Randomized algorithm

Tutorial 1

Joyce

2009-09-29

Introduction
- Monte Carlo methods
- Applications

Hints for assignment 1
- Question 1
- Question 2
- Question 3
- Question 4
- Question 5
- Question 6
- Question 7

Interesting quiz
- Shotgun game
- Shortest distance problem

Monte Carlo methods

The roots of randomized algorithms can be traced back to Monte Carlo methods used in numerical analysis, statistical physics, and simulation. <Rajeev Motwani>

Rajeev Motwani
1962/03/26-2009/06/05

[Monte Carlo methods]
John von Neumann and Stanislaw Ulam suggested that some problems in Manhattan Project can be solved by modeling the experiment on a computer using chance. The code name of this work is "Monte Carlo". <Wikipedia>

Stanislaw Ulam
1909/04/13-1984/05/13

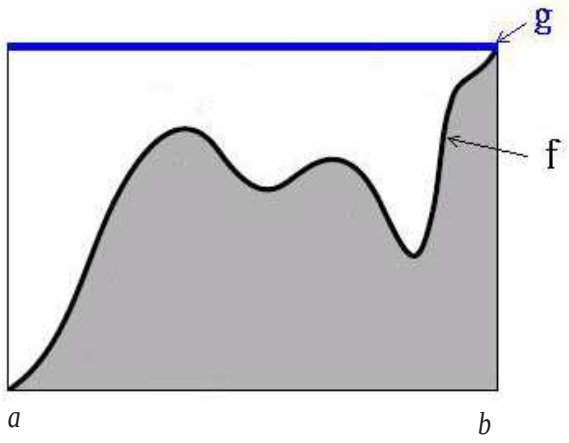
(a)

John von Neumann
1903/12/28-1957/02/08

(b)

"Hit and miss" integration is the simplest type of MC method to understand. <Joy Woller>

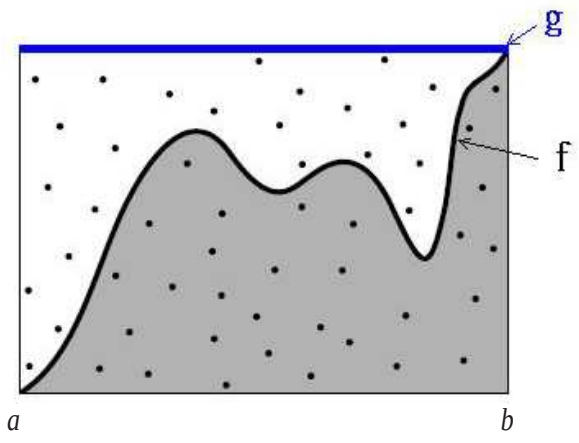
Let us introduce an example. Suppose we are given two functions $f(x)$ and $g(x)$. Now we want to get the area of the grey part. How?



Let I be the area, a be the left endpoint and b be right endpoints.

$$I = \int_a^b f(x)dx$$

We may choose n points in $g(x)$ like we did the figure.



Divide segment ab into n parts. x_1 means the leftmost part including a.
$I \sim \frac{b-a}{n} \sum_{i=1}^{n} f(x_i)dx$

[Applications]

1. Quicksort
2. Min-cut algorithm
3. Perfect matching in graphs
4. Pattern matching
5. Geometric algorithms, Data Structures, Approximate counting, etc..

You can find most of these applications in the book "Randomized Algorithms" written by Rajeev Motwani and Prabhakar Raghavan.

Hints for assignment 1

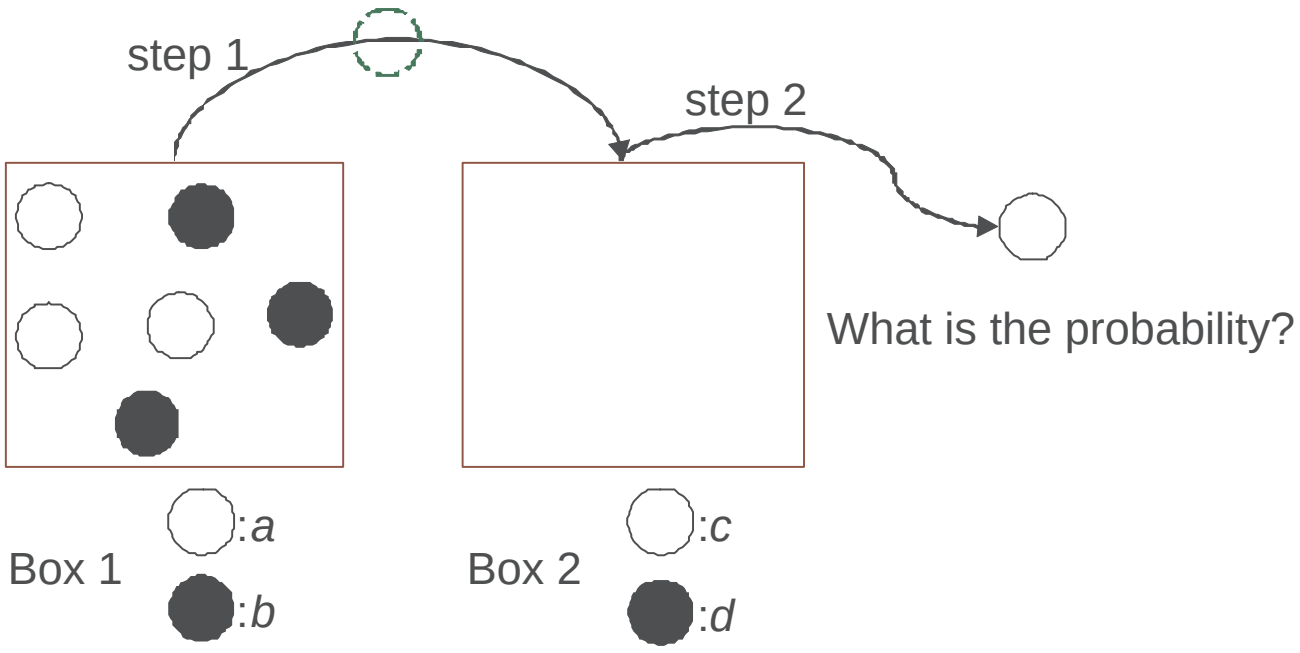
[Question 1]:
A biased coin is tossed till a head appears for the first time. What is the probability that the number of required tosses is odd?

[Hint]:
Let $A_i =$ "Head appears at the ih toss for the first time.

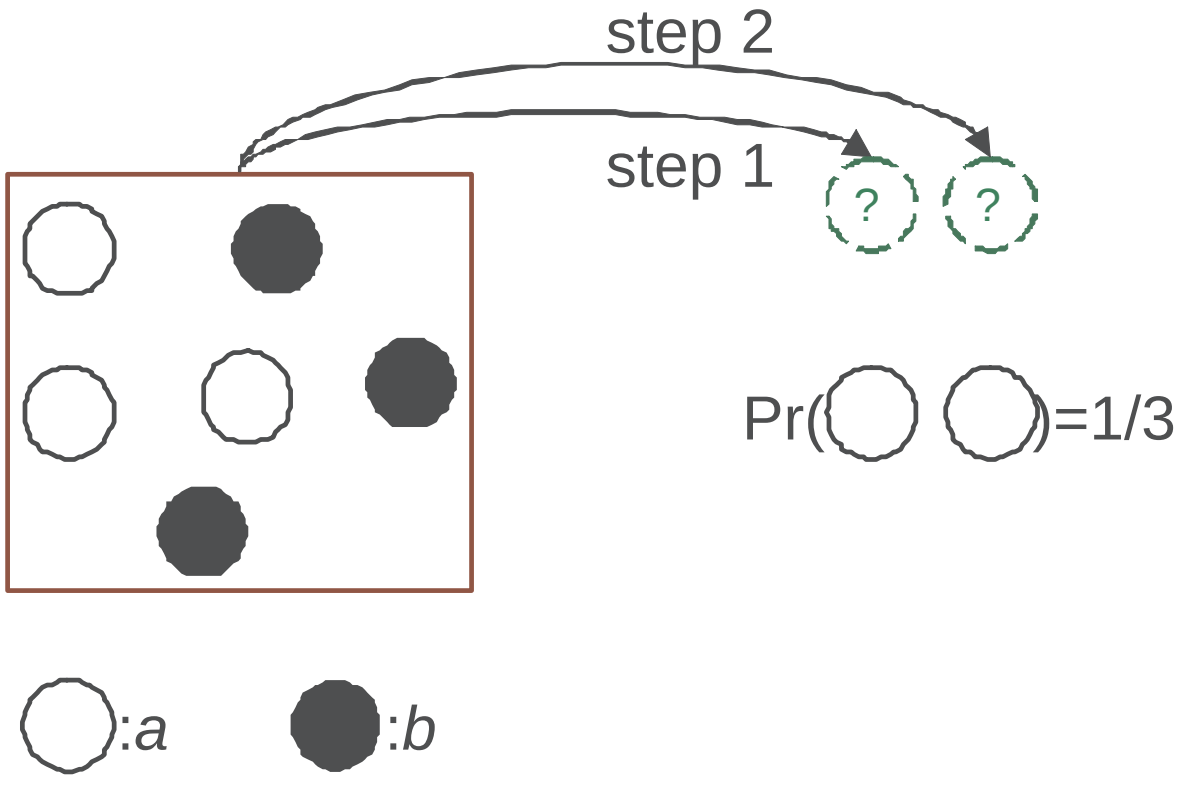
$$P(A_i) = P(T)P(T)P(T)\ldots P(T)P(H) = q^{i-1}q$$

where $P(H) = p$, $P(T) = q = 1 - p$

[Question 2]:



[Question 3]:

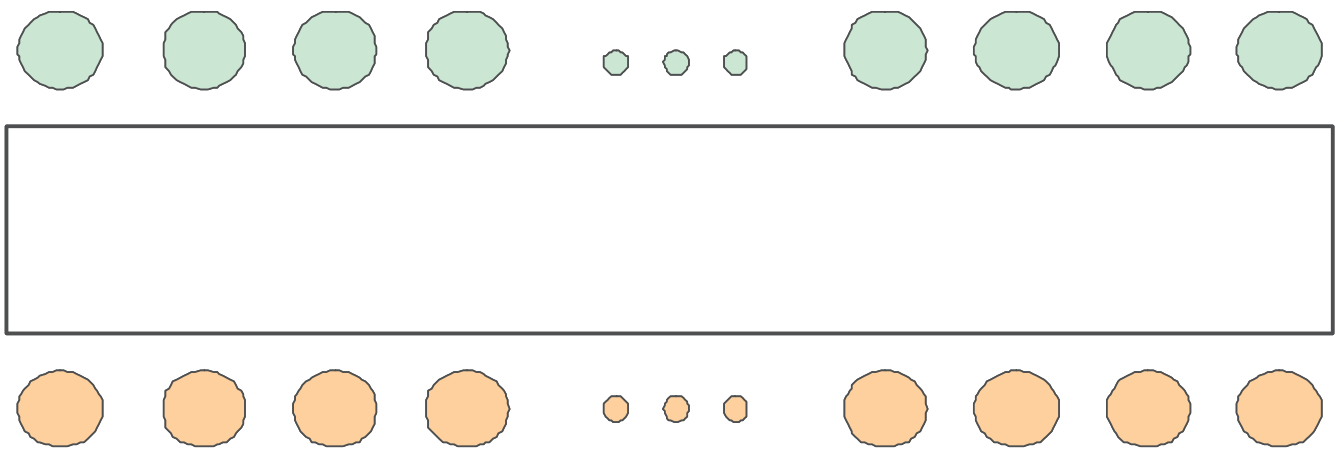


[Hint]:
The probability we get a white ball at the first time is $\frac{a}{a+b}$.
Tips: Conditional probability.

[Question 4]:
An inversion in a array, A, is defined as a pair of elements in A such that if there is a pair $(i, j), i < j$ and $k[i] > k[j]$. In bubble sort, we fix such inversions by swapping. What is the expected number of swaps in bubble sort?

[Hint]:
You may do this without considering permutation.
Tips: Indicator variable, Linearity of expectation.

[Question 5]:
20 couples are invited to a party. They are asked to be seated at a long table with 20 seats each side with husbands sitting at one side and wives the other side.
If the seating is done at random, what is the expected number of married couples that are seated face to face?



[Hint]:
Tips: Indicator variable, Linearity of expectation.

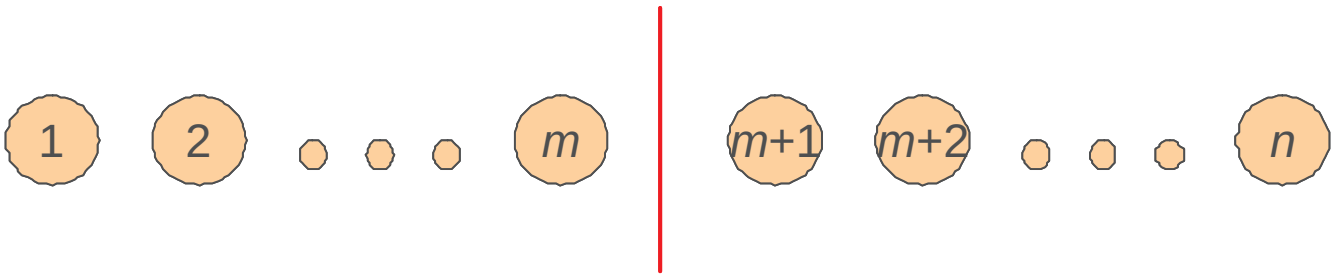
[Question 6]:
Let X and Y be independent geometric random variables, where X has parameter p and Y has parameter q.

1. What is the probability $\Pr(\min(X, Y) = k)$?
2. What is $\mathrm{E}[X \mid X \leq Y]$?
3. What is the probability that $X = Y$?
4. What is $\mathrm{E}[\max(X, Y)]$?

[Hint]:
Tips: Memory less property

[Question 7]:
You need a new staff assistant, and you have n people to interview. You want to hire the best candidate for this position. When you interview the candidates, you can give each of them a score, with the highest score will be the best and no ties being possible.

1. First, we give everyone a number card. The number card means the order of interview.



2. Interview first m people and choose the best grade from them as A.
3. Interview follow the ordering.
   - If someone got a better score than A, accept him.
4. After interviewing all remaining candidates but no one get better score, we choose the last one.

[Hint]:
In which cases, we may not get the best candidate?

- The best ones number card is less than $m + 1$.
- A, the best grade we chose from 1 to m, is not too good.

Thus, we need a nice m.
What is the "nice" m?
The second best candidate is in 1 to m.

[Hint]:
In which cases, we may not get the best candidate?

- The best ones number card is less than $m + 1$.
- A, the best grade we chose from 1 to m, is not too good.

Thus, we need a nice m.
What is the "nice" m?
The second best candidate is in 1 to m.

[Hint]:
In which cases, we may not get the best candidate?

- The best ones number card is less than $m + 1$.
- A, the best grade we chose from 1 to m, is not too good.

Thus, we need a nice m.
What is the "nice" m?
The second best candidate is in 1 to m.

Interesting quiz

There is a contest call "SHOTGUN". You've got into the last stage and your opponents are Robinhood and LongWu.
The rules are as follows:

1. Before the game, the ordering was given.
2. Each person will get to fire one shot sequentially.
3. Anybody who got hit is out and stops participating.

Follow the order of shots, the last one person who is alive is the winner.

Here we got a hitting ratio according past games.

LongWu
Record: 50/50
100%

Robinhood
Record: 40/50
80%

?

Record: 20/50
40%

It seems that you are already dead in this game.

Here we got a hitting ratio according past games.

LongWu
Record: 50/50
100%

Robinhood
Record: 40/50
80%

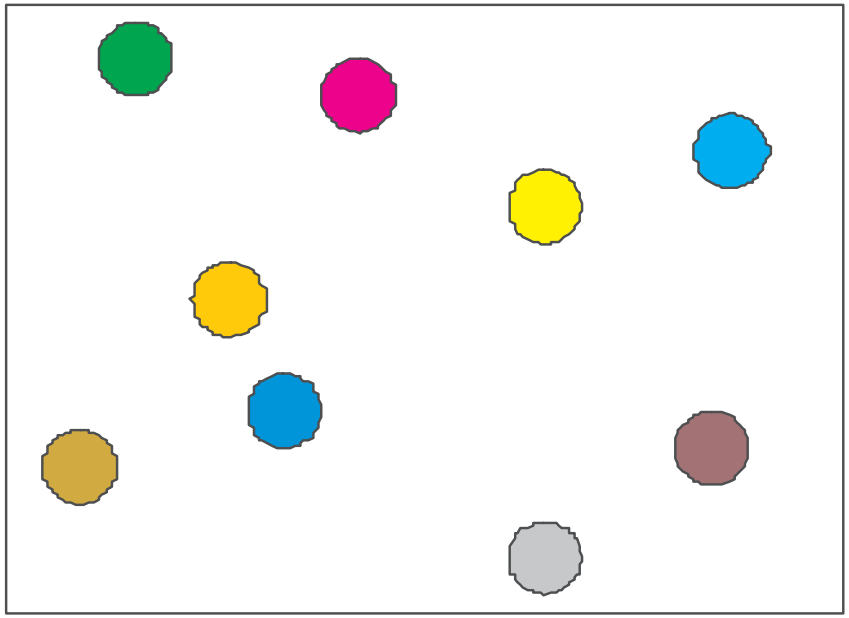
?

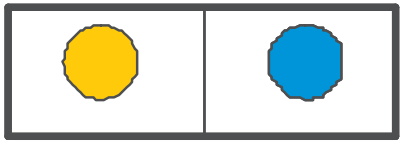
Record: 20/50
40%

It seems that you are already dead in this game.

But, you are so lucky that you are get the order of 1.
Now suppose all of you will us an optimal strategy.
Prove that your best strategy involves shooting into the air on your first shot; and, calculate each players probability of winning.

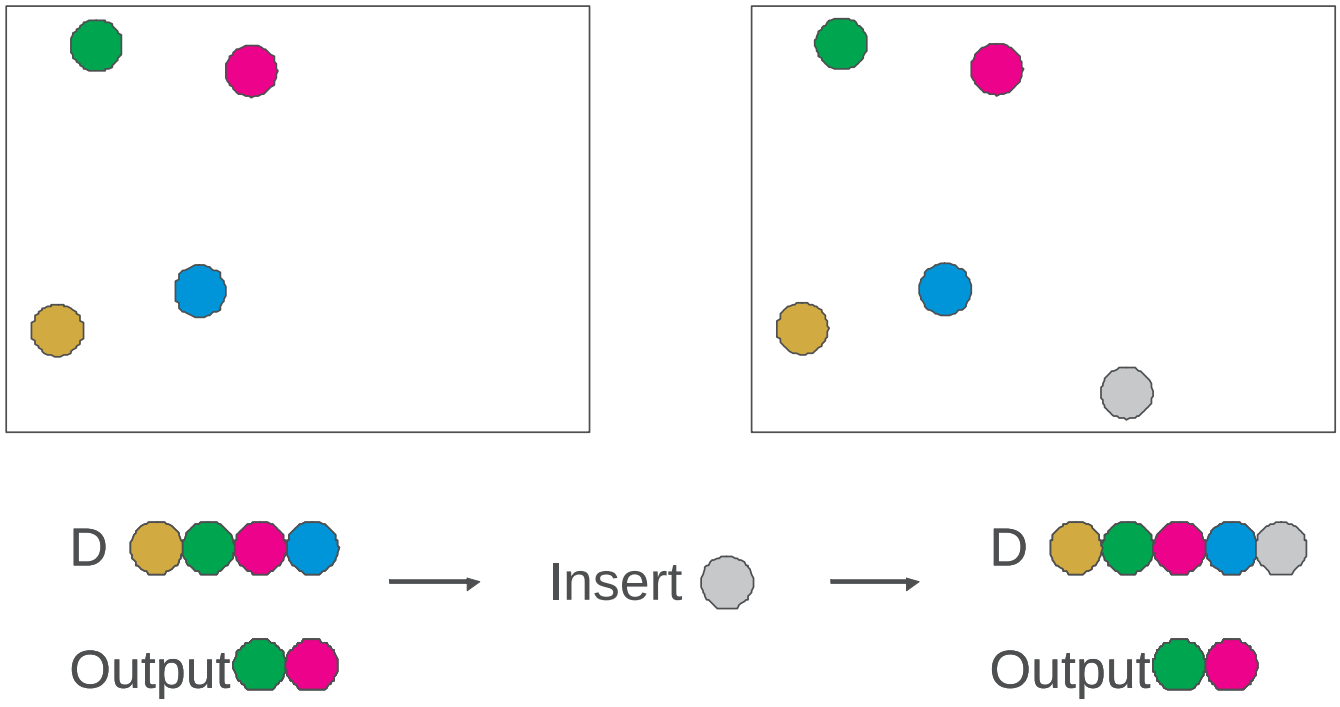
Determine the shortest distance between a pair of points in the array. (The points are in 2-d)



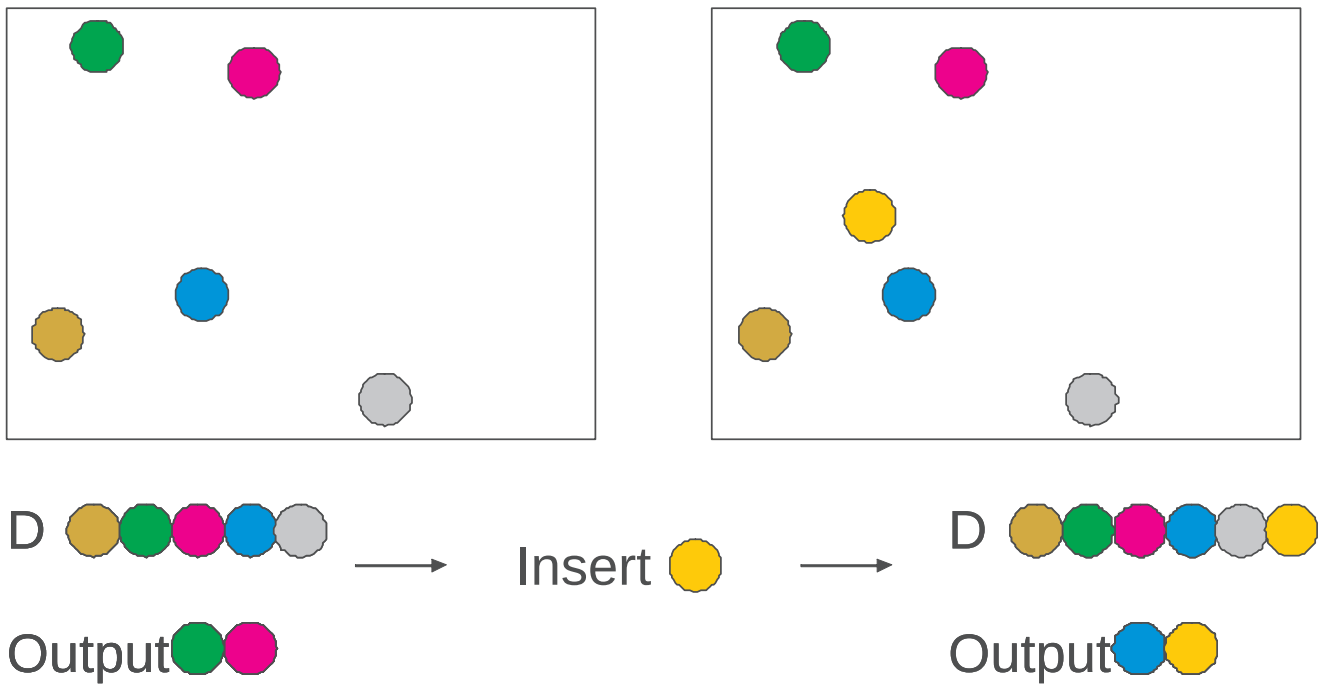


Now we have a storage data structure D.
Each time we insert a point.
When we give a point to D, it stores the point and answers the shortest distance of all points in D.
The time D takes depends on whether the answer change or not.

If output is the same: 1 clock tick.



If output is not the same: $|D|$ clock tick.



Show that the expected total number of clock ticks used by D is $O(n)$.

[Hint]
Let X_i be the clock ticks when inserting the ith point and X be the total clock ticks.

1. What is the probability that the ith point causes answer to change?
2. What is $\mathrm{E}[X_i]$?

[Solution]

1. Pr(ith point causes answer to change)
   $=$ Pr(ith point is one of the shortest pair among the first i points)
   $= 2/i$ –(why?)
2. E$[X_i]$
   $= i * 2/i + 1 - 2/i$
   < 3
3. E$[X] = 1 + \Sigma$E$[X_i] = O(n)$

Thank you