ABSTRACT

AN INVESTIGATION OF TECHNIQUES FOR THE REVERSE ENGINEERING OF SIMPLE GENETIC NETWORKS; CONNECTED INFORMATION AND GENETIC AND EVOLUTIONARY ALGORITHMS

Hugh Anthony Quintyne Atherley

Using simple models techniques are investigated which will allow the reverse engineering of the genetic pathways using microarray data. While the main focus in the reverse engineering process was the use of genetic and evolutionary algorithms, techniques developed in information theory were explored. The behaviour of the simple genetic algorithm was tested using a discrete gene network model (DGM) and the continuous gene network model (CGM). Data from each model was generated using a small number of genes (6) in order to test the algorithm. The discrete model posed little problems for the algorithms because of its limited nature. However it introduced the problem of non-uniqueness of solutions, where various networks were produced, which fit the transition data and network rules. An attempt was made to alleviate the issue with the development of the continuous model. The algorithm was again able to find solutions given the transition data. However there was no guarantee that a solution would be found for each run within the maximum generation terminating point. Given the issues of non-uniqueness for each model, other methods must be used to solve this problem. One possible method is the use of connected information. Methods in information theory are

being developed to discern the relative mixture of the different types of interactions in simple networks. By calculating the information uncovered or entropy lost by introducing each type of interaction researchers can discover the types of interactions in the structure. One such measure is the connected information, which gives the difference in the maximum entropy for k and $(k-1)$ interactions. Novel methodology is described which enables this calculation to be carried out more easily.

Keywords: Entropy; Genetic Algorithm; Microarrays; Connected Information.