

This the note for : 2011CB collective entity resolution in relational data

问题：将关系型实体进行集体对齐

如果把待对齐实体当做节点（节点可以具有属性），实体间有某种关系则连上一条关系边，则构成了reference graph。

属性相似度：每个属性计算相似度，然后线性加权。

类别相似度：对两个实体的关系实体集合计算相似度（比如jaccard相似度）

采用聚类的方式将指向同一对象的实体聚类。聚类采用层次聚类。

要初始化一些非常准确的对齐实体对。

算法基于的假设是：一个实体对应的潜在对象不仅仅由它的属性决定，还与它关联的实体的潜在对象决定。也就是说相关联的实体在进行对齐时是相关联的。