Dpto.de Sistemas Informáticos y Computación

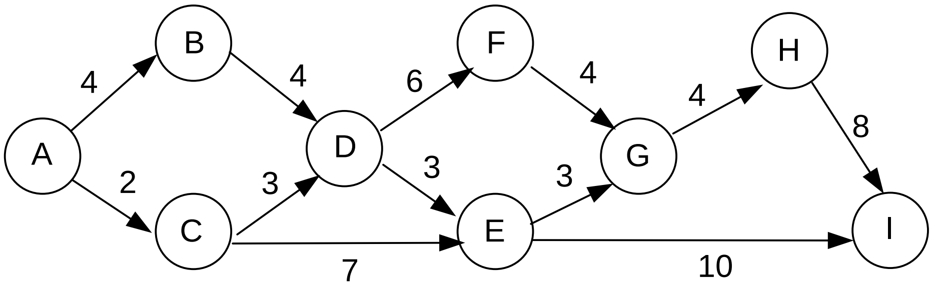
SISTEMAS INTELIGENTES (GII)
EXAMEN 28-01-2014
ETSINF

NOMBRE: **GRUPO:**

APELLIDOS:

Pregunta 1 (2 puntos, Tiempo Estimado 35’).

Sea el siguiente grafo donde cada arco indica su coste y la tabla indica la estimación del coste ‘h’ hasta la solución. El nodo ‘A’ es el estado inicial y el nodo ‘I’ es el estado final.



n	A	B	C	D	E	F	G	H
h(n)	15	14	13	12	11	10	9	8

1) Asumiendo que se aplica un algoritmo en anchura, que ante el mismo valor de la función ‘f(n)’ se expande antes el nodo alfabéticamente anterior y que se realiza control de nodos repetidos (descartar los nodos más profundos o nodos expandidos con anterioridad en caso del mismo nivel de profundidad), contesta a las siguientes preguntas:

   a) Escribe los nodos del camino solución desde el nodo A hasta el nodo I.
   A C E I

   b) ¿Cuántos nodos se han generado en total y cuántos nodos se han expandido en el árbol?
   12 nodos generados (3 de ellos repetidos), y 8 nodos expandidos (incluido el nodo I)

2) Asumiendo que se aplica la versión grafo de un algoritmo A con control de nodos repetidos, contesta a las siguientes preguntas:

   a) Escribe los nodos del camino solución desde el nodo A hasta el nodo I y el coste de dicho camino solución; el camino encontrado, ¿es la solución óptima?

   A C D E I; el coste del camino solución es 18. Sí, es la solución óptima porque no existe un camino de menor coste entre el nodo A y el nodo I

b) ¿Cuántos nodos se han generado en total y cuántos nodos se han expandido en el árbol?
10 nodos generados (2 de ellos repetidos), 6 nodos expandidos (incluido el nodo I)

c) Indica los nodos expandidos y su orden de expansión.
A C D B E I

3) Responde brevemente a las siguientes preguntas justificando las respuestas:

a) La función heurística de este problema, ¿es admisible?¿por qué?
No, porque h(E)=11 y h*(E)=10

b) La función heurística de este problema, ¿es consistente?¿por qué?
No, porque no se cumple h(E) <= h(I)+c(E,I), o sea, no se cumple 11 <= 0 + 10

c) Si aplicamos un algoritmo de profundización iterativa en este problema, ¿qué solución encontraría? Indica los nodos del camino solución así como el número total de nodos generados.

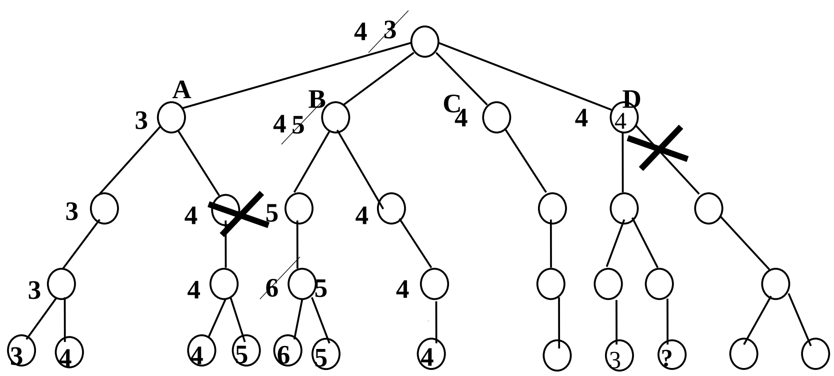
Encontraría la misma solución que anchura, A C E I.

Se generan 4 árboles en profundidad, desde el nivel 0 al nivel 3. El número de nodos generados sería, contabilizando los nodos por árbol generado: 1+3+6+12=22; o también, contabilizando el número de veces que se generan los nodos de cada nivel: 4 veces * 1 (el nodo de nivel 0) + 3 veces * 2 (dos nodos en nivel 1) + 2 veces * 3 (tres nodos en nivel 2) + 1 vez *6 (cuatro nodos de nivel 3) = 22.

<u>Pregunta 2 (1 punto, Tiempo Estimado 15’)</u>.

El siguiente árbol representa el resultado parcial de aplicar el algoritmo alfa-beta de izquierda a derecha al espacio de búsqueda de un juego dado. En el árbol se han eliminado los valores volcados de algunos de los nodos así como algunas de las podas. Determinar, razonando cada una de las respuestas:

a) Calcular en las dos ramas de la izquierda (A y B) los valores volcados aplicando alfa-beta.



b) Teniendo en cuenta la resolución del apartado a). ¿Qué rama (A, B, C o D) será la mejor jugada para el nodo raíz?. ¿por qué?

Se produce una meseta entre las ramas B, C y D, y por tanto cualquiera de ellas podría ser la mejor jugada. Sin embargo, si se ha producido alguna poda por debajo de una rama, los valores volcados podrían ser mayores que los obtenidos en caso de no producirse dichas podas.

La rama B será la primera de ellas en ser expandida, y no se producen podas, por tanto el valor volcado obtenido será el mismo que en el caso de utilizar MINIMAX y ésta será la mejor jugada en el caso de utilizar alfa-beta, es decir 'la rama de mayor valor volcado que haya sido expandida en primer lugar'. Por debajo del nodo C tampoco se producen podas, por lo que la rama C también podrá ser considerada mejor jugada. En la rama D no podemos garantizar que el valor volcado no sea menor que en caso de desarrollar completamente la búsqueda por debajo de la misma.

c) Teniendo en cuenta el corte que se produce por debajo del nodo de nivel 1 de la rama D, ¿qué valor deberá tener el nodo hoja marcado con **'?'**?

El valor será 4, si aplicamos el algoritmo alfa-beta, tal como se indica en la figura, podemos comprobar que este es el único valor posible.

<u>Pregunta 3 (2 puntos, Tiempo Estimado 40')</u>.

Supóngase el siguiente problema: "Un granjero quiere cruzar un río llevando consigo a una zorra, un ganso y un saco de trigo. Por desgracia, su bote es tan pequeño que solo puede transportar una de sus pertenencias en cada viaje. Además, la zorra, si no se le vigila, se come al ganso, y el ganso, si no se le cuida, se come el trigo. Así, el granjero no debe dejar a la zorra sola con el ganso o al ganso solo con el trigo."

Se desea diseñar un SBR que, a partir de un estado inicial en el que están todos en el lado A del río, determine la secuencia de acciones para que pasen todos al lado B. Se pide:

a) Asumiendo la siguiente especificación de la Base de Hechos:

(problema Granjero x^S Zorra x^S Ganso x^S Trigo x^S) $x^S \in \{A, B\}$

NOTA: En la parte derecha de las reglas utilizad solo expresiones assert

a.1) Especificar en CLIPS la regla para pasar al granjero solo, del lado-A al lado-B.

```
(defrule GranjeroAB  ; pasar hombre lado-B, no dejar ganso sólo con zorra o con trigo
    ?f  <-  (problema Granjero  A   Zorra  ?z   Ganso   ?g    Trigo ?t)
            (test   (or (eq ?g B)   (and   (eq ?t B) (eq ?z B))))
    =>
 (assert   (problema Granjero   B   Zorra   ?z   Ganso   ?g  Trigo   ?t)))
```

a.2) Especificar en CLIPS la regla para pasar a la zorra, del lado-A al lado-B.

```
(defrule ZorraAB   ;pasar zorra a Lado-B, no dejar ganso sólo con el trigo
      ?f  <- (problema Granjero  A   Zorra  A   Ganso   ?g    Trigo ?t)
             (test   (or (eq  ?g  B)  (eq  ?t  B)))
                                              ; tambien, (not (and (eq  ?g  A )  (eq  ?t  A)))
         =>
             (assert   (problema Granjero   B   Zorra  B    Ganso  ?g  Trigo  ?t)))
```

b) Supongamos que el río es tan ancho que no puede cruzarse sin hacer paradas sucesivas en islotes {I1, I2, …, In} que hay al medio del río,

b.1) sin modificar el patrón ya indicado en (b), añade la información necesaria en la Base de Hechos, para cubrir esta nueva información, tal que las reglas puedan ser independientes del origen/destino del movimiento que representan.

La BH quedaría:

(problema Granjero x^S Zorra x^S Ganso x^S Trigo x^S) $x^S \in$ {A, B, I1, I2, …, In}

Y un conjunto de hechos indicando las conexiones que hayan entre las orillas del río e islotes:

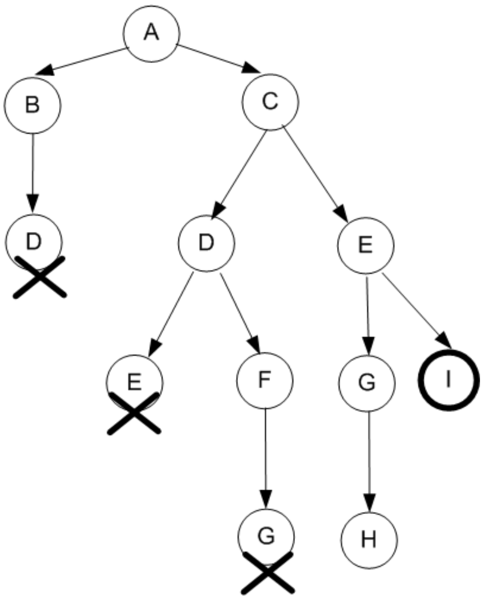
(conectado A I1)	(conectado I1 A)
(conectado I1 I2)	(conectado I2 I1)
(conectado I2 I3)	(conectado I3 I2)
....	....
(conectado I_{n-1} I_n)	(conectado I_n I_{n-1})
(conectado I_n B)	(conectado B I_n)

b.2) con la modificación anterior, modificad la regla (b.2) para pasar a la zorra, de una parte cualquiera a otra, sea lado del río o islote.

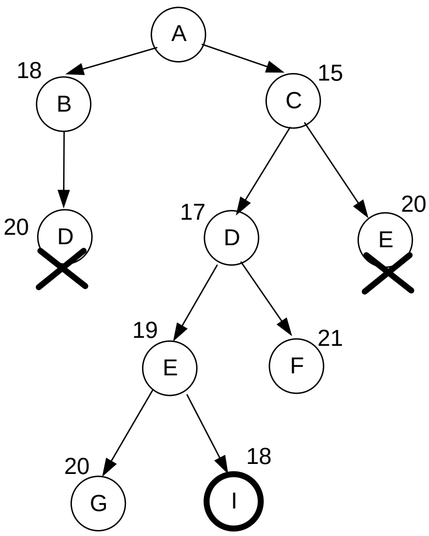
```
(defrule Zorra   ; pasar zorra de origen a destino, no dejar ganso sólo con el trigo"
      ?f  <- (problema  Granjero  ?origen  Zorra  ?origen  Ganso  ?g   Trigo ?t)
             (test  (not (and  (eq  ?g  ?origen )  (eq  ?t  ?origen))))
             (conectado  ?origen  ?destino)
         =>
             (assert  (problema  Granjero  ?destino  Zorra  ?destino  Ganso  ?g  Trigo  ?t)))
```

ANEXO

Árbol anchura



Árbol A



Examen Final de Sistemas Inteligentes: Bloque 2

ETSINF, Universitat Politècnica de València, 28 de enero de 2014

Apellidos: ______________________ **Nombre:** ____________

Grupo: ☐ **3A** ☐ **3B** ☐ **3C** ☐ **3D** ☐ **3E** ☐ **RE1** ☐ **RE2**

Cuestiones (2.5 puntos; tiempo estimado: 30 minutos)

Marca cada recuadro con una única opción de entre las dadas.

1 $\boxed{\text{D}}$ En un experimento de clasificación con 300 datos de *test* se han observado 15 errores. Con una confianza del 95 %, podemos afirmar que la verdadera probabilidad de error es:

A) $P(\text{error}) = 5\,\% \pm 0.3\,\%$
B) $P(\text{error}) = 0.05 \pm 0.3$
C) $P(\text{error}) = 0.05$, exactamente
D) $P(\text{error}) = 0.05 \pm 0.03$ $\quad 0.05 \pm 1.96\sqrt{\frac{0.05\,0.95}{300}} = 0.05 \pm 0.03 \quad (5\,\% \pm 3\,\%)$

2 $\boxed{\text{B}}$ En un problema de diagnóstico diferencial entre *Gripe* y *Resfriado*, se sabe que la incidencia relativa de la *Gripe* con respecto al *Resfriado* es del 30 % y se conocen las siguientes distribuciones de temperaturas corporales:

$t(^oC)$	36	37	38	39	40
$P(T=t \mid D=\text{GRIPE})$	0.05	0.10	0.20	0.30	0.35
$P(T=t \mid D=\text{RESFR})$	0.10	0.30	0.40	0.15	0.05

$$P(\text{GRIPE} \mid 37) = \frac{\frac{30}{130}\,0.10}{\frac{30}{130}\,0.10 + \frac{100}{130}\,0.30} = \frac{1}{11}$$

El diagnóstico de minimo riesgo de error para un paciente con 37º de fiebre es:

A) *Gripe*
B) *Resfriado*
C) Hay un empate entre ambos diagnósticos
D) Las probabilidades dadas son incorrectas ya que no suman 1; por tanto no es posible hacer un diagnóstico.

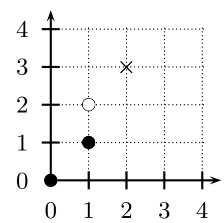
3 $\boxed{\text{B}}$ Sea un problema de clasificación en 2 clases, $c = A, B$, para objetos representados mediante vectores de características bidimensionales. Como resultado de la aplicación del algoritmo Perceptrón sobre un conjunto de entrenamiento, se han obtenido los vectores de pesos $\mathbf{w}_A = (1,1,0)^t$ y $\mathbf{w}_B = (-1,0,1)^t$. ¿En qué clases se clasifican $\mathbf{x}_1 = (-1,0)^t$ y $\mathbf{x}_2 = (0,3)^t$?

A) $\hat{c}(\mathbf{x}_1) = A$ y $\hat{c}(\mathbf{x}_2) = A$. $\quad \mathbf{x}_1 : \mathbf{w}_A^t \cdot (1,-1,0)^t = 0 \quad \mathbf{w}_B^t \cdot (1,-1,0)^t = -1 \Rightarrow \mathbf{x}_1 \in A$
B) $\hat{c}(\mathbf{x}_1) = A$ y $\hat{c}(\mathbf{x}_2) = B$.
C) $\hat{c}(\mathbf{x}_1) = B$ y $\hat{c}(\mathbf{x}_2) = A$. $\quad \mathbf{x}_2 : \mathbf{w}_A^t \cdot (1,0,3)^t = 1 \quad \mathbf{w}_B^t \cdot (1,0,3)^t = 2 \Rightarrow \mathbf{x}_2 \in B$
D) $\hat{c}(\mathbf{x}_1) = B$ y $\hat{c}(\mathbf{x}_2) = B$.

4 $\boxed{\text{B}}$ Sea un problema de clasificación en 2 clases, $c = 1, 2$, para objetos representados mediante vectores de características reales bidimensionales; esto es, de la forma $\mathbf{y} = (y_1, y_2) \in \mathbb{R}^2$. Sea T un árbol de clasificación para este problema y sea t un nodo interno de T. Sean B_1 y B_2 las cajas de mínima inclusión de los objetos de la clase 1 y 2 en t, respectivamente. Dichas cajas están caracterizadas por las coordenadas de sus esquinas inferior izquierda y superior derecha de la forma $[\min y_1, \min y_2] \times [\max y_1, \max y_2]$, siendo $B_1 = [1.5, 0.6] \times [2.3, 3.5]$ y $B_2 = [2.5, 1.3] \times [3.8, 3.2]$. En términos de decremento de impureza (medida como entropía), ¿cuál de las siguientes particiones de t es mejor?

A) $y_1 \leq 3.8$
B) $y_1 \leq 2.3$
C) $y_2 \leq 1.3$
D) $y_2 \leq 3.5$

5 $\boxed{\text{C}}$ La figura a la derecha muestra una partición de 4 puntos bidimensionales en 3 clústers (representados mediante los símbolos $\bullet$, $\circ$ y $\times$). La suma de errores cuadráticos de esta partición es $J = 1$. Si se ejecuta el algoritmo C-medias (de Duda y Hart) a partir de la misma:



A) No se realizará niguna transferencia de clúster.
B) Se transferirá un único punto, obteniéndose una partición de J entre $\frac{2}{3}$ y 1.
C) Se transferirá un único punto, obteniéndose una partición de J entre 0 y $\frac{2}{3}$. $\quad$ J=0.5
D) Se realizarán dos transferencias de clúster, obteniéndose una partición de J nula.

6 $\boxed{\text{B}}$ Dado un Modelo Oculto de Markov Θ 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 $\widetilde{P}(y|\Theta)$.

Examen Final de Sistemas Inteligentes: Bloque 2

ETSINF, Universitat Politècnica de València, 28 de enero de 2014

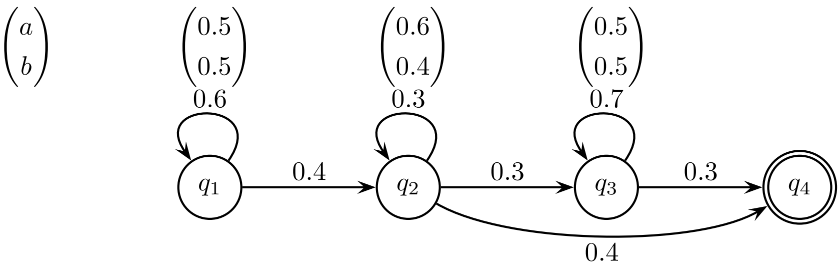
Apellidos: ______________________ **Nombre:** ______________

Grupo: ☐ **3A** ☐ **3B** ☐ **3C** ☐ **3D** ☐ **3E** ☐ **RE1** ☐ **RE2**

Problemas (2.5 puntos; tiempo estimado: 60 minutos)

1. **(1 punto)**

   Dado el siguiente modelo oculto de Markov M

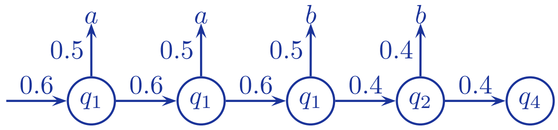


   con $\pi_{q_1} = 0.6, \pi_{q_2} = 0.4, \pi_{q_3} = \pi_{q_4} = 0$ y la cadena $aabb$, se pide:

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

   Solución

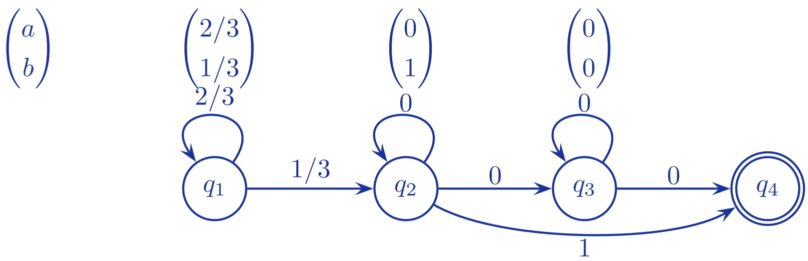
	a	a	b	b	
q_1	.30	.090	.0270	.0081	
q_2	.24	.072	.0144	.00432	
q_3		.036	.0126	.00441	
q_4					.001728



   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: $\pi_{q_1} = 1.0, \pi_{q_2} = \pi_{q_3} = \pi_{q_4} = 0$



2. **(1.5 puntos)**

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:

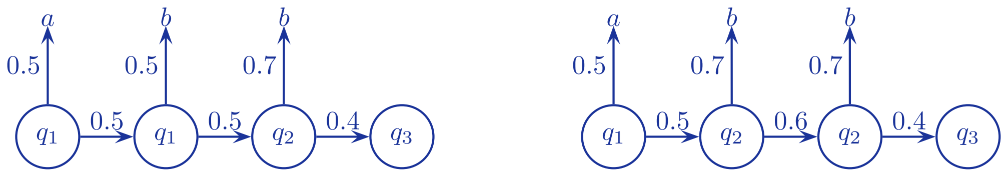
a) Calcula la probabilidad de la cadena *abb* mediante el algoritmo *forward* con ambos modelos.

b) Indica en qué clase quedaría clasificada dicha cadena por máxima probabilidad a *posteriori*.

Solución

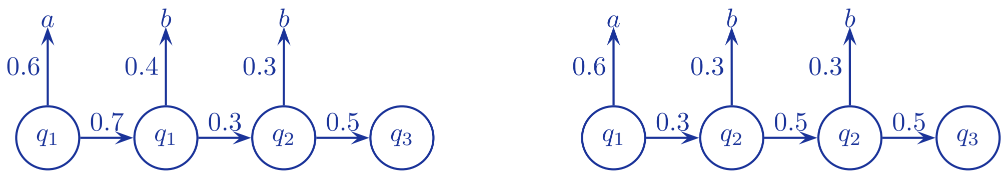
Para el primer modelo M_0 tenemos que:

	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.