Multilingual Knowledge Graph Completion with Zero Seed Alignment

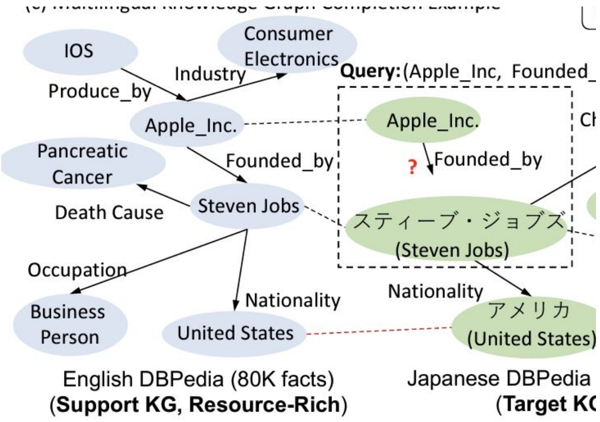
Group 11 – Rakesh Bal, Ashwath Radhachandran, Rahul Kapur, Rustem Aygun

Outline

- Background
- Our Goals
- Related Works
- SS-AGA Model
- Approach 1
  - Experiments + Results
- Approach 2
  - Experiments + Results
- Approach 3
- Conclusion + Future Work

Background

- The knowledge graph completion (KGC) task is an important research area
  - Most knowledge graphs (KGs) are often incomplete since human annotation is *expensive*
- The problem of KG incompleteness is amplified when working with multilingual data
  - Human annotations are challenging to acquire (For example - Greek KG in DBP-5L)
- The lack of seed alignment prevents effective knowledge transfer.
  - Seed alignment: mappings between similar concepts in two KGs
- The acquisition of alignment labels can be a costly and noisy process, but using unsupervised methods (correlation or similarity between entities) can help



Our Goals

- Investigate zero seed alignment setting
  - Involves eliminating pre-existing alignment pairs for knowledge transfer
  - More difficult, but more practical for real world situations
- Can we develop an effective unsupervised technique for generating reliable alignment pairs?
- Motivation: most current KGs are incomplete, and being able to fill in missing facts in these KGs can improve current technology (such as Google's SEOs).

Related Work

- Existing monolingual KG learning approaches perform poorly in low-resource language due to the sparseness
    - [Peng et al., 2021; Xu et al., 2021; Liang et al., 2021; Cao et al., 2021; Lovelace et al., 2021]
- KGs learned from multiple languages share some real-world entities and relations.
- Knowledge from a rich resource language can be transferred to low-resource language.

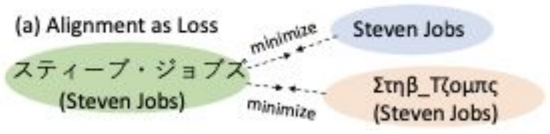
Related Work

Existing Multilingual Knowledge Transfer Approaches:

Embedding each language specific KG independently then employ an alignment loss to force pairs of aligned entities to be close maximally [Chen et al., 2018; Sun et al., 2018; Chen et al., 2017]

Problems with these approaches:

- Equal treatment of parallel entities cause KG inconsistency issue.
- Limited seed alignments results in poor knowledge transfer across languages.



SS-AGA Model

Current SoTA MKGC approach: Self-Supervised Adaptive Graph Alignment (SS- AGA) [Huang et al., 2022]

SS-AGA handles:

- the limited seed alignment,
- equal treatment of parallel entities

problems via GNN based self-supervised seed generation technique.

We use SS-AGA for the evaluation of our zero seed alignment technique in all experiments.

SS-AGA Model

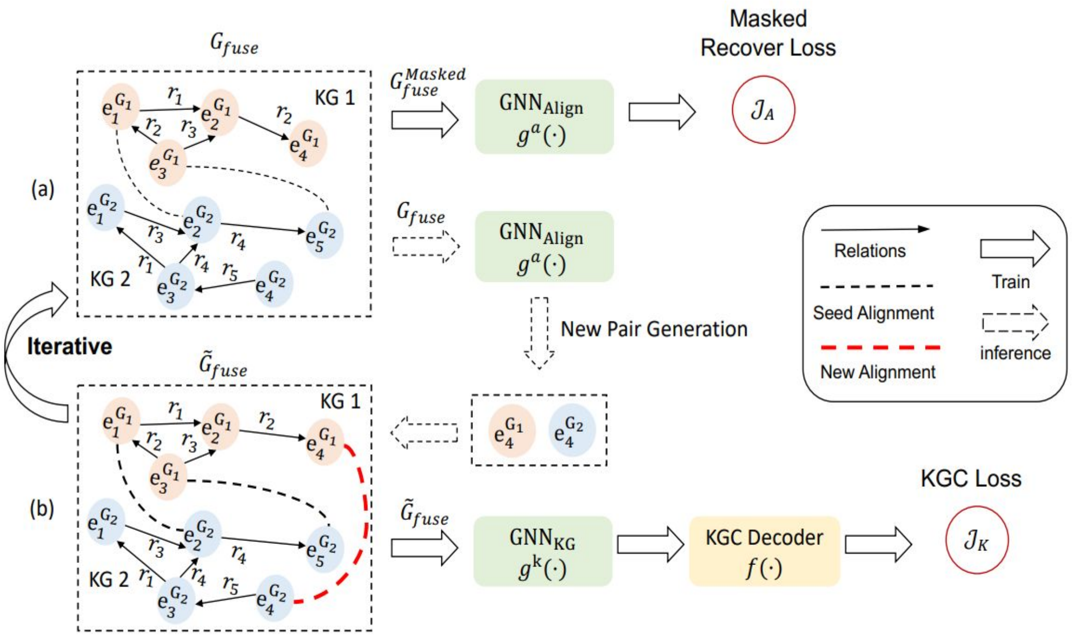


Figure 2: The overall framework of the Self-Supervised Adaptive Graph Alignment (SS-AGA).

- KGs from different languages are fused into a whole graph
- Seed alignments are considered as a new edge type between parallel entities across languages

SS-AGA Model

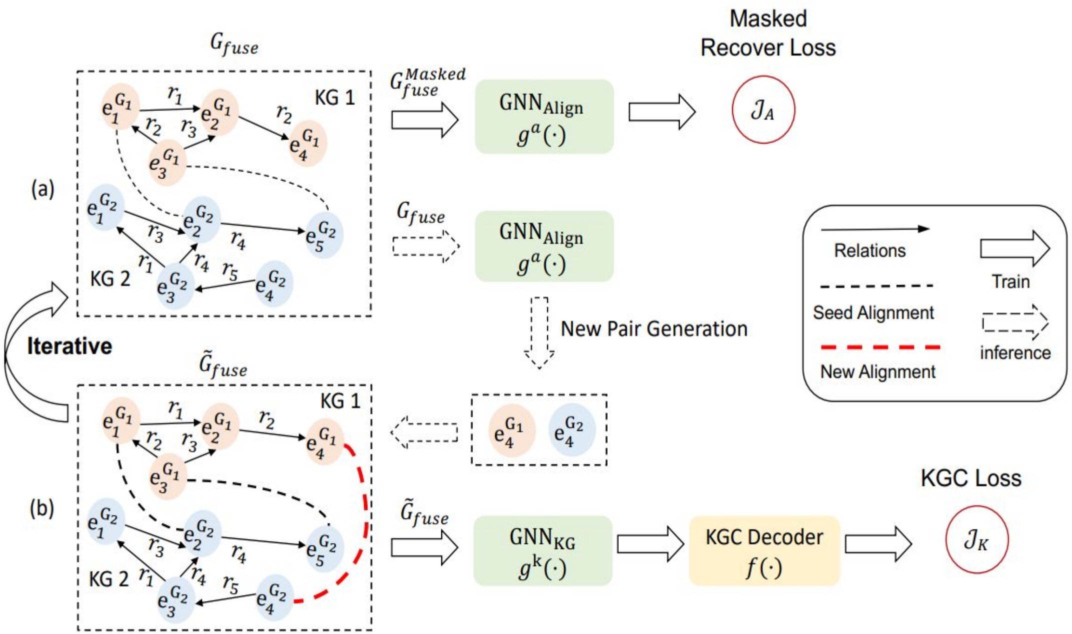
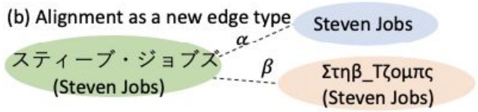


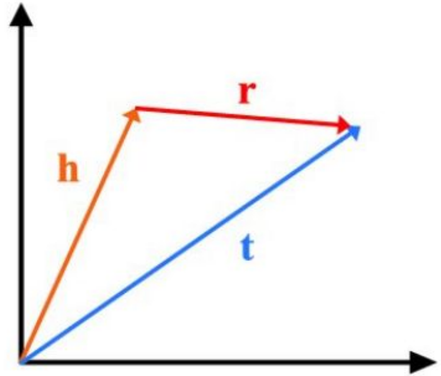
Figure 2: The overall framework of the Self-Supervised Adaptive Graph Alignment (SS-AGA).

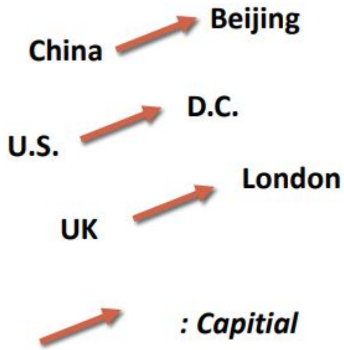
- GNN encoder with a relation-aware attention based aggregation function
- Generates new pairs iteratively via self-supervised learning



TransE

- Relation: translating embedding





- Score function
  - $f_r(\mathbf{h},\mathbf{t}) = -||\mathbf{h}+\mathbf{r}-\mathbf{t}|| = -d(\mathbf{h}+\mathbf{r},\mathbf{t})$

- Margin-based ranking loss
  - $L = \sum_{(h,r,t)\in S} \sum_{(h',r,t')\in S'_{(h,r,t)}} [\gamma + d(\mathbf{h}+\mathbf{r},\mathbf{t}) - d(\mathbf{h}'+\mathbf{r},\mathbf{t}')]_+$

$[x]_+$ denotes the positive part of xx, i.e., max(0, x)

$\gamma > 0$ denotes the margin hyperparameter

S: positive triple set; S' : corrupted triple set

Bordes et al., Translating embeddings for modeling multi-relational data, NeurIPS 2013

Dataset and Evaluation Metrics

- *DBP- 5L*: 5 language-specific KGs from DBpedia –English, French, Spanish, Japanese and Greek
- Dataset Statistics

Dataset	#Entity	#Relation	#Triple	#Aligned Links
Multilingual Academic KG (*DBP-5L*)				
EN	13,996	831	80,167	16,916
FR	13,176	178	49,015	16,877
ES	12,382	144	54,066	16,347
JA	11,805	128	28,774	16,263
EL	5,231	111	13,839	9,042

- Evaluation Metrics - Hits@1 (Accuracy) , Hits@10, MRR (Mean Reciprocal Rank)
- We use this dataset and evaluation metrics in all of our experiments
- TransE and SS-AGA are considered baseline in all experiments

Need for new methods for seed generation

- Generating new seed alignment pairs in an unsupervised fashion is key
- We propose three different ways of generating seeds without using any prior alignment.

Approach 1

- Use PreTrained multilingual BERT model for generation of Seed Alignments
- Textual representations of entities are used for calculating the bert embeddings
- The cosine similarity is calculated for all entities in one lang to all entities in another lang – 2D matrix for each pair of entities
- There will be 5C2 = 10 pairs of seed alignments
- Top ***k*** pairs are picked of entities in each pair of languages as seed alignments
- Selection of ***k*** – all pairs having ‘el’ as one lang: 2000, rest 6 pairs: 5000
- Feed these generated seed alignments to SS-AGA model instead of the annotated ones

Approach 1: SS-AGA Model + BERT Gen Seeds

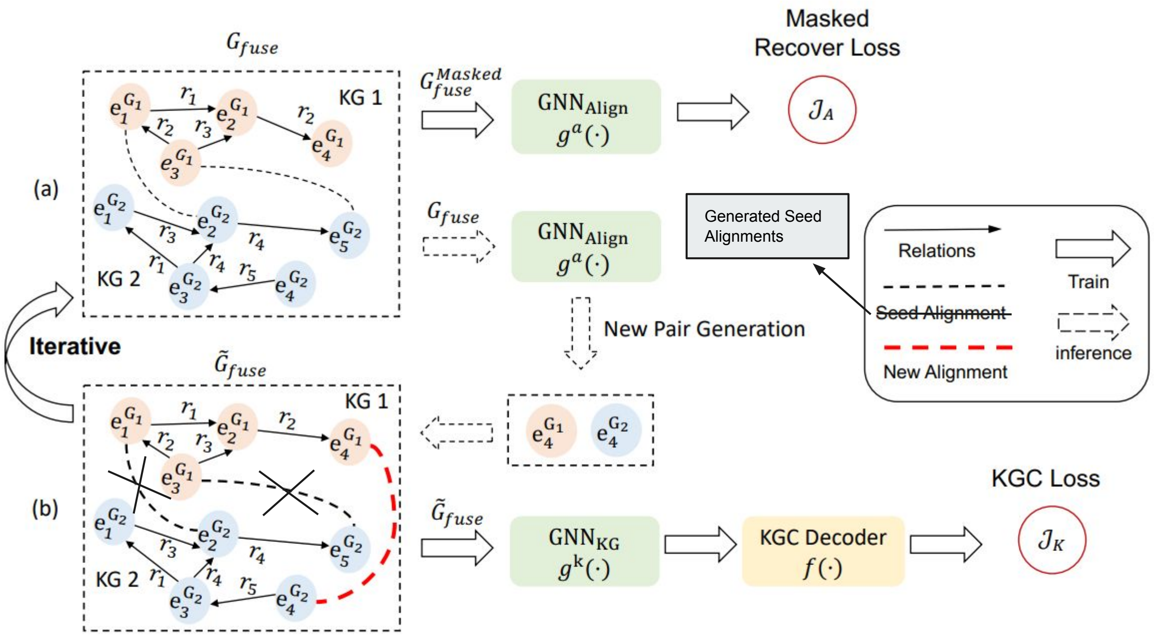
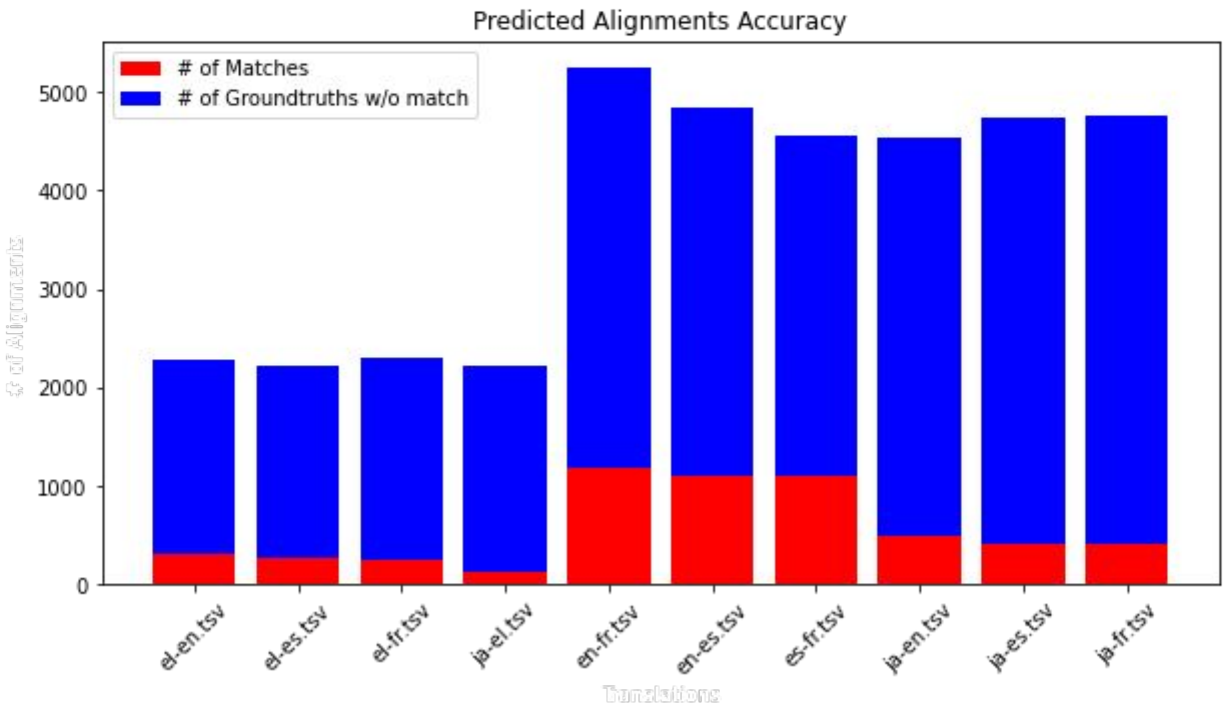


Figure 2: The overall framework of the Self-Supervised Adaptive Graph Alignment (SS-AGA).

Approach 1: Results

Method	Metric	JA	EL	ES	FR
Monolingual Baselines					
TransE	H@1	21.1	13.1	13.5	17.5
	H@10	48.5	43.7	45	48.8
	MRR	25.3	24.3	24.4	27.6
Multilingual Baselines					
SS-AGA	H@1	34.6	30.8	25.5	27.1
	H@10	66.9	58.6	61.9	65.5
	MRR	42.9	35.3	36.6	38.4
SS-AGA + Bert Gen Seeds (Approach 1)	H@1	17.9	16.8	14.1	17.2
	H@10	45.5	57.9	50.4	53
	MRR	27.7	30.9	26.4	29.1

Unsupervised Generation of Alignments - Results



Approach 2

- Remove all seeds from the seed alignments
- Let the model generate seeds for some epochs (hyperparameter - *m*)
- Freeze the alignment model for those *m* epochs
- Generates new seeds for *m* epochs and then starts training the alignment model
- Rest procedure same

Approach 2: SS-AGA Model without any seeds

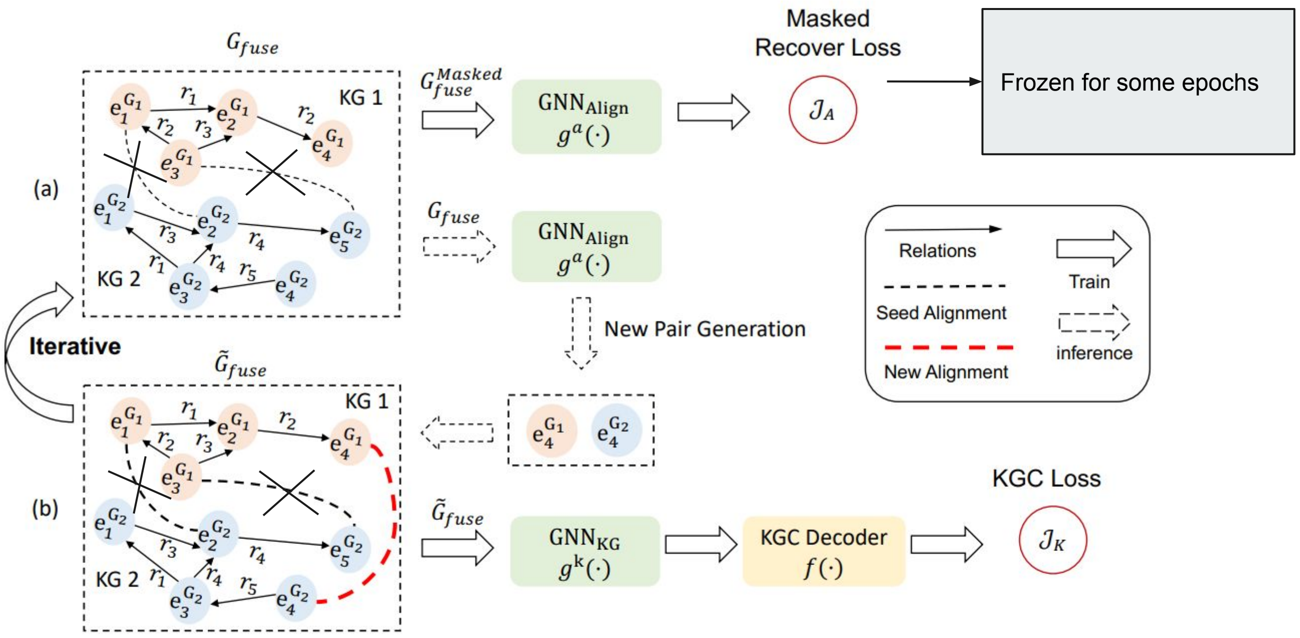


Figure 2: The overall framework of the Self-Supervised Adaptive Graph Alignment (SS-AGA).

Approach 2: Results

Method	Metric	JA	EL	ES	FR
Monolingual Baselines					
TransE	H@1	21.1	13.1	13.5	17.5
	H@10	48.5	43.7	45	48.8
	MRR	25.3	24.3	24.4	27.6
Multilingual Baselines					
SS-AGA	H@1	34.6	30.8	25.5	27.1
	H@10	66.9	58.6	61.9	65.5
	MRR	42.9	35.3	36.6	38.4
SS-AGA + Bert Gen Seeds	H@1	17.9	16.8	14.1	17.2
	H@10	45.5	57.9	50.4	53
	MRR	27.7	30.9	26.4	29.1
SS-AGA + No Seeds	H@1	20.4	17.3	9.7	13.3
	H@10	51.9	59.8	52.5	55.2
	MRR	31.1	31.1	23.4	27.1

Approach 2: Results on Hyperparameter m

Method	Metric	m = 0	m = 3	m = 6	m = 9	m = 12
SS-AGA + No Seeds (EL Dataset)	H@1	*14.9*	*17.3*	*17.5*	*14.7*	*15.7*
	H@10	*55.7*	*59.8*	*56.9*	*56.7*	*58.6*
	MRR	*29.6*	*31.1*	*30.8*	*29.8*	*30.5*

Approach 3: Work in Progress

- Add mBERT to the SS-AGA model and train the entire module end to end
- This will fine tune the mBERT model to produce better structural embeddings
- Theoretically can lead to better results

Approach 3: SS-AGA Model + mBERT fused

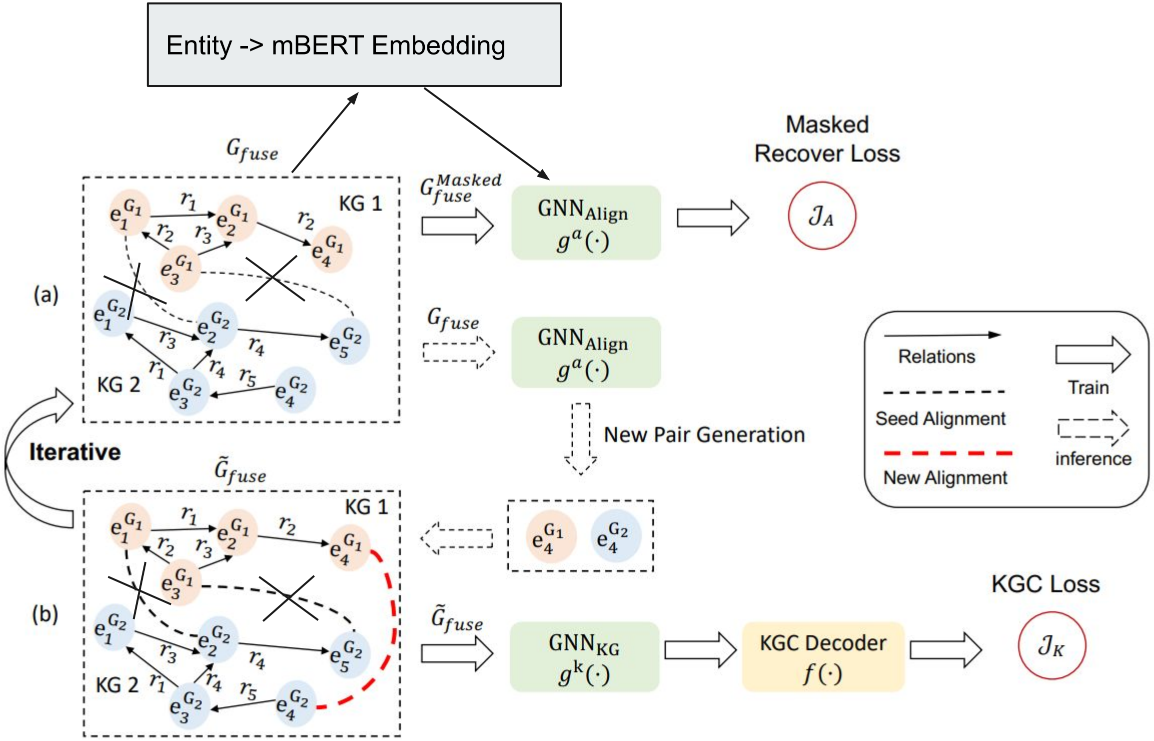


Figure 2: The overall framework of the Self-Supervised Adaptive Graph Alignment (SS-AGA).

Challenges

- GPU Resource problem – large graphs with very large models (SS-AGA, BERT)
- More compute could have helped us experiment faster

  – models took longer time to compute
- Couldn’t experiment with English dataset due to CUDA out of memory error
- Couldn’t experiment with mBERT model fused with SS-AGA due to *CUDA out of memory* error

Conclusion and Future Works

- Some of the unsupervised zero seed alignment methods are at par/better than the monolingual baseline *TransE*.
- Very few results at par with SS-AGA model (for el dataset which has least seeds)
- Getting better results can lead to great practical implications
- Future work includes experimenting with architectures without any seed generation
  - Getting rid of fused KG
  - Using a single encoder for all languages which will learn all lang KGs properly

Acknowledgments

We would like to thank ***Zijie*** and ***Roshni*** for their continued support throughout the project.

Thank You.
Q & A