最小表示法

定义

最小表示法是用于解决字符串最小表示问题的方法。

字符串的最小表示

循环同构

当字符串 S 中可以选定一个位置 i 满足

$$S[i \cdots n] + S[1 \cdots i-1] = T$$

则称 S 与 T 循环同构

最小表示

字符串 S 的最小表示为与 S 循环同构的所有字符串中字典序最小的字符串

simple 的暴力

我们每次比较 i 和 j 开始的循环同构，把当前比较到的位置记作 k，每次遇到不一样的字符时便把大的跳过，最后剩下的就是最优解。

实现

C++

```cpp
int k = 0, i = 0, j = 1;
while (k < n && i < n && j < n) {
  if (sec[(i + k) % n] == sec[(j + k) % n]) {
    ++k;
  } else {
    if (sec[(i + k) % n] > sec[(j + k) % n])
      ++i;
    else
      ++j;
    k = 0;
    if (i == j) i++;
  }
```

```
}
i = min(i, j);
```

Python

```python
k, i, j = 0, 0, 1
while k < n and i < n and j < n:
    if sec[(i + k) % n] == sec[(j + k) % n]:
        k += 1
    else:
        if sec[(i + k) % n] > sec[(j + k) % n]:
            i += 1
        else:
            j += 1
        k = 0
        if i == j:
            i += 1
i = min(i, j)
```

解释

该实现方法随机数据下表现良好，但是可以构造特殊数据卡掉。

例如：对于 $\mathtt{aaa}\cdots\mathtt{aab}$, 不难发现这个算法的复杂度退化为 $O(n^2)$。

我们发现，当字符串中出现多个连续重复子串时，此算法效率降低，我们考虑优化这个过程。

最小表示法

算法核心

考虑对于一对字符串 A, B, 它们在原字符串 S 中的起始位置分别为 i, j, 且它们的前 k 个字符均相同，即

$$S[i\cdots i+k-1] = S[j\cdots j+k-1]$$

不妨先考虑 $S[i+k] > S[j+k]$ 的情况，我们发现起始位置下标 l 满足 $i \le l \le i+k$ 的字符串均不能成为答案。因为对于任意一个字符串 S_{i+p}（表示以 $i+p$ 为起始位置的字符串，$p \in [0, k]$）一定存在字符串 S_{j+p} 比它更优。

所以我们比较时可以跳过下标 $l \in [i, i+k]$, 直接比较 S_{i+k+1}

这样，我们就完成了对于上文暴力的优化。

时间复杂度

$O(n)$

过程

1. 初始化指针 i 为 0，j 为 1；初始化匹配长度 k 为 0
2. 比较第 k 位的大小，根据比较结果跳转相应指针。若跳转后两个指针相同，则随意选一个加一以保证比较的两个字符串不同
3. 重复上述过程，直到比较结束
4. 答案为 i, j 中较小的一个

实现

C++

```cpp
int k = 0, i = 0, j = 1;
while (k < n && i < n && j < n) {
  if (sec[(i + k) % n] == sec[(j + k) % n]) {
    k++;
  } else {
    sec[(i + k) % n] > sec[(j + k) % n] ? i = i + k + 1 : j = j + k + 1;
    if (i == j) i++;
    k = 0;
  }
}
i = min(i, j);
```

Python

```python
k, i, j = 0, 0, 1
while k < n and i < n and j < n:
    if sec[(i + k) % n] == sec[(j + k) % n]:
        k += 1
    else:
        if sec[(i + k) % n] > sec[(j + k) % n]:
            i = i + k + 1
        else:
            j = j + k + 1
        if i == j:
            i += 1
        k = 0
i = min(i, j)
```

本页面最近更新：2023/10/4 21:50:08，更新历史
发现错误？想一起完善？ 在 GitHub 上编辑此页！
本页面贡献者：Ir1d, partychicken, StudyingFather, countercurrent-time, Early0v0, Enter-tainer, H-J-Granger, iamtwz, NachtgeistW, Suyun514, Xeonacid, AngelKitty, CCXXXI, cjsoft, diauweb, ezoixx130, fjzzq2002, GavinZhengOI, GekkaSaori, Gesrua, Junyan721113, karin0, Konano, ksyx, kxccc, LovelyBuggies, lychees, Makkiy, Menci, mgt, minghu6, P-Y-Y,

Peanut-Tang, PotassiumWings, SamZhangQingChuan, shawlleyw, sshwy, SukkaW, TrisolarisHD, weiyong1024