ENGINEERING TRIPOS PART IIB

---

Thursday 1 May 2008 9 to 10.30

---

Module 4G1

COMPUTATIONAL AND SYSTEMS BIOLOGY

*Answer **all** three questions.*

*All questions carry the same number of marks.*

*The **approximate** percentage of marks allocated to each part of a question is indicated in the right margin.*

*There are no attachments.*

STATIONERY REQUIREMENTS
Single-sided script paper

SPECIAL REQUIREMENTS
Engineering Data Book
CUED approved calculator allowed

**You may not start to read the questions printed on the subsequent pages of this question paper until instructed that you may do so by the Invigilator**

1 (a) Show that the complexity of the global alignment between two DNA sequences which are quite similar may be lower than the complexity of the Needleman Wunch algorithm. [35%]

(b) How does this relate to database searching algorithms? [30%]

(c) Discuss the complexity of the Sankoff algorithm for parsimony. [35%]

2 (a) Describe how you would set up a two-colour microarray experiment to find genes that are differentially expressed between animals in two different experimental conditions (such as wild-type versus a mutant mouse). Your answer should include details of how samples are prepared and how the microarray images are created. [35%]

(b) How would you analyse the images collected in Part (a) to find differentially-expressed genes? Include an explanation of how you would account for (i) noise in samples and (ii) the problem of multiple-comparisons. [35%]

(c) Give two examples of how clustering can be applied to gene expression datasets. What might the results tell us in each case? [30%]

(TURN OVER

3 (a) Consider a birth-death process described by

$$x \xrightarrow{\lambda} x+r, \quad x \xrightarrow{\beta x} x-1\ ,$$

where jump size $r > 0$ is a positive integer, $\lambda, \beta$ are positive constants and $x$ is the population size.

(i) Write down the master equation for the probability of $x$ taking value $k$ at time $t$. [20%]

(ii) Derive using the master equation a differential equation for the mean $\langle x \rangle$. Show detailed steps in your answer. [25%]

(iii) Using the differential equation for $\langle x^2 \rangle$ given below

$$\frac{d\langle x^2 \rangle}{dt} = \lambda r^2 - 2\beta \langle x^2 \rangle + (2\lambda r + \beta)\langle x \rangle$$

and your answer in Part (a)(ii), derive a differential equation for the variance of $x$. Hence write down an expression for the ratio $\frac{\sigma_x^2}{\langle x \rangle}$ at equilibrium ($\sigma_x^2$ denotes the variance of $x$). [25%]

(b) (i) Consider a gene (of concentration $y$) which is turned on and off with rates

$$\text{OFF} \underset{k_2 y}{\overset{k_1}{\rightleftharpoons}} \text{ON}\ .$$

Show that at equilibrium the probability of the ON state is $P(\text{ON}) = \frac{K}{K+y}$ where $K$ is a constant to be determined. [10%]

(ii) The birth-death process below is used as a model for the expression of this gene

$$y \xrightarrow{\frac{K}{K+y}} y+1, \quad y \xrightarrow{\beta y} y-1$$

By deriving a differential equation for the mean $\langle y \rangle$ of this process, determine if the differential equation $\dot{m} = \frac{K}{K+m} - \beta m$ is satisfied by $\langle y \rangle$. [20%]

**END OF PAPER**