

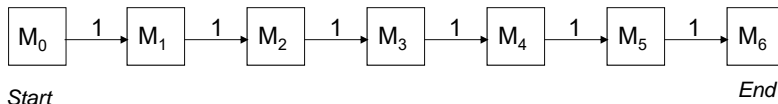
1. Modeling positional dependencies with HMMs
2. Profile HMM's model variable length motifs with insertions and deletions

### A PSSM for the WEIRD motif

|              |   |      |      |      |      |      |
|--------------|---|------|------|------|------|------|
| <b>WEIRD</b> | D | 0.08 | 0.08 | 0.08 | 0.08 | 0.33 |
|              | E | 0.08 | 0.54 | 0.08 | 0.08 | 0.08 |
| <b>WEIRD</b> | H | 0.08 | 0.08 | 0.08 | 0.08 | 0.25 |
| <b>WEIQH</b> | I | 0.08 | 0.08 | 0.54 | 0.08 | 0.08 |
|              | Q | 0.08 | 0.08 | 0.08 | 0.25 | 0.08 |
| <b>WEIRD</b> | R | 0.08 | 0.08 | 0.08 | 0.33 | 0.08 |
| <b>WEIQH</b> | W | 0.54 | 0.08 | 0.08 | 0.08 | 0.08 |

$$q[x, i] = \frac{c[x, i] + b}{k + |\Sigma| b}$$

An HMM that is equivalent to a PSSM for the WEIRD motif



Emission probabilities

|   | M1   | M2   | M3   | M4   | M5   |
|---|------|------|------|------|------|
| D | 0.08 | 0.08 | 0.08 | 0.08 | 0.33 |
| E | 0.08 | 0.54 | 0.08 | 0.08 | 0.08 |
| H | 0.08 | 0.08 | 0.08 | 0.08 | 0.25 |
| I | 0.08 | 0.08 | 0.54 | 0.08 | 0.08 |
| Q | 0.08 | 0.08 | 0.08 | 0.25 | 0.08 |
| R | 0.08 | 0.08 | 0.08 | 0.33 | 0.08 |
| W | 0.54 | 0.08 | 0.08 | 0.08 | 0.08 |

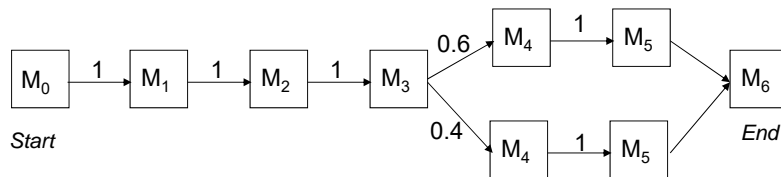
Does not capture positional dependencies

|       |   |      |      |      |      |      |
|-------|---|------|------|------|------|------|
| WEIRD | D | 0.08 | 0.08 | 0.08 | 0.08 | 0.33 |
| WEIRD | E | 0.08 | 0.54 | 0.08 | 0.08 | 0.08 |
| WEIRD | H | 0.08 | 0.08 | 0.08 | 0.08 | 0.25 |
| WEIQH | I | 0.08 | 0.08 | 0.54 | 0.08 | 0.08 |
| WEIRD | Q | 0.08 | 0.08 | 0.08 | 0.25 | 0.08 |
| WEIRD | R | 0.08 | 0.08 | 0.08 | 0.33 | 0.08 |
| WEIQH | W | 0.54 | 0.08 | 0.08 | 0.08 | 0.08 |

**Note:** We never see QD or RH, only RD and QH.  
 But,  $P(RH) = P(QD) = 0.083$ , while  $P(QH) = 0.063$

Branching topologies can model positional dependencies

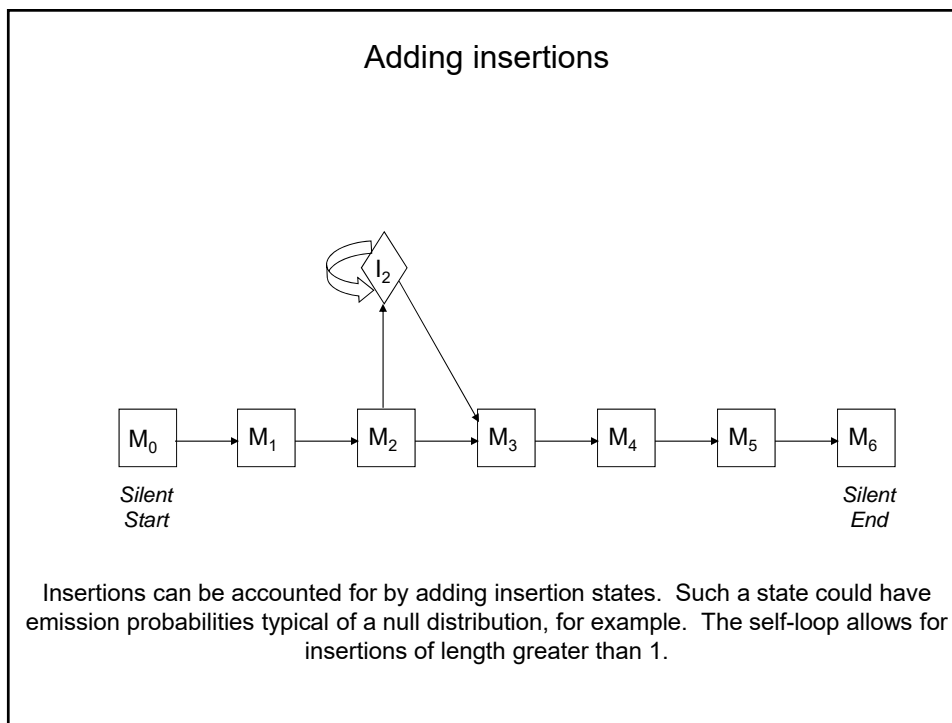
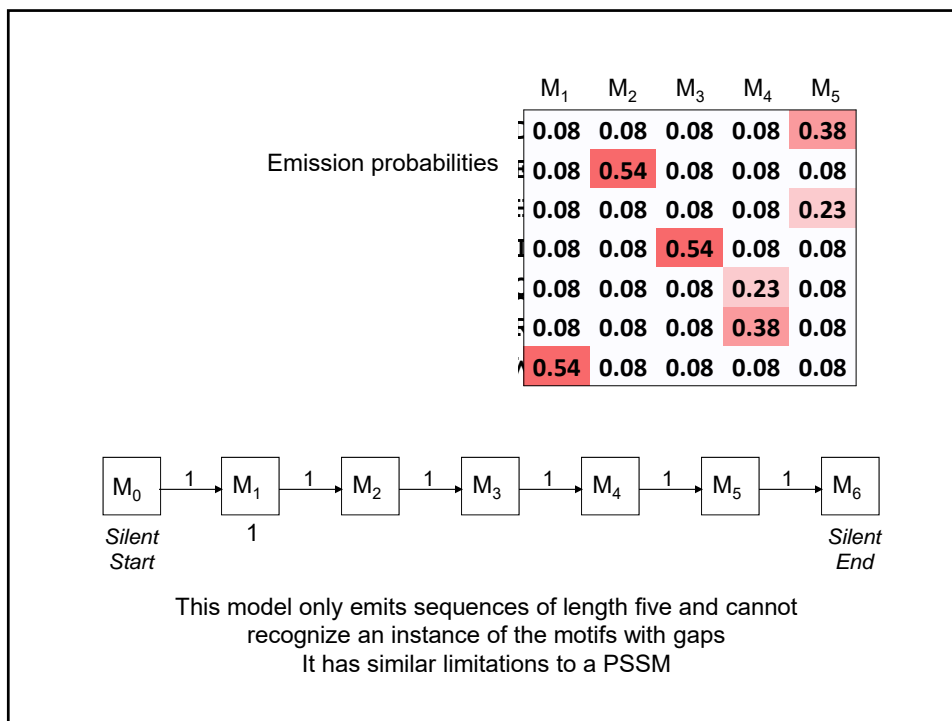
|   | M4    | M5    |
|---|-------|-------|
| D | 0.095 | 0.43  |
| H | 0.095 | 0.095 |
| Q | 0.095 | 0.095 |
| R | 0.43  | 0.095 |



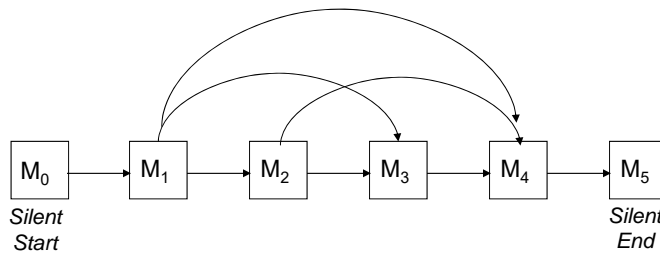
|   | M1   | M2   | M3   |
|---|------|------|------|
| D | 0.08 | 0.08 | 0.08 |
| E | 0.08 | 0.54 | 0.08 |
| H | 0.08 | 0.08 | 0.08 |
| I | 0.08 | 0.08 | 0.54 |
| Q | 0.08 | 0.08 | 0.08 |
| R | 0.08 | 0.08 | 0.08 |
| W | 0.54 | 0.08 | 0.08 |

|   | M4    | M5    |
|---|-------|-------|
| D | 0.095 | 0.095 |
| H | 0.095 | 0.395 |
| Q | 0.395 | 0.095 |
| R | 0.095 | 0.095 |

1. Modeling positional dependencies with HMMs
2. Profile HMM's model variable length motifs with insertions and deletions

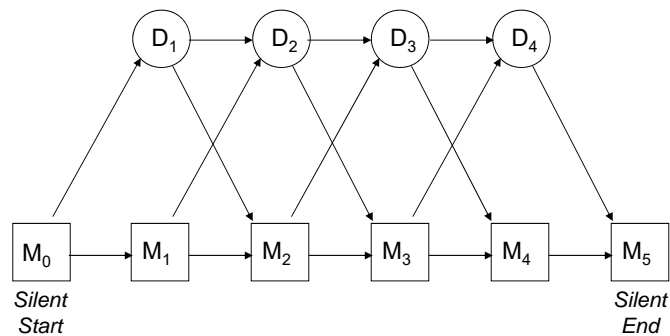


### Deletions



Deletions can be model by adding arcs that jump over one or more states. However, the number of such arcs will grow exponentially with the number of match states.

### Modeling deletions using explicit deletion states.



This approach avoids the combinatorial explosion in the number of transitions as the number of states increases. A state path cannot pass through both  $D_i$   $M_i$ , but not both.

