CS294-26/STAT260: Computational and Mathematical Population Genetics

January 24, 2008

The Yule-Harding Model

Recall that $\mathscr{T}_n^r$ denotes the set of all leaf-labeled rooted binary trees with n leaves. Suppose that the n leaves are bijectively labeled with $\mathcal{L} = \{1, 2, \ldots, n\}$. The Yule model (sometimes called the Yule-Harding model) generates a tree in $\mathscr{T}_n^r$ as follows:

1. Initialization:
   (a) With uniform probability, randomly pick two distinct labels i, j from $\mathcal{L}$ and let T be a 2-leaved rooted tree with i and j as leaf labels.
   (b) Redefine $\mathcal{L} \leftarrow \mathcal{L} \setminus \{i, j\}$.
2. Repeat the following procedure until $\mathcal{L} = \varnothing$:
   (a) With uniform probability, randomly pick a label k from $\mathcal{L}$.
   (b) With uniform probability, randomly pick a pendant edge (i.e., an edge attached to a leaf) from T, subdivide it to create a new vertex v, and join a new leaf labeled k to v, creating a new pendant edge.
   (c) Redefine $\mathcal{L} \leftarrow \mathcal{L} \setminus \{k\}$.

Remarks:

1. The edges in T are unweighted.
2. The Yule process is a pure birth process.
3. The Yule model produces the same probability distribution on $\mathscr{T}_n^r$ as the coalescent model.