# 摘要

知识图谱嵌入是将符号关系和实体投影到低维连续空间中的一种方法，是知识图谱补全的关键。近年来，基于翻译的嵌入模型（如TransE）以其简单、有效的特点越来越受到人们的关注。这些模型试图通过关系将语义从头实体转换为尾部实体，并在知识图之外推断出更丰富的事实。本文提出了一种新的知识图谱嵌入方法TransMS，它可以翻译和传输多方向语义：

1）头/尾实体的语义及其与具有非线性函数的尾/头实体的关系

2）从实体到线性偏量向量关系的语义。

对于每一个三元组，我们的模型只比TransE多了一个参数α，这使得它在大规模的知识图谱上具有更好的可扩展性。实验表明，TransMS相对于最先进的baselines（不知道怎么翻译）实现了实质性的改进，尤其是在FB15K数据集上，对于N-1关系的头实体预测和1-N关系的尾实体预测的Hit@10s分别提高了27.1%和24.8%。

# 1 简介

知识图谱是一种由三元组组成的语义图，在信息检索中有着广泛的应用，已成为人们研究的一个热点。作为知识图谱的基本单位，以（头实体，关系，尾实体）为形式的三元组，也可缩写为（h，r，t），由两个实体（头实体和尾实体）作为节点，由一个关系（关系）作为连接头和尾的边。尽管知识图谱可以由数百万个实体和数十亿个关系组成，但它们还远远不够完整，因此，根据知识图中已有的三元组来预测实体间的新关系，尤其是复杂关系，是知识补全过程中不可缺少的任务。近年来，知识图谱嵌入被认为是一种有效的知识图谱补全方法。特别是具有开创性的TransE模型因其简单有效而引起了越来越多的基于翻译模型的研究。TransE模型首先将实体和关系投影到低维连续空间中，以h,r,t作为形式。然后用关系将语义从头部映射到尾部，当三元组（h；r；t）成立时，需要h+r≈t。然而，这种约束过于严格，在处理1-N、N-1和N-N之间的复杂关系时是不现实的，为了克服transE模型的缺点，我们设计了基于翻译的后续方法。

具体而言，TransH将每个关系解释为一个具有平移向量和范数向量的超平面上的平移操作，并通过范数向量将实体向量投影到超平面上。然而，TransR模型认为不同的关系关注于实体的不同方面，这些实体具有多重性方面。TransR首先将实体和关系分别投影到实体空间和关系空间，然后通过关系投影矩阵将实体映射到关系空间。TransD方法是基于TransR的投影矩阵不仅与关系相关，而且与实体相关的假设。而TranSparse关注的是异构性，即一些关系只连接几个实体对，而另一些不连接；而不平衡性则意味着关系可能连接多个头/尾实体和较少的尾/头实体，因此TranSparse设置了与投影矩阵相关的实体数目的参数。Gtrans认为Trans（E，H，R，D）没有充分考虑实体的复杂性，而是用两个向量分别表示实体和关系的不同语义方面。

问题：虽然这些方法与以前的方法相比有了很大的改进，但是在处理复杂关系的链路预测方面仍然存在困难，我们将其归因于以下原因：

1. 实体之间的语义转换不是通过实体的头/尾传递，而是通过实体间的语义传递来实现的。以图1为例，有三个正确的三元组（老虎；吃；肉），（人；吃；肉）和（人；吃；水果）。这些方法只是把吃的语义传递给人、虎、果、肉，而从人类认知的角度来看，吃这个关系只在人与果、人与肉、虎与肉之间存在语义关系，这可能导致错误的推论：（老虎；吃；水果）。
2. 它们也没有通过将语义从实体传递到关系来转换关系嵌入。在一些研究中，三元组（h；r；t）可以看作（主语、谓语、宾语）的语法结构，即关系是每个三元组都有一个执行者h和一个被执行者t的谓语，这意味着主语（h）和宾语（t）到谓语（r）之间存在语义传递
3. 由于参数比TransE多得多，它们对大规模知识图的可扩展性较差。对于每个三元组，它们通常比TranE有一些额外的向量或矩阵，例如TransR对于每个关系都有一个额外的投影矩阵。
4. 它们通过线性变换来传递语义，这种转换可能受到限制，为了更好地进行语义翻译，必须用非线性变换来代替。
5. 从数学的角度看，当头/尾实体向量和关系向量固定时，h+r=t的左/右边也是固定的，即无论如何线性变换得到尾/头实体向量，最终的尾/头实体向量都是围绕一个中心分布的，因此当头/尾实体向量和关系向量固定时，尾/头实体向量彼此接近。

为了解决这些问题，我们提出了一个名为TransMS的新方法，对于每个三元组，它只比TransE有一个额外的参数α，如图2所示。

首先，头尾实体被投影为Ke维的向量，关系被投影为Kr维的向量和一维变量α，其中对于我们模型中的每一个三元组，Ke=Kr。然后，我们通过非线性函数p（h，r）将头部实体h和关系r的语义传递到尾部实体t来转换尾部实体嵌入，同时通过p（−t，r）将尾部实体t和关系r的语义传递到头部实体h来转换头部实体嵌入，因为它与关系的方向相反。由于关系到实体的非线性传输，我们只需在关系嵌入r中加入一个偏差向量，以防止非线性传输相互抵消，其中偏差向量是与h和t相关的α\*g（h，t）。因此，该模型不仅在语义上符合人类的认知，而且克服了以往模型中一个实体的向量学习是围绕当另一个实体和三元组之间的关系是固定的中心学习的。

贡献：本论文大体上的亮点如下

1. 我们提出了一个新的模型TransMS，它不仅考虑了从关系到实体的语义信息，而且考虑了从实体到关系和实体之间的语义信息
2. 我们的模型比最简单的TransE模型只增加了一个参数α，而其他模型对每个三元组都有更多的参数，例如TransR的参数比TransE多Ke X Kr个参数，因此我们的模型比其他模型具有更好的可扩展性。
3. 该模型对复杂关系的链路预测有显著的改进，特别是在基线模型中，该模型对复杂关系的链路预测有显著的改进，对于N-1关系的头实体预测和1-N关系的尾实体预测的Hit@10在FB15k数据集中分别提高了27.1%和24.8%。

# 3 提出的方法

## 3.1 动机

如引言所述，三元组的头实体可以解释为主语，关系可以解释为谓语，尾实体可以解释为宾语。因此，我们可以很容易地得到头实体、尾实体和关系与每个三元组的关系，就像主语、宾语和谓语对主谓句的关系一样。直观地说，对于每个三元组（h，r，t），不仅存在关系r到实体h（或t）的语义信息，而且h和t之间也存在语义信息。以往的方法主要考虑前一种语义信息，导致其对复杂关系的预测能力较差。为了克服这一缺点，我们的方法将实体h（或t）和关系r的语义信息传递给实体t（或h）。另外，我们认为非线性函数比线性函数更有效地进行语义转换，因此我们用非线性函数代替线性函数p（r），形式上为p（e，r）来实现h和r到t的语义传递，得到最终的尾部实体嵌入向量形式上为t=P’（P（h，r），t）。考虑到t到h的语义传递方向与r相反，我们将t和r到h的语义传递函数定义为p（−e，r），并将最终的头部实体嵌入向量形式化为h=P’（P（−t，r），h）。