补全是针对一个的KB/KG来的，挖掘这一个里面潜在的“边“。（link prediction）

然后可以从KB的内部结构里面推理（reason），也可以利用外部资源来找KB中的边关系。

**An Interpretable Knowledge Transfer Model for Knowledge Base Completion:**

**这个论文还是在如何更好的做embedding上下功夫（我好像知道为啥都是线性映射了，因为只要涉及翻译模型的embedding，就跑不开transE，transE就是线性映射，大家都跑不出来这个框）**。这个论文提出了一个ITransF，TransE是一个做embedding的模型，但是前提是假设entity和relation在一个vector space（h+r-t）；然后认为不同关系拥有不同的语言空间，发展出了TransR，把头实体和尾实体映射到relation-dependent space（未确定，应该是wh+r-wt）；然后又发展出了STransE，把头实体和尾实体用不同的映射方法映射到关系的空间（w1h+r-w2t）。然后这些都有数据稀疏的问题（sparsity problem），一个关系没那么多的数据，然后本文提出的这个ITransF相比之下解决的这个数据稀疏的问题。然后这个TransR和STransE还有个问题就是映射矩阵太多，参数太多，也阻碍了knowledge sharing，然后这个方案就要把relation-dependent space减小到smaller concept spaces（这个也没具体看怎么解决的）

这个用了“sparse attention mechanism”解决这些问题(没细看)

然后这个补全，大概是这个意思，给一个r和一个t，然后让你找h，就算这个energy function f(h,t)，然后做个rank。

**Feature-Rich Networks for Knowledge Base Completion**

**还是把这embedding来说的。只不过不是基于翻译模型（translation）了，是基于CNN这种深度神经网了，看这个论文第一个神经网络样子的插图就明白了。**这个是用额外的资源来辅助KB/KG进行补全，比如文档信息。KB/KG中的待发现的facts，用这些KB/KG外的信息来进行辅助发掘。Latent. MLP.

做KBC很好的一个模型是latent feature models，然后这个model是把KB做个embed映射到低维空间，然后给潜在的三元组一个score。这个没有用额外的比如text这种信息。然后一个用这些textual信息进行推理的办法是joint embed这个KB/KG的东西和textual的information。Joint需要一个universal schema。这种利用universal schema来joint embedding KB/KG的办法有一些局限性，KB和text relations的对应只能通过实体的共现隐式的对应起来，然后这个带了text就有各种的问题，noise，出现在一个地方有没有share语义关系巴拉巴拉……模型得在没有显式共现信息的前提下去学KB和text的对齐（alignment）……。然后这个文章就提出也是要结合KB和text，但是textual relations are not part of the same graph but are treated as side evidence。（虽然我不是很懂这个side evidence到底是咋操作的，**想问问老师，我每个论文看到它自己的思路的时候就下不去了，是应该深究下去还是别管了……**）(好像是这个text的embedding不用管，它只是KB做embedding的时候的一个辅助信息？)

**On the Equivalence of Holographic and Complex Embeddings for Link Prediction:**

**这个论文是基于两个论文的，这个两个论文的引用量都很高**

**Holographic Embedding of Knowledge Graphs:**

**这个做embedding用compositional vector space model（这个compositional到底啥意思我还没看明白,但是TransE是non-compositional的），然后这个compositional有很多种办法，其中有一种带了“non-linear”，说不定有用……**

**这个HoLE模型，结合了tensor product的expressive power和TransE的efficiency，simplicity。具体的没看，也没太看懂……**

**注：这个全息模型在表示学习的综述里提到了，说这个东西刚刚提出，尚未验证其效果，但是这个视角很新颖（综述是2016年，然后2017年这个日本人AAAI的论文说了这个东西）觉得这个全息模型可以看一看**

**这个非线性，好像都是带神经网的。几条大路，翻译模型做embedding就是线性的，然后什么神经网做embedding就是非线性的。现在又出了这个很新的全息模型，另一种做embedding的办法**

Concepts是什么意思？（好像是指这个实体在现实世界中是指什么概念）

ProjE：

减小参数的规模（前人工作有参数规模大的，还有多路径，这个只有length-1 paths，但是后面future里说要尝试更长的路径）；

self-contained，不需要预训练，不需要有各种pre的东西（前人的工作有）；

不用transformation matrices（我理解是这个论文搞得复杂点，利用的信息多一点？），用了一个combination operator，把e和r的信息融到了一起，然后再用一个模型（h(e,r)）算出来候选实体的score。

这些论文其实套路还是有的：

先说我要干啥-----再说之前的人做了啥------再说我的和之前的人比好在哪-----定义模型----定义损失函数

这个在具体的实现中，就把这个问题变成了多分类问题。（e，r）每个组合是一个类别，然后对于每个要预测的实体y，如果有(e,r,y)的关系，标签就是1，没有这个关系标签就是0，然后这个就是一个多分类问题了。（上面定义的h(e,r)中就相当于逻辑回归里的sigmoid函数的作用）

损失函数还用了softmax。

由于candidate-entities规模太大，用candidate sampling来减小这个规模。

数据用的FB15K，github上代码开源