Assignment 15, Data-Structures

Oleg Sivokon

<2016-05-14 Sat>

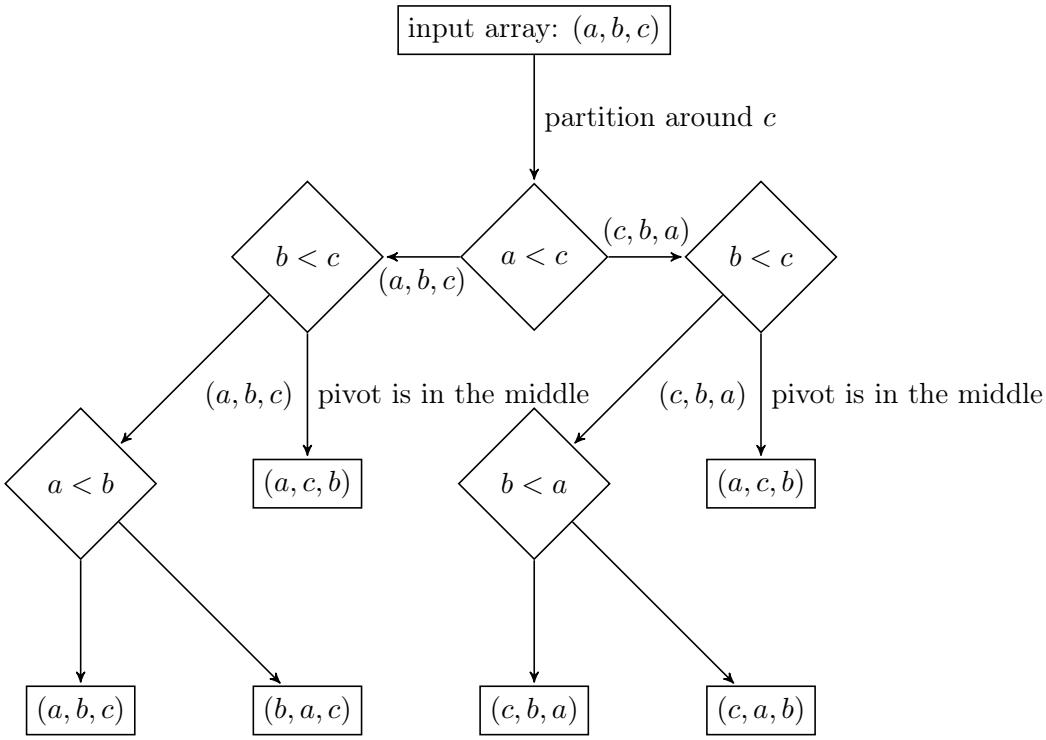
Contents

1 Problems

1.1 Problem 1

1. How many comparisons does the `quicksort` algorithm do on the input of size 3 in the best, worst and average cases?
2. Graph the decision tree for the above question, comment on how it represent the answers to the previous question.
3. How many comparisions does the `heapsort` do on the input of size 4 in the best, worst and average cases?
4. Show that for the input size 4, `heapsort` is sub-optimal. Explain why this doesn't contradict general optimality claims.

1.1.1 Answer 1



Which also answers the question of number of comparison that need to be made:

1. Best case scenario: 2 comparisons. If we are lucky, the pivot element ends up in the middle of the partitioned array, thus all resulting sub-arrays are of size 1 and need not be partitioned further.
2. Worst case scenario: 3 comparisons. Ohterwise, we'll need to compare both other elements with the pivot and then break the tie between the remaining elements.
3. As you can see above, we are "lucky" only 2 times out of 6. Taking weighted average gives us $2 \times 2 \times \frac{1}{6} + 3 \times 4 \times \frac{1}{6} = \frac{8}{3}$.

1.2 Problem 2

Design data-structure containing two independent queues both using the same "circular array" for storage. Define necessary operations: *insertion*, *deletion*, *boundary-checking*.

1.2.1 Answer 2

The idea is exactly the same as it was for the single queue, however in this case the elements of the first queue will be positioned at odd indices, while elements of the second queue will be positioned on the even indices. We will also need to keep four variables storing the position of the head and the tail of each of the two queues. *(See figure on the next page.)*

1.3 Problem 3

Given set S s.t. $S \subset \mathbb{N}, |S| = n$, $\max(S) = n^k - 1, k \geq 0$. Also given natural number z.

Algorithm 1 Double FIFO queue

```
procedure push(element, queue)
    if can-push(queue) then
        size ← size(queue)
        tail ← tail(queue)
        head ← head(queue)
        if is-even(queue) then
            queue[tail × 2] ← element
        else
            queue[tail × 2 + 1] ← element
        end if
        tail ← (tail + 1) mod size
    end if
end procedure
procedure pop(queue)
    if can-pop(queue) then
        size ← size(queue)
        tail ← tail(queue)
        head ← head(queue)
        if is-even(queue) then
            index ← head × 2
        else
            index ← head × 2 + 1
        end if
        head ← (head + 1) mod size
        return result
    end if
end procedure
procedure can-push(queue)
    size ← size(queue)
    tail ← tail(queue)
    head ← head(queue)
    return (tail + 1) mod size ≠ head
end procedure
procedure can-pop(queue)
    tail ← tail(queue)
    head ← head(queue)
    return tail ≠ head
end procedure
```

1. Write an algorithm for finding two distinct summands of z in S, s.t. its running time is $\Theta(n \times \min(k, \lg n))$.
2. Same as above, but find three distinct summands. Running time $\Theta(n^2)$.
3. Same as above, but for four distinct summands. Running time $\Theta(n^2 \times \min(k, \lg n))$.
4. Same as above, but for five distinct summands. Running time $\Theta(n^3)$.

1.3.1 Answer 3

One way of doing this would be, knowing k normalize the members of S by taking k^{th} root. Then, use bucket sort or radix sort. Then use binary search to find the summands. For three-summands algorithm we use the first algorithm as a sub-routine while looking at each element of the sorted array. Thus whenever we increase the number of summands we want, we will only gain a logarithmic increase in running time.

1.4 Problem 4

Given list of points $P = \{(x, y) \mid x^2 + y^2 \leq 0, x \geq 0\}$, assuming uniform random distribution of points across the semi-circle, write an algorithm for sorting them on $\tan\theta$, where θ is the angle between x axis and the line through origin and the given point.

1.4.1 Answer 4

The idea is to take `atan2` of x and y coordinates and use this value to represent the point. Then use `bucket-sort` or `counting-sort` to sort points as if they were floats. Below is the necessary code to do that:

Point structure definitions:

```c
typedef struct point {
    printable printable;
} point;

float atanxy(point* p) {
    pair coords = (pair)((printable*)p)->val;
    float x = *(float*)coords->first->val;
    float y = *(float*)coords->last->val;
    if (x == 0.0) return 1;
    if (y == 0.0) return 0;
    return y / sqrt(x * x + y * y + x);
}

char* point_to_string(point* p) {
    size_t size = ((printable*)p)->size;
    char* result = ALLOCATE(sizeof(char) * size);
    pair coords = (pair)((printable*)p)->val;

    snprintf(result, size, "{x: %.4f, y: %.4f, t: %.4f}",
             *(float*)coords->first->val,
             *(float*)coords->last->val,
             atanxy(p));
    return result;
}
```

Auxilary functions:

```c
printable* random_point_element_generator(void* elt) {
    point* result = ALLOCATE(sizeof(point));
    printable* presult = (printable*)result;
    pair coords = make_pair();
    float x, y;

    x = (float)rand() / (float)RAND_MAX;
    y = (float)rand() / (float)RAND_MAX;
    coords->first = (printable*)make_printable_float(x);
    coords->last = (printable*)make_printable_float(y);
    presult->val = ALLOCATE(sizeof(pair));
    presult->val = coords;
    presult->to_string = (printer)point_to_string;
    presult->size = (3 + FLT_DIG) * 3 + 16;
    return presult;
}

printable* printable_atanxy(printable* p) {
    point* pt = (point*)p;
    return (printable*)make_printable_float(atanxy(pt));
}
```

Callbacks for sorting algorithms:

```
int compare_points(const void* a, const void* b) {
    point* pa = *(point**)a;
    point* pb = *(point**)b;
    if (pa == pb) return 0;
    if (pa == NULL) return -1;
    if (pb == NULL) return 1;
    float fpa = atanxy(pa);
    float fpb = atanxy(pb);
    return (fpa > fpb) - (fpa < fpb);
}

size_t rationalize_point(printable* elt, printable* min,
                         printable* max, size_t range) {
    float n = atanxy((point*)min);
    float x = atanxy((point*)max);
    float e = atanxy((point*)elt);
    if (x == n) return 0;
    return (size_t)(range * ((e - n) / (x - n)));
}
```

Finally, example usage:

```
int main() {
    time_t t;
    srand((unsigned)time(&t));
    array test = make_random_array(
        13, 3, 97, random_point_element_generator);
    printf("Generated points array:\n%s\n",
           to_string((printable*)test));
    bucket_sort(test, rationalize_point, compare_points);
    printf("Sorted points array:\n%s\n",
           to_string((printable*)test));
    array tans = array_map(test, printable_atanxy);
    printf("Tangents:\n%s\n", to_string((printable*)tans));
    return 0;
}
```

For implementation of `array_map`, `make_random_array`, `bucket_sort` and `to_string` please see library code.

Example output:

```
Generated points array:
[{x: 0.8257, y: 0.8677, t: 0.7606}, {x: 0.1002, y: 0.3438, t: 0.2836},
{x: 0.9552, y: 0.9093, t: 0.8100}, {x: 0.0602, y: 0.3771, t: 0.1584},
{x: 0.6079, y: 0.8967, t: 0.5957}, {x: 0.2550, y: 0.9156, t: 0.2716},
{x: 0.6398, y: 0.0718, t: 1.4590}, {x: 0.7224, y: 0.6434, t: 0.8432},
{x: 0.9814, y: 0.6833, t: 0.9626}, {x: 0.7193, y: 0.7053, t: 0.7952},
{x: 0.3380, y: 0.5310, t: 0.5668}, {x: 0.4476, y: 0.6312, t: 0.6168},
{x: 0.3246, y: 0.5864, t: 0.5055}]

Sorted points array:
[{y: 0.3771, x: 0.0602, t: 0.1584}, {y: 0.9156, x: 0.255, t: 0.2716},
{y: 0.3438, x: 0.1002, t: 0.2836}, {y: 0.5864, x: 0.3246, t: 0.5055},
{y: 0.531, x: 0.338, t: 0.5668}, {y: 0.8967, x: 0.6079, t: 0.5957},
{y: 0.6312, x: 0.4476, t: 0.6168}, {y: 0.8677, x: 0.8257, t: 0.7606},
{y: 0.7053, x: 0.7193, t: 0.7952}, {y: 0.9093, x: 0.9552, t: 0.81},
{y: 0.6434, x: 0.7224, t: 0.8432}, {y: 0.6833, x: 0.9814, t: 0.9626},
{y: 0.0718, x: 0.6398, t: 1.459}]
```