BiQUE: Biquaternionic Embeddings of Knowledge Graphs

Jia Guo and Stanley Kok

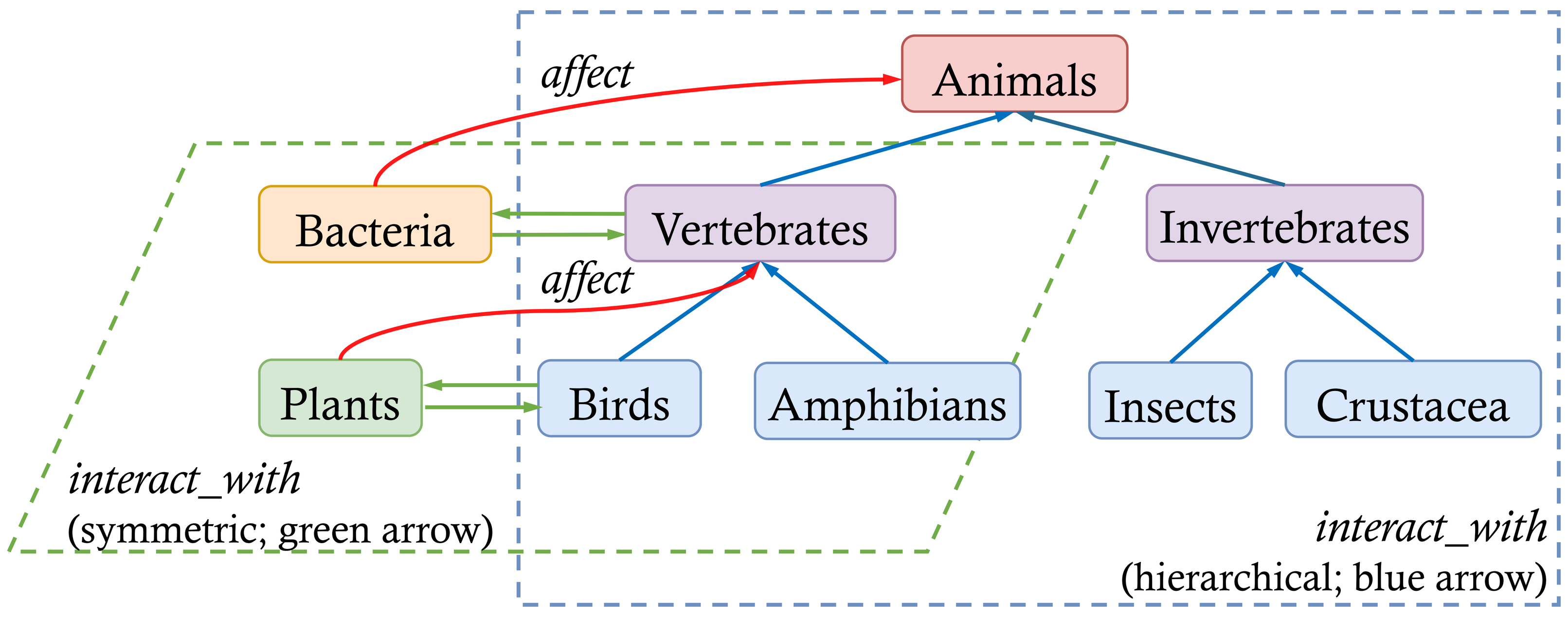
Department of Information Systems and Analytics, School of Computing
National University of Singapore

NUS National University of Singapore

Motivation

***Challenges*:**

- A Euclidean transformation (e.g., translation or rotation) is insufficient for knowledge graph embedding models.
  - Some relations (e.g., *interact_with*) are both hierarchical and symmetric → need to rely on the data to choose the sweet spot among various transformations.
  - Some relations (e.g., *affect*) require the composition of multiple transformations → need to learn the best composition of both representations jointly.



- We aim to build a coherent model to integrate multiple geometric operations and choose the best representation for each relation.

	Geometric Operations for KGE		
	Euclidean Translation	Euclidean Rotation	Hyperbolic Rotation
Can	asymmetry & inversion & composition	asymmetry & inversion & composition & symmetry	asymmetry & inversion & composition & symmetry & hierarchical structure
Can't	symmetry	hierarchical structure	incompatible to Euclidean properties/systems; hard to optimize

Background

- Complex numbers: $c = a + b\mathbf{i}$ $a, b \in \mathbb{R}$ (Real part, Imaginary part)
- Quaternions: $q = w + x\mathbf{i} + y\mathbf{j} + z\mathbf{k}$ $w, x, y, z \in \mathbb{R}$ (Real part, Imaginary parts)
- ***Biquaternions***: $q = w + x\mathbf{i} + y\mathbf{j} + z\mathbf{k}$ $w, x, y, z \in \mathbb{C}$ (Scalar part $s(q)$, Vector part $v(q)$)

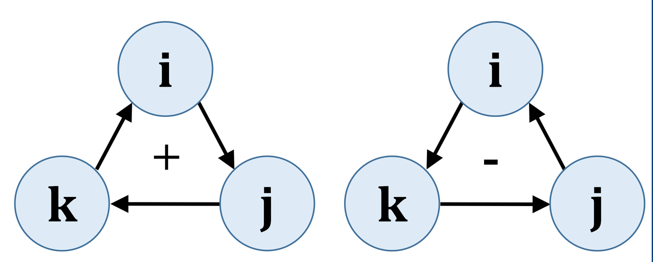
$$q = (w_r + w_i\mathbf{I}) + (x_r + x_i\mathbf{I})\mathbf{i} + (y_r + y_i\mathbf{I})\mathbf{j} + (z_r + z_i\mathbf{I})\mathbf{k}$$

$$q = q_r + q_i\mathbf{I} \quad q_r = w_r + x_r\mathbf{i} + y_r\mathbf{j} + z_r\mathbf{k},\ w_r, x_r, y_r, z_r \in \mathbb{R}$$
$$q_i = w_i + x_i\mathbf{i} + y_i\mathbf{j} + z_i\mathbf{k},\ w_i, x_i, y_i, z_i \in \mathbb{R}$$

Imaginary unit:

$$\mathbf{i}^2 = -1,\quad \mathbf{j}^2 = -1,\quad \mathbf{k}^2 = -1,\quad \mathbf{I}^2 = -1$$

$$\mathbf{iI} = \mathbf{Ii},\quad \mathbf{jI} = \mathbf{Ij},\quad \mathbf{kI} = \mathbf{Ik}$$



Basic operations on biquaternions:

$$q_1 = w_1 + x_1\mathbf{i} + y_1\mathbf{j} + z_1\mathbf{k}$$
$$q_2 = w_2 + x_2\mathbf{i} + y_2\mathbf{j} + z_2\mathbf{k}$$

- The addition/subtraction (element-wise):

$$q_1 \pm q_2 = (w_1 \pm w_2) + (x_1 \pm x_2)\mathbf{i} + (y_1 \pm y_2)\mathbf{j} + (z_1 \pm z_2)\mathbf{k}$$

- The multiplication (a.k.a. Hamilton product)
  - associative, not commutative

$$q_1 q_2 = (w_1w_2 - x_1x_2 - y_1y_2 - z_1z_2) + (w_1x_2 + x_1w_2 + y_1z_2 - z_1y_2)\mathbf{i} + (w_1y_2 - x_1z_2 + y_1w_2 + z_1x_2)\mathbf{j} + (w_1z_2 + x_1y_2 - y_1x_2 + z_1w_2)\mathbf{k}$$

$$\mathcal{M}(q_2)\,\mathcal{V}(q_1) = \begin{bmatrix} w_2 & -x_2 & -y_2 & -z_2 \\ x_2 & w_2 & z_2 & -y_2 \\ y_2 & -z_2 & w_2 & x_2 \\ z_2 & y_2 & -x_2 & w_2 \end{bmatrix} \cdot \begin{bmatrix} w_1 \\ x_1 \\ y_1 \\ z_1 \end{bmatrix}$$

Our BiQUE Model

BiQUE unifies Euclidean and hyperbolic rotations:

Theorem: The matrix $M(q)$ of a unit biquaternion q can be decomposed into the composition of two rotation matrices:
- $M(q) = M(h)\,M(u)$: Euclidean rotation, then hyperbolic rotation
- $M(q) = M(u)\,M(h')$: hyperbolic rotation, then Euclidean rotation

$$q = uh = \mathcal{M}(h)\mathcal{M}(u) \qquad u = \frac{q_r}{||q_r||} \qquad h = ||q_r|| + \mathbf{I}\frac{\overline{q_r}q_i}{||q_r||||q_i||}||q_i||$$

$$u = \cos\theta + \frac{x_r\sin\theta}{||v(q_r)||}\mathbf{i} + \frac{y_r\sin\theta}{||v(q_r)||}\mathbf{j} + \frac{z_r\sin\theta}{||v(q_r)||}\mathbf{k}$$
$$\theta = \cos^{-1}\frac{w_r}{||q_r||}$$

$$h = \cosh\phi + (a\mathbf{I}\sinh\phi)\mathbf{i} + (b\mathbf{I}\sinh\phi)\mathbf{j} + (c\mathbf{I}\sinh\phi)\mathbf{k}$$
$$||q_r|| = \cosh\phi,\quad ||q_i|| = \sinh\phi$$

$$M(u) = \begin{bmatrix} \cos\theta & -\tilde{x}_r\sin\theta & -\tilde{y}_r\sin\theta & -\tilde{z}_r\sin\theta \\ \tilde{x}_r\sin\theta & \cos\theta & \tilde{z}_r\sin\theta & -\tilde{y}_r\sin\theta \\ \tilde{y}_r\sin\theta & -\tilde{z}_r\sin\theta & \cos\theta & \tilde{x}_r\sin\theta \\ \tilde{z}_r\sin\theta & \tilde{y}_r\sin\theta & -\tilde{x}_r\sin\theta & \cos\theta \end{bmatrix}$$

$$M(h) = \begin{bmatrix} \cosh\phi & -a\mathbf{I}\sinh\phi & -b\mathbf{I}\sinh\phi & -c\mathbf{I}\sinh\phi \\ a\mathbf{I}\sinh\phi & \cosh\phi & c\mathbf{I}\sinh\phi & -b\mathbf{I}\sinh\phi \\ b\mathbf{I}\sinh\phi & -c\mathbf{I}\sinh\phi & \cosh\phi & a\mathbf{I}\sinh\phi \\ c\mathbf{I}\sinh\phi & b\mathbf{I}\sinh\phi & -a\mathbf{I}\sinh\phi & \cosh\phi \end{bmatrix}$$

Biquaternionic representations for knowledge graphs:

$$Q \in \{Q_h, Q_t, Q_r^+, Q_r^\times\},\ Q = w + x\mathbf{i} + y\mathbf{j} + z\mathbf{k}, w, x, y, z \in \mathbb{C}$$

(Head: Q_h; Tail: Q_t; Relation: $Q_r^+, Q_r^\times$)

Relation-specific transformations:

- Translation:

$$Q'_{h,r} = Q_h + Q_r^+ = (w_h + w_r^+) + (x_h + x_r^+)\mathbf{i} + (y_h + y_r^+)\mathbf{j} + (z_h + z_r^+)\mathbf{k} = w' + x'\mathbf{i} + y'\mathbf{j} + z'\mathbf{k}$$

- Unified rotation and scaling:

$$\widehat{Q_{h,r}} = Q'_{h,r} \circledast Q_r^\times = (w' \otimes w_r^\times - x' \otimes x_r^\times - y' \otimes y_r^\times - z' \otimes z_r^\times) + (w' \otimes x_r^\times + x' \otimes w_r^\times + y' \otimes z_r^\times - z' \otimes y_r^\times)\mathbf{i} + (w' \otimes y_r^\times - x' \otimes z_r^\times + y' \otimes w_r^\times + z' \otimes x_r^\times)\mathbf{j} + (w' \otimes z_r^\times + x' \otimes y_r^\times - y' \otimes x_r^\times + z' \otimes w_r^\times)\mathbf{k}$$

- Score function:

$$f(h, r, t) = \widehat{Q_{h,r}} \cdot Q_t = <\hat{w}, w_t> + <\hat{x}, x_t> + <\hat{y}, y_t> + <\hat{z}, z_t>$$

Experiments

- Datasets: WN18RR, FB15K-237, YAGO3-10, Concept100k and ATOMIC.
- Task: knowledge completion task (KGC), given (h, r, ?), predict t.
- Evaluation: mean reciprocal rank, Hits@k.

Results:

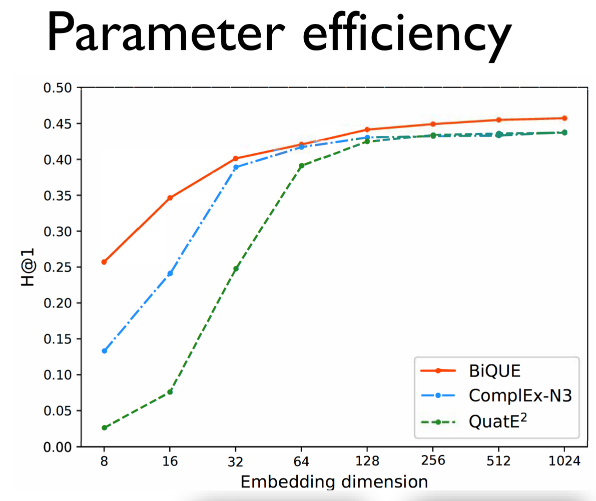
	CN-100K				ATOMIC			
Models	MRR	H@1	H@3	H@10	MRR	H@1	H@3	H@10
DistMult	0.090	0.045	0.098	0.174	0.124	0.092	0.152	0.183
ComplEx	0.114	0.074	0.125	0.190	0.142	0.133	0.141	0.160
ConvE	0.209	0.140	0.229	0.340	0.101	0.082	0.103	0.134
RotatE	0.247	-	0.282	0.454	0.112	-	0.115	0.156
ConvTransE	0.187	0.079	0.239	0.390	0.129	0.129	0.130	0.130
QuatE[2]	0.313	**0.217**	0.356	0.504	0.187	0.167	0.191	0.225
BiQUE	**0.320**	0.216	**0.359**	**0.553**	**0.191**	**0.171**	**0.196**	**0.230**

	WN18RR				FB15K-237				YAGO3-10			
Models	MRR	H@1	H@3	H@10	MRR	H@1	H@3	H@10	MRR	H@1	H@3	H@10
ComplEx-N3	0.480	0.435	0.495	0.572	0.357	0.264	0.392	0.547	0.569	0.498	0.609	0.701
RotatE	0.476	0.428	0.492	0.571	0.338	0.241	0.375	0.533	0.495	0.402	0.550	0.670
QuatE[2]	0.482	0.436	0.499	0.572	**0.366**	**0.271**	**0.401**	**0.556**	0.568	0.493	0.611	0.706
DualE[1]	0.482	0.440	0.500	0.561	0.330	0.237	0.363	0.518	-	-	-	-
MurP	0.481	0.440	0.495	0.566	0.335	0.243	0.367	0.518	0.354	0.249	0.400	0.567
ATTH	0.486	0.443	0.499	0.573	0.348	0.252	0.384	0.540	0.568	0.493	0.612	0.702
BiQUE	**0.504**	**0.459**	**0.519**	**0.588**	0.365	0.270	**0.401**	0.555	**0.581**	**0.509**	**0.624**	**0.713**

- Comparison of H@10 per relation

Relation Name	#Tripes	RotH	QuatE[2]	BiQUE	Lift
hypernym	1,251	0.276	0.283	**0.306**	8.13%
derivationally_related_form	1,074	0.968	0.969	**0.970**	0.10%
instance_hypernym	122	0.520	0.549	**0.602**	9.65%
also_see	56	0.705	0.688	**0.750**	6.38%
member_meronym	253	0.399	0.389	**0.453**	13.53%
synset_domain_topic_of	114	0.447	0.518	**0.539**	4.05%
has_part	172	0.346	0.337	**0.392**	13.29%
member_of_domain_usage	24	0.438	**0.563**	**0.563**	-
member_of_domain_region	26	0.365	0.327	**0.500**	36.99%
verb_group	39	**0.974**	**0.974**	**0.974**	-
similar_to	3	**1.000**	**1.000**	**1.000**	-

- Parameter efficiency



Scan Here

Paper Codes