**Course: Advance Bio Informatics**

**Module Title: Limitation of HMM**

**Module No: 105**

**Limitations of HMMs**

* HMM don't deal well with correlations between nucleotides.
* Conserved RNA base pairs which induce long-range pair-wise correlations

**What makes a good HMM problem space?**

**Classification problems**

There are two main types of output from an HMM:

* Scoring of sequences (Protein family modeling)
* Labeling of observations within a sequence (Gene Finding)

The observations in a sequence should have a clear and meaningful order. Unordered observations will not map easily to states. It’s beneficial, but not necessary for the observations follow some sort of grammar.

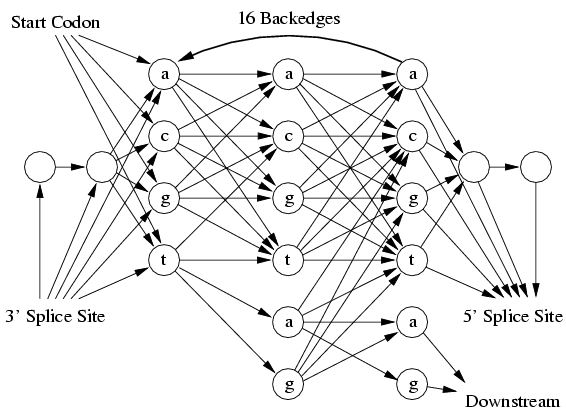
Makes it easier to design architecture

Gene Finding

Protein Family Modeling

**HMM Gene Finders:   
VEIL**

Grammatical structure & modular design. Uses many states that can only emit one symbol to get around state independence



**HMMGene**

Uses an extended HMM called a CHMM.

CHMM = HMM with classes

Takes full advantage of being able to modify the statistical algorithms.

Uses high-order states and trains everything at once.

**Genie**

Uses a generalized HMM (GHMM).

Edges in model are complete HMMs.

States can be any arbitrary program.

States are actually neural networks specially designed for signal finding.