Noam Rozov
Daniel Schreiber 06.06.2022

Biological Computation – Homework 1 Solution

Preparation

> Note: You can discuss the work with other students however you should write all code by yourself / with your exercise partner. Please document your code and also prepare a short (no more than 2 pages) explanation on how the code works and what were the main design and implementation decisions you made. If any code that a student did not write is used, explain where it is taken from and why it's needed.
>
> *In addition to submitting to Moodle prepare a GitHub repository with a readme on how to run the code.

Question 1

> 1) a) Write a program (in your favorite programming language) that gets as input a positive integer n and generates all connected sub-graphs of size n.
> The output should be a textual file of the following form:
>
> n=2
> count=2
> #1
> 1 2
> #2
> 1 2
> 2 1
>
> The first two lines output n and the total number (count) of different sub-graphs of size n. Then the sub-graphs themselves are given each starting with a line labelled #k for motif number followed by all edges, each line i j means an edge from source i to target j.
>
> b) Output the result of your program for n = 1 to 4.
>
> c) What is the maximal number n for which your program can complete successfully within no more than 1 hour of computing time?
>
> d) What is the maximal number n for which your program cancomplete successfully within 2,4,8 hours of computing time?

Answer 1

Our program works by generating all possible binary matrices (by creating all possible vectors of size n^2, recursively, and then constructing the $n \times n$ matrices). Then removing all matrices which have 1's on their diagonal (since we do not want vertices to have self-edges). After that we use DFS to get the matrices which are connected (by constructing undirected graphs and DFS'ing on them). Finally, we go over the remaining matrices, choose a representative for each isomorphic group, and return the representatives. In this section we used the $is\_isomorphic$ function from the networkx package.

Noam Rozov 06.06.2022
Daniel Schreiber

After running our program using Python

```
Please type '1' to start the solution for the first part of the homework.
Please type '2' to start the solution for the second part of the homework.
1
please enter n:
4
Press any key to continue . . .
```

we have the following output being generated under $text\_files\backslash q1\_output.txt$

#1:

1 2

1 3

1 4

#2:

1 3

1 4

2 1

...

#199:

1 2

1 3

1 4

2 1

2 3

2 4

3 1

3 2

3 4

4 1

4 2

4 3

Noam Rozov
Daniel Schreiber

06.06.2022

Question 2

2) Write a program that gets as input positive integer n and a graph of the format:

1 2

2 3

1 4

The graph in the example contains 4 vertices 1,2,3,4 and directed edges (1,2) (2,3) (1,4). The program should output all sub-graphs of size n and count how many instances appear of each motif. The format of the output of the identified sub-graphs should be like in question 1, where in the line after #k should appear the count of number of instances, count=m if the motif appears m times. Output count=0 if a motif does not appear in the graph.

Answer 2

In this section we generated all sub graphs of size n using the built-in $itertools.combinations$ class which returns successive r-length combinations of elements in an iterable, for example: $combinations(range(4), 3) \ --> \ (0,1,2), (0,1,3), (0,2,3), (1,2,3).$

Then, we also generated all motifs of size n, using the answer to question 1.

Finally, for each subgraph, we incremented a counter (per motif) each time the subgraph was isomorphic to one of the motifs.

After running our program using Python

```
Please type '1' to start the solution for the first part of the homework.
Please type '2' to start the solution for the second part of the homework.
2
Please enter n
3
Please enter graph edges and then type 'DONE'
1 2
2 3
1 4
DONE
Press any key to continue . . .
```

we have the following output being generated under $text\_files\backslash q2\_output.txt$

#1:

1 2

1 3

count=1

#2:

1 3

2 1

Noam Rozov
Daniel Schreiber 06.06.2022

count=1

#3:

1 2

1 3

2 1

count=0

#4:

1 3

2 3

count=0

#5:

1 2

1 3

2 3

count=0

#6:

1 2

1 3

2 1

2 3

count=0

#7:

1 2

2 1

3 1

count=0

#8:

1 2

1 3

2 1

3 1

count=0

#9:

1 2

2 3

3 1

count=0

#10:

1 2

1 3

2 3

3 1

count=0

#11:

1 3

2 1

2 3

3 1

count=0

#12:

1 2

1 3

2 1

2 3

3 1

count=0

#13:

1 2

1 3

2 1

2 3

3 1

Noam Rozov
Daniel Schreiber

06.06.2022

3 2

count=0