考虑角色之间有相似性，例如人肯定和人相似而不和时间相似。

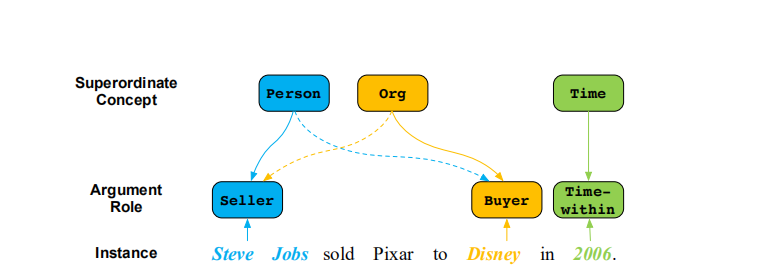
编码层

算某个高级概念（隐式 可能 人）对不同词的注意力，即c注意力。即这个概念可能更和什么词有关/

参数共享式按照上层关系来的。。（认为画了八种，即八种权值。）但有些有重叠

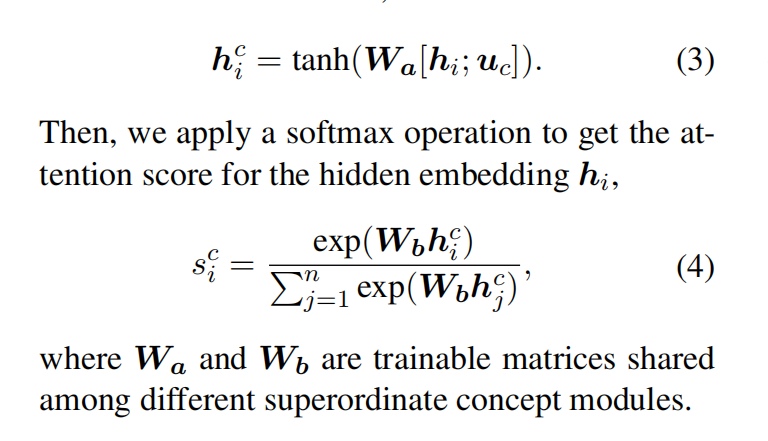
对每种高级事件，平均下来即为该词该取多大的角色占比。即role注意力

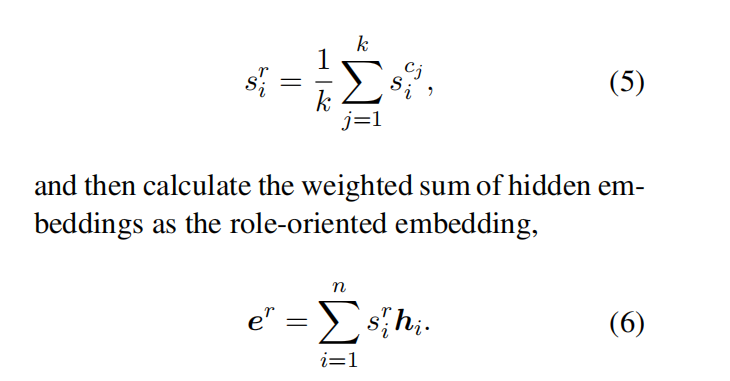
用role注意力平均。



也许0.6，0.4 0.3 0.7

于是变 1/8 1.1/8





一个角色要和其他的都有关。

要捕捉之间的关系必须注意力