

学界 | UCSB提出变分知识图谱推理：在KG中引入变分推理框架

2018-03-28 机器之心

选自arXiv

作者：Wenhu Chen等

机器之心编译

参与：张楚、思源

推理知识图谱中缺失的连接已经吸引了研究界的广泛关注。在本论文中，加州大学圣塔芭芭拉分校的王威廉等研究者在知识图谱推理中引入了变分推理框架，并将路径搜索和路径推理紧密结合从而进行联合推理，这种方法提升了知识图谱推理模型的稳定性。

自动推理（Automated reasoning）作为计算系统根据观察到的证据做出新推论的一种能力，已经引起了很多研究团体的关注。近年来，人们对为复杂推理任务设计机器学习算法的兴趣一浪高过一浪，尤其在大型知识图谱（KGs）方面，数不清的实体和连接让传统基于逻辑的算法面临巨大挑战。具体来说，我们将研究定位于这种大型知识图谱多跳（multi-hop）推理的情境，目标是设计一个自动推理模型，以完善大型知识图谱中现有实体间缺失的连接。例如，若知识图谱中包含「总统」（贝拉克·奥巴马，美国）及其配偶（米歇尔，贝拉克·奥巴马）这两个实体，我们要让机器来自动完善隐藏其中的连接（米歇尔，美国）。实现该任务的系统在解答复杂问题的应用中必不可少。

为了处理多跳连接推测的问题，人们提出过各种各样的方法。一些诸如 PRA（Path Ranking Algorithm，路径排序算法）（Lao et al., 2011; Gardner et al., 2014, 2013）的早期研究使用可重复启动的有界深度（bounded-depth）随机游走来获取路径。最近 DeepPath (Xiong et al., 2017) 和 MINERVA (Das et al., 2017) 将路径搜索问题设置为一个马尔可夫决策过程（MDP）并利用强化学习（RL）将预期的返回值最大化。与我们的工作同时进行的另一项工作是「推理链」（Chain-of-Reasoning, Das et al., 2016）和复合推理（Compositional Reasoning, Neelakantan et al., 2015）的研究，即以 PRA 学习到的多跳链为输入来推断其关系。

这里我们将 KG 推理任务设置为两个子步骤，即「路径搜索」和「路径推理」。我们发现大多数相关研究只专注于其中一步，且其主要缺陷在于忽视了两个步骤间的交互。更确切地说，DeepPath (Xiong et al., 2017) 和 MINERVA (Das et al., 2017) 可以被理解为「路径搜索」步骤的优化，而复合推理 (Neelakantan et al., 2015) 和推理链 (Das et al., 2016) 可被认为是「路径推理」这一步的优化。DeepPath 经训练以用来使两个给定实体间的路径搜索更有效率，但无法得知实体对之间的联系存在与否；相对地，给定一对查询

实体，MINERVA 会学习到目标节点，但又无法得知搜索路径的质量。相比来说，推理链和复合推理仅仅是推断给定关系的路径，但无法得知路径搜索程序的过程。缺乏交互使得模型无法理解更多样化的输入，而且对噪声和对抗样本非常敏感。

为了提高现有 KG 推理模型的稳定性并处理更多有噪声的环境，我们打算从潜变量图模型的视角，将两个步骤结合为一个整体。这个图模型在给定实体对的条件下将路径视为离散的潜变量，将关系看作显变量，这样路径搜寻模块可以被视为使用先验分布来推断 KG 中的底层连接。与之相对，路径推理模块可被视为把底层连接划分为不同类别的似然分布。在这一假设下，我们引入近似后验，并设计了一个变分自编码 (Kingma and Welling, 2013) 算法以最大化下界。这个变分架构将两个模块紧密结合为统一的整体，并对其同时进行训练。通过积极合作和交互，路径搜索可以考虑到搜索到路径的价值并使用更有意义的路径。与此同时，路径推理模块会收到路径搜索模块传来的多样性路径，从而更好地归纳未知的情境。我们的贡献在于以下三点：

- 在 KG 推理中引入了变分推理框架，将路径搜索和路径推理紧密结合从而进行联合推理。
- 成功地在训练中加入反面样本，同时提高了现有 KG 推理模型的稳定性。
- 本文的模型可以扩展到大型 KG，并在两项任务中得到最高水平的结果。

论文的其它部分结构如下：在第 2 部分我们将概述 KG 嵌入、多跳推理以及变分自编码的相关研究；在第 3 部分描述我们的变分知识推理工具 DIVA；第 4 部分展示了试验结果；第 5 部分为结论。

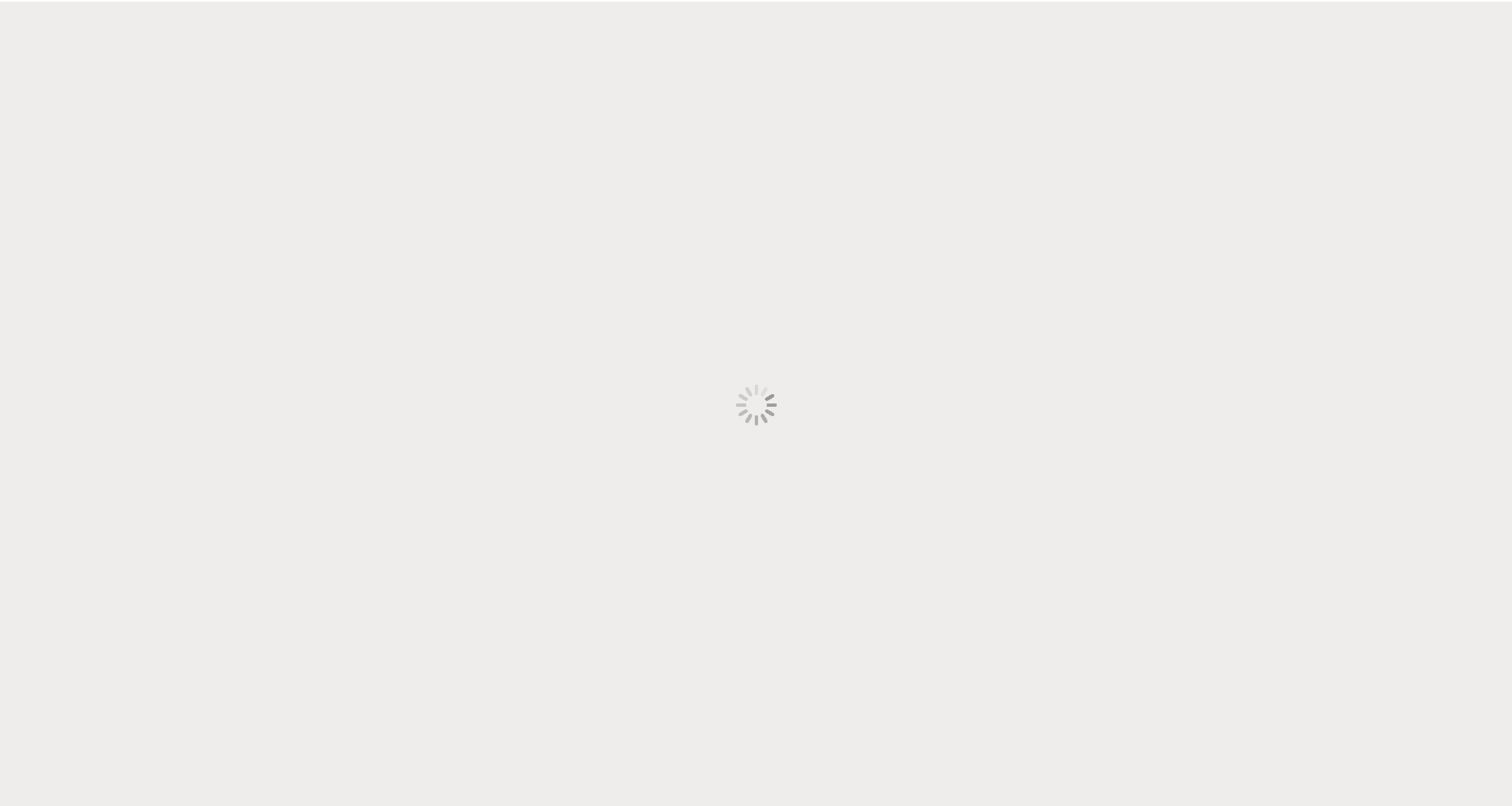


图 1：本文所提出的方法的概率图模型。虚线框内的箭头代表整个连接空间中以多项分布建模的近似后验。实线框内的箭头代表先验和似然分布。

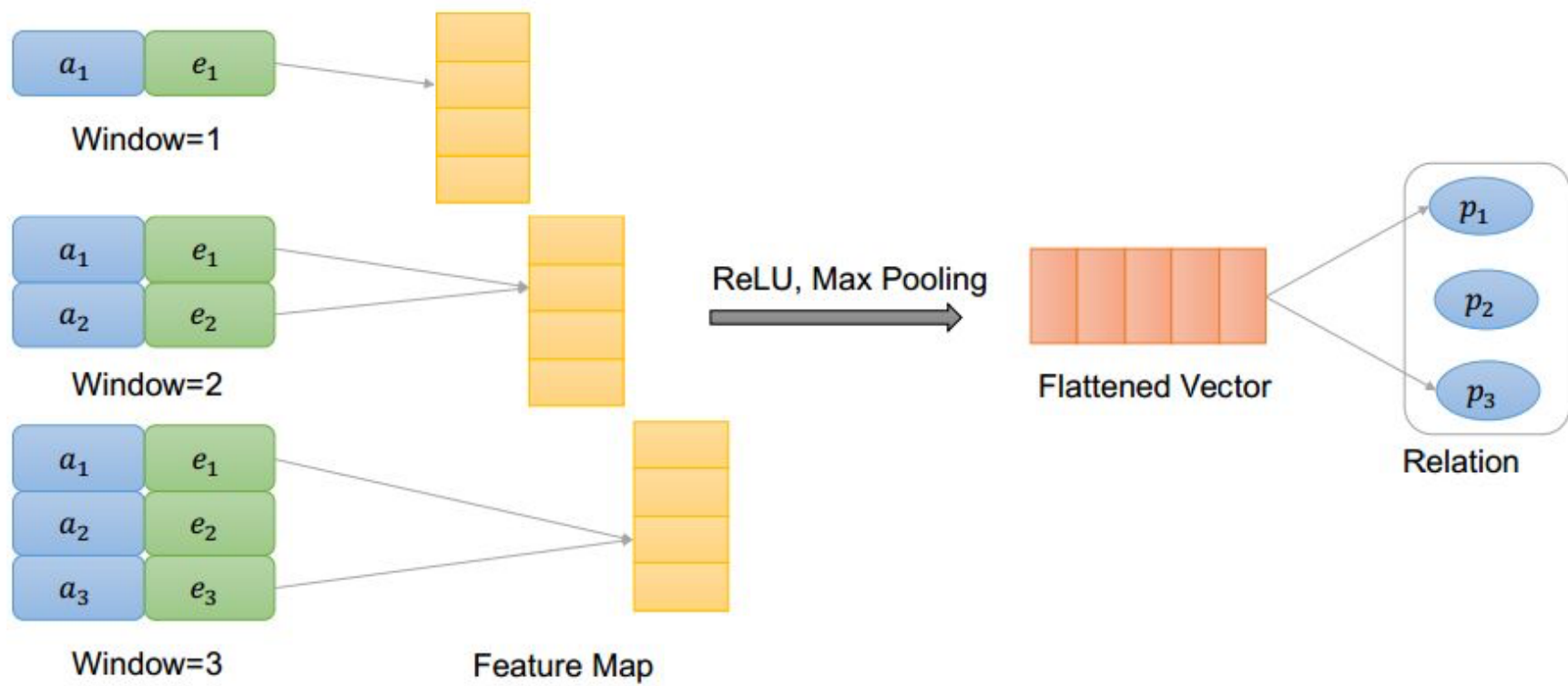


图 2：CNN 路径推理机概览。

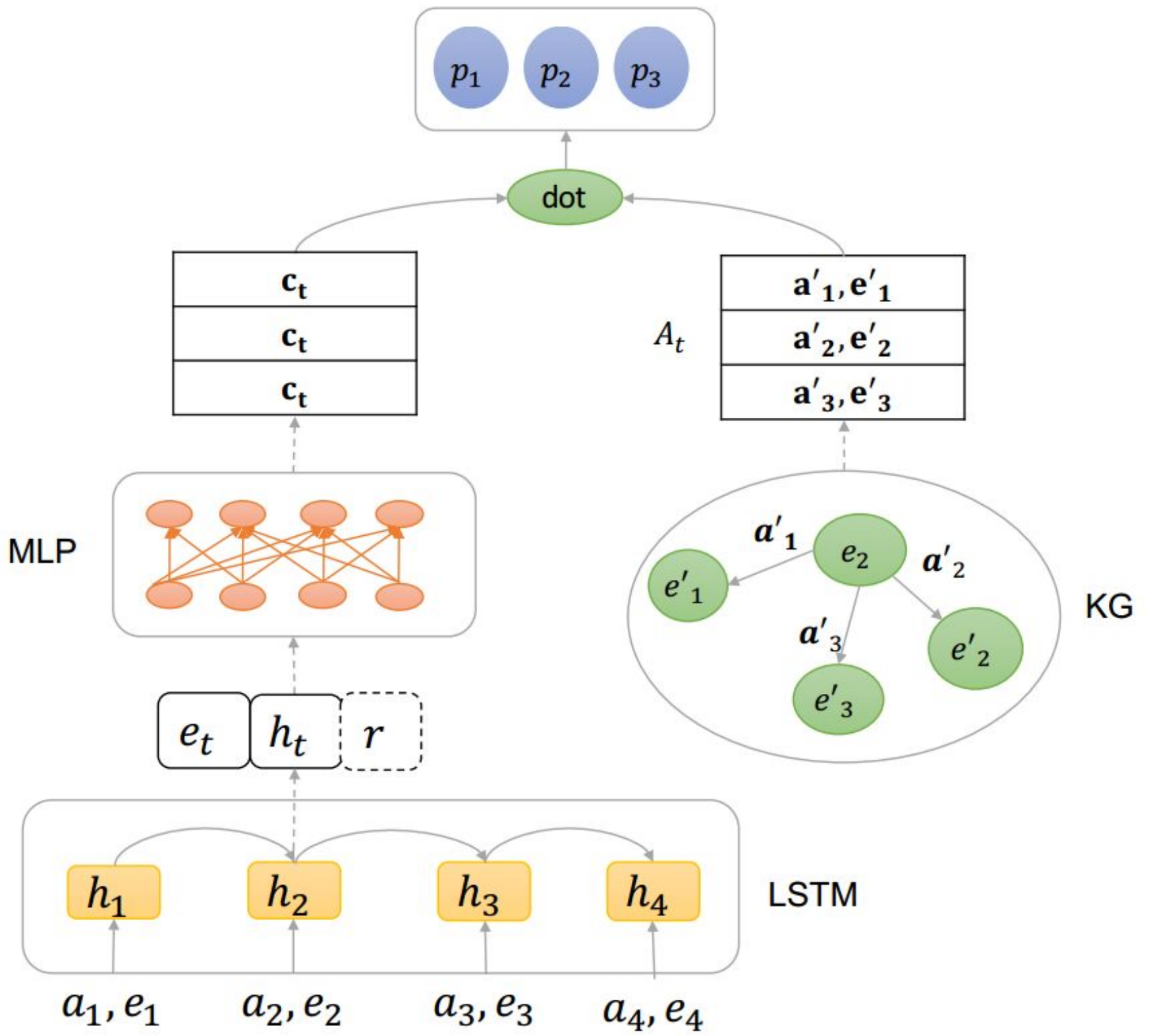


图 3：路径搜索模型概览。注意，在近似后验中存在 r_q 而在路径搜索模型中不存在 r_q



表 2：基于 NELL 数据集的 MAP 结果。鉴于 MINERVA (Das et al., 2017) 只取用了 12 个初始关系中的 9 个，我们在此展示了 NELL-995 数据集两个版本的结果。



表 3：基于 FB15k 数据集的结果，请注意，MINERVA 的结果是由我们实现获取的。



图 5: NELL 和 FB15k 连接推理任务的误差分析。由于 FB15k 数据集将占位符也作为实体，我们无法分析出误差是否来源于 KG 的噪声。

论文：变分知识图谱推理



论文链接：<https://arxiv.org/abs/1803.06581>

推理知识图谱中缺失的连接已经吸引了研究界的广泛关注。在本文中，我们处理了一类包含推理给定实体对间关系的实际查询任务。我们将这类问题设计为一个概率图模型下的推理问题并试图从变分推理的视角解决它。为了建立查询实体对的关系模型，我们假设 KG 中存在潜变量（所有连接这两个节点的路径集合），即它们之间的联系。但由于大型 KG 内的关联问题难以解决，我们提出用变分推理来使 ELBO 最大

化。更确切地说，我们的框架（DIVA）由三个模块组成——后验近似、先验（路径搜索）以及似然估计（路径推理）。通过变分推理方法，我们成功将三者紧密结合为一个统一的架构，同时对其联合优化以实现 KG 推理。随着伴随子模块间的积极交互，DIVA 可以更好地处理噪声并应付更复杂的推理情境。为评估我们的方法，我们基于 NELL-995 和 FB15k 数据集执行了连接推理任务试验，而在两个数据集下的表现都达到了很高的水准。

本文为机器之心编译，转载请联系本公众号获得授权。



加入机器之心（全职记者/实习生）：hr@jiqizhixin.com

投稿或寻求报道：editor@jiqizhixin.com

广告&商务合作：bd@jiqizhixin.com