UNIVERSITY OF SYDNEY

SCHOOL OF MATHEMATICS AND STATISTICS

Statistics Seminar

Friday 1 November, 2.00 p.m.

Lecture Room 173 (Level 1, Carslaw Building)

Biological sequence analysis

Professor Terry Speed
The Walter & Eliza Hall Institute of Medical Research
Royal Melbourne Hospital

Abstract

This talk will review a little over a decade's research on applying certain stochastic models to biological sequence analysis. The models themselves have a longer history, going back in some forms over 30 years, although many novel variants have arisen since that time. The function of the models in biological sequence analysis is to summarize the information concerning what is known as a motif or a domain in bioinformatics, and to provide a tool for discovering instances of that motif or domain in a separate sequence segment. We will introduce the motif models in stages, beginning from very simple, non-stochastic versions, progressively becoming more complex, until we reach modern profile HMMs for motifs. A second example will come from gene finding using sequence data from one or two species, where generalized HMMs or generalized pair HMMs have proved to be very effective.

M. P. Quine 9351 3012