Probabilistic Entity Representation Model for Reasoning over Knowledge Graphs

Authors: Nurendra Choudhary, Nikhil Rao, Sumeet Katariya, Karthik Subbian and Chandan K. Reddy
Presenter: Nurendra Choudhary

Outline

1. Introduction
2. Background
3. Our Solution
4. Evaluation
5. Conclusion

Introduction

Knowledge Graphs are ubiquitous data structures.

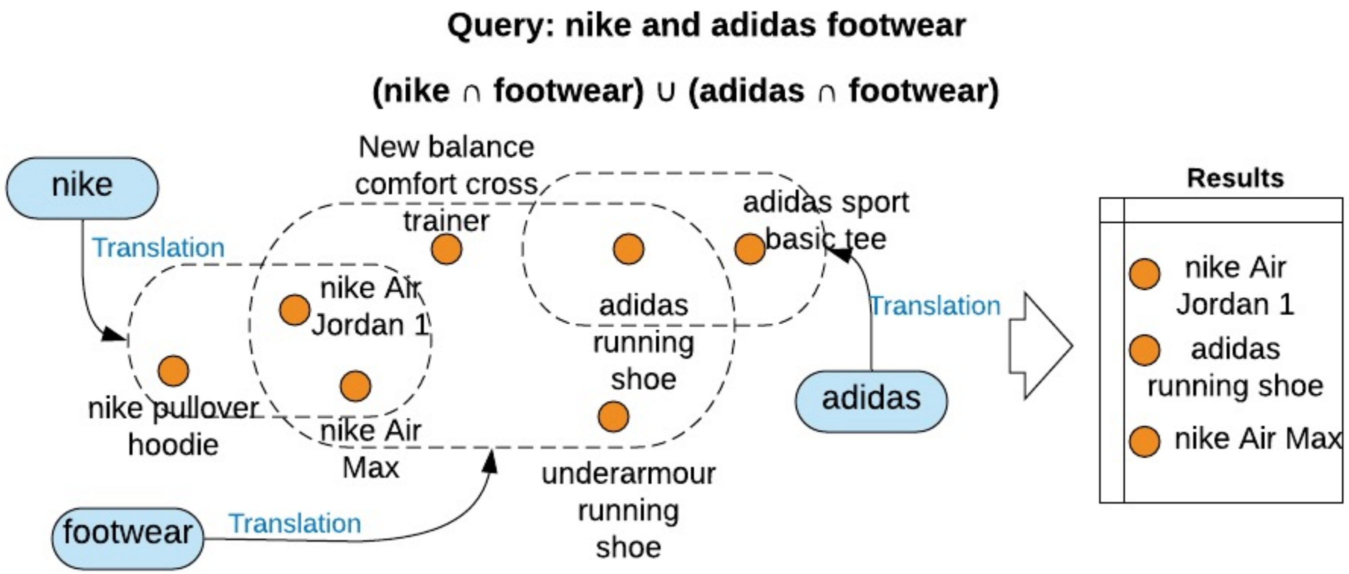
KG querying is computationally expensive due to its size (≈10M nodes with trillions of relations).

Introduction

Representation Learning can help!

Learn representations of entities and relations in a latent space.

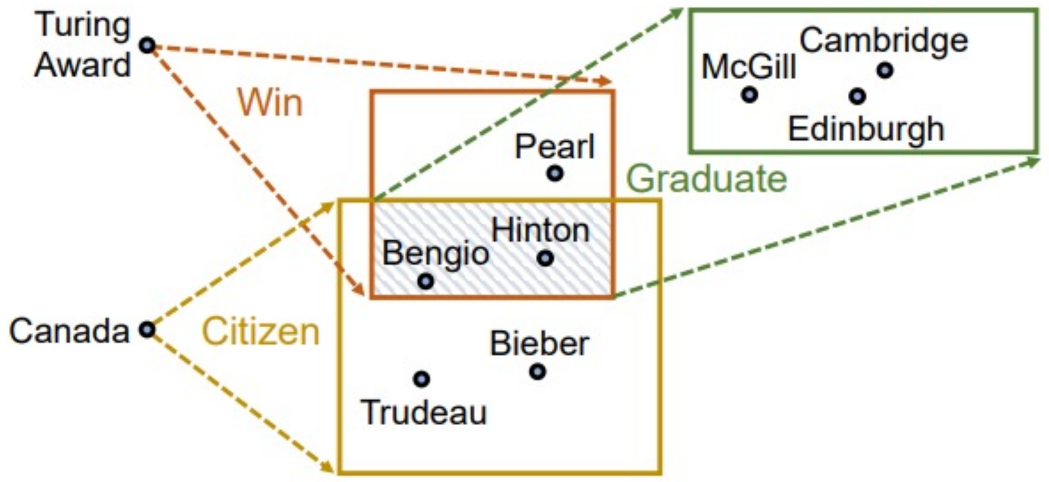
Apply logical operators to simulate querying behavior.



Background

Spatial Representations are better for modelling Knowledge Graphs.

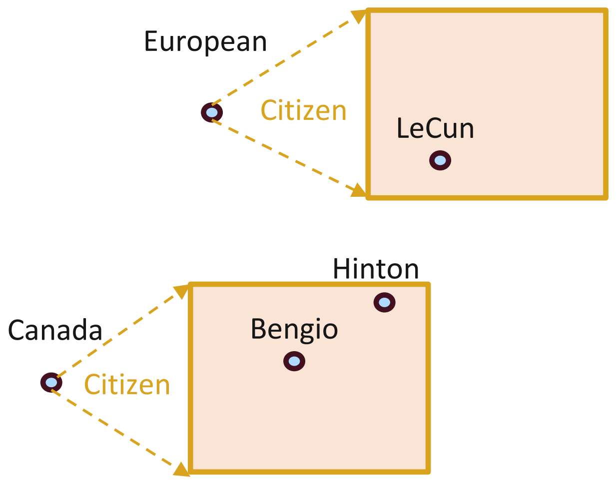
From which **universities** did the **Canadian Turing Award winners graduate**?



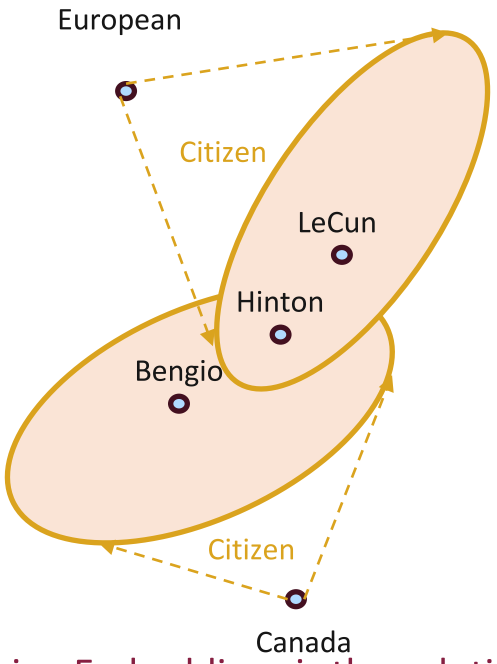
Box Embeddings (Query2Box)

Two challenges: (i) Non-smooth border (ii) Non-closure of union operation

Union of European-Canadian citizens



Box Embeddings

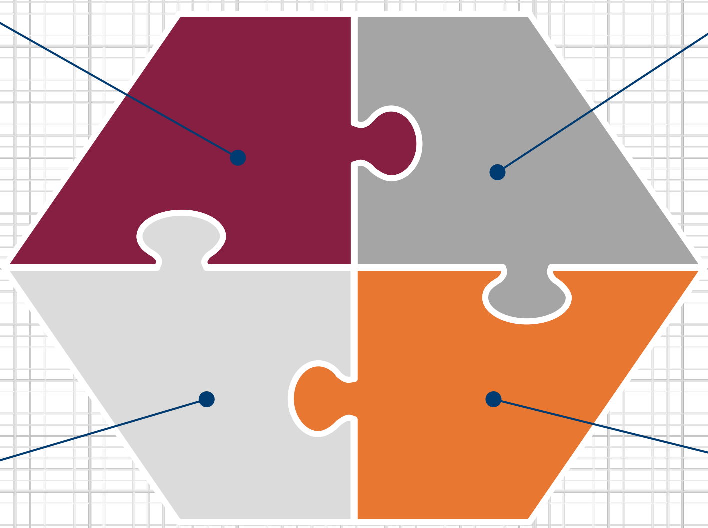


Gaussian Embeddings is the solution

Our Solution:
Multivariate Gaussian Representations

Capture spatial features of entities and relations.

Smooth border through Mahalanobis distance



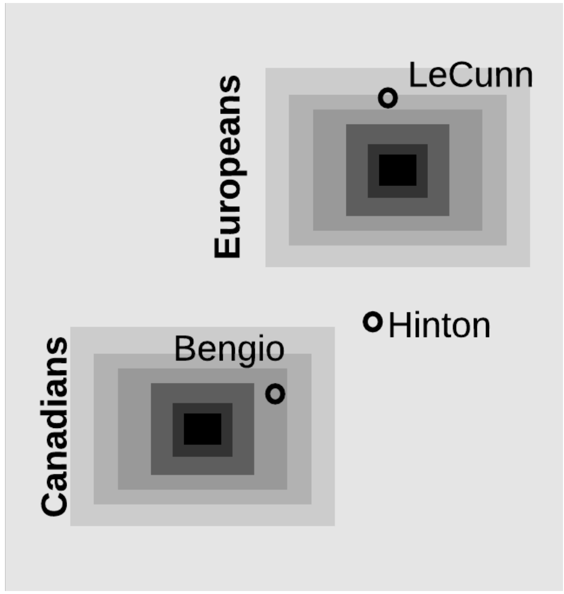
Closure of union operation by Gaussian mixtures

Scalability through chain reasoning

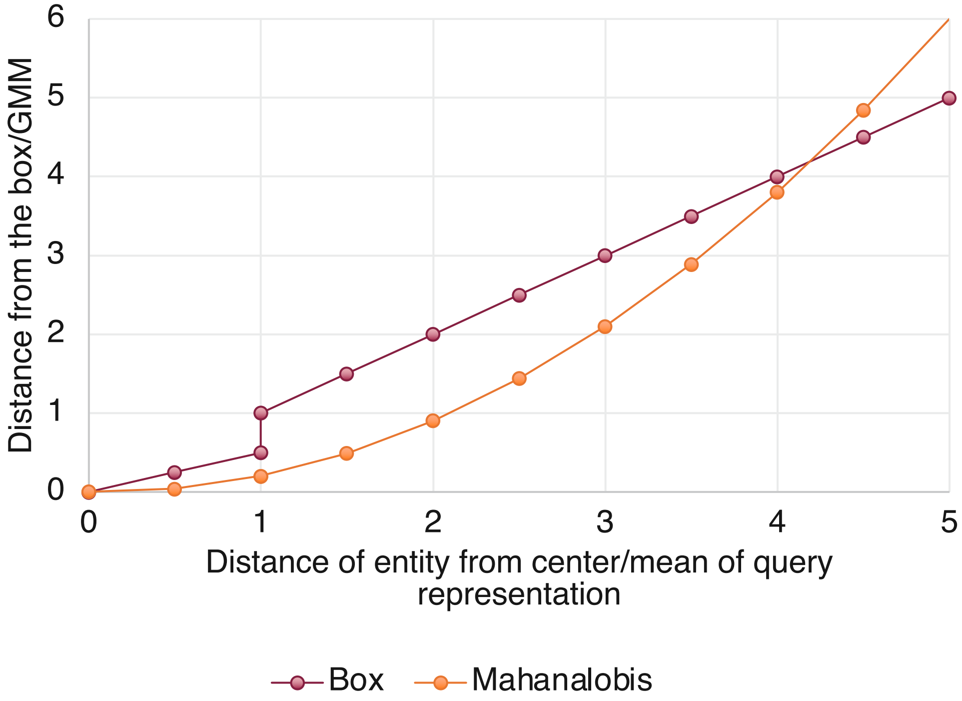
Our Solution

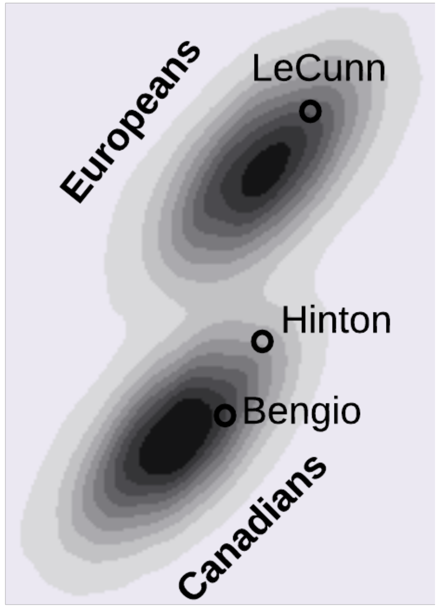
Challenge: Smooth Border

Mahalanobis Distance is a natural choice for a smooth distance function for Multivariate Gaussian Mixtures.



Box Embeddings
(Query2Box)



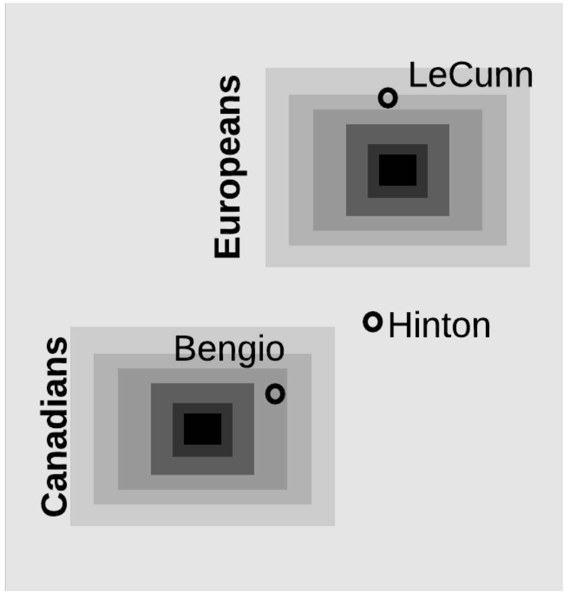


Gaussian Mixtures
(PERM)

Our Solution

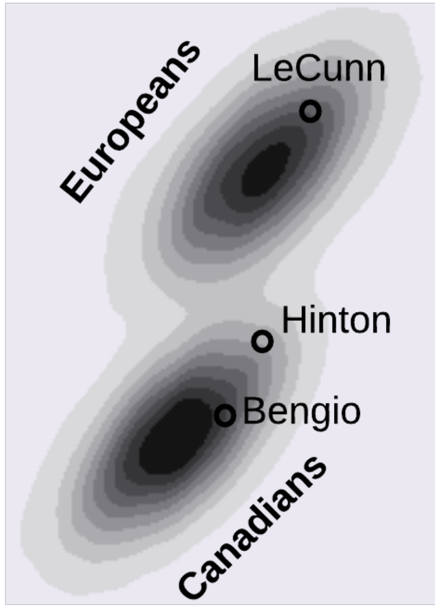
Challenge: Smooth Border

Mahalanobis Distance is a natural choice for a smooth distance function for **Multivariate Gaussian Mixtures**.



Box Embeddings
(Query2Box)

	Entity	Bengio	LeCun	Hinton
	Citizen	Canadian	French	British-Canadian
Union of Boxes	European	1	**0.211**	1.343
	Canadian	**0.146**	1.569	1.114
	Union	**0.146**	**0.211**	1.114
Gaussian Mixture	European	1	**0.215**	1.347
	Canadian	**0.146**	1.57	1.118
	Union	**0.255**	**0.626**	**0.766**



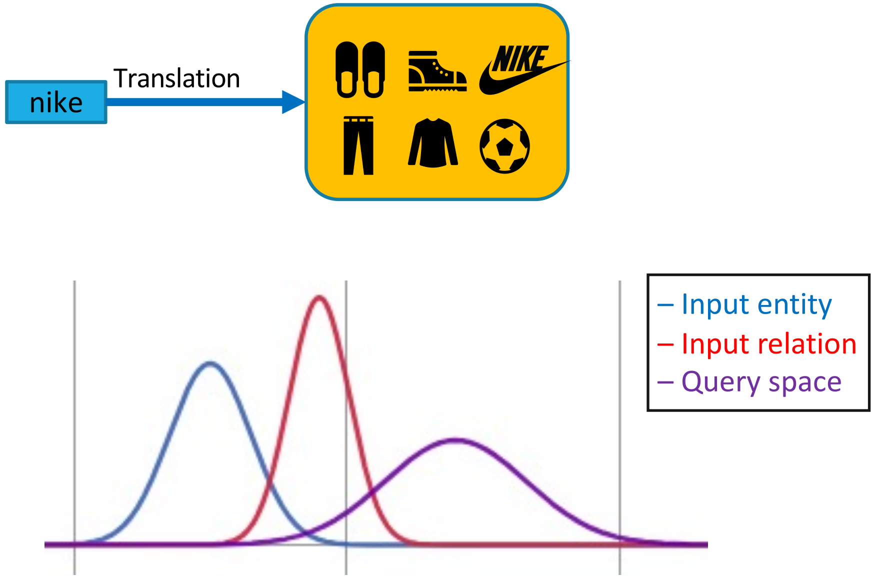
Gaussian Mixtures
(PERM)

Our Solution

Challenge: Closed solution to operations

Translation Queries: Gives all children of a query

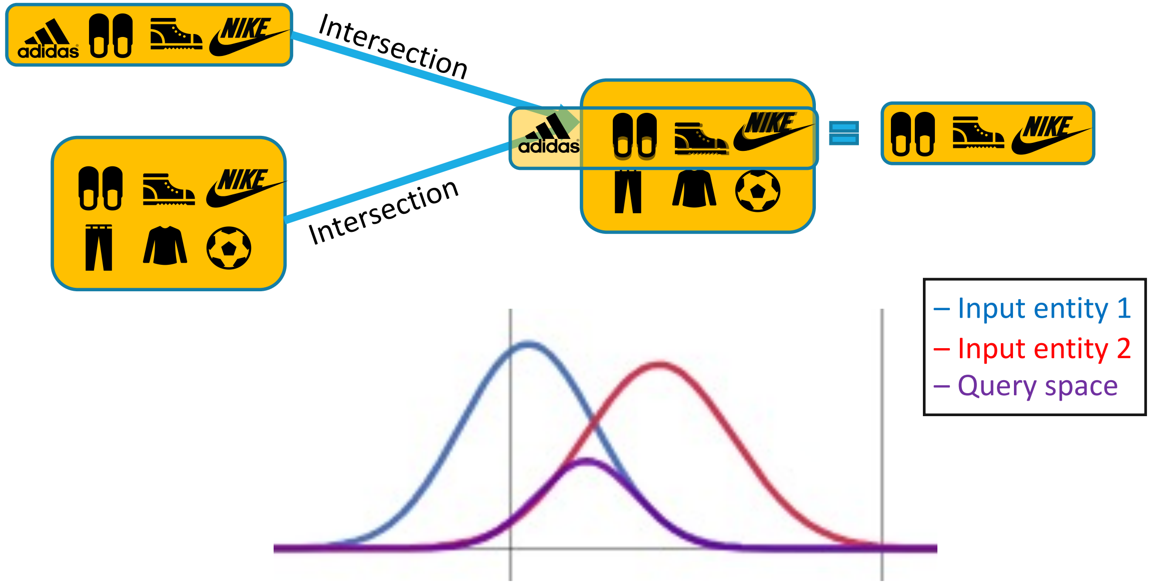
Q: nike



Gaussian Translation

Intersection: Gives intersection for two queries.

Q: nike footwear = nike ∩ footwear



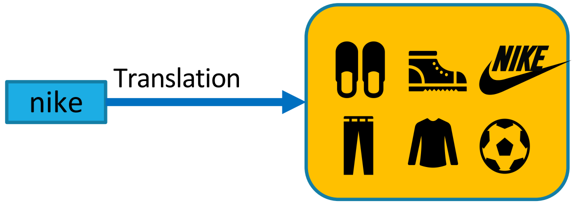
Gaussian Intersection

Our Solution

Challenge: Closed solution to operations

Translation Queries: Gives all children of a query

Q: nike

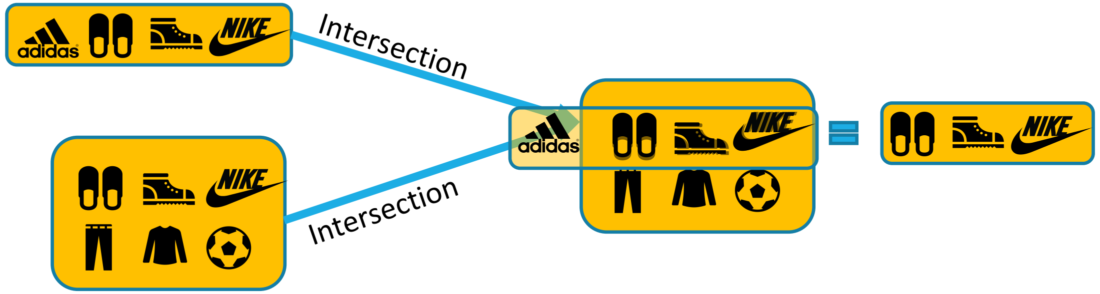


$$q_t = \mathcal{N}(\mu_e + \mu_r, (\Sigma_e^{-1} + \Sigma_r^{-1})^{-1});$$

Gaussian Translation

Intersection: Gives intersection for two queries.

Q: nike footwear = nike ∩ footwear



$$q_\cap = \mathcal{N}(\mu_{e_1}, \Sigma_{e_1})\mathcal{N}(\mu_{e_2}, \Sigma_{e_2}) = \mathcal{N}(\mu_3, \Sigma_3);$$

where, $\Sigma_3^{-1} = \Sigma_1^{-1} + \Sigma_2^{-1}$

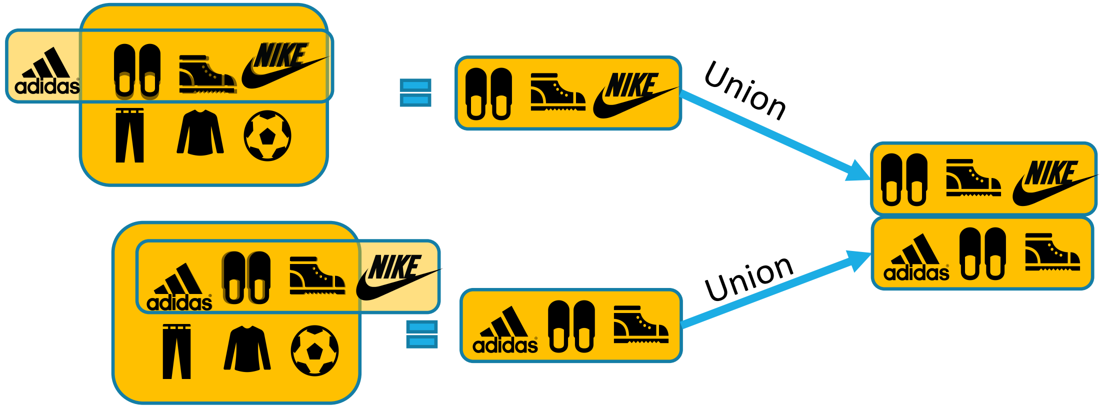
and $\mu_3 = \Sigma_3(\Sigma_2^{-1}\mu_1 + \Sigma_1^{-1}\mu_2) \implies \Sigma_3^{-1}\mu_3 = \Sigma_2^{-1}\mu_1 + \Sigma_1^{-1}\mu_2$

Gaussian Intersection

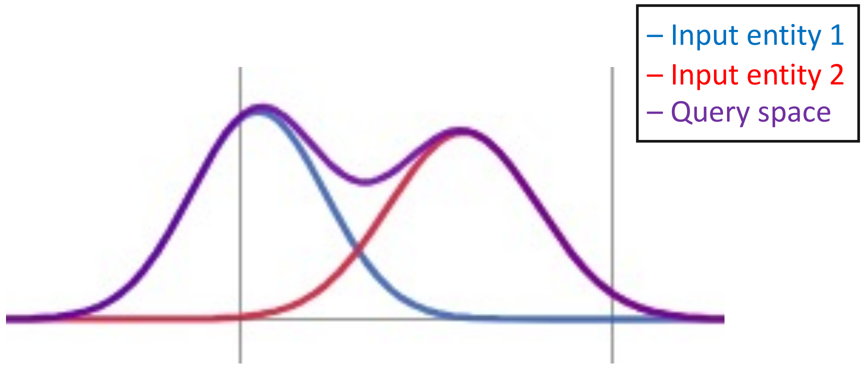
Our Solution

Challenge: Closed solution to operations

Union Queries: Gives union of two queries.
Q: (nike ∪ adidas) ∩ footwear = (nike ∩ footwear) ∪ (adidas ∩ footwear)



Unions can be modelled using Gaussian mixtures



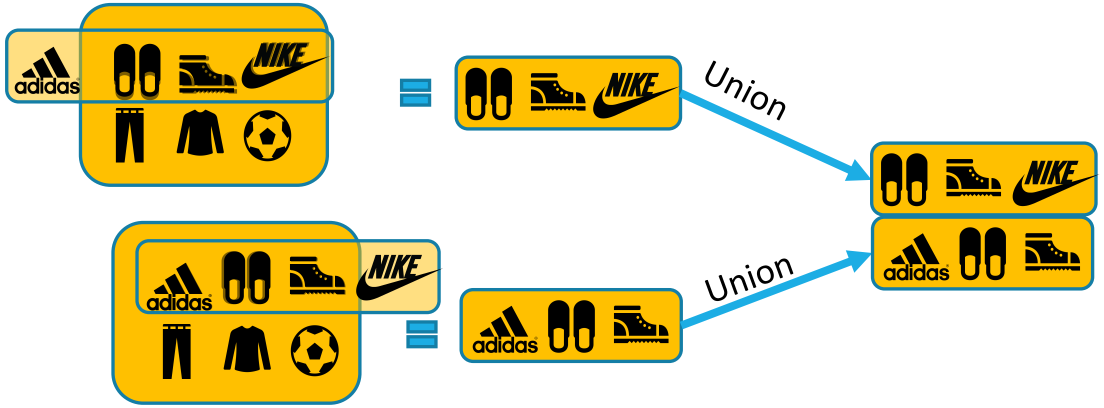
Gaussian Union

Our Solution

Challenge: Closed solution to operations

Union Queries: Gives union of two queries.

Q: (nike ∪ adidas) ∩ footwear = (nike ∩ footwear) ∪ (adidas ∩ footwear)



Unions can be modelled using Gaussian mixtures

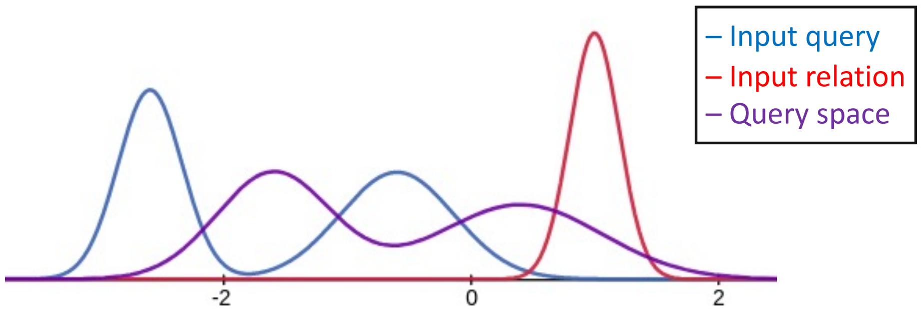
$$q_{\cup} = \sum_{i=1}^{n} \phi_i \mathcal{N}(\mu_{e_i}, \Sigma_{e_i});$$

$$\text{where, } \phi_i = \frac{exp\left(\mathcal{N}(\mu_{e_i}, \Sigma_{e_i})\right)}{\sum_{j=1}^{n} exp\left(\mathcal{N}(\mu_{e_j}, \Sigma_{e_j})\right)}$$

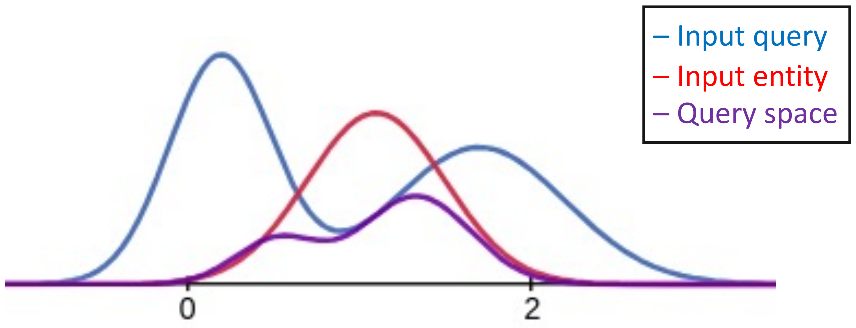
Gaussian Union

Our Solution

Challenge: Sequential chain operations.



Chain Translation:
Translation of a Gaussian mixture with another Gaussian distribution



Chain Intersection:
Intersection of a Gaussian mixture with another Gaussian distribution

Chain Union:
Union of a Gaussian mixture with another Gaussian distribution

Our Solution

Challenge: Sequential chain operations.

$$c_t = \sum_{i=1}^{n} \phi_i \mathcal{N}(\mu_i + \mu_r, (\Sigma_i^{-1} + \Sigma_r^{-1})^{-1})$$

Chain Translation:
Translation of a Gaussian mixture with another Gaussian distribution

$$\begin{aligned} c_\cap &= \cup_{i=1}^{n} \mathcal{N}(\mu_e, \Sigma_e)\mathcal{N}(\mu_i, \Sigma_i) \\ &= \sum_{i=1}^{n} \phi_i \mathcal{N}(\mu_{e\cap i}, \Sigma_{e\cap i}) \end{aligned}$$

Chain Intersection:
Intersection of a Gaussian mixture with another Gaussian distribution

$$c_\cup = \sum_{i=1}^{n} \phi_i \mathcal{N}(\mu_i, \Sigma_i) + \phi_e \mathcal{N}(\mu_e, \Sigma_e)$$

Chain Union:
Union of a Gaussian mixture with another Gaussian distribution

Our Solution

Challenge: Scalability

Points of note:

1. Single Gaussian/Box query operations require little memory individually.
2. Gaussian operation are closed under Gaussian Mixture models. Boxes are not.
3. Box operations require DNF transformation for union (needs the entire query). Hence, long queries not reducible to individual operations.

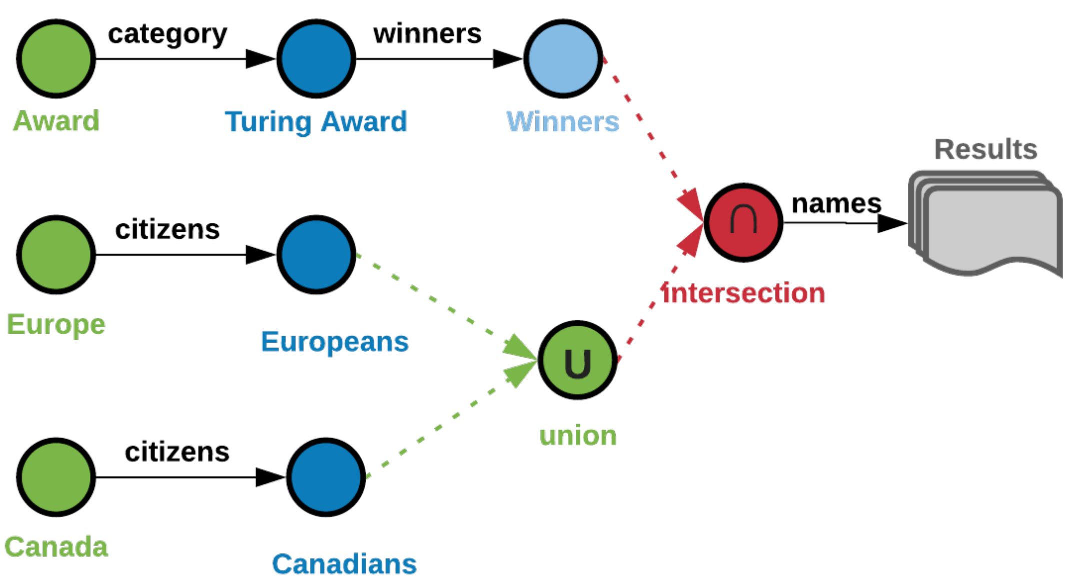
Gaussian queries are reducible to single query operations and can be merged together with sequential chain operations. Thus, they can be scaled to operate on any length of queries.

Our Solution

Processing a query in PERM architecture

Q: Who (X) are the Canadian (C) and European (E) Turing (T) Award (A) winners (W)?

$$?X : \exists X.names(X, \exists W.[winners(W, \exists T.category(T, A)) \\ \cap [citizen(W, Europe) \cup citizen(W, Canada)]])$$



Update Entity Representations with back-propagation

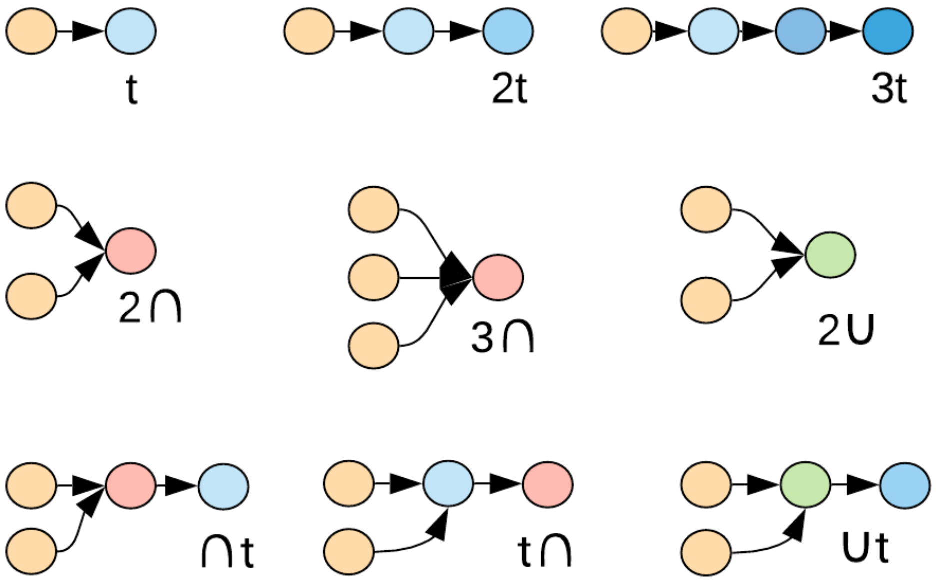
1. The entire query can be reduced to single operations due to PERM's closure under Gaussian mixture model.
2. The reducibility allows us to pre-process small queries, which can be aggregated during full inference.
3. This is not possible in the previous approaches, because the DNF transformation would process the entire query together.

Evaluation

1. Reasoning over KGs
2. Ablation Study
3. Drug Recommendation

Evaluation

Reasoning over KGs



Translation (t):

1t: "nike", "shoes", "adidas"

2t: "women shoes" ("shoes" → "women")

3t: "furniture" ("furniture"→"chair","table","dining" →"ikea","wayfair")

Intersection (∩):

2∩: "nike shoes" ("nike" AND "shoes")

3∩: "nike jordan laces" ("nike" AND "jordan" AND "laces")

∩t: "nike shoes" ("nike" AND "shoes" →products in the space)

t∩: "furniture ikea" ("furniture"→"chair","table","dining", etc AND "ikea")

Union (∪):

2∪: "nike and adidas" ("nike" OR "adidas")

∪t: "nike and adidas shoes" ("nike" OR "adidas"→products in the space)

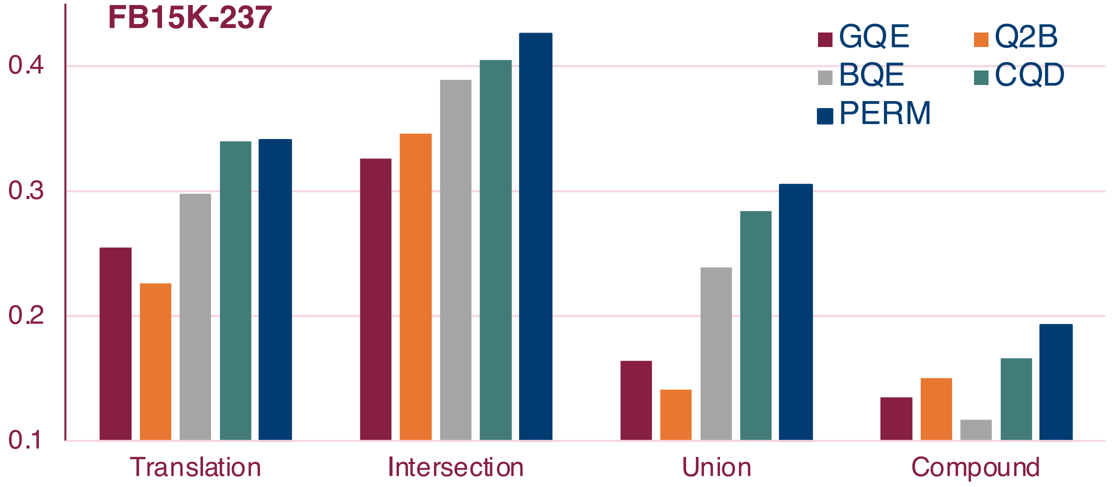
Evaluation

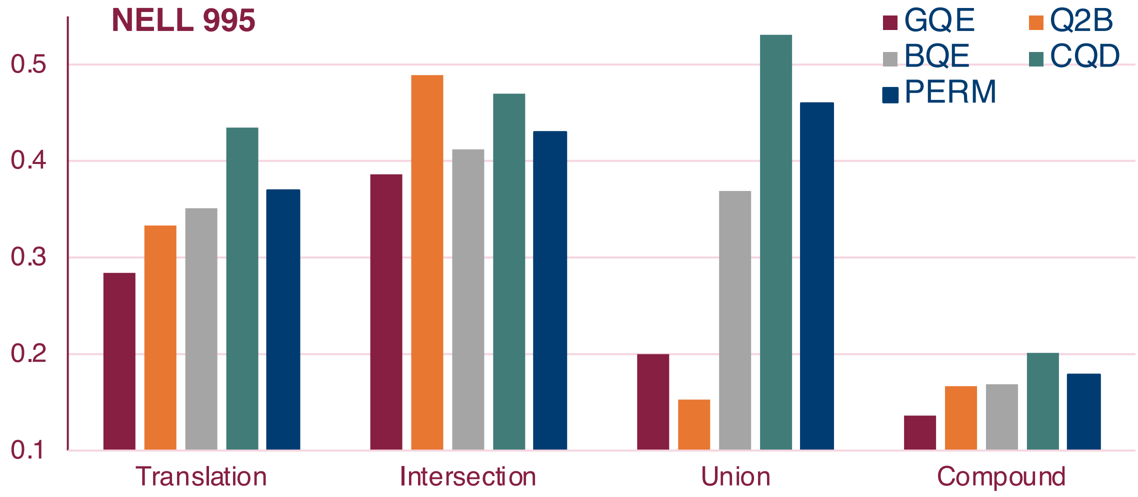
Reasoning over KGs

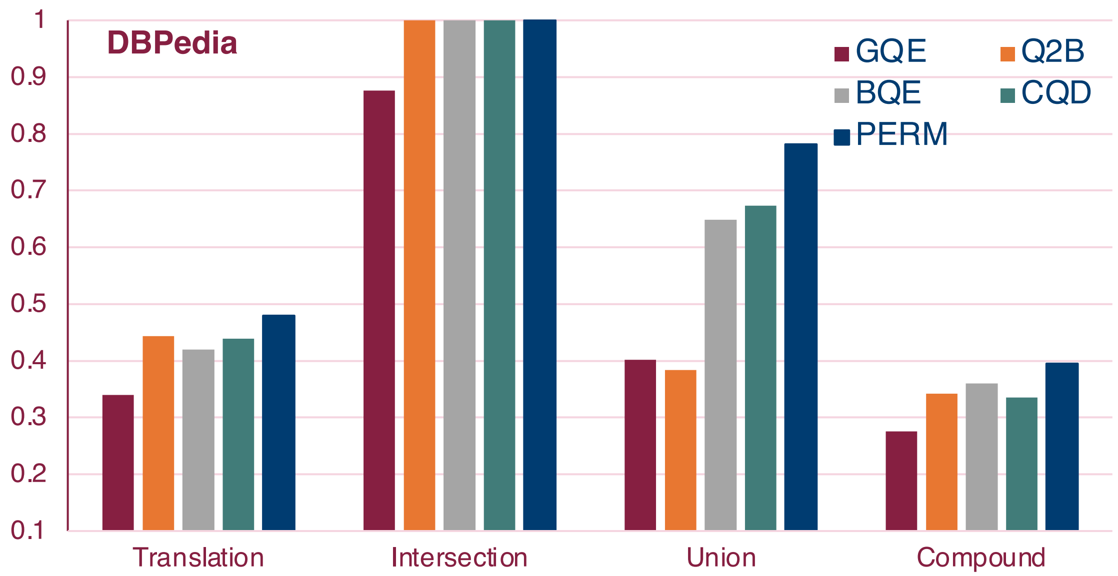
- Logical Query Reasoning
  - Dataset: FB15K-237, NELL995, DBPedia, DRKG
  - Baselines: GQE (Vectors), Q2B (Query2Box), BQE(Beta), CQD
  - Evaluation Metrics: HITS@3, Mean Reciprocal Rank

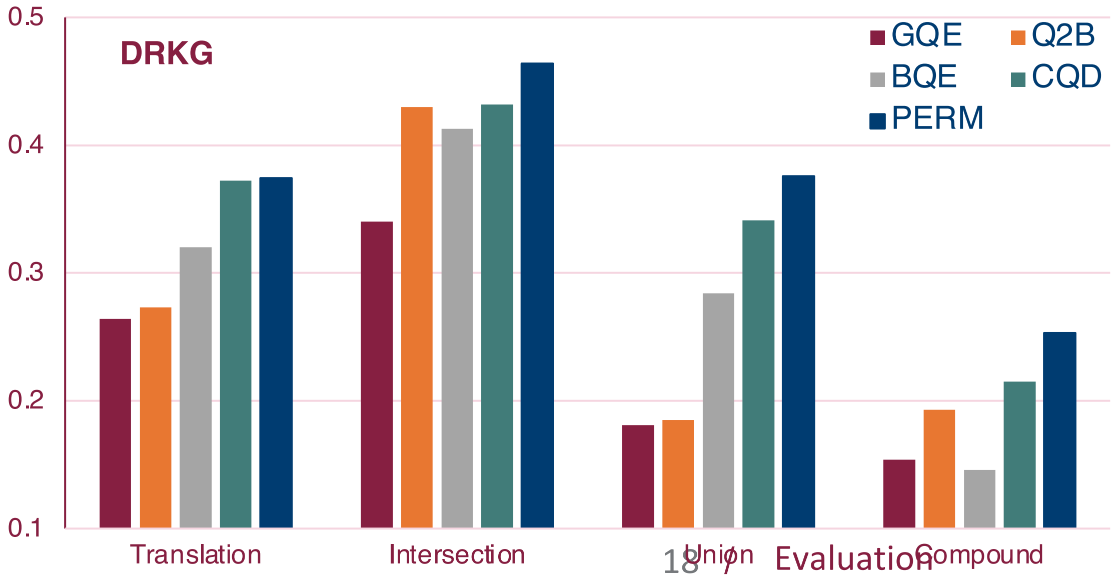
Evaluation

Reasoning over KGs (HITS@3)









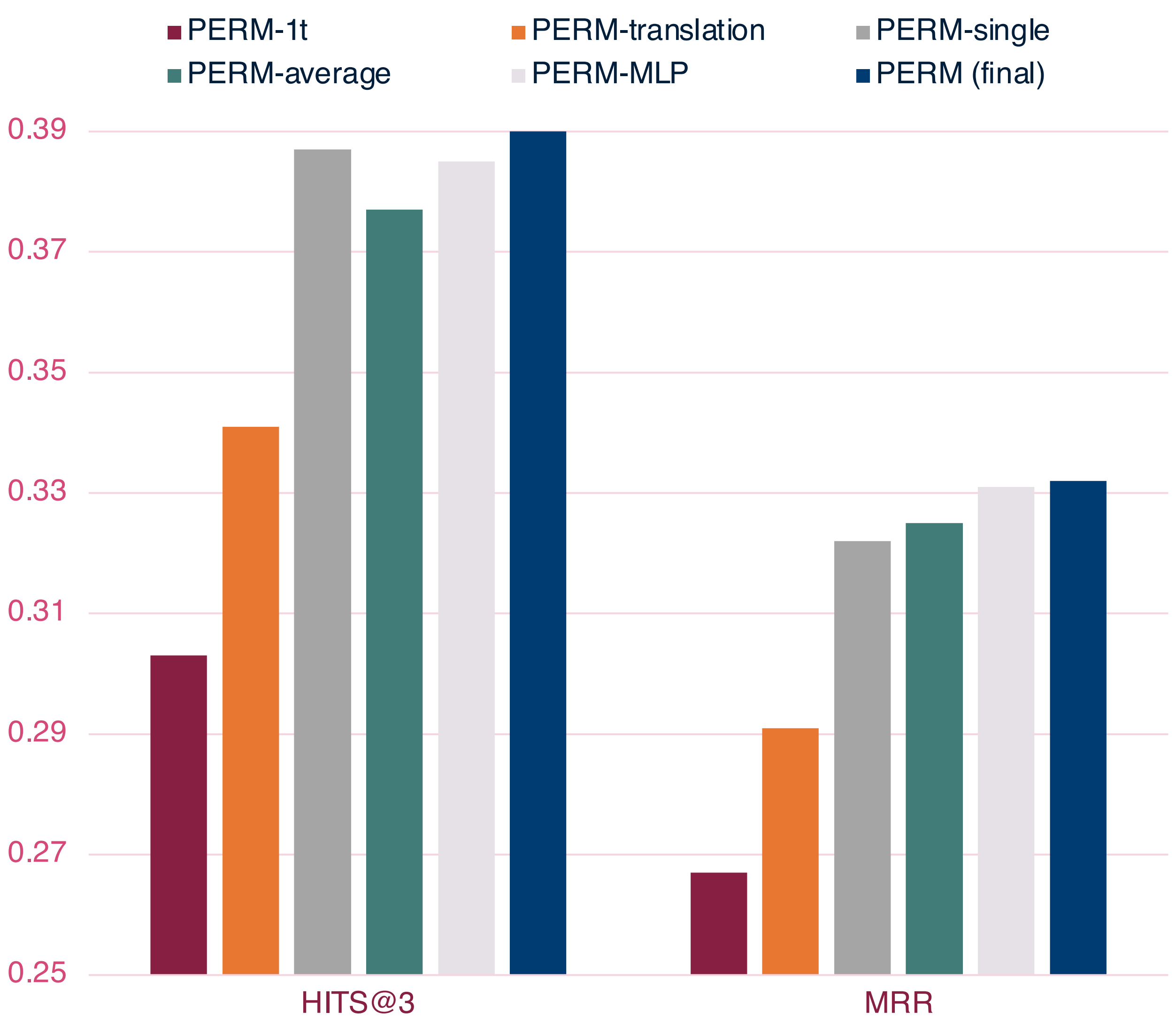
Evaluation

Reasoning over KGs (Qualitative Results)

Dataset	Query	Results
FB15K-237	Who are European and Canadian Turing awards winners?	Jeffrey Hinton, Yoshua Bengio, Andrew Yao
DBPedia	Which Actors and Football Players also became Governors?	Arnold Schwarzenegger, Heath Shuler, Frank White
DRKG	Which treatment drugs interact with all proteins associated with SARS diseases?	Ribavirin, Dexamethasone, Hydroxy-chloroquine

Evaluation

Ablation Study



- 1t: Only 1t queries
- translation: Using only translation queries
- single: Using single operator queries
- average: Average aggregation.
- MLP: MLP aggregation.

Evaluation

Case Study: Drug Recommendation

Model	P@10	R@10	F1	Top Recommended Drugs
GQE	0.119	0.174	0.141	Piclidenoson, Ibuprofen, Chloroquine
BQE	0.159	0.200	0.177	Ribavirin, Oseltamivir, Ruxolitinib
Q2B	0.194	0.255	0.221	Ribavirin, Dexamethasone, Deferoxamine
CQD	0.209	0.260	0.232	Ribavirin, Dexamethasone, Tofacitinib
PERM	**0.217**	**0.269**	**0.251**	Ribavirin, Dexamethasone, Hydroxychloroquine
PERM vs Q2B	11.9%	5.5%	13.6%	
PERM vs CQD	3.8%	3.5%	8.2%	

Conclusion and Broader Impact

1. We introduce Gaussian Embeddings with closed form solutions for scalable KG reasoning and smooth query borders.
2. The reducibility of Gaussian operations allow us to pre-process and store small operations, which can be aggregated for inference.
3. The basic idea behind the solution can be extended that needs to encode its basic units as probabilistic embeddings, e.g., model documents as union of topics.

Limitation: PERM depends on the integrity of the KG to learn representations. It cannot handle noisy training graphs, and hence its arbitrary application would be fatal in sensitive areas of research such as drug recommendations. It is necessary to maintain the integrity of training data before learning representations and querying with PERM.

Thanks!

Any questions?

Link to Implementation:

https://github.com/Akirato/PERM-GaussianKG

Find me at:

https://nurendra.me

nurendra@vt.edu