一个生成实体提及模型用于将实体链接到知识库

近年来，由于知识共享社区如维基百科1的扩散，以及Web上自动化知识库人群的众多研究工作，如“Web2”项目，越来越多的大型知识库可用。这些知识库包含有关世界实体的丰富知识，其语义属性以及彼此之间的语义关系。维基百科最知名的例子之一是：其2010年英语用文本数据填充这些知识库可以促进许多不同的任务，如实体搜索，信息提取和文本分类。例如，如图1所示，知道文档中的乔丹是指一名篮球运动员，而公牛队指的是NBA球队将有助于将本文件分类到运动/篮球类。

将知识库与文本数据相结合的一个关键问题是将文档中的实体与其通常称为实体链接任务的知识库中的对象进行链接。给定一组包含在文档中的名称M = {m1，m2，...，mk}以及包含一组实体E = {e1，e2，...，en}的知识库KB，实体链接系统是函数 ：ME将这些名称提及链接到其指定的实体（以KB为单位）。例如，在图1中，实体连接系统应该将名称提及约旦与实体迈克尔·杰弗里·乔丹（Michael Jeffrey Jordan）的名称相连，将公牛链接到芝加哥公牛队。

