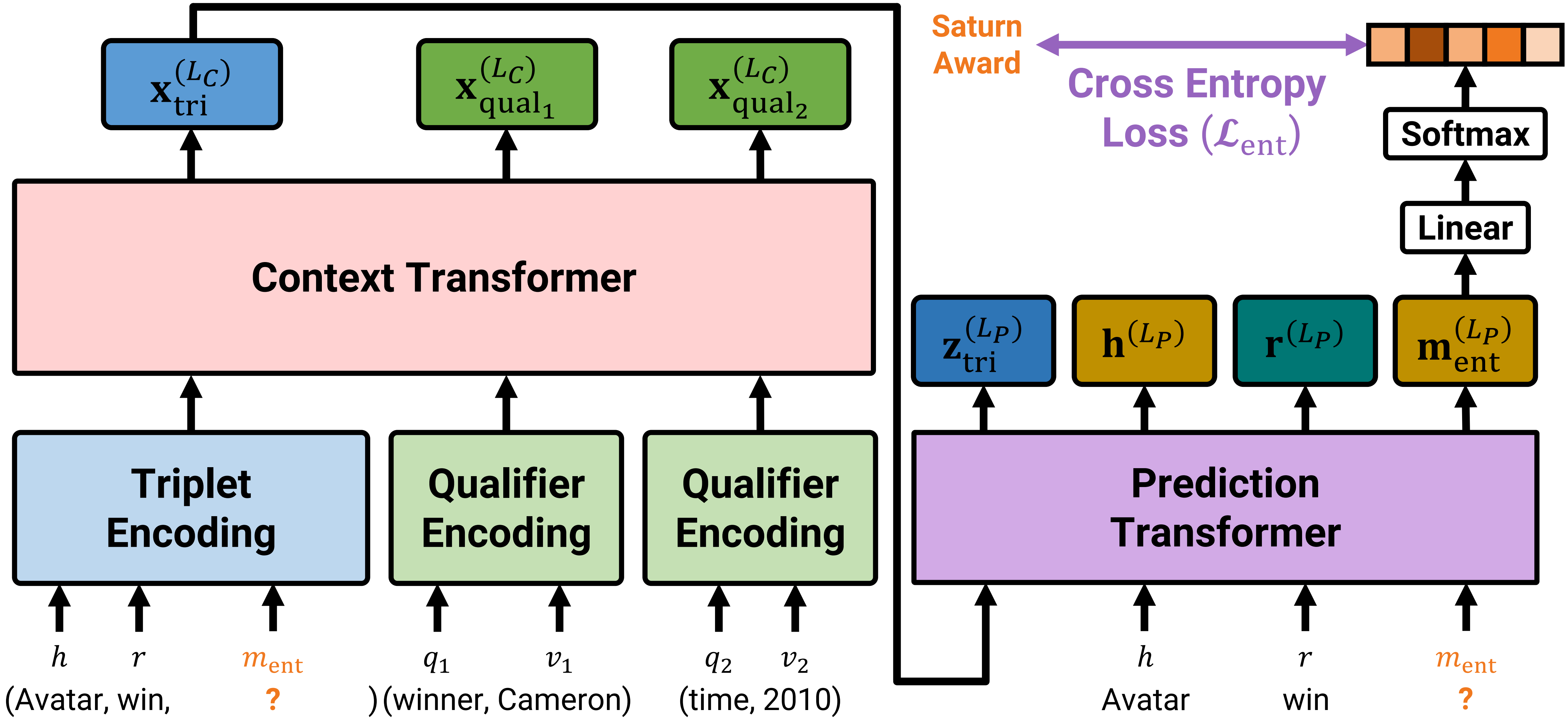
Context Transformer



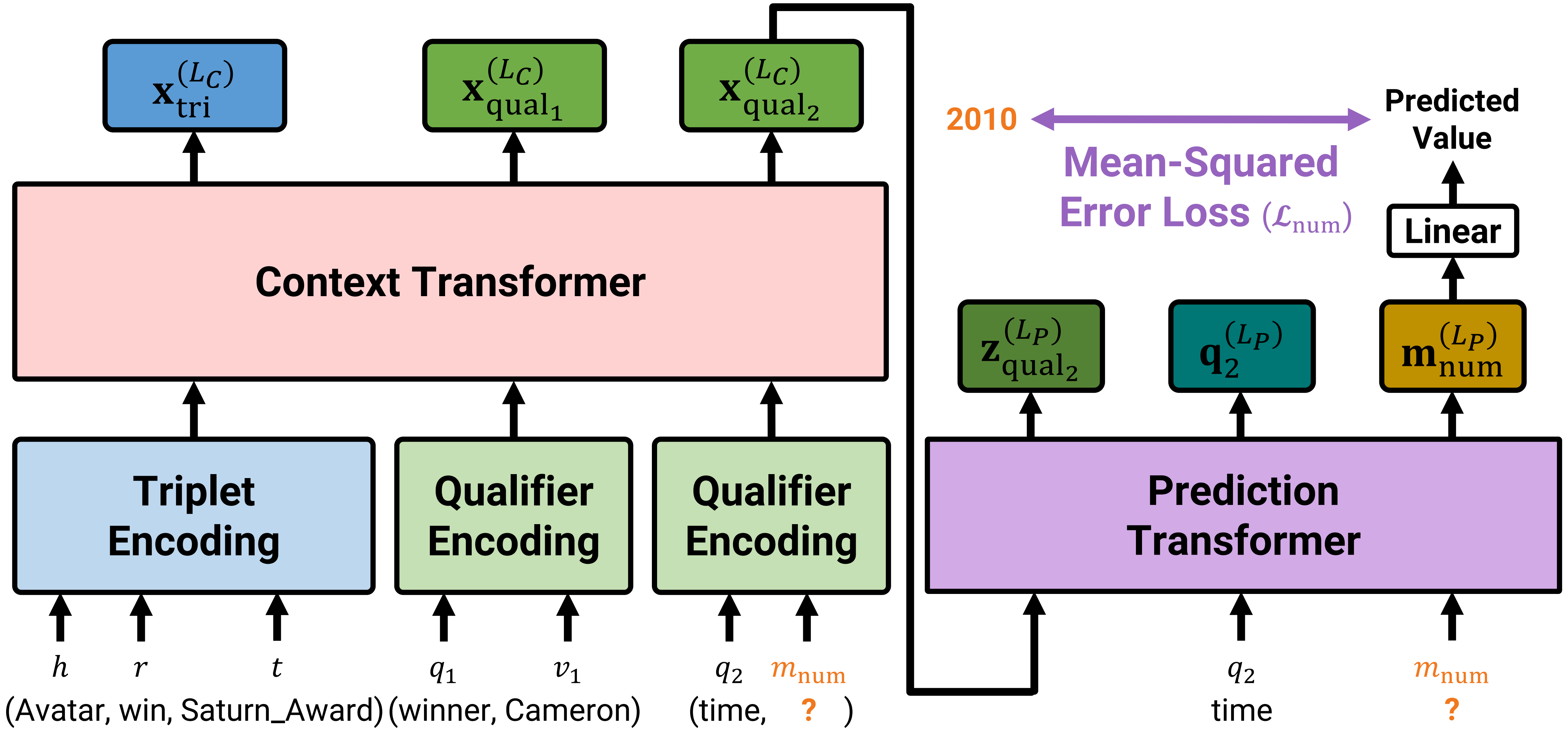
Prediction Transformer



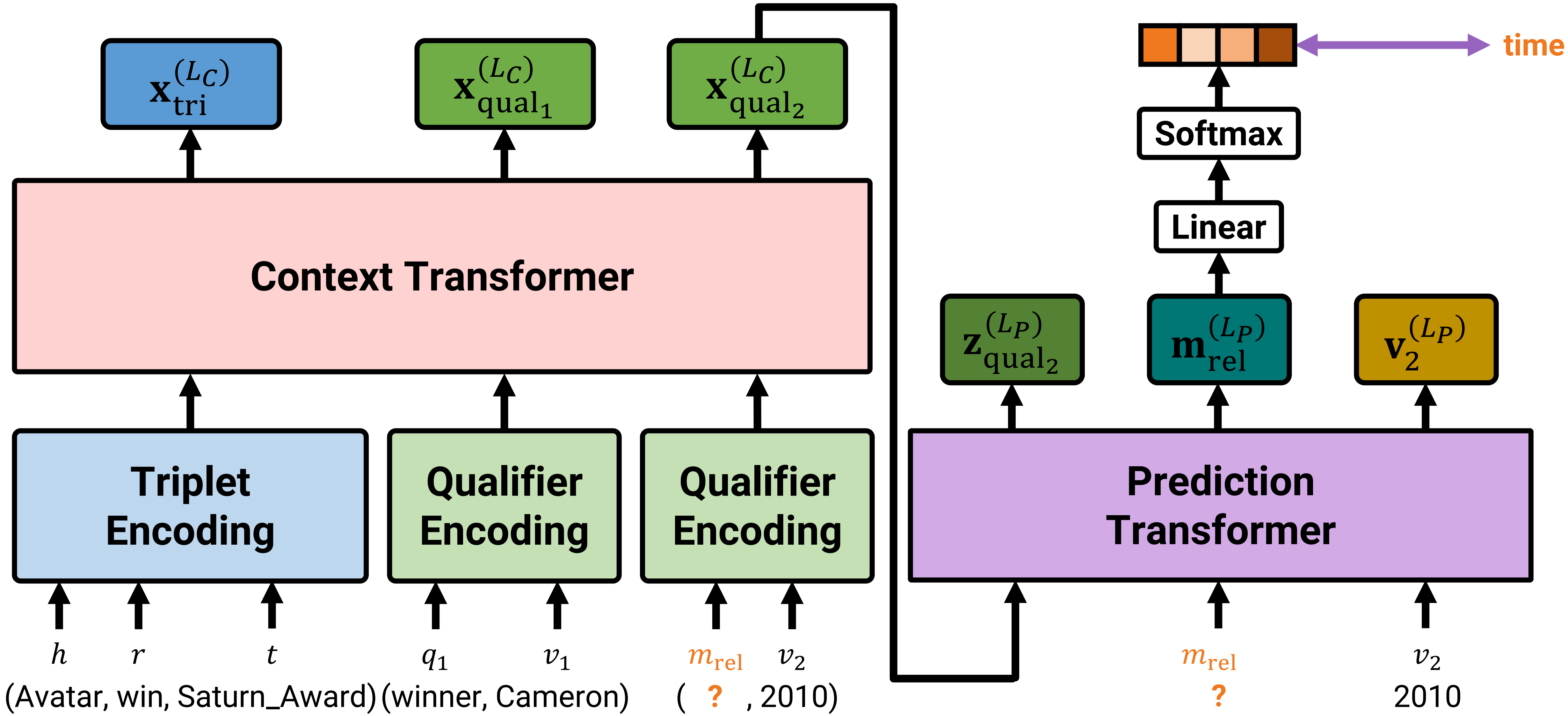
03 Link Prediction using HyNT



03 Numeric Value Prediction using HyNT

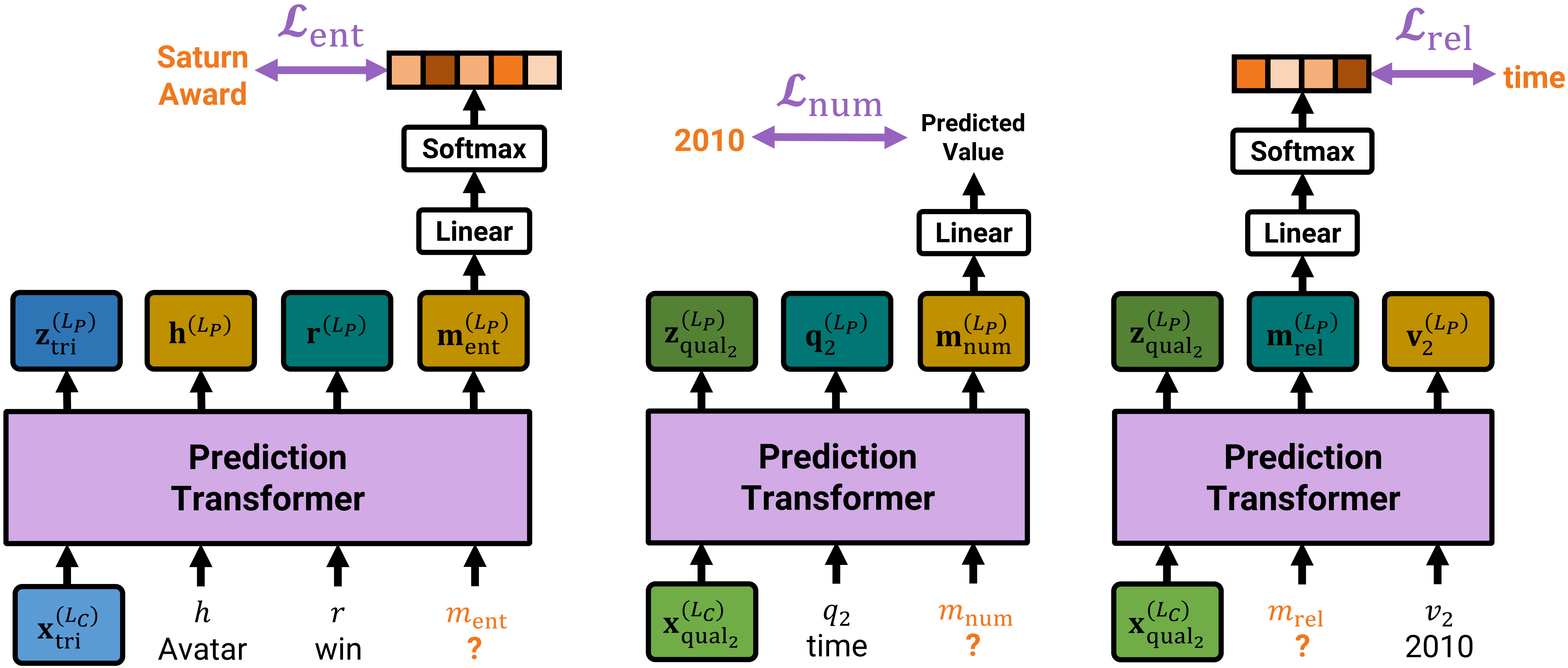


03 Relation Prediction using HyNT



03 Loss of HyNT

$$\mathcal{L} \coloneqq \mathcal{L}_{\mathrm{ent}} + \lambda_1 \cdot \mathcal{L}_{\mathrm{rel}} + \lambda_2 \cdot \mathcal{L}_{\mathrm{num}}$$

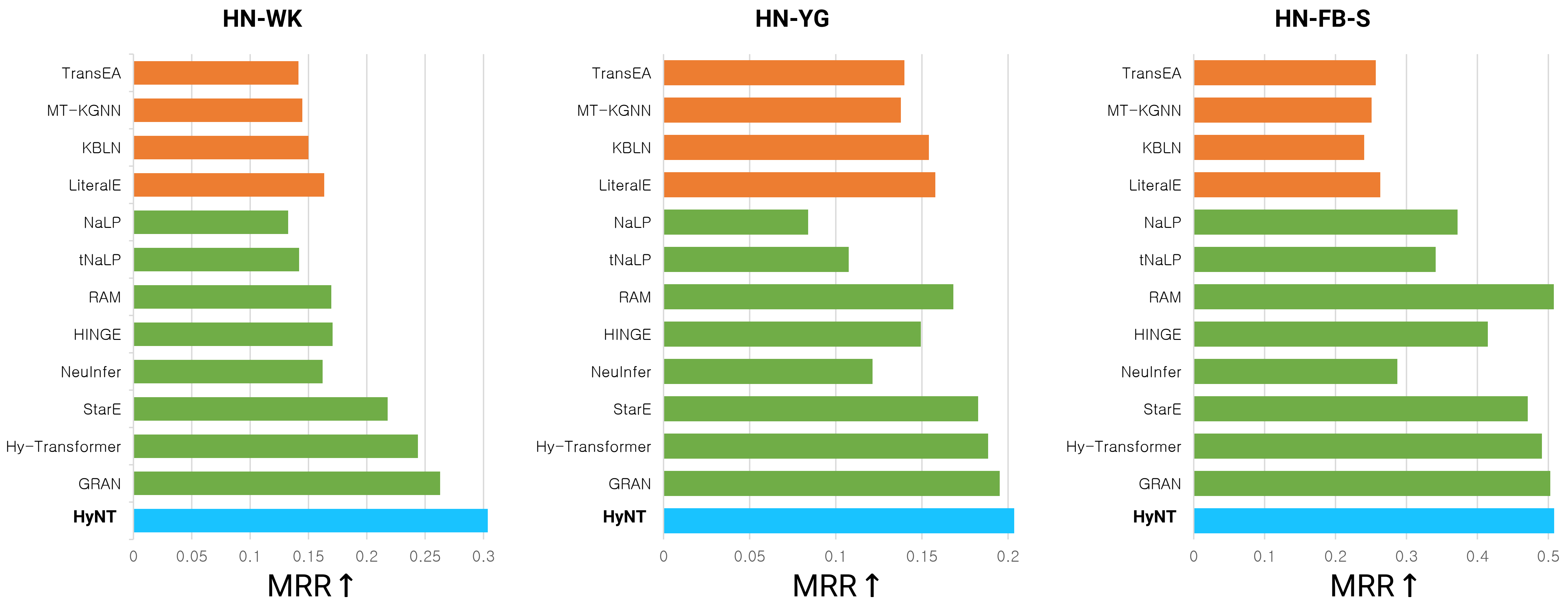


03 Experimental Results

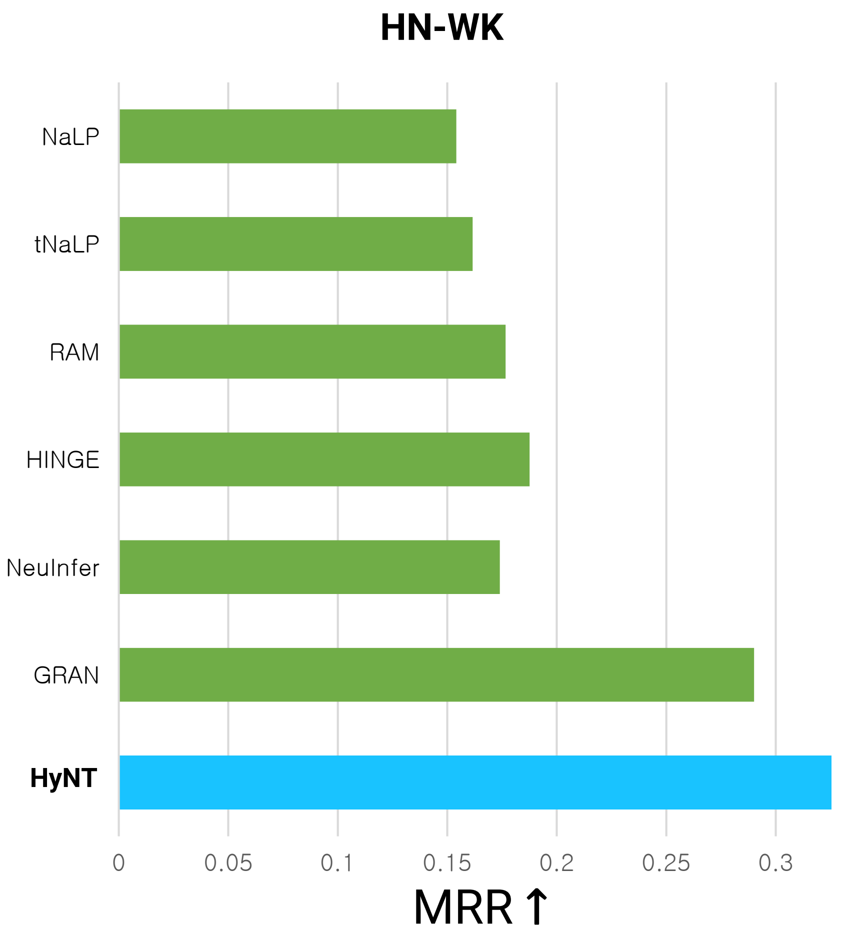
- Datasets
  - Based on Wikidata, YAGO, and Freebase
  - Create **4 Hyper-relational and Numeric Knowledge Graph (HN-KG)** datasets
    - HN-WK, HN-YG, HN-FB, HN-FB-S
- Comparison with **12 baseline methods**
  - Methods for handling numeric literals
    - TransEA, MT-KGNN, KBLN, LiteralE
  - Methods for handling hyper-relational facts
    - NaLP, tNaLP, RAM, HINGE, NeuInfer, StarE, Hy-Transformer, GRAN

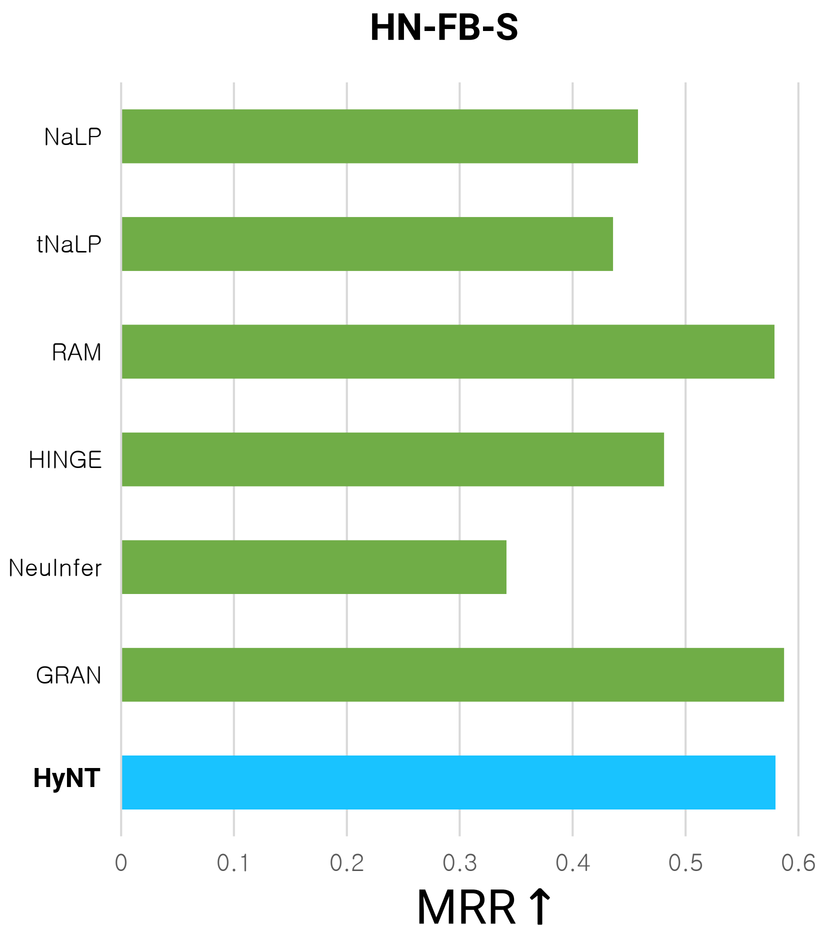
03

Link Prediction Results – Primary

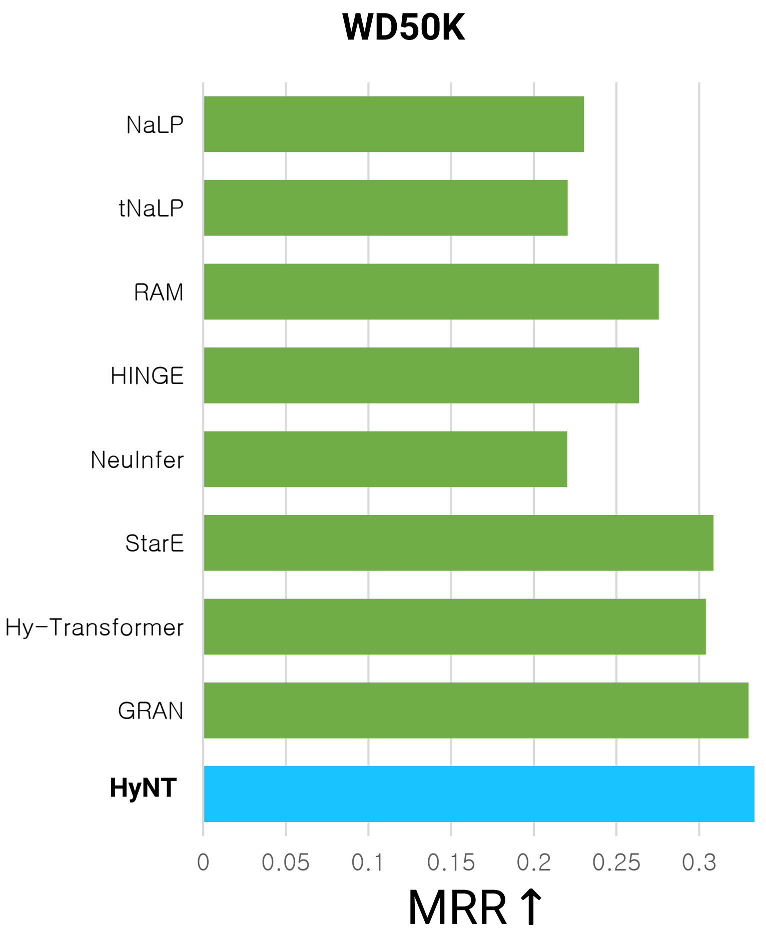


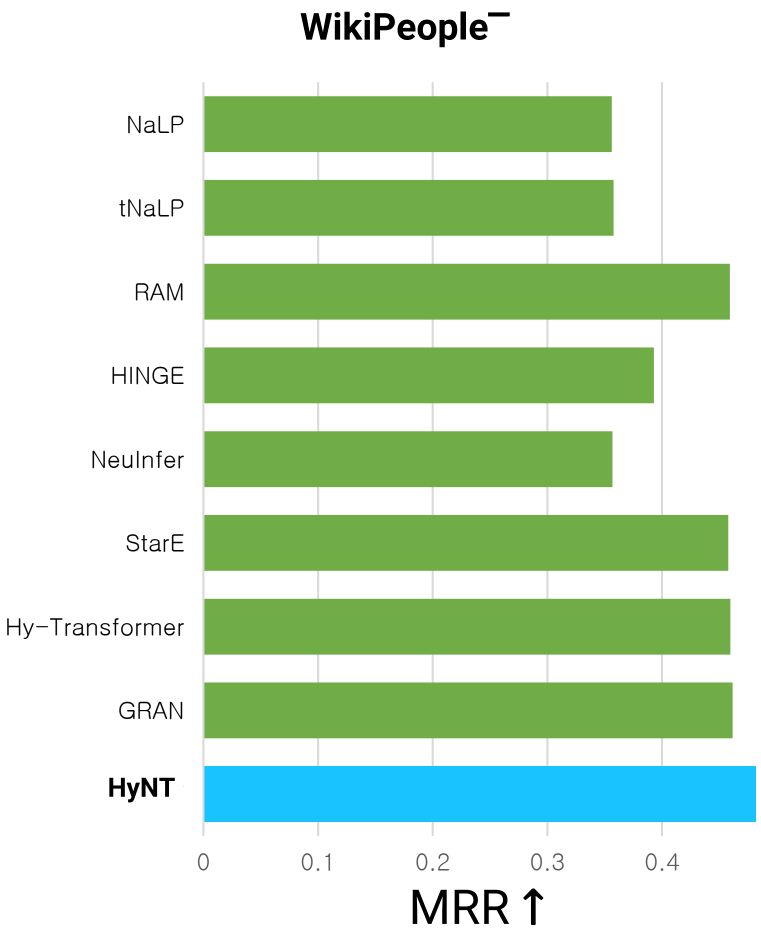
03 Link Prediction Results – All





Link Prediction Results – Primary (Benchmark Datasets)





03 Link Prediction Results of HyNT

((?, nominated_for, Best_Actor), {(for_work, Moneyball), (subject_of, 84^{th}_Oscars)})

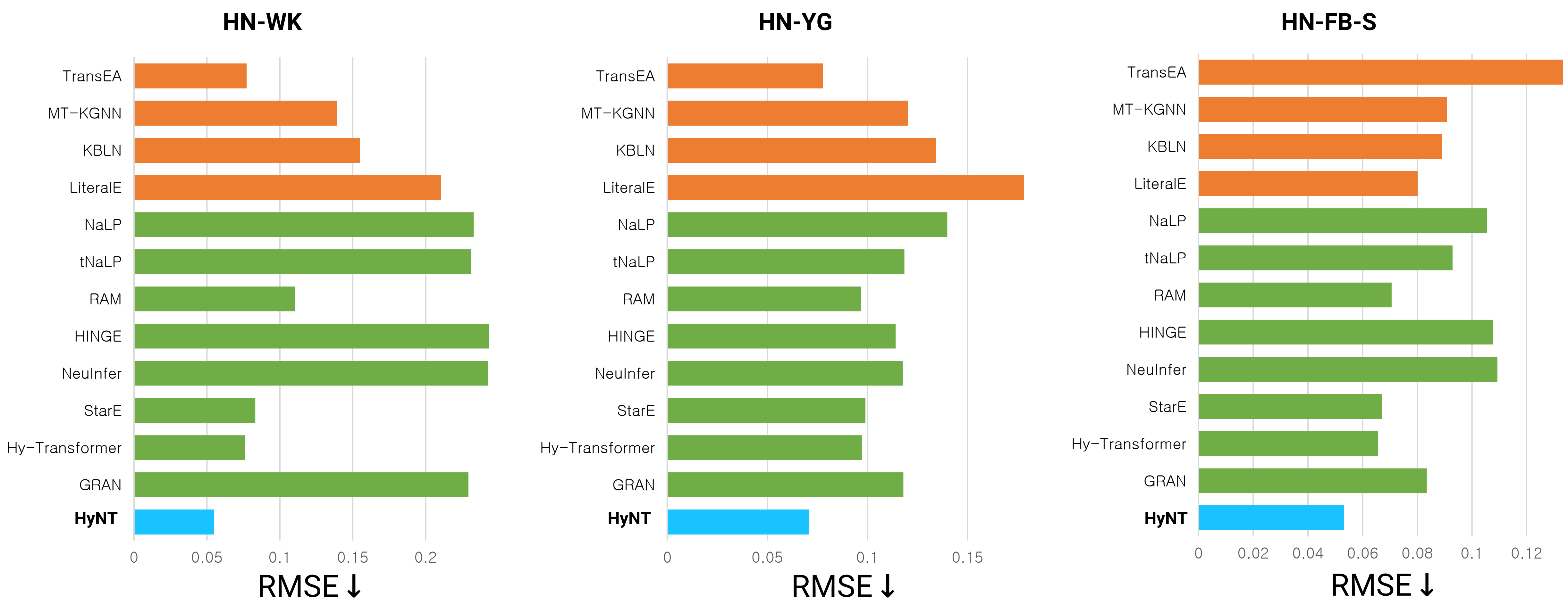
((?, nominated_for, Best_Actor), {(for_work, Forrest_Gump), (subject_of, 67^{th}_Oscars)})

03 Link Prediction Results of HyNT

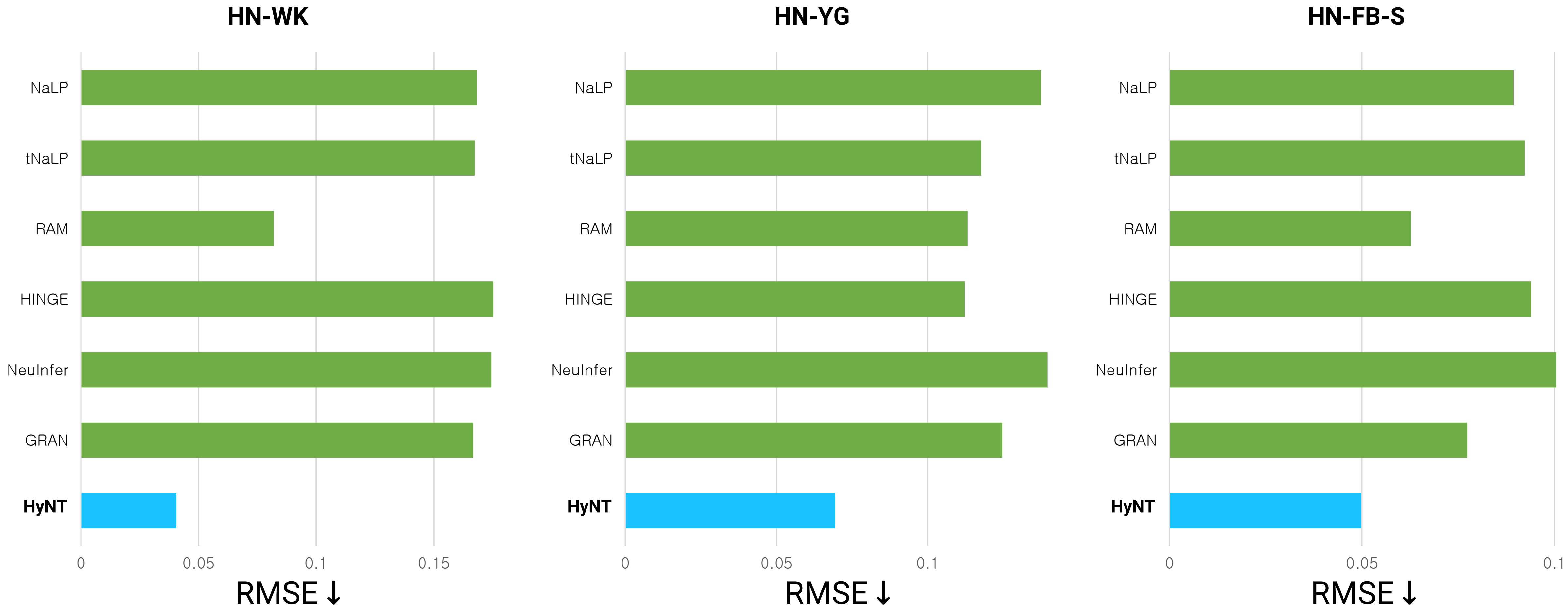
((Brad_Pitt, nominated_for, Best_Actor), {(for_work, Moneyball), (subject_of, 84^{th}_Oscars)})

((Tom_Hanks, nominated_for, Best_Actor), {(for_work, Forrest_Gump), (subject_of, 67^{th}_Oscars)})

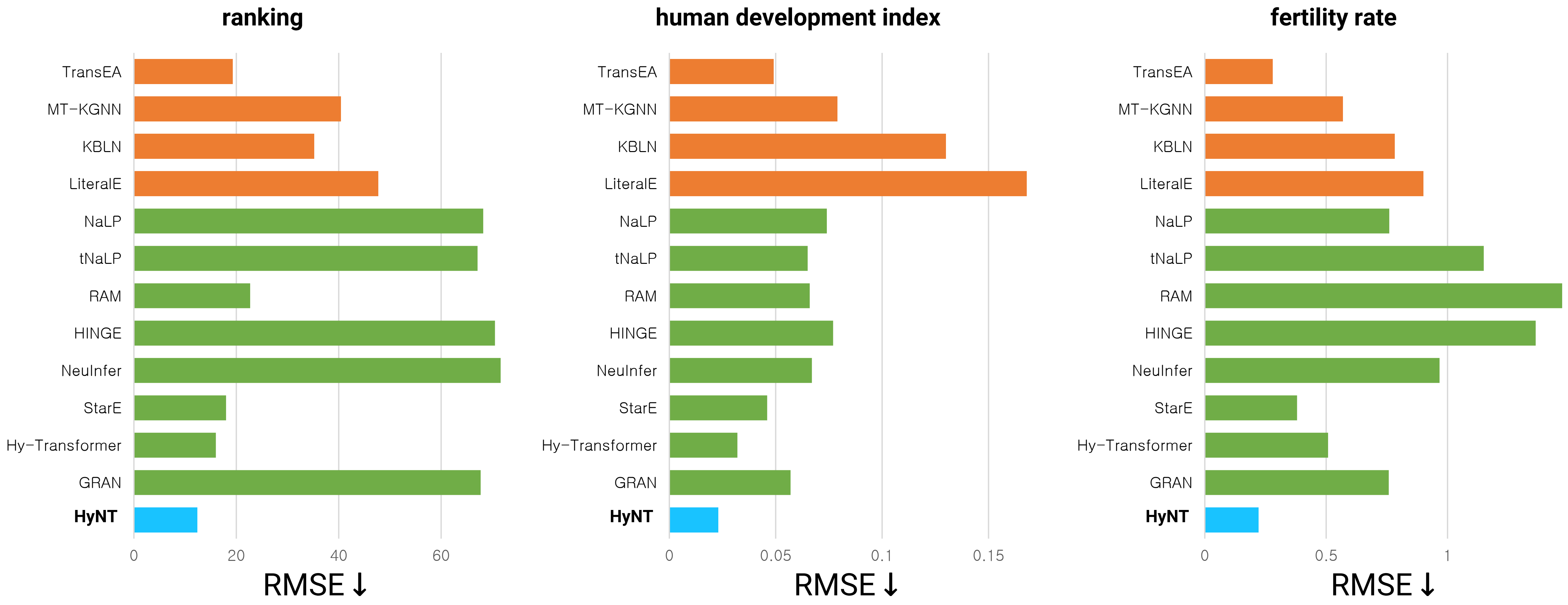
03 Numeric Value Prediction Results – Primary



03 Numeric Value Prediction Results – All



03 Numeric Value Prediction Results per Attribute Type in HN-WK



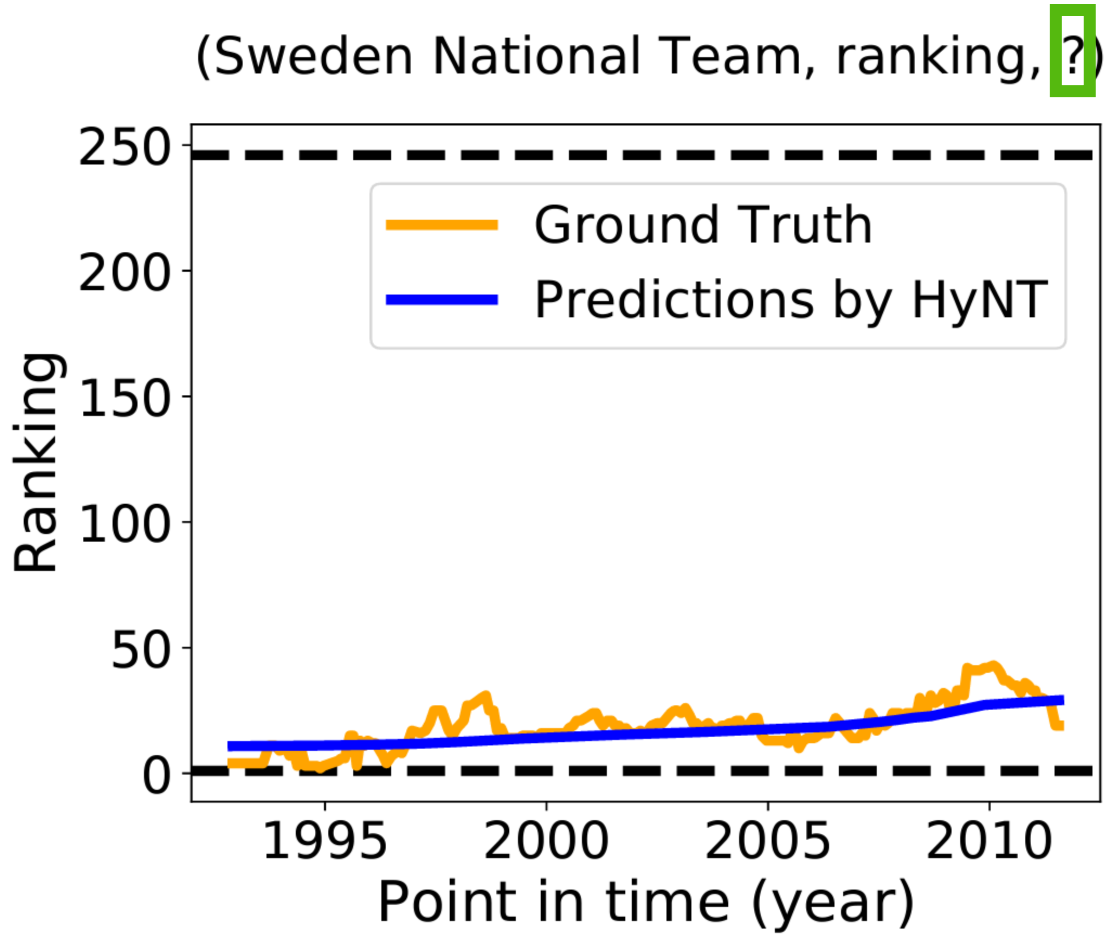
03 Visualization of the Predictions

Target Values

((Sweden National team, ranking, ?), {(point in time, 1995)})
((Sweden National team, ranking, ?), {(point in time, 1996)})
((Sweden National team, ranking, ?), {(point in time, 1997)})

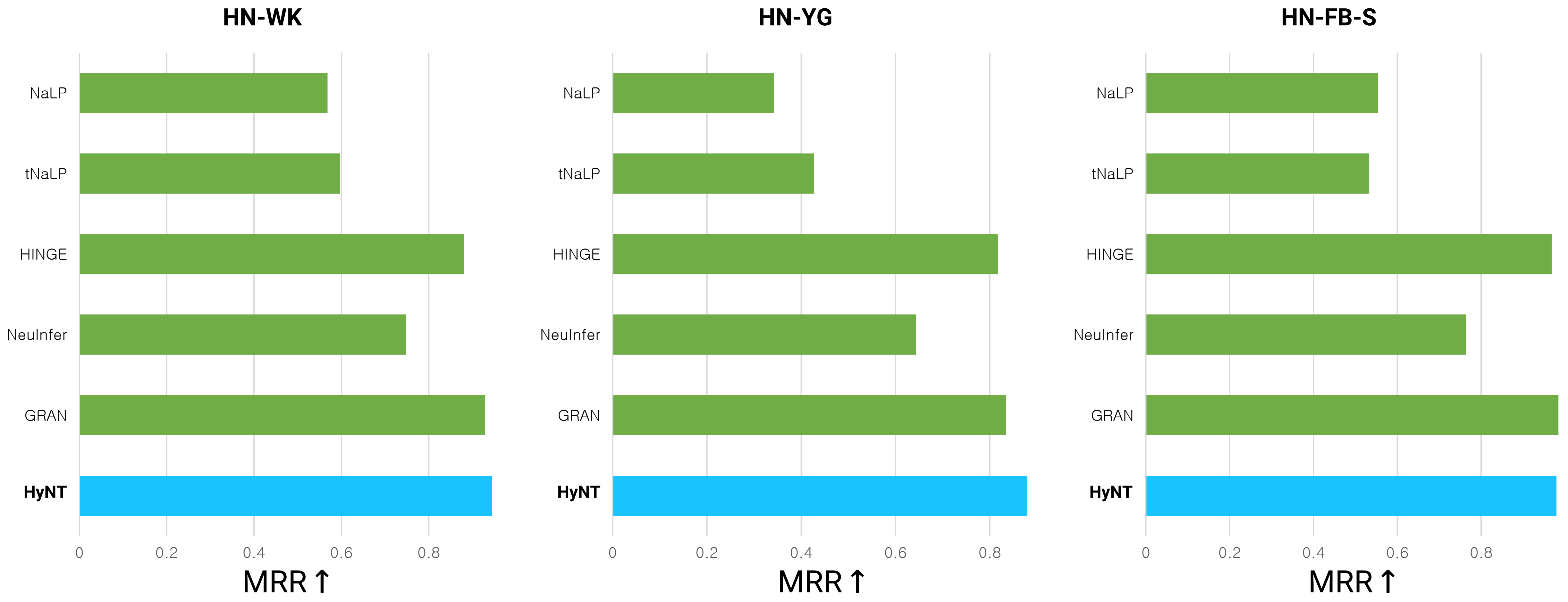
⋮

03 Visualization of the Predictions

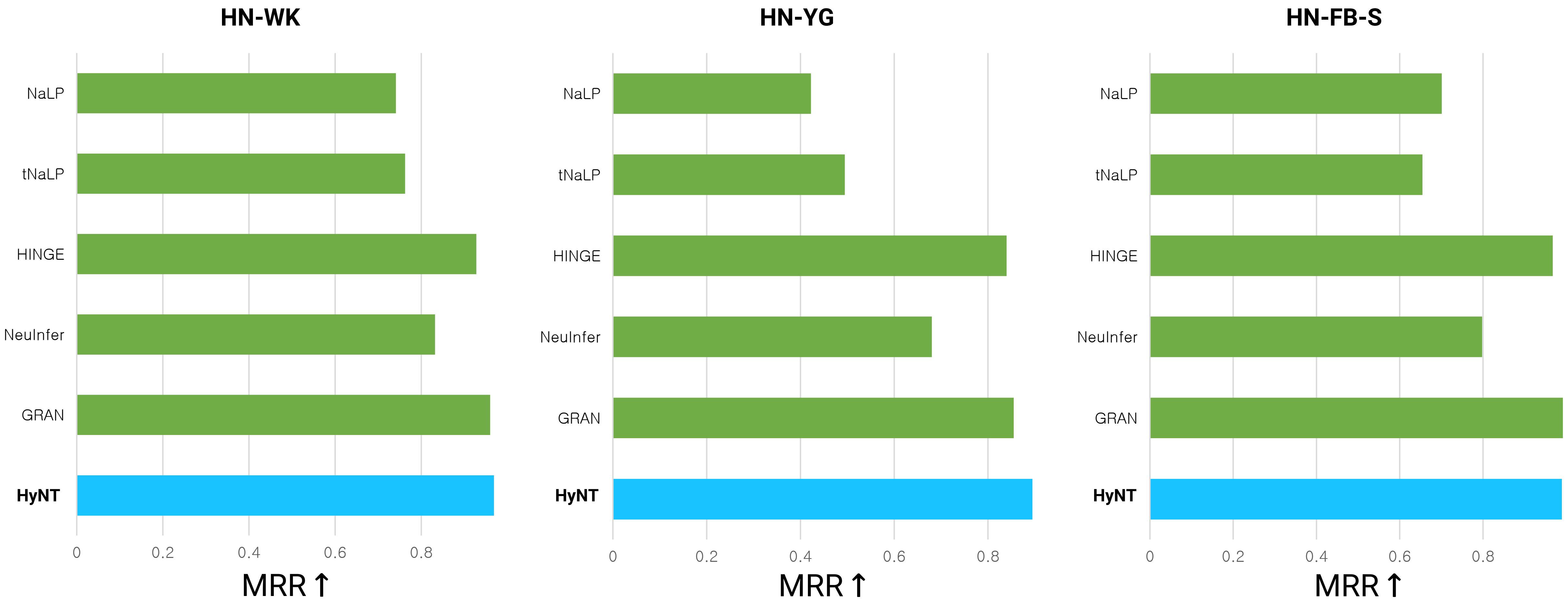


Target Values

03 Relation Prediction Results – Primary



03 Relation Prediction Results – All



03 Conclusion & Future Work

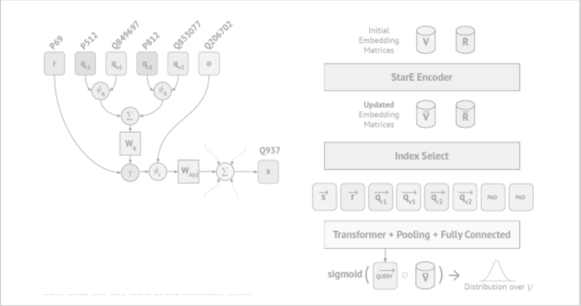
- **H**yper-relational and **N**umeric **K**nowledge **G**raphs **(HN-KGs)**
- Propose **HyNT** to solve **link prediction**, **numeric value prediction**, and **relation prediction** on **HN-KGs**
- HyNT significantly outperforms 12 different state-of-the-art methods
- Extend HyNT to **inductive learning scenarios**
  - New entities and relations appear at test time

I

Message Passing for Hyper-Relational Knowledge Graphs

Mikhail Galkin, Priyansh Trivedi, Gaurav Maheshwari, Ricardo Usbeck, and Jens Lehmann

EMNLP 2020

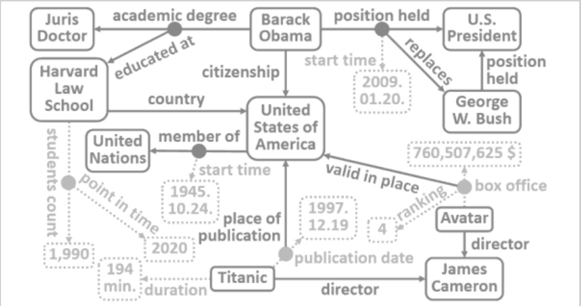


II

Representation Learning on Hyper-Relational and Numeric Knowledge Graphs with Transformers

Chanyoung Chung‡, Jaejun Lee‡, and Joyce Jiyoung Whang*

KDD 2023

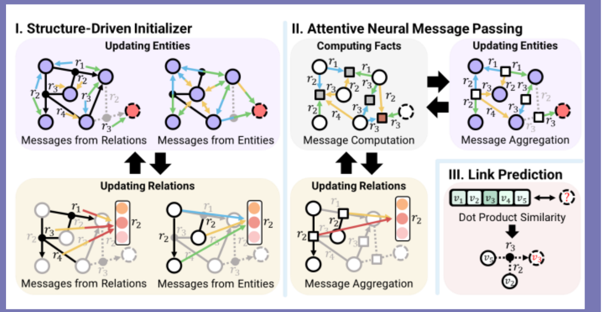


III

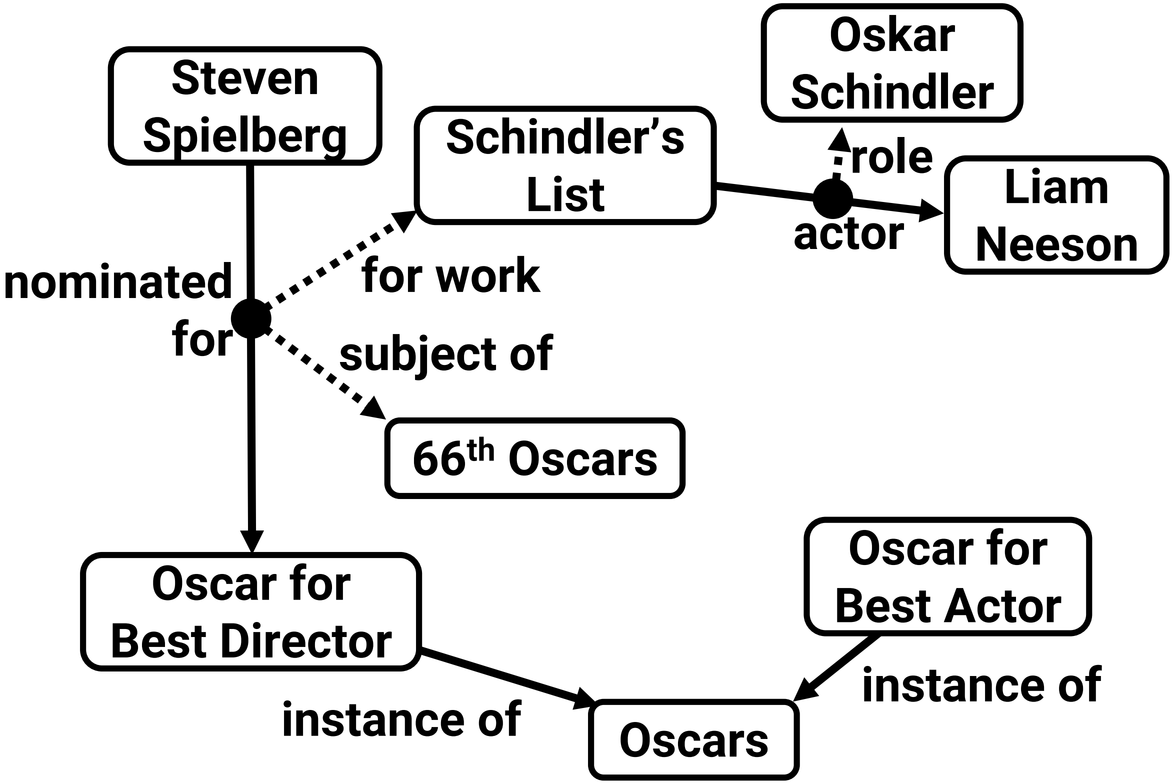
Structure Is All You Need: Structural Representation Learning on Hyper-Relational Knowledge Graphs

Jaejun Lee and Joyce Jiyoung Whang*

ICML 2025

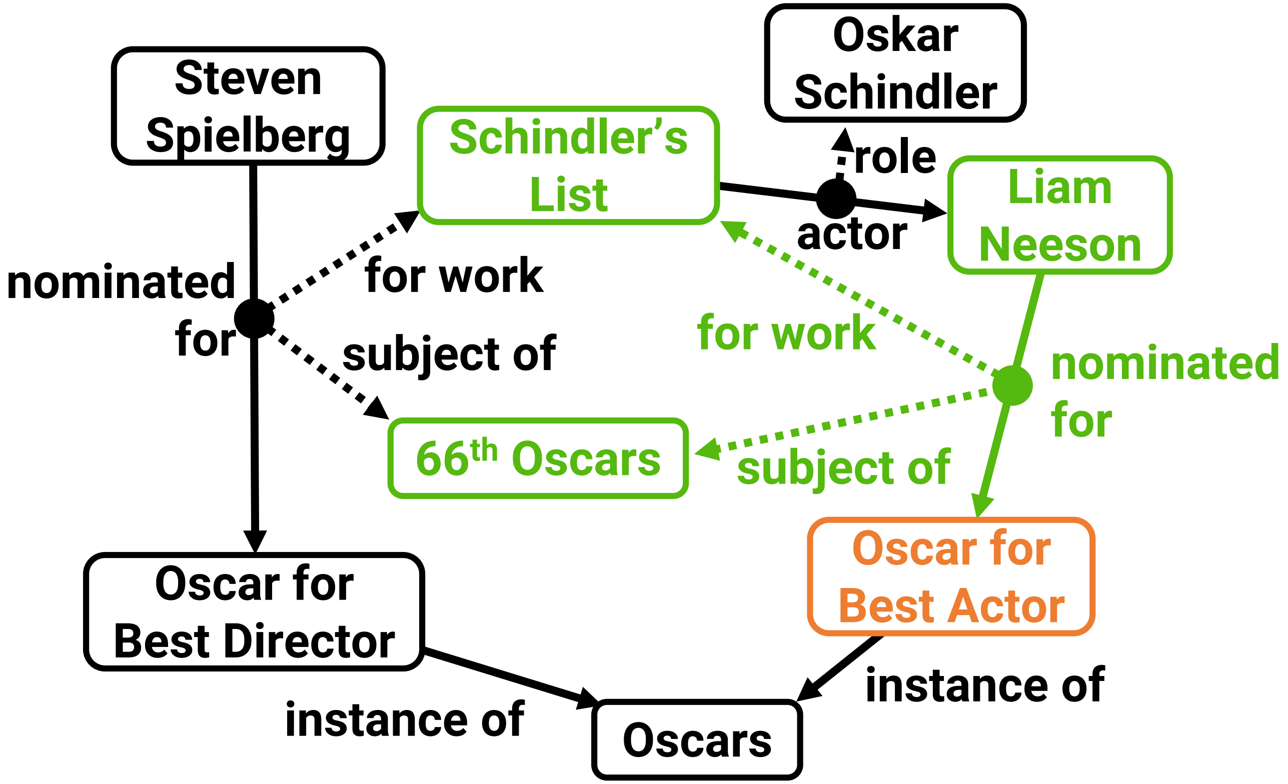


04 Hyper-relational Knowledge Graphs (HKGs)

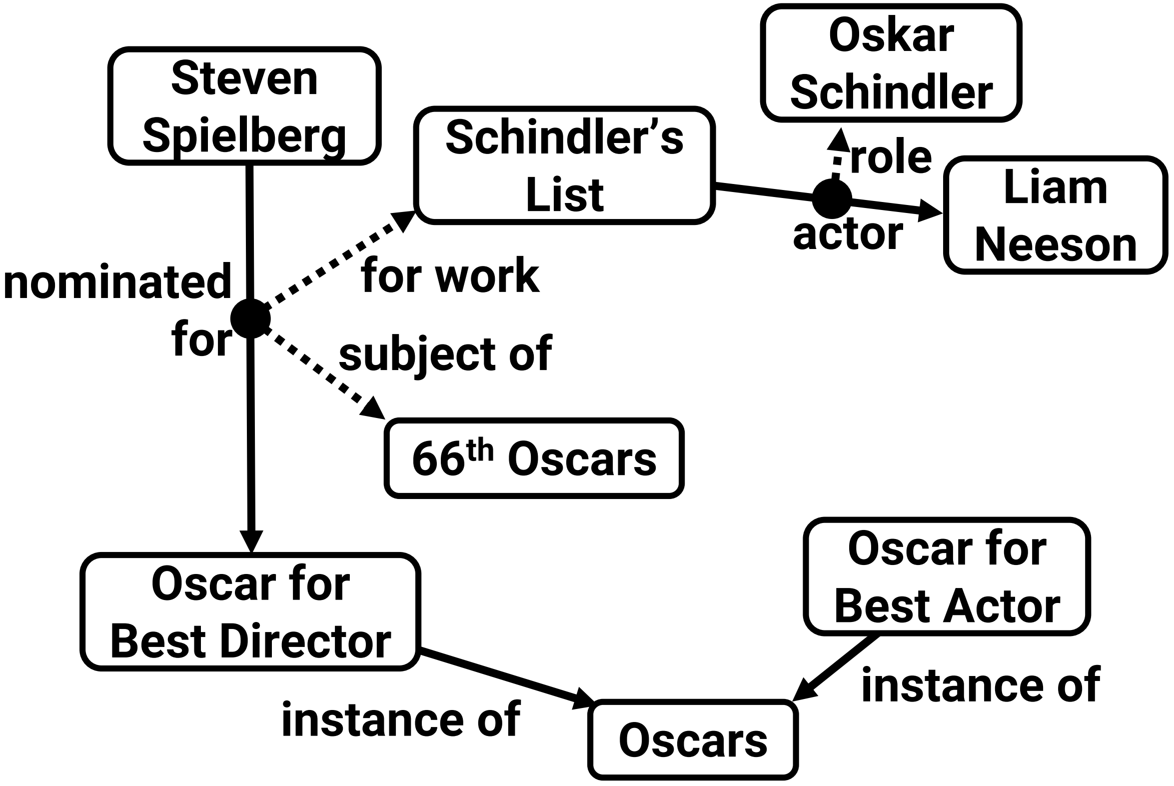


Transductive Link Prediction on HKGs

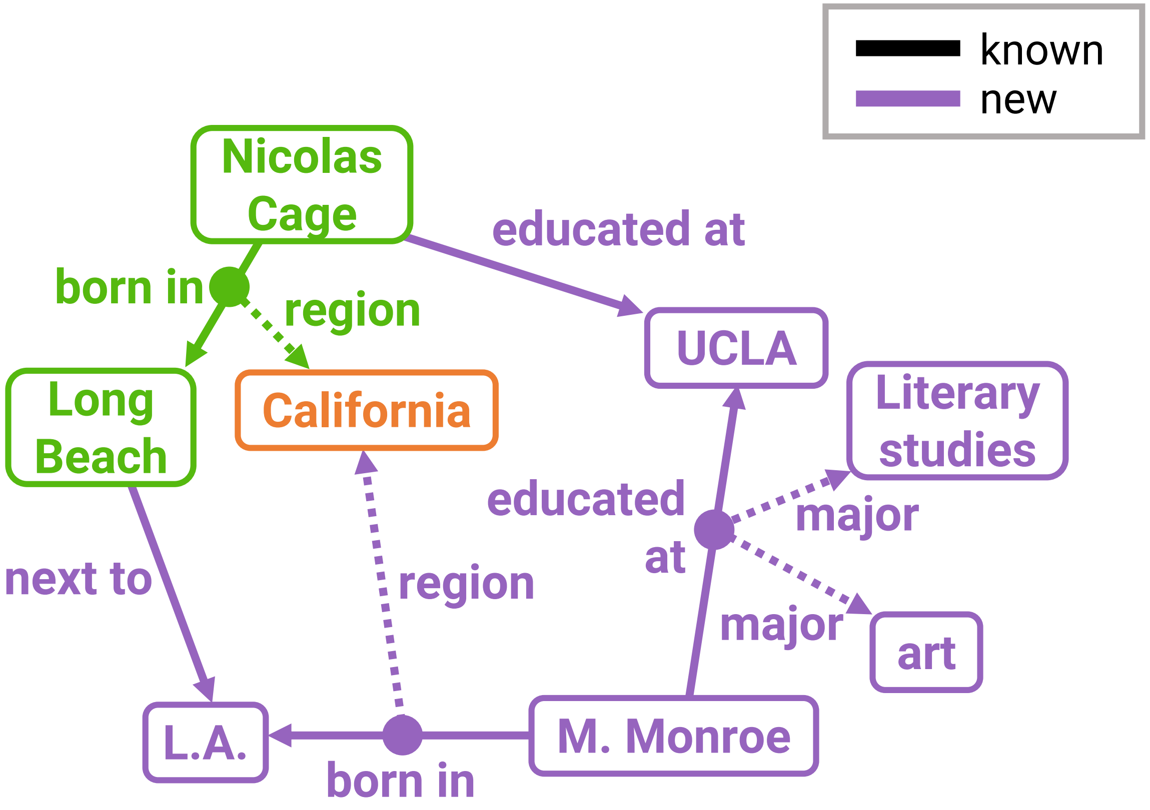
((Liam Neeson, nominated for, ?),
{(for work, Schindler's List), (subject of, 66th Oscars)})



04 Inductive Link Prediction on HKGs



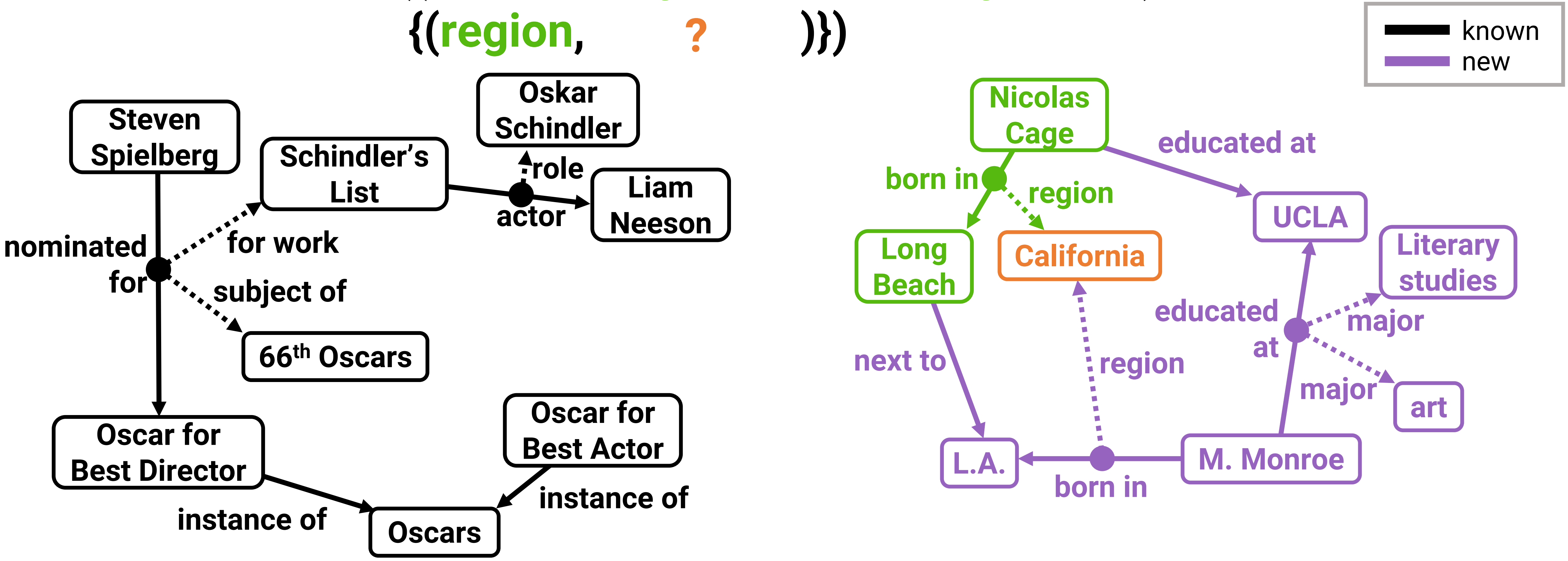
Training HKG



Inference HKG

Inductive Link Prediction on HKGs

((Nicolas Cage, born in, Long Beach),
{(region, ?)})

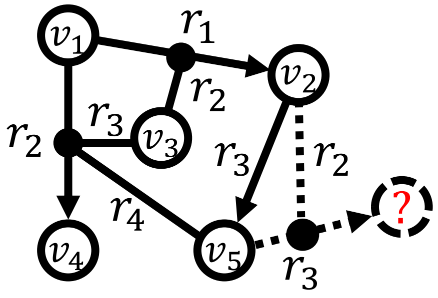


Training HKG

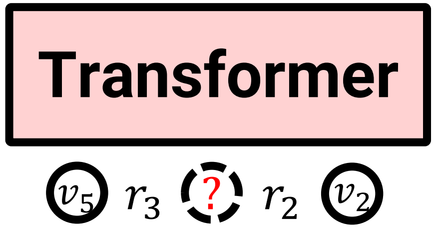
Inference HKG

04 Existing HKG Methods

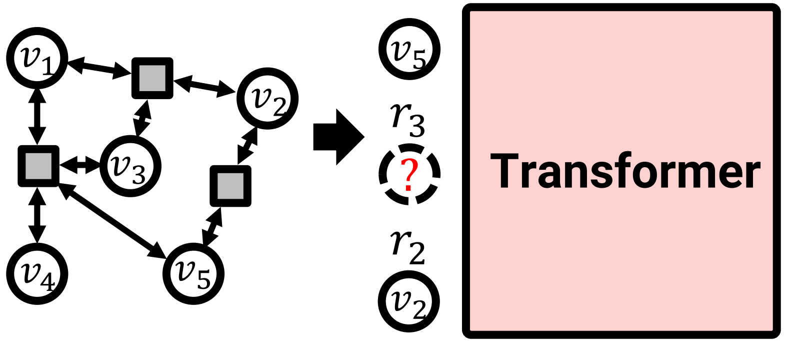
- Incorporate only limited structural information and fail to utilize HKG structures
  - Transformer-based methods: process each fact individually
  - GNN-based encoders: redundant / does not consider relations and positions of entities



Hyper-relational Knowledge Graph



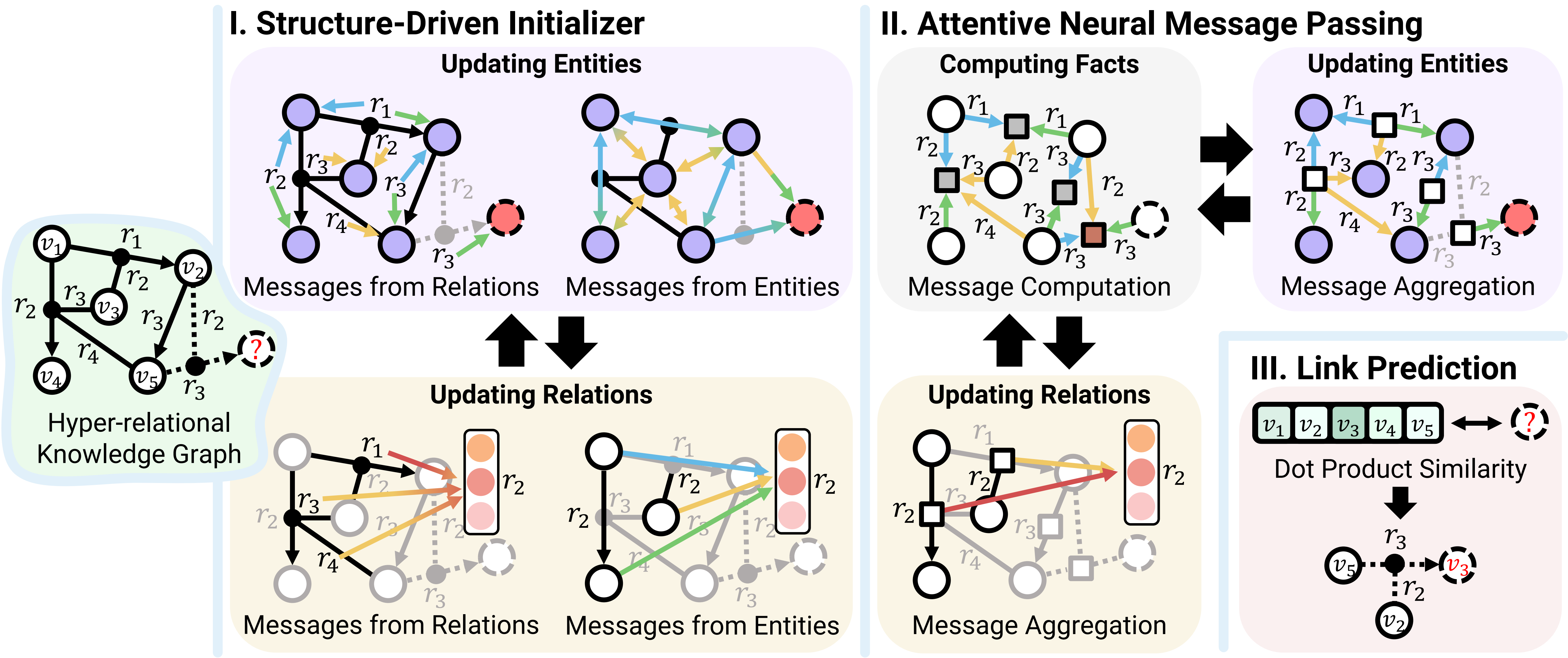
Transformer-based Methods



Existing GNN-based Methods

	Transformer-based Methods	Existing GNN-based Methods
Inductive Inference	X	X
Structure Utilization	X	△

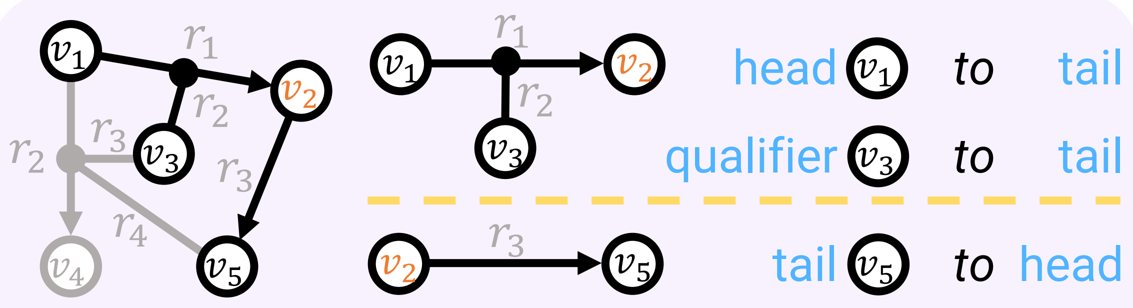
MAYPL: Message Passing Framework for Hyper-Relational Knowledge Graph Representation Learning



Structure-driven Initializer

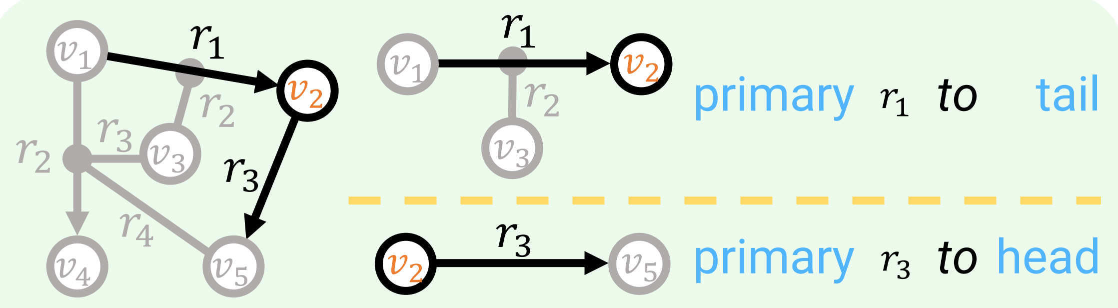
- Exploits the **interconnections**, **co-occurrence**, and **positions** of entities and relations

- Entity: Aggregate the messages of the **co-occurring entities** and **incident relations**
  - $\tilde{\mathbf{v}}^{(\tilde{l})} = \frac{1}{\sum_{u\in\mathcal{V}_v}|\mathcal{H}_u\cap\mathcal{H}_v|}\sum_{u\in\mathcal{V}_v}\sum_{h\in\mathcal{H}_u\cap\mathcal{H}_v}\widetilde{\boldsymbol{U}}^{(\tilde{l})}_{\lambda_h(v)}\widetilde{\boldsymbol{W}}^{(\tilde{l})}_{\lambda_h(u)}\tilde{\mathbf{u}}^{(l-1)} + \frac{1}{|\mathcal{R}_v|}\sum_{r\in\mathcal{R}_v}\widetilde{\boldsymbol{A}}^{(\tilde{l})}_{\tau_r(v)}\tilde{\mathbf{r}}^{(\tilde{l}-1)}$



Messages from co-occurring entities

+

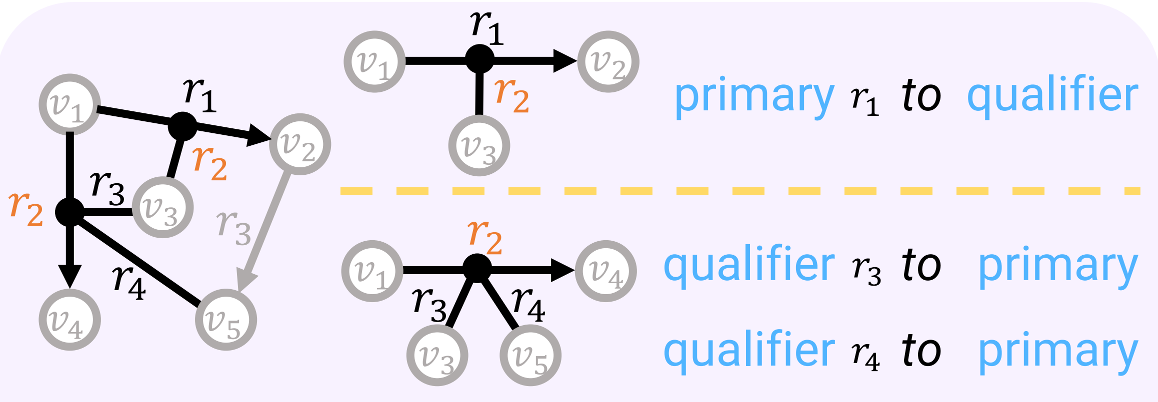


Messages from incident relations

04 Structure-driven Initializer

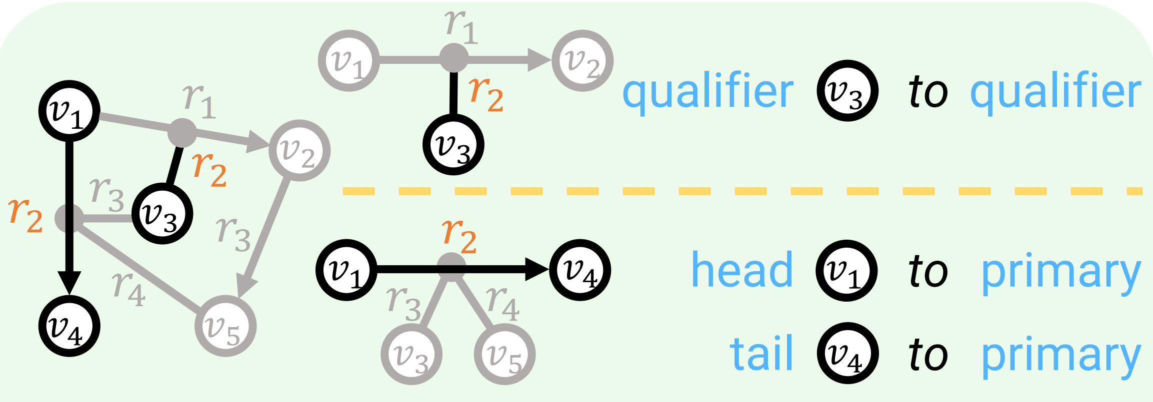
- Exploits the **interconnections**, **co-occurrence**, and **positions** of entities and relations

- Relation: Aggregate the messages of the **co-occurring relations** and **incident entities**
  - $\tilde{\mathbf{r}}^{(\tilde{l})} = \frac{1}{\sum_{y \in \mathcal{R}_r} |\mathcal{H}_y \cap \mathcal{H}_r|} \sum_{y \in \mathcal{R}_r} \sum_{h \in \mathcal{H}_y \cap \mathcal{H}_r} \widehat{\boldsymbol{U}}_{\lambda_h(r)}^{(\tilde{l})} \widehat{\boldsymbol{W}}_{\lambda_h(y)}^{(\tilde{l})} \tilde{\mathbf{y}}^{(l-1)} + \frac{1}{|\mathcal{V}_r|} \sum_{v \in \mathcal{V}_r} \widehat{\boldsymbol{A}}_{\tau_r(v)}^{(\tilde{l})} \tilde{\mathbf{v}}^{(\tilde{l}-1)}$



Messages from co-occurring relations

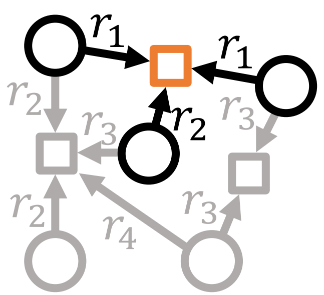
+

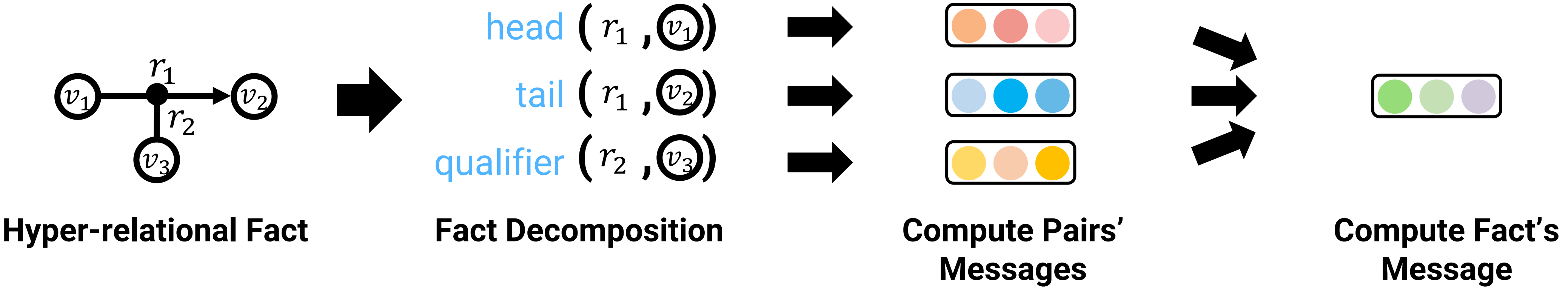


Messages from incident entities

04 Attentive Neural Message Passing

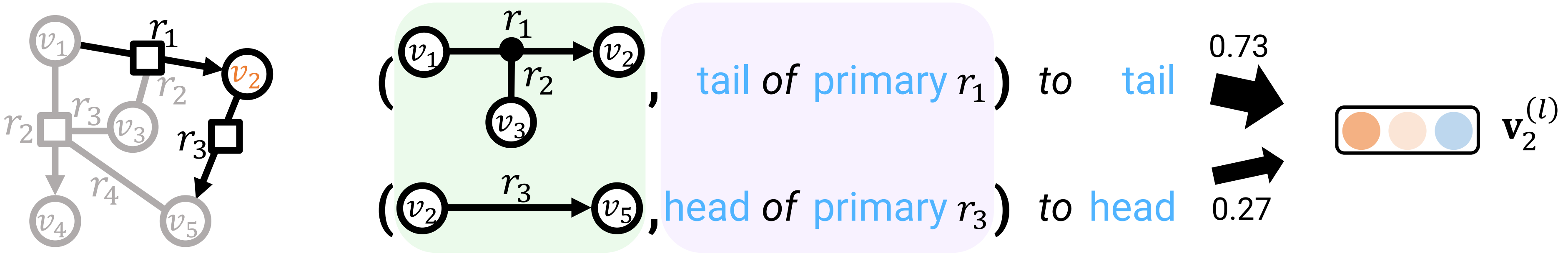
- Updates entity and relation representations by attentively aggregating facts' messages
- Computes a fact's message by considering which entities and relations comprise the fact
  - Decompose each fact as a set of its relation-entity pairs
  - compute a pair's message by considering the **entities** and **relations** and their **positions** within the fact
    - $\mathbf{p}^{(l)} = \boldsymbol{P}^{(l)}_{\lambda_h(v)}\left(\left(\boldsymbol{W}^{(l)}_{\lambda_h(v)}\mathbf{v}^{(l-1)}\right) \odot \left(\boldsymbol{U}^{(l)}_{\lambda_h(v)}\mathbf{r}^{(l-1)}\right)\right)$
  - Compute a fact's message by aggregating messages of its relation-entity pairs
    - $\mathbf{h}^{(l)} = \frac{1}{|(\mathcal{R}\times\mathcal{V})_h|}\sum_{p\in(\mathcal{R}\times\mathcal{V})_h}\mathbf{p}^{(l)}$





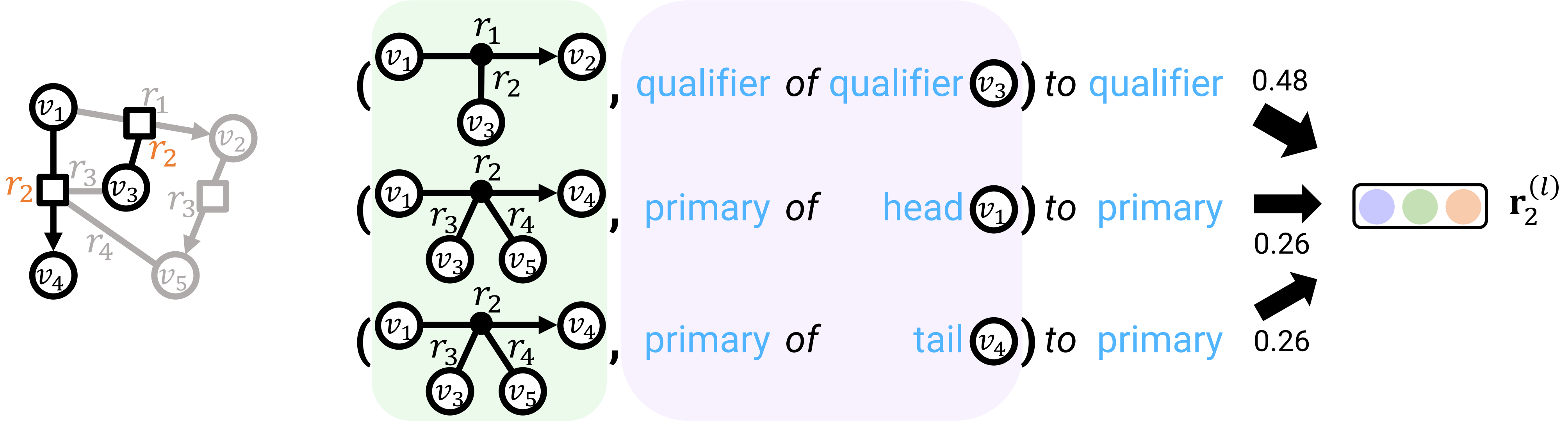
Attentive Neural Message Passing

- Updates an entity representation by considering which fact the entity belongs to
  - Considers pairs of incident relations and the corresponding facts
    - $\mathbf{q}^{(l)} = \boldsymbol{Q}^{(l)}_{\lambda_h(v)} \left(\mathbf{h}^{(l)} \odot \left(\boldsymbol{A}^{(l)}_{\lambda_h(v)} \mathbf{r}^{(l-1)} \right) \right)$
  - Attentively aggregate relation-fact pairs
    - $\mathbf{v}^{(l)} = \sum_{q \in (\mathcal{R} \times \mathcal{H})_v} \alpha^{(l)}_{q,v} \boldsymbol{B}^{(l)} \mathbf{q}^{(l)}$, where $\alpha^{(l)}_{q,v} = \frac{\exp(\mathbf{a}^{(l)} \cdot \sigma(\boldsymbol{Q}^{(l)} \mathbf{v}^{(l-1)} + \boldsymbol{K}^{(l)} \mathbf{q}^{(l)}))}{\sum_{k \in (\mathcal{R} \times \mathcal{H})_v} \exp(\mathbf{a}^{(l)} \cdot \sigma(\boldsymbol{Q}^{(l)} \mathbf{v}^{(l-1)} + \boldsymbol{K}^{(l)} \mathbf{k}^{(l)}))}$



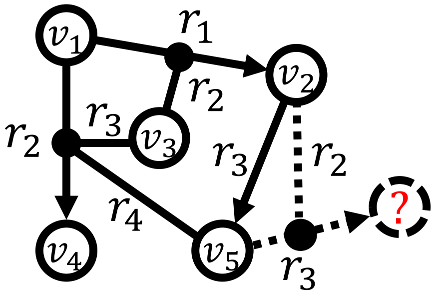
04 Attentive Neural Message Passing

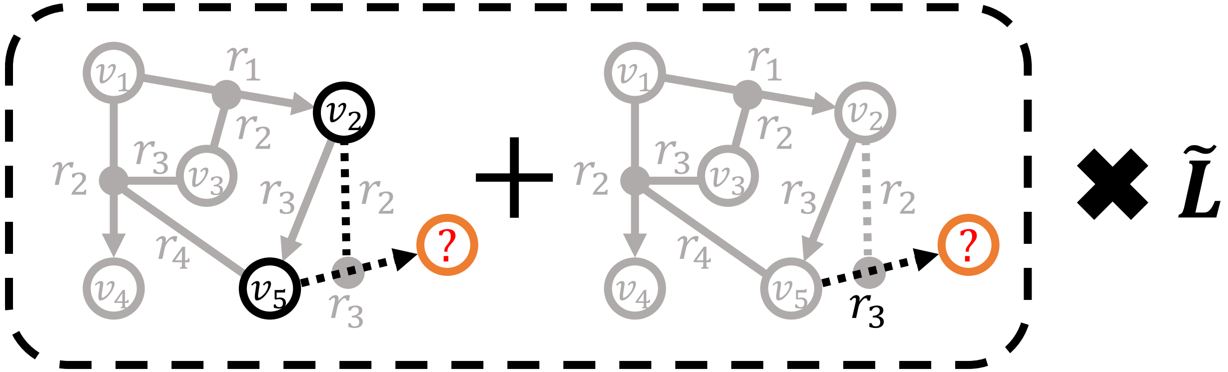
- Updates an relation representation by considering which fact the relation belongs to
  - Considers pairs of incident entities and the corresponding facts
    - $\bar{\mathbf{q}}^{(l)} = \bar{\boldsymbol{Q}}^{(l)}_{\lambda_h(r)}\left(\mathbf{h}^{(l)} \odot \left(\bar{\boldsymbol{A}}^{(l)}_{\lambda_h(v)}\mathbf{v}^{(l-1)}\right)\right)$
  - Attentively aggregate relation-fact pairs
    - $\mathbf{r}^{(l)} = \sum_{\bar{q}\in(\mathcal{V}\times\mathcal{H})_r} \bar{\alpha}^{(l)}_{\bar{q},r}\bar{\boldsymbol{B}}^{(l)}\bar{\mathbf{q}}^{(l)}$, where $\bar{\alpha}^{(l)}_{\bar{q},r} = \frac{\exp(\bar{\mathbf{a}}^{(l)}\cdot\sigma(\bar{\boldsymbol{Q}}^{(l)}\mathbf{r}^{(l-1)}+\bar{\boldsymbol{K}}^{(l)}\bar{\mathbf{q}}^{(l)}))}{\sum_{\bar{k}\in(\mathcal{V}\times\mathcal{H})_r}\exp(\bar{\mathbf{a}}^{(l)}\cdot\sigma(\bar{\boldsymbol{Q}}^{(l)}\mathbf{r}^{(l-1)}+\bar{\boldsymbol{K}}^{(l)}\bar{\mathbf{k}}^{(l)}))}$



04 Link Prediction on HKGs

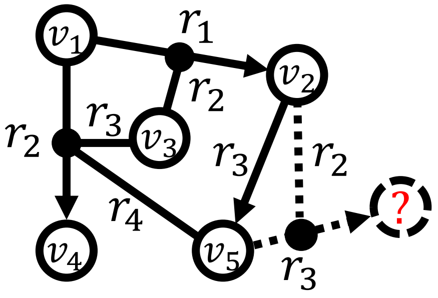
- Structure-driven initializer
  - $\tilde{\mathbf{x}}^{(\tilde{l})} = \sum_{v \in \mathcal{V}_x} \tilde{\boldsymbol{U}}^{(\tilde{l})}_{\lambda_h(x)} \tilde{\boldsymbol{W}}^{(\tilde{l})}_{\lambda_h(v)} \tilde{\mathbf{v}}^{(l-1)} + \tilde{\boldsymbol{A}}^{(\tilde{l})}_{\tau_r(v)} \tilde{\mathbf{r}}^{(\tilde{l}-1)}$
- Attentive Neural Message Passing
  - $\mathbf{x}^{(l)} = \boldsymbol{B}^{(l)} \left(\boldsymbol{Q}^{(l)}_{\lambda_h(x)} \left(\mathbf{h}^{(l)} \odot \left(\boldsymbol{A}^{(l)}_{\lambda_h(x)} \mathbf{r}^{(l-1)} \right) \right) \right)$
- Compute the dot product similarity between $\mathbf{x}^{(L)}$ and each entity representation
  - MAYPL predicts the missing entity as the entity with the highest similarity

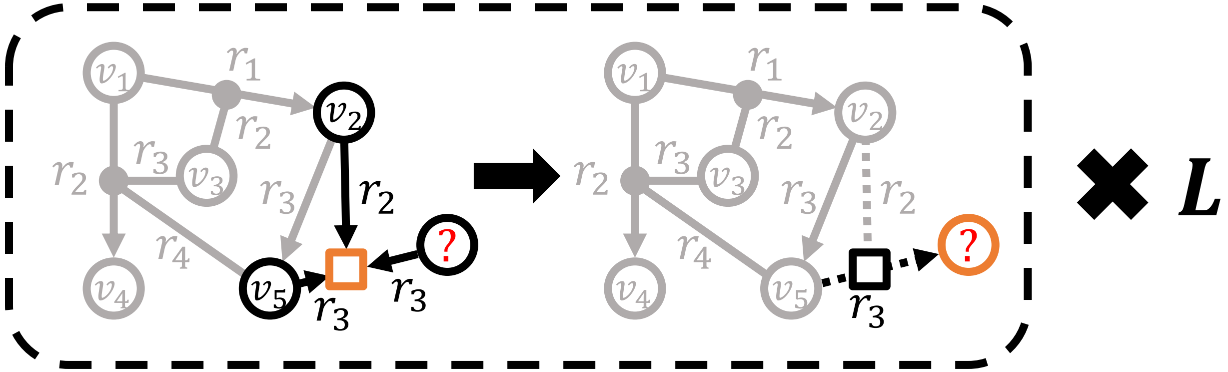




04 Link Prediction on HKGs

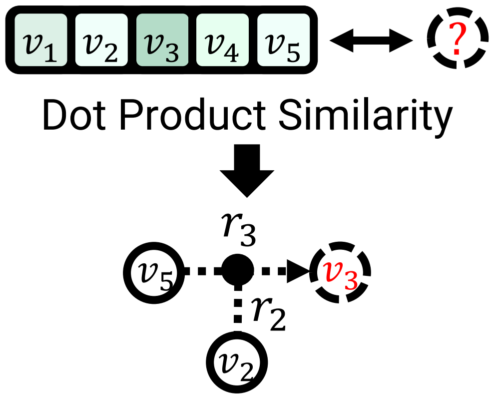
- Structure-driven initializer
  - $\tilde{\mathbf{x}}^{(\tilde{l})} = \sum_{v \in \mathcal{V}_x} \widetilde{\boldsymbol{U}}_{\lambda_h(x)}^{(\tilde{l})} \widetilde{\boldsymbol{W}}_{\lambda_h(v)}^{(\tilde{l})} \tilde{\mathbf{v}}^{(l-1)} + \widetilde{\boldsymbol{A}}_{\tau_r(v)}^{(\tilde{l})} \tilde{\mathbf{r}}^{(\tilde{l}-1)}$
- Attentive Neural Message Passing
  - $\mathbf{x}^{(l)} = \boldsymbol{B}^{(l)} \left(\boldsymbol{Q}_{\lambda_h(x)}^{(l)} \left(\mathbf{h}^{(l)} \odot \left(\boldsymbol{A}_{\lambda_h(x)}^{(l)} \mathbf{r}^{(l-1)} \right) \right) \right)$
- Compute the dot product similarity between $\mathbf{x}^{(L)}$ and each entity representation
  - MAYPL predicts the missing entity as the entity with the highest similarity





04 Link Prediction on HKGs

- Structure-driven initializer
  - $\tilde{\mathbf{x}}^{(\tilde{l})} = \sum_{v \in \mathcal{V}_x} \widetilde{\boldsymbol{U}}^{(\tilde{l})}_{\lambda_h(x)} \widetilde{\boldsymbol{W}}^{(\tilde{l})}_{\lambda_h(v)} \tilde{\mathbf{v}}^{(l-1)} + \widetilde{\boldsymbol{A}}^{(\tilde{l})}_{\tau_r(v)} \tilde{\mathbf{r}}^{(\tilde{l}-1)}$
- Attentive Neural Message Passing
  - $\mathbf{x}^{(l)} = \boldsymbol{B}^{(l)} \left(\boldsymbol{Q}^{(l)}_{\lambda_h(x)} \left(\mathbf{h}^{(l)} \odot \left(\boldsymbol{A}^{(l)}_{\lambda_h(x)} \mathbf{r}^{(l-1)} \right) \right) \right)$
- Compute the dot product similarity between $\mathbf{x}^{(L)}$ and each entity representation
  - MAYPL predicts the missing entity as the entity with the highest similarity

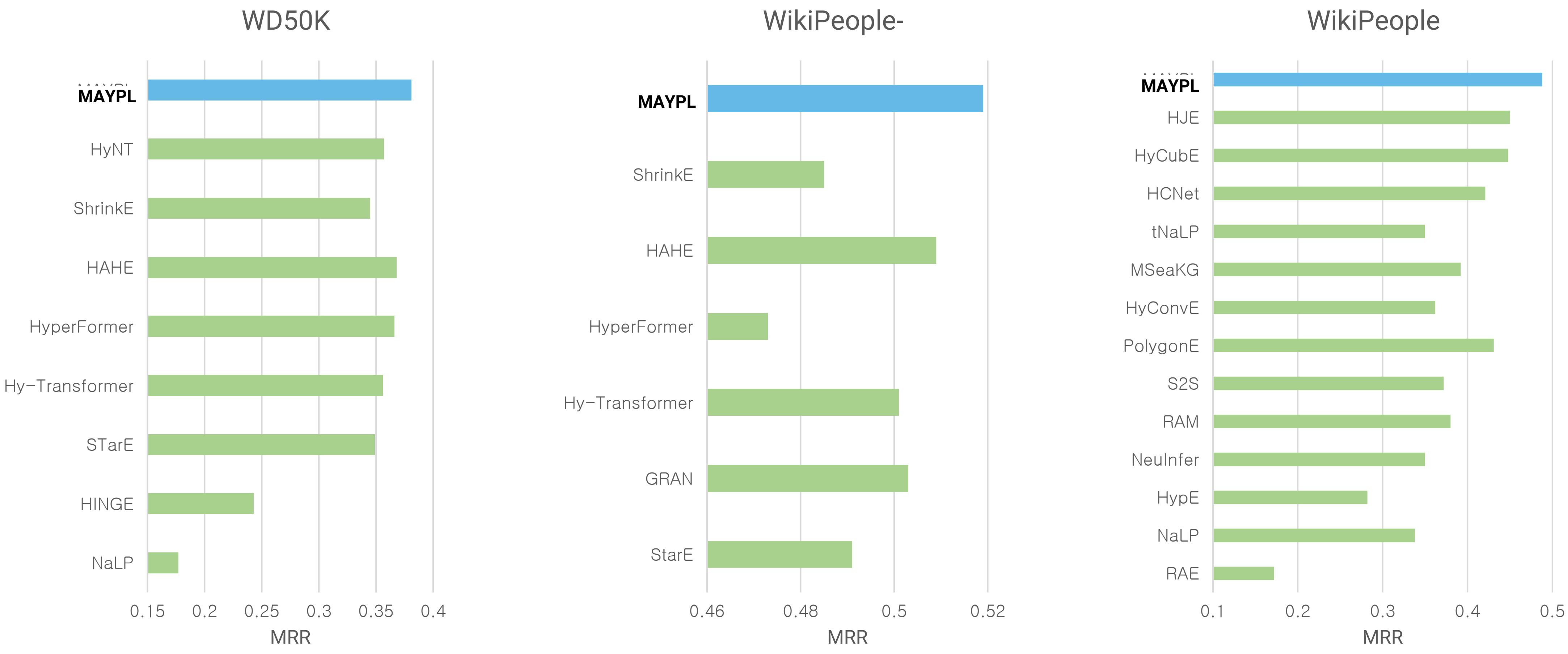


04 Experimental Results

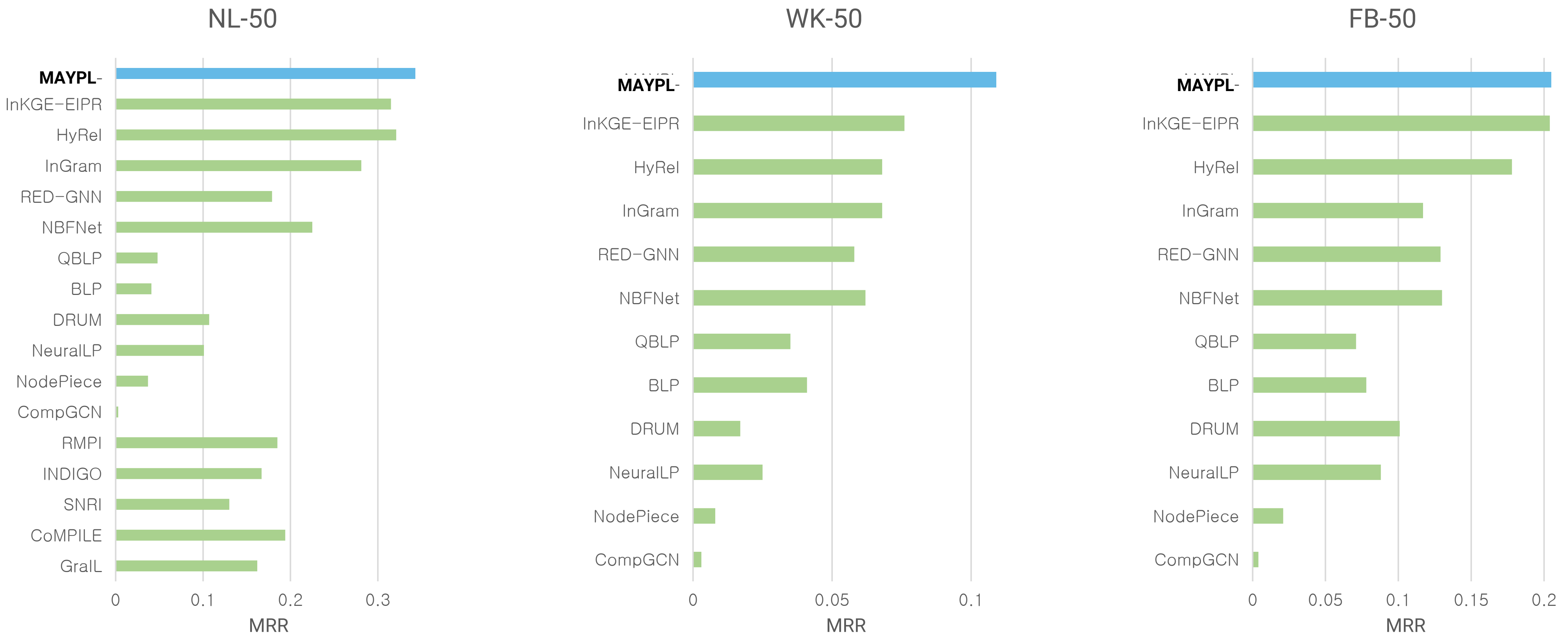
- Datasets
  - **3 Transductive HKG** datasets: WD50K, WikiPeople^{-}, WikiPeople
  - **12 Inductive KG** datasets from InGram (e.g., NL-100, WK-100, FB-100)
  - **4 Inductive HKG** datasets: WD20K(100)v1, WD20K(100)v2, WP-IND, MFB-IND

- Baselines
  - **41 knowledge graph completion methods**, compared with different baseline methods on different datasets

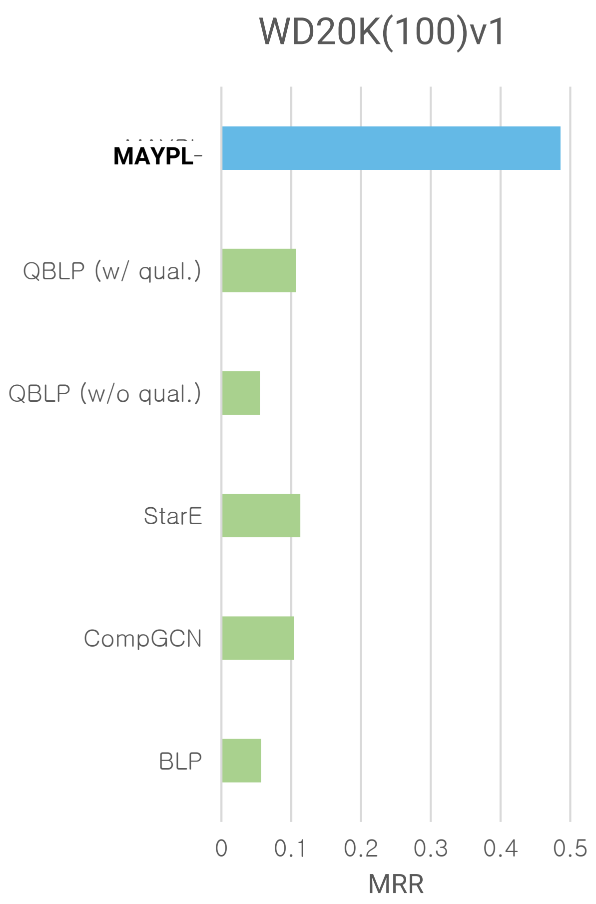
04 Transductive Link Prediction on HKGs

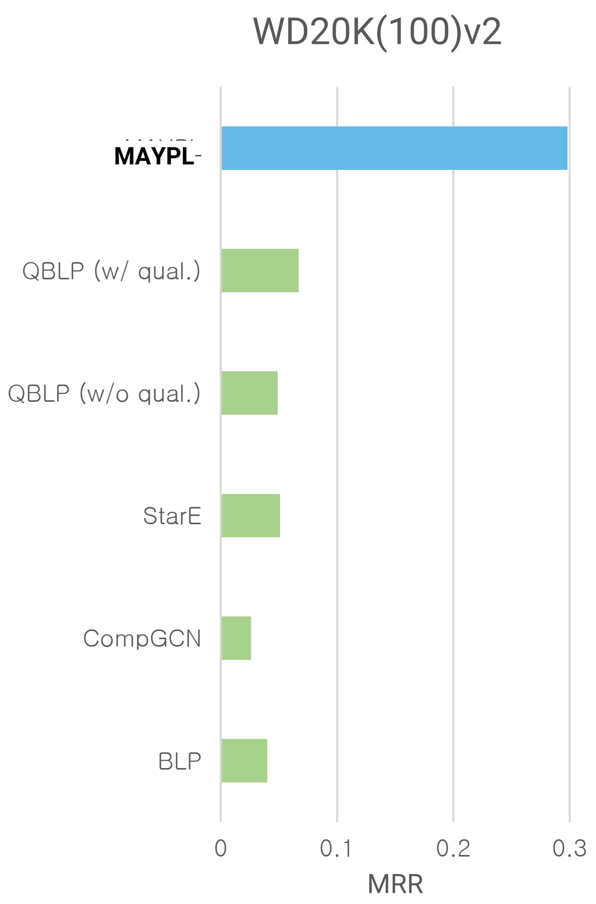


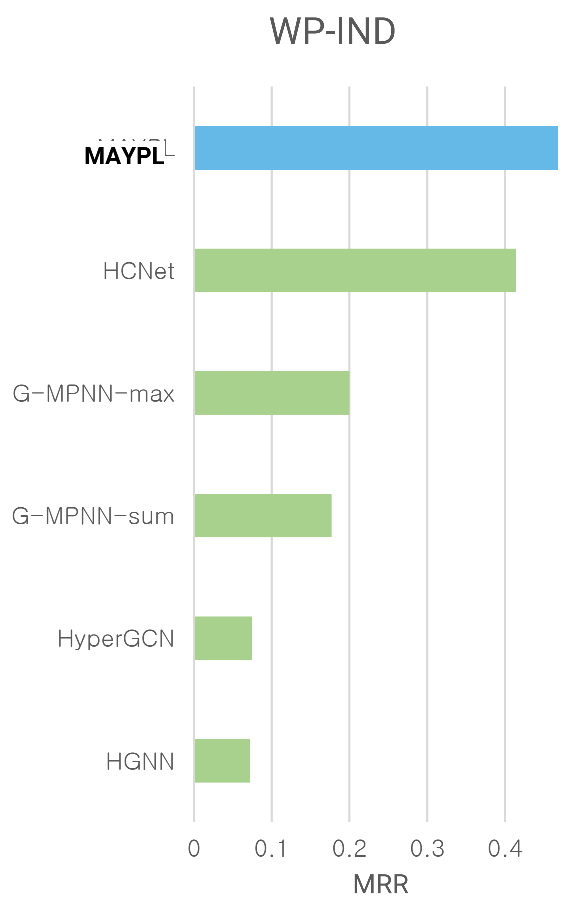
04 Inductive Link Prediction on KGs

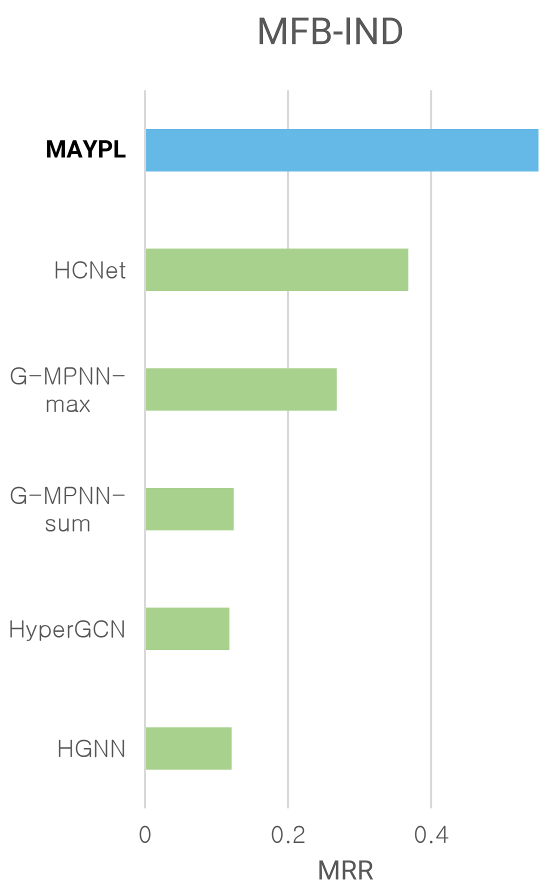


04 Inductive Link Prediction on HKGs









04 Conclusion

- Employing an **HKG's structure is crucial** for HKG reasoning
  - Thoroughly learning and exploiting the structure of an HKG is necessary and sufficient for learning representations on HKGs

- Propose **MAYPL**, the first structure-oriented representation learning method for HKGs
  - Can be applied in both transductive and inductive settings

- MAYPL can make inductive inferences with new entities and relations by **learning how to compute representations based solely on the structure** of a given HKG

- MAYPL outperforms **41 different methods on 19 benchmark datasets** in varied settings
  - Transductive link prediction on HKGs, inductive link prediction on KGs, and inductive link prediction on HKGs

05 References

- Some slides are made based on the following references.

- P. Rosso et al., “Beyond Triplets: Hyper-Relational Knowledge Graph Embedding for Link Prediction”, TheWebConf, 2020.
- M. Galkin et al., “Message Passing for Hyper-Relational Knowledge Graphs”, EMNLP, 2020.
- Q. Wang et al., “Link Prediction on N-ary Relational Facts: A Graph-based Approach”, ACL Findings, 2020.
- M. Ali et al., “Improving Inductive Link Prediction Using Hyper-relational Facts”, ISWC, 2021.
- H. Luo et al., “HAHE: Hierarchical Attention for Hyper-Relational Knowledge Graphs in Global and Local Level”, ACL, 2023.
- B. Xiong et al., “Shrinking Embeddings for Hyper-Relational Knowledge Graphs”, ACL, 2023.
- C. Chung et al., “Representation Learning on Hyper-Relational and Numeric Knowledge Graphs with Transformers”, KDD, 2023.
- Z. Hu et al., “HyperFormer: Enhancing Entity and Relation Interaction for Hyper-Relational Knowledge Graph Completion”, CIKM, 2023.
- J. Lee and J. J. Whang, “Structure Is All You Need: Structural Representation Learning on Hyper-Relational Knowledge Graphs”, ICML, 2025.