

# COMPUTER SCIENCE SEMINAR SERIES

Department of Computer Science  
City University of Hong Kong  
(Departmental Seminar Seminar 2006/2007 - No 68)

## Identification of Distinguishing Motifs

Mr WANG Zhanyong  
MPhil Student  
Department of Computer Science  
City University of Hong Kong

**Date :** 26 July 2007 (Thursday)

**Time :** 1:30pm - 2:30pm

**Venue :** CS Seminar Room, Rm Y6405, 6th Floor Yellow Zone, Academic Building, City University of Hong Kong, Tat Chee Avenue, Kowloon Tong.

### Abstract

Motivation: Motif identification for sequences has many important applications in biological studies, e.g., diagnostic probe design, locating binding sites and regulatory signals, and potential drug target identification. There are two versions. 1. Single Group: Given a group of  $n$  sequences, find a length- $l$  motif that appears in each of the given sequences and those occurrences of the motif are similar. 2. Two Groups: Given two groups of sequences  $B$  and  $G$ , find a length- $l$  (distinguishing) motif that appears in every sequence in  $B$  and does not appear in anywhere of the sequences in  $G$ . Here the occurrences of the motif in the given sequences have errors. Currently, most of existing programs can only handle the case of single group. Moreover, it is very difficult to use edit distance (allowing indels and replacements) for motif detection. In this paper, (1) We propose a randomized algorithm for the one group problem that can handle indels in the occurrences of the motif. (2) We give an algorithm for the two groups problem. (3) Extensive simulations have been done to evaluate the algorithms.

This paper is presented in the 18th Annual Symposium on Combinatorial Pattern Matching (CPM 2007), July 9 - 11, 2007

Supervisor: Dr Lusheng Wang (CS)  
Research Interests: Bioinformatics, algorithm

**All are welcome!**

\* \* \* \* \*

*In case of questions, please contact Dr Lusheng Wang at Tel: 2788 9820, E-mail: [cswangl@cityu.edu.hk](mailto:cswangl@cityu.edu.hk), or visit the CS Departmental Seminar Web at <http://www.cs.cityu.edu.hk/>.*