

Final Week Project and Report (2018)

In the final week of COSMOS, we would like you to do a short project and write a report on the Computational Biophysics part. In the report, you need to first state and answer the following two questions(individually):

[1.] Programming: In this cluster, you have written programs in both C and python. Which language do you prefer to program in based on your experience to date (C, python, no preference)? Why do you make that choice? How do C and python compare to each other or to another language you may know?

[2.] Biophysics: Explain, in what sense, is the following statement true: ‘Everything that animals do, atoms do.’

Complete a miniproject on a problem of your choice (can be team effort for up to three people). Your problem should involve a random walk simulation program to experimentally explore an hypothesis you have. Below are listed samples of projects from past years, which you take as is, or modify. Talk to one of us if you want to do ”something completely different” (in the words of Monty Python).

- A. Study quantitatively the Gambler’s ruin problem.
 - (i) For $p=0.5$, show that probability of Gambler 1 winning is $M_1/(M_1+M_2)$. Make a suitable plot to illustrate your result
 - (ii) For $p=0.5$, show that average number of steps to ruin is M_1*M_2 . Make a suitable plot to illustrate your result.
 - (iii) Make a histogram for the number of steps till ruin for a few different values of M_1 and M_2 .
 - (iv) Study the Gambler’s ruin problem for p different from 0.5, with $M_1=M_2=M$. Make a plot of winning probability for Gambler 1 or 2 as a function of p for several different values of M . Allow p to range from 0 to 1. Make the same plot for fixed p as a function of M . Can you guess the dependence on M for large M ?
- B. Study the distribution of one and two (and possibly three) dimensional random walks as a function of number of steps:
 - (i) What is the mean and standard deviation of the position of the walker after N steps? Make a plot versus N for different dimensions.
 - (ii) Make a plot of many walks in two or three dimension. Does the resulting shape resemble something?
 - (iii) Make a histogram of the distance traveled after N steps. Show the histogram for several N (say 10, 50, 100, 500, 1000) in one plot.
- C. Study and plot target-finding inside 2D or 3D cells. Analyze how average number of steps to target changes based on varying a single parameter and leaving all other choices the same. For example, you might select a specific shape and size(volume) for the boundary and target and then vary the distance from the start of the random walk to the target. Then graph the average number of steps vs distance from target and draw conclusion. Other single parameters you may consider studying are:

- (i) size of cell (same shape , different volume),
- (ii) shape of cell(different shapes, same volume)
- (iii) shape of target,
- (iv) size of target,
- (v) rule used for random walk moves (grid vs angle) *etc..*

Remember to vary only one parameter at a time in a study; otherwise, it will be difficult to draw meaningful conclusions.

- D. Pass the broccoli random walk

Suppose there are $n + 1$ people, $0, 1, \dots, n$, sitting in a circle. Person 0 has a big stalk of nutritious broccoli, which provides 250% of the US recommended daily allowance of vitamin C and is also a good source of vitamin A and iron. Person 0 passes the broccoli either to the person on his left or the person on his right with equal probability. Then, that person also passes the broccoli left or right at random and so on. After a while, everyone in an arc of the circle has touched the broccoli and everyone outside that arc has not. Eventually, the arc grows until all but one person has touched the broccoli. That final person is declared the winner because they have avoided the broccoli for the longest time.

Suppose that you are allowed to position yourself anywhere in the circle. Where should you stand in order to maximize the probability that you win? You shouldn't be person 0; you can't win in that position.

Make a reasoned guess about what position you would take to win. Then devise a random walk experiment to test your assumption, that is write a program to explore the likelihood of winning based on different positions on the circle and the average number of steps it takes to declare a winner. Then write up your conclusions.

- E. Random walks on a circular DNA (Finding origin of replication).

Generate a (circular) list of 4,000-10,000 of random DNA base pairs ('A','T','G','C'). Insert a new "primer" string of DNA: 'AAAAAAAAAA' into the list above at a random location. This will represent the origin of replication or the point at which DNA cloning will start. Investigate the best strategy (minimum number of steps) for the DNA polymerase to find the origin of replication in the larger DNA string. Possible strategies include (a) single step walks starting from random location or (b) walks with steps of arbitrary length (randomly jumping around list) or (c) some combination of a and b (your choice). Be careful with indexes so you don't overrun the start or end of the DNA string.

Present a brief writeup on your project, with suitable figures and conclusions. Attach your program to the project report and explain key lines of the code. Use graphs to make your report and conclusion visually compelling.