Monday, September 10, 2012

0.1 Dynamic programming and DPP

For concreteness, we focus on Nussinov's DP algorithm. The ideas should work for general DPs. Suppose we have a fixed DNA/RNA sequence S of length L. Let y denote a particular secondary folding structure of the sequence. We try to follow the notations of the Supplementary Methods of Taskar's SDPP paper.

The factor graph has $T = L(L-1)$ number of variable nodes, indexed by $t = (i,j), i < j$. The possible values for the variable node (i,j) is the set of energies possible for the subsequence $S(i,j)$. There are also $L(L-1)$ factor nodes indexed by $(i,j), i < j$. The factor node $F(i,j)$ is connected to the variable nodes $\{(i+1,j-1), (i+k,j), (k+1,j)\}, i < k < j$. For $(i,j) \neq (1,L)$, the factor $F(i,j) = 1$ if the substructure $y(i,j)$ has a consistent folding (no mismatched base-pairs, etc), and $F(i,j) = 0$ if the substructure is not consistent. For the "last" factor, $F(1,L) =$ energy of the structure y