

Telegraph: A New Dynamic Programming Template Library.

Ian Holmes † ihh@fruitfly.org
Guy St.C. Slater ‡ guy@ebi.ac.uk
Ewan Birney ‡ birney@ebi.ac.uk
Gerald M. Rubin † gerry@fruitfly.berkeley.edu

† *Berkeley Drosophila Genome Project,*
Rubin Lab, LSA Rm. 539, U.C. Berkeley,
Berkeley, CA 94720-3200

‡ *European Bioinformatics Institute,*
Wellcome Trust Genome Campus,
Hinxton, Cambridge, CB10 1SD, UK

Abstract

Dynamic Programming (DP) is the foundation of a large number of algorithms in bioinformatics, including those used in protein comparison, gene prediction and database search using profile hidden Markov models. Each of these algorithms align sequences to a *Finite State Machine* (FSM), although the architecture of the FSM is different in each case. The transitions of the FSM have associated *weights*, which determine the scoring scheme of the algorithm (*e.g.* gap penalties and substitution matrices). If the weights are normalised, then the FSM is a probabilistic model and the alignment scores may be treated as log-likelihoods. This representation is convenient for *training* the model from data.

The ubiquitous nature of DP in bioinformatics has motivated development of general purpose libraries for specifying algorithms in terms of the underlying FSM [1] Using ideas from several of these libraries (most notably Dynamite [2]) we have built a new object-oriented library for DP in bioinformatics, called Telegraph.

Telegraph offers the following advantages over existing libraries:

- **Scriptable.** The outer interface to the library is written in Perl and integrates smoothly into a Bioperl installation.
- **Intuitive.** The object model combines the specialised *data structure*-oriented features of a language like Dynamite, with *functional programming*-inspired features such as binding of model parameters to user-defined functions.
- **Portable.** The core algorithms have a well-defined ANSI C API, hence many platforms are supported.
- **Probabilistic.** Built-in and integrated support for training algorithms, including Baum-Welch with Dirichlet mixture priors.

This poster illustrates Telegraph using a number of real life examples.

¹David B. Searls and Kevin P. Murphy. *Automata-Theoretic Models of Mutation and Alignment*. Proc. ISMB 1995, pp 341-349

²Ewan Birney and Richard Durbin. *Dynamite: A flexible code generating language for dynamic programming methods used in sequence comparison*. Proc. ISMB 1997, pp 56-64