# Sistemas Inteligentes
# Cuestiones y ejercicios del bloque 2, temas 5, 6 y 7
# Modelos de Markov

Escola Tècnica Superior d'Informàtica
Dep. de Sistemes Informàtics i Computació
Universitat Politècnica de València

10 de noviembre de 2014

## 1. Cuestiones

### 1.1. Modelos de Markov: definición, topologías, usos, etc.

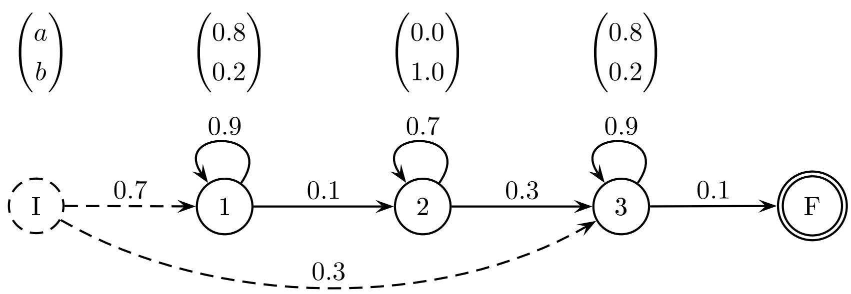
1 B Los siguientes enunciados se refieren al clasificador basado en modelos de Markov. ¿Cuál de ellos es incorrecto?

A) Puede verse como una particularización del clasificador de Bayes para cadenas, en la que las funciones de probabilidad condicionales de las clases vienen dadas por modelos de Markov.
B) En reconocimiento por Viterbi, aplicaremos la regla de Bayes sustituyendo probabilidades exactas por aproximaciones de Viterbi; esto es, la cadena a clasificar se asignará a la clase cuyo modelo obtenga mayor probabilidad de generarla según el algoritmo de Viterbi.
C) En aprendizaje por Viterbi, aplicaremos el algoritmo de re-estimación por Viterbi en cada clase por separado. Cada modelo tendrá un probabilidad comparativamente alta de generar las cadenas de entrenamiento de su clase.
D) Un problema de diseño crucial será la elección del número de estados y topología de cada modelo de Markov. Asimismo, tendremos que escoger un criterio adecuado de inicialización para el algoritmo de re-estimación por Viterbi.

2 C Identifica el enunciado incorrecto: los modelos de Markov...

A) son equivalentes a las gramáticas regulares estocásticas.
B) son adecuados para procesos unidimensionales variantes con el tiempo.
C) no son aplicables a tareas de OCR ya que las imágenes son objetos bidimensionales.
D) son útiles en reconocimiento del habla para lo que generalmente se usan topologías izquierda-derecha.

3 D Sean $M$ un modelo de Markov de representación gráfica:



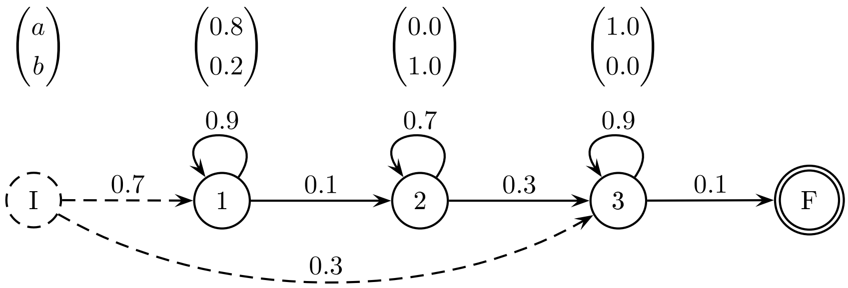

¿Cuántas cadenas distintas de longitud 3 puede generar $M$?

A) Ninguna.
B) Entre 1 y 3.
C) Entre 3 y 6.
D) Más de 6.

4 C Sea $M$ un modelo de Markov de probabilidades de transición no nulas, $A_{q,q'} > 0$, para todo estado no final $q$ y todo estado $q'$ (incluido el final). Decimos que la topología de $M$ es...

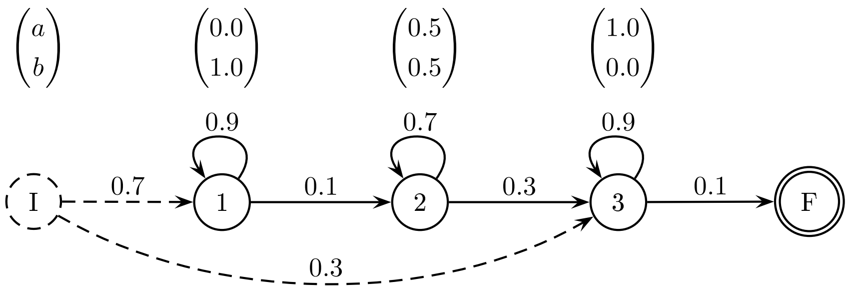
A) Lineal.
B) Estrictamente lineal.
C) Ergódica.
D) No se puede decir nada mientras no conozcamos también las probabilidades iniciales y las de emisión.

5 $\boxed{\text{C}}$ Sean $M$ un modelo de Markov de representación gráfica:




A) No puede generar la cadena *abaa*.
B) Existe más de un camino que genera la cadena *aaabaa*.
C) Existe sólo un camino que genera la cadena *bbaa*.
D) No puede generar cadenas de longitud mayor que 5 y que empiecen por *b*.

6 $\boxed{\text{C}}$ Sean $M$ un modelo de Markov de representación gráfica:




A) No puede generar cadenas que empiecen por el símbolo *a*
B) Existe sólo un camino que genera la cadena *bbaa*
C) Existe sólo un camino que genera la cadena *baba*
D) No puede generar cadenas de longitud mayor que 5 y que empiecen por *a*

7 $\boxed{\text{D}}$ Sea $M$ un modelo de Markov, con conjunto de estados $Q = \{1, 2, ..., F\}$, de probabilidades de transición no nulas, $A_{q,q'} > 0$ si y solo si $q < q'$, para todo estado no final $q$ y todo estado $q'$ (incluido el final $q_F$, $q_F > q, \forall q$). Además se tiene que $\pi_1 = 1$. Decimos que la topología de $M$ es...

A) Lineal.
B) Estrictamente lineal.
C) Ergódica.
D) Izquierda-derecha

8 $\boxed{\text{C}}$ Los siguientes enunciados se refieren al clasificador basado en modelos de Markov. ¿Cuál de ellos es incorrecto?

A) Un problema de diseño crucial será la elección del número de estados y topología de cada modelo de Markov.
B) Puede verse como una particularización del clasificador de Bayes para cadenas, en la que las funciones de probabilidad condicionales de las clases vienen dadas por modelos de Markov.
C) En reconocimiento por Viterbi, aplicaremos la regla de Bayes y la cadena a clasificar se asignará a la clase cuyo modelo la genere con mayor probabilidad, según la aproximación de Viterbi.
D) En aprendizaje por Viterbi, aplicaremos el algoritmo de re-estimación por Viterbi en cada clase por separado. Este algoritmo garantiza que la verosimilitud de las muestras de entrenamiento no decrece en las sucesivas iteraciones

## 1.2. Modelos de Markov: probabilidad de (generar) una cadena y aproximación de Viterbi

1 $\boxed{\text{C}}$ Identifica el enunciado incorrecto: dado un modelo de Markov $M$ y una cadena $x \in \Sigma^\star$, la probabilidad $P(x|M)$ es...

A) igual a la suma de probabilidades de analizar $x$ mediante todos los caminos posibles para $x$ en $M$.
B) igual a 0 si $\bar{x}$ no se puede analizar mediante $M$.
C) menor que $P(xa|M)$ para cualquier $a \in \Sigma$.
D) siempre mayor que $\tilde{P}(x|M)$ (aproximación de Viterbi).

2 $\boxed{B}$ Los siguientes enunciados se refieren a la aproximación por Viterbi a la probabilidad de que un modelo de Markov genere una cadena dada, $\tilde{P}(y|M)$,

A) $\tilde{P}(y|M)$ es siempre mayor que 0.
B) $\tilde{P}(y|M)$ es menor que la verdadera probabilidad de generación de la cadena, $P(y|M)$.
C) $\tilde{P}(y|M)$ tiene mucho interés porque es una probabilidad muy cercana a 1.
D) El coste de calcular $\tilde{P}(y|M)$ es exponencial con el número de estados del modelo.

3 $\boxed{B}$ Sea el siguiente modelo de Markov $M$: $Q = \{1, 2, F\}$, $\pi_1 = 0$, $\pi_2 = 1$, $\Sigma = \{a, b\}$,

| A | 1 | 2 | F |
|---|---|---|---|
| 1 | 0.2 | 0.8 | 0.0 |
| 2 | 0.1 | 0.7 | 0.2 |

| B | a | b |
|---|---|---|
| 1 | 0 | 1 |
| 2 | 1 | 0 |

Calcula $P(aba|M)$ e indica el resultado:

A) 0.128
B) 0.016
C) 0.002
D) 0.0

4 $\boxed{A}$ Los siguientes enunciados se refieren a la aproximación por Viterbi $\tilde{P}(y|M)$ a la probabilidad $P(y|M)$ con que un modelo de Markov genere una cadena dada. Indica cuál es correcto.

A) $\tilde{P}(y|M) \leq P(y|M)$
B) $\tilde{P}(y|M) \approx P(y|M)$
C) $\tilde{P}(y|M) \approx 1$
D) $\tilde{P}(y|M)$ tiene mucho interés ya que su coste computacional es polinómico con el número de estados del modelo, mientras que el coste de calcular $P(y|M)$ es exponencial en el peor de los casos.

5 $\boxed{C}$ Sea el siguiente modelo de Markov $M$: $Q = \{1, 2, F\}$, $\pi_1 = 0$, $\pi_2 = 1$, $\Sigma = \{a, b\}$,

| A | 1 | 2 | F |
|---|---|---|---|
| 1 | 0.2 | 0.8 | 0.0 |
| 2 | 0.1 | 0.7 | 0.2 |

| B | a | b |
|---|---|---|
| 1 | 0 | 1 |
| 2 | 1 | 0 |

Calcula $P(abba|M)$ e indica el resultado:

A) 0.0128
B) 0.0016
C) 0.0032
D) 0.0

6 $\boxed{D}$ Sea el siguiente modelo de Markov $M$: $Q = \{1, 2, F\}$, $\pi_1 = 0$, $\pi_2 = 1$, $\Sigma = \{a, b\}$,

| A | 1 | 2 | F |
|---|---|---|---|
| 1 | 0.4 | 0.6 | 0.0 |
| 2 | 0.5 | 0.4 | 0.1 |

| B | a | b |
|---|---|---|
| 1 | 0 | 1 |
| 2 | 1 | 0 |

Calcula $P(aba|M)$ e indica en qué intervalo de los siguientes se halla:

A) Menor que 0.001.
B) Mayor que 0.001 y menor que 0.005.
C) Mayor que 0.005 y menor que 0.010.
D) Mayor que 0.010.

7 $\boxed{D}$ Sea $p$ la probabilidad con la que un modelo de Markov genera cierta cadena $x$ y $q$ la correspondiente aproximación de Viterbi. Se puede afirmar que:

A) $p < q$ en cualquier caso.
B) $p > q$ si y solo si $x$ es suficientemente corta.
C) $p \leq q$ si y solo si $x$ es suficientemente larga.
D) $p \geq q$ en cualquier caso.

8 $\boxed{B}$ Sean $M$ un model de Markov de alfabeto $\Sigma$, $x \in \Sigma^\star$ una cadena arbitraria, $p_M(x)$ la probabilidad de que $M$ genere $x$ y $\tilde{p}_M(x)$ la aproximación de Viterbi a $p_M(x)$. Se cumple que:

A) $\tilde{p}_M(x)$ es siempre menor que $p_M(x)$.
B) $\tilde{p}_M(x)$ es menor que $p_M(x)$ cuando en $M$ existe más de un camino que genera $x$; es igual cuando sólo existe un camino generador de $x$.
C) $\tilde{p}_M(x)$ es siempre menor o igual que $p_M(x)$, pero el ser una cosa u otra no depende del número de caminos distintos que generan $x$.
D) Ninguna de las anteriores.

9 $\boxed{\text{D}}$ La probabilidad exacta de que un modelo de Markov $M$ genere una cadena $x$…

A) no se suele calcular ya que su coste de cálculo es al menos exponencial con la longitud de $x$.
B) se calcula mediante el algoritmo de Viterbi.
C) se calcula mediante el algoritmo de re-estimación por Viterbi.
D) Ninguna de las anteriores.

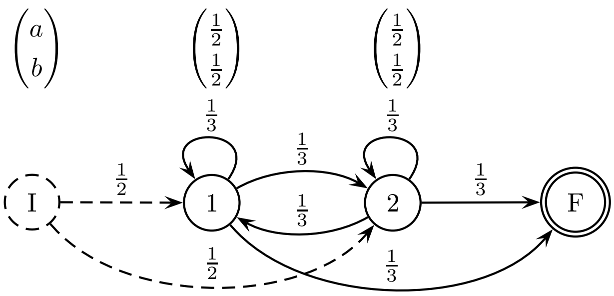
10 $\boxed{\text{D}}$ La aproximación por Viterbi a la probabilidad exacta de que un modelo de Markov $M$ genere una cadena $x$, $\hat{P}(x \mid M)$:

A) Sólo se puede calcular en modelos de Markov de topología lineal o ergódica.
B) Siempre sera mayor o igual que la verdadera probabilidad $P(x \mid M)$.
C) Siempre sera menor que la verdadera probabilidad $P(x \mid M)$.
D) Será nula sólo si la probabilidad $P(x, z \mid M)$ también es nula para cualquier secuencia de estados $z$.

11 $\boxed{\text{D}}$ La aproximación por Viterbi a la probabilidad exacta de que un modelo de Markov $M$ genere una cadena $x$, $\hat{P}(x \mid M)$:

A) Será nula si la probabilidad $P(x, z \mid M)$ es nula para una secuencia cualquiera de las secuencias de estados $z$.
B) Siempre sera mayor o igual que la verdadera probabilidad $P(x \mid M)$.
C) Será no nula sólo si la probabilidad $P(x, z \mid M)$ es no nula para toda secuencia de estados $z$.
D) Siempre sera menor o igual que la verdadera probabilidad $P(x \mid M)$.

12 $\boxed{\text{C}}$ Sea $M$ el modelo de Markov representado en la figura a la derecha. Considérese una cadena arbitraria de longitud $T$, $x = x_1 x_2 \cdots x_T \in \{a, b\}^T$, la probabilidad de que $M$ la genere, $p_M(x)$, así como la aproximación de Viterbi a esta probabilidad, $\tilde{p}_M(x)$. Entonces:



A) $p_M(x) = \tilde{p}_M(x)$
B) $\dfrac{\tilde{p}_M(x)}{p_M(x)} < 1$, aumentando esta relación al aumentar $T$
C) $\dfrac{\tilde{p}_M(x)}{p_M(x)} < 1$, disminuyendo esta relación al aumentar $T$
D) Ninguna de los anteriores

13 $\boxed{\text{C}}$ Sea el siguiente modelo de Markov $M$: $Q = \{1, 2, F\}$, $\pi_1 = 0$, $\pi_2 = 1$, $\Sigma = \{a, b\}$,

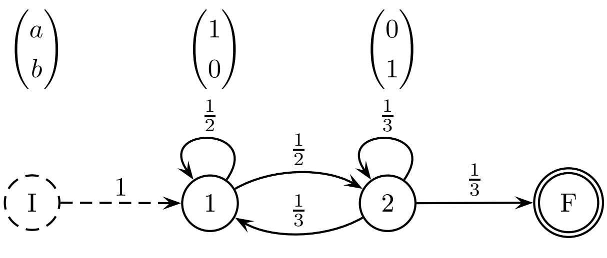
| A | 1 | 2 | F |
|---|---|---|---|
| 1 | 0.2 | 0.8 | 0.0 |
| 2 | 0.1 | 0.7 | 0.2 |

| B | a | b |
|---|---|---|
| 1 | 0 | 1 |
| 2 | 1 | 0 |

Calcula $P(abba|M)$:

A) 0.0128
B) 0.0016
C) 0.0032
D) 0.0

14 $\boxed{\text{D}}$ Sea $M$ el modelo de Markov representado en la figura de la derecha. Sea $\tilde{P}(y \mid M)$ la aproximación por Viterbi a la verdadera probabilidad $P(y \mid M)$. Cuál de las siguientes afirmaciones es cierta:



A) $\tilde{P}(a^n b^n \mid M) < P(a^n b^n \mid M), n > 0$
B) $P(a^n|M) > 0, n > 0$
C) $P(b^n|M) > 0, n > 0$
D) $\tilde{P}((abab)^n \mid M) = P((abab)^n \mid M), n > 0$

15 $\boxed{\text{B}}$ Los siguientes enunciados se refieren a la aproximación por Viterbi a la probabilidad de que un modelo de Markov $M$ genere una cadena $y$, $\tilde{P}(y|M)$, indica cuál de las siguientes afirmaciones es verdadera:

A) $\tilde{P}(y|M)$ es siempre mayor que 0
B) $\tilde{P}(y|M)$ nunca es mayor que la verdadera probabilidad de generación de la cadena, $P(y|M)$
C) $\tilde{P}(y|M)$ nunca puede llegar a 1
D) El coste de calcular $\tilde{P}(y|M)$ es exponencial con el número de estados del modelo, pero mediante técnicas de programación dinámica el coste es lineal con el número de estados de $M$ para cualquier $M$

16 $\boxed{\text{A}}$ Sea $P(y)$ la probabilidad con que un modelo de Markov genere una cadena dada y sea $\tilde{P}(y)$ la aproximación de Viterbi a $P(y)$. ¿Cuál de los siguientes enunciados es cierto?

A) $\tilde{P}(y) \leq \sum_z P(z)P(y \mid z) = P(y)$, donde $z$ denota una secuencia de estados del modelo
B) $\tilde{P}(y) > 0 \; \forall y$
C) $\tilde{P}(y) \approx 0$
D) El coste de calcular $\tilde{P}(y)$ es cuadrático con la longitud (número de símnolos) de $y$

17 $\boxed{\text{B}}$ Sea $\mathcal{M}$ un modelo de Markov ergódico con $Q = \{1, 2, F\}$, $\Sigma = \{\text{a,b}\}$, y probabilidades equiprobables de *estado inicial*, de *transición* y de *generación*.

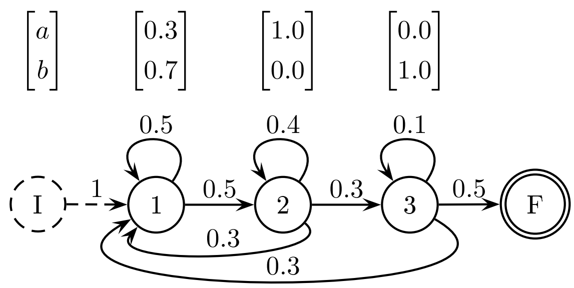
Calcula $P(\text{“ab”} \mid \mathcal{M})$:

A) 0.0333
B) 1/18
C) 1/3
D) 4/27

18 $\boxed{\text{D}}$ ¿Cuál de los siguientes enunciados sobre la aproximación por Viterbi, $\tilde{P}(y|M)$, a la probabilidad de que un modelo de Markov genere una cadena dada es correcto?

A) $\tilde{P}(y|M)$ es siempre mayor que 0
B) $\tilde{P}(y|M)$ tiene mucho interés porque es una probabilidad muy cercana a 1
C) El coste de calcular $\tilde{P}(y|M)$ es exponencial con el número de estados del modelo
D) $\tilde{P}(y|M)$ es siempre algo menor o igual que la verdadera probabilidad de generación de la cadena, $P(y|M)$

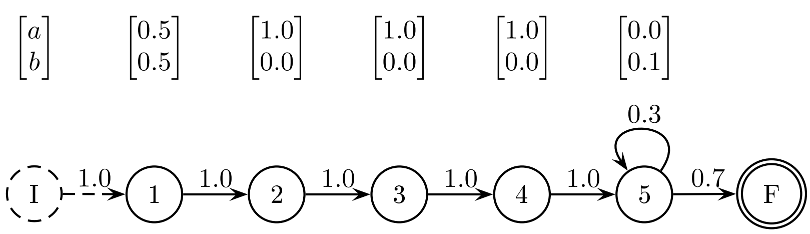
19 $\boxed{\text{A}}$ (Examen de SIN del 18 de enero de 2013)
Al analizar la cadena *aabb* con el modelo de Markov adjunto, ¿Cuál es la probabilidad de generación de la cadena por el modelo y cuál es la probabilidad de generar la cadena por la mejor secuencia de estados en el modelo?:



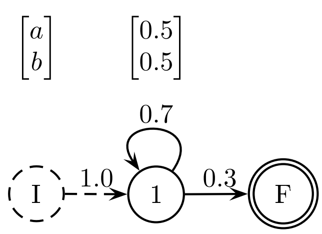

A) Total 0.00225; mejor camino 0.00225.
B) Total 0.0045; mejor camino 0.00225.
C) Total 0.00225; mejor camino 0.000225.
D) Total 0.00225; mejor camino 0.0045.

20 $\boxed{\text{B}}$ (Examen de SIN del 18 de enero de 2013)
Dados los dos modelos ocultos de Markov $M_1$ y $M_2$ siguientes:

Modelo $M_1$




Modelo $M_2$



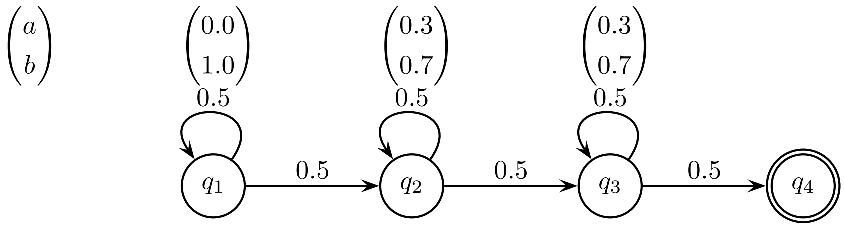

¿Qué modelo producirá la mayor probabilidad para la cadena “aaaabb”? ¿Y para “aaab”?

A) Los dos modelos igual para cada una de las cadenas.
B) $M_1$ para la primera y $M_2$ para la segunda.
C) $M_2$ para la primera y $M_1$ para la segunda.
D) $M_1$ para las dos cadenas.

21 $\boxed{\text{C}}$ (Examen de SIN del 15 de enero de 2014; examen del bloque 2; cuestión 15)
Dado un Modelo Oculto de Markov $\Theta$ y una cadena $y$ reconocida por dicho modelo, indica cuál de las siguientes afirmaciones es cierta:

A) Siempre se cumple que $P(y|\Theta) = \widetilde{P}(y|\Theta)$.
B) Siempre se cumple que $P(y|\Theta) \leq \widetilde{P}(y|\Theta)$.
C) Siempre se cumple que $P(y|\Theta) \geq \widetilde{P}(y|\Theta)$.
D) Siempre se cumple que $P(y|\Theta) \neq \widetilde{P}(y|\Theta)$.

22 $\boxed{\text{C}}$ (Examen de SIN del 15 de enero de 2014; examen del bloque 2; cuestión 16)
Dado el Modelo Oculto de Markov $\Theta$




con $\pi_{q_1} = 1, \pi_{q_2} = \pi_{q_3} = \pi_{q_4} = 0$ y las cadenas $y_1 =$ "babb" y $y_2 =$ "aaaa", indica cuál de las siguientes afirmaciones es cierta:

A) $P(y_1|\Theta) = P(y_2|\Theta)$.
B) $P(y_1|\Theta) < P(y_2|\Theta)$.
C) $P(y_1|\Theta) > P(y_2|\Theta)$.
D) $P(y_1|\Theta) = P(y_2|\Theta) = 0$.

23 $\boxed{\text{C}}$ (Examen de SIN del 15 de enero de 2014; examen del bloque 2; cuestión 18)
En relación al algoritmo *forward* definido para Modelos Ocultos de Markov indica cuál de la siguientes afirmaciones es verdadera:

A) Calcula la probabilidad de una cadena teniendo en cuenta sólo la secuencia de análisis de máxima probabilidad.
B) Calcula la probabilidad de una cadena sin incluir la probabilidad de la secuencia de estados más probable.
C) Calcula la probabilidad de una cadena incluyendo todas las secuencias de estados.
D) Nunca calcula la probabilidad de una cadena.

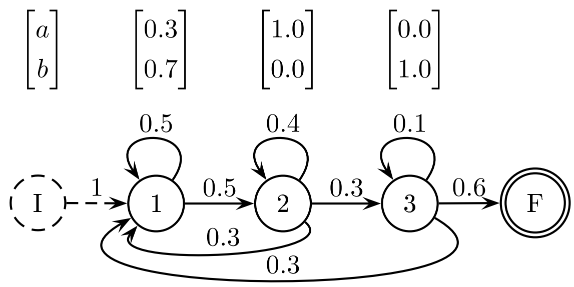
24 $\boxed{\text{B}}$ (Examen de SIN del 28 de enero de 2014; examen final; cuestión 6)
Dado un Modelo Oculto de Markov $\Theta$ y una cadena $y$ aceptada por dicho modelo, ¿cuál de las siguientes afirmaciones sobre los algoritmos *forward* y Viterbi es verdadera?

A) *Forward* y Viterbi calculan $P(y|\Theta)$.
B) *Forward* calcula $P(y|\Theta)$ y Viterbi $\widetilde{P}(y|\Theta)$.
C) *Forward* calcula $\widetilde{P}(y|\Theta)$ y Viterbi $P(y|\Theta)$.
D) *Forward* y Viterbi calculan $\widetilde{P}(y|\Theta)$.

## 1.3. Modelos de Markov: re-estimación por Viterbi

1 $\boxed{\text{C}}$ (Examen de SIN del 30 de enero de 2013)
Si se utiliza sólo la cadena $aabb$ para entrenar un modelo de Markov partiendo del modelo adjunto, ¿Cuántas iteraciones sin contar la de verificación final dará el algoritmo de estimación por Viterbi?:

A) Ninguna
B) 2
C) 1
D) 5




2 $\boxed{\text{D}}$ (Examen de SIN del 15 de enero de 2014; examen del bloque 2; cuestión 17)
Dado el Modelo Oculto de Markov $\Theta$ de la cuestión 22 de la sección 1.2, si lo estimamos con una sola iteración con la muestra $Y = \{baba, abab\}$ utilizando el algoritmo de Viterbi indica cuál de las siguientes afirmaciones es cierta:

A) El modelo estimado tiene todos los parámetros a 0.0.
B) Ninguna probabilidad de transición entre estados en el modelo estimado toma valor 0.0.
C) El modelo obtenido tiene todos los parámetros igual al modelo incial.
D) El modelo obtenido queda con varios parámetros con valor 0.

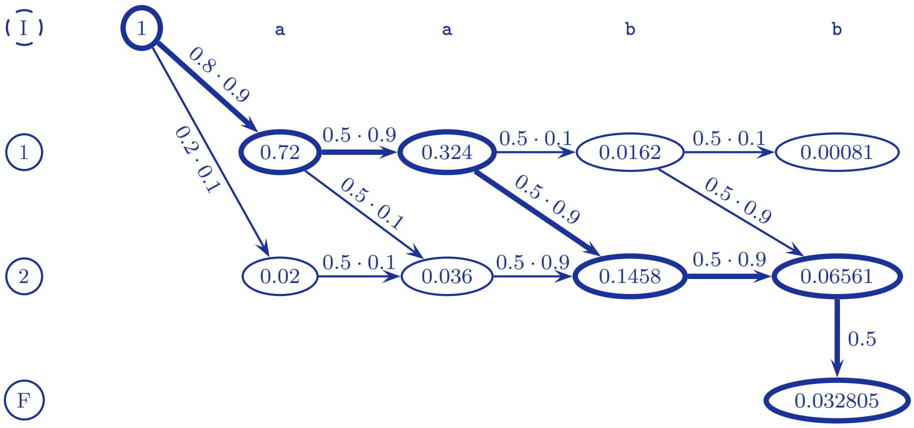
## 2. Problemas

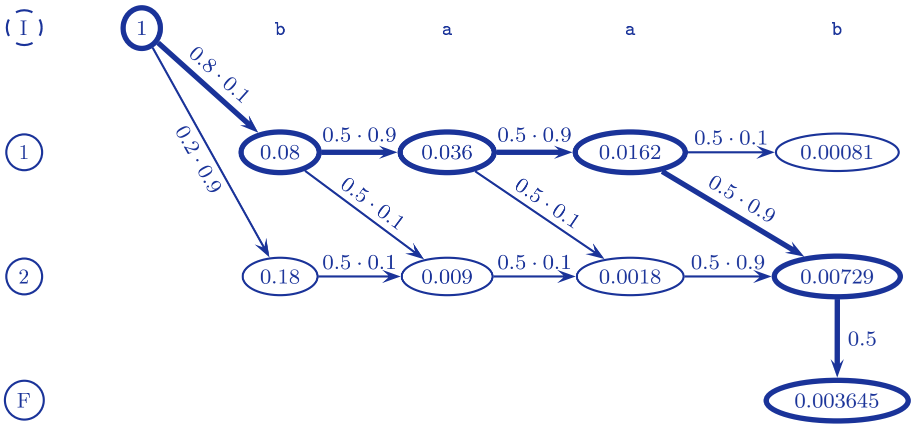
1. Sea $M$ un modelo de Markov de conjunto de estados $Q = \{1, 2, F\}$; alfabeto $\Sigma = \{a, b\}$; probabilidades iniciales $\pi_1 = 0.8, \pi_2 = 0.2, \pi_F = 0.0$; y probabilidades de transición entre estados y de emisión de símbolos:

| $A$ | 1 | 2 | $F$ |
|---|---|---|---|
| 1 | 0.5 | 0.5 | 0 |
| 2 | 0 | 0.5 | 0.5 |

| $B$ | $a$ | $b$ |
|---|---|---|
| 1 | 0.9 | 0.1 |
| 2 | 0.1 | 0.9 |

Reestima los parámetros de $M$ mediante *una* iteración de reestimación por Viterbi, a partir de las cadenas de entrenamiento "`aabb`" y "`baab`".




Los pares *cadena-secuencia óptima de estados* a considerar son:

| aabb | baab |
|---|---|
| 1122F | 1112F |

Los parámetros reestimados son:

$$\hat{\pi}_1 = \frac{2}{2} = 1$$

$$\hat{\pi}_2 = \frac{0}{2} = 0$$

| $A$ | 1 | 2 | $F$ |
|---|---|---|---|
| 1 | $\frac{3}{5}$ | $\frac{2}{5}$ | 0 |
| 2 | 0 | $\frac{1}{3}$ | $\frac{2}{3}$ |

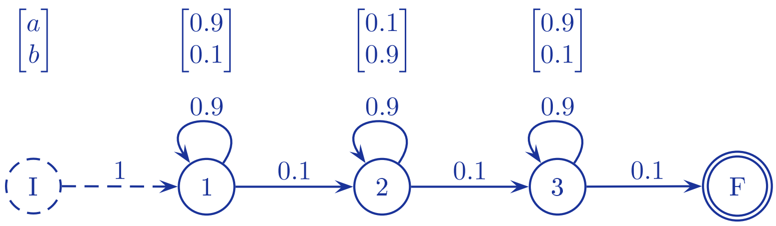
| $B$ | $a$ | $b$ |
|---|---|---|
| 1 | $\frac{4}{5}$ | $\frac{1}{5}$ |
| 2 | 0 | 1 |

2. Sea $M$ un modelo de Markov de conjunto de estados $Q = \{1, 2, 3, F\}$; alfabeto $\Sigma = \{a, b\}$; probabilidades iniciales $\pi_1 = 1$, $\pi_2 = \pi_3 = \pi_F = 0$; y probabilidades de transición y de emisión:

| $A$ | 1 | 2 | 3 | $F$ |
|---|---|---|---|---|
| 1 | 0.9 | 0.1 | 0.0 | 0.0 |
| 2 | 0.0 | 0.9 | 0.1 | 0.0 |
| 3 | 0.0 | 0.0 | 0.9 | 0.1 |

| $B$ | $a$ | $b$ |
|---|---|---|
| 1 | 0.9 | 0.1 |
| 2 | 0.1 | 0.9 |
| 3 | 0.9 | 0.1 |

*a*) Representa gráficamente el modelo $M$.




*b*) Determina las secuencias de estados que generan cadenas de 4 ó menos símbolos.

Cadenas de menos de 3 de símbolos: no pueden generarse

Cadenas de 3 símbolos: $(1, 2, 3)$

Cadenas de 4 símbolos: $(1, 1, 2, 3)$, $(1, 2, 2, 3)$ y $(1, 2, 3, 3)$

*c*) Calcula las probabilidades con las que $M$ genera las cadenas "aaa" y "aab".

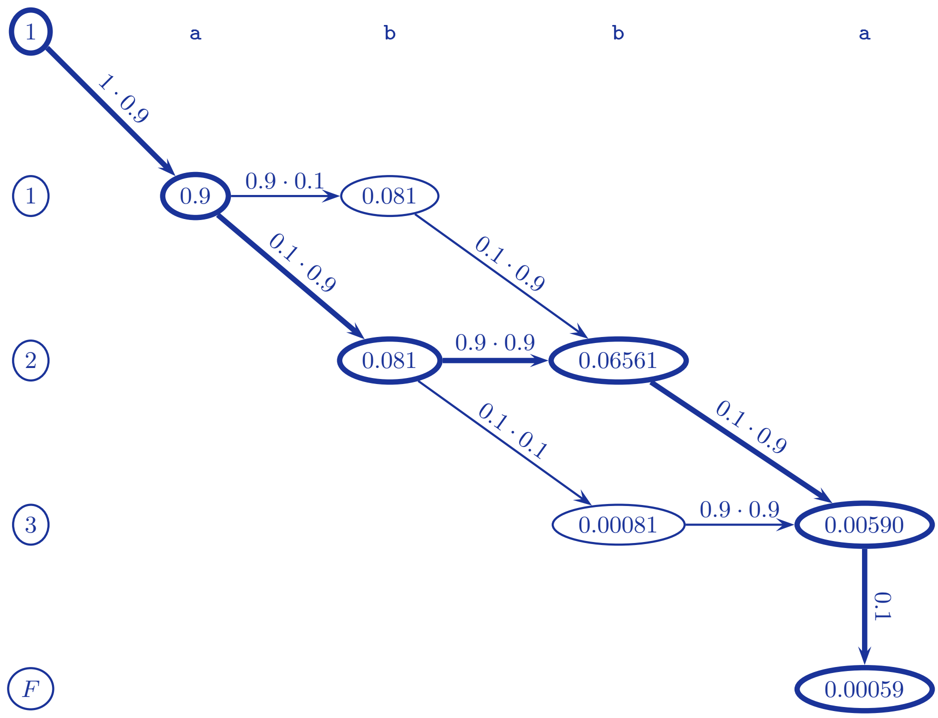
$$P(\texttt{aaa}|M) = (1 \cdot 0.9)\ (0.1 \cdot 0.1)\ (0.1 \cdot 0.9)\ 0.1 = 0.000081$$
$$P(\texttt{aab}|M) = (1 \cdot 0.9)\ (0.1 \cdot 0.1)\ (0.1 \cdot 0.1)\ 0.1 = 0.000009$$

*d*) Calcula la probabilidad de que $M$ genere una cadena de 3 símbolos (arbitraria).

Puede calcularse exhaustivamente, sumando las probabilidades con las que $M$ genera cada una de las 8 cadenas posibles. Más directamente:

$$\begin{aligned}\sum_{x_1x_2x_3 \in \Sigma^+} P(x_1x_2x_3|M) &= \sum_{x_1x_2x_3 \in \Sigma^+} \pi_1\ B_{1,x_1}\ A_{1,2}\ B_{2,x_2}\ A_{2,3}\ B_{3,x_3}\ A_{3,F} \\ &= \pi_1\ A_{1,2}\ A_{2,3}\ A_{3,F} \sum_{x_1x_2x_3 \in \Sigma^+} B_{1,x_1}\ B_{2,x_2}\ B_{3,x_3} \\ &= 1\ 0.1\ 0.1\ 0.1\ 1 = 0.001\end{aligned}$$

*e*) Realiza una traza del algoritmo de Viterbi para obtener la secuencia de estados más probable con la que $M$ genera la cadena "abba".



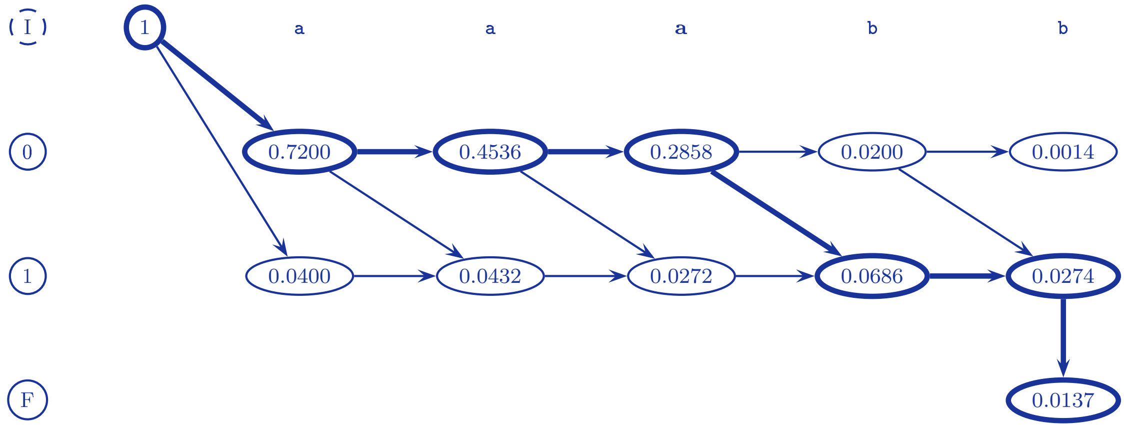

$$\tilde{Q} = (1, 2, 2, 3, F)$$

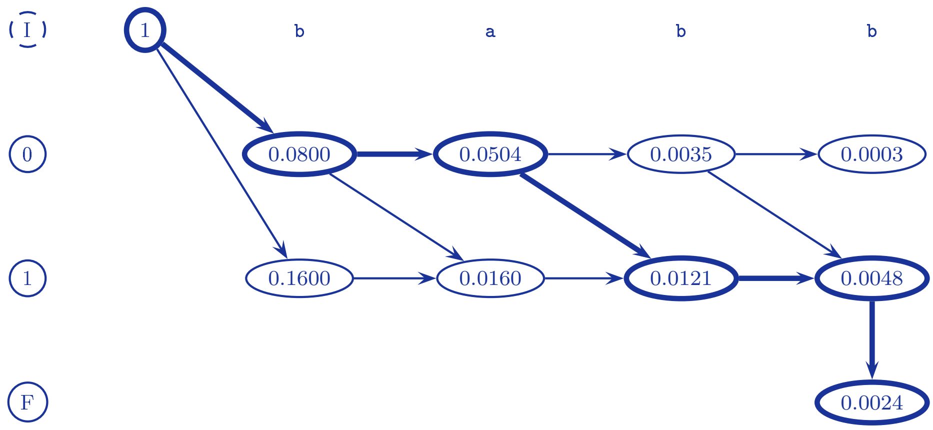
3. Sea $M$ un modelo de Markov de conjunto de estados $Q = \{0, 1, F\}$; alfabeto $\Sigma = \{a, b\}$; probabilidades iniciales $\pi_0(0) = 0.8, \pi_0(1) = 0.2, \pi_0(F) = 0.0$; y probabilidades de transición entre estados y de emisión de símbolos:

| $A$ | 0 | 1 | $F$ |
|---|---|---|---|
| 0 | 0.7 | 0.3 | 0 |
| 1 | 0 | 0.5 | 0.5 |

| $B$ | $a$ | $b$ |
|---|---|---|
| 0 | 0.9 | 0.1 |
| 1 | 0.2 | 0.8 |

Reestima los parámetros de $M$ mediante *una* iteración de reestimación por Viterbi, a partir de las cadenas de entrenamiento "`aaabb`" y "`babb`".




Los pares *cadena-secuencia óptima de estados* a considerar son:

| | |
|---|---|
| aaabb | babb |
| 00011F | 0011F |

Los parámetros reestimados son:

$\hat{\pi}_0(0) = 1$
$\hat{\pi}_0(1) = 0$

| $A$ | 0 | 1 | $F$ |
|---|---|---|---|
| 0 | $\frac{3}{5}$ | $\frac{2}{5}$ | 0 |
| 1 | 0 | $\frac{2}{4}$ | $\frac{2}{4}$ |

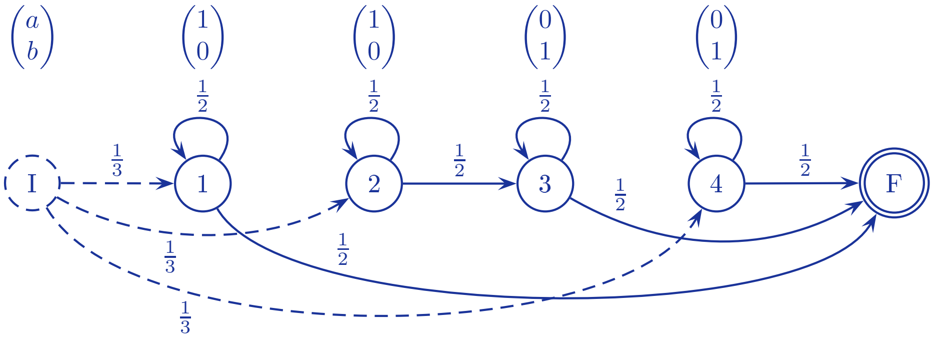
| $B$ | $a$ | $b$ |
|---|---|---|
| 0 | $\frac{4}{5}$ | $\frac{1}{5}$ |
| 1 | 0 | 1 |

4. Sea $M$ un modelo de Markov de conjunto de estados $Q = \{1, 2, 3, 4, F\}$; alfabeto $\Sigma = \{a, b\}$; probabilidades iniciales $\pi_0(1) = \pi_0(2) = \pi_0(4) = \frac{1}{3}$, $\pi_0(3) = \pi_0(F) = 0$; y probabilidades de transición entre estados y de emisión de símbolos:

| $A$ | 1 | 2 | 3 | 4 | $F$ |
|---|---|---|---|---|---|
| 1 | $\frac{1}{2}$ | 0 | 0 | 0 | $\frac{1}{2}$ |
| 2 | 0 | $\frac{1}{2}$ | $\frac{1}{2}$ | 0 | 0 |
| 3 | 0 | 0 | $\frac{1}{2}$ | 0 | $\frac{1}{2}$ |
| 4 | 0 | 0 | 0 | $\frac{1}{2}$ | $\frac{1}{2}$ |

| $B$ | $a$ | $b$ |
|---|---|---|
| 1 | 1 | 0 |
| 2 | 1 | 0 |
| 3 | 0 | 1 |
| 4 | 0 | 1 |

*a*) Representa gráficamente el modelo $M$.




*b*) Sean $n$ y $m$ dos enteros positivos. Halla las probabilidades: $P(a^n \mid M)$, $P(a^n\, b^m \mid M)$ y $P(b^n \mid M)$.

$$P(a^n \mid M) = \frac{1}{3}\frac{1}{2^n}$$
$$P(a^n\, b^m \mid M) = \frac{1}{3}\frac{1}{2^{n+m}}$$
$$P(b^n \mid M) = \frac{1}{3}\frac{1}{2^n}$$

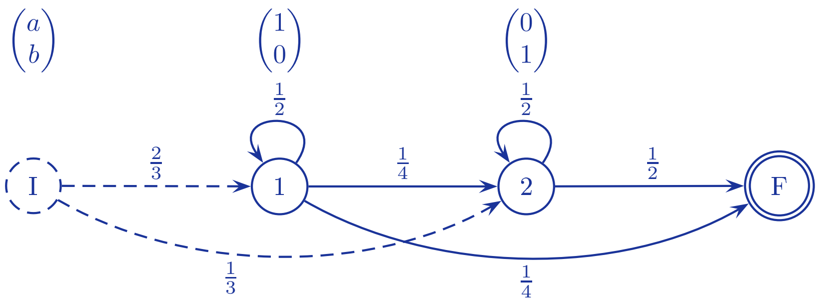
*c*) Calcula la probabilidad de que $M$ genere una cadena cualquiera de longitud 7.
El conjunto de cadenas de longitud 7 que $M$ puede generar es:

$$S = \{a^7, a^6b, a^5b^2, a^4b^3, a^3b^4, a^2b^5, ab^6, b^7\}$$

Cada cadena de $S$ se genera con probabilidad $\frac{1}{3}\frac{1}{2^7}$.
Por tanto,

$$P(x \in S \mid M) = 8\ \frac{1}{3}\ \frac{1}{2^7} = 0.0208$$

*d*) Diseña un modelo de Markov $M'$ de conjunto de estados $Q' = \{1, 2, F\}$ y alfabeto $\Sigma' = \Sigma$ que genere el mismo lenguaje probabilístico que $M$.

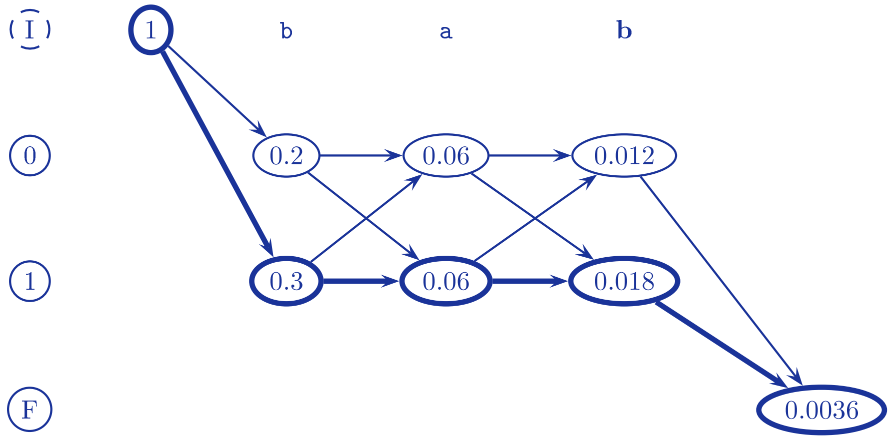


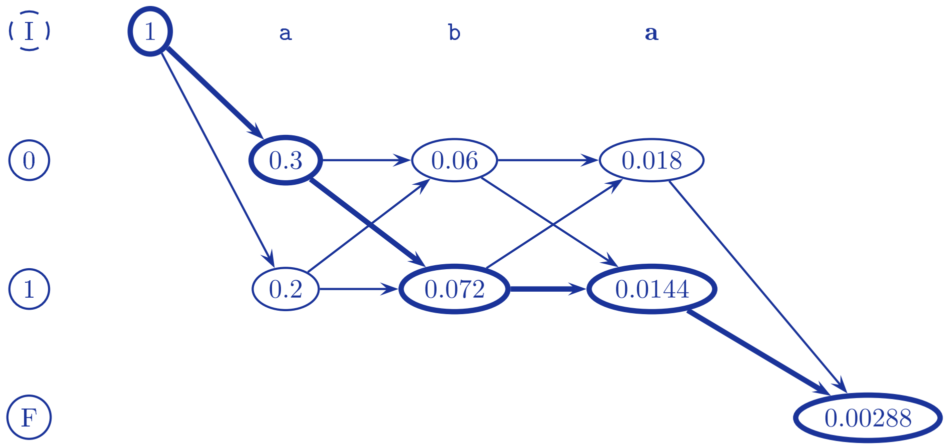
5. Sea $M$ un modelo de Markov de conjunto de estados $Q = \{0, 1, F\}$; alfabeto $\Sigma = \{a, b\}$; probabilidades iniciales $\pi_0(0) = 0.5, \pi_0(1) = 0.5, \pi_0(F) = 0.0$; y probabilidades de transición entre estados y de emisión de símbolos:

| $A$ | 0 | 1 | $F$ |
|---|---|---|---|
| 0 | 0.5 | 0.4 | 0.1 |
| 1 | 0.3 | 0.5 | 0.2 |

| $B$ | $a$ | $b$ |
|---|---|---|
| 0 | 0.6 | 0.4 |
| 1 | 0.4 | 0.6 |

Reestima los parámetros de $M$ mediante *una* iteración de reestimación por Viterbi, a partir de las cadenas de entrenamiento "`bab`" y "`aba`".




Los pares *cadena-secuencia óptima de estados* a considerar son:

| bab | aba |
|---|---|
| 111F | 011F |

Los parámetros reestimados son:

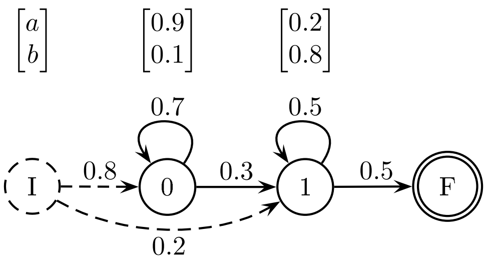
$\hat{\pi}_0(0) = 0.5$
$\hat{\pi}_0(1) = 0.5$

| $A$ | 0 | 1 | $F$ |
|---|---|---|---|
| 0 | 0 | 1 | 0 |
| 1 | 0 | $\frac{3}{5}$ | $\frac{2}{5}$ |

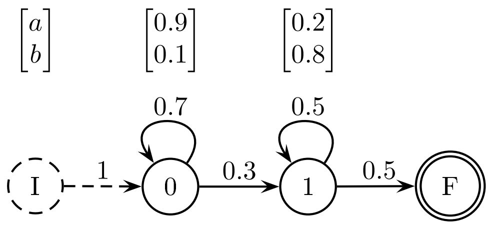
| $B$ | $a$ | $b$ |
|---|---|---|
| 0 | 1 | 0 |
| 1 | $\frac{2}{5}$ | $\frac{3}{5}$ |

6. Se tiene un problema de clasificación en dos clases ($A$ y $B$) de objetos representados mediante cadenas de símbolos en el alfabeto $\Sigma = \{a, b\}$. Las probabilidades a priori de las clases son $P(A) = 0.6$ y $P(B) = 0.4$. Las funciones de probabilidad condicional de las clases vienen caracterizadas por los modelos de Markov:

Modelo $M_A$: $P(x \mid A) = P(x \mid M_A)$




Modelo $M_B$: $P(x \mid B) = P(x \mid M_B)$




*a*) Calcula las probabilidades de que $M_A$ y $M_B$ generen la cadena "`aab`".

$$\begin{aligned}
P(aab \mid M_A) &= P(aab, q_1 = 0, q_2 = 0, q_3 = 1 \mid M_A) \\
&+ P(aab, q_1 = 0, q_2 = 1, q_3 = 1 \mid M_A) \\
&+ P(aab, q_1 = 1, q_2 = 1, q_3 = 1 \mid M_A) \\
&= (0.8 \cdot 0.9)\ (0.7 \cdot 0.9)\ (0.3 \cdot 0.8)\ 0.5 \\
&+ (0.8 \cdot 0.9)\ (0.3 \cdot 0.2)\ (0.5 \cdot 0.8)\ 0.5 \\
&+ (0.2 \cdot 0.2)\ (0.5 \cdot 0.2)\ (0.5 \cdot 0.8)\ 0.5 \\
&= 0.0544 + 0.0086 + 0.0008 \\
&= 0.0638
\end{aligned}$$

$$\begin{aligned}
P(aab \mid M_B) &= P(aab, q_1 = 0, q_2 = 0, q_3 = 1 \mid M_B) \\
&+ P(aab, q_1 = 0, q_2 = 1, q_3 = 1 \mid M_B) \\
&= (1 \cdot 0.9)\ (0.7 \cdot 0.9)\ (0.3 \cdot 0.8)\ 0.5 \\
&+ (1 \cdot 0.9)\ (0.3 \cdot 0.2)\ (0.5 \cdot 0.8)\ 0.5 \\
&= 0.0680 + 0.0108 \\
&= 0.0788
\end{aligned}$$

*b*) Determina las probabilidades a posteriori de que la cadena "`aab`" pertenezca a las clases $A$ y $B$.

$$P(aab \mid A)\, P(A) = 0.0638\ 0.6 = 0.0383$$
$$P(aab \mid B)\, P(B) = 0.0788\ 0.4 = 0.0315$$

$$P(x) = P(aab \mid M_A)\, P(A) + P(aab \mid M_B)\, P(B) = 0.0383 + 0.0315 = 0.0698$$

$$P(A \mid aab) = \frac{P(aab \mid A)\, P(A)}{P(x)} = \frac{0.0383}{0.0698} = 0.5487$$

$$P(B \mid aab) = \frac{P(aab \mid B)\, P(B)}{P(x)} = \frac{0.0315}{0.0698} = 0.4513$$

*c*) Clasifica la cadena "`aab`" por mínima probabilidad de error.

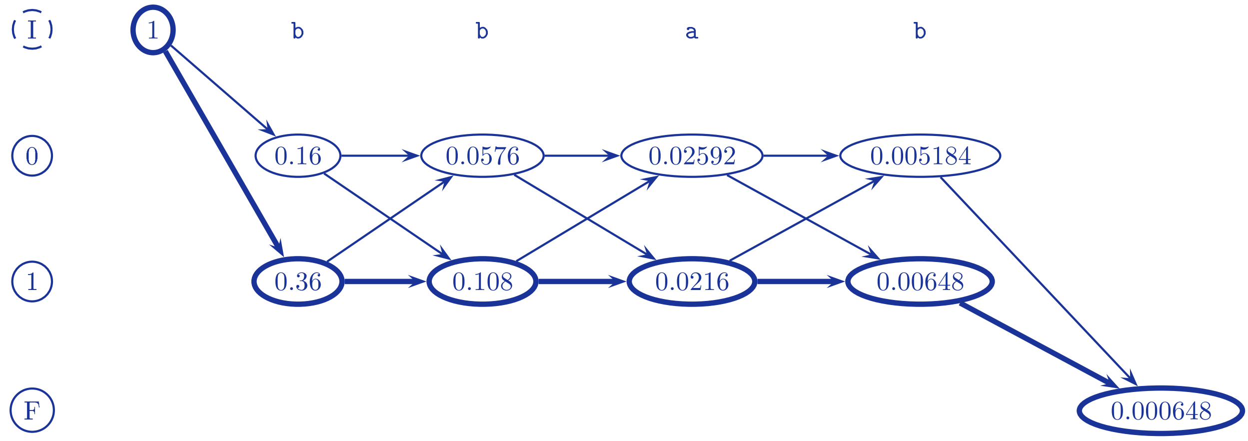
$$\omega^*(aab) = \operatorname*{arg\,max}_{\omega = A, B} P(\omega \mid aab) = A$$

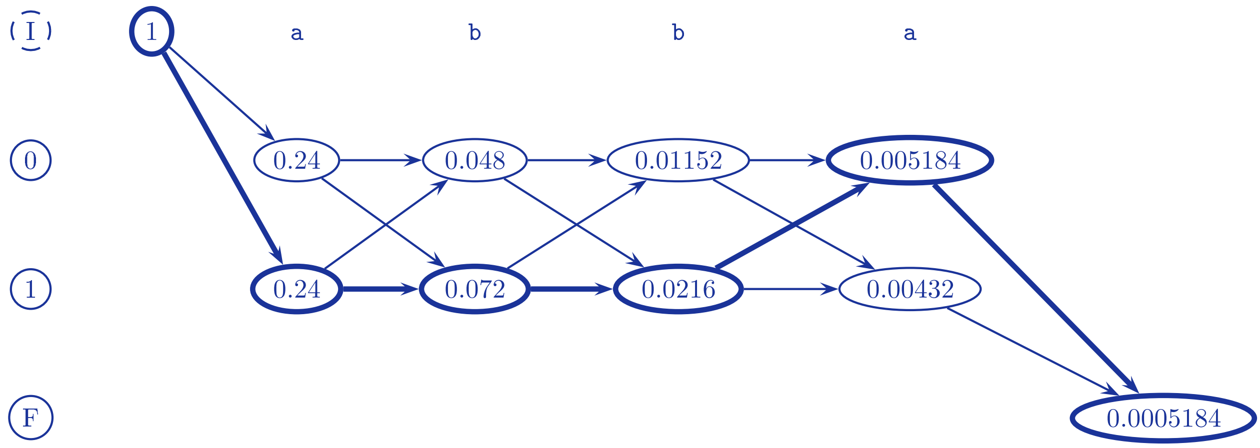
7. Sea $M$ un modelo de Markov de conjunto de estados $Q = \{0, 1, F\}$; alfabeto $\Sigma = \{a, b\}$; probabilidades iniciales $\pi_0(0) = 0.4, \pi_0(1) = 0.6, \pi_0(F) = 0.0$; y probabilidades de transición entre estados y de emisión de símbolos:

| $A$ | 0 | 1 | $F$ |
|---|---|---|---|
| 0 | 0.5 | 0.4 | 0.1 |
| 1 | 0.4 | 0.5 | 0.1 |

| $B$ | $a$ | $b$ |
|---|---|---|
| 0 | 0.6 | 0.4 |
| 1 | 0.4 | 0.6 |

Reestima los parámetros de $M$ mediante *una* iteración de reestimación por Viterbi, a partir de las cadenas de entrenamiento "`bbab`" y "`abba`".




Los pares *cadena-secuencia óptima de estados* a considerar son:

| bbab | abba |
|---|---|
| 1111F | 1110F |

Los parámetros reestimados son:

$$\hat{\pi}_0(0) = 0$$
$$\hat{\pi}_0(1) = \frac{2}{2}$$

| $A$ | 0 | 1 | $F$ |
|---|---|---|---|
| 0 | 0 | 0 | 1 |
| 1 | $\frac{1}{7}$ | $\frac{5}{7}$ | $\frac{1}{7}$ |

| $B$ | $a$ | $b$ |
|---|---|---|
| 0 | 1 | 0 |
| 1 | $\frac{2}{7}$ | $\frac{5}{7}$ |

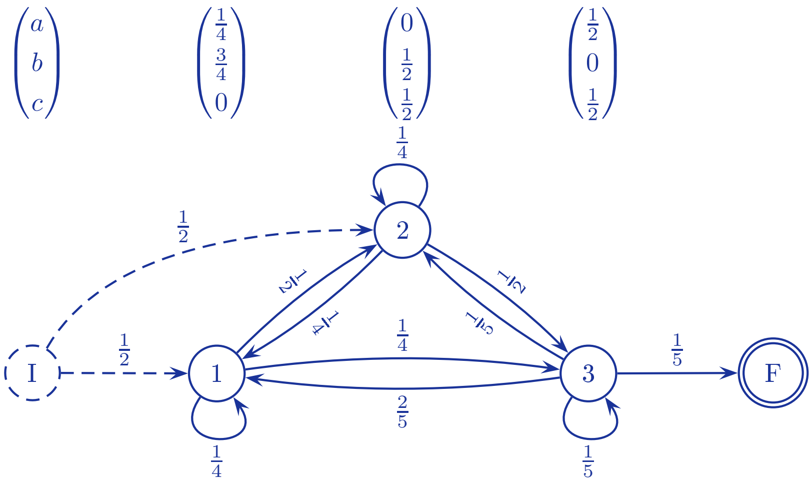
8. Sea $M$ un modelo de Markov de conjunto de estados $Q = \{1, 2, 3, F\}$; alfabeto $\Sigma = \{a, b, c\}$; probabilidades iniciales $\pi_0(1) = \pi_0(2) = \frac{1}{2}$ y $\pi_0(3) = \pi_0(F) = 0$; y probabilidades de transición entre estados y de emisión de símbolos:

| $A$ | 1 | 2 | 3 | $F$ |
|---|---|---|---|---|
| 1 | $\frac{1}{4}$ | $\frac{1}{2}$ | $\frac{1}{4}$ | 0 |
| 2 | $\frac{1}{4}$ | $\frac{1}{4}$ | $\frac{1}{2}$ | 0 |
| 3 | $\frac{2}{5}$ | $\frac{1}{5}$ | $\frac{1}{5}$ | $\frac{1}{5}$ |

| $B$ | $a$ | $b$ | $c$ |
|---|---|---|---|
| 1 | $\frac{1}{4}$ | $\frac{3}{4}$ | 0 |
| 2 | 0 | $\frac{1}{2}$ | $\frac{1}{2}$ |
| 3 | $\frac{1}{2}$ | 0 | $\frac{1}{2}$ |

Se pide:

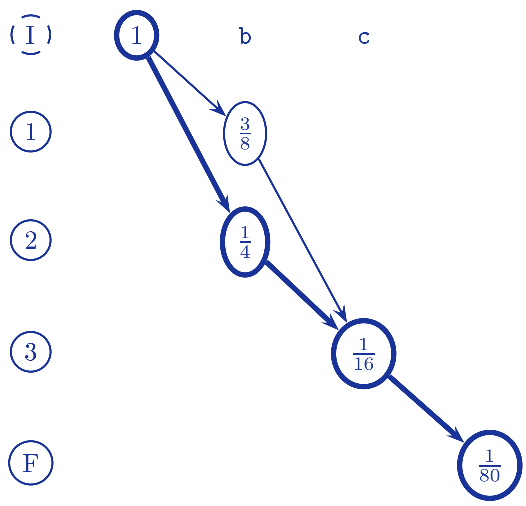
*a*) Representa gráficamente el modelo $M$.

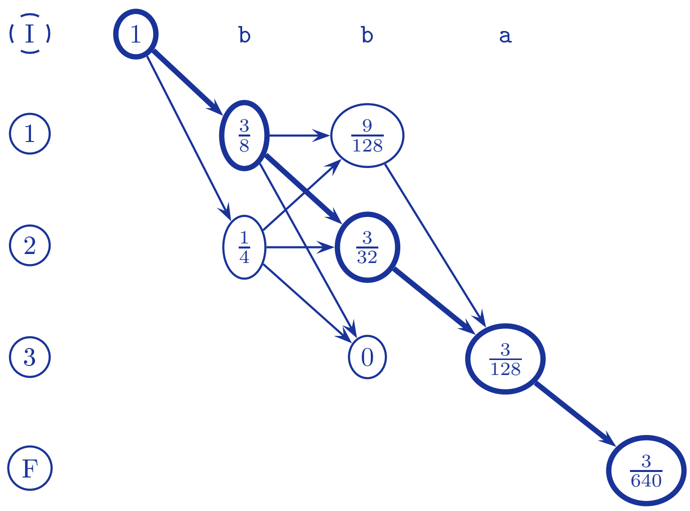


*b*) Calcula la probabilidad de que $M$ genere la cadena "`bc`".

$$p(bc \,|\, M) = p(bc, q_1 = 1, q_2 = 3 \,|\, M) + p(bc, q_1 = 2, q_2 = 3 \,|\, M) = \frac{1}{2}\frac{3}{4}\,\frac{1}{4}\frac{1}{2}\,\frac{1}{5} + \frac{1}{2}\frac{1}{2}\,\frac{1}{2}\frac{1}{2}\,\frac{1}{5} = \frac{3}{640} + \frac{1}{80} = \frac{11}{640}$$

*c*) Realiza una iteración del proceso de reestimación por Viterbi. con el conjunto de entrenamiento $R = \{$"bc","bba"$\}$.





Los pares *cadena-secuencia óptima de estados* a considerar son:

| bc | bba |
|---|---|
| 23F | 123F |

Los parámetros reestimados son:

$$\pi_0(1) = \frac{1}{2}$$
$$\pi_0(2) = \frac{1}{2}$$
$$\pi_0(3) = 0$$

| $A$ | 1 | 2 | 3 | $F$ |
|---|---|---|---|---|
| 1 | 0 | 1 | 0 | 0 |
| 2 | 0 | 0 | 1 | 0 |
| 3 | 0 | 0 | 0 | 1 |

| $B$ | $a$ | $b$ | $c$ |
|---|---|---|---|
| 1 | 0 | 1 | 0 |
| 2 | 0 | 1 | 0 |
| 3 | $\frac{1}{2}$ | 0 | $\frac{1}{2}$ |

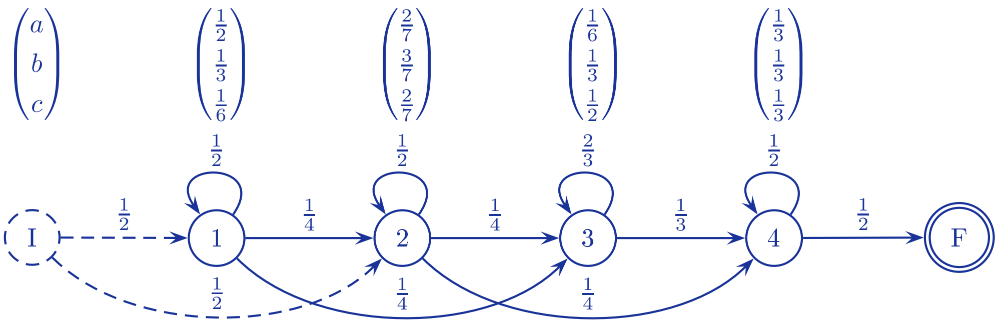
9. Sea $M$ un modelo de Markov de conjunto de estados $Q = \{1, 2, 3, 4, F\}$; alfabeto $\Sigma = \{a, b, c\}$; probabilidades iniciales $\pi_0(1) = \pi_0(2) = \frac{1}{2}$ y $\pi_0(3) = \pi_0(4) = \pi_0(F) = 0$; y probabilidades de transición entre estados y de emisión de símbolos:

| $A$ | 1 | 2 | 3 | 4 | $F$ |
|---|---|---|---|---|---|
| 1 | $\frac{1}{2}$ | $\frac{1}{4}$ | $\frac{1}{4}$ | | |
| 2 | | $\frac{1}{2}$ | $\frac{1}{4}$ | $\frac{1}{4}$ | |
| 3 | | | $\frac{2}{3}$ | $\frac{1}{3}$ | |
| 4 | | | | $\frac{1}{2}$ | $\frac{1}{2}$ |

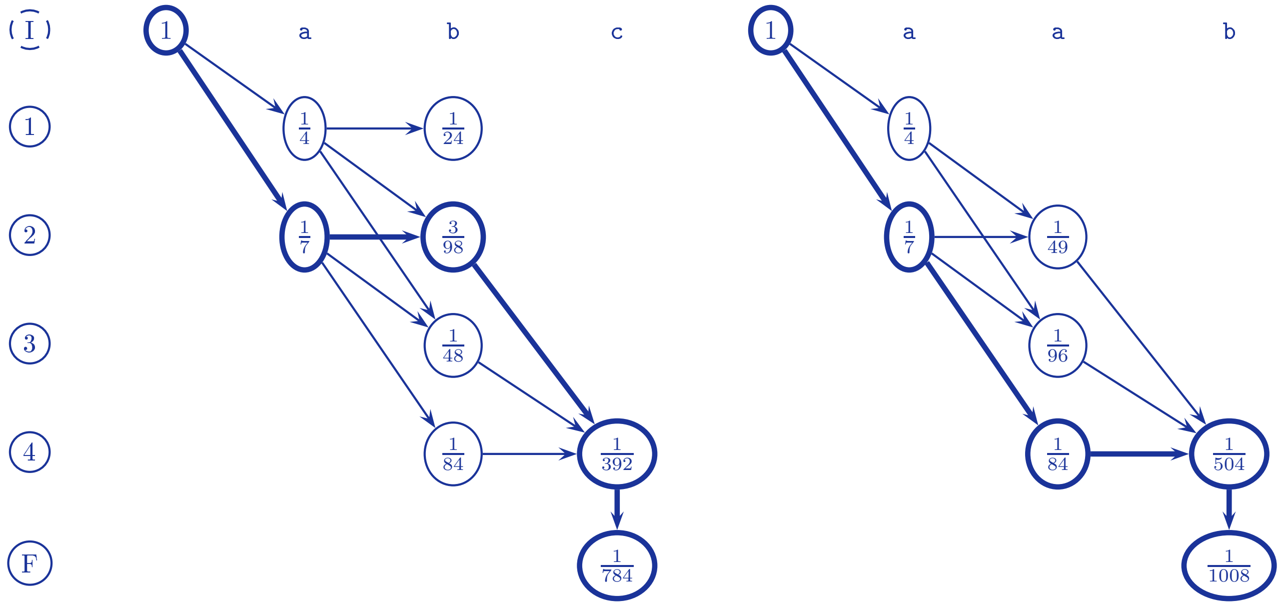
| $B$ | $a$ | $b$ | $c$ |
|---|---|---|---|
| 1 | $\frac{1}{2}$ | $\frac{1}{3}$ | $\frac{1}{6}$ |
| 2 | $\frac{2}{7}$ | $\frac{3}{7}$ | $\frac{2}{7}$ |
| 3 | $\frac{1}{6}$ | $\frac{1}{3}$ | $\frac{1}{2}$ |
| 4 | $\frac{1}{3}$ | $\frac{1}{3}$ | $\frac{1}{3}$ |

Se pide:

*a*) Representa gráficamente el modelo $M$.



*b*) Realiza una iteración del proceso de reestimación por Viterbi con el conjunto de entrenamiento $R = \{\text{"abc","aab"}\}$.




Los pares *cadena-secuencia óptima de estados* a considerar son:

| abc | aab |
|---|---|
| 224F | 244F |

Los parámetros reestimados son:

$\pi_0(2) = 1$
$\pi_0(1) = \pi_0(3) = \pi_0(4) = 0$

| $A$ | 1 | 2 | 3 | 4 | $F$ |
|---|---|---|---|---|---|
| 1 | 0 | 0 | 0 | 0 | 0 |
| 2 | 0 | $\frac{1}{3}$ | 1 | $\frac{2}{3}$ | 0 |
| 3 | 0 | 0 | 0 | 0 | 0 |
| 4 | 0 | 0 | 0 | $\frac{1}{3}$ | $\frac{2}{3}$ |

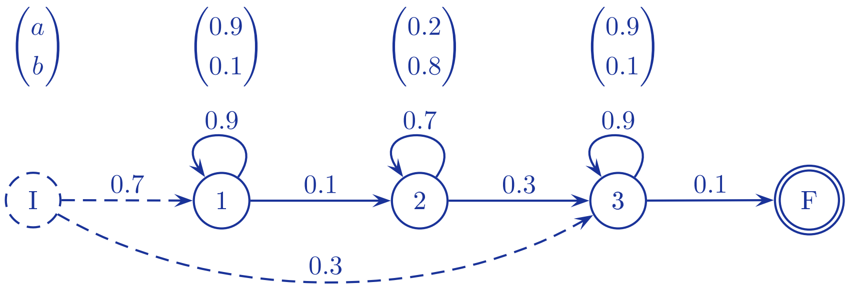
| $B$ | $a$ | $b$ | $c$ |
|---|---|---|---|
| 1 | 0 | 0 | 0 |
| 2 | $\frac{2}{3}$ | $\frac{1}{3}$ | 0 |
| 3 | 0 | 0 | 0 |
| 4 | $\frac{1}{3}$ | $\frac{1}{3}$ | $\frac{1}{3}$ |

10. Sea $\mathcal{M}$ un modelo de Markov de cuatro estados sobre $\Sigma = \{a, b\}$, definido por $\pi_1 = 0.7, \pi_2 = \pi_F = 0, \pi_3 = 0.3$ y las matrices A y B:

| A | 1 | 2 | 3 | F |
|---|---|---|---|---|
| 1 | 0.9 | 0.1 | 0.0 | 0.0 |
| 2 | 0.0 | 0.7 | 0.3 | 0.0 |
| 3 | 0.0 | 0.0 | 0.9 | 0.1 |

| B | a | b |
|---|---|---|
| 1 | 0.9 | 0.1 |
| 2 | 0.2 | 0.8 |
| 3 | 0.9 | 0.1 |

*a*) Representa gráficamente el modelo.




*b*) Determina cuántas cadenas diferentes, de longitud 3, puede generar $\mathcal{M}$ con probabilidad mayor de cero.

Las posibles secuencias de estados son: $z_1$ =“1,2,3,F” y $z_2$ =“3,3,3,F”. Como en cada estado se pueden emitir dos símbolos, por cada secuencia de estados se pueden generar $2^3 = 8$ cadenas distintas; sin embargo, estas 8 cadenas serán las mismas en las dos secuencias de estados. Por tanto, el número de cadenas diferentes es 8.

*c*) Determina las probabilidades de que $\mathcal{M}$ produzca las secuencias de estados $z_1$=“1,2,3,F” y $z_2$=“3,3,3,F”.

$$Pr(``1, 2, 3, F'') = \pi_1 A_{1,2} A_{2,3} A_{3,F} = 0.7 \cdot 0.1 \cdot 0.3 \cdot 0.1 = 0.0021$$
$$Pr(``3, 3, 3, F'') = \pi_3 A_{3,3} A_{3,3} A_{3,F} = 0.3 \cdot 0.9 \cdot 0.9 \cdot 0.1 = 0.0243$$

*d*) Determina la probabilidad de que $\mathcal{M}$ genere una cadena de longitud 3.

$$\begin{aligned} Pr(x|\ long(x) = 3) &= \sum_{i=1}^{2} Pr(z_i) \\ &= \pi_1 A_{1,2} A_{2,3} A_{3,F} + \pi_3 A_{3,3} A_{3,3} A_{3,F} \\ &= 0.7 \cdot 0.1 \cdot 0.3 \cdot 0.1 + 0.3 \cdot 0.9 \cdot 0.9 \cdot 0.1 \\ &= 0.0021 + 0.0243 = 0.0264 \end{aligned}$$

*e*) Realiza una traza del algoritmo de Viterbi para obtener la secuencia de estados mas probable con la que $\mathcal{M}$ genera la cadena `"b a a b"`.

Valores en los nodos del *trelis*:

| | b | a | a | b | |
|---|---|---|---|---|---|
| 1 | 0.07 | 0.0567 | 0.0459 | 0.00414 | |
| 2 | | 0.0014 | 0.0011 | 0.00368 | |
| 3 | 0.03 | 0.0243 | 0.0197 | 0.00177 | |
| F | | | | | 0.000177 |

Secuencia óptima de estados:

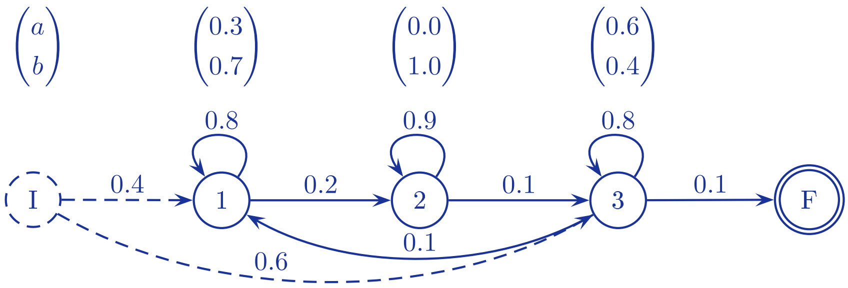
$\tilde{q} =$ 3, 3, 3, 3, F

11. Sea $\mathcal{M}$ un modelo de Markov de cuatro estados sobre $\Sigma = \{a, b\}$, definido por $\pi_1 = 0.4, \pi_2 = \pi_F = 0, \pi_3 = 0.6$ y las matrices A y B:

| A | 1 | 2 | 3 | F |
|---|---|---|---|---|
| 1 | 0.8 | 0.2 | 0.0 | 0.0 |
| 2 | 0.0 | 0.9 | 0.1 | 0.0 |
| 3 | 0.1 | 0.0 | 0.8 | 0.1 |

| B | a | b |
|---|---|---|
| 1 | 0.3 | 0.7 |
| 2 | 0 | 1 |
| 3 | 0.6 | 0.4 |

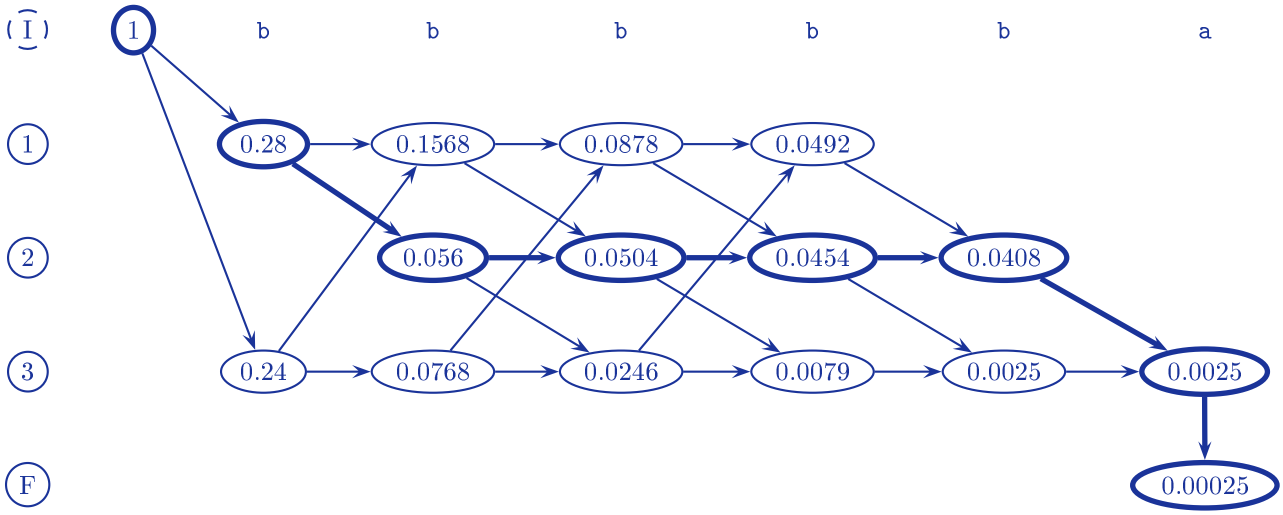
*a*) Representa gráficamente el modelo.




*b*) Determina cuál es la probabilidad de que el modelo genere la cadena "`b b b b b a`" siguiendo la secuencia de estados $112233F$.

$$\begin{aligned} p(bbbbba|112233F) &= \pi_1\, B_{1b} \cdot A_{11}\, B_{1b} \cdot A_{12}\, B_{2b} \cdot A_{22}\, B_{2b} \cdot A_{23}\, B_{3b} \cdot A_{33}\, B_{3a} \cdot A_{3F} \\ &= 0.4\, 0.7 \cdot 0.8\, 0.7 \cdot 0.2\, 1 \cdot 0.9\, 1 \cdot 0.1\, 0.4 \cdot 0.8\, 0.6 \cdot 0.1 \\ &= 0.00005419 \end{aligned}$$

*c*) Realiza una traza del algoritmo de Viterbi para obtener la secuencia de estados mas probable con la que $\mathcal{M}$ genera la cadena "`b b b b b a`".




$$\tilde{Q} = (1, 2, 2, 2, 2, 3, F)$$

*d*) Determina la probabilidad de que una cadena comience por $a$.

$$p(x_1 = \mathtt{a}) = \pi_1\, B_{1a} + \pi_3\, B_{3a} = 0.4\, 0.3 + 0.6\, 0.6 = 0.48$$

*e*) Determina la probabilidad de que una cadena termine por $b$.

$$p(x_{|x|} = \mathtt{b}) = B_{3b} = 0.4$$

12. En un problema de clasificación en dos clases $u$ y $v$, cada objeto a clasificar está representado mediante una cadena, $y$, sobre un alfabeto de dos primitivas; es decir $y \in \Sigma = \{a, b\}$. Las probabilidades a priori de las clases son $P(u) = 0.7$ y $P(v) = 0.3$. Las probabilidades condicionales, $P(y|u)$ y $P(y|v)$, vienen dadas por modelos de Markov de tres estados, $\mathcal{M}_u$, $\mathcal{M}_v$:

$$\mathcal{M}_u = (Q_u, \Sigma, \pi_u, A_u, B_u), \quad \mathcal{M}_v = (Q_v, \Sigma, \pi_v, A_v, B_v), \quad \text{donde:}$$

$$Q_u = Q_v = \{1, 2, F\}, \qquad \pi_u[1] = 0.8,\ \pi_u[2] = 0.2, \qquad \pi_v[1] = 0.2,\ \pi_v[2] = 0.8$$

$$A_u = A_v = \begin{bmatrix} 0.3 & 0.2 & 0.5 \\ 0 & 0.5 & 0.5 \end{bmatrix}, \quad B_u = \begin{bmatrix} 0.5 & 0.5 \\ 0.4 & 0.6 \end{bmatrix}, \quad B_v = \begin{bmatrix} 0.5 & 0.5 \\ 0.6 & 0.4 \end{bmatrix}$$

Dada la cadena $y = abb$,

*a*) Calcula la aproximación de Viterbi a las probabilidades de generar $y$ mediante $\mathcal{M}_u$ y $\mathcal{M}_v$. En ambos casos, realiza una traza del algoritmo de Viterbi y determina la secuencia de estados más probable. En los apartados siguientes supondremos que las aproximaciones calculadas son suficientemente precisas.

$P(y|u) \approx \tilde{P}(y|\mathcal{M}_u) = 0.0072,$ $\qquad$ $P(y|v) \approx \tilde{P}(y|\mathcal{M}_v) = 0.0096$

| $\mathcal{M}_u$ | a | b | b | |
|---|---|---|---|---|
| 1 | 0.40 | 0.060 | 0.0090 | |
| 2 | 0.08 | 0.048 | 0.0144 | |
| F | | | | 0.0072 |
| $\hat{z}$ | 1 | 2 | 2 | F |

| $\mathcal{M}_v$ | a | b | b | |
|---|---|---|---|---|
| 0 | 0.10 | 0.015 | 0.00225 | |
| 1 | 0.48 | 0.096 | 0.01920 | |
| F | | | | 0.0096 |
| $\hat{z}$ | 2 | 2 | 2 | F |

*b*) Halla la probabilidad incondicional de $y$.

$P(y) = P(u)P(y|u) + P(v)P(y|v) \ = \ 0.00504 + 0.00288 \ = \ 0.00792$

*c*) Calcula las probabilidades a posteriori de las clases $u$ y $v$ tras observar $y$.

$P(u|y) \ = \ P(u)P(y|u)/P(y) \ = \ 0.636$ $\qquad$ $P(v|y) \ = \ P(v)P(y|v)/P(y) \ = \ 0.364$

*d*) Clasifica $y$ mediante la regla de Bayes.

$\hat{w} \ = \ argmax_{w \in \{u,v\}} P(w|y) \ = \ u$

*e*) Determina la probabilidad de que la clasificación del apartado anterior sea errónea.

$P(error|y) \ = \ 1 \ - \ max_{w \in \{u,v\}} P(w|y) \ = \ 0.364$

13. Sea $\mathcal{M}$ un modelo de Markov de cuatro estados sobre $\Sigma = \{a, b\}$, definido por $\pi_1 = 0.4, \pi_2 = \pi_F = 0, \pi_3 = 0.6$ y las matrices A y B:

| A | 1 | 2 | 3 | F |
|---|---|---|---|---|
| 1 | 0.8 | 0.2 | 0.0 | 0.0 |
| 2 | 0.0 | 0.4 | 0.6 | 0.0 |
| 3 | 0.4 | 0.0 | 0.5 | 0.1 |

| B | a | b |
|---|---|---|
| 1 | 0.3 | 0.7 |
| 2 | 0 | 1 |
| 3 | 0.6 | 0.4 |

*a*) Determina la probabilidad de que el modelo genere la cadena $y =$ "`a b b b`".

Hay tres secuencias de estados que pueden generar $y$ con probabilidad no nula:

$z_1 =< 3\ 1\ 2\ 3 >,\ z_2 =<\ 1\ 2\ 3\ 3 >,\ z_3 =< 1\ 1\ 2\ 3 >.$

$P(y, z_1) = 0.000484,\ P(y, z_2) = 0.0001152,\ P(y, z_3) = 0.00013824,$

$P(y \mid \mathcal{M}) = 0.000484 + 0.0001152 + 0.00013824 = 0.00073744.$

*b*) Realiza una traza del algoritmo de Viterbi para determinar la secuencia de estados que con mayor probabilidad genera la cadena $y$. Obtén también la probabilidad correspondiente (aproximación de Viterbi) y el error con respecto a la probabilidad obtenida en el punto anterior.

```
y  =      a          b          b          b

1  0.120000  0.100800  0.056448  0.031611
2            0.024000  0.020160  0.011290
3  0.360000  0.072000  0.014400  0.004838
F                                          0.000484

z  =      3          1          2          3          F
```

$\tilde{P}(y \mid \mathcal{M}) = P(y\,, < 3\,1\,2\,3 >) \approx 0.000484$

Error $= 0.00073744 - 0.000484 = 0.00025344$ (34 %)

*c*) Se dispone del siguiente conjunto de cadenas de entrenamiento, $Y$, cada una de las cuales acompañada de la secuencia de estados con la que ha sido generada por $\mathcal{M}$ con mayor probabilidad (secuencia óptima de Viterbi):

```
cadenas:    a b b b a    a b b    b b b a    b b a
estados:    3 1 1 2 3    3 3 3    1 1 2 3    1 2 3
```

Usando estas cadenas (y secuencias de estados) y $\mathcal{M}$ como modelo inicial, realiza una iteración del algoritmo de reestimación de Viterbi para obtener el modelo reestimado $\mathcal{M}'$

El modelo $\mathcal{M}'$ tiene la misma topología que $\mathcal{M}$ y los siguientes parámetros:

$\pi_1 = 2/4,\ \pi_2 = \pi_F = 0,\ \pi_3 = 2/4$

| A | 1 | 2 | 3 | F |
|---|---|---|---|---|
| 1 | 2/5 | 3/5 | 0 | 0 |
| 2 | 0 | 0 | 1 | 0 |
| 3 | 1/7 | 0 | 2/7 | 4/7 |

| B | a | b |
|---|---|---|
| 1 | 1/6 | 5/6 |
| 2 | 1/4 | 3/4 |
| 3 | 5/7 | 2/7 |

*d*) Calcula la (aproximación de Viterbi a la) verosimilitud del conjunto de cadenas de entrenamiento con $\mathcal{M}$.

$$\tilde{P}(Y \mid \mathcal{M}) = \tilde{P}(abbba \mid \mathcal{M}) \cdot \tilde{P}(abb \mid \mathcal{M}) \cdot \tilde{P}(bbba \mid \mathcal{M}) \cdot \tilde{P}(bba \mid \mathcal{M}) =$$

$$0.000407 \cdot 0.001440 \cdot 0.001129 \cdot 0.002017 \approx 0.00000000000133$$

14. Sea $\mathcal{M}$ un modelo de Markov de cuatro estados sobre $\Sigma = \{a, b\}$, definido por $\pi_1 = 0.4, \pi_2 = \pi_F = 0, \pi_3 = 0.6$ y las matrices A y B:

| A | 1 | 2 | 3 | F |
|---|---|---|---|---|
| 1 | 0.8 | 0.2 | 0.0 | 0.0 |
| 2 | 0.0 | 0.6 | 0.3 | 0.1 |
| 3 | 0.4 | 0.0 | 0.5 | 0.1 |

| B | a | b |
|---|---|---|
| 1 | 0.3 | 0.7 |
| 2 | 0 | 1 |
| 3 | 0.6 | 0.4 |

*a*) Determina la probabilidad de que el modelo genere una cadena de longitud 2.

Hay solo dos secuencias de dos estados con probabilidad mayor que cero:

$z_1 =< 1\ 2 >,\ z_2 =< 3\ 3 >;\quad P(z_1) = 0.4 \cdot 0.2 \cdot 0.1 = 0.008,\ P(z_2) = 0.6 \cdot 0.5 \cdot 0.1 = 0.03.$

Hay cuatro posibles cadenas de $\{a, b\}$ de longitud 2: $aa, ab, ba, bb$. Por tanto, la probabilidad de que $\mathcal{M}$ genere una cadena de longitud 2 es:

$$P(aa, z_1) + P(aa, z_2) + P(ab, z_1) + P(ab, z_2) + P(ba, z_1) + P(ba, z_2) + P(bb, z_1) + P(bb, z_2)$$
$$= 0.008 \cdot 0.3 \cdot 0 \cdot + 0.03 \cdot 0.6 \cdot 0.6 + 0.008 \cdot 0.3 \cdot 1 \cdot + 0.03 \cdot 0.6 \cdot 0.4 + 0.008 \cdot 0.7 \cdot 0 \cdot + 0.03 \cdot 0.4 \cdot 0.6 + 0.008 \cdot 0.7 \cdot 1 \cdot + 0.03 \cdot 0.4 \cdot 0.4$$
$$= P(z_1) + P(z_2) = 0.038$$

*b*) Realiza una traza del algoritmo de Viterbi para determinar la secuencia de estados que con mayor probabilidad genera la cadena $abbb$. Obtén también la probabilidad correspondiente (aproximación de Viterbi).

```
y  =      a          b          b          b

1   0.120000   0.100800   0.056448   0.031611
2              0.024000   0.020160   0.012096
3   0.360000   0.072000   0.014400   0.002880
F                                               0.0012096

z  =      3          1          2          2          F
```

$\tilde{P}(y \mid \mathcal{M}) = P(y\,, < 3\,1\,2\,2 >) \approx 0.00121$

*c*) Se dispone del siguiente conjunto de cadenas de entrenamiento, $Y$, cada una de las cuales acompañada de la secuencia de estados con la que ha sido generada por $\mathcal{M}$ con mayor probabilidad (secuencia óptima de Viterbi):

```
cadenas:     a b b b a     a b b     b b b a     b b a
estados:     3 1 2 2 3     3 1 2     1 2 2 3     3 3 3
```

Usando estas cadenas (y secuencias de estados) y $\mathcal{M}$ como modelo inicial, realiza una iteración del algoritmo de reestimación de Viterbi para obtener el modelo reestimado $\mathcal{M}'$

El modelo $\mathcal{M}'$ tiene la misma topología que $\mathcal{M}$ y los siguientes parámetros:

$\pi_1 = 1/4,\ \pi_2 = \pi_F = 0,\ \pi_3 = 3/4$

| A | 1 | 2 | 3 | F |
|---|---|---|---|---|
| 1 | 0 | 1 | 0 | 0 |
| 2 | 0 | 2/5 | 2/5 | 1/5 |
| 3 | 2/7 | 0 | 2/7 | 3/7 |

| B | a | b |
|---|---|---|
| 1 | 1/4 | 3/4 |
| 2 | 1/6 | 5/6 |
| 3 | 5/7 | 2/7 |

*d*) Calcula la (aproximación de Viterbi a la) verosimilitud del conjunto de cadenas de entrenamiento con $\mathcal{M}$.

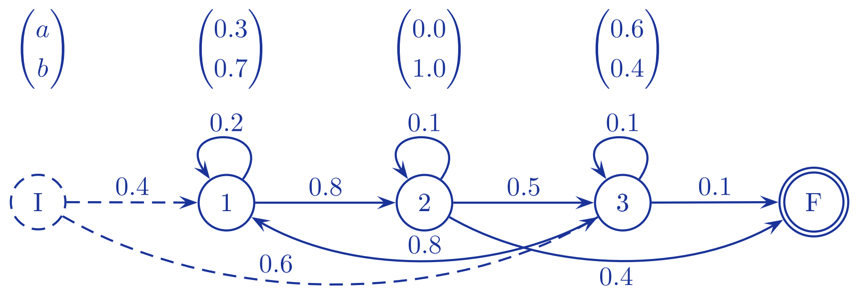
$$\tilde{P}(Y \mid \mathcal{M}) = \tilde{P}(abbba \mid \mathcal{M}) \cdot \tilde{P}(abb \mid \mathcal{M}) \cdot \tilde{P}(bbba \mid \mathcal{M}) \cdot \tilde{P}(bba \mid \mathcal{M}) =$$
$$0.000218 \cdot 0.00144 \cdot 0.000605 \cdot 0.00144 \approx .0000000000002734871$$

15. Sea $\mathcal{M}$ un modelo de Markov de cuatro estados sobre $\Sigma = \{a, b\}$, definido por $\pi_1 = 0.4, \pi_2 = \pi_F = 0, \pi_3 = 0.6$ y las matrices A y B:

| A | 1 | 2 | 3 | F |
|---|---|---|---|---|
| 1 | 0.2 | 0.8 | 0.0 | 0.0 |
| 2 | 0.0 | 0.1 | 0.5 | 0.4 |
| 3 | 0.8 | 0.0 | 0.1 | 0.1 |

| B | a | b |
|---|---|---|
| 1 | 0.3 | 0.7 |
| 2 | 0.0 | 1.0 |
| 3 | 0.6 | 0.4 |

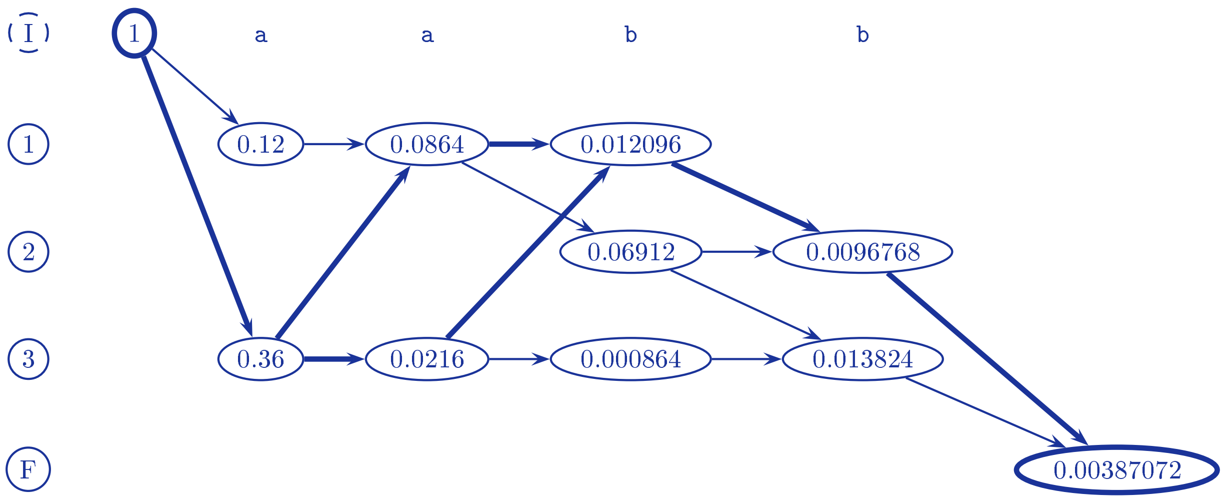
*a*) Representa gráficamente el modelo.




*b*) Determina cuál es la probabilidad de que el modelo genere la cadena “abbbba” siguiendo la secuencia de estados 112233F.

$$\begin{aligned} p(abbbba, 112233F) &= \pi_1\, B_{1a} \cdot A_{11}\, B_{1b} \cdot A_{12}\, B_{2b} \cdot A_{22}\, B_{2b} \cdot A_{23}\, B_{3b} \cdot A_{33}\, B_{3a} \cdot A_{3F} \\ &= 0.4\, 0.3 \cdot 0.2\, 0.7 \cdot 0.8\, 1.0 \cdot 0.1\, 1.0 \cdot 0.5\, 0.4 \cdot 0.1\, 0.6 \cdot 0.1 \\ &= 0.0000016128 \end{aligned}$$

*c*) Realiza una traza del algoritmo de Viterbi para obtener la secuencia de estados mas probable con la que $\mathcal{M}$ genera la cadena "aabb".




Hay dos caminos óptimos (cualquiera de ellos es solución al problema):

$$\tilde{z} = (3, 1, 1, 2, F) \qquad \text{y} \qquad \tilde{z}' = (3, 3, 1, 2, F)$$
$$p(aabb, \tilde{z}) = p(aabb, \tilde{z}') = 0.00387072$$

*d*) Realiza una iteración de reestimación por Viterbi de $\mathcal{M}$ a partir de las cadenas de aprendizaje

`ab, ba, aab, aba,`

sabiendo que las secuencias de estados óptimas para estas cadenas son, respectivamente,

`12F, 33F, 312F, 123F.`

Contadores:

| $\pi'$ | 1 | 2 | 3 |
|---|---|---|---|
| | 2 | 0 | 2 |

| A’ | 1 | 2 | 3 | F |
|---|---|---|---|---|
| 1 | 0 | 3 | 0 | 0 |
| 2 | 0 | 0 | 1 | 2 |
| 3 | 1 | 0 | 1 | 2 |

| B’ | a | b |
|---|---|---|
| 1 | 3 | 0 |
| 2 | 0 | 3 |
| 3 | 3 | 1 |

Normalizando:

| $\pi$ | 1 | 2 | 3 |
|---|---|---|---|
| | 1/2 | 0 | 1/2 |

| A | 1 | 2 | 3 | F |
|---|---|---|---|---|
| 1 | 0 | 1 | 0 | 0 |
| 2 | 0 | 0 | 1/3 | 2/3 |
| 3 | 1/4 | 0 | 1/4 | 1/2 |

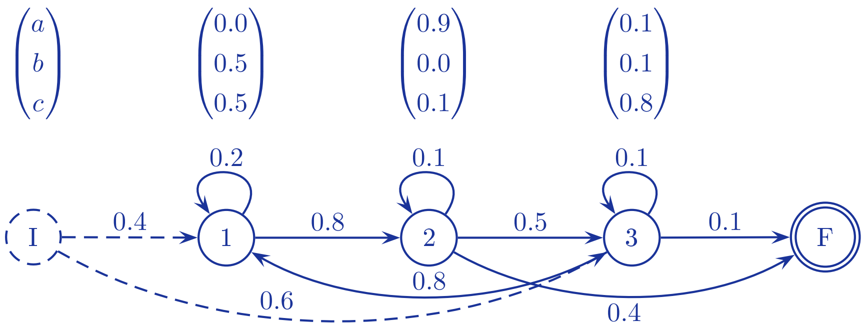
| B | a | b |
|---|---|---|
| 1 | 1 | 0 |
| 2 | 0 | 1 |
| 3 | 3/4 | 1/4 |

16. Sea $\mathcal{M}$ un modelo de Markov de cuatro estados sobre $\Sigma = \{a, b\}$, definido por $\pi_1 = 0.4, \pi_2 = \pi_F = 0, \pi_3 = 0.6$ y las matrices A y B:

| A | 1 | 2 | 3 | F |
|---|---|---|---|---|
| 1 | 0.2 | 0.8 | 0.0 | 0.0 |
| 2 | 0.0 | 0.1 | 0.5 | 0.4 |
| 3 | 0.8 | 0.0 | 0.1 | 0.1 |

| B | a | b | c |
|---|---|---|---|
| 1 | 0.0 | 0.5 | 0.5 |
| 2 | 0.9 | 0.0 | 0.1 |
| 3 | 0.1 | 0.1 | 0.8 |

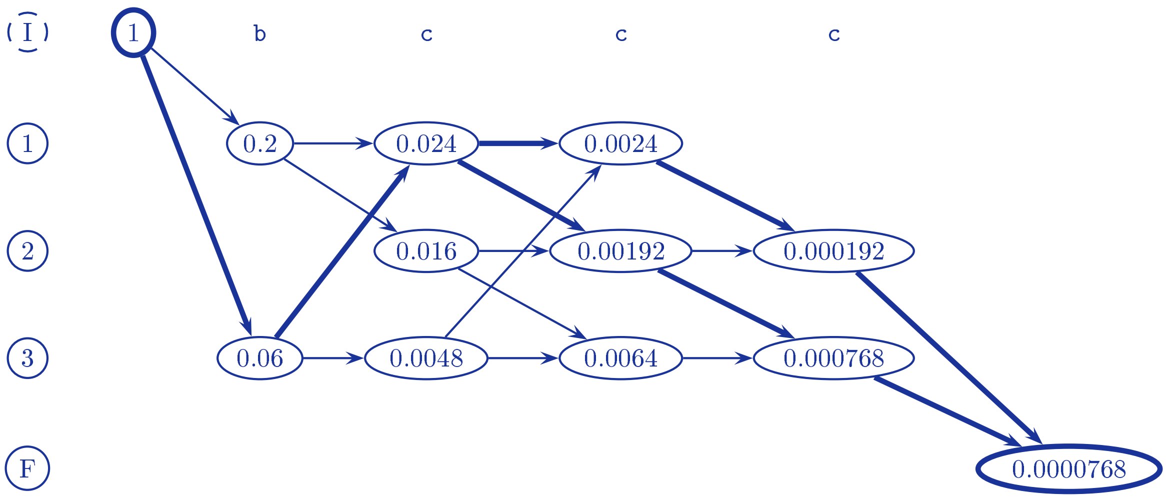
*a*) Representa gráficamente el modelo.




*b*) Determina cuál es la probabilidad de que el modelo genere la cadena "`abab`".

$$
\begin{aligned}
p(abab) &= p(abab, 3123F) + p(abab, 3333F)\\
&= 0.6\,0.1 \cdot 0.8\,0.5 \cdot 0.8\,0.9 \cdot 0.5\,0.1 \cdot 0.1 + 0.6\,0.1 \cdot 0.1\,0.1 \cdot 0.1\,0.1 \cdot 0.1\,0.1 \cdot 0.1\\
&= 0.0000864 + 0.000000006\\
&= 0.000086406
\end{aligned}
$$

*c*) Realiza una traza del algoritmo de Viterbi para obtener la secuencia de estados mas probable con la que $\mathcal{M}$ genera la cadena "`bccc`".




Hay dos caminos óptimos (cualquiera de ellos es solución al problema):

$$\tilde{z} = (3, 1, 1, 2, F) \qquad \text{y} \qquad \tilde{z}' = (3, 1, 2, 3, F)$$

$$p(bccc, \tilde{z}) = p(bccc, \tilde{z}') = 0.0000768$$

17. Se pretende aprender un Modelo de Markov lineal de conjunto de estados $Q = \{0, 1, 2, F\}$ y conjunto de símbolos $\Sigma = \{a, b, c\}$, a partir del siguiente conjunto de cadenas de aprendizaje:

$$A = \{acbbcc, aabbbc, abcc\}$$

*a*) Obtén un modelo inicial $M_0$ por segmentación lineal de las cadenas de $A$,

Una segmentación lineal reparte uniforme y secuencialmente los símbolos de cada cadena entre los 3 estados disponibles. Esto corresponde a la secuencia de estados 112233 para $acbbcc, aabbbc$ y 1223 para $abcc$, de longitudes 6 y 4, respectivamente. Como 4 no es múltiplo exacto 3, dependiendo de como se haga el redondeo, hay otras secuencias de estados posibles para $abcc$. Por ejemplo, 1123, 1233 son adecuadas, e incluso 1122 o 2233, siempre que las probabilidades iniciales y/o de estados finales se calculen correctamente. Los resultados que siguen corresponden a la secuencia 1223.

$M_0 = (\pi_0, A_0, B_0)$:

$\pi_0 = (1, 0, 0)$

| $A_0$ | 1 | 2 | 3 | F |
|---|---|---|---|---|
| 1 | 2/5 | 3/5 | 0 | 0 |
| 2 | 0 | 1/2 | 1/2 | 0 |
| 2 | 0 | 0 | 2/5 | 3/5 |

| $B_0$ | a | b | c |
|---|---|---|---|
| 1 | 4/5 | 0 | 1/5 |
| 2 | 0 | 5/6 | 1/6 |
| 3 | 0 | 1/5 | 4/5 |

*b*) Calcula la aproximación de Viterbi a la probabilidad de que $M_0$ genere el conjunto de aprendizaje $A$.

$\tilde{P}(A \mid M_0) = \tilde{P}(acbbcc \mid M_0) \cdot \tilde{P}(aabbbc \mid M_0) \cdot \tilde{P}(abcc \mid M_0) \approx 0.0011 \cdot 0.0053 \cdot 0.0307 \approx 1.7 \cdot 10^{-7}$

$\tilde{P}(acbbcc \mid M_0) \approx 0.0011$:

```
          a      c      b      b      c      c
    0 0.800  0.064
    1        0.080 0.0333 0.0139 0.0012 0.0001
    2              0.0080 0.0033 0.0056 0.0018 0.0011
    Q:    1      2      2      2      3      3      F
```

$\tilde{P}(aabbbc \mid M_0) \approx 0.0053$:

```
          a      a      b      b      b      c
    0   0.8  0.256
    1               0.128 0.0530 0.0222 0.0018
    2                     0.0128 0.0053 0.0088 0.0053
    Q:    0      0      1      1      1      2      F
```

$\tilde{P}(abcc \mid M_0) \approx 0.0307$:

```
          a      b      c      c
    0   0.8
    1          0.4 0.0333 0.0028
    2              0.1600 0.0512 0.0307
    Q:    0      1      2      2      F
```

*c*) Obtén un nuevo modelo $M_1$, mediante *una* iteración de re-estimación por Viterbi a partir de $M_0$,

$M_1 = (\pi_1, A_1, B_1)$:

$\pi_1 = (1, 0, 0)$

| $A_1$ | 1 | 2 | 3 | F |
|---|---|---|---|---|
| 1 | 1/4 | 3/4 | 0 | 0 |
| 2 | 0 | 4/7 | 3/7 | 0 |
| 2 | 0 | 0 | 2/5 | 3/5 |

| $B_1$ | a | b | c |
|---|---|---|---|
| 1 | 1 | 0 | 0 |
| 2 | 0 | 6/7 | 1/7 |
| 3 | 0 | 0 | 1 |

*d*) Calcula la aproximación de Viterbi a la probabilidad de que $M_1$ genere el conjunto de aprendizaje $A$.

$\tilde{P}(A \mid M_1) = \tilde{P}(acbbcc \mid M_1) \cdot \tilde{P}(aabbbc \mid M_1) \cdot \tilde{P}(abcc \mid M_1) \approx 0.0026 \cdot 0.0099 \cdot 0.0661 \approx 1.7 \cdot 10^{-6}$

$\tilde{P}(acbbcc \mid M_1) \approx 0.0026$:

```
          a      c      b      b      c      c
    0   1.0
    1       0.1071 0.0525 0.0257 0.0021 0.0002
    2                            0.0110 0.0044 0.0026
    Q:    0      1      1      1      2      2      F
```

$\tilde{P}(aabbbc \mid M_1) \approx 0.0099$:

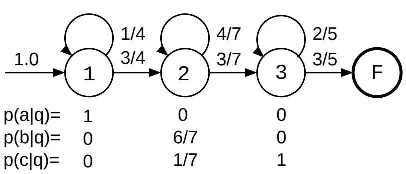
```
          a      a      b      b      b      c
    0   1.0   0.25
    1              0.1607 0.0787 0.0386 0.0031
    2                                   0.0165 0.0099
    Q:    0      0      1      1      1      2      F
```

$\tilde{P}(abcc \mid M_1) \approx 0.0661$:

```
          a      b      c      c
    0   1.0
    1       0.6429 0.0524 0.0042
    2              0.2755 0.1102 0.0661
    Q:    0      1      2      2      F
```

*e*) Representa gráficamente $M_1$ e indica cuál ha sido el cambio cualitativo más significativo que ha ocurrido entre $M_0$ y $M_1$.

$M_1$ genera las cadenas de aprendizaje con una verosimilitud 10 veces mayor que $M_0$. Además de esto, el cambio cualitativo más significativo de $M_0$ a $M_1$ ha sido un aumento de la especialización de los estados en la emisión de símbolos: En $M_0$ el estado 1 emite $a$'s y $c$'s, y el 2 y 3 emiten $b$'s y $c$'s. Sin embargo, en $M_1$ el estado 1 emite solo $a$'s el 2 (casi) solo $b$'s y el 3 solo $c$'s. La representación gráfica se muestra a continuación:



18. Se tiene un problema de clasificación en dos clases $X$ e $Y$, con probabilidades a priori respectivas de 0.3 y 0.7. Se dispone de modelos de Markov para cada clase, $M_X$ y $M_Y$, cuyos parámetros son:

$\pi_{X1} = \pi_{Y1} = 1$
$\pi_{X2} = \pi_{Y2} = 0$

| $A_X$ | 1 | 2 | F |
|---|---|---|---|
| 1 | 0.8 | 0.2 | 0 |
| 2 | 0.7 | 0 | 0.3 |

| $B_X$ | a | b |
|---|---|---|
| 1 | 0.9 | 0.1 |
| 2 | 0.1 | 0.9 |

| $A_Y$ | 1 | 2 | F |
|---|---|---|---|
| 1 | 0.8 | 0.2 | 0 |
| 2 | 0.7 | 0 | 0.3 |

| $B_Y$ | a | b |
|---|---|---|
| 1 | 0.1 | 0.9 |
| 2 | 0.9 | 0.1 |

*a*) Calcular la aproximación de Viterbi a las probabilidades de generar la cadena $y$ = "`b b a b`" por $M_X$ y $M_Y$. Presentar las trazas de ejecución correspondientes.

*b*) Trazas del algoritmo de Viterbi para el cálculo de $\tilde{P}(y \mid M_X)$ y $\tilde{P}(y \mid M_Y)$:

```
Traza Mx:   b       b       a       b
    1 0.10000 0.00800 0.01134 0.00091
    2         0.01800 0.00016 0.00204 0.00061
    Q:      1       2       1       2       F
```

```
Traza My:   b       b       a       b
    1 0.90000 0.64812 0.05185 0.07352
    2         0.01800 0.11667 0.00104 0.00031
    Q:      1       1       1       2       F
```

$\tilde{P}(y \mid M_X) \approx 0.00061\,, \quad \tilde{P}(y \mid M_Y) \approx 0.00031$

*c*) Clasificar $y$ usando esta aproximación.
$P(X \mid y) = P(y \mid M_X)P(X)/P(y) \approx 0.46\,, \quad P(Y \mid y) = P(y \mid M_y)P(Y)/P(y) \approx 0.54.$
Por tanto, $y$ se clasifica en el clase $Y$.

*d*) Calcular la verdadera probabilidad de generar $y$ mediante $M_X$ y el error de la aproximación de Viterbi.
Aparte de la secuencia de estados óptima (de Viterbi) $< 1, 2, 1, 2 >$, hay otra secuencia, $< 1, 1, 1, 2 >$, mediante la que $M_X$ también puede generar $y$ con probabilidad $0.00031 > 0$. Por tanto la verdadera probabilidad de generación de $y$ por $M_X$ es $0.00061 + 0.00031 = 0.00092$ y el error de la aproximación es 0.00031 (34 %).

*e*) Determinar cuáles son las dos cadenas (de 3 o menos símbolos) que se generan con mayor probabilidad (verdadera) mediante cada uno de los modelos $M_X$ y $M_Y$. Calcular el error de la aproximación de Viterbi para cada una de estas cuatro cadenas.
Para $M_X$ : {"`a b`", "`a a b`"}; Para $M_Y$ : {"`b a`", "`b b a`"}
$P(\texttt{a b} \mid M_X) = P(\texttt{b a} \mid M_Y) \approx 0.049, \quad P(\texttt{a a b} \mid M_X) = P(\texttt{b b a} \mid M_Y) \approx 0.035$
El error de la aproximación de Viterbi es 0 en todos los casos, ya que las cuatro cadenas se generan mediante secuencias de estados únicas.

19. Considérese una tarea de clasificación de cadenas según sus longitudes en dos clases, $\mathcal{S}$ (cadenas cortas) y $\mathcal{L}$ (cadenas largas), con probabilidades a priori $P(\mathcal{S}) = 0.8$ y $P(\mathcal{L}) = 0.2$, respectivamente. Para simplificar, supongamos que las cadenas constan de un único símbolo, $a$. Para ello, se utilizan dos modelos de Markov de un solo estado, $M_{\mathcal{S}}$ y $M_{\mathcal{L}}$, cuyos parámetros son:

| | $\pi_{\mathcal{S}}$ |
|---|---|
| 1 | 1.0 |

| $A_{\mathcal{S}}$ | 1 | F |
|---|---|---|
| 1 | 0.4 | 0.6 |

| $B_{\mathcal{S}}$ | a |
|---|---|
| 1 | 1.0 |

| | $\pi_{\mathcal{L}}$ |
|---|---|
| 1 | 1.0 |

| $A_{\mathcal{L}}$ | 1 | F |
|---|---|---|
| 1 | 0.6 | 0.4 |

| $B_{\mathcal{L}}$ | a |
|---|---|
| 1 | 1.0 |

*a*) Sean $y_1 =$"$aaaaa$", $y_2 =$"$aaaaaa$". Determinar las probabilidades con la que cada modelo genera estas cadenas, $P(y \mid c)$, $y \in \{y_1, y_2\}$, $c \in \{\mathcal{S}, \mathcal{L}\}$. Indicar también los valores de las correspondientes aproximaciones de Viterbi a estas probabilidades.

Las longitudes de $y_1$ e $y_2$ son 5 y 6, respectivamente. Por tanto:

$P(y_1 \mid \mathcal{S}) = 0.4^4 \cdot 0.6 = 0.015360$ $\qquad P(y_2 \mid \mathcal{S}) = 0.4^5 \cdot 0.6 = 0.006144$

$P(y_1 \mid \mathcal{L}) = 0.6^4 \cdot 0.4 = 0.051840$ $\qquad P(y_2 \mid \mathcal{L}) = 0.6^5 \cdot 0.4 = 0.031104$

Como estos modelos son no-ambiguos, la aproximación de Viterbi produce estas mismas probabilidades exactas.

*b*) Determinar las probabilidades a posteriori $P(\mathcal{S} \mid y)$, $P(\mathcal{L} \mid y)$, $y \in \{y_1, y_2\}$.

$P(y_1) = 0.8 \cdot P(y_1 \mid \mathcal{S}) + 0.2 \cdot P(y_1 \mid \mathcal{L}) = 0.022656$

$P(y_2) = 0.8 \cdot P(y_2 \mid \mathcal{S}) + 0.2 \cdot P(y_2 \mid \mathcal{L}) = 0.011136$

$P(\mathcal{S} \mid y_1) = \frac{0.8}{P(y_1)} \cdot P(y_1 \mid \mathcal{S}) = 0.542373$ $\qquad P(\mathcal{S} \mid y_2) = \frac{0.8}{P(y_2)} \cdot P(y_2 \mid \mathcal{S}) = 0.441379$

$P(\mathcal{L} \mid y_1) = \frac{0.2}{P(y_1)} \cdot P(y_1 \mid \mathcal{L}) = 0.457627$ $\qquad P(\mathcal{L} \mid y_2) = \frac{0.2}{P(y_2)} \cdot P(y_2 \mid \mathcal{L}) = 0.558621$

*c*) Clasificar $y_1$ e $y_2$ mediante el clasificador de mínimo riesgo de error o de Bayes. Determinar las probabilidades de que estas clasificaciones sean erróneas.

La clasificación de mínimo riesgo de error de $y_1$ es $\mathcal{S}$ y la de $y_2$ es $\mathcal{L}$. La probabilidad de que estas clasificaciones sean erróneas son, respectivamente, $1 - 0.542373 = 0.457627$ y $1 - 0.558621 = 0.441379$.

*d*) Calcular la probabilidad de error de este clasificador. Para este cálculo, pueden resultar útiles las siguientes fórmulas de sumas de series geométricas (en las que las $i$es se corresponden con las longitudes de las cadenas):

$$\sum_{i=m}^{n} r^i = \frac{r^m - r^{(n+1)}}{1-r} \qquad \sum_{i=1}^{\infty} r^i = \frac{r}{1-r}$$

Sea $i$ la longitud de una cadena $y$. La probabilidad de error viene dada por:

$$P(error) = \sum_{y:i=1}^{\infty} P(error \mid y) \cdot P(y)$$

Para $i \leq 5$, todas las cadenas se clasifican en la clase $\mathcal{S}$, mientras que para $i > 5$, la clasificación es $\mathcal{L}$. Por tanto:

$$\begin{aligned}
P(error) &= \sum_{y:i=1}^{5} P(\mathcal{L} \mid y) \cdot P(y) + \sum_{y:i=6}^{\infty} P(\mathcal{S} \mid y) \cdot P(y) \\
&= \sum_{i=1}^{5} 0.2 \cdot 0.6^{i-1} \cdot 0.4 + \sum_{i=6}^{\infty} 0.8 \cdot 0.4^{i-1} \cdot 0.6 \\
&= 0.2 \cdot 0.4 \cdot \sum_{i=0}^{4} 0.6^i + 0.8 \cdot 0.6 \cdot \sum_{i=5}^{\infty} 0.4^i \\
&= 0.2 \cdot 0.4 \cdot \frac{1 - 0.6^5}{1 - 0.6} + 0.8 \cdot 0.6 \cdot \left( \sum_{i=0}^{\infty} 0.4^i - \sum_{i=0}^{4} 0.4^i \right) \\
&= 0.2 \cdot (1 - 0.6^5) + 0.8 \cdot 0.6 \cdot \left( \frac{1}{1-0.4} - \frac{1 - 0.4^5}{1 - 0.4} \right) \\
&= 0.2 \cdot (1 - 0.6^5) + 0.8 \cdot 0.4^5 \\
&= 0.184448 + 0.36 = 0.544448
\end{aligned}$$

20. Considérese una tarea de clasificación de cadenas de símbolos del alfabeto $\Sigma = \{a, b\}$ en dos clases, $V$ y $W$, con probabilidades a priori $P(V) = 0.6$ y $P(W) = 0.4$. Para ello, se utilizan dos modelos de Markov, $M_V$ y $M_W$, ambos con dos estados ($Q = \{1, 2, F\}$), cuyos parámetros son:

| $q$ | $\pi_V$ |
|---|---|
| 1 | 0.8 |
| 2 | 0.2 |

| $A_V$ | 1 | 2 | F |
|---|---|---|---|
| 1 | 0.7 | 0.3 | 0.0 |
| 2 | 0.0 | 0.5 | 0.5 |

| $B_V$ | a | b |
|---|---|---|
| 1 | 0.9 | 0.1 |
| 2 | 0.8 | 0.2 |

| $q$ | $\pi_W$ |
|---|---|
| 1 | 1.0 |
| 2 | 0.0 |

| $A_W$ | 1 | 2 | F |
|---|---|---|---|
| 1 | 0.7 | 0.3 | 0.0 |
| 2 | 0.0 | 0.5 | 0.5 |

| $B_W$ | a | b |
|---|---|---|
| 1 | 0.9 | 0.1 |
| 2 | 0.8 | 0.2 |

*a*) Calcula las probabilidades exactas (*no* la aproximación de Viterbi) de que $M_V$ y $M_W$ generen la cadena "`aab`".
Hay tres secuencias de estados en $M_V$ que generan la cadena $aab$: $z_1 = 1, 1, 2, F$, $z_2 = 1, 2, 2, F$, $z_3 = 2, 2, 2, F$; análogamente, en $M_W$ hay dos secuencias: $z_1 = 1, 1, 2, F$, $z_2 = 1, 2, 2, F$. Por tanto:

$$\begin{aligned} P(aab \mid V) &= (0.8 \cdot 0.9)\ (0.7 \cdot 0.9)\ (0.3 \cdot 0.2)\ 0.5 \\ &+ (0.8 \cdot 0.9)\ (0.3 \cdot 0.8)\ (0.5 \cdot 0.2)\ 0.5 \\ &+ (0.2 \cdot 0.8)\ (0.5 \cdot 0.8)\ (0.5 \cdot 0.2)\ 0.5 \\ &= 0.01361 + 0.00864 + 0.0032 \ = \ \mathbf{0.02545} \end{aligned}$$

$$\begin{aligned} P(aab \mid W) &= (1 \cdot 0.9)\ (0.7 \cdot 0.9)\ (0.3 \cdot 0.2)\ 0.5 \\ &+ (1 \cdot 0.9)\ (0.3 \cdot 0.8)\ (0.5 \cdot 0.2)\ 0.5 \\ &= 0.01701 + 0.0108 \ = \ \mathbf{0.02781} \end{aligned}$$

*b*) Clasifica la cadena "`aab`" por mínima probabilidad de error.

$$c^*(aab) \ = \ \arg\max_{c \in \{V,W\}} P(c \mid aab) \ = \ \arg\max_{c \in \{V,W\}} P(aab \mid c) \cdot P(c)$$

$$P(aab \mid V)\, P(V) = 0.02545 \cdot 0.6 = \mathbf{0.015270} \ > \ P(aab \mid W)\, P(W) = 0.02781 \cdot 0.4 = \mathbf{0.011124}$$

Por tanto, $c^*(aab) \ = \ \boldsymbol{V}$.

*c*) Determina la probabilidad de error de la clasificación obtenida en el apartado anterior.

$$P(\text{error} \mid aab) \ = \ 1 - \max_{c \in \{V,W\}} P(c \mid aab) \ = \ 1 - P(V \mid aab) \ = \ P(W \mid aab) \ = \ \frac{P(aab \mid W)\, P(W)}{P(aab)}$$

$$P(aab) \ = \ P(aab \mid V)\, P(V) + P(aab \mid W)\, P(W) \ = \ 0.011124 + 0.015270 = 0.026394$$

$$P(\text{error} \mid aab) \ = \ \frac{0.011124}{0.026394} \ = \ \mathbf{0.42145}$$

*d*) Repite los cálculos de los dos apartados anteriores, utilizando las probabilidades obtenidas mediante la aproximación de Viterbi en lugar de las probabilidades exactas.

$$\tilde{P}(aab \mid V) \ = \ \max_{z \in Q^+} P(aab, z \mid V) \ = \ \max(0.01361, 0.00864, 0.0032) \ = \ \mathbf{0.01361}$$

$$\tilde{P}(aab \mid W) \ = \ \max_{z \in Q^+} P(aab, z \mid W) \ = \ \max(0.01701 + 0.0108) \ = \ \mathbf{0.01701}$$

$$\tilde{c}^*(aab) \ = \ \arg\max_{c \in \{V,W\}} P(c \mid aab) \ \approx \ \arg\max_{c \in \{V,W\}} \tilde{P}(aab \mid c) \cdot P(c)$$

$$\tilde{P}(aab \mid V)\, P(V) = 0.01361 \cdot 0.6 = \mathbf{0.008166} \ > \ \tilde{P}(aab \mid W)\, P(W) = 0.01701 \cdot 0.4 = \mathbf{0.006804}$$

Por tanto, $\tilde{c}^*(aab) \ = \ \boldsymbol{V}$.

$$P(\text{error} \mid aab) \ = \ 1 - \max_{c \in \{V,W\}} P(c \mid aab) \ = \ 1 - P(V \mid aab) \ = \ P(W \mid aab) \ \approx \ \frac{\tilde{P}(aab \mid W)\, P(W)}{\tilde{P}(aab)}$$

$$\tilde{P}(aab) \ = \ \tilde{P}(aab \mid V)\, P(V) + \tilde{P}(aab \mid W)\, P(W) \ = \ 0.008166 + 0.006804 = 0.01497$$

$$P(\text{error} \mid aab) \ \approx \ \frac{0.006804}{0.01497} \ = \ \mathbf{0.45451}$$

21. Los modelos de Markov discretos que se estudian en SIN pueden extenderse fácilmente para que, en lugar de generar secuencias de *símbolos*, generen secuencias de *vectores* definidos en $\mathbb{R}^d$. Para ello basta asociar a cada estado no final una densidad condicional sobre $\mathbb{R}^d$. Así pues, estos modelos se definen mediante una cuádrupla $(Q, \pi, A, D)$, donde $Q, \pi$ y $A$ son los usuales y $D$ (que sustituye a la matriz de probabilidades de emisión de símbolos, $B$) representa ahora a las *densidades condicionales* $p(\mathbf{y} \mid q)$, $\mathbf{y} \in \mathbb{R}^d$, $q \in Q - \{F\}$. La probabilidad con la que uno de estos modelos genera una secuencia de *vectores* $Y = \mathbf{y}_1, \mathbf{y}_2, \ldots, \mathbf{y}_m$ se calcula de forma similar al caso de modelos discretos, con la diferencia de que la probabilidad de emisión de cada vector $\mathbf{y}_t$ se obtienen mediante $p(\mathbf{y}_t \mid q_t)$, en vez de $B_{q_t, y_t}$.

Sea $M$ uno de estos modelos definidos en $\mathbb{R}^1$ (recta real), con $Q = \{1, 2, F\}$, $\pi(1) = 1.0$, $\pi(2) = 0.0$,

| A | 1 | 2 | F |
|---|---|---|---|
| 1 | 0.0 | 1.0 | 0.0 |
| 2 | 0.5 | 0.2 | 0.3 |

$$p(\mathbf{y} \mid 1) = \begin{cases} 1/4 & -3 \le y < 1 \\ 0 & \text{en otros casos} \end{cases} \qquad p(\mathbf{y} \mid 2) = \begin{cases} 1/5 & -1 \le y < 4 \\ 0 & \text{en otros casos} \end{cases}$$

Por ejemplo, $M$ genera la secuencia $X = 0.0,\ 0.3,\ 0.0$ mediante una *única* secuencia de estados $z = 1, 2, 2, F$, con probabilidad no nula:

$$\begin{aligned} p(X \mid M) = p(X, z) = p(z) \cdot p(X \mid z) &= (\pi(1) \cdot A_{1,2} \cdot A_{2,2} \cdot A_{2,F}) \cdot (p(0.0 \mid 1) \cdot p(0.3 \mid 2) \cdot p(0.0 \mid 2)) \\ &= (1.0 \cdot 1.0 \cdot 0.2 \cdot 0.3) \cdot (1/4 \cdot 1/5 \cdot 1/5) \\ &= 0.06 \cdot 0.01 = 0.0006 \end{aligned}$$

*a*) Encontrar dos secuencias de estados $z', z''$ que generen la secuencia $Y = 0.0,\ 0.0,\ 0.0,\ 0.0$ con probabilidad no nula. ¿Hay más secuencias de estados que generen $Y$ con probabilidad no nula?

$z' = 1, 2, 1, 2, F$; $z'' = 1, 2, 2, 2, F$. Cualquier otra secuencia de estados tiene probabilidad nula de generar $Y$.

*b*) Calcular $p(Y \mid z')$, $p(Y \mid z'')$, $\tilde{p}(Y \mid M)$ y $p(Y \mid M)$ (las dos últimas son la aproximación de Viterbi y la verdadera probabilidad de generar $Y$ mediante $M$).

$p(z') = 1.0 \cdot 1.0 \cdot 0.5 \cdot 1.0 \cdot 0.3 = 0.15$; $p(z'') = 1.0 \cdot 1.0 \cdot 0.2 \cdot 0.2 \cdot 0.3 = 0.012$;
$p(Y \mid z') = 1/4 \cdot 1/5 \cdot 1/4 \cdot 1/5 = 0.0025$; $p(Y \mid z'') = 1/4 \cdot 1/5 \cdot 1/5 \cdot 1/5 = 0.0020$;
$\tilde{p}(Y \mid M) = \text{máx}(0.15 \cdot 0.0025,\ 0.012 \cdot 0.002) = 0.000375$
$p(Y \mid M) = 0.15 \cdot 0.0025 + 0.012 \cdot 0.002 = 0.000399$.

*c*) Sea $Y' = -2.0,\ -0.5,\ 0.0,\ 0.5,\ 2.0$. ¿Cuántas secuencias de estados de $M$ generan $Y'$ con probabilidad no nula?

Sólo hay tres secuencias de estados que pueden generar cadenas de longitud 5 (número de vectores de $Y'$): $z^1 = 1, 2, 1, 2, 2$, $z^2 = 1, 2, 2, 1, 2$, $z^3 = 1, 2, 2, 2, 2$, . Todas ellas generan $Y'$ con probabilidad mayor que cero.

*d*) Calcular la aproximación de Viterbi a la probabilidad con la que $M$ genera $Y'$.

Podría utilizarse el algoritmo de Viterbi pero, dado que solo hay tres secuencias de estados, es preferible hacer el cálculo exhaustivo:

$p(Y' \mid z^1) = 1.0 \cdot 1.0 \cdot 0.5 \cdot 1.0 \cdot 0.2 \cdot 0.3 \cdot 1/4 \cdot 1/5 \cdot 1/4 \cdot 1/5 \cdot 1/5 = 0.03 \cdot 0.0005 = 0.000015$
$p(Y' \mid z^2) = 1.0 \cdot 1.0 \cdot 0.2 \cdot 0.5 \cdot 1.0 \cdot 0.3 \cdot 1/4 \cdot 1/5 \cdot 1/5 \cdot 1/4 \cdot 1/5 = 0.03 \cdot 0.0005 = 0.000015$
$p(Y' \mid z^3) = 1.0 \cdot 1.0 \cdot 0.2 \cdot 0.2 \cdot 0.2 \cdot 0.3 \cdot 1/4 \cdot 1/5 \cdot 1/5 \cdot 1/5 \cdot 1/5 = 0.0024 \cdot 0.0004 = 0.00000096$
$\tilde{p}(Y' \mid M) = \text{máx}(0.000015,\ 0.000015,\ 0.00000096) = 0.000015$

22. Sea $M$ un modelo de Markov de conjunto de estados $Q = \{1, 2, 3, 4, F\}$; alfabeto $\Sigma = \{a, b\}$; probabilidades iniciales $\pi_0(1) = \pi_0(2) = \pi_0(4) = \frac{1}{3}$, $\pi_0(3) = \pi_0(F) = 0$; y probabilidades de transición entre estados y de emisión de símbolos:

| $A$ | 1 | 2 | 3 | 4 | $F$ |
|---|---|---|---|---|---|
| 1 | $\frac{1}{2}$ | 0 | 0 | 0 | $\frac{1}{2}$ |
| 2 | 0 | $\frac{1}{2}$ | $\frac{1}{2}$ | 0 | 0 |
| 3 | 0 | 0 | $\frac{1}{2}$ | 0 | $\frac{1}{2}$ |
| 4 | 0 | 0 | 0 | $\frac{1}{2}$ | $\frac{1}{2}$ |

| $B$ | $a$ | $b$ |
|---|---|---|
| 1 | 1 | 0 |
| 2 | 1 | 0 |
| 3 | 0 | 1 |
| 4 | 0 | 1 |

*a*) Sean $n$ y $m$ dos enteros positivos. Halla las probabilidades $P(a^n \mid M)$, $P(a^n\, b^m \mid M)$ y $P(b^n \mid M)$.

$$
\begin{aligned}
P(a^n \mid M) &= \frac{1}{3}\frac{1}{2^n} \\
P(a^n\, b^m \mid M) &= \frac{1}{3}\frac{1}{2^{n+m}} \\
P(b^n \mid M) &= \frac{1}{3}\frac{1}{2^n}
\end{aligned}
$$

*b*) Calcula la probabilidad de que $M$ genere una cadena cualquiera de longitud comprendida entre 5 y 7.

El conjunto de cadenas de longitud 7 que $M$ puede generar es:

$$S_7 = \{a^7, a^6 b, a^5 b^2, a^4 b^3, a^3 b^4, a^2 b^5, a b^6, b^7\}$$

Cada una de las 8 cadenas de $S_7$ se genera con probabilidad $\frac{1}{3}\frac{1}{2^7}$.

Análogamente:

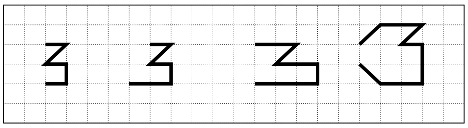
hay 7 cadenas de longitud 6 que se generan con probabilidad $\frac{1}{3}\frac{1}{2^6}$

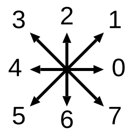
hay 6 cadenas de longitud 5 que se generan con probabilidad $\frac{1}{3}\frac{1}{2^5}$

Por tanto, si $S$ es el conjunto de cadenas de longitudes comprendidas entre 5 y 7,

$$P(x \in S \mid M) = 8\,\frac{1}{3}\,\frac{1}{2^7} + 7\,\frac{1}{3}\,\frac{1}{2^6} + 6\,\frac{1}{3}\,\frac{1}{2^5} \approx 0.02083 + 0.03646 + 0.06250 \approx 0.11979$$

23. En la figura se pueden ver muestras generadas por un Modelo de Markov de 7 estados con topología lineal. Cada trazo elemental de estas muestras corresponde a un símbolo del *código de contorno de 8 direcciones* $\Sigma = \{“0”, “1”, \ldots, “8”\}$, que también se puede ver en la figura.






Sea $\mathcal{M}_0$ un Modelo de Markov lineal de 7 estados $Q = \{1, 2, 3, 4, 5, 6, 7, F = 8\}$, con los siguientes parámetros:

$$\pi_q = \begin{cases} 1/2 & \text{si } q \in \{1, 2\} \\ 0 & \text{en otros casos} \end{cases}$$

$$A_{q'q} = \begin{cases} 1 & \text{si } q' = 1, q = 2 \\ 1/2 & \text{si } q \in \{q', q'+1\},\ 2 \le q' \le 5 \\ 1/3 & \text{si } q' = 6, q \ge 6 \\ 1 & \text{si } q' = 7, q = 8 \\ 0 & \text{en otros casos} \end{cases}$$

$$B_{q\sigma} = \begin{cases} 1 & \text{si } (q, \sigma) \in \{(1,“1”),\ (2,“0”),\ (3,“5”),\ (4,“0”),\ (5,“6”),\ (6,“4”),\ (7,“3”)\} \\ 0 & \text{en otros casos} \end{cases}$$

*a*) Calcular la aproximación de Viterbi a la probabilidad con la que $\mathcal{M}_0$ genera cada una de las muestras de la figura.
Cadenas de entrenamiento: “0 5 0 6 4” , “0 5 0 6 4 4” , “0 0 5 0 0 6 4 4 4” , “1 0 0 5 0 6 6 4 4 3”.

HMM inicial $\mathcal{M}_0$:

| $q$ | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8=F |
|---|---|---|---|---|---|---|---|---|
| $\pi_q$ | 1/2 | 1/2 | 0 | 0 | 0 | 0 | 0 | – |

| $q' \backslash q$ | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8=F |
|---|---|---|---|---|---|---|---|---|
| 1 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 |
| 2 | 0 | 1/2 | 1/2 | 0 | 0 | 0 | 0 | 0 |
| 3 | 0 | 0 | 1/2 | 1/2 | 0 | 0 | 0 | 0 |
| 4 | 0 | 0 | 0 | 1/2 | 1/2 | 0 | 0 | 0 |
| 5 | 0 | 0 | 0 | 0 | 1/2 | 1/2 | 0 | 0 |
| 6 | 0 | 0 | 0 | 0 | 0 | 1/3 | 1/3 | 1/3 |
| 7 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 |

| $q \backslash \sigma$ | “0” | “1” | “2” | “3” | “4” | “5” | “6” | “7” |
|---|---|---|---|---|---|---|---|---|
| 1 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 |
| 2 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 3 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 |
| 4 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 5 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 |
| 6 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 |
| 7 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 |

Aproximación de Viterbi a la probabilidad con la que $\mathcal{M}_0$ genera las muestras: $0.01042, 0.00347, 0.00029, 0.00087$.

*b*) Calcular la verdadera probabilidad con la que $\mathcal{M}_0$ genera cada una de las muestras de la figura.
Las verdaderas probabilidades con la que $\mathcal{M}_0$ genera las muestras son las mismas que en el apartado a), ya que $\mathcal{M}_0$ es determinista.

*c*) Calcular la verosimilitud con la $\mathcal{M}_0$ genera todas las muestras.
Verosimilitud de todas las muestras: $9.1 \cdot 10^{-12}$.

*d*) Partiendo de $\mathcal{M}_0$, realizar una iteración del algoritmo de reestimación por Viterbi con estas muestras.
Secuencias de estados correspondientes a los cálculos de a):
$< 2, 3, 4, 5, 6, F >$, $< 2, 3, 4, 5, 6, 6, F >$, $< 2, 2, 3, 4, 4, 5, 6, 6, 6, F >$, $< 1, 2, 2, 3, 4, 5, 5, 6, 6, 7, F >$

HMM reestimado a partir de estas secuencias:

| $q$ | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8=F |
|---|---|---|---|---|---|---|---|---|
| $\pi_q$ | 1/2 | 1/2 | 0 | 0 | 0 | 0 | 0 | – |

| $q' \backslash q$ | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8=F |
|---|---|---|---|---|---|---|---|---|
| 1 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 |
| 2 | 0 | 1/3 | 2/3 | 0 | 0 | 0 | 0 | 0 |
| 3 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 |
| 4 | 0 | 0 | 0 | 1/5 | 4/5 | 0 | 0 | 0 |
| 5 | 0 | 0 | 0 | 0 | 1/5 | 4/5 | 0 | 0 |
| 6 | 0 | 0 | 0 | 0 | 0 | 1/2 | 1/8 | 3/8 |
| 7 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 |

| $q \backslash \sigma$ | “0” | “1” | “2” | “3” | “4” | “5” | “6” | “7” |
|---|---|---|---|---|---|---|---|---|
| 1 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 |
| 2 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 3 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 |
| 4 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 5 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 |
| 6 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 |
| 7 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 |

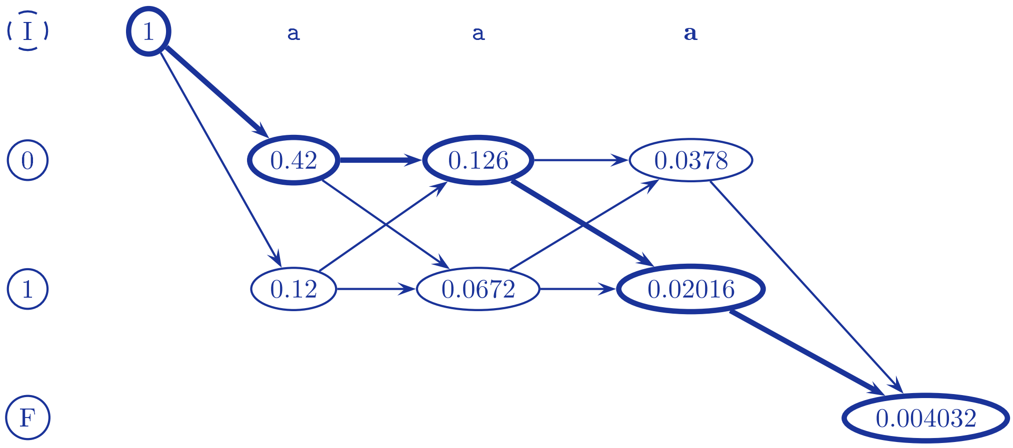
24. Sea $M$ un modelo de Markov de conjunto de estados $Q = \{0, 1, F\}$; alfabeto $\Sigma = \{a, b\}$; probabilidades iniciales $\pi_0(0) = 0.7, \pi_0(1) = 0.3$; y probabilidades de transición entre estados y de emisión de símbolos:

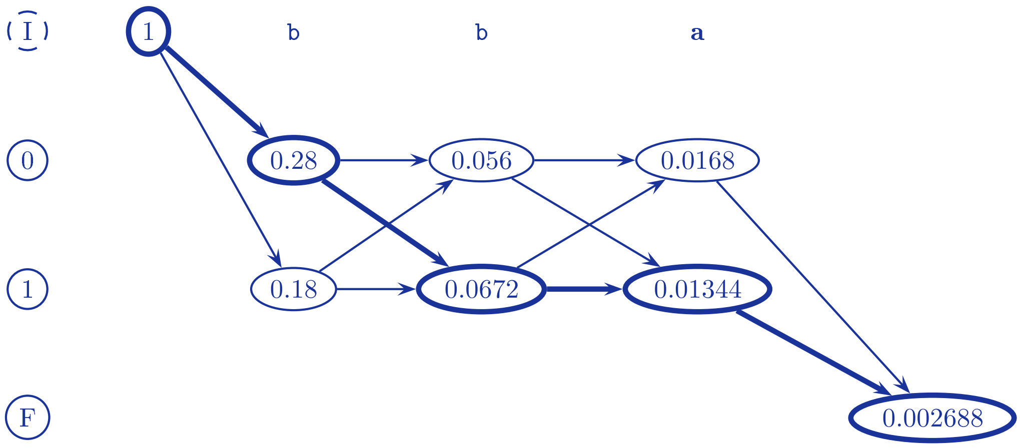
| $A$ | 0 | 1 | $F$ |
|---|---|---|---|
| 0 | 0.5 | 0.4 | 0.1 |
| 1 | 0.3 | 0.5 | 0.2 |

| $B$ | $a$ | $b$ |
|---|---|---|
| 0 | 0.6 | 0.4 |
| 1 | 0.4 | 0.6 |

Reestima los parámetros de $M$ mediante *una* iteración de reestimación por Viterbi, a partir de las cadenas de entrenamiento "`a a a`" y "`b b a`".

Primero determinamos la secuencia de estados mas probable para cada cadena mediante el algoritmo de Viterbi:




Los pares *cadena-secuencia óptima de estados* obtenidos son:

| a a a | b b a |
|---|---|
| 0 0 1 F | 0 1 1 F |

Los parámetros reestimados a partir de estos pares son:

$$\hat{\pi}_0(0) = \frac{2}{2} = 1$$

$$\hat{\pi}_0(1) = \frac{0}{2} = 0$$

| $A$ | 0 | 1 | $F$ |
|---|---|---|---|
| 0 | $\frac{1}{3}$ | $\frac{2}{3}$ | 0 |
| 1 | 0 | $\frac{1}{3}$ | $\frac{2}{3}$ |

| $B$ | $a$ | $b$ |
|---|---|---|
| 0 | $\frac{2}{3}$ | $\frac{1}{3}$ |
| 1 | $\frac{2}{3}$ | $\frac{1}{3}$ |

25. A continuación se dan 4 cadenas generadas por un modelo de Markov, cada una junto a una posible secuencias de estados:

| | | | |
|---|---|---|---|
| “a c c” | “c c a” | “c b b a” | “a b b c” |
| “0 1 2” | “3 4 5” | “3 4 4 5” | “0 1 1 2” |

Usando probabilidades aproximadas por Viterbi cuando sea necesario:

a) Inicializa un modelo de markov $\mathcal{M}_0$ a partir de estos datos. Representa gráficamente $\mathcal{M}_0$.
$\mathcal{M}_0 = (Q\Sigma, \pi, A, B)$; $Q = \{0, 1, 2, 3, 4, 5\}$, $\pi(0) = \pi(3) = 1/2$, $\pi_1 = \pi_2 = \pi_4 = \pi_5 = 0$,

| $A$ | 0 | 1 | 2 | 3 | 4 | 5 | $F$ |
|---|---|---|---|---|---|---|---|
| 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 |
| 1 | 0 | 1/3 | 2/3 | 0 | 0 | 0 | 0 |
| 2 | 0 | 0 | 0 | 0 | 0 | 0 | 1 |
| 3 | 0 | 0 | 0 | 0 | 1 | 0 | 0 |
| 4 | 0 | 0 | 0 | 0 | 1/3 | 2/3 | 0 |
| 5 | 0 | 0 | 0 | 0 | 0 | 0 | 1 |

| $B$ | a | b | c |
|---|---|---|---|
| 0 | 1 | 0 | 0 |
| 1 | 0 | 2/3 | 1/3 |
| 2 | 0 | 0 | 1 |
| 3 | 0 | 0 | 1 |
| 4 | 0 | 2/3 | 1/3 |
| 5 | 1 | 0 | 0 |

b) Obtén la versoimilitud con la que $\mathcal{M}_0$ genera las dos primeras cadenas (“a c c”, “c c a”).
$P(\text{“acc”}, \text{“cca”} \mid \mathcal{M}_0) \approx \tilde{P}(\text{“acc”} \mid \mathcal{M}_0) \cdot \tilde{P}(\text{“cca”} \mid \mathcal{M}_0) = 0.111111 \cdot 0.111111 = 0.012345679$

| | a | c | c | |
|---|---|---|---|---|
| 0 | 0.5 | - | - | |
| 1 | - | 0.166666 | 0.018518 | |
| 2 | - | - | 0.111111 | |
| 3 | - | - | - | |
| 4 | - | - | - | |
| 5 | - | - | - | |
| z: | 0 | 1 | 2 | 0.111111 |

| | c | c | a | |
|---|---|---|---|---|
| 0 | - | - | - | |
| 1 | - | - | - | |
| 2 | - | - | - | |
| 3 | 0.5 | - | - | |
| 4 | - | 0.166666 | - | |
| 5 | - | - | 0.111111 | |
| z: | 3 | 4 | 5 | 0.111111 |

c) Obtén la versoimilitud con la que las dos primeras cadenas (“a c c”, “c c a”) son generadas por un modelo $\mathcal{M}'$, definido como:
$Q' = \{0, 1, F\}$; alfabeto $\Sigma = \{a, b, c\}$; probabilidades iniciales $\pi'(0) = 4/5, \pi'_0(1) = 1/5$;
y probabilidades de transición entre estados y de emisión de símbolos:

| $A'$ | 0 | 1 | $F$ |
|---|---|---|---|
| 0 | 0 | 1/2 | 1/2 |
| 1 | 1/5 | 0 | 4/5 |

| $B'$ | a | b | c |
|---|---|---|---|
| 0 | 1/5 | 1/5 | 3/5 |
| 1 | 3/5 | 1/5 | 1/5 |

$P(\text{“acc”}, \text{“cca”} \mid \mathcal{M}') \approx \tilde{P}(\text{“acc”} \mid \mathcal{M}') \cdot \tilde{P}(\text{“cca”} \mid \mathcal{M}') = 0.001152 \cdot 0.001152 = 0.000001327$

| | a | c | c | |
|---|---|---|---|---|
| 0 | 0.16 | 0.0144 | 0.00192 | |
| 1 | 0.12 | 0.0160 | 0.00144 | |
| z: | 1 | 0 | 1 | 0.001152 |

| | c | c | a | |
|---|---|---|---|---|
| 0 | 0.48 | 0.0048 | 0.00192 | |
| 1 | 0.04 | 0.0480 | 0.00144 | |
| z: | 1 | 0 | 1 | 0.001152 |

26. Sea $M$ un modelo de Markov de conjunto de estados $Q = \{0, 1, F\}$; alfabeto $\Sigma = \{a\}$; probabilidades iniciales $\pi_0 = 0.6$, $\pi_1 = 0.4$; y probabilidades de transición entre estados:

| $A$ | 0 | 1 | $F$ |
|---|---|---|---|
| 0 | 0.0 | 0.3 | 0.7 |
| 1 | 0.7 | 0.0 | 0.3 |

a) Calcular la probabilidad con la que $M$ genera la cadena "`a a`"

Como solo hay un símbolo, la probabilidad de generarlo en cualquier estado ha de ser 1.0.

Solo hay dos caminos que generan $x$ ="`a a`": $z_1$ =`0,1,F` y $z_2$ =`1,0,F`. Por tanto, la probabilidad de generar $x$ mediante $M$ es:

$$\begin{aligned} P(x \mid M) &= P(x, z_1) + P(x, z_2) \\ &= P(z_1)P(x \mid z_1) + P(z_2)P(x \mid z_2) \\ &= P(z_1) \cdot 1 + P(z_2) \cdot 1 \\ &= 0.6 \cdot 0.3 \cdot 0.3 + 0.4 \cdot 0.7 \cdot 0.7 \\ &= 0.054 + 0.196 \\ &= 0.25 \end{aligned}$$

b) Calcular la probabilidad aproximada por Viterbi con la que $M$ genera la cadena "`a a a`". Presentar una traza de ejecución del algoritmo de Viterbi con la que se obtiene esta probabilidad.

Solo hay dos caminos que generan $y$ ="`a a a`": $z_1$ =`0,1,0,F` y $z_2$ =`1,0,1,F`. En ambos caminos hay un bucle formado por los arcos "`0,1`" y `1,0`. Por tanto el camino de mayor probabilidad será aquél en el que la probabilidad inicial por la final sea mayor; es decir, $z =$ `0,1,0,F`:

$$\begin{aligned} \tilde{P}(y \mid M) &= P(y, z) \\ &= P(z) \cdot 1 \\ &= 0.6 \cdot 0.3 \cdot 0.7 \cdot 0.7 \\ &= 0.0882 \end{aligned}$$

Mediante el algoritmo de Viterbi:

```
     a     a      a
0    0.6   0.28   0.126
1    0.4   0.18   0.084
F                        0.0882
z:   0     1      0      F
```

c) Obtener una expresión de la probabilidad aproximada por Viterbi con la que $M$ genera una cadena de longitud *par* en función de dicha longitud, $n$.

Hacemos los cálculos para varias longitudes y generalizamos por inducción:

$$\begin{aligned} \tilde{P}(\mathrm{a}^2 \mid M) &= 0.4 \cdot 0.7 \cdot 0.7 \\ \tilde{P}(\mathrm{a}^4 \mid M) &= 0.4 \cdot 0.7 \cdot (0.3 \cdot 0.7)^1 \cdot 0.7 \\ \tilde{P}(\mathrm{a}^6 \mid M) &= 0.4 \cdot 0.7 \cdot (0.3 \cdot 0.7)^2 \cdot 0.7 \\ \tilde{P}(\mathrm{a}^8 \mid M) &= 0.4 \cdot 0.7 \cdot (0.3 \cdot 0.7)^3 \cdot 0.7 \\ \ldots \quad & \ldots \\ \tilde{P}(\mathrm{a}^n \mid M) &= 0.4 \cdot 0.7 \cdot (0.3 \cdot 0.7)^{(n/2-1)} \cdot 0.7 \\ &= 0.196 \cdot (0.3 \cdot 0.7)^{(n/2-1)} \end{aligned}$$

d) Reestimar los parámetros de $M$ mediante *una* iteración de reestimación por Viterbi, a partir de las cadenas de entrenamiento "`a a`" y "`a a a`".

Las secuencias de estados óptimas para estas cadenas son `0,1,0,F` y `1,0,F`. Los correspondientes parámetros reestimados de $M$ son: $\pi_0 = \pi_1 = 1/2$, $B_{0,\text{"a"}} = B_{1,\text{"a"}} = 1$, y matriz de probabilidades de transición:

| $A$ | 0 | 1 | $F$ |
|---|---|---|---|
| 0 | 0 | 1/3 | 2/3 |
| 1 | 1 | 0 | 0 |

27. Se dispone del siguiente conjunto de cadenas de entrenamiento, $Y$, cada una de las cuales acompañada de la secuencia de estados con la que ha sido generada por $\mathcal{M}$ con mayor probabilidad (secuencia óptima de Viterbi):

```
cadenas:    a b b b a    a b b    b b b a    b b a
estados:    3 1 1 2 3    3 3 3    1 1 2 3    1 2 3
```

a) Usando estas cadenas y secuencias de estados, realiza una iteración del algoritmo de reestimación de Viterbi para obtener un modelo $\mathcal{M}$

El modelo $\mathcal{M}$ tiene tres estados además del estado final y un alfabeto de dos símbolos, $\Sigma = \{\mathtt{a},\mathtt{b}\}$. Los restantes parámetros se estiman como:

$$\pi_1 = \pi_3 = 1/2, \quad \pi_2 = \pi_F = 0$$

| A | 1 | 2 | 3 | F |
|---|---|---|---|---|
| 1 | 2/5 | 3/5 | 0 | 0 |
| 2 | 0 | 0 | 1 | 0 |
| 3 | 1/7 | 0 | 2/7 | 4/7 |

| B | a | b |
|---|---|---|
| 1 | 0 | 1 |
| 2 | 0 | 1 |
| 3 | 5/7 | 2/7 |

b) Calcula la verosimilitud (aproximada por Viterbi) del conjunto de cadenas de entrenamiento con $\mathcal{M}$.

$$\begin{aligned}\tilde{P}(Y \mid \mathcal{M}) &= \tilde{P}(\text{abbba} \mid \mathcal{M}) \cdot \tilde{P}(\text{abb} \mid \mathcal{M}) \cdot \tilde{P}(\text{bbba} \mid \mathcal{M}) \cdot \tilde{P}(\text{bba} \mid \mathcal{M}) \\ &\approx 0.005 \cdot 0.00136 \cdot 0.049 \cdot 0.1225 \quad \approx \quad 0.000000041\end{aligned}$$

c) Calcula la verdadera probabilidad de genración de la cadena "`b b a`" y el error cometido por la aproximación de Viterbi.

Solo hay dos secuencias de estados que generan la cadena "`b b a`": `<1,2,3,F>` y `<3,3,3,F>`. Por tanto,

$$P(\text{“b b a”} \mid \mathcal{M}) = \frac{1}{2}\cdot\frac{3}{5}\cdot 1\cdot\frac{4}{7} \;\cdot\; 1\cdot 1\cdot\frac{5}{7} + \frac{1}{2}\cdot\frac{2}{7}\cdot\frac{2}{7}\cdot\frac{4}{7} \;\cdot\; \frac{2}{7}\cdot\frac{2}{7}\cdot\frac{5}{7} \approx 0.171\cdot 0.714 + 0.023\cdot 0.0583 \approx 0.124$$

El error de la aproximación de Viterbi es aproximadamente: $0.124 - 0.1225 = 0.0015$

28. Considera un problema de clasificación en dos clases $U$ y $V$, con probabilidades a prioi $P(U) = 0.7, P(V) = 0.3$, representadas mediante modelos de Markov $\mathcal{M}_U, \mathcal{M}_V$, ambos definidos sobre el conjunto de primitivas $\{\mathtt{a}, \mathtt{b}\}$:

| $A_U$ | 1 | 2 | F |
|---|---|---|---|
| 1 | 0.6 | 0.4 | 0 |
| 2 | 0 | 0.5 | 0.5 |

| $B_U$ | a | b |
|---|---|---|
| 1 | 0.9 | 0.1 |
| 2 | 0.2 | 0.8 |

$\pi_{U\,1} = 0.8, \quad \pi_{U\,2} = 0.2$

| $A_V$ | 1 | 2 | F |
|---|---|---|---|
| 1 | 0.4 | 0.6 | 0 |
| 2 | 0 | 0.5 | 0.5 |

| $B_V$ | a | b |
|---|---|---|
| 1 | 0.9 | 0.1 |
| 2 | 0.2 | 0.8 |

$\pi_{V\,1} = 1, \quad \pi_{V\,2} = 0$

*a*) Calcula la aproximación por Viterbi a la probabilidad de que $\mathcal{M}_U$ genere la cadena "aab" y la correspondiente secuencia óptima de estados.

$\tilde{P}(\mathtt{aab} \mid \mathcal{M}_U) = 0.062208 \qquad \tilde{q}_1 = 1,\ \tilde{q}_2 = 1,\ \tilde{q}_3 = 2,\ \tilde{q}_4 = F \qquad$ (Ver respuesta al apartado b)

*b*) Calcula las verdaderas probabilidades de que $\mathcal{M}_U$ y $\mathcal{M}_V$ generen la cadena "aab".

$$\begin{aligned}
P(\mathtt{aab} \mid \mathcal{M}_U) &= P(\mathtt{aab}, q_1 = 1, q_2 = 1, q_3 = 2 \mid \mathcal{M}_U) \\
&+ P(\mathtt{aab}, q_1 = 1, q_2 = 2, q_3 = 2 \mid \mathcal{M}_U) \\
&+ P(\mathtt{aab}, q_1 = 2, q_2 = 2, q_3 = 2 \mid \mathcal{M}_U) \\
&= (0.8 \cdot 0.9)\ (0.6 \cdot 0.9)\ (0.4 \cdot 0.8)\ 0.5 \\
&+ (0.8 \cdot 0.9)\ (0.4 \cdot 0.2)\ (0.5 \cdot 0.8)\ 0.5 \\
&+ (0.2 \cdot 0.2)\ (0.5 \cdot 0.2)\ (0.5 \cdot 0.8)\ 0.5 \\
&= 0.062208 + 0.0115200 + 0.0008 \\
&= 0.074528
\end{aligned}$$

$$\begin{aligned}
P(\mathtt{aab} \mid \mathcal{M}_V) &= P(\mathtt{aab}, q_1 = 1, q_2 = 1, q_3 = 2 \mid \mathcal{M}_V) \\
&+ P(\mathtt{aab}, q_1 = 1, q_2 = 2, q_3 = 2 \mid \mathcal{M}_V) \\
&= (1 \cdot 0.9)\ (0.4 \cdot 0.9)\ (0.6 \cdot 0.8)\ 0.5 \\
&+ (1 \cdot 0.9)\ (0.6 \cdot 0.2)\ (0.5 \cdot 0.8)\ 0.5 \\
&= 0.07776 + 0.021600 \\
&= 0.09936
\end{aligned}$$

*c*) Determina las probabilidades a posteriori de que la cadena "aab" pertenezca a las clases $U$ y $V$.

$$P(\mathtt{aab} \mid U)\, P(U) = 0.074528\ \cdot\ 0.7 = 0.0521696$$
$$P(\mathtt{aab} \mid V)\, P(V) = 0.09936\ \cdot\ 0.3 = 0.029808$$

$$P(x) = P(\mathtt{aab} \mid \mathcal{M}_U)\, P(U) + P(\mathtt{aab} \mid \mathcal{M}_V)\, P(V) = 0.0521696 + 0.029808 \ = \ 0.0819776$$

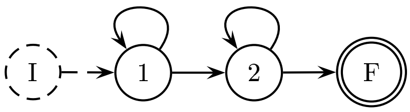
$$P(U \mid \mathtt{aab}) = \frac{P(\mathtt{aab} \mid U)\, P(U)}{P(x)} = \frac{0.0521696}{0.0819776} \ \approx\ 0.6364$$

$$P(V \mid \mathtt{aab}) = \frac{P(\mathtt{aab} \mid V)\, P(V)}{P(x)} = \frac{0.029808}{0.0819776} \ \approx\ 0.3636$$

*d*) Clasifica la cadena "aab" por mínima probabilidad de error.

$$c^*(\mathtt{aab}) = \underset{c \in \{U,V\}}{\arg\max}\, P(c \mid \mathtt{aab}) \ = \ U$$

29. (Examen de SIN del 18 de enero de 2013; tiempo estimado: 30 minutos)
Sea $M$ un modelo de Markov de conjunto de estados $Q = \{1, 2, F\}$; topología:




alfabeto $\Sigma = \{a, b\}$; y "aaabb" una muestra de entrenamiento. Obtener el correspondiente modelo oculto de Markov estimado inicializando con una segmentación lineal.

SOLUCIÓN:

Segmentación inicial

a a a b b
1 1 2 2 2 F

Inicialización a partir de la segmentación inicial

| $\pi$ | 1 | 2 |
|---|---|---|
| | 1 | 0 |

| A | 1 | 2 | F |
|---|---|---|---|
| 1 | 0.5 | 0.5 | 0 |
| 2 | 0 | 0.66 | 0.34 |

| B | a | b |
|---|---|---|
| 1 | 1 | 0 |
| 2 | 0.34 | 0.66 |

Primera iteración

Análisis por Viterbi

| | | a | a | a | b | b | |
|---|---|---|---|---|---|---|---|
| I | 1 | | | | | | |
| 1 | | 1 | 0.5 | 0.25 | 0.0 | 0.0 | |
| 2 | | 0 | 0.16 | 0.085 | 0.085 | 0.037 | |
| F | | | | | | | 0.012 |

Secuencia de estados óptima: 1 1 1 2 2

Restimación

| $\pi$ | 1 | 2 |
|---|---|---|
| | 1 | 0 |

| A | 1 | 2 | F |
|---|---|---|---|
| 1 | 0.66 | 0.34 | 0 |
| 2 | 0 | 0.5 | 0.5 |

| B | a | b |
|---|---|---|
| 1 | 1 | 0 |
| 2 | 0 | 1 |

Segunda iteración

Análisis por Viterbi

| | | a | a | a | b | b | |
|---|---|---|---|---|---|---|---|
| I | 1 | | | | | | |
| 1 | | 1 | 0.66 | 0.436 | 0.0 | 0.0 | |
| 2 | | 0 | 0 | 0 | 0.148 | 0.07 | |
| F | | | | | | | 0.035 |

Secuencia de estados óptima: 1 1 1 2 2

Restimación

| $\pi$ | 1 | 2 |
|---|---|---|
| | 1 | 0 |

| A | 1 | 2 | F |
|---|---|---|---|
| 1 | 0.66 | 0.34 | 0 |
| 2 | 0 | 0.5 | 0.5 |

| B | a | b |
|---|---|---|
| 1 | 1 | 0 |
| 2 | 0 | 1 |

Final porque el modelo no ha cambiado

30. (Examen de SIN del 30 de enero de 2013; tiempo estimado: 30 minutos)
Supongamos que disponemos de tres cadenas de símbolos de entrenamiento para estimar las probabilidades de un modelo de Markov y que en una iteración determinada del algoritmo de reestimación por Viterbi se obtienen las siguientes secuencias de estados óptimas:

a a a d d d c c c c a a a a a c c b
1 1 1 2 2 2 2 3 3 3 2 2 2 4 4 4 4 4 F

a a d d c a b a b a a b c c c d c b b
1 1 2 2 2 4 4 4 4 4 2 2 2 2 2 1 1 1 1 F

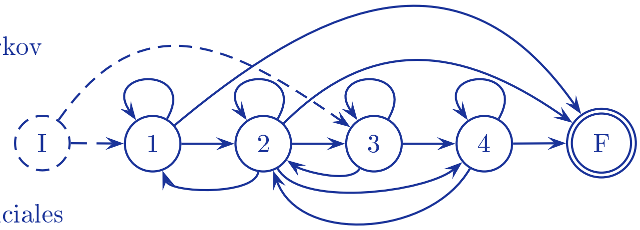
a a a a d c d c d a b a b a b c c c a a b
3 3 3 3 4 4 4 4 4 2 2 2 2 2 2 4 4 4 4 2 2 2 F

A partir de esas secuencias, se pide:

a) dibujar la topología (estados y transiciones con probabilidad distinta de cero) del modelo de Markov que se está utilizando;
b) calcular las probabilidades de ser estado inicial;
c) calcular las probabilidades de transición entre estados; y
d) calcular las probabilidades de emisión en los estados.

SOLUCIÓN:

a) Topología del modelo de Markov




b) Probabilidades de estados iniciales

| $\pi$ | 1 | 2 | 3 | 4 |
|---|---|---|---|---|
| | 2/3 | 0 | 1/3 | 0 |

c) Probabilidades de transición entre estados

| # | 1 | 2 | 3 | 4 | F | Total |
|---|---|---|---|---|---|---|
| 1 | 2+4+0=6 | 1+1+0=2 | | | 0+1+0=1 | 9 |
| 2 | 0+1+0=1 | 5+6+6=17 | 1+0+0=1 | 1+1+1=3 | 0+0+1=1 | 23 |
| 3 | | 1+0+0=1 | 2+0+3=5 | 0+0+1=1 | | 7 |
| 4 | | 0+1+2=3 | | 4+4+7=15 | 0+0+0=1 | 19 |

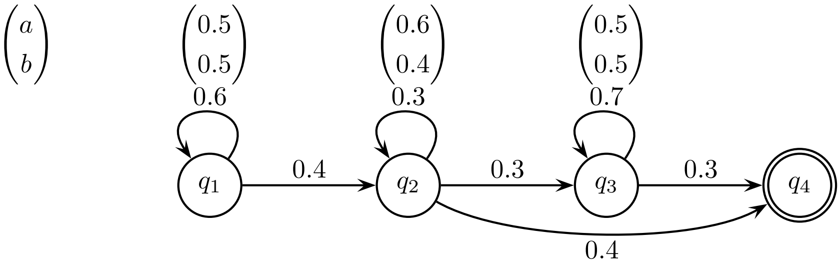
| A | 1 | 2 | 3 | 4 | F |
|---|---|---|---|---|---|
| 1 | 2/3 | 2/9 | 0 | 0 | 1/9 |
| 2 | 1/23 | 17/23 | 1/23 | 3/23 | 1/23 |
| 3 | 0 | 1/7 | 5/7 | 1/7 | 0 |
| 4 | 0 | 3/19 | 0 | 15/19 | 1/19 |

d) Probabilidades de emisión

| # | a | b | c | d | Total |
|---|---|---|---|---|---|
| 1 | 3+2+0=5 | 0+2+0=2 | 0+1+0=1 | 0+1+0=1 | 9 |
| 2 | 3+1+5=9 | 0+1+4=4 | 1+4+0=5 | 3+2+0=5 | 23 |
| 3 | 0+0+4=4 | 0+0+0=0 | 3+0+0=3 | 0+0+0=0 | 7 |
| 4 | 2+3+0=5 | 1+2+1=4 | 2+2+1=7 | 0+0+2=3 | 19 |

| B | a | b | c | d |
|---|---|---|---|---|
| 1 | 5/9 | 2/9 | 1/9 | 1/9 |
| 2 | 9/23 | 4/23 | 5/23 | 5/23 |
| 3 | 4/7 | 0 | 3/7 | 0 |
| 4 | 5/19 | 4/19 | 7/19 | 3/19 |

31. (Examen de SIN del 28 de enero de 2014; problema 1; tiempo estimado: 30 minutos)
Dado el siguiente modelo oculto de Markov M

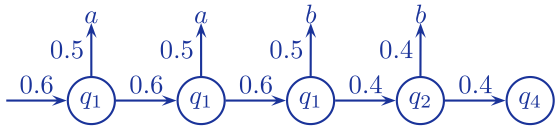



con $\pi_{q_1} = 0.6, \pi_{q_2} = 0.4, \pi_{q_3} = \pi_{q_4} = 0$ y la cadena $aabb$, se pide:

a) Obtén la secuencia de estados que con mayor probabilidad genera dicha cadena aplicando el algoritmo de Viterbi.

**Solución**

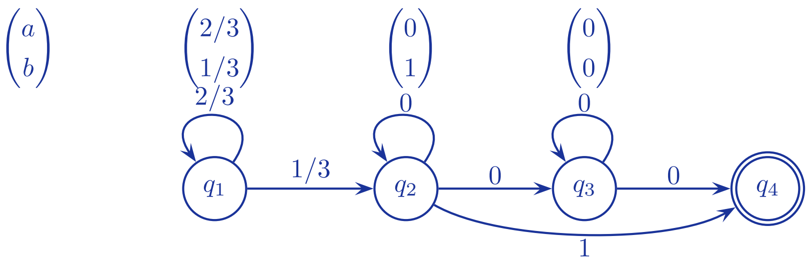
| | a | a | b | b | |
|---|---|---|---|---|---|
| $q_1$ | .30 | .090 | .0270 | .0081 | |
| $q_2$ | .24 | .072 | .0144 | .00432 | |
| $q_3$ | | .036 | .0126 | .00441 | |
| $q_4$ | | | | | .001728 |




b) Dibuja cómo quedaría el modelo de Markov y sus probabilidades después de estimarlo mediante una iteración con los resultados obtenidos en el apartado anterior.

**Solución**

La probabilidades iniciales quedarían como: $\pi_{q_1} = 1.0, \pi_{q_2} = \pi_{q_3} = \pi_{q_4} = 0$



32. (Examen de SIN del 28 de enero de 2014; problema 2; tiempo estimado: 30 minutos)
Tenemos un problema de clasificación de cadenas en dos clases equiprobables $c_0$ y $c_1$. Las cadenas son de tres símbolos $x_0x_1x_2$, tal que $x_0, x_1, x_2 \in \{a, b\}$. Dado el modelo oculto de Markov $M_0$ asociado a la clase $c_0$ que aparece a la izquierda y el modelo oculto de Markov $M_1$ asociado a la clase $c_1$ que aparece a la derecha:

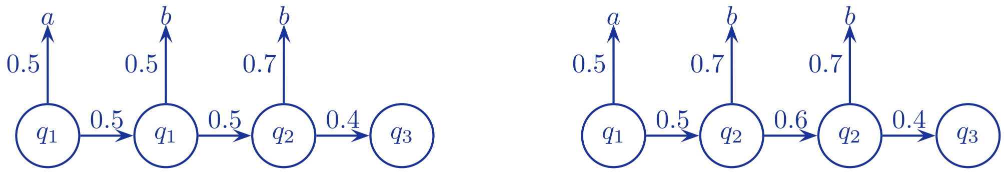

tal que en ambos modelos $\pi_{q_1} = 1, \pi_{q_2} = \pi_{q_3} = 0$, se pide:

a) Calcula la probabilidad de la cadena *abb* mediante el algoritmo *forward* con ambos modelos.
b) Indica en qué clase quedaría clasificada dicha cadena por máxima probabilidad a *posteriori*.

**Solución**

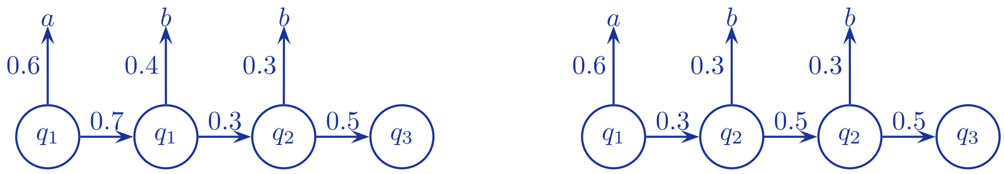
Para el primer modelo $M_0$ tenemos que:

| | a | b | b | |
|---|---|---|---|---|
| $q_1$ | .5 | .125 | .03125 | |
| $q_2$ | | .175 | .11725 | |
| $q_4$ | | | | .0469 |




con una probabilidad total $p(abb|M_0) = 0.0469$. Mientras que para el segundo modelo $M_1$ tenemos que:

| | a | b | b | |
|---|---|---|---|---|
| $q_1$ | .6 | .168 | .04704 | |
| $q_2$ | | .054 | .02322 | |
| $q_4$ | | | | .01161 |




con una probabilidad total $p(abb|M_1) = 0.0116$. Por lo que la cadena quedaría clasificada en la clase $c_0$.