

算法基础

第十次作业 (DDL: 2024 年 12 月 19 日 23:59)

解答过程中请写出必要的计算和证明过程

Q1.(20 分) 对字母表 $\Sigma = \{a, b, c\}$, 模式 $P = ababcbabababc$,

1. 画出 P 对应的字符串匹配自动机的状态转换图。
2. 计算 P 的前缀函数 π 。

Q2.(15 分) 定义一个字符串 s_1, s_2, \dots, s_n 的循环变换为 $s_2, s_3, \dots, s_n, s_1$ 。给的两个字符串 P_1, P_2 , 给出 $O(n)$ 的算法确定 P_2 是否可以通过对 P_1 进行若干次循环变换得到。

Q3.(15 分) 定义一个字符串 $S = s_1, s_2, \dots, s_n$ 的最小循环节为最短的字符串 $T = s_1, s_2, \dots, s_r$, 且满足:

1. 是 S 的一个前缀。
2. 可以将 T 复制有限次得到字符串 $S' = TT\dots T$, 使得 S 是 S' 的前缀。

例如, $ababa$ 的最小循环节是 ab , $abcabc$ 的最小循环节是 abc 。给定字符串 P , 给出 $O(n)$ 的算法确定 P 的最小循环节。

Q4.(20 分) 将 3-SAT 问题归约到独立集问题。

note. 3-SAT 问题: 输入一个合取范式形式的命题公式 Φ , 其中每个子句恰好包含 3 个文字, Φ 是否可满足? 独立集问题: 输入一个图 G 和一个整数 k , 图 G 是否存在一个大小至少为 k 的独立集?

Q5.(30 分) 考虑集合覆盖问题 (set cover problem):

令 U 是一个包含 n 个元素的集合。令 $S = \{S_1, S_2, \dots, S_m\}$ 是 U 的子集集合, 且满足 $\bigcup_{i=1}^m S_i = U$ 。我们的目标是从 S 中选择尽可能少的子集, 使得它们的并集覆盖 U 。

考虑以下算法 SETCOVER 用于解决这个问题:

Algorithm 1: SETCOVER

Input: A set U and a collection of subsets $\mathcal{S} = \{S_1, S_2, \dots, S_m\}$
 such that $\bigcup_{i=1}^m S_i = U$

Output: A collection $C \subseteq \mathcal{S}$ that covers U

```

1 Initialize  $C \leftarrow \emptyset$ ;
2 while  $U$  contains elements not covered by  $C$  do
3   Find the set  $S_i$  containing the largest number of uncovered
   elements;
4   Add  $S_i$  to  $C$ ;
5 return  $C$ ;
```

为了分析上述算法, 令 $k = \text{OPT}$, 即最优解中所需的集合数。令 $E_0 = U$, 且令 E_t 表示在第 t 步之后尚未被覆盖的元素集合。

(a) (15 分) 证明 $|E_{t+1}| \leq |E_t| - |E_t|/k$ 。

(b) (15 分) 证明算法 SETCOVER 是一个用于解决集合覆盖问题的 $\ln n$ -近似算法。

提示: 证明算法 SETCOVER 在 $\text{OPT} \cdot \ln n$ 步内结束。