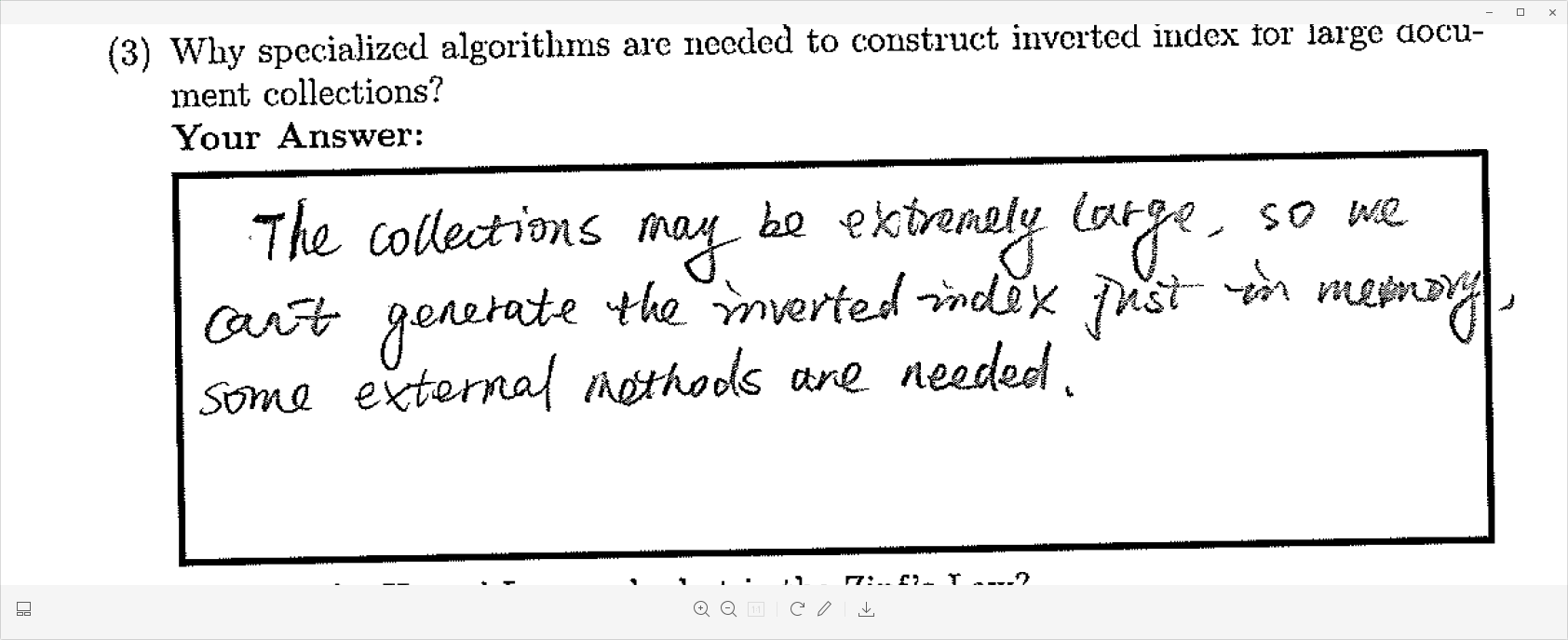
简答题：

1.Consin similarity 和 eucldean哪个更好

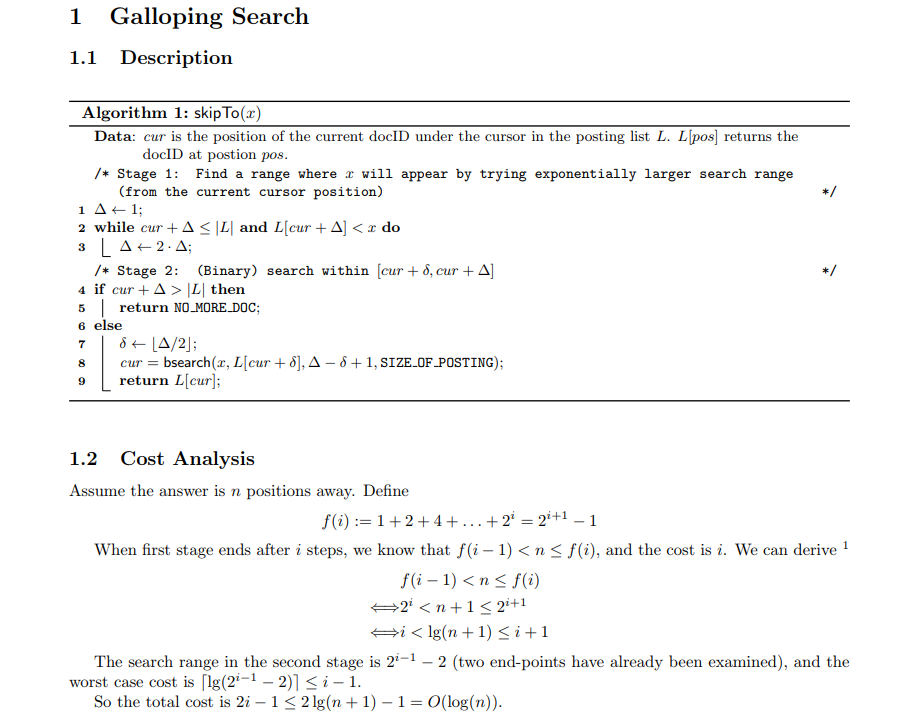
2.How does stemming typically affect recall ? why?

3.

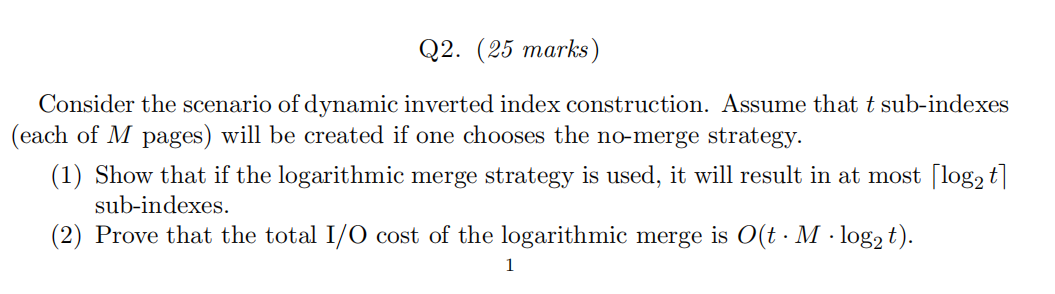


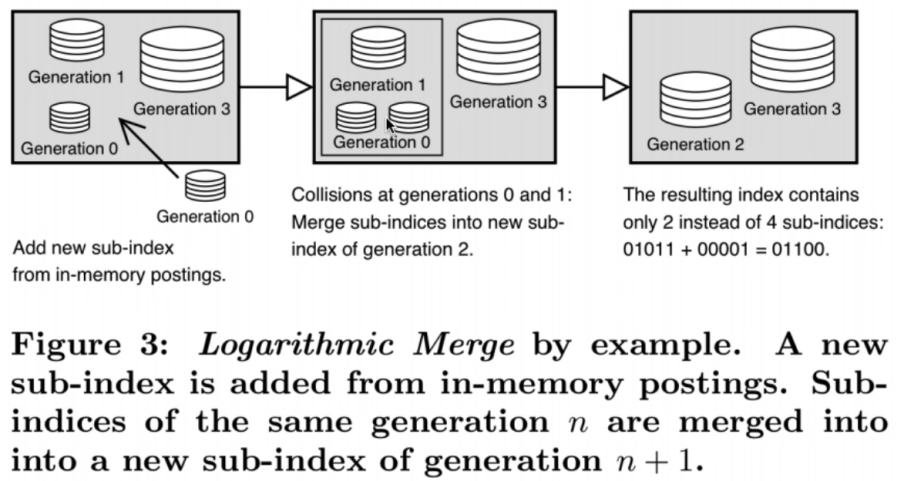
4.recall和precision 会不会一定有一个时间点相等？为什么？

Q2： 简述题和伪代码



Q3: log merge 题



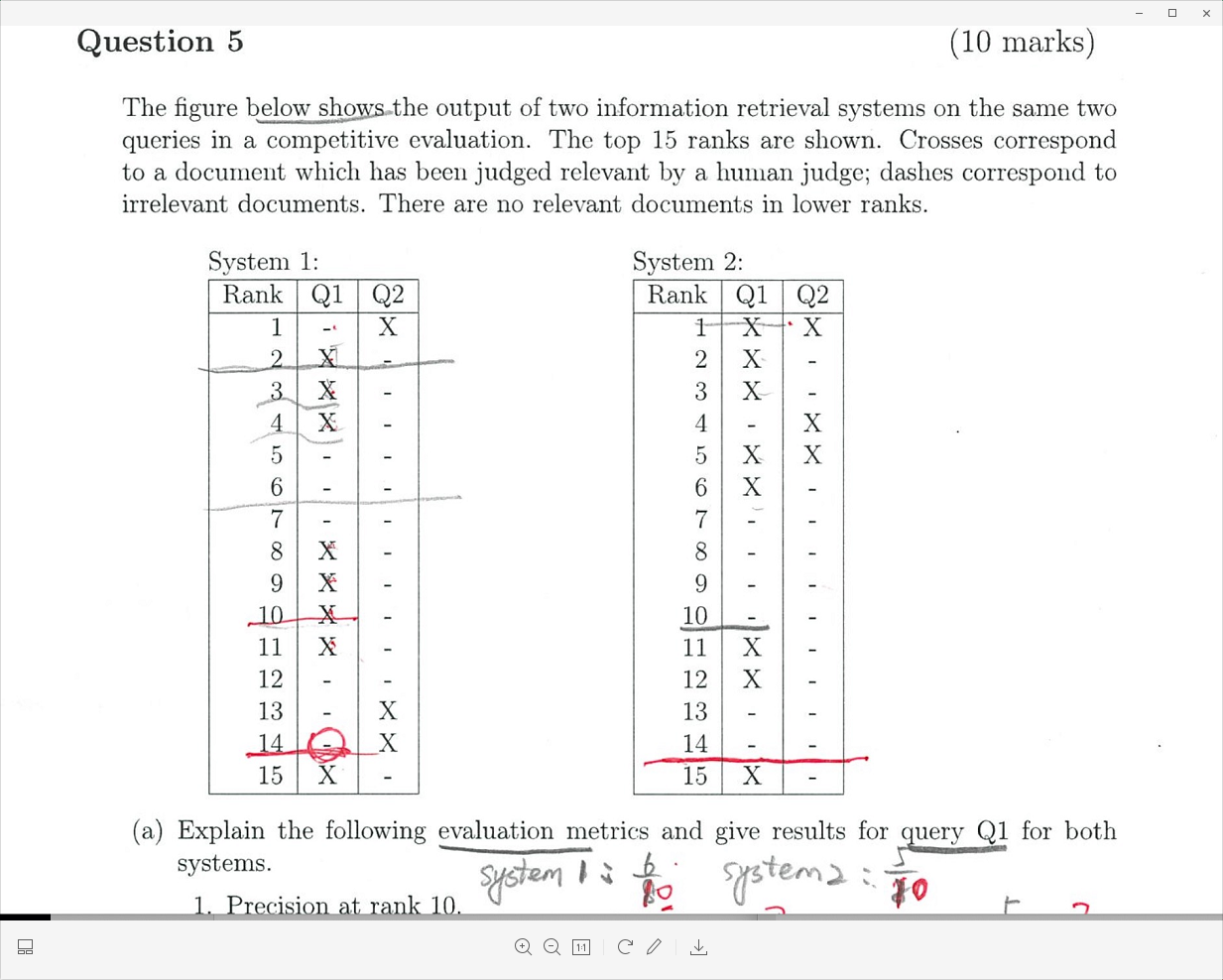


第一问merge结果，memory满了就变化merge，求sub-index的变化

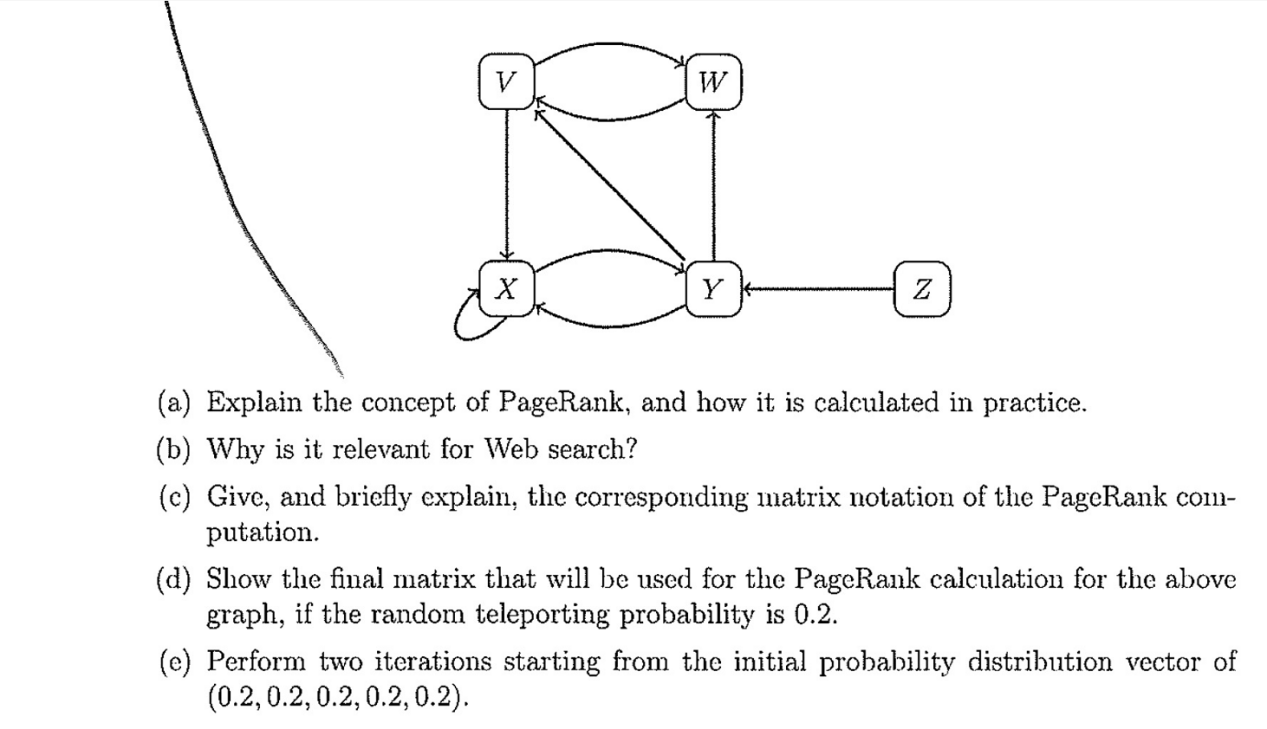
第二小问，assignment q2第一问原题

第三小问，求merge次数，递推出来，先找到当前sub-index有哪些，用递推的公式求merge次数的总和

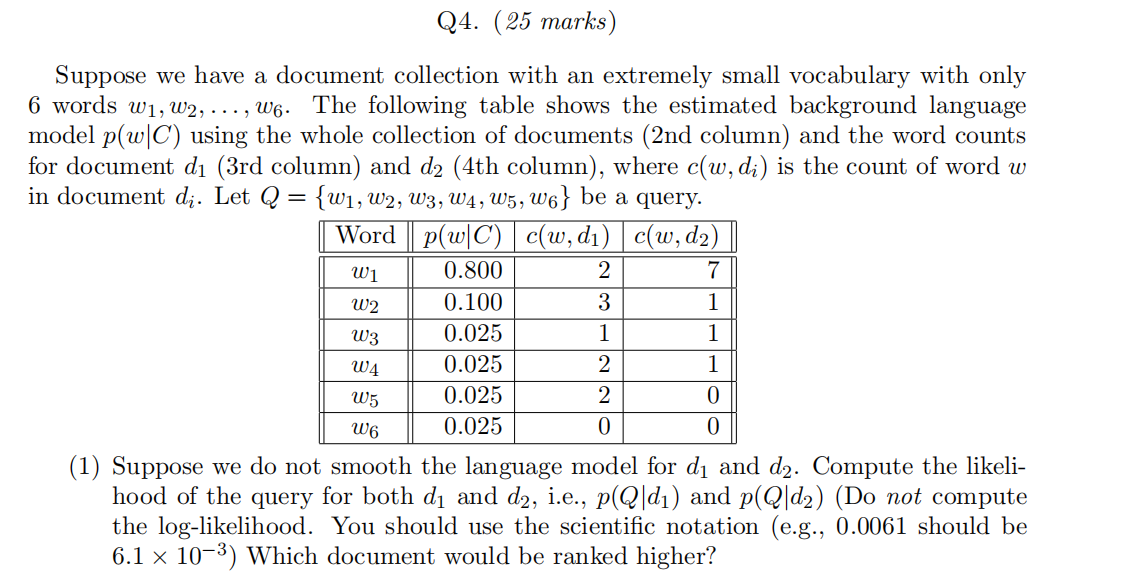
Q4:

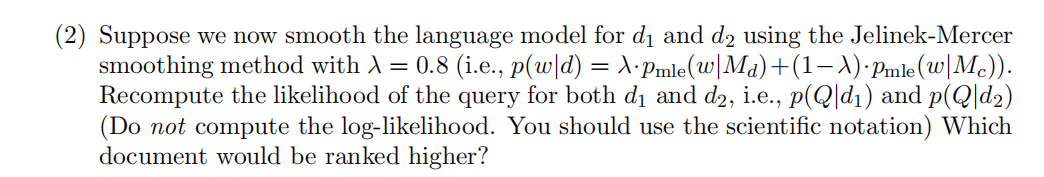


Q5：原题

第四问考试概率是0.5，最后只要迭代一次。

Q6:assignment q4原题





Q7：ppmi（阿尔法）

Q8:ipm