

基于组合关系翻译的知识表示学习模型

区恩海

广东工业大学计算机学院, 广东 广州

收稿日期: 2022年2月18日; 录用日期: 2022年3月14日; 发布日期: 2022年3月21日

摘要

目前, 基于表示学习的知识图谱嵌入方法旨在将知识图谱中的实体和关系映射到一个低维的向量空间中, 例如, 翻译模型如TransE及其变体在近年来已显示出可喜的结果, 它们将实体与实体之间的关系表示为实体向量之间的平移操作。这些工作中的大多数都专注于对单一关系进行建模, 因此没有充分利用知识图谱的图结构信息。在本文中, 我们提出了TransE的扩展, 它通过添加相应的翻译向量来对关系的组合进行建模。我们的实验结果表明, 我们的方法可以提高预测单个关系以及它们的组合的性能。

关键词

知识图谱, 表示学习, 翻译模型, 知识嵌入

Knowledge Representation Learning Model Based on Composing Relations with Translation

Enhai Ou

School of Computer, Guangdong University of Technology, Guangzhou Guangdong

Received: Feb. 18th, 2022; accepted: Mar. 14th, 2022; published: Mar. 21st, 2022

Abstract

Currently, knowledge graph embedding methods based on representation learning aim to map entities and relations in knowledge graphs into a low-dimensional vector space. For example, translation model such as TransE and its variants have shown promising results in recent years. In translation model, relations of each entity pair are regarded as translation operations between entity vectors. Most of these works focus on modeling a single relation and thus do not fully exploit the graph structure information of knowledge graphs. In this paper, an extension of TransE is

purposed, which models the composition of relations by adding the corresponding translation vectors. Our experimental results show that our method can improve the performance of predicting single relations as well as compositions of them.

Keywords

Knowledge Graph, Representation Learning, Translation Model, Knowledge Embedding

Copyright © 2022 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

知识图谱是由实体作为节点和关系作为边而组成的有向图。它们以三元组(*head, relation, tail*)的形式存储事实信息，其中 *head* 和 *tail* 分别是头实体和尾实体，*relation* 表示从头到尾的关系。虽然目前的知识图谱由大量的三元组组成，但还远远不够完整，而完整性是决定知识图谱可用性的关键参数。为了补全知识库中大量的缺失信息，对多关系知识数据进行链接预测变得至关重要。本文旨在提供一种高效、可扩展的方法来对知识图进行建模以补全它们。

最近，现有知识图谱的规模在快速扩大，这导致传统的基于逻辑的方法难以处理。在这种背景下，学者们试图建立实体和关系的低维向量或矩阵表示嵌入模型。近年来，各种知识图谱嵌入方法被提出，其中基于翻译的模型简单有效，性能良好。受 word2vec [1] [2] 的启发，给定一个三元组 (h, r, t) ，TransE [3] 学习满足 $h + r \approx t$ 的向量嵌入 h 、 r 和 t 。之后提出的 TransH [4]、TransR/CTransR [5] 和 TransD [6] 等来解决 TransE 建模时无法解决 1-N、N-1 和 N-N 关系的问题。作为双线性模型 RESCAL [7] 的扩展，HolE [8]、DistMult [9] 和 ComplEx [10] 实现了出色的多关系建模能力。同时，也有一些不同的方法利用实体类型[11]、文本描述[12]多种外部信息，以及加强知识图谱表示的逻辑规则[13] [14]。

这些基于嵌入的方法将知识图中的每个对象编码到一个连续的向量空间中，显示出很强的可行性和健壮性。因此，这种方法受到了广泛的关注。尤其，最早出的翻译模型 TransE [3]，已经被证明具有优秀的嵌入能力。该模型将每个关系视为对嵌入实体进行操作的翻译向量。因此，对于三元组(*head, relation, tail*)，模型通过学习头实体和尾实体的向量嵌入以便它们通过关系向量进行翻译连接。目前，学者们已经提出了许多 TransE 的扩展模型来提高其表示能力，同时仍然保持其简单性，例如在翻译过程中增加一些投影操作[5] [6]。

在本文中，我们提出了 TransE 的扩展模型，该扩展模型尝试通过学习关系组合在向量空间中表现出的翻译序列来改进翻译模型对多关系知识图的表示。这个想法是通过学习简单的推理来训练实体的嵌入，例如，给定关系 *people/nationality*，应该给出 *people/city_of_birth* 加上 *city/country*，这样类似的关系组合结果。我们这个称为 TransECR 的模型会通过在知识图中执行受约束的游走来增加此类组合在训练集中的相关示例，并训练这些组合以使翻译序列产生所需的结果。之前 Neelakantan 等人[15]提出了用非线性的递归神经网络(RNN)对多关系数据的组合性进行建模的想法。然而，我们展示了一种通过简单地添加翻译向量而不需要额外参数来表示组合关系的便捷方法，这使得它具有良好的可扩展性。

本文的实验结果表明 TransECR 在链接预测方面优于 TransE。实验评估根据测试实例和训练数据的相似性，把测试实例分为了 hard 和 easy 两个类别，并突出 TransECR 在这两个类别上的改进。我们的实

验还包含了一个回答有关组合关系问题的评估，例如回答这个关系组合($head, relation_1, relation_2, ?$)的 tail 是一个什么实体。TransECR 在这个评估任务上的表现也明显优于 TransE。

2. 相关工作

近年来，多关系数据建模在社交网络和知识库上取得了巨大进展。许多工作将关系学习作为一个多关系表示学习问题来处理，其中有基于不同基线方法的模型在低维潜在空间中编码实体和关系，有基于贝叶斯聚类的模型[16] [17] [18] [19]，有基于矩阵分解的模型[7] [20] [21]。在现有的表示模型中，TransE [3]将关系视为头尾实体之间的翻译进行优化，在预测精度和计算效率之间取得了很好的效果。现有的知识库表示学习方法都只使用实体之间的直接关系，而忽略了组合关系路径中的丰富信息。

关系路径已经在社交网络和推荐系统中得到广泛关注。这些工作中的大多数将每个关系和路径视为离散符号，并使用图算法来处理它们。关系路径也被用于大规模知识图谱的推理，例如路径排名算法(PRA) [22]。PRA 也被用于基于知识图结构的关系提取。Neelakantan 等人[15]通过进一步研究，提出用循环神经网络(RNN)来学习知识图所涉及的关系，从而表示潜在关系路径。我们注意到，这些方法专注于为关系提取建模关系路径，而不考虑实体的任何信息。相比之下，TransECR 把整合了建模实体和关系路径这两方面优点。我们的实验展示了 TransECR 可以学习实体和关系的更丰富的表示，以补全知识图谱和进行组合关系推理。

3. TransECR 知识表示方法

我们提出的模型受到 TransE 的启发。在 TransE 中，知识库的实体和关系被映射到低维空间中，这些被映射到低维空间的实体和关系向量称为嵌入。对于知识库中的每个事实 (h, r, t) ，这些嵌入经过学习后，在向量空间中表示为 $h + r \approx t$ 。因此，对关系进行翻译操作自然会导致将两个关系的组合嵌入为它们的嵌入之和：在知识图谱的某条路径上，由 (h, r, t) 和 (t, r', t') 这两个事实三元组，我们可以在嵌入空间中得出 $h + r + r' \approx t'$ 。原先的 TransE 并没有严格约束实体嵌入能准确再现这样的组合。我们提出的 TransECR，修改了训练阶段的约束条件，使得嵌入能准确再现这样的组合。这样可以允许 TransECR 模型在知识图谱中建模简单的推理。

3.1. 组合关系建模

我们在本节中将全面描述了我们的模型，它允许处理任意数量的关系的组合。首先，作为组合结果的三元组会被表示为 $(h, \{r_i\}_{i=1}^p, t)$ ，其中 p 是从 h 到 t 路径上的组合关系数。这条路径意味着存在实体 e_1, \dots, e_{p+1} ，其中 $e_1 = h$ 且 $e_{p+1} = t$ 使得对于所有 k ， (e_k, r_k, e_{k+1}) 是知识图谱中的一个事实。我们的模型 TransECR 用组合关系表示知识图中路径的每一步 $s_k(h, \{r_i\}_{i=1}^p, t)$ ：

$$s_1(h, \{r_i\}_{i=1}^p, t) = \mathbf{h} \quad (1)$$

$$s_{k+1}(h, \{r_i\}_{i=1}^p, t) = s_k(h, \{r_i\}_{i=1}^p, t) + r_k \quad (2)$$

其中 \mathbf{h} 是头实体 h 嵌入到低维空间中的向量。然后，我们遵循 TransE [3] 的三元组得分计算方法，因此三元组的计算公式如下：

$$d(h, \{r_i\}_{i=1}^p, t) = \|s_p(h, \{r_i\}_{i=1}^p, t) - t\|_2 \quad (3)$$

3.2. 路径的构造和过滤

本文的实验希望通过关系的组合在知识库中学习简单的推理。因此，我们只分析路径长度为 2 的关

系组合，下面会详细介绍。

首先，对于每个事实 (h, r, t) ，检索所有路径 $(h, \{r_1, r_2\}, t)$ 是否有使 (h, r_1, e) 和 (e, r_2, t) 都在知识图中的 e ，然后我们过滤后，将得到这些使得 $(h, r_1, e) = (h, r, t)$ 或者 $(e, r_2, t) = (h, r, t)$ ，以及 $r_1 = r_2$ 和 $h = e = t$ 的路径。我们关注明确的路径，以便使得推理真正有意义。特别是，我们只考虑由 r_1 是一对一或一对多关系，而 r_2 是一对一或多对一关系。在我们的实验中，为训练模型所创建的路径只是用于训练一部分事实三元组。另外，在本文的其余部分，这种含有路径长度为 2 的事实称为四元组。

3.3. 模型训练

我们的训练目标分为两部分：第一部分是用基于间隔的损失函数训练模型，训练数据忽略了四元组。第二部分通过增加正则项的方式，将路径加入到训练中。用 S 表示知识图中的一组事实，训练目标的第一部分是用以下基于间隔的损失函数训练三元组的嵌入

$$\sum_{\substack{(h, r, t) \in S \\ (h', r', t') \in S_{(h, r, t)}}} [\gamma + d(h, r, t) - d(h', r', t')]_+ \quad (4)$$

其中 $[x]_+ = \max(x, 0)$ 是取 x 为正数的时候， γ 是间隔大小， $S_{(h, r, t)}$ 是通过替换事实三元组 (h, r, t) 中的 h 或者 t 而创建的负例三元组集合。

这个损失函数能有效地训练了尾实体，以使得它的嵌入是最接近翻译后的头实体嵌入，但是它并不能保证尾实体嵌入和翻译后的头实体嵌入之间的达到极小的距离。根据最近邻准则足够对简单的三元组进行推断，但要确保组合关系规则准确，必须确保它们距离很小。考虑到关系的组合性，我们添加了两个额外的正则项：

$$\lambda \sum_{(h, r, t) \in S} d(h, r, t)^2 \quad (5)$$

$$\alpha \sum_{(h, \{r_1, r_2\}, t) \in S} N_{r \rightarrow \{r_1, r_2\}} d(h, \{r_1, r_2\}, t)^2 \quad (6)$$

公式(4)是基于间隔的损失函数，适用于训练原始数据集中的事实三元组。而公式(5)(6)考虑通过添加的约束条件，使得模型能用于训练四元组。公式(6)中 $r \rightarrow \{r_1, r_2\}$ 意思是组合关系 r 蕴含 r_1, r_2 这两个关系。 $N_{r \rightarrow \{r_1, r_2\}}$ 是一个涉及到 r_1, r_2 和由它们组合而成的关系 r 的参数。它是一个包含 $\{r_1, r_2\}$ 这两个关系的路径数，然后由关系 r 的事实创建的四元组总数进行归一化而来的参数。这个约束条件能赋予可替代关系 r 的可靠路径 $\{r_1, r_2\}$ 更高权重。最后，为每个实体的嵌入添加正则项 $\mu \|e\|_2^2$ 。

4. 实验结果与分析

本节阐述了在基准数据集 FB15K [3] 和 Family [23] 的实验。

4.1. 数据集分析

FB15K 是 Freebase 的一个子集，它是一个非常大的通用事实数据库，收集了超过 12 亿个三元组和 8000 万个实体。Family 是另外一个数据集，其中包含表示 6 代的 5 个家庭的 7 种关系(cousin of, has ancestor, married to, parent of, related to, sibling of, uncle of)的三元组。这个是人工数据集，每个家庭都组织在一个分层的树结构中，其中每一层都代表一代。家庭之间通过两个成员之间的婚姻联系连接起来，这些成员是从不同家庭的同一层中随机抽取的。有趣的是，在这个数据集上，有明显的组合关系，比如 uncle of \approx sibling of + parent of 或者 parent of \approx married to + parent of 等组合关系。表 1 给出了数据集的相关信息。

Table 1. Statistics of the datasets used**表 1.** 数据集的统计情况

数据集	Family	FB15K
实体数	721	14,951
关系数量	7	1345
训练三元组数量	8461	4,843,142
训练四元组数量	0	30,252
验证集三元组数量	2820	50,000
测试集三元组数量	2821	59,071
测试集四元组数量	0	1852

4.2. 实验参数设置

我们的实验参数主要参考的是 TransE，因此我们遵循与 TransE 中相同的实验设置，使用排名指标进行评估。对于每个测试三元组，我们依次用每个实体替换头部，然后计算每个候选者的分数并对它们进行排序。由于其他正候选实体(即形成事实三元组的实体)的排名可能高于目标，因此我们从排名中过滤掉训练集、验证集和测试集中存在的所有正候选实体，然后保留目标实体的排名。重复相同的过程，但去除尾部实体而不是头部。1) Mean rank 是这些实体排名的平均值，该指标的值越小，说明模型的性能越好。2) Hits@10 是前 10 个预测中有目标实体的命中率，该指标的值越大，说明模型的表示学习的能力越强，实体嵌入越准确。

模型调试过程中，如何获取最优的参数是非常重要的。对于 Family 数据集，嵌入维度设置为 20，FB15K 的嵌入维度设置为 100。训练采用随机梯度下降进行，在 500 个 epoch 之后停止。在 FB15K 上，我们使用 TransE 的学习好的嵌入来初始化 TransECR，我们将学习率 α 设置为 0.001 来微调 TransECR。在 Family 上，两种算法都是随机初始化实体嵌入的，并使用 0.01 的学习率。平均排名 Meanrank 被用作检验模型的标准。 γ, λ, α 和 μ 分别在如下参数集合中选取， $\gamma \in \{0.25, 0.5, 1\}$ ， $\lambda \in \{1e^{-4}, 1e^{-5}, 0\}$ ， $\alpha \in \{0.1, 0.05, 0.01, 0.001, 0.005\}$ 和 $\mu \in \{1e^{-4}, 1e^{-5}, 0\}$ 。

4.3. 实验结果分析

为了验证本文所提出的模型综合性能，将选择 TransE 和我们提出的模型 TransECR 在链接预测任务上进行参与评估。实验结果展示在表 2。其中 w. COMP. 表示在训练集中存在与其关系匹配的组合关系。其中 w/o COMP. 表示在训练集中不存在与其关系匹配的组合关系。

在 FB15K 数据集上，表 2(最后一行)显示，在较长的路径(此处长度为 2)上进行训练始终可以提高性能，同时仅预测三元组的头或尾：TransECR 的整体 Hits@10 从 TransE 的 71.5% 提高近 5% 到 76.2%。

为了更好地理解 TransECR 的优势，我们对 FB15K 进行了详细的评估，通过沿两个维度对测试三元组进行分类：Easy 与 Hard 以及有组合关系与无组合关系。一个测试三元组 (h, r, t) ，如果它的头实体和尾实体通过训练集中的三元组连接起来，那么这个三元组被标记为 Easy。否则，这个三元组被标记为 Hard。此外，如果对于 r 至少存在一条路径 $\{r_1, r_2\}$ ，则测试三元组 (h, r, t) 被标记为组合关系。如果不存在这样的路径，则 (h, r, t) 没有这样的组合关系。

详细结果如表 2 所示。我们可以看到，与 TransE 相比，TransECR 在 Hits@10 方面特别提高了模型对组合关系三元组的性能，在 Easy 三元组上提高了 4.2% (从 78.8% 到 83.0%) 和 Hard 三元组上提高了 2.5%

(从 46.8% 到 9.3%)。性能提高最多的仍然是对 Easy 三元组, 实际上没有组合关系的 Easy 三元组的 Hits@10 增加了 4%, 从 71.3% 增加到 75.3%。平均排名在 Easy 的三元组上也有显著提高, 而在 Hard 三元组上则保持不变。总而言之, 结果表明, 在训练期间考虑路径可以显著提高性能, 并且组合关系三元组的结果表明 TransECR 确实能够推理出存在于较长路径中的事实。

Table 2. Detailed performances on FB15k of TransE and TransECR (%)
表 2. TransE 和 TransECR 在 FB15k 上的详细性能

模型	TransE		TransECR	
	MR	Hits@10	MR	Hits@10
Easy	17.7	76.8	12.5	82.2
Hard	191.0	48.9	205.7	51.0
Easyw. COMP.	16.4	78.8	11.6	83.0
Easyw/o COMP.	21.6	71.3	16.0	75.3
Hard w. COMP.	208.1	46.8	212.2	49.3
Hard w/o COMP.	122.9	57.0	123.8	57.5
Overall	50.7	71.5	49.5	76.2

4.4. 四元组预测任务结果分析

在知识图中进行链接预测侧重于预测测试三元组中的缺失元素。为了更全面地对模型进行评估, 我们在此将扩展到更复杂的任务, 例如预测四元组 $(h, \{r_1, r_2\}, ?, ?)$ 或 $(?, \{r_1, r_2\}, t)$ 中的尾实体或者头实体。

表 3 给出了 TransE 和我们的 TransECR 模型对此类四元组的预测示例。TransECR 能够从这些信息中正确推断出国家或者地区, 因为 TransECR 具有组合关系的推断能力, 我们在建模的时候就考虑对组合关系的学习。另一方面, TransE 受到级联误差的影响, 因为排名损失函数并不能保证 $h + r_1$ 和实体 “Malaysia” 之间的距离很小, 所以当对关系 r_2 求和时, 它最终会更接近实体 “Taiwan” 而不是实体 “asia”。

Table 3. Examples of quadruples predictions of TransECR and TransE

表 3. TransECR 和 TransE 的四元组预测示例

3 个最邻近 $h + r_1 + r_2$ 的实体		
	TransECR	TransE
$h: malay$ $r_1: language/main country$ $r_2: continent$	southeast asia Malaysia asia	taiwan southeast asia philippines
$h: university of victoria$ $r_1: institution/campuses$ $r_2: location/citytown$	victoria kurniaby kelownat	kelownar toronato ottawat
$h: stargate atlantis$ $r_1: regular TV appearance$ $r_2: nationality$	Hawaii Scotland U.S.A.	Scotland Hawaii U.K.

为了进行更定量的分析, 我们在 FB15K 上生成了一个新的四元组预测测试数据集。这个测试集是通

过从三元组测试集中生成路径并删除那些用于训练的四元组来创建的。我们获得了 1852 个四元组。整体实验方案和前面一样，依次尝试预测这些四元组的头部或尾部。

在该评估协议中，TransECR 的 MR 为 113.0, Hits@10 为 69.4%，而 TransE 的 MR 为 158.9, Hits@10 为 65.4%。我们可以看到，递推关系的学习提高了两个指标的性能，在 Hits@10 方面获得了 3% 的提升，在平均排名 MR 上提高大约 46，相对提升了接近 30%。

5. 结论

本文提出了 TransECR，一种新的知识图谱表示学习方法，它对组合关系的路径进行编码，以将实体和关系都嵌入到低维空间中。为了利用组合关系路径，我们提出在训练中加入了关于路径的正则项来衡量关系路径的可靠性，并使用关系组合来表示路径以进行优化。我们评估 TransECR 在三元组和四元组链接预测这两个任务上的性能。实验结果表明，与 TransE 相比，TransECR 实现了显着的改进。

未来，我们将探索以下研究方向：1) 本文的模型仅通过组合关系之间的路径关联推理模式进行学习。关系之间有很多复杂的模式。例如，由“公主”可以推理出“女性”，但是目前的模型无法作出这样的推理。我们可以利用一阶逻辑对这些推理模式进行编码以进行表示学习。2) TransE 有一些扩展，例如 TransH 和 TransR。他们采用 TransECR 的思想并非易事，我们将探索将 TransECR 扩展到这些模型，以更好地处理知识图谱表示学习中的复杂场景。

参考文献

- [1] Mikolov, T., Sutskever, I., Chen, K., Corrado, G.S. and Dean, J. (2013) Distributed Representations of Words and Phrases and Their Compositionality. In: Burges, C., Bottou, L., Welling, M., Ghahramani, Z. and Weinberger, K., Eds., *Advances in Neural Information Processing Systems*, Curran Associates, Inc., Red Hook, 3111-3119. <https://proceedings.neurips.cc/paper/2013/file/9aa42b31882ec039965f3c4923ce901b-Paper.pdf>
- [2] Mikolov, T., Chen, K., Corrado, G. and Dean, J. (2013) Efficient Estimation of Word Representations in Vector Space. <https://arxiv.org/abs/1301.3781>
- [3] Yang, B., Yih, W.T., He, X., Gao, J. and Deng, L. (2014) Embedding Entities and Relations for Learning and Inference in Knowledge Bases. <https://arxiv.org/abs/1412.6575>
- [4] Wang, Z., Zhang, J., Feng, J. and Chen, Z. (2014) Knowledge Graph Embedding by Translating on Hyperplanes. *Proceedings of the AAAI Conference on Artificial Intelligence*, Québec, 27-31 July 2014, 1112-1119.
- [5] Lin, Y., Liu, Z., Sun, M., Liu, Y. and Zhu, X. (2015) Learning Entity and Relation Embeddings for Knowledge Graph Completion. *Twenty-Ninth AAAI Conference on Artificial Intelligence*, Austin, 25-30 January 2015, 2181-2187.
- [6] Ji, G., He, S., Xu, L., Liu, K. and Zhao, J. (2015) Knowledge Graph Embedding via Dynamic Mapping Matrix. *Proceedings of the 53rd Annual Meeting of the Association for Computational Linguistics and the 7th International Joint Conference on Natural Language Processing*, Beijing, 26-31 July 2015, 687-696.
- [7] Nickel, M., Tresp, V. and Kriegel, H.P. (2011) A Three-Way Model for Collective Learning on Multi-Relational Data. *Proceedings of the 28th International Conference on International Conference on Machine Learning*, Bellevue, 28 June-2 July 2011, 809-816.
- [8] Nickel, M., Rosasco, L. and Poggio, T. (2016) Holographic Embeddings of Knowledge Graphs. *Proceedings of the Thirtieth AAAI Conference on Artificial Intelligence*, Phoenix Arizona, 12-17 February 2016, 1955-1961.
- [9] Yang, B., Yih, W.T., He, X., Gao, J. and Deng, L. (2014) Embedding Entities and Relations for Learning and Inference in Knowledge Bases. arXiv:1412.6575. <https://arxiv.org/abs/1412.6575>
- [10] Trouillon, T., Welbl, J., Riedel, S., Gaussier, É. and Bouchard, G. (2016) Complex Embeddings for Simple Link Prediction. *Proceedings of the 33rd International Conference on Machine Learning*, New York, 19-24 June 2016, 2071-2080. <https://arxiv.org/pdf/1606.06357.pdf>
- [11] Lv, X., Hou, L., Li, J. and Liu, Z. (2018) Differentiating Concepts and Instances for Knowledge Graph Embedding. arXiv:1811.04588.
- [12] Dai, S., Liang, Y., Liu, S., Wang, Y., Shao, W., Lin, X. and Feng, X. (2018) Learning Entity and Relation Embeddings with Entity Description for Knowledge Graph Completion. *Proceedings of the 2nd International Conference on Artificial Intelligence: Technologies and Applications (ICAITA)*, Chengdu, 25-26 March 2018, 202-205.

-
- [13] Rocktäschel, T., Singh, S. and Riedel, S. (2015) Injecting Logical Background Knowledge into Embeddings for Relation Extraction. *Proceedings of the 2015 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, Denver, 31 May-5 June 2015, 1119-1129.
 - [14] Wang, Q., Wang, B. and Guo, L. (2015) Knowledge Base Completion Using Embeddings and Rules. *Twenty-Fourth International Joint Conference on Artificial Intelligence*, Buenos Aires, 25-31 July 2015, 1859-1865.
 - [15] Neelakantan, A., Roth, B. and McCallum, A. (2015) Compositional Vector Space Models for Knowledge Base Completion. arXiv:1504.06662.
 - [16] Kemp, C., Tenenbaum, J.B., Griffiths, T.L., Yamada, T. and Ueda, N. (2006) Learning Systems of Concepts with an Infinite Relational Model. *Proceedings of the Twenty-First National Conference on Artificial Intelligence*, Boston, 16–20 July 2006, 381-388.
 - [17] Miller, K., Jordan, M. and Griffiths, T. (2009) Nonparametric Latent Feature Models for Link Prediction. *Advances in Neural Information Processing Systems*, **22**, 1276-1284.
 - [18] Sutskever, I., Tenenbaum, J. and Salakhutdinov, R.R. (2009) Modelling Relational Data Using Bayesian Clustered Tensor Factorization. *Advances in Neural Information Processing Systems*, **22**, 1821-1828.
 - [19] Zhu, J., Song, J. and Chen, B. (2016) Max-Margin Nonparametric Latent Feature Models for Link Prediction. arXiv:1602.07428.
 - [20] Singh, A.P. and Gordon, G.J. (2008) Relational Learning via Collective Matrix Factorization. *Proceedings of the 14th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, Las Vegas, 24-27 August 2008, 650-658. <https://doi.org/10.1145/1401890.1401969>
 - [21] Nickel, M., Tresp, V. and Kriegel, H.P. (2012) Factorizing YAGO: Scalable Machine Learning for Linked Data. *Proceedings of the 21st International Conference on World Wide Web*, Lyon, 16-20 April 2012, 271-280. <https://doi.org/10.1145/2187836.2187874>
 - [22] Lao, N. and Cohen, W.W. (2010) Relational Retrieval Using a Combination of Path-Constrained Random Walks. *Machine Learning*, **81**, 53-67. <https://doi.org/10.1007/s10994-010-5205-8>
 - [23] García-Durán, A., Bordes, A. and Usunier, N. (2014) Effective Blending of Two and Three-Way Interactions for Modeling Multi-Relational Data. In: Calders, T., Esposito, F., Hüllermeier, E. and Meo R., Eds., *Machine Learning and Knowledge Discovery in Databases*, Springer, Berlin, 434-449. https://doi.org/10.1007/978-3-662-44848-9_28