

1 融合多源信息

多源信息能够帮助构建更加精准的知识表示, 经典的知识表示学习模型往往仅关注知识图谱自身固有的结构信息而忽略了蕴含在多源信息中的丰富知识。这些多源信息包括: 类型信息, 实体描述信息以及关系路径等, 本章中将讨论如何将外部信息源融合进知识表示空间中学习以增强知识图谱中知识的表示。

$$L = \sum_{(h,r,t) \in S} \sum_{(h',r',t') \in S'} \max(0, f(h,r,t) + \gamma - f(h',r',t')) \quad (1)$$

1.1 类别信息

最常用的类别信息就是实体类别, Guo 等人^[3]提出语义平滑嵌入模型 (Semantically Smooth Embedding, SSE), 基于的思想是属于同一语义类型的实体在嵌入空间中距离比较近。SSE 利用两种流行学算法 Laplacian eigenmaps 和 Locally linear embedding 来约束这种平滑性假设。Lacian eigenmaps 算法要求一个实体与它具有同类别的实体在嵌入空间中距离更近, 这种约束形式用公式表示为:

$$\mathcal{R}_1 = \|\mathbf{h} - \mathbf{t}\|_2^2 w_{ht}^{(1)} \quad (2)$$

其中 \mathbf{h}, \mathbf{t} 分别是实体 h, t 的嵌入表示, $w_{ht}^{(1)}$ 代表两个实体之间的邻接矩阵, 当两个实体 h 和 t 属于同一类型时, $w_{ht}^{(1)} = 1$, 反之为 0。Locally linear embedding 算法要求一个实体可以由它临近的实体经过线性组合表示出来。这里临近的实体就是指属于同一个类别的实体, 这种约束形式用公式表示为:

$$\mathcal{R}_2 = \left\| \mathbf{h} - \sum_{t \in \mathbb{N}(h)} w_{ht}^{(2)} \mathbf{t} \right\|_2^2 \quad (3)$$

其中 $\mathbb{N}(h)$ 代表与实体 h 类型相同的 K 个实体集合, 集合的个数 K 是一个超参数。当实体 t 在集合 $\mathbb{N}(h)$ 中, 则 $w_{ht}^{(2)}$ 为 1, 反之为 0。SSE 模型的能量函数与 TransE^[TransE] 模型一致, 将 $\mathcal{R}_1, \mathcal{R}_2$ 加到最大间隔方法里作为整个模型损失函数的一个正则化项从而达到约束嵌入空间语义平滑的作用。实验表明, SSE 模型学习后的嵌入空间中, 语义类别一致的实体很好的聚在了一起。

SSE 模型的缺点在于默认所有实体只有一个类别, 而现实世界中的实体不仅有多个类别而且每一个类别是有层次结构的。如图 1 所示, 从图中可以

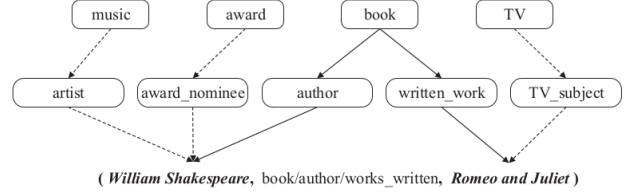


图 1 取自 TKRL^[15]

看到头实体 William Shakespeare 存在很多实体类型, 而且每一个实体类型都有子类型。基于以上问题, Xie 等人^[15]提出一种融合实体层次类型信息的模型 TKRL (type-embodied knowledge representation learning), 引入具有层次结构的实体类别信息以及与关系之间的约束信息, TKRL 模型的能量函数定义为:

$$f(h, r, t) = \|M_{r,h}h + r - M_{rt}t\| \quad (4)$$

其中 $M_{r,h}, M_{r,t}$ 分别表示的是头实体和尾实体的层次类型映射矩阵, 由于一个类别有多个子类别, TKRL 为每一个子类型也设置了映射矩阵, 使用这些子类型的映射矩阵构建层次类型映射矩阵。以头实体 h 为例, h 类别集合是 $c = \{c_1, c_2, \dots, c_n\}$, n 代表类别数目, c_i 是 h 的第 i 个类别, 每一个类别有很多个子类别, $c_i = \{c_i^1, c_i^2, \dots, c_i^m\}$, m 代表类别 i 的子类别的个数, 其中 c_i^j 代表 h 的第 i 个类别的第 j 个子类别。头实体 h 的类型映射矩阵表示为:

$$M_{rh} = \alpha_1 M_{c_1^h} + \alpha_2 M_{c_2^h} + \dots + \alpha_n M_{c_n^h} \quad (5)$$

其中 $M_{c_i^h}$ 代表 h 在它的第 i 个类型下的映射矩阵, α_i 就是权重, 表示的就是 h 在当前第 i 个类别下的凸显程度, 尾实体 h 的类型映射矩阵计算方式同理。为了进一步的处理层次类型信息, TKRL 采用两种层次编码器。第一种是递归层次编码器:

$$M_{c_i^h} = \prod_{j=1}^m M_{c_i^h}^{(j)} = M_{c_i^h}^{(1)} M_{c_i^h}^{(2)} \dots M_{c_i^h}^{(m)} \quad (6)$$

第二种是加权层次编码器：

$$M_{c_i^h} = \sum_{j=1}^m \beta_j M_{c_i^h}^{(j)} \quad (7)$$

$$= \beta_1 M_{c_i^h}^{(1)} + \beta_2 M_{c_i^h}^{(2)} + \dots + \beta_m M_{c_i^h}^{(m)}$$

其中 $M_{c_i^h}^{(j)}$ 代表的是头实体 h 的第 i 个类型的第 j 个子类型的映射矩阵。TRKL 的损失函数采用做大间隔方法。

实体类别信息还能在模型中作为类别限制，帮助模型学习更好的知识表示。不同于 Krompaß 等人^[6] 使用硬类型限制控制负例三元组的生成，强制替换的实体与原来的实体具有同样的类别，TKRL 模型采用一种软类型限制 (Soft Type Constraint, STC) 机制，STC 的策略是以某一定的比例选取那些与被替换实体有相同类型的实体，如下公式

$$P(\tilde{e} \in E_c) = \frac{(k+1)|E_c|}{|E| + k|E_c|} \quad (8)$$

其中 E 是三元组集合， $E_c \in E$ 代表的一个实体集合其中所有实体都有类别 c ， k 是一个超参数。

在测试时对于给定的一种关系，TKRL 模型限定这种关系的头尾实体的具体类别，这种做法泛化型不高并且并不是所有的实体都具有层次类别。针对这些问题，Jin 等人^[TEKRL] 提出的 TEKRL 模型解决了其它模型在引入实体类别信息时额外添加规则的问题，TEKRL 直接利用数据集中最底层的实体类别，通过引入注意力机制学习实体类别与关系之间的相关性，通过对实体类别表示加权求和取代制定实体类别和关系之间约束的操作。TEKRL 模型为每一个实体构建基于结构和基于类别的两种实体表示，模型的能量函数定义为：

$$f = f_{ss} + \beta f_{cc}$$

其中 β 用来调整基于类别的表示在模型中的作用大小， f_{ss} 和 f_{cc} 分别代表基于结构的能量函数和基于类别的能量函数。

$$f_{ss} = \|\mathbf{h}_s + \mathbf{r} - \mathbf{t}_s\|$$

$$f_{cc} = \|\mathbf{h}_c + \mathbf{r} - \mathbf{t}_c\|$$

\mathbf{h}_s 和 \mathbf{t}_s 是基于结构的实体表示，它们通过知识图谱中结构化信息学习。 \mathbf{h}_c 和 \mathbf{t}_c 是基于类别的实体表

示，通过注意力机制得到实体类别表示与三元组中关系的相关性，然后对所有类别的表示加权求和作为实体的基于类别的表示。

除了考虑实体类别信息外，Zhang 等人^[19] 提出一种利用关系层次结构的知识表示模型 HRS (Hierarchical Relation Structure)。HRS 将知识图谱中的关系划分为一个三层的层次结构，顶层关系 r_c 是由语义相近的关系聚类而成，中间层指特定的关系 r' ，底层关系 r_s 是将每一个关系再进一步的划分成多个子关系。在训练时，对于每一个三元组 (h, r, t) ，其中关系 r 的嵌入表示 \mathbf{r} 为三层关系的嵌入表示之和 $\mathbf{r} = \mathbf{r}_c + \mathbf{r}' + \mathbf{r}_s$ 。HRS 模型的损失函数定义为

$$L = L_{Orig} + L_{HRS} \quad (9)$$

其中 L_{Orig} 采用经典模型 TransE, TransR 等的最大间隔方法， L_{HRS} 作为正则化项约束模型学习层次关系的表示。

$$L_{HRS} = \lambda_1 \sum_{\mathbf{r}_c \in \mathcal{C}} \|\mathbf{r}_c\|_2^2 + \lambda_2 \sum_{\mathbf{r}' \in \mathcal{R}} \|\mathbf{r}'\|_2^2 + \lambda_3 \sum_{\mathbf{r}_s \in \mathcal{S}} \|\mathbf{r}_s\|_2^2 \quad (10)$$

其中 \mathcal{C} , \mathcal{R} , \mathcal{S} 分别代表关系聚类集合，关系集合，子关系聚类集合。此外在知识图谱中某些关系代表的是实体的属性信息，而属性与关系的区别在于属性有明显的一对多关系。Lin 等人^[7] 提出一种知识表示模型从所有的关系分离出关系与属性，然后再建模实体表示、关系表示、属性表示。

1.2 文本描述

知识图谱中很多实体都是带有描述信息的，这些信息能够作为知识图谱中结构化信息的辅助，帮助模型学习更精准的知识表示，知识库的构建资源也往往是从文本中获取，因此实体描述文本能天然的与知识空间进行交互。如图 2 所示。这些描述文

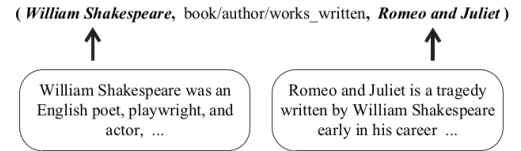


图 2 取自 DKRL^[17]

本是对实体相关信息的详细描述。那些仅仅基于知

识图谱结构化信息知识表示模型无法处理不在知识图谱中的实体 (out of KG)，而联合文本嵌入的方式可以做到互补使得模型学习到那些在文本中出现而不在知识图谱中的实体。

Wang 等人^[12] 首先提出联合知识图谱和实体描述文本的知识表示学习模型，模型基于 TransE^[TransE] 和 Skip-gram[9] 模型的思想。整个模型分成三个部分：知识模型、文本模型、对齐模型。知识模型基于平移模型的思想学习实体和关系的表示，损失函数定义如下：

$$\begin{aligned} \mathcal{L}_K = & - \sum_{(h,r,t)} \log \Pr(h|r, t) \\ & + \log \Pr(t|h, r) + \log \Pr(r|h, t) \end{aligned} \quad (11)$$

以 $\Pr(h|r, t)$ 为例：

$$\Pr(h|r, t) = \frac{\exp z(h, r, t)}{\sum_{\tilde{h}} \exp z(\tilde{h}, r, t)} \quad (12)$$

$$z(h, r, t) = b - \frac{1}{2} \|\mathbf{h} + \mathbf{r} - \mathbf{t}\|_2^2 \quad (13)$$

文本模型用来学习实体描述文本中单词的表示，基于 Skip-gram 的思想，定义文本模型的损失函数如下：

$$\mathcal{L}_T = - \sum_{(w,v) \in \mathcal{C}} \log \Pr(w|v) \quad (14)$$

其中 $\Pr(w|v)$ 代表两个单词在滑动窗口内的共现概率， \mathcal{C} 代表滑动窗口内共现单词的集合。

$$\Pr(w|v) = \frac{\exp z(w, v)}{\sum_{\tilde{w}} \exp z(\tilde{w}, v)} \quad (15)$$

$$z(w, v) = b - \frac{1}{2} \|\mathbf{w} - \mathbf{v}\|_2^2 \quad (16)$$

最后对齐模型实现知识空间和文本空间对齐，对齐原则是利用实体名称或者维基百科锚文本。模型的损失函数为这三个子模块之和

$$\mathcal{L} = \mathcal{L}_K + \mathcal{L}_T + \mathcal{L}_A \quad (17)$$

\mathcal{L}_A 是对齐模型的损失函数。但是由于实体名称歧义性较大的问题利用实体名称对齐的原则会打乱文本原有的语义空间，而利用维基百科锚文本对齐的原则依赖于特定的数据源，这使得这种方式无法应用到其它数据源领域。为了解决以上问题，Zhong 等

人^[20] 提出利用实体描述文本作为对齐原则，定义对齐模块的损失函数为：

$$\mathcal{L}_A = - \sum_{e \in E} \sum_{w \in D_e} \log \Pr(w|e) + \log \Pr(e|w) \quad (18)$$

E 代表知识图谱中的实体集合， D_e 代表实体 e 的描述文本， $\Pr(w|e)$ 和 $\Pr(e|w)$ 的计算方式见公式 (15)，(16)。类似的还有 Zhang 等人^[18] 也尝试使用实体名称和实体描述文本中词向量的均值作为实体的文本表示。

以上的模型考虑的是单词级别的文本信息而没有利用整个文本的语序语义信息。Xie 等人^[17] 提出一种融合实体描述的知识表示模型 (Description-embodied knowledge representation learning, DKRL)。DKRL 在 TransE 模型的基础上融合实体描述的文本信息，为每一个实体设置两种知识表示。第一种是基于知识图谱结构化信息的表示，可以随机初始化或者利用 TransE 模型预训练好的实体表示。第二种是基于文本描述的表示，使用连续词袋 (continuous bag-of-words, CBOW) 模型和卷积神经网络 (Convolutional neural network, CNN) 模型从文本中构建。给定三元组 (h, r, t) ，DKRL 模型的能量函数定义如公式：

$$\begin{aligned} E(h, r, t) = & \|\mathbf{h}_S + \mathbf{r} - \mathbf{t}_S\| + \|\mathbf{h}_S + \mathbf{r} - \mathbf{t}_D\| \\ & + \|\mathbf{h}_D + \mathbf{r} - \mathbf{t}_S\| + \|\mathbf{h}_D + \mathbf{r} - \mathbf{t}_D\| \end{aligned} \quad (19)$$

其中 $\mathbf{h}_S/\mathbf{h}_D$ ， $\mathbf{t}_S/\mathbf{t}_D$ 分别代表头实体/尾实体的基于结构的表示和基于实体描述的表示， \mathbf{r} 代表关系的知识表示。通过混合项 $\|\mathbf{h}_S + \mathbf{r} - \mathbf{t}_D\|$ ， $\|\mathbf{h}_D + \mathbf{r} - \mathbf{t}_S\|$ 的限制，DKRL 模型将实体的两种不同的表示向量映射到同一个语义空间中。实验结果表明 DKRL 模型在很多知识图谱任务上效果表现很好。然而 DKRL 模型是一种弱关联建模，在融合实体基于结构的表示和基于文本的表示时没有足够的交互过程。Xiao 等人^[14] 提出语义空间投影模型 (Semantic space projection, SSP)，将三元组的嵌入表示投影到语义子空间上，在语义子空间上学习实体的两种表示。与 DKRL 不同的是 SSP 采用主题 (topic) 模型建模实体的文本表示。

Wang 等人^[13] 提出一种增强型文本知识嵌入模型 (Text-enhanced knowledge embedding, TEKE)。

TEKE 不仅将文本信息融入到知识图谱的空间中, 并且每一个关系对于不同的头实体和尾实体都有不同的表示, 这样可以更好的解决 1-N, N-1 以及 N-N 的关系问题。首先使用实体连接工具标注出文本中是知识图谱中实体的单词并且基于这些实体单词构建共现网络 (co-occurrence network)。对于文本中的实体单词 e , 定义它的上下文集合为 $n(e)$, 它的文本向量表示为 $\mathbf{n}(e)$, $\mathbf{n}(e)$ 是 $n(e)$ 中所有单词词向量的加权求和, 权重为两个实体单词共同出现的频次归一化后的值。给定三元组 (h, t, r) , TEKE 模型认为两个实体 h, t 的上下文集合 $n(h)$, $n(t)$ 中共同出现的单词往往反映着两个实体之间的关系, 定义 $n(h, t) = n(h) \cap n(t)$ 作为两个实体上下文的匹配集合, 其向量表示为 $\mathbf{n}(h, t)$ 。最后将文本的语义空间映射到知识图谱的语义空间:

$$\begin{aligned}\hat{h} &= \mathbf{n}(h)A + \mathbf{h} \\ \hat{r} &= \mathbf{n}(h, t)B + \mathbf{r} \\ \hat{t} &= \mathbf{n}(t)A + \mathbf{t}\end{aligned}\quad (20)$$

其中 A, B 是映射矩阵, $\mathbf{h}, \mathbf{r}, \mathbf{t}$ 是偏置项, 它们都是可学习的参数, 模型的能量函数定义为 $f(h, r, t) = \|\hat{h} + \hat{r} - \hat{t}\|_2^2$, 利用最大间隔方法训练。

然而 TEKE 模型并没有考虑到实体和关系的模糊性, 所谓模糊性是指关系所表达的精确语义依赖于头尾实体的语义。此外两个实体单词的上下文集合中共同出现的单词并不能总是表达两个实体单词之间的关系, 这种方式会给某些三元组带来噪声信息。基于以上问题, An 等人^[1] 提出 ATEKE 模型, 利用实体描述文本和特定三元组的关系提及 (relation mention) 来增强实体和关系的表示。为了处理噪声的问题, 在提取关系提及时添加两个约束条件: (1) 包含特定三元组的两个实体以及关系的下义词或同义词; (2) 关系和句子中某一单词的相似性超过阈值。满足两个约束条件之一的句子被认为是准确的关系提及。通过注意力机制计算关系提及和实体描述的相关性, 得到三元组 (h, r, t) 的文本表示 $(\mathbf{h}_D, \mathbf{r}_D, \mathbf{t}_D)$ 。最后将基于文本的表示和基于结构的表示加权求和

作为实体表示。

$$\begin{aligned}Re(\mathbf{h}) &= \alpha Re(\mathbf{h}_S) + (1 - \alpha)\mathbf{h}_D \\ Re(\mathbf{r}) &= \alpha Re(\mathbf{r}_S) + (1 - \alpha)\mathbf{r}_D \\ Re(\mathbf{t}) &= \alpha Re(\mathbf{t}_S) + (1 - \alpha)\mathbf{t}_D\end{aligned}$$

其中 $0 \leq \alpha \leq 1$, Re 表示取向量的实值部分。

1.3 关系路径

关系路径是指两个实体之间的多步关系而不仅仅限于两个实体之间直接相连的关系。多步关系包含了两个实体之间丰富的语义关系, 多步推理。

Lin 等人^[8] 提出 PTransE 模型, 在 TransE 模型的基础上将两个实体之间的多步关系路径看做两个实体之间相连的关系。由于两个实体之间存在大量的关系路径, PTransE 模型采用路径约束资源分配 (path-constraint resource allocation, PCRA) 算法提取出可以用来建模两个实体之间关系的可靠的关系路径。PTransE 模型的优化目标分为两部分: 第一部分是利用 TransE 模型建模两个实体之间直接相连的关系; 第二部分是建模两个实体之间多关系路径。定义 $P(h, t) = \{p_1, p_2, \dots, p_N\}$ 代表实体 h 和 t 之间的多关系路径的集合, 模型的能量函数定义如下

$$G(h, r, t) = f(h, r, t) + f(h, p, t) \quad (21)$$

其中 $f(h, r, t)$ 即为公式 (1) 中 TransE 模型的能量函数。 $p = (r_1, r_2, \dots, r_l)$, 它是给定头实体 h 和尾实体 t 下的某一条多步关系路径, r_i 代表路径 p 上的第 i 个关系, l 是路径上关系的个数。 $f(h, p, t)$ 代表给定实体 h 和 t 以及多跳关系路径 p 的能量函数。

$$f(h, p, t) = \|\mathbf{p} - \mathbf{r}\| \quad (22)$$

\mathbf{p} 代表关系路径 p 的向量表示。PTransE 模型采用三种方式计算 \mathbf{p} 。

$$\begin{aligned}\mathbf{p} &= \mathbf{r}_1 + \mathbf{r}_2 + \dots + \mathbf{r}_l \quad \text{ADD} \\ \mathbf{p} &= \mathbf{r}_1 \cdot \mathbf{r}_2 \cdot \dots \cdot \mathbf{r}_l \quad \text{MUL} \\ \mathbf{c}_i &= f(W[\mathbf{c}_{i-1}; \mathbf{r}_i]) \quad \text{RNN}\end{aligned}$$

ADD 表示将路径上所有关系的向量表示相加, MUL 表示相乘, RNN 代表利用循环神经网络, 此时 \mathbf{p} 为

路径上最后一个关系的隐藏状态。模型的目标函数为：

$$L = L(h, r, t) + \frac{1}{Z} \sum_{p \in P(h, t)} R(p|h, t) L(h, p, t) \quad (23)$$

$L(h, r, t)$ 为公式 (1) 中 TransE 模型的目标函数, $R(p|h, t)$ 是利用 PCRA 算法计算的给定头尾实体 h, t 下关系路径 p 的可信度, Z 是规范化因子, $L(h, p, t)$ 用来建模实体基于关系路径的表示, 能量函数为 $f(h, p, t)$ 。

Guu 等人^[4] 提出另一种融合多步关系路径的知识表示学习模型, 利用关系路径构建新的三元组并且对 TransE 模型和 RESCAL^[RESCAL] 模型进行了扩展。给定头实体 h 和尾实体 t 以及长度为 l 的多跳关系路径 p , 扩展 TransE 模型后能量函数为

$$f(h, p, t) = \|\mathbf{h} + \mathbf{r}_1 + \dots + \mathbf{r}_l - \mathbf{t}\|$$

扩展 RESCAL 模型后的能量函数为：

$$f(h, p, t) = \mathbf{h}^T (\mathbf{M}_1 \circ \dots \circ \mathbf{M}_l) \mathbf{t}$$

上面两个模型的共同点都是从众多关系路径中选择一部分子集同时限制路径的长度为 2 步或者 3 步, Toutanova 等人^[11] 利用动态规划算法建模两个实体之间所有可能的关系路径。

两条头尾实体相同的多步关系路径, 即使路径中间的关系也相同但是由于路径中间实体的不同, 两条路径表达的也可能是不同的含义。基于这个问题 Seo 等人^[10] 提出 RKRL 模型, 同时考虑关系路径上中间的实体和关系信息。

其中 \mathbf{r}_i 代表多步关系路径上第 i 个关系的表示。RKRL 模型的能量函数和损失函数与 PTransE 模型一致, 见公式 (21), (23)。其中基于关系路径的能量函数定义为：

$$f(h, p, t) = \|\mathbf{h} + \mathbf{p} - \mathbf{t}\|$$

$$\mathbf{p} = W_{e_1} \mathbf{r}_1 + \dots + W_{e_{l-1}} \mathbf{r}_{l-1} + \mathbf{r}_l$$

RKRL 采用两种特征来衡量路径 p 的可靠性, 第一种特征是 PTransE 中提出的资源分配算法 PCRA, 但是由于该算法仅仅依靠资源流动的量来衡量可靠性可能不足以判别路径的语义关系, 因此额外添加路径上与实体相连的关系出现的频率这一特征。

1.4 其它信息

除了可以融入上面三种信息源外, 还有一些其它的信息也可以融入到知识图谱的语义空间中, 增强知识的语义表示。如 Xie 等人^[16] 提出的 IKRL 模型, 尝试融合实体图像信息学习实体跨模态的知识表示, 与 TKRL 模型类似仍然是为每一个实体设置两种知识表示, 基于结构的和基于图像的表示, 采用 AlexNet^[5] 网络作为图像的特征提取器, 模型的能量函数和损失函数与 TKRL 一致。

以上介绍的知识表示学习模型都是将知识图谱看做三元组的集合, 孤立的学习每个三元组的表示, 尽管融合多步关系路径的模型引入了其它的实体和关系, 但是绝大部分模型也只是将关系路径看做是两个实体之间的新关系, 知识图谱的图结构信息并没有被完全利用。Feng 等人^[2] 提出一种图感知的表示学习模型 GAKE (graph aware knowledge embedding), 模型给出一个新的概念 graph context。Graph context 定义为知识图谱中与一个实体有关联的三种集合: neighbor context, path context 和 edge context。Neighbor context 是一个实体的三元组集合, path context 是与一个实体相连的关系路径集合, edge context 是一个实体所有的关系集合。由于不同的实体和关系对于一个实体的影响不同, GAKE 利用注意力机制赋予不同关系和实体不同的权重。实验证明这种考虑图结构的知识表示模型在元组分类和链接预测任务上达到很好的效果。

References

- [1] Bo An et al. "Accurate Text-Enhanced Knowledge Graph Representation Learning". In: *Proceedings of the 2018 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long Papers)*. New Orleans, Louisiana: Association for Computational Linguistics, June 2018, pp. 745–755. DOI: 10.18653/v1/N18-1068. URL: <https://www.aclweb.org/anthology/N18-1068>.

- [2] Jun Feng et al. “GAKE: Graph Aware Knowledge Embedding”. In: *COLING 2016, 26th International Conference on Computational Linguistics, Proceedings of the Conference: Technical Papers, December 11-16, 2016, Osaka, Japan*. Ed. by Nicoletta Calzolari, Yuji Matsumoto, and Rashmi Prasad. ACL, 2016, pp. 641–651. URL: <https://www.aclweb.org/anthology/C16-1062/>.
- [3] Shu Guo et al. “Semantically Smooth Knowledge Graph Embedding”. In: *Proceedings of the 53rd Annual Meeting of the Association for Computational Linguistics and the 7th International Joint Conference on Natural Language Processing of the Asian Federation of Natural Language Processing, ACL 2015, July 26-31, 2015, Beijing, China, Volume 1: Long Papers*. The Association for Computer Linguistics, 2015, pp. 84–94. DOI: 10.3115/v1/p15-1009. URL: <https://doi.org/10.3115/v1/p15-1009>.
- [4] Kelvin Guu, John Miller, and Percy Liang. “Traversing Knowledge Graphs in Vector Space”. In: *Proceedings of the 2015 Conference on Empirical Methods in Natural Language Processing, EMNLP 2015, Lisbon, Portugal, September 17-21, 2015*. Ed. by Lluís Màrquez et al. The Association for Computational Linguistics, 2015, pp. 318–327. DOI: 10.18653/v1/d15-1038. URL: <https://doi.org/10.18653/v1/d15-1038>.
- [5] Alex Krizhevsky, Ilya Sutskever, and Geoffrey E. Hinton. “ImageNet Classification with Deep Convolutional Neural Networks”. In: *Advances in Neural Information Processing Systems 25: 26th Annual Conference on Neural Information Processing Systems 2012. Proceedings of a meeting held December 3-6, 2012, Lake Tahoe, Nevada, United States*. Ed. by Peter L. Bartlett et al. 2012, pp. 1106–1114. URL: <http://papers.nips.cc/paper/4824-imagenet-classification-with-deep-convolutional-neural-networks>.
- [6] Denis Krompaß, Stephan Baier, and Volker Tresp. “Type-Constrained Representation Learning in Knowledge Graphs”. In: *The Semantic Web - ISWC 2015 - 14th International Semantic Web Conference, Bethlehem, PA, USA, October 11-15, 2015, Proceedings, Part I*. Ed. by Marcelo Arenas et al. Vol. 9366. Lecture Notes in Computer Science. Springer, 2015, pp. 640–655. DOI: 10.1007/978-3-319-25007-6_37. URL: https://doi.org/10.1007/978-3-319-25007-6_37.
- [7] Yankai Lin, Zhiyuan Liu, and Maosong Sun. “Knowledge Representation Learning with Entities, Attributes and Relations”. In: *Proceedings of the Twenty-Fifth International Joint Conference on Artificial Intelligence, IJCAI 2016, New York, NY, USA, 9-15 July 2016*. Ed. by Subbarao Kambhampati. IJCAI/AAAI Press, 2016, pp. 2866–2872. URL: <http://www.ijcai.org/Abstract/16/407>.
- [8] Yankai Lin et al. “Modeling Relation Paths for Representation Learning of Knowledge Bases”. In: *Proceedings of the 2015 Conference on Empirical Methods in Natural Language Processing, EMNLP 2015, Lisbon, Portugal, September 17-21, 2015*. Ed. by Lluís Màrquez et al. The Association for Computational Linguistics, 2015, pp. 705–714. DOI: 10.18653/v1/d15-1082. URL: <https://doi.org/10.18653/v1/d15-1082>.
- [9] Tomas Mikolov et al. “Distributed Representations of Words and Phrases and their Compositionality”. In: *Advances in Neural Information Processing Systems 26: 27th Annual Conference on Neural Information Processing Systems 2013. Proceedings of a meeting held December 5-8, 2013, Lake Tahoe, Nevada,*

- United States*. Ed. by Christopher J. C. Burges et al. 2013, pp. 3111–3119. URL: <http://papers.nips.cc/paper/5021-distributed-representations-of-words-and-phrases-and-their-compositionality>.
- [10] Seungmin Seo, Byungkook Oh, and Kyong-Ho Lee. “Reliable Knowledge Graph Path Representation Learning”. In: *IEEE Access* 8 (2020), pp. 32816–32825. DOI: 10.1109/ACCESS.2020.2973923. URL: <https://doi.org/10.1109/ACCESS.2020.2973923>.
- [11] Kristina Toutanova et al. “Compositional Learning of Embeddings for Relation Paths in Knowledge Base and Text”. In: *Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics, ACL 2016, August 7-12, 2016, Berlin, Germany, Volume 1: Long Papers*. The Association for Computer Linguistics, 2016. DOI: 10.18653/v1/p16-1136. URL: <https://doi.org/10.18653/v1/p16-1136>.
- [12] Zhen Wang et al. “Knowledge Graph and Text Jointly Embedding”. In: *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP)*. Doha, Qatar: Association for Computational Linguistics, Oct. 2014, pp. 1591–1601. DOI: 10.3115/v1/D14-1167. URL: <https://www.aclweb.org/anthology/D14-1167>.
- [13] Zhigang Wang and Juan-Zi Li. “Text-Enhanced Representation Learning for Knowledge Graph”. In: *Proceedings of the Twenty-Fifth International Joint Conference on Artificial Intelligence, IJCAI 2016, New York, NY, USA, 9-15 July 2016*. Ed. by Subbarao Kambhampati. IJCAI/AAAI Press, 2016, pp. 1293–1299. URL: <http://www.ijcai.org/Abstract/16/187>.
- [14] Han Xiao et al. “SSP: Semantic Space Projection for Knowledge Graph Embedding with Text Descriptions”. In: *Proceedings of the Thirty-First AAAI Conference on Artificial Intelligence, February 4-9, 2017, San Francisco, California, USA*. Ed. by Satinder P. Singh and Shaul Markovitch. AAAI Press, 2017, pp. 3104–3110. URL: <http://aaai.org/ocs/index.php/AAAI/AAAI17/paper/view/14306>.
- [15] Ruobing Xie, Zhiyuan Liu, and Maosong Sun. “Representation Learning of Knowledge Graphs with Hierarchical Types”. In: *Proceedings of the Twenty-Fifth International Joint Conference on Artificial Intelligence, IJCAI 2016, New York, NY, USA, 9-15 July 2016*. Ed. by Subbarao Kambhampati. IJCAI/AAAI Press, 2016, pp. 2965–2971. URL: <http://www.ijcai.org/Abstract/16/421>.
- [16] Ruobing Xie et al. “Image-embodied Knowledge Representation Learning”. In: *CoRR* abs/1609.07028 (2016). arXiv: 1609.07028. URL: <http://arxiv.org/abs/1609.07028>.
- [17] Ruobing Xie et al. “Representation Learning of Knowledge Graphs with Entity Descriptions”. In: *Proceedings of the Thirtieth AAAI Conference on Artificial Intelligence. AAAI’ 16*. Phoenix, Arizona: AAAI Press, 2016, pp. 2659–2665.
- [18] Dongxu Zhang et al. “Joint Semantic Relevance Learning with Text Data and Graph Knowledge”. In: *Proceedings of the 3rd Workshop on Continuous Vector Space Models and their Compositionality*. Beijing, China: Association for Computational Linguistics, July 2015, pp. 32–40. DOI: 10.18653/v1/W15-4004. URL: <https://www.aclweb.org/anthology/W15-4004>.
- [19] Zhao Zhang et al. “Knowledge Graph Embedding with Hierarchical Relation Structure”. In: *Proceedings of the 2018 Conference on Empirical Methods in Natural Language Process-*

ing, Brussels, Belgium, October 31 - November 4, 2018. Ed. by Ellen Riloff et al. Association for Computational Linguistics, 2018, pp. 3198–3207. DOI: 10.18653/v1/d18-1358. URL: <https://doi.org/10.18653/v1/d18-1358>.

- [20] Huaping Zhong et al. “Aligning Knowledge and Text Embeddings by Entity Descriptions”. In: *Proceedings of the 2015 Conference on Empirical Methods in Natural Language Processing, EMNLP 2015, Lisbon, Portugal, September 17-21, 2015*. Ed. by Lluís Màrquez et al. The Association for Computational Linguistics, 2015, pp. 267–272. DOI: 10.18653/v1/d15-1031. URL: <https://doi.org/10.18653/v1/d15-1031>.