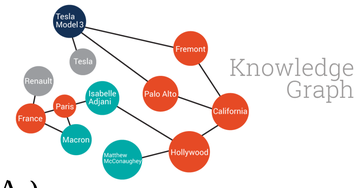
Expanding Holographic Embeddings for Knowledge Completion

Yexiang Xue[1], Yang Yuan[2], Zhitian Xu[1], Ashish Sabharwal[3]

[1]Purdue University; [2]Cornell University; [3]Allen Institute for Artificial Intelligence

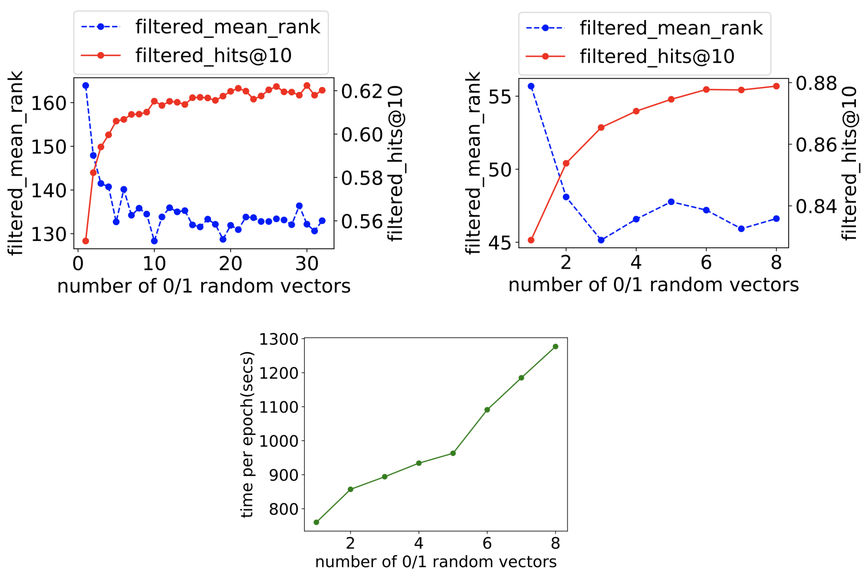
Knowledge Graph Completion



- **Knowledge graph completion**
  (Bill Gates, Lives in, ??)
  (??, the capital of, the U.S.A.)
- **Tail prediction:** given subject *s*, relation *r*, predict object *o*.
- **Head prediction:** given object *o*, relation *r*, predict subject *s*.

- **Compositional embedding:**
  embed subject s, relation r, object o into high dimensional space: $s \in R^{d_s}$, o $\in R^{d_o}$, r $\in R^{d_r}$
  learn composition function $\phi_r(s, o)$ given dataset Ω
  $\Pr(\phi_r(s, o) = 1 \mid \Omega)$

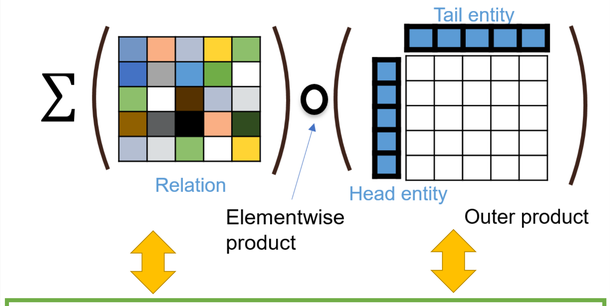
Impact on Varying the Number of Perturbation Vectors



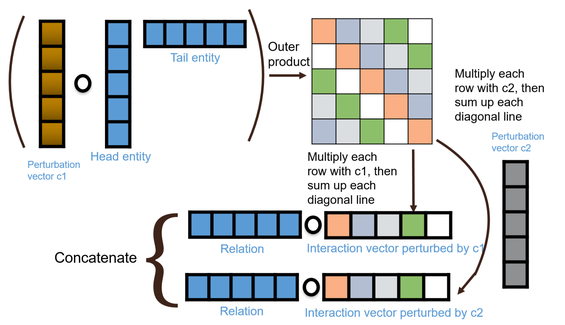
- **(Top left) Full interpolation for embedding dimension 32; (Top right) Similar trend for embedding dimension 256; (Bot) Training time per epoch for embedding dimension 256**

Interpolation Between RESCAL and Holographic Embedding (HolE)

RESCAL **#Params:** $O(d^2|R|)$

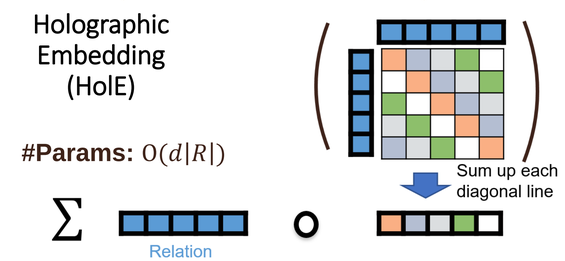


Our Contribution:
Expanded Holographic Embedding (HolEX)
- **Full interpolation between RESCAL and HolE**
- **Combines the strength of the two models.**



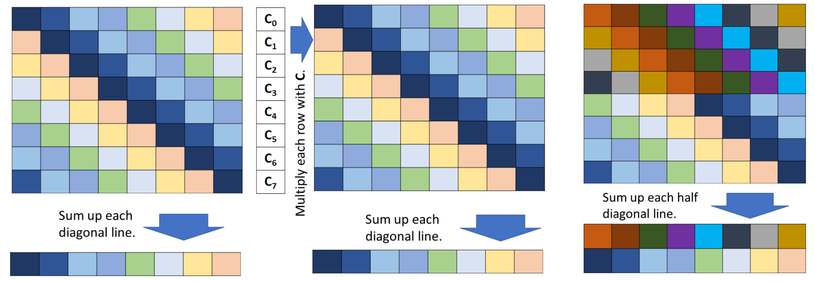
Holographic Embedding (HolE)

#Params: $O(d|R|)$



Perturbation with Haar and Random 0/1 Vectors

$$H_2 = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \qquad H_4 = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & 0 & 0 \\ 0 & 0 & 1 & -1 \end{pmatrix} \qquad H_{2n} = \begin{pmatrix} H_n & \otimes_k & [1,1] \\ I_n & \otimes_k & [1,-1] \end{pmatrix}$$



Experimental Results on FB15K

Knowledge Completion Method	Mean Rank	HITS@10 (%)	MRR	HITS@5 (%)	HITS@1 (%)
EXISTING METHODS					
RESCAL [18]	683	44.1	-	-	-
TransE [5]	125	47.1	-	-	-
TransR [16]	77	68.7	-	-	-
TransE + Rev [15]	63	70.2	-	-	-
HOLE (original, dim=150) [19]	-	73.9	0.524	-	40.2
HOLE (reimplementation, dim=150)	70	78.4	0.588	72.0	47.7
ProjE_pointwise* (dim=256) [22]	71	80.2	0.650	74.8	56.7
ProjE_wlistwise* (dim=256) [22]	64	82.1	0.666	76.8	57.9
ProjE_listwise* (dim=256) [22]	53	82.9	0.665	78.1	56.8
HolE (reimplementation, dim=256)	51	83.0	0.665	77.9	56.9
ComplEx [26]	-	84.0	0.692	-	59.9
PTransE (ADD, len-2 path) [15]	54	83.4	-	-	-
PTransE (ADD, len-3 path) [15]	58	84.6	-	-	-
DistMult [28], re-tuned by Kadlec et al. [13]	**42**	**89.3**	**0.798**	-	-
PROPOSED METHOD (dim=256)					
HolE (reimplemented baseline from above)	51	83.0	0.665	77.9	56.9
HOLEX, 8 Haar vectors	51	86.7	-	-	-
HOLEX, 2 random 0/1 vectors	48	85.4	0.720	81.4	64.0
HOLEX, 4 random 0/1 vectors	47	87.1	0.763	83.9	69.8
HOLEX, 8 random 0/1 vectors	**47**	87.9	0.786	85.0	73.1
HOLEX, 16 random 0/1 vectors	49	**88.6**	**0.800**	**86.0**	**75.0**