

자료구조 기말고사 (105점 만점)

Spring 2015
Open book, 70분

* 거의 없겠지만, 혹시 문제에 가정이 필요하거나, 문제에 추가적인 정보가 필요하다면 스스로 가정하고 하라. (합리적인 가정은 가산점을 줄 수 있고, 불필요한 가정은 감점 요인이 될 수 있음.)

1. (20점) 아래에 답하라. 당신의 대답에 대한 이유도 같이 밝혀야 한다. O-X 문제도 대답의 이유를 밝혀야 한다.

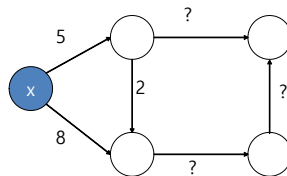
1.1 Quicksort의 수행시간은 $O(n^2)$ 이다. (O, X)

1.2 Insertion sort에 이미 제대로 정렬되어 있는 입력이 들어가면 수행시간은 $O(n)$ 이다. (O, X)

1.3 서울대생의 수가 충분히 많다고 가정하자. 학번으로 서울대생들을 sorting하는 $\Theta(n)$ 알고리즘이 존재한다. (O, X)

1.4 Merge sort는 asymptotic running time 관점에서 보면 이론적으로 거의 완벽한 sorting 알고리즘이다. 이것이 왜 필드에서 그리 환영받지 못할까?

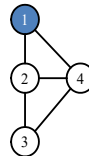
2. (10점. 질문 불허) 아래 그래프의 각 edge에 Dijkstra의 shortest-path 알고리즘이 optimal solution을 구하지 못하도록 weight을 할당해 보라. Vertex x가 시작 vertex이다. 그림을 똑같이 그린 다음 물음표의 weight 값들을 할당해 보라.



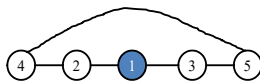
3. (10점) 아래 그래프 중 DFS traversal과 BFS traversal의 방문 순서가 동일한 것은 O, 다른 것은 X로 표하라. Vertex 1부터 방문하고 다음에 방문할 수 있는 vertex가 2개 이상이면 vertex의 index가 낮은 것을 우선적으로 방문한다. 이유 설명 필요 없음.



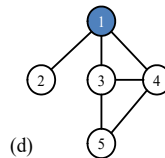
(a)



(b)



(c)



(d)

4. (15점. 질문 불허. 질문을 이해하는지를 보는 문제임.) Array $A[1...n]$ 에 들어있는 원소들이 min-heap을 이루고 있다. Heap의 루트 노드에 있는 원소를 제거하여 heap의 크기를 $n-1$ 로 줄이려 한다. 아래 물음에 답하라.(간단한 이유와 함께)
- 4.1 $A[1...n]$ 의 원소들이 모두 동일할 때, 이 작업에 소요되는 시간은 얼마인가? $\Theta()$ 로 대답하라.
- 4.2 $A[1...n]$ 에 원소들이 크기가 작은 순서로 정렬되어 있을 때, 이 작업에 소요되는 시간은 얼마인가? $\Theta()$ 로 대답하라.
5. (15점) 한 노드에 최대 1000개까지의 key를 담을 수 있는 B-tree에 대해 물음에 답하라.
- 5.1 “leaf node들에 있는 key들의 총 수 ÷ internal node들에 있는 key들의 총 수”의 가능한 최댓값과 최솟값은 얼마인가? B-tree를 만족하기만 하면 가능하다고 보고, 그런 상태가 실제로 만들어질 수 있는지는 생각하지 않아도 좋다. B-tree의 높이는 2 이상 20 이하라 가정하라.
- 5.2 방금 한 원소의 삭제 결과 B-tree의 높이가 하나 낮아졌다. 방금 진행된 삭제 직전의 B-tree는 어떤 모양이었음을 알 수 있는가? 삭제 작업과 관련된 곳들의 상황만 알 수 있으므로 방금 진행된 삭제의 전 과정에서 알 수 있는 부분을 최대한 설명하라.
6. (15점. 기출) Connected graph $G=(V,E)$ 에서 $|V|=1$ 억 이다. 각 vertex는 평균 10만 개의 다른 vertex와 adjacent하다. Adjacency matrix는 10^{16} 에 비례하는 공간의 크기 때문에 사용하기 주저된다. Adjacency list로 표현하려 하니 각 vertex당 평균 10만 개의 노드를 가진 linked list를 만들어야 하는데, 임의의 두 vertex의 adjacency 여부를 알아보기 위해 링크를 따라가면서 두 수를 비교하는 작업을 평균 5만번(adjacent할 경우) 또는 평균 10만번(adjacent하지 않을 경우) 정도 해야 하므로 그것도 마음에 들지 않는다. 더구나 vertex들의 adjacency 체크는 알고리즘이 수시로 해야 하는 기본 작업이다. $O(|E|)$ 의 공간을 사용하면서 빠른 시간에 adjacency 체크를 할 수 있도록 하려 한다. 임의의 두 vertex간에 edge가 존재하는지를 평균 $\Theta(\log k)$ 시간에 체크할 수 있는 방법을 제시하라.(여기서 k 는 각 vertex의 평균 degree) 채점자를 위해 그림을 그려서 설명하라.
7. (15점) 나라가 어떤 유행에 휩쓸리다 보니 부모의 성을 다 사용하여 성을 지어야 하도록 법령이 개정되었다. 예를 들어, ‘최전 두환’과 ‘박이 순자’가 결혼하여 재용을 낳으면 이 사람은 ‘최전박이 재용’이 된다. ‘최전박이 재용’과 ‘박김이전 지숙’이 결혼하여 소라를 낳으면 이 사람은 ‘최전박이박김이전 소라’가 된다. 이렇게 k 세대를 거치니 성의 길이 n 이 2^k 가 되었다. 몇백년이 흐르고 나니 성이 너무 길어져 도저히 호적을 관리하기가 힘들어졌다. 이에 국가에서는 모든 국민들의 성에서 중복이 일어난 부분을 모두 제거하고 남은 성들을 나열한 것으로 성을 삼기로 결정하였다. 예를 들어, ‘김박골동최전이전천방박박마박천민’이란 성을 가진 사람의 성에서 중복을 제거하면 ‘김박골동최전이전천방마민’이 된다. ‘남궁’ 같은 2자 성은 없다고 가정한다.
- 모든 국민에 대해 이 작업을 하고자 하니 효율적인 알고리즘이 필요하게 되었다. 충분히 큰 $n=2^k$ 길이의 성을 가진 어떤 한 사람에 대해 이 작업을 평균 $\Theta(n)$ 의 시간에 끝내는 방법을 제안하라. 우리나라에 존재하는 성의 총 수는 n 보다 충분히 클 수도 있다고 가정하라.
- 여러분은 이 과정을 이해하고 있다는 점을 채점자가 알 수 있을 정도로 묘사하면 된다. 수행시간에 대한 설명도 있어야 한다.