



Strathmore
UNIVERSITY

STRATHMORE INSTITUTE OF MATHEMATICAL SCIENCES
END OF SEMESTER EXAMINATION
MASTER OF SCIENCE IN BIOMATHEMATICS

BMA 8305: METHODS & MODELS IN GENOMICS

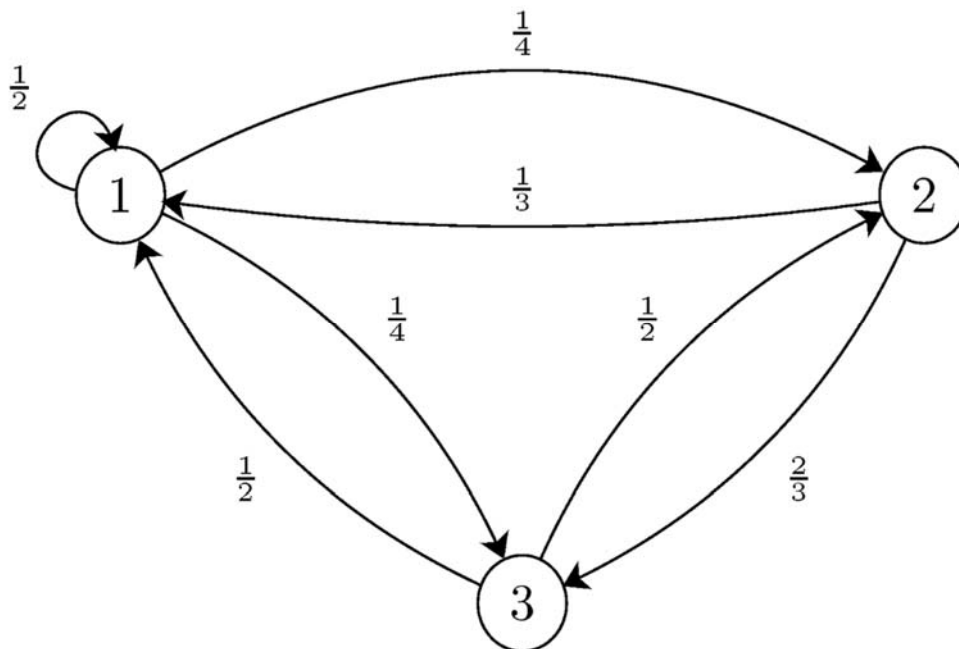
Date: 18th April, 2024

Time: 3 Hours

Instruction: Answer Question one and any other two questions

Question ONE (30 Marks)

- a. Which of the following is true about a codon? (1 mark)
- It is a group of three nucleotides in mRNA that codes for an amino acid.
 - It is a group of three nucleotides in tRNA that codes for an amino acid.
 - It is a group of three nucleotides in DNA that codes for an amino acid.
 - It is a group of three amino acids that form a protein.
- b. If the sequence of one polynucleotide of a DNA double helix is 5'-ATAGCAATGCAA-3', what is the sequence of the complementary polynucleotide? (2 marks)
- c. How is the genetic code decoded during translation? (4 marks)
- d. Briefly describe the types of sequence alignment in searching for common character patterns in genomic research (6 marks)
- e. Briefly describe the databases and tools used in bioinformatics, and how do they facilitate biological research? (6 marks)
- f. What do you understand by the central dogma of molecular biology? (4 marks)
- g. Consider the Markov chain shown in the figure below



- i. Is this chain irreducible? (1 mark)
- ii. Is this chain aperiodic? (1 mark)
- iii. Find the stationary distribution for this chain. (3 marks)
- iv. Is the stationary distribution a limiting distribution for the chain? (1 mark)

Question TWO (15 marks)

Define an HMM \mathcal{H} with three states $\{A, B, C\}$ and alphabet $\{0, 1, 2\}$. The initial state probabilities are $\pi_A = 1$ and $\pi_B = \pi_C = 0$. The transition and emission probabilities are as follows:

	A	B	C	0	1	2
A	0.2	0.8	0.0	0.8	0.2	0.0
B	0.0	0.8	0.2	0.0	0.6	0.4
C	0.4	0.0	0.6	0.2	0.0	0.8

- (a) Draw the state diagram of this HMM and show the transition probabilities. (3 Marks)
- (b) Give all state paths with non-zero probability for the sequence $O = 0, 1, 2, 0$. (3 Marks)
- (c) What is $P(O)$? (Use the brute force approach, not the Forward algorithm.) (5 Marks)
- (d) What is the most probable path, Q^* ? What is $P(O, Q^*)$, the probability of transitioning through this path and emitting O ? (Again, you do not need to use the Viterbi algorithm here.) (3 Marks)

Question THREE (15 marks)

a. Elaborate what is genome annotation, and what information does it provide?
(10 marks)

b. Explain the importance of genome visualization tools in genomics research.
(5 marks)

Question FOUR (15 marks)

- a. Describe the process of genome assembly in genome sequencing. (5 marks)
- b. Hidden Markov models (HMM) are used to identify genes in genome sequencing projects. Describe how you would build a hidden Markov model to identify genes in a genome sequence. (9 marks)
- c. Give one other application of hidden Markov models. (1 marks)