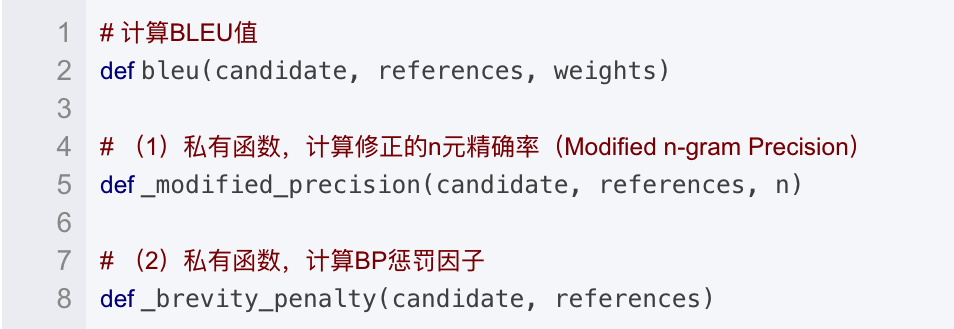
BLEU

BLEU(Bilingual Evaluation Understudy),原论文是BLEU: a Method for automatic evaluation of machine translation.

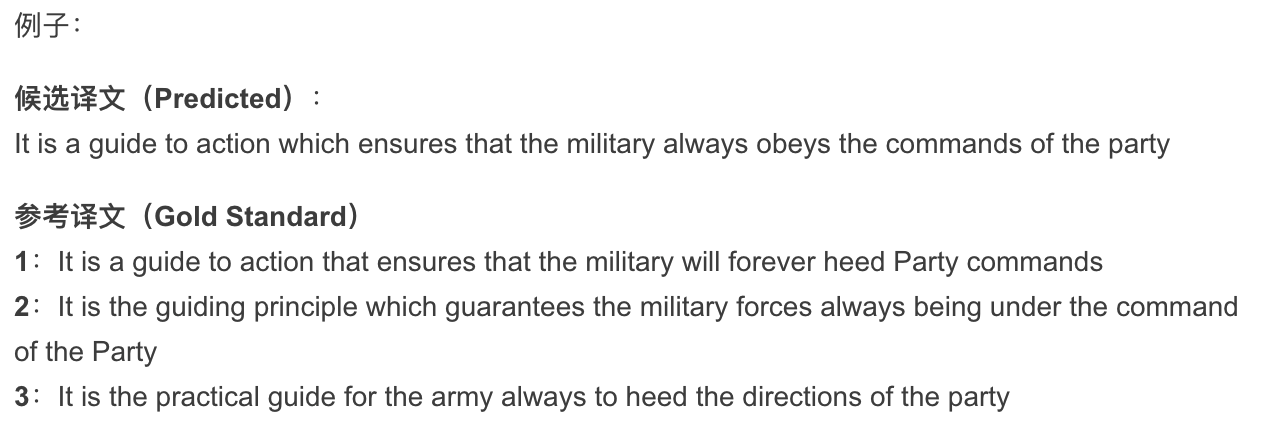
下面我们通过一个例子来详细介绍BLEU是如何计算以及NLTK中的nltk.align.bleu\_score 模块的源码，首先祭出公式：

其中是Uniform weights，也就是权重，是修正后的n-gram计算的分值，BP是Brevity Penalty，这个是一个惩罚因子，用来对候选句子过短时进行惩罚。

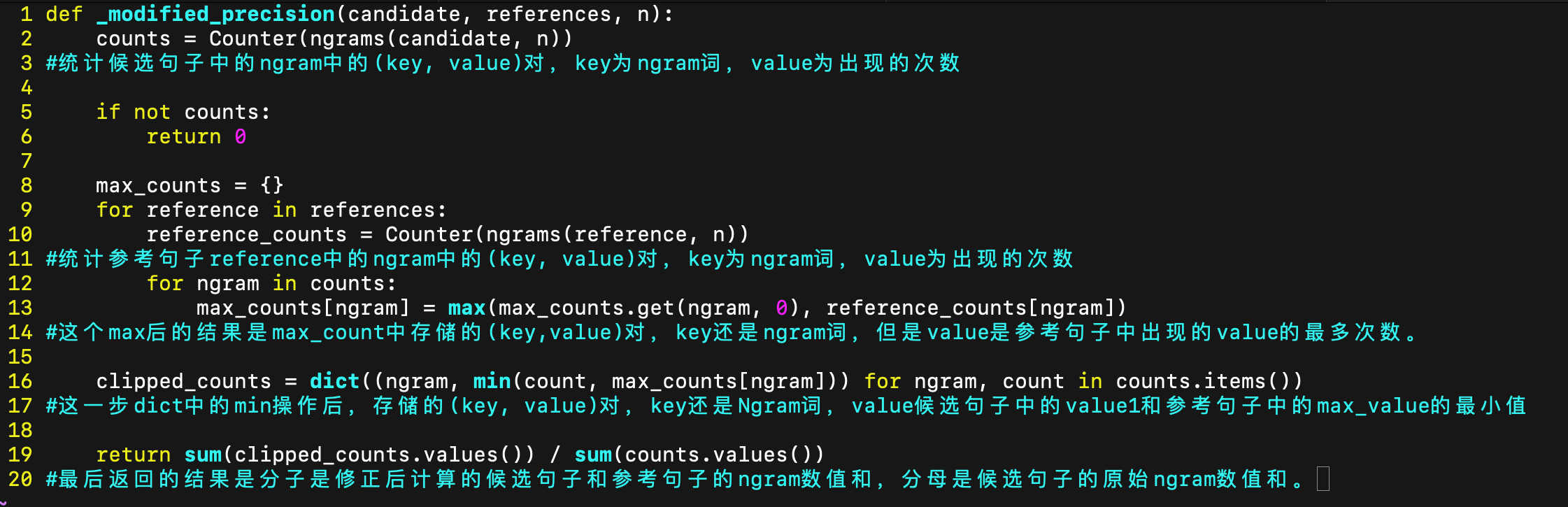
Nltk.align.bleu\_score模块实现了这里的公式，主要包括三个函数，两个私有函数分别计算P和BP，一个函数整合计算BLEU分值。



一个例子如下：

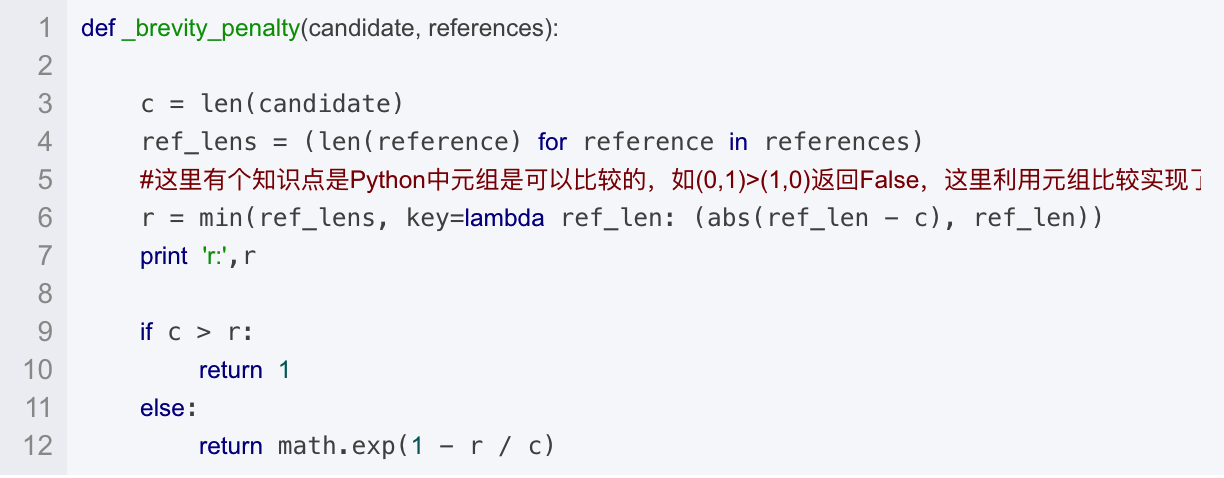


计算的代码如下：



这里的n取值为4，也就是从1-gram计算到4-gram。

下面是计算惩罚因子BP的代码：



其实BP主要是“过短惩罚”，它的取值范围是，候选句子越短，越接近0。

假如候选翻译句子长度为18，参考翻译长度分别为:16，18，16.

所以c=18，r=18（参考翻译中选取长度最接近候选翻译的作为r）。

所以

最后计算出BP,,以后，代入公式计算即可。