

论文阅读报告

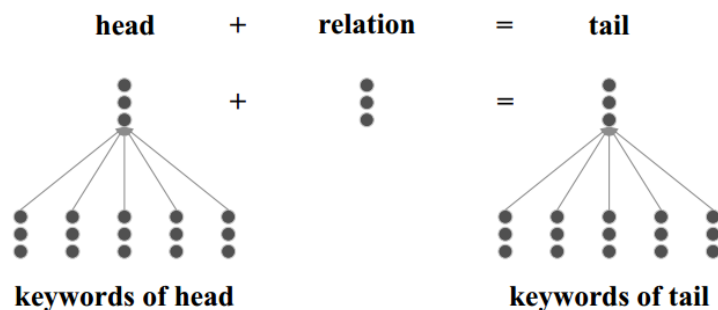
黄焱晖

2019 年 3 月 6 日

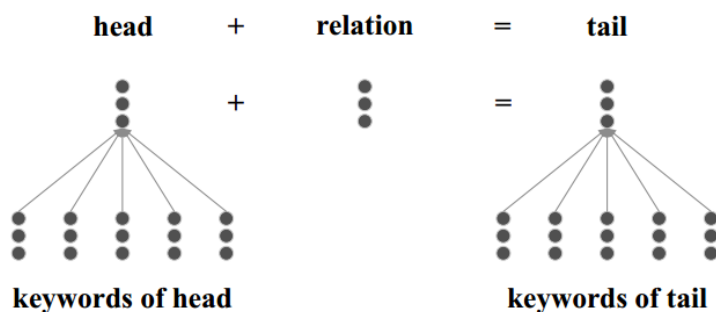
0.1 阅读论文: Representation Learning of Knowledge Graphs with Entity Descriptions

动机: 实体描述中含有丰富的语义信息, 以往的知识表示只是用到了图的结构, 而没有用到实体描述中的语义信息

解决思路: 作者的思路比较简单, 将图的结构信息和实体的语义信息相加, 获得实体表示。其中图的结构信息只是简单利用TransE模型, 而对于实体的语义信息, 作者利用了两种方式去得到实体表示, 一种是简单的词袋模型 (CBOW), 另一种是CNN的卷积模型。



上图是词袋模型示意图



上图是CNN卷积网络示意图

在实体表示中，实体表示的势能函数由三部分组成，第一部分为 $h_d + r - t_d$ ，第二部分为 $h_d + r - t_s$ ，第三部分为 $h_s + r - t_d$ ， h_s 是结构信息表示的头结点， h_d 是语义信息表示的头结点， t_s 是结构信息表示的尾结点， t_d 是语义信息表示的头结点，作者这样定义势能函数的原因是将两种类型的实体表示能够放在同一向量空间下训练(也包括关系向量)。

不足：对于实体描述中的词语，作者是通过Wikipedia训练好的词向量来进行语义表示，这里面可以再深化一层，除了两种类型的实体表示放在同一向量空间中，对于词的训练是不是也可以通过Line这种模型，将词向量的训练和实体表示的训练放在同一向量空间中呢。