



INTERNATIONAL CENTRE
OF EXCELLENCE FOR
EDUCATION IN
MATHEMATICS

Statistical methods in an area of mathematical physics
Alan Huang, School of Mathematics and Statistics, University of Sydney

Given genetic sequences for a number of species of animals, a phylogenetic tree was reconstructed using a maximum likelihood approach. A phylogenetic tree is similar to a "family" tree, where each pair of species (known as end nodes) are joined to a common ancestor (known as a internal node) and then all internal nodes are joined to each other until all species have been linked.

A computer program was written to efficiently and accurately reconstruct these trees, up to the so-called "Felsenstein 94 (F94)" model. The algorithm used was such that it was invariant with respect to the order in which the nodes were joined and the labelling system used.

Throughout my project, I worked through the models that were already existent and examined some of the software available to do such reconstructions. I wrote some short routines to permute the order of node joining and to evaluate the transition/transversion matrix for a given model, which were implemented in the overall software program.

Although I have finished my part of the vacation scholarship, the project is continuing on to models more general than the F94 model, which is the most general model in the software available to date.