HousE: Knowledge Graph Embedding with Householder Parameterization

Rui Li[1], Jianan Zhao[2], Chaozhuo Li[3], Di He[4], Yiqi Wang[5], Yuming Liu[3], Hao Sun[3], Senzhang Wang[6], Weiwei Deng[3], Yanming Shen[1], Xing Xie[3], Qi Zhang[3]

1 大连理工大学 DALIAN UNIVERSITY OF TECHNOLOGY
2 UNIVERSITY OF NOTRE DAME
3 Microsoft
4 北京大学 PEKING UNIVERSITY
5 MICHIGAN STATE UNIVERSITY
6 中南大学 CENTRAL SOUTH UNIVERSITY

- **Knowledge Graph (KG)**:
  - Nodes represent **Entities**.
  - Edges represent **Relations**.
  - A collection of **factual triplets** → $(head\ entity,\ relation,\ tail\ entity)$.

$$\boldsymbol{h} \xrightarrow{\boldsymbol{r}} \boldsymbol{t}$$

- KGs suffer from the **Incompleteness**.

- **Knowledge Graph Embedding** excels as an effective tool for predicting missing links.
  - Learns low-dimensional representations of entities and relations.

Relations Are Important

- The effectiveness of KGE largely depends on the ability to model and infer:

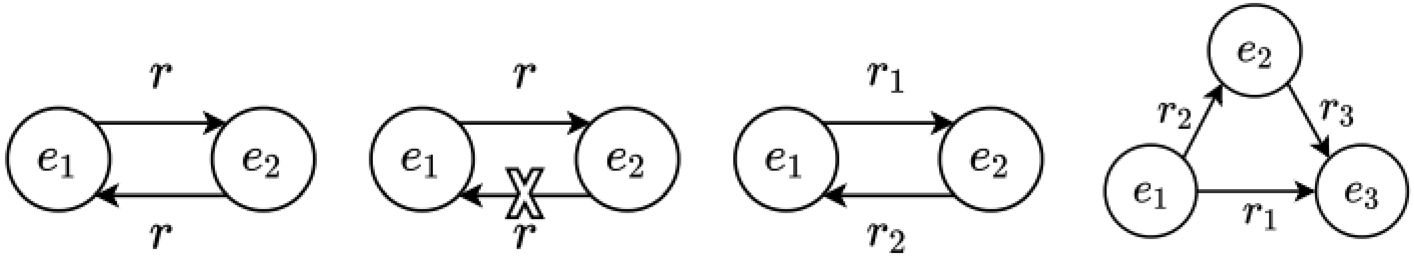
  - **Relation Patterns**:

    Symmetry

    Antisymmetry

    Inversion

    Composition



(a) Symmetry (b) Antisymmetry (c) Inversion (d) Composition

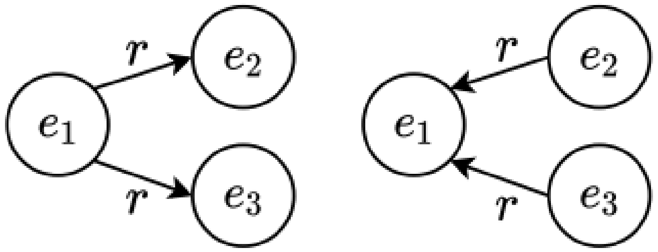
  - **Relation Mapping Properties** (RMPs):

    1-to-1

    1-to-N

    N-to-1

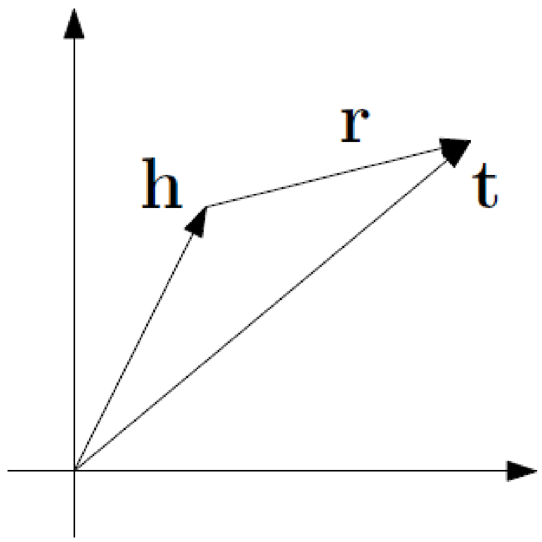
    N-to-N



(e) 1-to-N (f) N-to-1

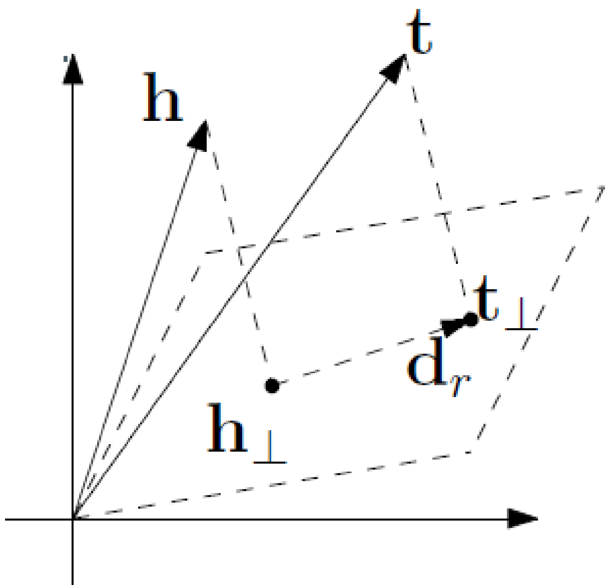
Translation-based Models

- TransE:
  - Regards each relation as a **translation** from a head entity to a tail entity.
  - Expects $\mathbf{h} + \mathbf{r} \approx \mathbf{t}$ if triple (h, r, t) holds.
  - Fails to model symmetric relations and RMPs.



[Bordes et al., 2013] Translating embeddings for modeling multi-relational data. NIPS 2013.

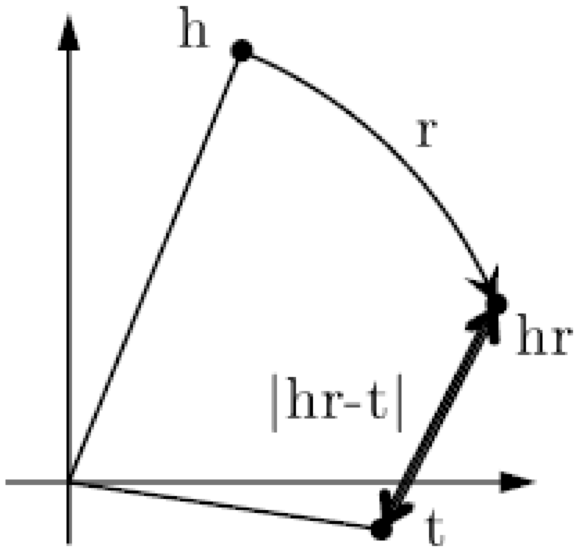
- TransH:
  - Projects entities to a relation-specific hyperplane. (**irreversible projection**)
  - Performs translation on this hyperplane.
  - **Loses the capability of modeling inversion and composition patterns.**



[Wang et al., 2014] Knowledge graph embedding by translating on hyperplanes. AAAI 2014.

RotatE:

- Represents each relation as a **2-dimensional rotation** by using complex multiplication.
- RotatE is capable of modeling all the four relation patterns, but **Fails to model RMPs**.



[Sun et al., 2019] RotatE: Knowledge graph embedding by relational rotation in complex space. ICLR 2019.

Rotation-based Models

- Rotations based on **quaternions**:
  - Rotate3D[1] $\Longrightarrow$ **3**-dimensional rotations
  - QuatE[2] $\Longrightarrow$ **4**-dimensional rotations

[1][Gao et al., 2020] Rotate3D: Representing Relations as Rotations in Three-Dimensional Space for Knowledge Graph Embedding. CIKM 2020.

[2][Zhang et al., 2019] Quaternion knowledge graph embeddings. NIPS 2019.

Modeling Capability

Model	Symmetry	Antisymmetry	Inversion	Composition	Mapping Properties	Dimension of Rotation
TransE	---	✓	✓	✓	---	---
TransX	✓	✓	---	---	✓	---
DistMult	✓	---	---	---	✓	---
ComplEx	✓	✓	✓	---	✓	---
RotatE	✓	✓	✓	✓	---	2
Rotate3D	✓	✓	✓	✓	---	3
QuatE	✓	✓	✓	---	✓	4

Limitations

1. Existing rotations are restricted to **fixed and low-dimensional spaces**, which greatly limits the modeling capacity;
2. None of existing models is capable of modeling all the four relation patterns and **relation mapping properties** simultaneously.

Model	Symmetry	Antisymmetry	Inversion	Composition	Mapping Properties	Dimension of Rotation
TransE	---	✓	✓	✓	---	---
TransX	✓	✓	---	---	✓	---
DistMult	✓	---	---	---	✓	---
ComplEx	✓	✓	✓	---	✓	---
RotatE	✓	✓	✓	✓	---	2
Rotate3D	✓	✓	✓	✓	---	3
QuatE	✓	✓	✓	---	✓	4

Householder Framework

- To tackle these limitations, this work proposes **HousE**, which is a more powerful and general KGE framework based on the **Householder reflections**.

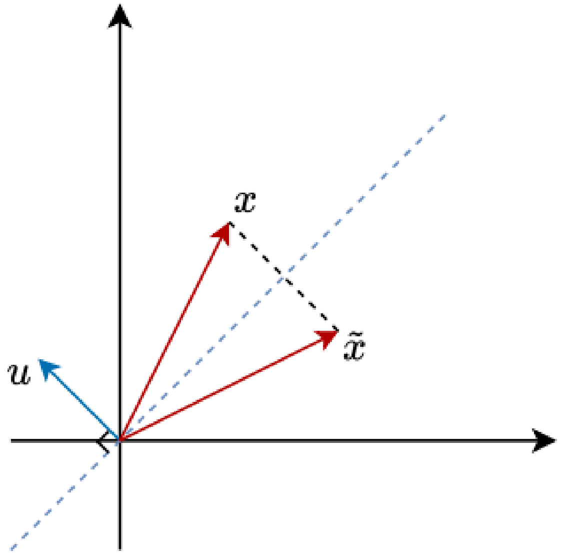
Model	Symmetry	Antisymmetry	Inversion	Composition	Mapping Properties	Dimension of Rotation
TransE	---	✓	✓	✓	---	---
TransX	✓	✓	---	---	✓	---
DistMult	✓	---	---	---	✓	---
ComplEx	✓	✓	✓	---	✓	---
RotatE	✓	✓	✓	✓	---	2
Rotate3D	✓	✓	✓	✓	---	3
QuatE	✓	✓	✓	---	✓	4
HousE	✓	✓	✓	✓	✓	***k***

- HousE
  - A framework based on the **Householder reflections**.
  - Designs two kinds of linear transformations:
    - **Householder rotation** composed of Householder reflections.
      $\Longrightarrow$ **High-dimensional Rotations**
    - **Householder projection** modified from Householder reflections.
      $\Longrightarrow$ **Modeling RMPs**

- **Householder reflection** describes **a reflection about a hyperplane**:

$$\tilde{x} = H(u)x = x - 2 < x, u > u$$

$$H(u) = I - 2uu^T$$



  - u is a unit vector that is orthogonal to the hyperplane.

- The composition of $2\lfloor\frac{k}{2}\rfloor$ Householder reflections can represent any ***k*-dimensional rotations**.

$$\Downarrow$$

Householder Rotation

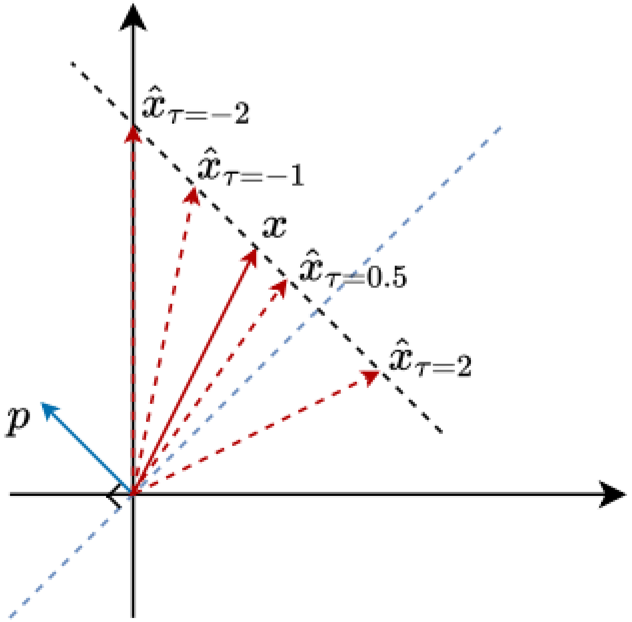
HousE-r: Relational Householder Rotations

- Defines each relation as a k-dimensional **Householder rotation** from the head entity h to the tail entity t.
- Theoretically, HousE-r is capable of modeling all the four relation patterns.
- However, HousE-r still **Fails to model RMPs**. ($\Longrightarrow$ **Invertible Projection** is needed)

- We modify the original Householder reflection:

$$\hat{x} = M(p,\tau) = x - \tau \langle x, p \rangle p,$$

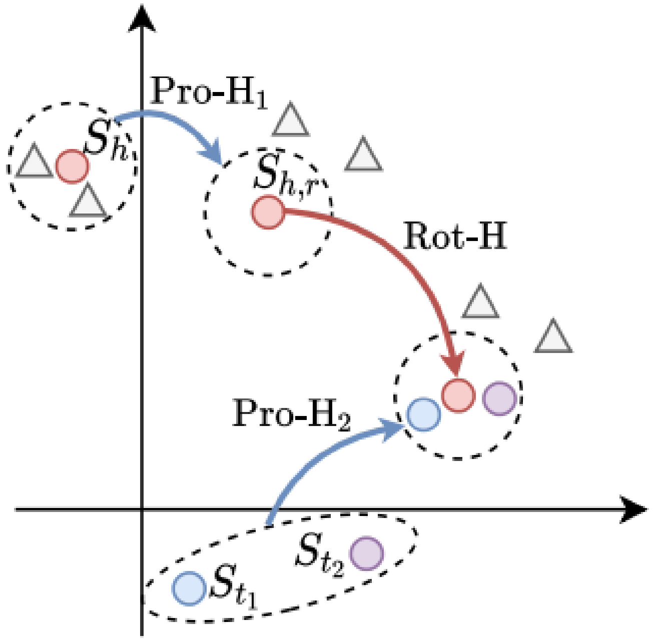
$$M(p,\tau) = I - \tau p p^\top.$$



- p is a unit vector that is orthogonal to the hyperplane.
- τ is a real scalar and $\tau \neq 1$.
- $M(p,\tau)$ is **invertible**.

- **Householder Projection**: the projection composed of m modified Householder reflections.

- For each triple (h, r, t):
  - First, HousE uses **relational Householder projections** to generate r-specific representations h_r and t_r for h and t respectively.
  - Then, HousE applies **relational Householder rotations** on h_r and expects the transformed result to be close to t_r.
  - Theoretically, HousE is capable of modeling all the four relation patterns and RMPs simultaneously.



Datasets:

WN18, FB15k, WN18RR, FB15k-237, YAGO3-10

Dataset	#entity	#relation	#training	#validation	#test
WN18	40,943	18	141,442	5,000	5,000
FB15k	14,951	1,345	483,142	50,000	59,071
WN18RR	40,943	11	86,835	3,034	3,134
FB15k-237	14,541	237	272,115	17,535	20,466
YAGO3-10	123,182	37	1,079,040	5,000	5,000

4 Experiment | Main Results

	WN18					FB15k				
Model	MR	MRR	H@1	H@3	H@10	MR	MRR	H@1	H@3	H@10
TransE†	-	.495	.113	.888	.943	-	.463	.297	.578	.749
DistMult◇	655	.797	-	-	.946	42.2	.798	-	-	.893
ComplEx	-	.941	.936	.945	.947	-	.692	.599	.759	.84
ConvE	374	.943	.935	.946	.956	51	.657	.558	.723	.831
RotatE	309	.949	.944	.952	.959	40	.797	.746	.830	.884
Rotate3D	214	.951	.945	.953	.961	39	.789	.728	.832	.887
QuatE	388	.949	.941	.954	.960	41	.770	.700	.821	.878
DualE	-	.951	.945	.956	.961	-	.790	.734	.829	.881
HousE-r	155	.953	.947	.956	.964	39	.807	.758	.839	.893
HousE	**137**	**.954**	**.948**	**.957**	**.964**	**38**	**.811**	**.759**	**.847**	**.898**

	WN18RR					FB15k-237					YAGO3-10				
Model	MR	MRR	H@1	H@3	H@10	MR	MRR	H@1	H@3	H@10	MR	MRR	H@1	H@3	H@10
TransE†	3384	.226	-	-	.501	357	.294	-	-	.465	-	-	-	-	-
DistMult◇	5110	.43	.39	.44	.49	254	.241	.155	.263	.419	5926	.34	.24	.38	.54
ComplEx◇	5261	.44	.41	.46	.51	339	.247	.158	.275	.428	6351	.36	.26	.4	.55
ConvE◇	4187	.43	.40	.44	.52	224	.325	.237	.356	.501	1671	.44	.35	.49	.62
RotatE	3340	.476	.428	.492	.571	177	.338	.241	.375	.533	1767	.495	.402	.55	.67
Rotate3D	3328	.489	.442	.505	.579	165	.347	.250	.385	.543	-	-	-	-	-
QuatE	3472	.481	.436	.500	.564	176	.311	.221	.342	.495	-	-	-	-	-
DualE	-	.482	.440	.500	.561	-	.330	.237	.363	.518	-	-	-	-	-
Rot-Pro	2815	.457	.397	.482	.577	201	.344	.246	.383	.540	1797	.542	.443	.596	.669
HousE-r	1885	.496	.452	.511	.585	165	.348	.254	.384	.534	1449	.565	.487	.616	.703
HousE	**1303**	**.511**	**.465**	**.528**	**.602**	**153**	**.361**	**.266**	**.399**	**.551**	**1415**	**.571**	**.491**	**.620**	**.714**

Table 4. MRR for the models tested on each relation of WN18RR.

Relation Name	RotatE	QuatE	HousE-r	HousE
hypernym	0.154	0.172	0.182	**0.207**
instance_hypernym	0.324	0.362	0.395	**0.440**
member_meronym	0.255	0.236	0.275	**0.312**
synset_domain_topic_of	0.334	0.395	0.396	**0.428**
has_part	0.205	0.210	0.217	**0.232**
member_of_domain_usage	0.277	0.372	0.415	**0.453**
member_of_domain_region	0.243	0.140	0.281	**0.395**
derivationally_related_form	0.957	0.952	0.958	**0.958**
also_see	0.627	0.607	0.638	**0.640**
verb_group	0.968	0.930	0.968	**0.968**
similar_to	1.000	1.000	1.000	**1.000**

Table 5. MRR for the models tested on RMPs in FB15k-237.

Task	RMPs	RotatE	HousE
Predicting Head (MRR)	1-to-1	0.498	**0.514**
	1-to-N	0.475	**0.479**
	N-to-1	0.088	**0.114**
	N-to-N	0.260	**0.286**
Predicting Tail (MRR)	1-to-1	0.490	**0.502**
	1-to-N	0.071	**0.086**
	N-to-1	0.747	**0.778**
	N-to-N	0.367	**0.392**

5 Conclusions

- The effectiveness of KGE largely depends on the ability to model intrinsic relation patterns and mapping properties;

- Advantages: HousE
  - is capable of modeling all the four relation patterns and **relation mapping properties**.
  - can naturally model relations as high-dimensional rotations for **better modeling capacity**.
  - is a **generalization** of existing rotation-based models.

Thanks