Week 9 : Assignment 9

The due date for submitting this assignment has passed.

Due on 2025-03-26, 23:59 IST.

Assignment submitted on 2025-03-18, 19:44 IST

1) Which of the following statement best describes why knowledge graphs (KGs) are considered more powerful than a traditional relational knowledge base (KB)? ***1 point***

- KGs require no schema, whereas KBs must have strict schemas.
- KGs store data only in the form of hypergraphs, eliminating redundancy.
- KGs allow flexible, graph-based connections and typed edges, enabling richer relationships and inferences compared to KBs.
- KGs completely replace the need for textual sources by storing all possible facts.

Yes, the answer is correct.
Score: 1
Accepted Answers:
KGs allow flexible, graph-based connections and typed edges, enabling richer relationships and inferences compared to KBs.

2) Entity alignment and relation alignment are crucial between KGs of different languages. Which of the following factors contribute to effective alignment? ***1 point***

- Aligning relations solely by their lexical similarity, ignoring semantic context
- Transliteration or language-based string matching for entity labels
- Ensuring all language aliases are represented identically in each KG
- Matching neighbours, or connected entities, across different KGs

Yes, the answer is correct.
Score: 1
Accepted Answers:
Transliteration or language-based string matching for entity labels
Matching neighbours, or connected entities, across different KGs

3) In the context of knowledge graph completion (KGC), which statement best describes the role of the scoring function f(s,r,o)? ***1 point***

- ○ It determines whether two entities refer to the same real-world concept.
- ◉ It produces a raw confidence score indicating how plausible a triple (s,r,o) is.
- ○ It explicitly encodes only the subject's embedding, ignoring the relation and object embeddings.
- ○ It ensures that every negative triple gets a higher score than any positive triple.

Yes, the answer is correct.
Score: 1
Accepted Answers:
It produces a raw confidence score indicating how plausible a triple (s,r,o) is.

4) One key difference between the differentiable KG approach and the semantic interpretation approach to KGQA is: ***1 point***

- ○ Differentiable KG approaches are fully rule-based, while semantic interpretation is purely neural.
- ○ Differentiable KG approaches do not require any graph embeddings, relying instead on explicit logical forms.
- ○ Semantic interpretation is more transparent or interpretable, whereas differentiable KG is end-to-end trainable but less interpretable.
- ◉ Both approaches use logical forms; the primary difference is the type of question they can answer.

No, the answer is incorrect.
Score: 0
Accepted Answers:
Semantic interpretation is more transparent or interpretable, whereas differentiable KG is end-to-end trainable but less interpretable.

5) Considering the differentiable KG approach, which elements are typically learned jointly when training an end-to-end KGQA model? ***1 point***

- ☑ The textual question representation (e.g., BERT embeddings)
- ☑ The graph structure encoding (e.g., GCN or transformer-based graph embeddings)
- ☐ Predefined logical forms to ensure interpretability
- ☑ The final answer selection mechanism that identifies which node(s) in the graph satisfy the question

Yes, the answer is correct.
Score: 1
Accepted Answers:
The textual question representation (e.g., BERT embeddings)
The graph structure encoding (e.g., GCN or transformer-based graph embeddings)
The final answer selection mechanism that identifies which node(s) in the graph satisfy the question

6) Uniform negative sampling can have high variance and may require large number of samples. Why is that the case? ***1 point***

- ○ Because the margin-based loss cannot converge without big mini-batches.
- ◉ Because randomly picking negative entities does not guarantee close or challenging negatives, causing unstable training estimates.
- ○ Because negative sampling must ensure every possible negative triple is covered.
- ○ Because the number of relations in the KG is too large for small number of samples.

Yes, the answer is correct.
Score: 1
Accepted Answers:
Because randomly picking negative entities does not guarantee close or challenging negatives, causing unstable training estimates.

7) In testing embedding and score quality for KG completion, mean rank and hits@K are typical metrics. What does hits@K specifically measure in this context? ***1 point***

- (●) The percentage of queries for which the correct answer appears in the top-K of the ranked list.
- ( ) The reciprocal of the rank of the correct answer.
- ( ) The probability of the correct answer appearing as the highest scored candidate.
- ( ) The margin of the correct triple score relative to all negative triples.

Yes, the answer is correct.
Score: 1
Accepted Answers:
The percentage of queries for which the correct answer appears in the top-K of the ranked list.

8) In the TransE model, the scoring function for a triple (s,r,o) is typically defined as ***1 point***

$f(s,r,o)=\|e_s + e_r - e_o\|$

where e_s, e_r, e_o are embeddings of the subject, relation, and object, respectively. Which statement best explains what a low value of f(s,r,o) indicates in this context?

- ( ) That (s,r,o) is an invalid triple according to the learned embeddings.
- ( ) That e_s and e_o must be orthogonal.
- ( ) That the relation embedding e_r is zero.
- (●) That (s,r,o) has a high likelihood of being a true fact in the knowledge graph.

Yes, the answer is correct.
Score: 1
Accepted Answers:
That (s,r,o) has a high likelihood of being a true fact in the knowledge graph.

9) In RotatE, if a relation r is intended to be symmetric, how would that typically manifest in the complex plane? ***1 point***

- ( ) The relation embedding e_r must always equal zero.
- (●) The angle of e_r must be $\pi/2$.
- ( ) The relation embedding e_r is its own inverse (i.e., a 180° rotation when squared).
- ( ) The magnitude of e_r must be greater than 1.

No, the answer is incorrect.
Score: 0
Accepted Answers:
The relation embedding e_r is its own inverse (i.e., a 180° rotation when squared).

10) Which main advantage do rotation-based models (like RotatE) have over translation-based ones (like TransE) when it comes to complex multi-relational patterns in a KG? ***1 point***

- ( ) Rotation-based models cannot model any symmetry or inverse patterns, so they are simpler.
- (●) Rotation-based models handle a broader set of relation properties (symmetry, anti-symmetry, inverses, composition) more naturally.
- ( ) Rotation-based models have no hyperparameters to tune, unlike TransE.
- ( ) Rotation-based models are guaranteed to yield perfect link prediction.

Yes, the answer is correct.
Score: 1
Accepted Answers:
Rotation-based models handle a broader set of relation properties (symmetry, anti-symmetry, inverses, composition) more naturally.