

# Homework 3: Conditional Random Fields

CIS 620

February 11, 2005

This homework is due Tuesday, February 29. It should be submitted by email to me ([pereira@cis.upenn.edu](mailto:pereira@cis.upenn.edu)) as a tar or zip file. There is a single question, which is to reproduce the label bias experiment discussed in class and described in Section 5.1 of the original CRF paper <http://www.cis.upenn.edu/~epereira/papers/crf.pdf>. This is doable in MATLAB, since the data is pretty small. You can represent input symbols labels with small integers to make things easier. To generate training data, encode the simple HMM described in the paper, and generate observation-state pairs with a succession of random draws with the specified probabilities (the only nontrivial transition probabilities should be 0.5, 0.5 from the start state). For Viterbi (needed for MEMMs and CRFs) and forward-backward (needed for CRFs), you can use matrices with one row for each label and one column for each input position. For optimization, you can use iterative scaling (the data is that small), or a convenient optimization routine, such as the BFGS code in the MATLAB optimization toolbox or a free alternative such as <http://ftp.math.ncsu.edu/FTP/kelley/optimization/matlab/bfgswopt.m>. You can also try to use Kevin Murphy's CRF MATLAB code <http://www.cs.ubc.ca/~murphyk/Software/CRFall.zip> but I've never tried it myself.