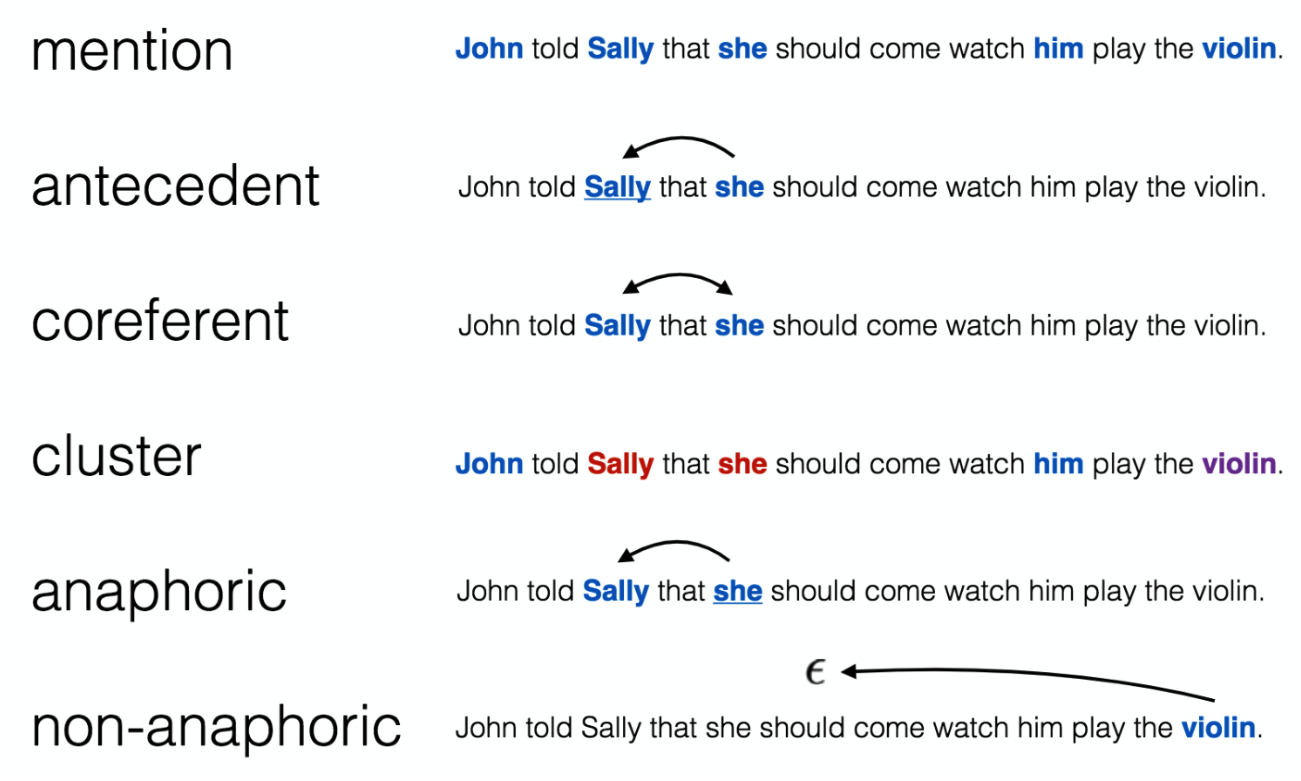
共指消解基本知识

共指消解(coreference resolution)，指的是在一文本中几个表达段指向现实世界中的同一个实体。

**相关术语**



其中各个名称代表的含义是：

* mention：指文本中具有实体含义的表述
* antecedent：前指，如上图中“Sally”和“she”具有共指关系，他们都指向“Sally”这个人。Sally在she前面，Sally也就是she的前指。
* coreferent：这就是共指关系，上文中的Sally和she就是具有共指关系。
* cluster：理解为簇，共指消解其实就是将文本中具有共指关系的mention聚到一起，形成一个簇。
* anaphoric：回指的，可以理解为与前指相对应。
* non-anaphoric：指文中没有与这个mention具有回指的Mention。

**相关特征**

共指消解需要考虑的特征主要有以下几类：词汇、距离、一致性、语法、语义等。

* 词汇特征：主要考虑两个mention的字符串的匹配程度，一般而言字符串相同程度越高的mention共指概率越大。
* 距离特征：主要考虑两个mention的句子距离，一般而言，两个mention的距离越近，他们之间共指的概率也就越大。
* 一致性特征：一致性特征可以分为性别、单复数、语义类别等是否一致，这组特征主要起到筛选作用。
* 语法特征：语法关系可以用来判断两个mention的语法角色之间的关系，可以采用一些基于特定模板的方法，例如判断两个mention之间是否被逗号隔开或者相邻等来决定是否具有同位关系。
* 语义特征：语义特征主要是考察两个mention在语义类别不一致时是否满足上下位或者同义、近义关系。