APELLIDOS: PL:
NOMBRE: DNI:

ESCUELA DE INGENIERÍA INFORMÁTICA
SISTEMAS INTELIGENTES
Examen Final. Lunes 17 de enero de 2022.

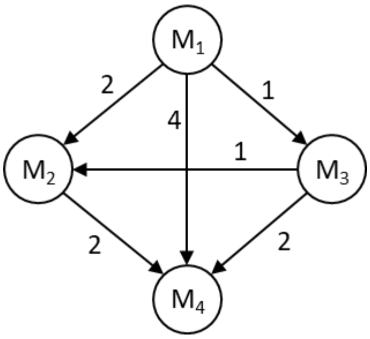
Primer Parcial

Ejercicio 1.

a) **[0,25 puntos]** Dar la definición de heurístico monótono e indicar qué consecuencias tiene su uso en el algoritmo A*.

b) **[0,25 puntos]** Si un heurístico h es monótono, ¿lo será también el heurístico h' definido como h'(n) = h(n)/2 para todo n?

Consideremos el siguiente espacio de búsqueda abstracto, siendo el estado inicial M_1 y M_4 el único objetivo, y los heurísticos cuyos valores se indican por extensión en la tabla



	h1	h2	h3	h*
M1	2	3	1,5	
M2	1	0,5	1	
M3	4	2	1	
M4	0	0	0	

Se pide responder de forma razonada a las siguientes cuestiones

c) **[0,25 puntos]** ¿Cuál es el valor de C*, y el de h* para cada uno de los estados?

d) **[0,75 puntos]** ¿Qué propiedades tienen los heurísticos h1, h2 y h3? ¿Se puede establecer algún tipo de comparación entre ellos?

e) **[1 punto]** De acuerdo con las propiedades anteriores, ¿Qué se puede decir con respecto a la expansión del nodo M_2 con cada uno de los tres heurísticos?

f) **[1 punto]** ¿Qué solución se encontrará A* con cada uno de los heurísticos h1, h2, h3 y h*?

Solución Ejercicio 1

a) Un heurístico es monótono si, para todo par de estados n1, n2 conectados por un arco en el espacio de búsqueda cumple h(n1)<=h(n2)+c(n1,n2), siendo c(n1,n2) el coste del arco. Si un heurístico es monótono también es consistente (son propiedades equivalentes) y también es admisible. Además, cuando A* expande un estado n se cumplirá que g(n) = g*(n), es decir ya se conoce el camino óptimo desde el inicial a n. Otra consecuencia es que la secuencia de valores de f() de los nodos expandidos es no decreciente.

b) Si h es monótono entonces para todo par de nodos n1,n2 h(n1)<=h(n2)+c(n1,n2). Luego h(n1)/2 <= h(n2)/2 + c(n1,n2)/2 < h(n2)/2 + c(n1,n2). Luego h' también es monótono.

c) Los valores de h* son 3 para M1, 2 para M2 y M3, y 0 para M4. C*=h*(M1) = 3.

d) Los tres heurísticos están bien definidos; h1 no es admisible ya que 4=h1(M_3)>h*(M_3)=3; h2 es admisible pero no monótono ya que h2(M3) > h2(M2) + c(M3,M2); h3 es monótono ya que h3(n) = h*(n)/2 para todo n (ver la cuestión a)) y por lo tanto admisible. En cuanto a la comparación, h1 no es comparable a ninguno ya que no es admisible, y h2 y h3 siendo los dos admisibles tampoco son comparables ya que h2(M1)>h3(M1) y h2(M2)<h3(M2).

e) Sobre la expansión de M2, con h1 no podemos decir nada a priori ya que no es admisible. Con h2, tenemos que considerar todos los caminos desde el inicial a M2 y comprobar si alguno de ellos es C*-acotado o estrictamente C*-acotado. Podemos ver que los caminos M1 -> M2, y M1 -> M3 -> M2 son C*-acotados (no estrictamente). En el caso del primero tenemos que gP(M1)+h2(M1) = 0 + 3 <= C*=3 y gP(M2)+h2(M2) = 2 + 0,5 < C*=3 luego cumple la condición necesaria de expansión pero no la suficiente, con lo cual puede expandirse o no.

En el caso de h3, que es monótono, podemos ver que g*(M2)+h3(M2) = 2 + 1 <= C* = 3, luego también cumple la condición necesaria pero no la suficiente.

f) Con h2, h3 y h* se encontrará la solución óptima ya que son admisibles, es decir el camino M1 -> M3 -> M4 de coste C*=3. Para saber la solución que se encuentra con h1 es necesario hacer una traza del algoritmo A*. Al expandir el estado inicial, en frontier tendremos un nodo con el estado M3 con f(M3)=5 y otro M4 con f(M4)=4. Después, no se introducirá ningún otro nodo con M3, con lo que se expandirá antes un nodo con M4 que el nodo con M3 con lo que la solución encontrada será una de estas dos: M1 -> M4 o bien M1 -> M2 -> M4, es decir una solución no óptima con coste 4 > C*=3.

Ejercicio 2. Considera la siguiente base de conocimiento:

Identificador	Regla
R1	$h8, h6, h5 \to h4$
R2	$h6, h3 \to h8$
R3	$h7, h4 \to h3$
R4	$h8 \to h1$
R5	$h6 \to h5$
R6	$h4, h1 \to h2$
R7	$h7 \to h6$
R8	$h1, h7 \to h2$
R9	$h1, h8 \to h6$

a) **[0,75 puntos]** Aplica el algoritmo de encadenamiento hacia adelante para demostrar $h2$ cuando inicialmente $BH = \{h7, h3\}$. Para la resolución de conflictos escoge la regla con menor identificador.

b) **[0,75 puntos]** Aplica el algoritmo de encadenamiento hacia atrás para demostrar $h4$ cuando inicialmente $BH = \{h8, h2\}$. Para la resolución de conflictos escoge la regla con menor identificador.

Solución Ejercicio 2

a)
- Iteración 1 $CC = \{R7\}$, se selecciona $R7$ y se añade $h6$ a BH con lo que $BH_1 = \{h7, h3, h6\}$.
- Iteración 2 $CC = \{R2, R5\}$, se selecciona $R2$ y se añade $h8$ a BH con lo que $BH_2 = \{h7, h3, h6, h8\}$.
- Iteración 3 $CC = \{R4, R5\}$, se selecciona $R4$ y se añade $h1$ a BH con lo que $BH_3 = \{h7, h3, h6, h8, h1\}$.
- Iteración 4 $CC = \{R5, R8, R9\}$, se selecciona $R5$ y se añade $h5$ a BH con lo que $BH_4 = \{h7, h3, h6, h8, h1, h5\}$.
- Iteración 5 $CC = \{R1, R8, R9\}$, se selecciona $R1$ y se añade $h4$ a BH con lo que $BH_5 = \{h7, h3, h6, h8, h1, h5, h4\}$.
- Iteración 6 $CC = \{R3, R6, R8, R9\}$, se selecciona $R3$ pero $h3$ ya está en BH con lo que $BH_6 = \{h7, h3, h6, h8, h1, h5, h4\}$.
- Iteración 7 $CC = \{R6, R8, R9\}$, se selecciona $R6$ y se añade $h2$ a BH con lo que $BH_7 = \{h7, h3, h6, h8, h1, h5, h4, h2\}$ y la meta queda demostrada.

b) $h4 \notin BH$, $CC = \{R1\}$, se escoge $R1$ y como sus antecedentes $h6, h5 \notin BC$, hay que verificarlos.
  Submeta $h6$: $CC = \{R7, R9\}$, se escoge $R7$.
    Regla $R7$, como su antecedente $h7 \notin BH$, hay que verificarlo.
      Submeta $h7$: $CC = \emptyset$, no se puede verificar.
    Regla $R9$, como $h1 \notin BH$ hay que verificarlo.
      Submeta $h1$: $CC = \{R4\}$, se escoge $R4$
        Regla $R4$, $h8 \in BH$ entonces la regla se puede aplicar y $BH = \{h8, h2, h1\}$
    La regla $R9$ se puede aplicar y $BH = \{h8, h2, h1, h6\}$
  Submeta $h5$: $CC = \{R5\}$, se escoge esta regla
    Regla $R5$, $h6 \in BH$ entonces se puede aplicar y $BH = \{h8, h2, h1, h6, h5\}$
La regla $R1$ se puede aplicar y $BH = \{h8, h2, h1, h6, h5, h4\}$

Segundo Parcial

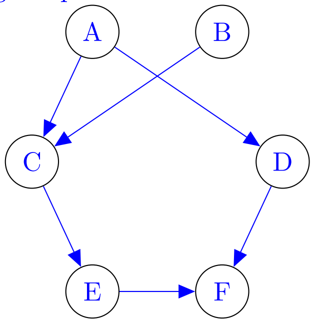
Ejercicio 3. [0.75 puntos] Considera el caso en que la probabilidad conjunta de las variables $\{A, B, C, D, E, F\}$ se puede escribir de la siguiente forma:

$$P(A, B, C, D, E, F) = P(A) \cdot P(B) \cdot P(C|A, B) \cdot P(D|A) \cdot P(E|C) \cdot P(F|D, E)$$

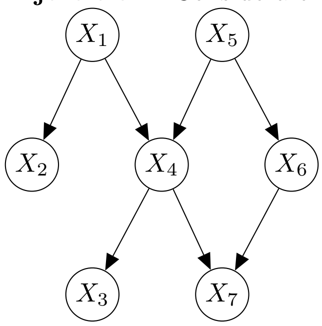
Dibuja el grafo de la red bayesiana que modela la relación entre estas 6 variables.

Solución Ejercicio 3

En una red bayesiana, aplicando la factorización, se puede escribir la probabilidad conjunta como el producto de las probabilidades condicionales de cada variable dados sus padres. Por tanto, en este caso el grafo para esta factorización es el siguiente:



Ejercicio 4. Considera el siguiente grafo de una red bayesiana



a) **[0.25 puntos]** Usando el criterio de D-separación, indica un par de variables que sean marginalmente dependientes, pero se necesite conocer exactamente el valor de otras dos para que sean independientes.

b) **[0.25 puntos]** Indica el manto de Markov para la variable X_4.

c) **[0.25 puntos]** Indica un orden topológico para este grafo.

Solución Ejercicio 4

a) En esta red tenemos que las variables X_5 y X_7 son dependientes cuando no hay ninguna evidencia ya que hay dos caminos con una cadena causal. Sin embargo, cuando se conocen precisamente X_4 y X_6 que son los nodos intermedios de cada uno de los dos caminos, estos se bloquean.

b) El manto de Markov de X_4 es $\{X_1, X_5, X_3, X_7, X_6\}$.

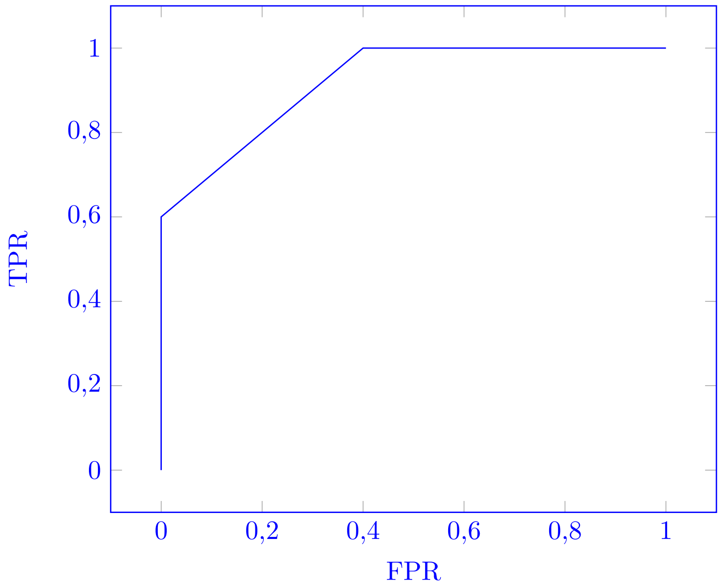
c) Un posible orden topológico es $\{X_1, X_5, X_2, X_4, X_3, X_6, X_7\}$.

Ejercicio 5. **[0.75 puntos]** Dibuja la curva ROC para el clasificador del cual se han obtenido los siguientes resultados:

Real	pos	neg	pos	pos	neg	neg	neg	pos	neg	pos
Score(pos)	0.6	0.1	0.4	0.8	0.6	0.3	0.4	0.7	0.2	0.7

Solución Ejercicio 5

Coordenadas FPR=(0.0,0.0,0.0,0.2,0.4,0.6,0.8,1.0); TPR=(0.0,0.2,0.6,0.8,1.0,1.0,1.0,1.0)



Ejercicio 6. [0.75 puntos] Para un problema de clasificación en el que la variable de interés es Y, que tiene los valores "sí" y "no", contamos con un conjunto de entrenamiento con 100 instancias y las variables A y B. La siguiente tabla muestra la distribución de valores de cada una de estas variables en función de la variable de interés.

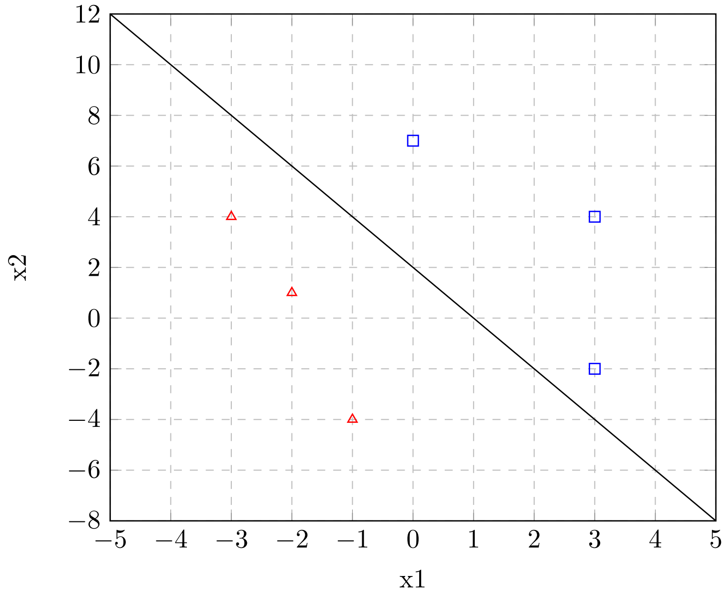
	Y=sí	Y=no
A	a1:20 a2:40	a1:**x** a2:**y**
B	b1:40 b2:20	b1:5 b2:35

Razona qué valores dar a **x** e **y**, es decir cuantas instancias para A=a1 y A=a2 cuando Y=no, de forma que la variable A sea elegida por delante de B como nodo de un árbol de decisión.

Solución Ejercicio 6

Se ha de conseguir que A separe mejor que B los casos para Y=sí e Y=no. Como para Y=sí ambas variables tienen una distribución simétrica, todo se reduce a poner valores para x e y de forma que sean más extremos que los que se consiguen con B para Y=no. Por ejemplo, con x=40 e y=0, aunque sería suficiente con x=36 e y=4

Ejercicio 7. [1 punto] Calcula los valores de los pesos w_0, w_1, w_2 para el perceptrón cuya frontera de decisión se ilustra en la siguiente figura.



Solución Ejercicio 7

Buscamos la recta que muestra la figura y esta es $x2 = -2x1 + 2$. Dado que el valor umbral del perceptrón es 0, dejamos la ecuación de la recta con un 0 a un lado del igual, por ejemplo $2x1 + x2 - 2 = 0$ y de aquí tenemos los valores de los pesos asociados a cada variable: $w_0 = -2$; $w_1 = 2$; $w_2 = 1$. Otra alternativa con $-2x1 - x2 + 2$: $w_0 = 2$; $w_1 = -2$; $w_2 = -1$

Ejercicio 8. [1 punto] Considera un modelo SVM lineal con margen duro cuyos parámetros son

$w = \begin{pmatrix} -0{,}8 \\ -0{,}4 \end{pmatrix}$ y $b = 6{,}2$. Este modelo se ha obtenido del conjunto de entrenamiento siguiente. Razona cuáles son los vectores soporte y a qué clase pertenecen.

x1	3	6	8	8	3	5	9	5
x2	7	6	5	5	3	3	6	1
y	1	-1	-1	-1	1	1	-1	1

Solución Ejercicio 8

Con un margen duro, las instancias que son vectores soporte deben dar como resultado en la ecuación del hiperplano 1 ó -1. De los casos anteriores cumplen esta condición los puntos 1, 2 y 6.

Para el punto 1: $\begin{pmatrix} 3 & 7 \end{pmatrix} \cdot \begin{pmatrix} -0{,}8 \\ -0{,}4 \end{pmatrix} + 6{,}2 = 1$

Para el punto 2: $\begin{pmatrix} 6 & 6 \end{pmatrix} \cdot \begin{pmatrix} -0{,}8 \\ -0{,}4 \end{pmatrix} + 6{,}2 = -1$

Para el punto 6: $\begin{pmatrix} 5 & 3 \end{pmatrix} \cdot \begin{pmatrix} -0{,}8 \\ -0{,}4 \end{pmatrix} + 6{,}2 = 1$

Por tanto hay dos vectores soporte de la clase positiva, las instancias 1 y 6, y un vector soporte de la clase negativa, la instancia 2.

$$P(x_1,\ldots,x_i) = \sum_{x_{i+1},\ldots,x_n} P(x_1,\ldots,x_i,x_{i+1},\ldots,x_n) = \sum_{x_{i+1}} P(x_1,\ldots,x_i,x_{i+1})$$

$$P(x|y) \equiv P(X=x|Y=y) = \frac{P(x,y)}{P(y)} \quad // \quad P(X,Y) = P(X|Y)P(Y) = P(Y|X)P(X)$$

$$P(Y) = \sum_{i=1}^{m} P(Y|x_i)P(x_i) \quad // \quad P(A|MB(A),B) = P(A|MB(A))$$

$$P(X_1=x_1,\ldots,X_n=x_n) = \prod_{i=1}^{n} P(X_i=x_i|X_{i+1}=x_{i+1},\ldots,X_n=x_n)$$

$$P(x_1,\ldots,x_n) = \prod_{i=1}^{n} P(X_i=x_i|Padres(X_i))$$

$$P(X|Padres(X)) = P(X|Padres(X),NoDescendientes(X))$$

$$H(X) = -\sum_{i=1}^{k} p_i \log_2 p_i \quad // \quad H(X|Y) = \sum_{y\in Y} p(y)\,H(X|Y=y) \quad // \quad IG(X,Y) = H(X) - H(X|Y)$$

$$GR(X,Y) = \frac{IG(X,Y)}{H(Y)} \quad // \quad Gini(X) = 1 - \sum_{x\in X} p(x)^2 \quad // \quad Gini(X,Y) = Gini(X) - Gini(X|Y)$$

$$dist(\mathbf{x},\mathbf{y}) = \left(\sum_i |x_i - y_i|^p\right)^{1/p} \quad // \quad g(\cdot) = \begin{cases} 1 \text{ si } wx > 0 \\ -1 \text{ en otro caso} \end{cases} \quad // \quad E(w) = \frac{1}{2}\sum_d \sum_k (y_{kd} - o_{kd})^2$$

$$\Delta w_i = \eta(y_d - wx_d)x_{di} \quad // \quad \delta_k = o_k(1-o_k)(y_k - o_k) \quad // \quad \delta_h = o_h(1-o_h)\textstyle\sum_k w_{kh}\delta_k$$

$$E(w) = \tfrac{1}{2}(y_d - wx_d)^2 \quad // \quad d = \tfrac{2}{\|w\|} \quad // \quad K(x,y) = (xy+1)^d \quad // \quad g(x) = \tfrac{1}{1+e^{-x}}$$

$$b = \frac{1}{|VS|}\sum_{i\in VS} y_i - wx_i \quad // \quad w = \sum_{i\in VS} \alpha_i y_i x_i \quad // \quad g(\mathbf{x}) = \sum_{i\in VS} \alpha_i y_i K(\mathbf{x_i},\mathbf{x}) + b$$