

## **Computational Methods for Cluster Detection of Patterns in Biological Sequences**

W.Y. Wendy Lou

University of Toronto, Toronto, ON, Canada M5S 3M7.

### **Abstract**

The detection of clusters of patterns, either identical or non-identical, in a long biological sequence is often of interest in establishing genetic signals. The probability of occurrence of a cluster (event) for a given sequence, assuming some underlying distribution for the bases (outcomes) that make up the biological sequence, can be used as a tool for this purpose. In this talk, a motivating example involving the detection of palindrome clusters in DNA sequences will be first presented, followed by methods we propose to achieve the computational objectives. The total number of palindrome clusters, defined through a general concept of runs and patterns, is finite Markov chain imbeddable in the sense of Fu and Koutras (1994). Numerical examples will be given to illustrate our results.