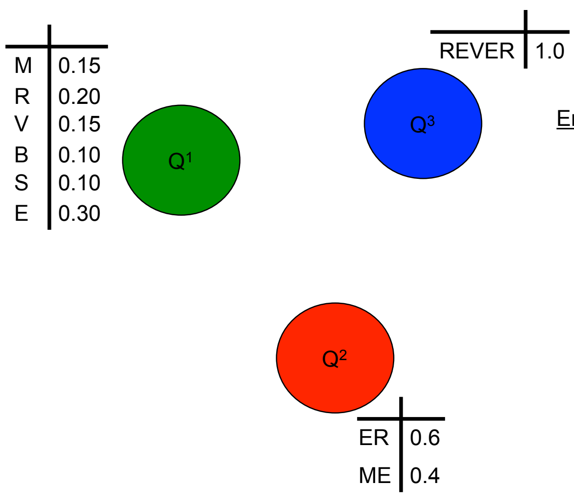
Gene Finding

Letters and Words

REVERES VERVE SEER BEES REMEMBER ERE VEER SEE ME MERE REVERSE VERSE EVER BEER VERBS SERVE

HMM: Emission Probabilities, *B*



Emission Probabilities, *B*

$b_1(M) = 0.15$

$b_1(R) = 0.20$

$b_1(V) = 0.15$

$b_1(B) = 0.10$

$b_1(S) = 0.10$

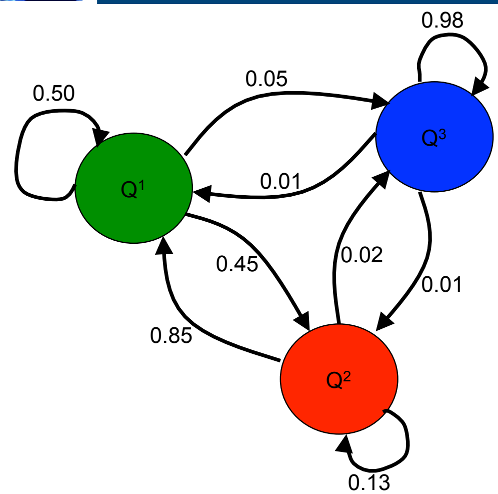
$b_1(E) = 0.30$

$b_2(ER) = 0.6$

$b_2(ME) = 0.4$

$b_3(REVER) = 1.0$

HMM: Transition Probability, *A*



Transition Probabilities, *A*

$a_{11} = 0.50$

$a_{12} = 0.45$

$a_{13} = 0.05$

$a_{21} = 0.85$

$a_{22} = 0.13$

$a_{23} = 0.02$

$a_{31} = 0.01$

$a_{32} = 0.01$

$a_{33} = 0.98$

HMM: Model, λ

$$\lambda = (A, B)$$

Transition Probabilities, A

$a_{11} = 0.50$

$a_{12} = 0.45$

$a_{13} = 0.05$

$a_{21} = 0.85$

$a_{22} = 0.13$

$a_{23} = 0.02$

$a_{31} = 0.01$

$a_{32} = 0.01$

$a_{33} = 0.98$

Emission Probabilities, B

$b_1(M) = 0.15$

$b_1(R) = 0.20$

$b_1(V) = 0.15$

$b_1(B) = 0.10$

$b_1(S) = 0.10$

$b_1(E) = 0.30$

$b_2(ER) = 0.6$

$b_2(ME) = 0.4$

$b_3(REVER) = 1.0$

Probability of Observation Sequence

$O =$	M	E	R	E	V	E	R	B	S
	Q^1	Q^1	Q^1	Q^1	Q^1	Q^1	Q^1	Q^1	Q^1

$P(O) = b_1(M)a_{11}\, b_1(E)a_{11}\, b_1(R)a_{11}\, b_1(E)a_{11}\, b_1(V)a_{11}\, b_1(E)a_{11}\, b_1(R)\, a_{11}\, b_1(B)a_{11}\, b_1(S)$

0.15 0.50 0.30 0.50 0.20 0.50 0.30 0.50 0.15 0.50 0.30 0.50 0.20 0.50 0.10 0.50 0.10

0.00000000094921875

* Assuming we start in state 1

Probability of Observation Sequence

$O =$	M	E	R	E	V	E	R	B	S
	Q^1	Q^1	Q^3					Q^1	Q^1

$P(O) = b_1(M)a_{11}\, b_1(E)a_{13}\, b_3(REVER)a_{31} \quad b_1(B)a_{11}\, b_1(S)$

0.15 0.50 0.30 0.05 1.00 0.01 0.10 0.50 0.10

0.00000005625

* Assuming we start in state 1

Probability of Observation Sequence

$O =$	M	E	R	E	V	E	R	B	S
	Q^1	Q^2		Q^1	Q^1	Q^2		Q^1	Q^1

$P(O) = b_1(M)a_{12}\, b_2(ER)\, a_{21} \quad b_1(E)a_{11}\, b_1(V)a_{12}\, b_2(ER)a_{21} \quad b_1(B)\, a_{11}\, b_1(S)$

0.15 0.45 0.60 0.85 0.30 0.50 0.15 0.45 0.60 0.85 0.10 0.50 0.10

0.0000008881046875

* Assuming we start in state 1

Probability of Observation Sequence

O =	M	E	R	E	V	E	R	B	S
	Q^1	Q^1	Q^2		Q^3				

$P(O) = b_1(M) a_{11}\, b_1(E)\, a_{12}\, b_2(RE)\, a_{23} \quad b_3(VERBS)$

0.15 0.50 0.30 0.45 0.00 0.02 0.00

0.00

* Assuming we start in state 1

Dynamic Programming Table

$$\delta_t(j) = \begin{cases} b_1(O_1) & \text{if } t = 1, j = 1 \quad \text{// base case, we start in state \#1} \\ 0.0 & \text{if } t = 1, j \neq 1 \quad \text{// base case, we cannot start in // a state other than state \#1} \\ \max\limits_{1 \le i \le N} \left(\delta_{t-1}(i) * a_{ij}\right) * b_j(O_t) & \text{if } 2 \le t \le T \quad \text{// recursive case} \end{cases}$$

$\delta_1(1)$	$\delta_2(1)$	$\delta_3(1)$	$\delta_4(1)$	$\delta_5(1)$	$\delta_6(1)$	$\delta_7(1)$	$\delta_8(1)$	$\delta_9(1)$	$\delta_{10}(1)$
$\delta_1(2)$	$\delta_2(2)$	$\delta_3(2)$	$\delta_4(2)$	$\delta_5(2)$	$\delta_6(2)$	$\delta_7(2)$	$\delta_8(2)$	$\delta_9(2)$	$\delta_{10}(2)$
$\delta_1(3)$	$\delta_2(3)$	$\delta_3(3)$	$\delta_4(3)$	$\delta_5(3)$	$\delta_6(3)$	$\delta_7(3)$	$\delta_8(3)$	$\delta_9(3)$	$\delta_{10}(3)$

Base Case*

$$\delta_t(j) = \begin{cases} b_1(O_1) & \text{if } t = 1, j = 1 \quad \text{// base case, we start in state \#1} \\ 0.0 & \text{if } t = 1, j \neq 1 \quad \text{// base case, we cannot start in // a state other than state \#1} \end{cases}$$

$b_1(O_1)$									
0.0									
0.0									

* Assuming we start in state 1 and state 1 emits characters of length 1

When State Outputs Character of Length 1

$$\delta_t(j) = \begin{cases} \\ \max\limits_{1 \le i \le N} \left(\delta_{t-1}(i) * a_{ij}\right) * b_j(O_t) & \text{if } 2 \le t \le T \quad \text{// recursive case} \end{cases}$$

$$\boldsymbol{\delta_7(1)} = \max\{\boldsymbol{\delta_6(1)} * a_{11}, \boldsymbol{\delta_6(2)} * a_{21}, \boldsymbol{\delta_6(3)} * a_{31}\} * b_1(O_7)$$

					$\delta_6(1)$	$\delta_7(1)$			
					$\delta_6(2)$				
					$\delta_6(3)$				

When State Outputs Character of Length 2

$$\delta_t(j) = \begin{cases} \\ \max\limits_{1 \le i \le N} \left(\delta_{t-1}(i) * a_{ij} \right) * b_j(O_t) & \text{if } 2 \le t \le T \end{cases}$$ // recursive case

$\boldsymbol{\delta_7(2)} = \max\{\boldsymbol{\delta_5(1)} * a_{12}, \boldsymbol{\delta_5(2)} * a_{22}, \boldsymbol{\delta_5(3)} * a_{32}\} * b_2(O_6O_7)$

				$\delta_5(1)$					
				$\delta_5(2)$		$\delta_7(2)$			
				$\delta_5(3)$					

When State Outputs Character of Length 5

$$\delta_t(j) = \begin{cases} \\ \max\limits_{1 \le i \le N} \left(\delta_{t-1}(i) * a_{ij} \right) * b_j(O_t) & \text{if } 2 \le t \le T \end{cases}$$ // recursive case

$\boldsymbol{\delta_7(3)} = \max\{\boldsymbol{\delta_2(1)} * a_{13}, \boldsymbol{\delta_2(2)} * a_{23}, \boldsymbol{\delta_2(3)} * a_{33}\} * b_3(O_3O_4O_5O_6O_7)$

	$\delta_2(1)$								
	$\delta_2(2)$								
	$\delta_2(3)$					$\delta_7(3)$			

Recurrence For When States Output Characters of Length Greater Than 1

$$\delta_t(j) = \begin{cases} b_1(O_1) & \text{if } t = 1, j = 1 \\ 0.0 & \text{if } t = 1, j \ne 1 \\ \max\limits_{1 \le i \le N} \left(\delta_{t-|b_j|}(i) * a_{ij} \right) * b_j(O_{t-|b_j|+1} \ldots O_t) & \text{if } 2 \le t \le T \end{cases}$$

// base case, we start in state #1

// base case, we cannot start in // a state other than state #1

// recursive case, where $|b_j|$ is // the length of characters // emitted by state j

$\delta_1(1)$	$\delta_2(1)$	$\delta_3(1)$	$\delta_4(1)$	$\delta_5(1)$	$\delta_6(1)$	$\delta_7(1)$	$\delta_8(1)$	$\delta_9(1)$	$\delta_{10}(1)$
$\delta_1(2)$	$\delta_2(2)$	$\delta_3(2)$	$\delta_4(2)$	$\delta_5(2)$	$\delta_6(2)$	$\delta_7(2)$	$\delta_8(2)$	$\delta_9(2)$	$\delta_{10}(2)$
$\delta_1(3)$	$\delta_2(3)$	$\delta_3(3)$	$\delta_4(3)$	$\delta_5(3)$	$\delta_6(3)$	$\delta_7(3)$	$\delta_8(3)$	$\delta_9(3)$	$\delta_{10}(3)$

* Assuming we start in state 1 and state 1 emits characters of length 1

Using Logarithms To Compute Optimal State Sequence

Given an observation sequence $O = O_1O_2O_3 \ldots O_T$ and a model $\lambda = (A, B)$, what is the optimal state sequence?

- We want to uncover the hidden part of the model. We want to maximize $P(Q|O, \lambda)$.

$$\arg\max_{Q_1,Q_2,\cdots,Q_T} \left(b_{Q_1}(O_1) a_{Q_1Q_2} b_{Q_2}(O_2) a_{Q_2Q_3} b_{Q_3}(O_3) \cdots a_{Q_{T-1}Q_T} b_{Q_T}(O_T) \right)$$

$$\arg\max_{Q_1,Q_2,\cdots,Q_T} \left(\ln\left(b_{Q_1}(O_1) a_{Q_1Q_2} b_{Q_2}(O_2) a_{Q_2Q_3} b_{Q_3}(O_3) \cdots a_{Q_{T-1}Q_T} b_{Q_T}(O_T) \right) \right)$$

$$\arg\max_{Q_1,Q_2,\cdots,Q_T} \left(\ln(b_{Q_1}(O_1)) + \ln(a_{Q_1Q_2}) + \ln(b_{Q_2}(O_2)) + \ln(a_{Q_2Q_3}) + \cdots \ln(a_{Q_{T-1}Q_T}) + \ln(b_{Q_T}(O_T)) \right)$$

Recurrence With Logarithms

$$\delta_t(j) = \begin{cases} \ln(b_1(O_1)) & \text{if } t = 1, j = 1 \\ -\infty & \text{if } t = 1, j \neq 1 \\ \max\limits_{1 \le i \le N} \left(\delta_{t-|b_j|}(i) + \ln(a_{ij})\right) + \ln(b_j(O_{t-|b_j|+1} \ldots O_t)) & \text{if } 2 \le t \le T \end{cases}$$

$\delta_1(1)$	$\delta_2(1)$	$\delta_3(1)$	$\delta_4(1)$	$\delta_5(1)$	$\delta_6(1)$	$\delta_7(1)$	$\delta_8(1)$	$\delta_9(1)$	$\delta_{10}(1)$
$\delta_1(2)$	$\delta_2(2)$	$\delta_3(2)$	$\delta_4(2)$	$\delta_5(2)$	$\delta_6(2)$	$\delta_7(2)$	$\delta_8(2)$	$\delta_9(2)$	$\delta_{10}(2)$
$\delta_1(3)$	$\delta_2(3)$	$\delta_3(3)$	$\delta_4(3)$	$\delta_5(3)$	$\delta_6(3)$	$\delta_7(3)$	$\delta_8(3)$	$\delta_9(3)$	$\delta_{10}(3)$

* Assuming we start in state 1 and state 1 emits characters of length 1

Example HMM Model, λ

$\lambda = (A, B)$

Transition Probabilities, *A*

$a_{11} = 0.50$
$a_{12} = 0.45$
$a_{13} = 0.05$

$a_{21} = 0.85$
$a_{22} = 0.13$
$a_{23} = 0.02$

$a_{31} = 0.01$
$a_{32} = 0.01$
$a_{33} = 0.98$

Emission Probabilities, *B*

$b_1(M) = 0.15$
$b_1(R) = 0.20$
$b_1(V) = 0.15$
$b_1(B) = 0.10$
$b_1(S) = 0.10$
$b_1(E) = 0.30$
$b_2(ER) = 0.6$
$b_2(ME) = 0.4$
$b_3(REVER) = 1.0$

Example Observation Sequence: MEREVERBS

$$\delta_t(j) = \begin{cases} \ln(b_1(O_1)) & \text{if } t = 1, j = 1 \\ -\infty & \text{if } t = 1, j \neq 1 \\ \max\limits_{1 \le i \le N} \left(\delta_{t-|b_j|}(i) + \ln(a_{ij})\right) + \ln(b_j(O_{t-|b_j|+1} \ldots O_t)) & \text{if } 2 \le t \le T \end{cases}$$

* Assuming we start in state 1 and state 1 emits characters of length 1

Example Observation Sequence: MEREVERBS

$$\delta_t(j) = \begin{cases} \ln(b_1(O_1)) & \text{if } t = 1, j = 1 \\ -\infty & \text{if } t = 1, j \neq 1 \\ \max\limits_{1 \le i \le N} \left(\delta_{t-|b_j|}(i) + \ln(a_{ij})\right) + \ln(b_j(O_{t-|b_j|+1} \ldots O_t)) & \text{if } 2 \le t \le T \end{cases}$$

$\boldsymbol{\delta_1(1)} = \ln(b_1(O_1)) = \ln(b_1(M)) = \ln(0.15) \approx -1.9$

-1.9								

* Assuming we start in state 1 and state 1 emits characters of length 1

Example Observation Sequence: MEREVERBS

$$\delta_t(j) = \begin{cases} \ln(b_1(O_1)) & \text{if } t=1, j=1 \\ -\infty & \text{if } t=1, j\neq 1 \\ \max\limits_{1\le i\le N}\left(\delta_{t-|b_j|}(i)+\ln(a_{ij})\right)+\ln(b_j(O_{t-|b_j|+1}\ldots O_t)) & \text{if } 2\le t\le T\end{cases}$$

$\delta_1(2)$ ≈ -∞

-1.9								
-∞								

* Assuming we start in state 1 and state 1 emits characters of length 1

Example Observation Sequence: MEREVERBS

$$\delta_t(j) = \begin{cases} \ln(b_1(O_1)) & \text{if } t=1, j=1 \\ -\infty & \text{if } t=1, j\neq 1 \\ \max\limits_{1\le i\le N}\left(\delta_{t-|b_j|}(i)+\ln(a_{ij})\right)+\ln(b_j(O_{t-|b_j|+1}\ldots O_t)) & \text{if } 2\le t\le T\end{cases}$$

$\delta_1(3)$ ≈ -∞

-1.9								
-∞								
-∞								

* Assuming we start in state 1 and state 1 emits characters of length 1

Example Observation Sequence: MEREVERBS

$$\delta_t(j) = \begin{cases} \ln(b_1(O_1)) & \text{if } t=1, j=1 \\ -\infty & \text{if } t=1, j\neq 1 \\ \max\limits_{1\le i\le N}\left(\delta_{t-|b_j|}(i)+\ln(a_{ij})\right)+\ln(b_j(O_{t-|b_j|+1}\ldots O_t)) & \text{if } 2\le t\le T\end{cases}$$

$\delta_2(1)$ = max{$\delta_1(1)$+ln(a_{11}), $\delta_1(2)$+ln(a_{21}), $\delta_1(3)$+ln(a_{31})}+ln(b_1(E))

-1.9								
-∞								
-∞								

* Assuming we start in state 1 and state 1 emits characters of length 1

Example Observation Sequence: MEREVERBS

$$\delta_t(j) = \begin{cases} \ln(b_1(O_1)) & \text{if } t=1, j=1 \\ -\infty & \text{if } t=1, j\neq 1 \\ \max\limits_{1\le i\le N}\left(\delta_{t-|b_j|}(i)+\ln(a_{ij})\right)+\ln(b_j(O_{t-|b_j|+1}\ldots O_t)) & \text{if } 2\le t\le T\end{cases}$$

$\delta_2(1)$ = max{-1.9+ln(0.5),-∞+ln(0.85),-∞+ln(0.01)}+ln(0.3) ≈ -3.8

-1.9	-3.8							
-∞								
-∞								

* Assuming we start in state 1 and state 1 emits characters of length 1

Example Observation Sequence: MEREVERBS

$$\delta_t(j) = \begin{cases} \ln(b_1(O_1)) & \text{if } t = 1, j = 1 \\ -\infty & \text{if } t = 1, j \neq 1 \\ \max\limits_{1 \le i \le N}\left(\delta_{t-|b_j|}(i) + \ln(a_{ij})\right) + \ln(b_j(O_{t-|b_j|+1} \ldots O_t)) & \text{if } 2 \le t \le T \end{cases}$$

$\boldsymbol{\delta_2(2)} = \max\{\delta_0(1)+\ln(a_{12}), \delta_0(2)+\ln(a_{22}), \delta_0(3)+\ln(a_{32})\}+\ln(b_2(\text{ME}))$

-1.9	-3.8							
-∞								
-∞								

* Assuming we start in state 1 and state 1 emits characters of length 1

Example Observation Sequence: MEREVERBS

$$\delta_t(j) = \begin{cases} \ln(b_1(O_1)) & \text{if } t = 1, j = 1 \\ -\infty & \text{if } t = 1, j \neq 1 \\ \max\limits_{1 \le i \le N}\left(\delta_{t-|b_j|}(i) + \ln(a_{ij})\right) + \ln(b_j(O_{t-|b_j|+1} \ldots O_t)) & \text{if } 2 \le t \le T \end{cases}$$

$\boldsymbol{\delta_2(2)} = \max\{-\infty+\ln(0.45),-\infty+\ln(0.13),-\infty+\ln(0.01)\}+\ln(0.4) \approx -\infty$

-1.9	-3.8							
-∞	-∞							
-∞								

* Assuming we start in state 1 and state 1 emits characters of length 1

Example Observation Sequence: MEREVERBS

$$\delta_t(j) = \begin{cases} \ln(b_1(O_1)) & \text{if } t = 1, j = 1 \\ -\infty & \text{if } t = 1, j \neq 1 \\ \max\limits_{1 \le i \le N}\left(\delta_{t-|b_j|}(i) + \ln(a_{ij})\right) + \ln(b_j(O_{t-|b_j|+1} \ldots O_t)) & \text{if } 2 \le t \le T \end{cases}$$

$\boldsymbol{\delta_2(3)} = \max\{\delta_{-3}(1)+\ln(a_{13}), \delta_{-3}(2)+\ln(a_{23}), \delta_{-3}(3)+\ln(a_{33})\}+\ln(b_3(\text{ME}))$

-1.9	-3.8							
-∞	-∞							
-∞								

* Assuming we start in state 1 and state 1 emits characters of length 1

Example Observation Sequence: MEREVERBS

$$\delta_t(j) = \begin{cases} \ln(b_1(O_1)) & \text{if } t = 1, j = 1 \\ -\infty & \text{if } t = 1, j \neq 1 \\ \max\limits_{1 \le i \le N}\left(\delta_{t-|b_j|}(i) + \ln(a_{ij})\right) + \ln(b_j(O_{t-|b_j|+1} \ldots O_t)) & \text{if } 2 \le t \le T \end{cases}$$

$\boldsymbol{\delta_2(3)} = \max\{-\infty+\ln(0.05),-\infty+\ln(0.02),-\infty+\ln(0.98)\}+\ln(0.0) \approx -\infty$

-1.9	-3.8							
-∞	-∞							
-∞	-∞							

* Assuming we start in state 1 and state 1 emits characters of length 1

Example Observation Sequence: MEREVERBS

$$\delta_t(j) = \begin{cases} \ln(b_1(O_1)) & \text{if } t = 1, j = 1 \\ -\infty & \text{if } t = 1, j \neq 1 \\ \max_{1 \le i \le N}\left(\delta_{t-|b_j|}(i) + \ln(a_{ij})\right) + \ln(b_j(O_{t-|b_j|+1} \ldots O_t)) & \text{if } 2 \le t \le T \end{cases}$$

$\delta_3(1)$ = max{$\delta_2(1)$+ln(a_{11}), $\delta_2(2)$+ln(a_{21}), $\delta_2(3)$+ln(a_{31})}+ln(b_1(R))

-1.9	-3.8							
-∞	-∞							
-∞	-∞							

* Assuming we start in state 1 and state 1 emits characters of length 1

Example Observation Sequence: MEREVERBS

$$\delta_t(j) = \begin{cases} \ln(b_1(O_1)) & \text{if } t = 1, j = 1 \\ -\infty & \text{if } t = 1, j \neq 1 \\ \max_{1 \le i \le N}\left(\delta_{t-|b_j|}(i) + \ln(a_{ij})\right) + \ln(b_j(O_{t-|b_j|+1} \ldots O_t)) & \text{if } 2 \le t \le T \end{cases}$$

$\delta_3(1)$ = max{-3.8+ln(0.50),-∞+ln(0.85),-∞+ln(0.01)}+ln(0.2) ≈ -6.1

-1.9	-3.8	-6.1						
-∞	-∞							
-∞	-∞							

* Assuming we start in state 1 and state 1 emits characters of length 1

Example Observation Sequence: MEREVERBS

$$\delta_t(j) = \begin{cases} \ln(b_1(O_1)) & \text{if } t = 1, j = 1 \\ -\infty & \text{if } t = 1, j \neq 1 \\ \max_{1 \le i \le N}\left(\delta_{t-|b_j|}(i) + \ln(a_{ij})\right) + \ln(b_j(O_{t-|b_j|+1} \ldots O_t)) & \text{if } 2 \le t \le T \end{cases}$$

$\delta_3(2)$ = max{$\delta_1(1)$+ln(a_{12}), $\delta_1(2)$+ln(a_{22}), $\delta_1(3)$+ln(a_{32})}+ln(b_2(ER))

-1.9	-3.8	-6.1						
-∞	-∞							
-∞	-∞							

* Assuming we start in state 1 and state 1 emits characters of length 1

Example Observation Sequence: MEREVERBS

$$\delta_t(j) = \begin{cases} \ln(b_1(O_1)) & \text{if } t = 1, j = 1 \\ -\infty & \text{if } t = 1, j \neq 1 \\ \max_{1 \le i \le N}\left(\delta_{t-|b_j|}(i) + \ln(a_{ij})\right) + \ln(b_j(O_{t-|b_j|+1} \ldots O_t)) & \text{if } 2 \le t \le T \end{cases}$$

$\delta_3(2)$ = max{-1.9+ln(0.45),-∞+ln(0.13),-∞+ln(0.01)}+ln(0.6) ≈ -3.2

-1.9	-3.8	-6.1						
-∞	-∞	-3.2						
-∞	-∞							

* Assuming we start in state 1 and state 1 emits characters of length 1

Example Observation Sequence: MEREVERBS

$$\delta_t(j) = \begin{cases} \ln(b_1(O_1)) & \text{if } t = 1, j = 1 \\ -\infty & \text{if } t = 1, j \neq 1 \\ \max\limits_{1 \le i \le N}\left(\delta_{t-|b_j|}(i) + \ln(a_{ij})\right) + \ln(b_j(O_{t-|b_j|+1} \ldots O_t)) & \text{if } 2 \le t \le T \end{cases}$$

$\boldsymbol{\delta_3(3)} = \max\{\delta_{-2}(1)+\ln(a_{13}), \delta_{-2}(2)+\ln(a_{23}), \delta_{-2}(3)+\ln(a_{33})\}+\ln(b_3(\text{MER}))$

-1.9	-3.8	-6.1						
-∞	-∞	-3.2						
-∞	-∞							

* Assuming we start in state 1 and state 1 emits characters of length 1

Example Observation Sequence: MEREVERBS

$$\delta_t(j) = \begin{cases} \ln(b_1(O_1)) & \text{if } t = 1, j = 1 \\ -\infty & \text{if } t = 1, j \neq 1 \\ \max\limits_{1 \le i \le N}\left(\delta_{t-|b_j|}(i) + \ln(a_{ij})\right) + \ln(b_j(O_{t-|b_j|+1} \ldots O_t)) & \text{if } 2 \le t \le T \end{cases}$$

$\boldsymbol{\delta_3(3)} = \max\{-\infty+\ln(0.05),-\infty+\ln(0.02),-\infty+\ln(0.98)\}+\ln(0.0) \approx -\infty$

-1.9	-3.8	-6.1						
-∞	-∞	-3.2						
-∞	-∞	-∞						

* Assuming we start in state 1 and state 1 emits characters of length 1

Example Observation Sequence: MEREVERBS

$$\delta_t(j) = \begin{cases} \ln(b_1(O_1)) & \text{if } t = 1, j = 1 \\ -\infty & \text{if } t = 1, j \neq 1 \\ \max\limits_{1 \le i \le N}\left(\delta_{t-|b_j|}(i) + \ln(a_{ij})\right) + \ln(b_j(O_{t-|b_j|+1} \ldots O_t)) & \text{if } 2 \le t \le T \end{cases}$$

$\boldsymbol{\delta_4(1)} = \max\{\delta_3(1)+\ln(a_{11}), \delta_3(2)+\ln(a_{21}), \delta_3(3)+\ln(a_{31})\}+\ln(b_1(\text{E}))$

-1.9	-3.8	-6.1						
-∞	-∞	-3.2						
-∞	-∞	-∞						

* Assuming we start in state 1 and state 1 emits characters of length 1

Example Observation Sequence: MEREVERBS

$$\delta_t(j) = \begin{cases} \ln(b_1(O_1)) & \text{if } t = 1, j = 1 \\ -\infty & \text{if } t = 1, j \neq 1 \\ \max\limits_{1 \le i \le N}\left(\delta_{t-|b_j|}(i) + \ln(a_{ij})\right) + \ln(b_j(O_{t-|b_j|+1} \ldots O_t)) & \text{if } 2 \le t \le T \end{cases}$$

$\boldsymbol{\delta_4(1)} = \max\{-6.1+\ln(0.50),-3.2+\ln(0.85),-\infty+\ln(0.01)\}+\ln(0.3) \approx -4.6$

-1.9	-3.8	-6.1	-4.6					
-∞	-∞	-3.2						
-∞	-∞	-∞						

* Assuming we start in state 1 and state 1 emits characters of length 1

Example Observation Sequence: MEREVERBS

$$\delta_t(j) = \begin{cases} \ln(b_1(O_1)) & \text{if } t=1, j=1 \\ -\infty & \text{if } t=1, j\neq 1 \\ \max\limits_{1\le i\le N}\left(\delta_{t-|b_j|}(i)+\ln(a_{ij})\right)+\ln(b_j(O_{t-|b_j|+1}\dots O_t)) & \text{if } 2\le t\le T \end{cases}$$

$\boldsymbol{\delta_4(2)} = \max\{\delta_2(1)+\ln(a_{12}), \delta_2(2)+\ln(a_{22}), \delta_2(3)+\ln(a_{32})\}+\ln(b_2(\text{RE}))$

-1.9	-3.8	-6.1	-4.6					
-∞	-∞	-3.2						
-∞	-∞	-∞						

* Assuming we start in state 1 and state 1 emits characters of length 1

Example Observation Sequence: MEREVERBS

$$\delta_t(j) = \begin{cases} \ln(b_1(O_1)) & \text{if } t=1, j=1 \\ -\infty & \text{if } t=1, j\neq 1 \\ \max\limits_{1\le i\le N}\left(\delta_{t-|b_j|}(i)+\ln(a_{ij})\right)+\ln(b_j(O_{t-|b_j|+1}\dots O_t)) & \text{if } 2\le t\le T \end{cases}$$

$\boldsymbol{\delta_4(2)} = \max\{-3.8+\ln(0.45),-\infty+\ln(0.13),-\infty+\ln(0.01)\}+\ln(0.0) \approx -\infty$

-1.9	-3.8	-6.1	-4.6					
-∞	-∞	-3.2	-∞					
-∞	-∞	-∞						

* Assuming we start in state 1 and state 1 emits characters of length 1

Example Observation Sequence: MEREVERBS

$$\delta_t(j) = \begin{cases} \ln(b_1(O_1)) & \text{if } t=1, j=1 \\ -\infty & \text{if } t=1, j\neq 1 \\ \max\limits_{1\le i\le N}\left(\delta_{t-|b_j|}(i)+\ln(a_{ij})\right)+\ln(b_j(O_{t-|b_j|+1}\dots O_t)) & \text{if } 2\le t\le T \end{cases}$$

-1.9	-3.8	-6.1	-4.6	-7.2	-9.1	-11.4	-10.9	-13.9
-∞	-∞	-3.2	-∞	-∞	-∞	-8.5	-∞	-∞
-∞	-∞	-∞	-∞	-∞	-∞	-6.8	-∞	-∞

* Assuming we start in state 1 and state 1 emits characters of length 1

Example Observation Sequence: MEREVERBS

$$\delta_t(j) = \begin{cases} \ln(b_1(O_1)) & \text{if } t=1, j=1 \\ -\infty & \text{if } t=1, j\neq 1 \\ \max\limits_{1\le i\le N}\left(\delta_{t-|b_j|}(i)+\ln(a_{ij})\right)+\ln(b_j(O_{t-|b_j|+1}\dots O_t)) & \text{if } 2\le t\le T \end{cases}$$

Backtracking Table

-1	1	1	2	1	1	1	2	1
-1	-1	1	1	2	1	1	1	2
-1	-1	-1	-1	-1	1	1	2	1

* Assuming we start in state 1 and state 1 emits characters of length 1

Example Observation Sequence: MEREVERBS

$Q^1Q^2Q^1Q^1Q^2Q^1Q^1$

Backtracking Table

-1	1	1	2	1	1	1	2	1
-1	-1	1	1	2	1	1	1	2
-1	-1	-1	-1	-1	1	1	2	1

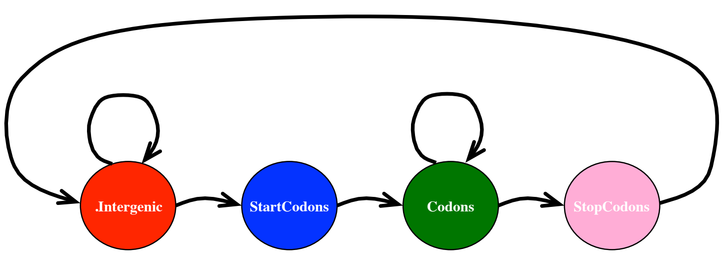
* Assuming we start in state 1 and state 1 emits characters of length 1

Application of HMMs: Finding Genes in a Sequenced Genome

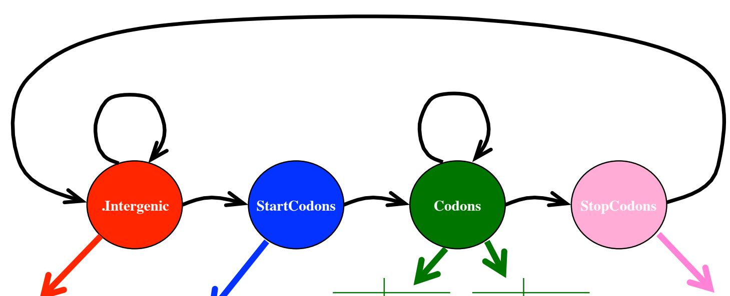
AGCTGTACATCGGCGTATCGCGATGCGATCGGCATGTACGGCGCGTATATGCGCATTAGA
TTTAGCGAGTCTCTCGATTGTGTACTCTCTGCTGCATGATCTATCATACTTAGATTAGTA
CGTCGCCCATGAG**ATGCGCGCGCAGAGCGAATCTATACTACTACTACCTGCTGCATGATG**
GACGTATGCATGTATCGAGAGGATCGATGACTGAGAGGAGGAGTCTCTCTATGATATATA
GGTATGCTTCGCATGCATCGATTAGCTAGCATCTGATCTAAGCCTCGATTGTGTACTCTG
AAAAACACCCAATCGGCGTATCGCGATGCGATCGGCATGTACGGCGCGTATATGCGCATT
ATACAATGCGTCTCTCGATTGTGTACTCTCTGCTGCATGATCTATCATACTTAGATTAGT
GTAATCGCAATGCAATGAGCGACGCGCGCAGAGCGAATCTATACTACTACATGCTGCTGC
GTACTGTGCGTAGTACTGAGCATGTATCGAGAGGATCGATGACTGAGAGGAGGAGTCTCT
TACGATTCGATTGAGCTTCGCATGCATCGATTAGCTAGCATCTGATCGATGCCTCGATTG
TACTTACGATCTGACGTATGCTGTGACTGATGCATCGTATCGATGCATCATGCGTATGAC
ATATCATAGTACTGAGTTCTCTTCTCTCTGATCGGGGAGAGGGGGGGCGTATATCGGAGA
GTAAGTACGCATTGGCATCGATTGCAGGACTTAGCGAGAGAGAGCTTCTAGCGTCTAGTA
ATCCCATGATCTACGAGATGCATGCATGCTGATCGACTGATGTATGCTACTGACTGATGT
ATCATCAGATCTGACTGATGCGCTCTGCATGATGCATCGATCGATGCTATCGGATATACG
CGATACGCTGATACGTATGCATGGCATATTCTCTCTCTCGCTGCTGCTTCGTCTGGAAGA
TTTTCAGAGGGCGTATATATACTCTTCTCTCTATAGCTATACGCTGATCAATACGATCGT
GCATGAGACTATGCATGCTGATCGATATCTCTCTGATATCGCATTAGCATCATGCTAGCA

Example HMM Used for Gene Finding

- Here, our HMM has $N = 4$ states. The emission alphabet corresponds to the four DNA nucleotides: *A*, *C*, *G*, *T*. What are our model parameters?



Emission Probabilities



A	0.28
C	0.22
G	0.22
T	0.28

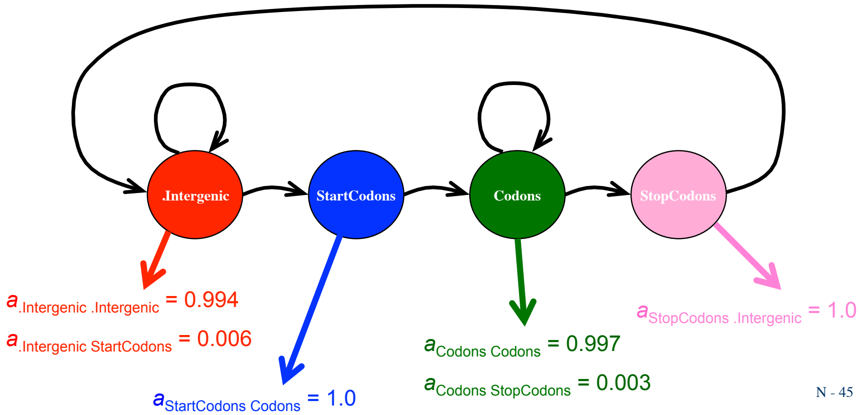
ATG	0.90
TTG	0.02
GTG	0.08

AAA	0.03	ACC	0.02
AAC	0.02	...	
AAG	0.01	TTC	0.02
AAT	0.02	TTG	0.01
ACA	0.01	TTT	0.02

TAA	0.65
TGA	0.28
TAG	0.07

Transition Probabilities

- Suppose that the average length of intergenic regions is 160 nucleotdies
- Suppose that the average length of genes is 300 amino acids (i.e., 900 nucleotides or 300 codons)



Observation Sequence

AGCTGTACATCGGCGTATCGCGATGCGATCGGCATGTACGGCGCGTATATGCGCATTAGA
TTTAGCGAGTCTCTCGATTGTGTACTCTCTGCTGCATGATCTATCATACTTAGATTAGTA
CGTCGCCCATGAGATGCGCGCGCAGAGCGAATCTATACTACTACTACCTGCTGCATGATG
GACGTATGCATGTATCGAGAGGATCGATGACTGAGAGGAGGAGTCTCTCTATGATATATA
GGTATGCTTCGCATGCATCGATTAGCTAGCATCTGATCTAAGCCTCGATTGTGTACTCTG
AAAAACACCCAATCGGCGTATCGCGATGCGATCGGCATGTACGGCGCGTATATGCGCATT
ATACAATGCGTCTCTCGATTGTGTACTCTCTGCTGCATGATCTATCATACTTAGATTAGT
GTAATCGCAATGCAATGAGCGACGCGCGCAGAGCGAATCTATACTACTACATGCTGCTGC
GTACTGTGCGTAGTACTGAGCATGTATCGAGAGGATCGATGACTGAGAGGAGGAGTCTCT
TACGATTCGATTGAGCTTCGCATGCATCGATTAGCTAGCATCTGATCGATGCCTCGATTG
TACTTACGATCTGACGTATGCTGTGACTGATGCATCGTATCGATGCATCATGCGTATGAC
ATATCATAGTACTGAGTTCTCTTCTCTCTGATCGGGGAGAGGGGGGGCGTATATCGGAGA
GTAAGTACGCATTGGCATCGATTGCAGGACTTAGCGAGAGAGAGCTTCTAGCGTCTAGTA
ATCCCATGATCTACGAGATGCATGCATGCTGATCGACTGATGTATGCTACTGACTGATGT
ATCATCAGATCTGACTGATGCGCTCTGCATGATGCATCGATCGATGCTATCGGATATACG
CGATACGCTGATACGTATGCATGGCATATTCTCTCTCTCGCTGCTGCTTCGTCTGGAAGA
TTTTCAGAGGGCGTATATATACTCTTCTCTCTATAGCTATACGCTGATCAATACGATCGT
GCATGAGACTATGCATGCTGATCGATATCTCTCTGATATCGCATTAGCATCATGCTAGCA

Viterbi Algorithm

$$\delta_t(j) = \begin{cases} \ln(b_1(O_1)) & \text{if } t = 1, j = 1 \\ -\infty & \text{if } t = 1, j \neq 1 \\ \max\limits_{1 \le i \le N} \left(\delta_{t-|b_j|}(i) + \ln(a_{ij})\right) + \ln(b_j(O_{t-|b_j|+1} \dots O_t)) & \text{if } 2 \le t \le T \end{cases}$$

$\delta_1(1)$	$\delta_2(1)$	$\delta_3(1)$	$\delta_4(1)$	$\delta_5(1)$	$\delta_6(1)$	$\delta_7(1)$	$\delta_8(1)$	$\delta_9(1)$	…
$\delta_1(2)$	$\delta_2(2)$	$\delta_3(2)$	$\delta_4(2)$	$\delta_5(2)$	$\delta_6(2)$	$\delta_7(2)$	$\delta_8(2)$	$\delta_9(2)$	…
$\delta_1(3)$	$\delta_2(3)$	$\delta_3(3)$	$\delta_4(3)$	$\delta_5(3)$	$\delta_6(3)$	$\delta_7(3)$	$\delta_8(3)$	$\delta_9(3)$	…
$\delta_1(4)$	$\delta_2(4)$	$\delta_3(4)$	$\delta_4(4)$	$\delta_5(4)$	$\delta_6(4)$	$\delta_7(4)$	$\delta_8(4)$	$\delta_9(4)$	…

* Assuming we start in state 1 and state 1 emits characters of length 1

Observation Sequence

AGCTGTACATCGGCGTATCGCGATGCGATCGGCATGTACGGCGCGTATATGCGCATTAGA

TTTAGCGAGTCTCTCGATTGTGTACTCTCTGCTGCATGATCTATCATACTTAGATTAGTA

CGTCGCCCATGAGATGCGCGCGCAGAGCGAATCTATACTACTACTACCTGCTGCATGATG

GACGTATGCATGTATCGAGAGGATCGATGACTGAGAGGAGGAGTCTCTCTATGATATATA

GGTATGCTTCGCATGCATCGATTAGCTAGCATCTATCTAAGCCTGCGATTGTGTACTCTG

AAAAACACCCAATCGGCGTATCGCGATGCGATCGGCATGTACGGCGCGTATATGCGCATT

ATACAATGCGTCTCTCGATTGTGTACTCTCTGCTGCATGATCTATCATACTTAG

Optimal State Sequence

```
AGCTGTACATCGGCGTATCGCGATGCGATCGGCATGTACGGCGCGTATATGCGCATTAGA
............................................................

TTTAGCGAGTCTCTCGATTGTGTACTCTCTGCTGCATGATCTATCATACTTAGATTAGTA
............................................................

CGTCGCCCATGAGATGCGCGCGCAGAGCGAATCTATACTACTACTACCTGCTGCATGATG
.............S  C  C  C  C  C  C  C  C  C  C  C  C  C  C  C

GACGTATGCATGTATCGAGAGGATCGATGACTGAGAGGAGGAGTCTCTCTATGATATATA
 C  C  C  C  C  C  C  C  C  C  C  C  C  C  C  C  C  C  C  C

GGTATGCTTCGCATGCATCGATTAGCTAGCATCTATCTAAGCCTGCGATTGTGTACTCTG
 C  C  C  C  C  C  C  C  C  C  C  C  S  ....................

AAAAACACCCAATCGGCGTATCGCGATGCGATCGGCATGTACGGCGCGTATATGCGCATT
............................................................

ATACAATGCGTCTCTCGATTGTGTACTCTCTGCTGCATGATCTATCATACTTAG
......................................................
```


Optimal Annotation

```
AGCTGTACATCGGCGTATCGCGATGCGATCGGCATGTACGGCGCGTATATGCGCATTAGA
............................................................

TTTAGCGAGTCTCTCGATTGTGTACTCTCTGCTGCATGATCTATCATACTTAGATTAGTA
............................................................

CGTCGCCCATGAGATGCGCGCGCAGAGCGAATCTATACTACTACTACCTGCTGCATGATG
.............SSSCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCC

GACGTATGCATGTATCGAGAGGATCGATGACTGAGAGGAGGAGTCTCTCTATGATATATA
CCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCC

GGTATGCTTCGCATGCATCGATTAGCTAGCATCTATCTAAGCCTGCGATTGTGTACTCTG
CCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCSSS....................

AAAAACACCCAATCGGCGTATCGCGATGCGATCGGCATGTACGGCGCGTATATGCGCATT
............................................................

ATACAATGCGTCTCTCGATTGTGTACTCTCTGCTGCATGATCTATCATACTTAG
......................................................
```


Parallel Observation Sequence

11	11	11	11	11	11	32	32	32	32	32	32	32	32	32	32
T	G	T	C	C	G	A	G	A	T	G	C	T	T	A	T
.	.	.	.	.	.	.	.	S	S	S	C	C	C	C	C
32	32	84	84	84	84	84	84	84	84	84	84	84	84	84	84
G	G	G	A	C	A	A	T	T	C	G	G	A	C	A	A
C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C
84	47	47	47	47	47	47	47	47	47	08	08	08	08	08	08
T	C	G	T	A	A	A	G	T	G	C	G	C	A	T	T
C	C	C	S	S	S	.	.	.	.	.	.	.	.	.	.
08	08	08	08	08	08	25	25	25	25	25					
G	C	A	G	G	C	C	T	G	A	T					
.	.	.	.	.	.	.	.	.	.	.					

Which Model Parameters Require Updating?

$$\lambda = (A, B)$$

Transition Probabilities, *A*

$a_{.Intergenic\ .Intergenic} = 0.994$

$a_{.Intergenic\ StartCodons} = 0.006$

$a_{StartCodons\ Codons} = 1.0$

$a_{Codons\ Codons} = 0.997$

$a_{Codons\ StopCodons} = 0.003$

$a_{StopCodons\ .Intergenic} = 1.0$

Emission Probabilities, *B*

A	0.28
C	0.22
G	0.22
T	0.28

ATG	0.90
TTG	0.02
GTG	0.08

TAA	0.65
TGA	0.28
TAG	0.07

AAA	0.03
AAC	0.02
AAG	0.01
AAT	0.02
ACA	0.01

ACC	0.02
...	
TTC	0.02
TTG	0.01
TTT	0.02

$b_{.Intergenic}(10) = 0.08$

$b_{.Intergenic}(40) = 0.05$

$b_{.Intergenic}(70) = 0.01$

$b_{StartCodons}(10) = 0.01$

$b_{StartCodons}(40) = 0.03$

$b_{StartCodons}(70) = 0.09$

$b_{Codons}(10) = 0.01$

$b_{Codons}(40) = 0.03$

$b_{Codons}(70) = 0.09$

$b_{StopCodons}(10) = 0.01$

$b_{StopCodons}(40) = 0.03$

$b_{StopCodons}(70) = 0.09$

Parallel Emission Probabilities

For example, the probability that the `Codons` state outputs the parallel observation sequences

P:	**40**
O:	**A**

is given by $b_{Codons}(\texttt{A}) * b_{Codons}(\texttt{40})$

The probability that the `Codons` state outputs the parallel observation sequences

P:	**40**	**40**	**40**	**40**	**40**	**40**
O:	**A**	**C**	**C**	**T**	**T**	**G**

is given by

$b_{Codons}(\texttt{ACC}) * b_{Codons}(\texttt{ACC}) * b_{Codons}(\texttt{40}) * b_{Codons}(\texttt{40}) *$
$b_{Codons}(\texttt{40}) * b_{Codons}(\texttt{40}) * b_{Codons}(\texttt{40}) * b_{Codons}(\texttt{40}) =$

0.02*0.01*0.03*0.03*0.03*0.03*0.03*0.03

Parallel Emission Probabilities

More generally, the probability that state *j* outputs the parallel observation sequences O_x ... O_y and P_x ... P_y is given by

$$b_j(O_x) * \ldots * b_j(O_y) * b_j(P_x) * \ldots * b_j(P_y)$$

Recurrence With Logarithms

$$\delta_t(j) = \begin{cases} \ln(b_1(O_1)) & \text{if } t = 1, j = 1 \\ -\infty & \text{if } t = 1, j \neq 1 \\ \max\limits_{1 \le i \le N}\left(\delta_{t-|b_j|}(i) + \ln(a_{ij})\right) + \ln(b_j(O_{t-|b_j|+1} \ldots O_t)) + \ln(b_j(P_{t-|b_j|+1} \ldots P_t)) & \text{if } 2 \le t \le T \end{cases}$$

$\delta_1(1)$	$\delta_2(1)$	$\delta_3(1)$	$\delta_4(1)$	$\delta_5(1)$	$\delta_6(1)$	$\delta_7(1)$	$\delta_8(1)$	$\delta_9(1)$	$\delta_{10}(1)$
$\delta_1(2)$	$\delta_2(2)$	$\delta_3(2)$	$\delta_4(2)$	$\delta_5(2)$	$\delta_6(2)$	$\delta_7(2)$	$\delta_8(2)$	$\delta_9(2)$	$\delta_{10}(2)$
$\delta_1(3)$	$\delta_2(3)$	$\delta_3(3)$	$\delta_4(3)$	$\delta_5(3)$	$\delta_6(3)$	$\delta_7(3)$	$\delta_8(3)$	$\delta_9(3)$	$\delta_{10}(3)$

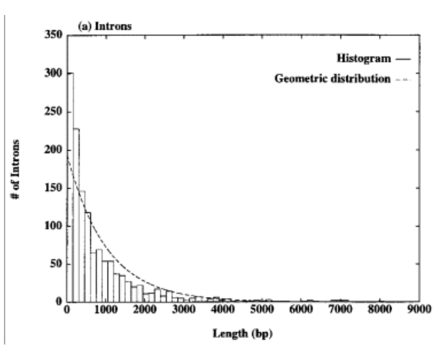
* Assuming we start in state 1 and state 1 emits characters of length 1

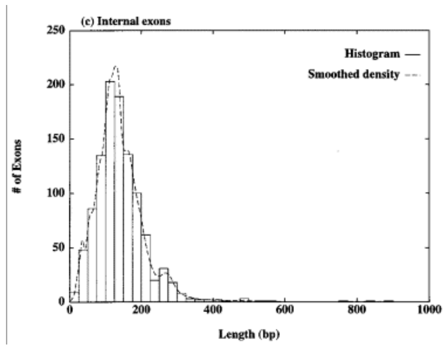
EXTENSION: GMM vs. HMM

Duration of Time Spent in State

Probability of d consecutive observations in state j:

$$p_j(d) = (a_{jj})^{d-1} * (1 - a_{jj})$$





Example HMM Model, λ

$$\lambda = (A, B, C)$$

Transition Probabilities, A

$a_{\text{Intergenic Intergenic}} = 0.994$

$a_{\text{Intergenic StartCodons}} = 0.006$

$a_{\text{StartCodons Codons}} = 1.0$

$a_{\text{Codons Codons}} = 0.997$

$a_{\text{Codons StopCodons}} = 0.003$

$a_{\text{StopCodons Intergenic}} = 1.0$

Emission Probabilities, B

A	0.28
C	0.22
G	0.22
T	0.28

ATG	0.90
TTG	0.02
GTG	0.08

TAA	0.65
TGA	0.28
TAG	0.07

AAA	0.03
AAC	0.02
AAG	0.01
AAT	0.02
ACA	0.01

ACC	0.02
...	
TTC	0.02
TTG	0.01
TTT	0.02

Duration Probabilities, C

$c_{\text{Codons}}(50) = 0.001$

$c_{\text{Codons}}(100) = 0.002$

$c_{\text{Codons}}(150) = 0.003$

$c_{\text{Codons}}(200) = 0.004$

$c_{\text{Codons}}(250) = 0.004$

$c_{\text{Codons}}(300) = 0.003$

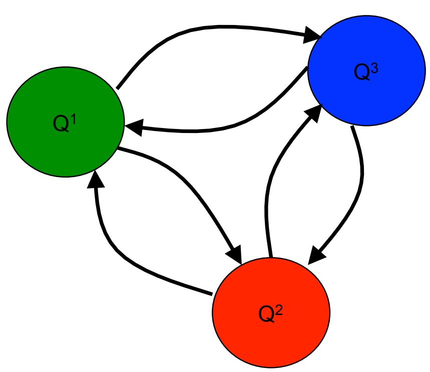
$c_{\text{Codons}}(350) = 0.002$

$c_{\text{Codons}}(400) = 0.001$

OR

$c_{\text{Codons}}(d) = \text{Gamma}(d, \textit{shape}, \textit{scale})$

Generating an Observation Sequence



1. Begin in initial state
2. Determine duration, i.e., how many characters to output (emit) in the current state
3. Emit the determined number of characters in the current state
4. Transition to a *different* state
5. Return to step #2 and repeat

Recurrence With Logarithms

$$\delta_t(j) = \begin{cases} \ln(b_1(O_1)) & \text{if } t = 1, j = 1 \\ -\infty & \text{if } t = 1, j \neq 1 \\ \max\limits_{\text{minlength}_j \le d \le \text{maxlength}_j} \left(\max\limits_{1 \le i \le N} \left(\delta_{t-d}(i) + \ln(a_{ij}) \right) + \ln(c_j(d)) + \ln(b_j(O_{t-d+1} \ldots O_t)) \right) & \text{if } 2 \le t \le T \end{cases}$$

$\delta_1(1)$	$\delta_2(1)$	$\delta_3(1)$	$\delta_4(1)$	$\delta_5(1)$	$\delta_6(1)$	$\delta_7(1)$	$\delta_8(1)$	$\delta_9(1)$	$\delta_{10}(1)$
$\delta_1(2)$	$\delta_2(2)$	$\delta_3(2)$	$\delta_4(2)$	$\delta_5(2)$	$\delta_6(2)$	$\delta_7(2)$	$\delta_8(2)$	$\delta_9(2)$	$\delta_{10}(2)$
$\delta_1(3)$	$\delta_2(3)$	$\delta_3(3)$	$\delta_4(3)$	$\delta_5(3)$	$\delta_6(3)$	$\delta_7(3)$	$\delta_8(3)$	$\delta_9(3)$	$\delta_{10}(3)$

* Assuming we start in state 1 and state 1 emits characters of length 1