



東南大學
SOUTHEAST UNIVERSITY

东南大学东大开题答辩通用模板

General PowerPoint Template Of Work report Or Scientific Report Or Work Plans

心於至善

金陵情报站

2023-09-06



PART 06

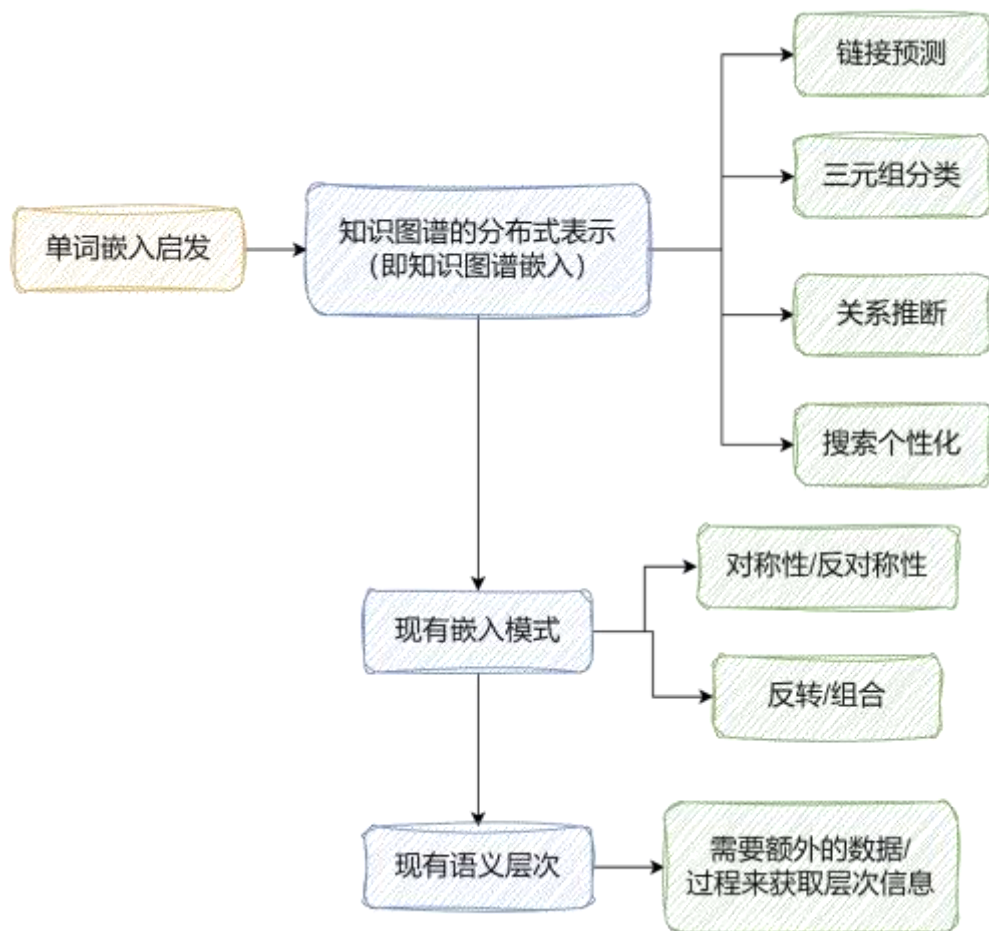
学习层次感知型知识图谱 嵌入以进行链接预测

止於至善

JUNE 12th

- **知识图谱**：知识图谱是一种结构化的数据库，它用**三元组**（头实体、关系、尾实体）的形式存储知识。比如，“苹果（头实体）是（关系）一种（尾实体）水果”就是一个三元组。
- **知识图谱嵌入**：这是一种技术，它把实体和关系转换成**低维的向量**（可以想象成在数学空间中的一个点），这样可以更容易地在计算机中处理和分析。
- **链接预测**：由于知识图谱往往是不完整的，链接预测的任务就是**预测实体之间缺失的关系**。比如，如果我们知道“苹果是一种水果”，我们可能想预测“香蕉也是一种水果”。

- **层次感知**：在现实世界中，**知识是有层次的**。比如，“水果”是一个高层次的概念，而“苹果”和“香蕉”是低层次的具体实例。这篇论文提出的方法能够理解这种层次结构。
- **HAKE模型**：论文提出了一个叫做HAKE（Hierarchy-Aware Knowledge Graph Embedding）的模型。这个模型**把实体映射到极坐标系**中，用径向坐标（从原点到点的距离）来表示实体在层次结构中的级别，用角度坐标来区分同一层次中的不同实体。



- 层次感知知识图谱嵌入（**HAK**E）
- 区分两类实体：（a）处于不同层次的实体；（b）处于同一层次的实体。
- 具有层次属性的实体可以被视为树→使用节点（实体）的深度来模拟层次结构的不同层次→使用模数信息来模拟类别（a）中的实体。
- 同一圆上的点可以具有不同相位→使用相位信息来模拟类别（b）中的实体。
- 结合模数和相位信息将实体映射到极坐标系统中，径向坐标对应模数信息，角度坐标对应相位信息。
- 效果：能清晰地区分实体的语义层次结构，能在基准数据集上显著且一致地优于几种最先进的方法。

➤ 知识图谱嵌入模型:

➤ 平移距离模型; 双线性模型; 基于神经网络的模型

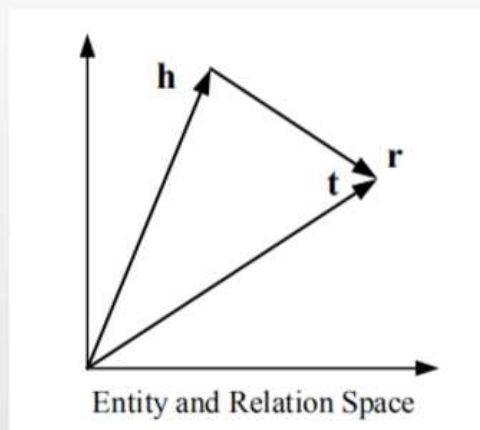
Model	Score Function $f_r(\mathbf{h}, \mathbf{t})$	Parameters
TransE (Bordes et al. 2013)	$-\ \mathbf{h} + \mathbf{r} - \mathbf{t}\ _{1/2}$	$\mathbf{h}, \mathbf{r}, \mathbf{t} \in \mathbb{R}^k$
TransR (Lin et al. 2015)	$-\ \mathbf{M}_r \mathbf{h} + \mathbf{r} - \mathbf{M}_r \mathbf{t}\ _2$	$\mathbf{h}, \mathbf{t} \in \mathbb{R}^d, r \in \mathbb{R}^k, \mathbf{M}_r \in \mathbb{R}^{k \times d}$
ManifoldE (Xiao, Huang, and Zhu 2016)	$-(\ \mathbf{h} + \mathbf{r} - \mathbf{t}\ _2^2 - \theta_r^2)^2$	$\mathbf{h}, \mathbf{r}, \mathbf{t} \in \mathbb{R}^k$
RotatE (Sun et al. 2019)	$-\ \mathbf{h} \circ \mathbf{r} - \mathbf{t}\ _2$	$\mathbf{h}, \mathbf{r}, \mathbf{t} \in \mathbb{C}^k, r_i = 1$
RESCAL (Nickel, Tresp, and Kriegel 2011)	$\mathbf{h}^\top \mathbf{M}_r \mathbf{t}$	$\mathbf{h}, \mathbf{t} \in \mathbb{R}^k, \mathbf{M}_r \in \mathbb{R}^{k \times k}$
DistMult (Yang et al. 2015)	$\mathbf{h}^\top \text{diag}(\mathbf{r}) \mathbf{t}$	$\mathbf{h}, \mathbf{r}, \mathbf{t} \in \mathbb{R}^k$
Complex (Trouillon et al. 2016)	$\text{Re}(\mathbf{h}^\top \text{diag}(\mathbf{r}) \bar{\mathbf{t}})$	$\mathbf{h}, \mathbf{r}, \mathbf{t} \in \mathbb{C}^k$
ConvE (Dettmers et al. 2018)	$f(\text{vec}(f([\bar{\mathbf{r}}, \bar{\mathbf{h}}] * \omega))) \mathbf{W} \mathbf{t}$	$\mathbf{h}, \mathbf{r}, \mathbf{t} \in \mathbb{R}^k$
HAKE	$-\ \mathbf{h}_m \circ \mathbf{r}_m - \mathbf{t}_m\ _2 - \lambda \ \sin((\mathbf{h}_p + \mathbf{r}_p - \mathbf{t}_p)/2)\ _1$	$\mathbf{h}_m, \mathbf{t}_m \in \mathbb{R}^k, \mathbf{r}_m \in \mathbb{R}_+^k,$ $\mathbf{h}_p, \mathbf{r}_p, \mathbf{t}_p \in [0, 2\pi)^k, \lambda \in \mathbb{R}$

CSDN @OYLZ

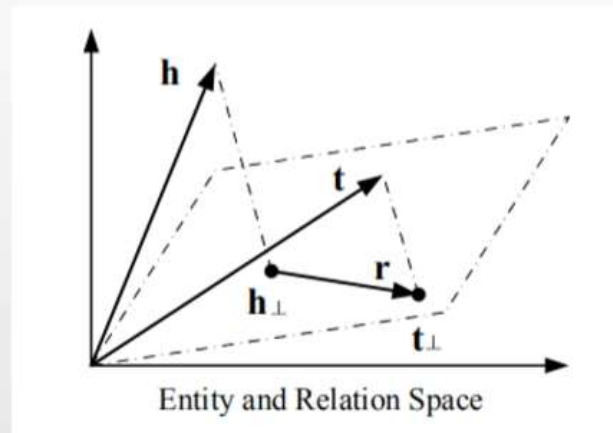
相关工作

➤ 知识图谱嵌入模型:

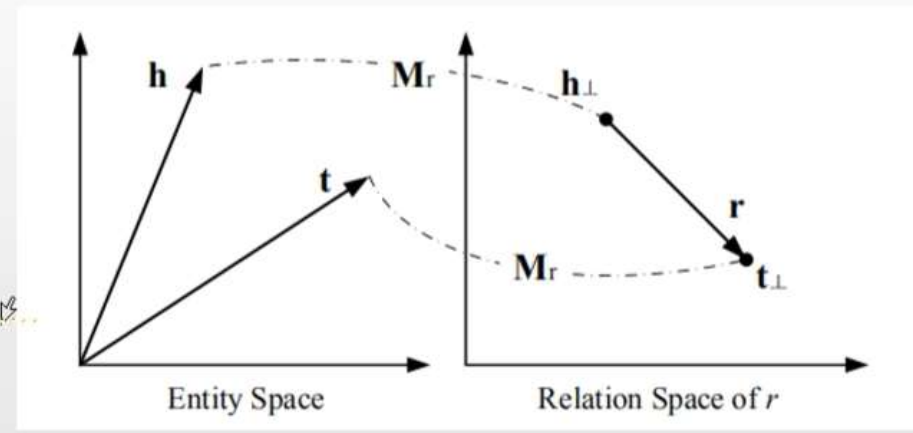
➤ **平移距离模型**; 双线性模型; 基于神经网络的模型



TransE



TransH



TransR

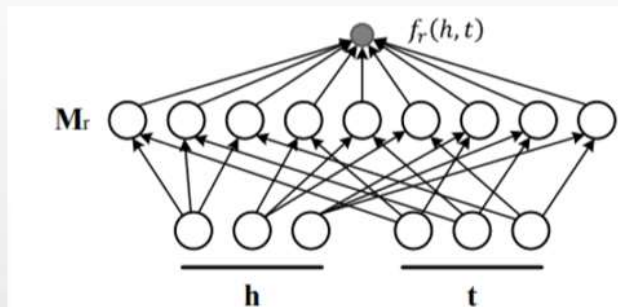
➤ 对于一个三元组 (h, r, t) , 头实体向量 h 和关系向量 r 的和应该接近尾实体向量 t , 即 $h + r \approx t$

➤ 解决 TransE 在处理一对多、多对一和多对多关系时的问题, 引入了关系特定的投影。

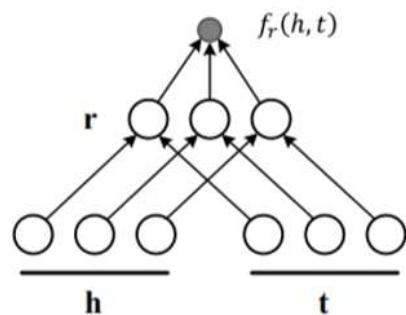
➤ 将头实体 h 从其原始向量空间映射到关系 r 的特定向量空间。

➤ 知识图谱嵌入模型:

➤ 平移距离模型; **双线性模型**; **基于神经网络的模型**

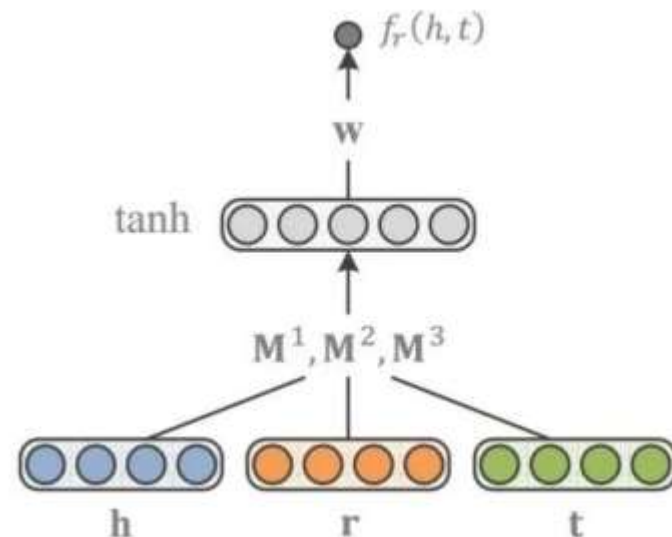


Rescal模型

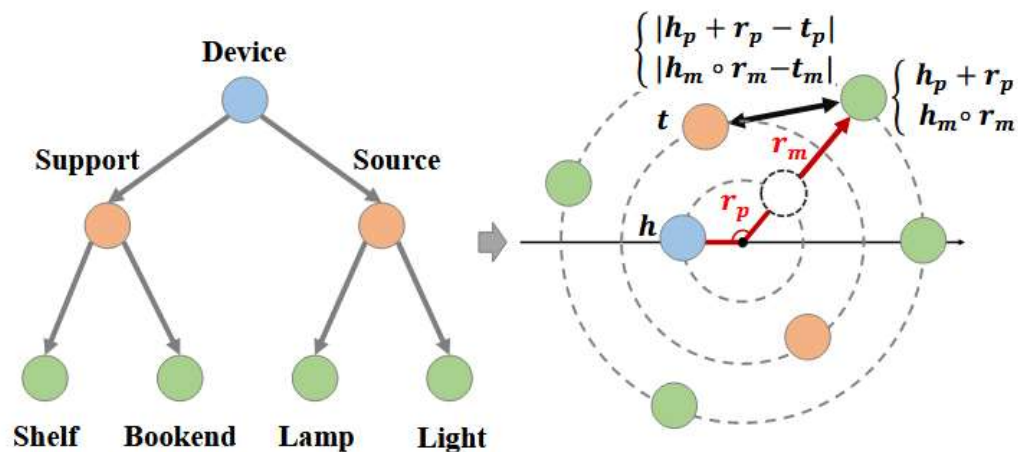


DistMult模型

- 将每个关系表示为一个满秩矩阵 M_r ，并将得分函数定义为 $f_r(h, t) = h^T M_r t$
- 简化了 RESCAL，假设关系矩阵 M_r 是一个对角矩阵，大大减少了参数的数量。



- 早期的工作使用**全连接神经网络**来对给定的三元组进行打分。
- 近年使用卷积神经网络来定义打分函数。也有一些工作将图神经网络 应用于知识图谱嵌入。



- 径向坐标用于建模层次结构中不同层次的实体，半径较小的实体位于较高的层次；角度坐标用于区分同一层次的实体，这些实体的半径大致相同，但角度不同

➤ 评价指标

- 平均互反排名 (MRR)
- 对于一个查询，模型会对所有可能的候选实体进行排序。MRR 计算的是所有查询的平均互反排名。

$$MRR = \frac{1}{n} \sum_{i=1}^n \frac{1}{rank_i}.$$

- 命中率 @N (H@N)
- 在前N个排序结果中是否包含正确答案。

$$H@N = \frac{1}{m} \sum_{j=1}^m hit_j.$$

➤ 负采样损失函数+自我对抗训练

- 对于每个正样本三元组，负采样会随机替换头实体或尾实体，生成负样本三元组。

$$L = -\log \sigma(\gamma - d_r(\mathbf{h}, \mathbf{t})) \\ - \sum_{i=1}^n p(h'_i, r, t'_i) \log \sigma(d_r(\mathbf{h}'_i, \mathbf{t}'_i) - \gamma),$$

where γ is a fixed margin, σ is the sigmoid function, and (h'_i, r, t'_i) is the i th negative triple. Moreover,

- 通过让模型自身生成的“对抗样本”来增强模型的泛化能力和对复杂关系的建模能力。
- 选择那些当前模型认为比较合理（即评分较高）但实际上是错误的负样本进行训练。

$$p(h'_j, r, t'_j | \{(h_i, r_i, t_i)\}) = \frac{\exp \alpha f_r(\mathbf{h}'_j, \mathbf{t}'_j)}{\sum_i \exp \alpha f_r(\mathbf{h}'_i, \mathbf{t}'_i)}$$

➤ 数据集

Dataset	#E	#R	#TR	#VA	#TE
WN18RR	40,493	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

➤ 优化器

- Adam (Adaptive Moment Estimation) 计算梯度的一阶矩估计 (均值) 和二阶矩估计 (方差) 来调整学习率, 在训练过程中能够更快地收敛到较好的解, 对不同参数的适应性更强, 相比传统的优化方法 (如随机梯度下降), 能够在更短的时间内获得较好的模型性能。

➤ 超参数

- 网格调参, 避免过拟合: 在验证集上计算模型在每个组合下的评估指标 (如准确率、损失值等), 最后选择使评估指标最优的超参数组合作为最终用于训练模型的超参数值。

➤ 距离函数

- 引入额外系数更好地平衡对不同层次语义信息的学习

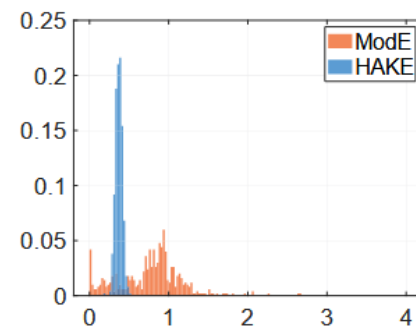
$$d_r(h, t) = \lambda_1 d_{r,m}(h_m, t_m) + \lambda_2 d_{r,p}(h_p, t_p)。$$

Main results

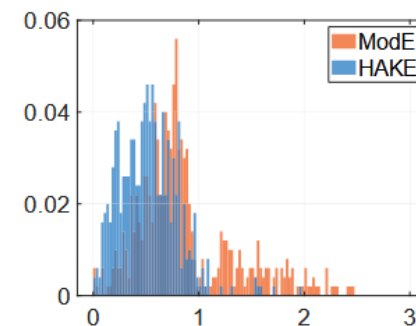
	WN18RR				FB15k-237				YAGO3-10			
	MRR	H@1	H@3	H@10	MRR	H@1	H@3	H@10	MRR	H@1	H@3	H@10
TransE	.226	-	-	.501	.294	-	-	.465	-	-	-	-
DistMult	.43	.39	.44	.49	.241	.155	.263	.419	.34	.24	.38	.54
ConvE	.43	.40	.44	.52	.325	.237	.356	.501	.44	.35	.49	.62
ComplEx	.44	.41	.46	.51	.247	.158	.275	.428	.36	.26	.40	.55
RotatE	.476	.428	.492	.571	.338	.241	.375	.533	.495	.402	.550	.670
ModE	.472	.427	.486	.564	.341	.244	.380	.534	.510	.421	.562	.660
HAKE	.497	.452	.516	.582	.346	.250	.381	.542	.545	.462	.596	.694

关系嵌入分析

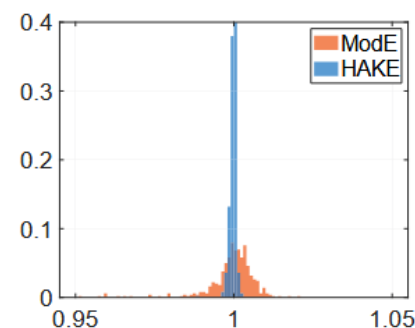
- ModE 和 HAKE 两种模型在六个不同关系上的模量分布直方图
- 连接同一层次实体 (`_similar_to` 和 `friend`)
- 尾实体层次高于头实体 (`_hypernym` 和 `isLocatedIn`)
- 尾实体层次低于头实体 (`_member_meronym` 和 `_has_part`)
- HAKE 的关系嵌入模量方差更低，可以更准确捕捉实体在语义层次中的位置关系。



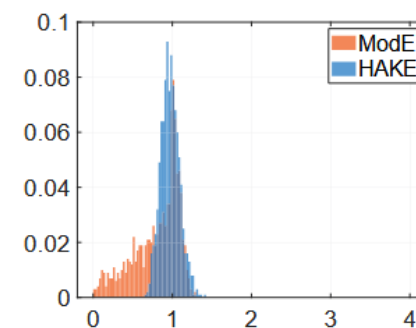
(a) `_hypernym`



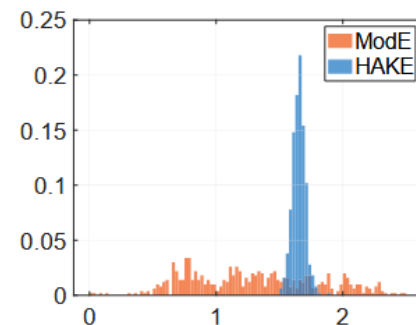
(b) `isLocatedIn`



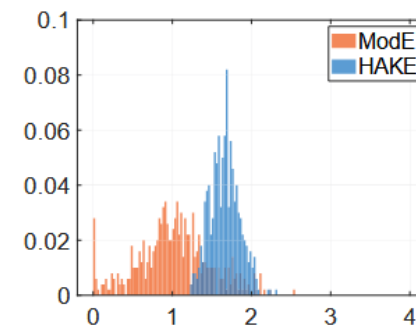
(c) `_similar_to`



(d) `friend`



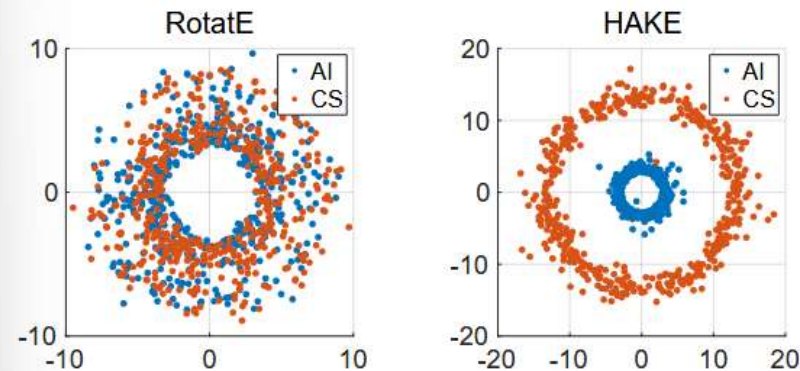
(e) `_member_meronym`



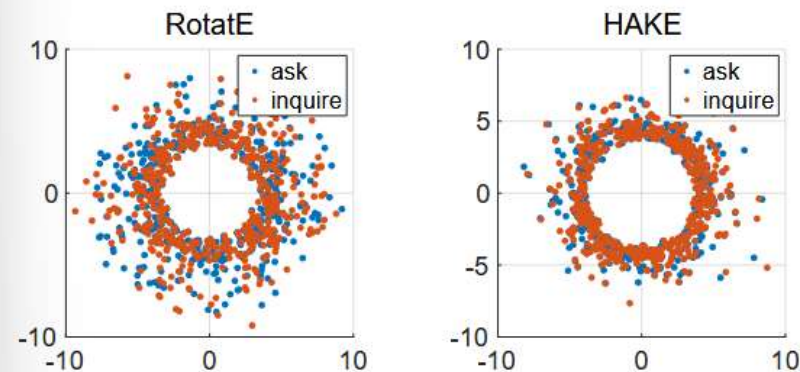
(f) `_has_part`

实体嵌入分析

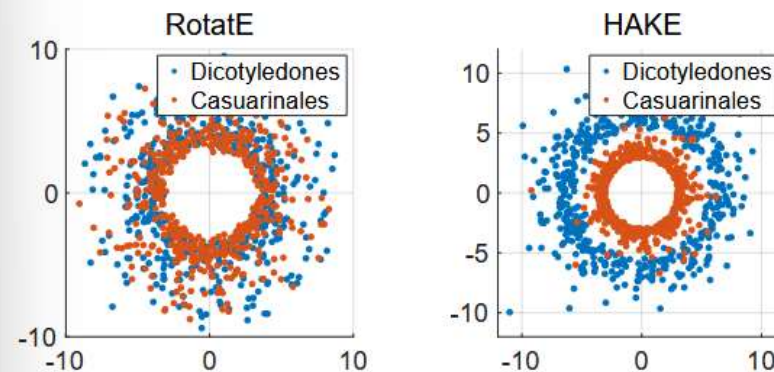
- 可视化实体向量的分布情况，观察模型是否能够将不同层次或同一层次的实体在向量空间中进行合理的区分和聚类，从而判断模型对实体语义层次的捕捉能力。
- A (“人工智能”，“上位词”，“计算机科学”) 实体化
- B (“询问”，“动词组”，“打听”)
- C (“双子叶植物纲”，“部分成员”，“木麻黄目”)



(a) (AI, *hypernym*, CS)



(b) (ask, *verb_group*, inquire)



(c) (Dicotyledones, *member_meronym*, Casuarinales)

Table 4: Ablation results on WN18RR, FB15k-237 and YAGO3-10 datasets. The symbols **m**, **p**, and **b** represent the modulus part, the phase part, and the mixture bias term, respectively.

			WN18RR				FB15k-237				YAGO3-10			
			MRR	H@1	H@3	H@10	MRR	H@1	H@3	H@10	MRR	H@1	H@3	H@10
✓			.240	.047	.404	.527	.258	.121	.333	.508	.476	.374	.541	.658
	✓		.465	.423	.480	.550	.324	.226	.361	.519	.480	.383	.532	.664
✓	✓		.496	.449	.517	.584	.336	.239	.373	.533	.522	.429	.581	.693
✓	✓	✓	.497	.452	.516	.582	.346	.250	.381	.542	.545	.462	.596	.694

- 混合偏差项有助于提升模型性能
- 模量部分和相位部分单独使用时存在局限性
- 两者结合能够发挥最大优势

Table 5: Comparison results with TKRL models (Xie, Liu, and Sun 2016) on FB15k dataset. RHE, WHE, RHE+STC, and WHE+STC are four versions of TKRL model, of which the results are taken from the original paper.

	HAKE	RHE	WHE	RHE+STC	WHE+STC
H@10	.884	.694	.696	.731	.734

- 长尾分布：是指一些实体或关系出现的频率较低，而传统的模型在处理这些低频项时往往效果不佳，因为缺乏足够的训练数据来学习它们的有效表示。
- 它们所属的层次类型可能与其他高频实体或关系存在关联。TKRL 模型可以利用这种层次类型的关联信息，通过共享或迁移学习的方式，从高频部分获取一些有用的信息来帮助表示低频部分。

- TKRL (Type-embodied Knowledge Representation Learning) 模型：
- 将实体的层次类型信息视为实体的投影矩阵
- 捕捉知识图谱中的实体在不同关系中可能具有不同的语义角色
- 优点：考虑了实体的层次类型信息，将层次类型视为实体的投影矩阵，能够建模实体在不同场景中的多种表示形式；在知识图补全和三重分类任务上有较好的表现，特别是在长尾分布情况下。

- JOIE (Jointly Optimizing Interaction and Entity Embeddings)
- 结合了外部文本信息来增强实体和关系的表示。通过联合学习知识图谱中的三元组信息和外部文本中与实体相关的描述信息，利用文本中的语义信息来丰富实体类型表示。例如，对于一个知识图谱中的“苹果公司”实体，JOIE 会利用新闻报道、公司简介等外部文本中关于“苹果公司”的描述，提取语义特征并融入到实体的类型表示中。

- HRFN (Hierarchical Relation-aware Factorization Network)
- 构建了一个层次关系感知的分解网络，通过将关系分解为多个层次的组件来建模语义层次结构。在每个层次上，学习实体和关系组件的表示，并通过层次结构将它们组合起来，从而捕捉不同层次的语义关系。