

计算机问题求解 – 论题2-14

- 用于动态等价关系的数据结构

课程研讨

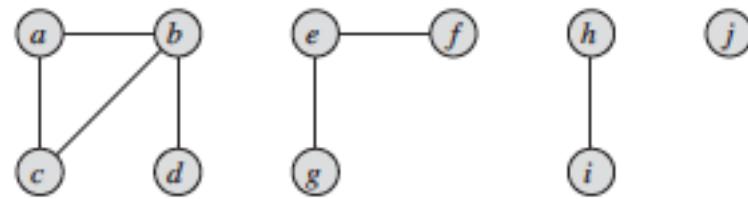
- TC第21章
- SB第2章

问题1：connected components

- 你能否借助这个例子，解释Union-Find的三个操作？
 - MAKE-SET(x)
 - UNION(x,y)
 - FIND-SET(x)
- 你会证明CONNECTED-COMPONENTS的正确性吗？

CONNECTED-COMPONENTS (G)

```
1 for each vertex  $v \in G.V$ 
2   MAKE-SET( $v$ )
3 for each edge  $(u, v) \in G.E$ 
4   if FIND-SET( $u$ )  $\neq$  FIND-SET( $v$ )
5     UNION( $u, v$ )
```

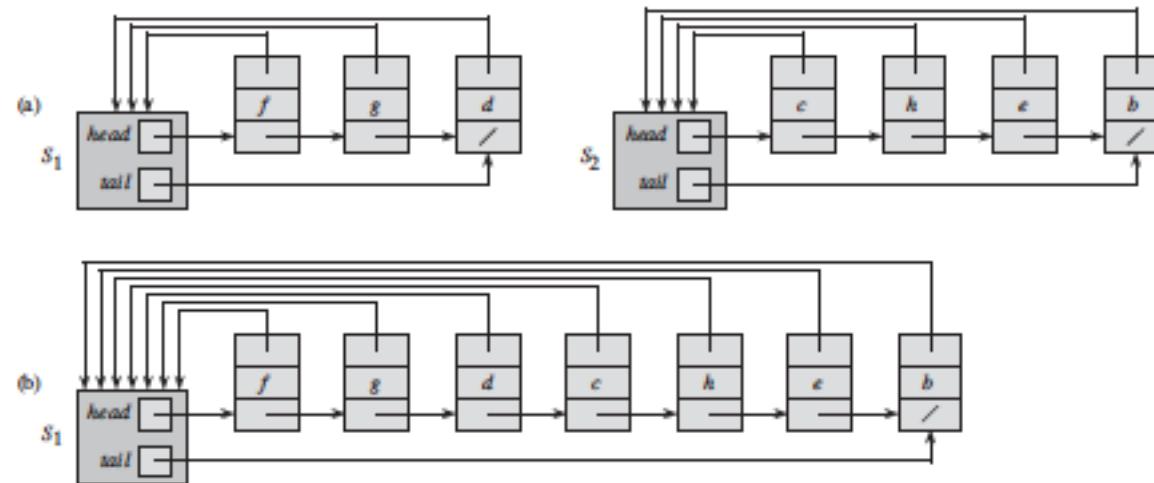


Union-Find的实现

- Matrix (relation matrix)
 - Space in $\Theta(n^2)$, and worst-case cost in $\Theta(mn)$ (mainly for row copying for MAKE).
- Array (for equivalence class id.)
 - Space in $\Theta(n)$, and worst-case cost in $\Theta(mn)$ (mainly for search and change for MAKE).

问题2：linked-list representation

- 请简要解释这种实现方式
 - Union-Find的三个操作如何实现？
 - 运行时间分别是？其中的瓶颈是？
 - 如何利用weighted-union heuristic加以改进？
 - 能改进到O(1)吗？（需要付出什么其它代价？）
 - head和tail能不能省掉一个？（不改变运行时间）

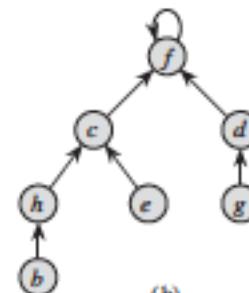


问题3：disjoint-set forests

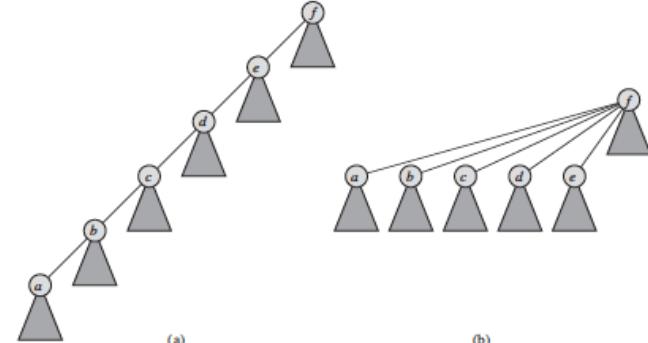
- 请简要解释这种实现方式
 - Union-Find的三个操作如何实现？
 - 运行时间分别是？其中的瓶颈是？
 - 如何利用union by rank加以改进？
 - 如何利用path compression加以改进？



(a)



(b)



问题4： how many tables

- Today is Ignatius' birthday. He invites a lot of friends. Now it's dinner time. Ignatius wants to know how many tables he needs at least. You have to notice that not all the friends know each other, and all the friends do not want to stay with strangers.

问题5： the suspects

- Severe acute respiratory syndrome (SARS), an atypical pneumonia of unknown etiology, was recognized as a global threat in mid-March 2003. To minimize transmission to others, the best strategy is to separate the suspects from others.
- In the Not-Spreadng-Your-Sickness University (NSYSU), there are many student groups. Students in the same group intercommunicate with each other frequently, and a student may join several groups. To prevent the possible transmissions of SARS, the NSYSU collects the member lists of all student groups, and makes the following rule in their standard operation procedure (SOP).
 - Once a member in a group is a suspect, all members in the group are suspects.
 - However, they find that it is not easy to identify all the suspects when a student is recognized as a suspect. Your job is to write a program which finds all the suspects.

问题6：partition refinement

- In the design of algorithms, **partition refinement** is a technique for representing a partition of a set as a data structure that allows the partition to be refined by splitting its sets into a larger number of smaller sets. In that sense it is **dual** to the union-find data structure, which also maintains a partition into disjoint sets but in which the operations merge pairs of sets together.
- 请给出partition refinement的一种实现，并分析 refinement的运行时间。
- 利用你的实现，解决distinguishing prefixes问题：
Given a set of strings S , the distinguishing prefix of a string s in S is the shortest prefix of s that is a prefix of no other string in S .
Determine the distinguishing prefixes of all the strings in S .