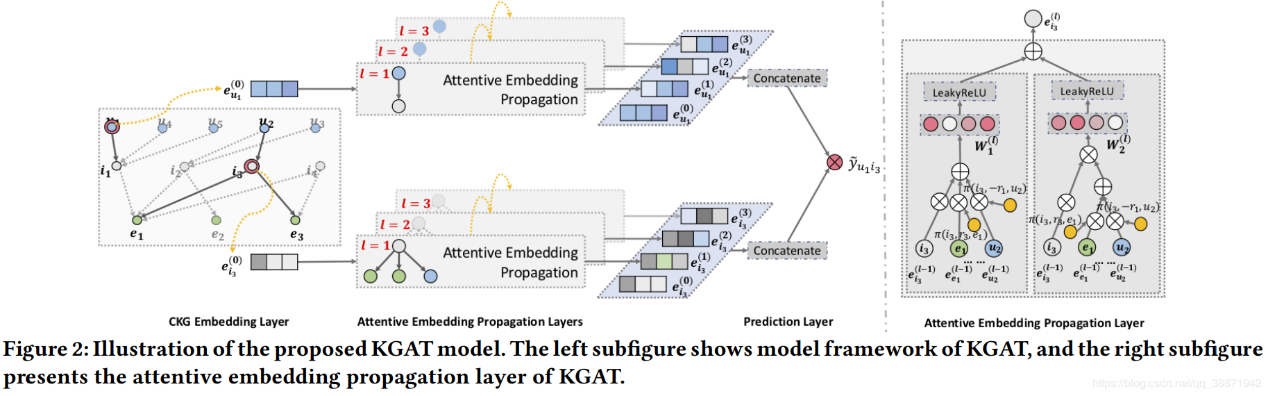
# 一、KGAT总体模型

主要包含三个部分：  
　　1、Embedding Layer：通过保留CKG的结构将每个节点参数化为一个向量  
　　2、Attentive Embedding Propogation Layer：递归地传播节点邻居的Embedding信息以更新其表示，并利用知识感知的注意力机制在传播过程中学习每个邻居的权值  
　　3、Prediction Layer：集成来自所有传播层的用户和物品的表示，并输出相应的预测评分



# 二、KGAT模型按部分详情

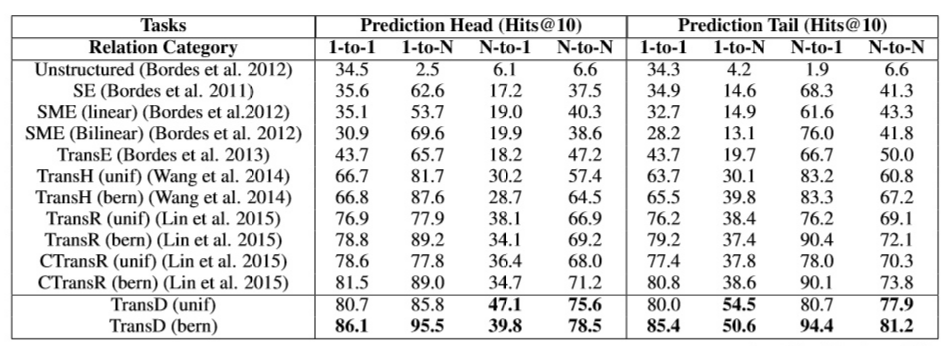
## ****1、Embedding Layer****

图嵌入部分。知识图谱的嵌入，是参数化实体和关系作为向量表示的一种有效方法，同时能够保留图的结构信息。（保留CKG论文的方法）

常用的知识表示方法可以分为两类：**基于结构的方法**和**基于语义的方法**。

1. 基于结构的嵌入表示方法包括TransE, TransH, TransR&CTransR, TransD等，这类方法从三元组的结构出发学习KG的实体和联系的表示；
2. 基于语义的嵌入表示方法包括NTN、 SSP和 DKRL等，这类方法从文本语义的角度出发学习KG的实体和联系的表示。

改进意见：



【基于结构的知识表示方法：<https://zhuanlan.zhihu.com/p/43436288】>

1. **Trans系列在原始的改进比较中，TransR和CTransR效果是相当的，我们的改进需要向着TransD进行改进**
2. **改进为用BPR来学习关系可信度-前期需要对数据集进行处理**

## ****2、Attentive Embedding Propagation Layers****

这个部分首先在单层上（如模型图中Attentive Embedding Propogation Layer部分是通过三个叠加layer组成，这里为了便于理解先从单层的角度出发）介绍其三个组成部分，再将其推广到多层上。

### ****（1）Information Propagation****

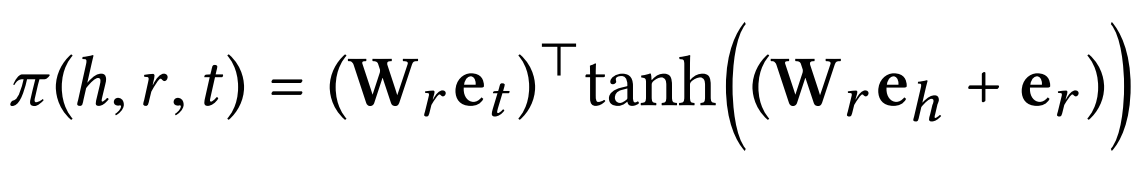
ego-network网络，构建异构信息网络。

改进意见：

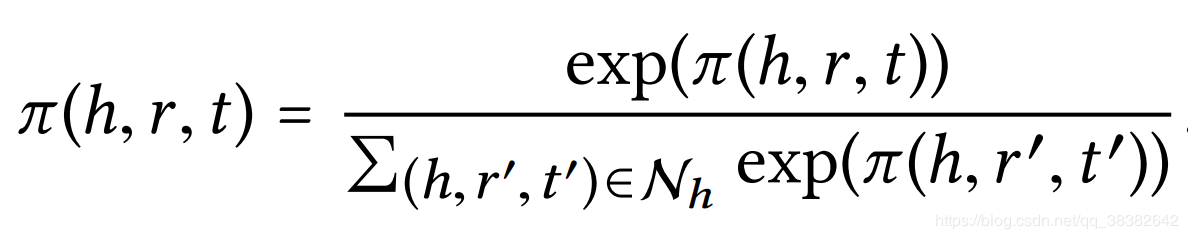
无

### ****（2）Knowledge-aware Attention****

注意力函数*eij*​也有很多种变体。四种注意力变体：加性注意力（additive attention）、乘法（点积）注意力（multiplicative attention）、自注意力（self-attention）和关键值注意力（key-value attention）。这里给出的是用tanh非线性激励函数+加性注意力（additive attention）,可以使得注意力得分依赖于超平面*r*空间中*eh*​和*et*​之间的距离，为更近的实体传播更多信息。



采用softmax函数对所有与刚刚说的与*h*头实体相连的三元组的系数进行归一化处理:



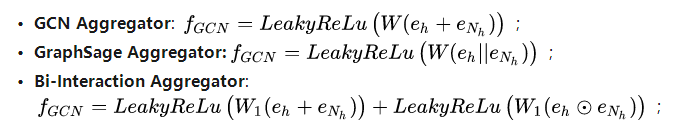
改进意见：

1. **注意力机制改进：改为自注意力机制**
2. **Tanh非线性激励函数改进（小）-激励函数有多种**
3. **softmax归一化函数改进（小）-归一化3种方法**

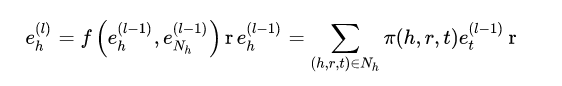
### ****（3）Information Aggregation****

该模块的作用：将前两层的结果（实体表示*eh*​ 和 ego-network表示 *eNh*​​）进行集成。

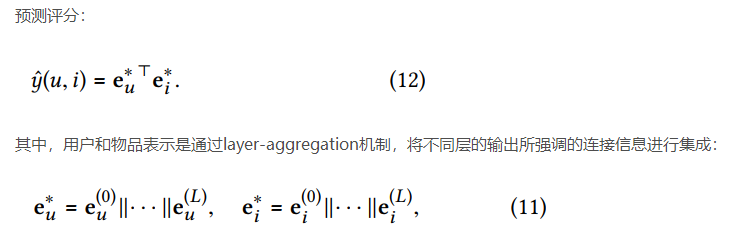
信息的聚合有3种方法（论文采用第三种）：

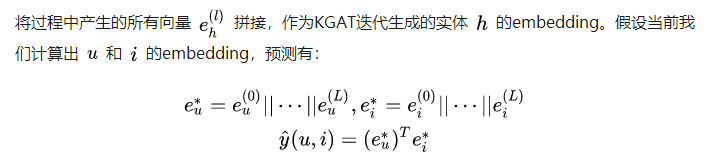
**多层传播**

重复前三步骤（三个层次）多次，对于第L次：

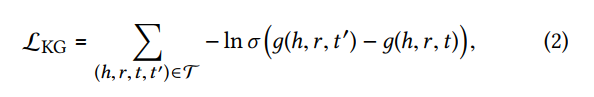


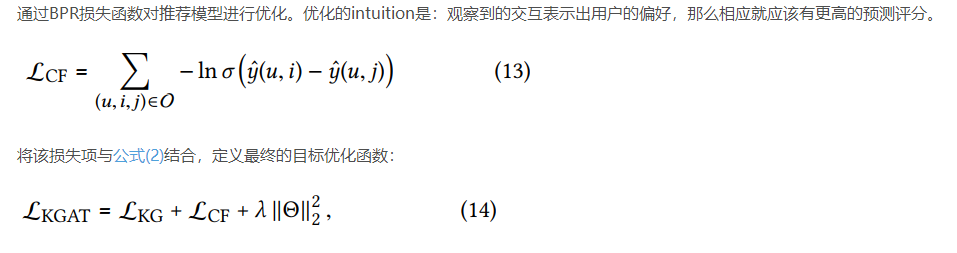
## 3、Prediction Layer





进行优化：





改进意见：

1. **论文采用的损失函数为：pairwise BPR loss+Lkg建立了新的损失函数**