GCR

Citation: 42

Contributions of the paper

The GCR model introduces a novel approach to **link prediction** by transforming it into a **logical reasoning** problem, enabling the use of both associative learning and logical connections between links. This offers a fresh perspective on leveraging neighboring link information and could strengthen Elisa's work by providing a logical, collaborative framework that refines relational reasoning tasks within KGs.

Logical Connections Among Triplets: (x, capitalOf, y) → (x, locatedIn, y)

Triplets as Predicate Expressions: `capitalOf(x, y) → locatedIn(x, y)

Strengths

GCR's ability to handle uncertainty in logical reasoning, and its use of soft-logic constraints, aligns well with the goals of Elisa's TransHySeCo method. By incorporating neighbor link information without predefined logic rules, GCR's approach would allow Elisa's model to better handle cases with sparse or noisy data, ultimately leading to richer and more reliable embeddings.

soft-logic constraints: Soft logic allows rules to have flexibility by assigning probabilities, so rules can occasionally be violated without strict consequences.

In GCR, for example, if a user bought a phone, it likely—but not certainly—implies they might buy a phone case. Soft logic handles this uncertainty by treating the rule as probable rather than absolute.

Overview of the method

GCR employs **Graph Neural Networks (GNNs)** and **modularized neural logic networks**, translating link prediction tasks into **logical expressions**. These expressions allow for efficient learning by applying logical constraints across neighboring links, which Elisa's work could incorporate to enhance both prediction accuracy and training efficiency
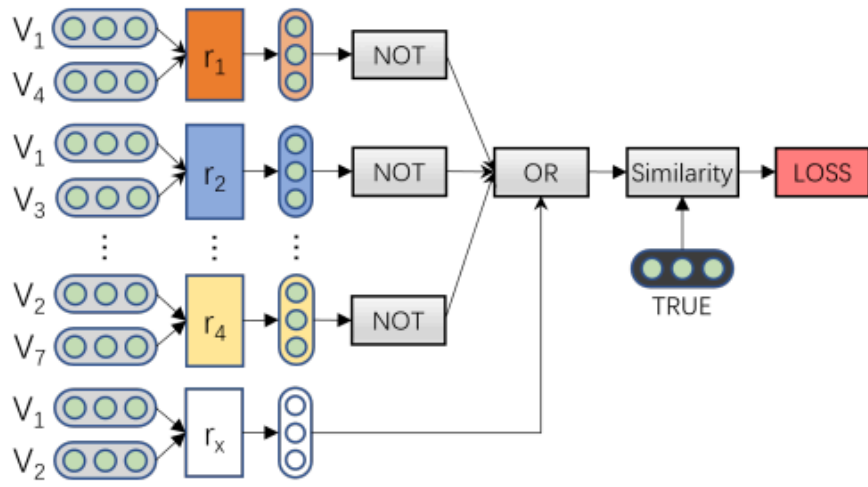
Figure 2: The logical network structure of the link prediction task given in Figure 1. The network is assembled using the logical equivalent expression which is converted via De Morgan's Law.

$$\neg(A \wedge B) = \neg A \vee \neg B$$

Weaknesses

GCR, however, may face scalability issues with very large datasets, particularly if high-order logical implications are necessary. This could limit its use in broader datasets like those Elisa may intend to apply her methods on. Addressing computational overheads would be crucial in combining these methodologies.

Areas of improvement / adjustments

Integrating GCR's collaborative reasoning approach with TransHySeCo could benefit from exploring multi-hop logic reasoning. This would improve predictions by capturing more complex, indirect relationships between entities, a potential enhancement to Elisa's ontology-based framework. Additionally, adjusting the regularization settings for sparsity (as seen in GCR) might optimize performance for sparse datasets

Code

Code is not available!!

KALE

Contributions of the paper

Unified Framework for Embedding: KALE proposes a framework that jointly embeds knowledge triples and logical rules. It treats triples as atomic formulae and logical rules as complex formulae, creating a unified embedding space that accommodates both

Incorporation of t-norm Fuzzy Logic: To handle the truth values of complex logical rules, KALE uses t-norm fuzzy logic, allowing a continuous range of truth values rather than binary outcomes. This is particularly useful for managing partial rule satisfaction

Enhanced Prediction for New Facts: By embedding triples and rules together, KALE improves its ability to predict new facts that might not be directly inferable through logical rules alone. This capability is crucial for knowledge graph completion
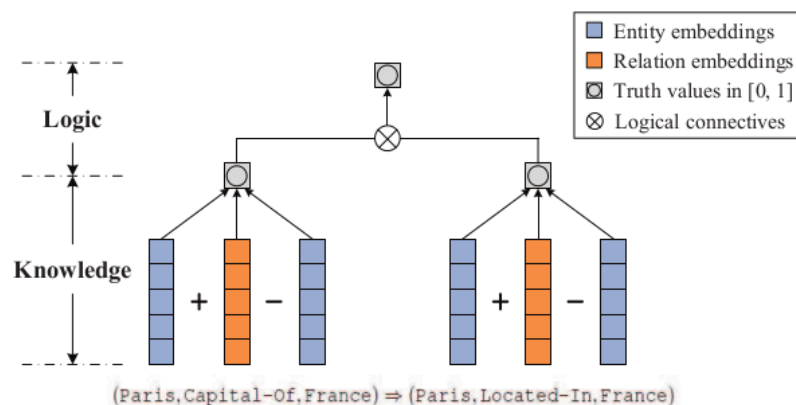

Figure 1: Simple illustration of KALE.

Strengths

Joint Modeling of Knowledge and Logic: KALE's ability to embed both triples and rules in a single framework allows it to better capture the relational structure of knowledge graphs, leading to more semantically meaningful embeddings

Overview of the method

Triple Modeling: KALE follows the translation-based approach used in TransE, where each relation is represented as a translation between entities. A triple's truth value is computed by measuring the distance between the head and tail entity embeddings after translation

Rule Modeling with t-norm Fuzzy Logic: Logical rules are represented as complex formulae. KALE uses t-norm fuzzy logic to calculate truth values of logical rules, combining atomic truth values (from triples) with logical operations like conjunction and implication

Joint Loss Function: KALE minimizes a global loss function over both triples and rules. Positive samples include observed triples and ground rules, while negative samples are generated by corrupting entities or relations. This loss function ensures that the embeddings are compatible with both observed facts and logical rules

and define its <mark>soft truth value</mark> as

$$I\left(e_i, r_k, e_j\right) = 1 - \frac{1}{3\sqrt{d}} \left\|\mathbf{e}_i + \mathbf{r}_k - \mathbf{e}_j\right\|_1, \quad (1)$$

t-norm fuzzy logics:

$$I(f_1 \wedge f_2) = I(f_1) \cdot I(f_2),$$
$$I(f_1 \vee f_2) = I(f_1) + I(f_2) - I(f_1) \cdot I(f_2),$$
$$I(\neg f_1) = 1 - I(f_1),$$

Joint Learning:

$$\min_{\{\mathbf{e}\},\{\mathbf{r}\}} \sum_{f^+ \in \mathcal{F}} \sum_{f^- \in \mathcal{N}_{f^+}} \left[\gamma - I(f^+) + I(f^-)\right]_+,$$

$$\text{s.t.} \quad \|\mathbf{e}\|_2 \leq 1, \forall e \in \mathcal{E}; \quad \|\mathbf{r}\|_2 \leq 1, \forall r \in \mathcal{R}. \quad (4)$$

| | | Test-I | | | | | Test-II | | | | | Test-ALL | | | | |
| | | MRR | MED | HITS@N (%) | | | MRR | MED | HITS@N (%) | | | MRR | MED | HITS@N (%) | | |
				3	5	10			3	5	10			3	5	10
FB122	TransE	0.220	29.0	25.7	32.4	40.6	0.296	5.0	40.0	50.8	57.8	0.262	10.0	33.6	42.5	50.0
	TransH	0.218	29.0	25.0	31.3	39.2	0.297	6.0	37.5	48.5	56.3	0.249	12.0	31.9	40.7	48.6
	TransR	0.219	31.0	24.7	30.8	38.9	0.273	9.0	32.4	42.8	51.6	0.261	15.0	28.9	37.4	45.9
	KALE-Trip	0.201	25.0	23.9	31.6	40.1	0.309	5.0	40.9	51.3	58.0	0.261	11.0	33.3	42.4	50.0
	KALE-Pre	0.203	25.0	24.1	31.7	40.2	**0.368**	**4.0**	**47.3**	**55.4**	**61.4**	**0.294**	**9.0**	**36.9**	**44.8**	**51.9**
	KALE-Joint	**0.229**	**21.0**	**26.3**	**33.8**	**42.2**	0.357	**4.0**	44.0	53.0	59.3	0.299	**9.0**	36.1	44.3	51.6
WN18	TransE	0.248	**4.0**	40.9	60.6	77.0	0.363	3.0	59.4	70.8	81.4	0.331	3.0	54.3	67.9	80.2
	TransH	0.242	**4.0**	39.2	60.1	75.9	0.482	**2.0**	63.5	70.8	79.3	0.415	3.0	56.7	67.8	78.3
	TransR	0.240	**4.0**	40.1	57.7	74.3	0.449	3.0	55.7	64.5	74.3	0.391	3.0	51.3	62.6	73.5
	KALE-Trip	0.250	**4.0**	40.6	62.3	78.1	0.393	**2.0**	61.9	71.2	80.6	0.353	3.0	56.0	68.7	79.9
	KALE-Pre	0.248	**4.0**	40.4	61.5	78.2	0.451	3.0	**69.6**	**77.5**	**85.3**	0.395	3.0	**61.4**	**73.0**	**83.3**
	KALE-Joint	**0.260**	**4.0**	**43.6**	**64.1**	79.2	**0.563**	**2.0**	67.6	73.8	81.0	**0.478**	**2.0**	60.9	71.1	80.5

Table 4: Link prediction results on the test-I, test-II, and test-all sets of FB122 and WN18 (raw setting).

Weaknesses

Complexity in Rule Grounding: KALE requires grounding of universally quantified rules, which can be computationally intensive, especially with large datasets

Dependency on Rule Quality: Since KALE relies on logical rules for embedding, its performance depends heavily on the quality of these rules. Poorly defined or incomplete rules may lead to suboptimal embeddings

Areas of improvement / adjustments

KALE's main innovation over Elisa's approach lies in its unified framework for embedding both knowledge triples and logical rules, its use of t-norm fuzzy logic for flexible truth value handling, and a joint loss function that enhances predictive capabilities even for complex, non-directly inferable facts. These innovations allow KALE to generate embeddings that are more robust and semantically informed than those from methods that rely solely on triples or on strict ontological consistency.

Code