



APELLIDOS:

PL:

NOMBRE:

DNI:

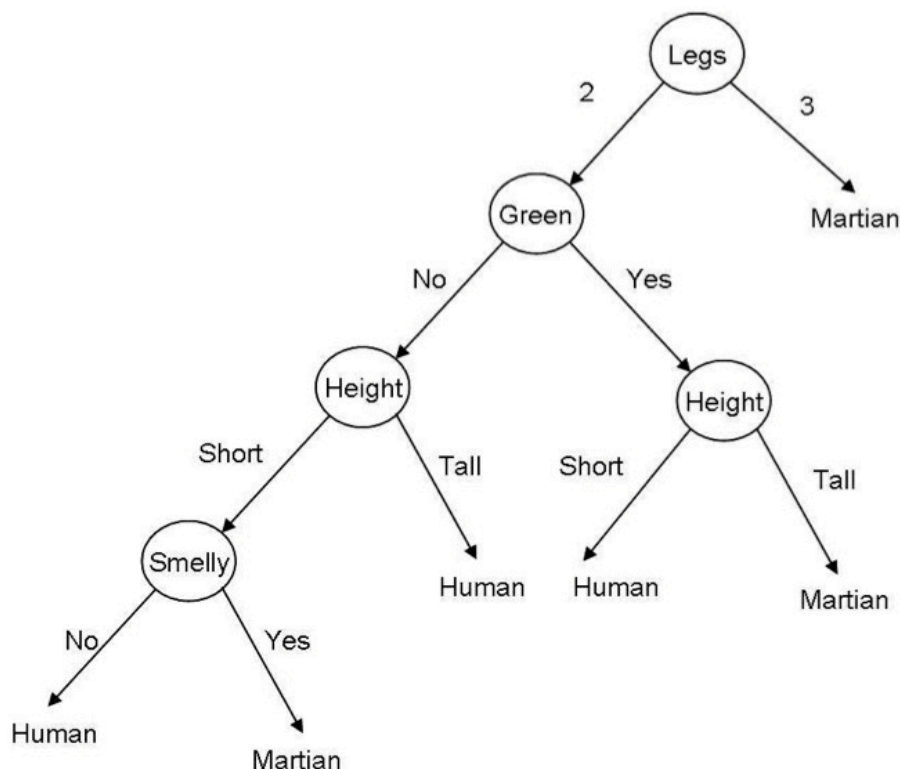
ESCUELA DE INGENIERÍA INFORMÁTICA**SISTEMAS INTELIGENTES****Examen de Teoría, convocatoria extraordinaria. Viernes 01 de Julio de 2016.**

1.-a) [2 puntos] La NASA quiere construir un sistema para discriminar entre Humanos (H) y Marcianos (M) basándose en las siguientes características: $Green \in \{N, Y\}$, $Legs \in \{2, 3\}$, $Height \in \{S, T\}$, $Smelly \in \{N, Y\}$. El conjunto de entrenamiento es el siguiente:

Ejemplo	Especie	Green	Legs	Height	Smelly
1	M	N	3	S	Y
2	M	Y	2	T	N
3	M	Y	3	T	N
4	M	N	2	S	Y
5	M	Y	3	T	N
6	H	N	2	T	Y
7	H	N	2	S	N
8	H	N	2	T	N
9	H	Y	2	S	N
10	H	N	2	T	Y

Construye un árbol de decisión para este problema siguiendo los pasos del método ID3.

RESPUESTA: El atributo cuya IG es menor es Legs, por tanto la raíz del árbol es Legs. Todos los ejemplos que cumplen $Legs == 3$ son Marcianos, así que se etiqueta. En cambio para los ejemplos que cumplen $Legs == 2$ hay que seguir profundizando en el árbol puesto que hay ejemplos de los dos tipos. Por tanto se vuelve a buscar el atributo más informativo, que en este caso es Green. Siguiendo recursivamente el procedimiento, se obtiene el árbol adjunto



b) [0,5 puntos] Pon un ejemplo de una matriz de confusión para un problema de clasificación en TRES clases {A,B,C} de modo que la precisión para la clase A sea 1 y el recall para la clase B sea 0.5.

RESPUESTA: Considerando que en las filas aparece el resultado de la clasificación del árbol y en las columnas la clasificación real, cualquier matriz que satisfaga:

- En la que la fila asociada a la clase A solo haya elementos distintos de 0 excepto el de la diagonal.
- En la columna asociada a la clase B hay tantos elementos fuera de la diagonal como en la diagonal.

Es solución al problema.

c) [0,5 puntos] Considera el siguiente conjunto de datos

<i>Ejemplo</i>	<i>x</i>	<i>y</i>	<i>Clase</i>
1	-1	1	-
2	0	1	+
3	0	2	-
4	1	-1	-
5	1	0	+
6	1	2	+
7	2	2	-
8	2	3	+

¿Cuál es la predicción para el punto (1,1) con 3-NN y 5-NN?

RESPUESTA: La predicción es + en el caso 3-NN, puesto que los tres puntos más cercanos a (1,1) son 2, 5 y 6. En el caso de 5-NN ocurre lo mismo, puesto que la mayoría de los 5 puntos más cercanos (entre los que se incluyen el 2,5 y 6) son de la clase +.

¿Y cuando ponderas por el inverso de la distancia? En este caso la clase es –

2.- a) [1 punto] En el contexto de un algoritmo de búsqueda como el famoso algoritmo A*, ¿qué es un heurístico y qué propiedades debe cumplir para estar bien definido? ¿en qué consiste el método de relajación del problema?

RESPUESTA: Un heurístico h es una función definida entre el conjunto de estados de un espacio de búsqueda y los números reales no negativos, es decir $h: \text{Estados} \rightarrow \mathbb{R}^+$. Para cada estado n , $h(n)$ es una estimación del coste óptimo del subproblema representado por el estado n , $h^*(n)$. Para que esté bien definido tiene que ser una estimación positiva, es decir $h(n) \geq 0$ para todo estado n , y para todo objetivo obj debe ser $h(obj) = 0$.

El método de relajación del problema es un método sistemático para definir heurísticos. Consiste en relajar alguna de las restricciones del problema original hasta obtener un problema que se pueda resolver en tiempo polinomial de forma óptima (clase P). El coste de la solución óptima del problema relajado es el valor que retorna el heurístico para el estado correspondiente del problema real. Si n es un estado que representa un subproblema p y pR es la versión relajada del problema p , entonces $h(n) = \text{coste óptimo del problema } pR$.

b) Consideremos los conocidos heurísticos h_1 (número de fichas descolocadas) y h_2 (suma de distancias ortogonales) para el problema del 8-puzzle y la instancia del problema definida por los siguientes estados inicial y objetivo. Los estados n_1 , n_2 y n_3 son obviamente los sucesores del inicial:

1	2	3
7	8	5
4		6

inicial

1	2	3
7		5
4	8	6

n_1

1	2	3
7	8	5
	4	6

n_2

1	2	3
7	8	5
4	6	

n_3

1	2	3
8		4
7	6	5

objetivo

Sin necesidad de ejecutar el algoritmo A*, responde de forma razonada a las siguientes cuestiones:

b1) [0,5 puntos] Calcula los valores que faltan en la siguiente tabla, algunos de estos valores podrás deducirlos después de responder a las siguientes cuestiones:

	inicial	n1	n2	n3	objetivo
g^*	0	1	1	1	7
h_1	5	5	5	4	0
h_2	7	8	6	6	0
h^*	7	8	6	8	0

RESPUESTA: Los valores de las filas h_1 y h_2 se obtienen de forma trivial, basta con evaluar h_1 y h_2 en los estados. $g^*(\text{inicial}) = 0$ y $g^*(\text{objetivo}) = h^*(\text{inicial}) = C^* = 7$. $h^*(\text{objetivo}) = 0$. Dado que n_1 , n_2 y n_3 son los sucesores del inicial, g^* es 1 para todos ellos. El resto de los valores se pueden obtener a partir de las respuestas a las cuestiones siguientes.

b2) [0,75 puntos] ¿Qué se puede decir con respecto a la expansión de los nodos n_1 , n_2 y n_3 , con cada uno de los heurísticos h_1 y h_2 ?

RESPUESTA: h_1 y h_2 son heurísticos consistentes, en consecuencia las condiciones de expansión se pueden expresar a partir de los valores $g^*(n) + h(n)$.

Si calculamos estos valores para cada heurístico y nodo anteriores, tenemos:

	n_1	n_2	n_3
$g^*(n) + h_1(n)$	6	6	5
$g^*(n) + h_2(n)$	9	7	7

De estos valores se deduce que el nodo n_1 no se expandirá con h_2 , ya que $g^*(n_1) + h_2(n_1) = 1 + 8 = 9 > C^* = 7$; y que los nodos n_1 , n_2 y n_3 se expandirán fijo con h_1 ya que $g^*(n_1) + h_1(n_1) = g^*(n_2) + h_1(n_2) = 6 < C^* = 7$, y $g^*(n_3) + h_1(n_3) = 5 < C^* = 7$. Los nodos n_2 y n_3 podrán expandirse o no con h_2 ya que en estos casos se da la igualdad.

b3) [0,5 puntos] ¿A la vista de lo anterior, podemos asegurar que alguno de los nodos n_1 , n_2 o n_3 no forma parte de una solución óptima?

RESPUESTA: El nodo n_1 no puede formar parte de la solución óptima ya que $h^*(n_1) \geq h_2(n_1) = 8$ y así la solución que pasa por n_1 tiene un valor de al menos $1 + 8 = 9 > C^* = 7$. Por otra parte, dado que $h^*(\text{inicial}) = 7$, n_1 tiene una solución de $1 + 7 = 8$, luego $h^*(n_1) \leq 8$, en consecuencia $h^*(n_1) = 8$.

b4) [0,25 puntos] Observa la estructura del nodo n_3 y compárala con la estructura del nodo objetivo, ¿se te ocurre algún argumento para justificar que $h^*(n_3) > h_2(n_3)$?

RESPUESTA: A la vista del nodo n_3 , con $h_2(n_3) = 6$, las fichas 5, 8 y 7 requieren al menos un movimiento para llegar al objetivo, la ficha 6 no requiere ninguno, y la ficha 4 requiere 3 al menos. Pero es imposible que la ficha 4 salga de su posición sin que la 6 se mueva o sin que la 7 se mueva más de una vez. Esto nos lleva a concluir que es imposible colocar todas las fichas con solo 6 movimientos y que al menos habrá que hacer 8, luego $h^*(n_3) \geq 8$. En consecuencia la solución óptima debe pasar por n_2 , con lo que $h^*(n_2) = 6$. Además, $h^*(n_3) = 8$ ya que éste es el coste de la solución de n_3 a través del inicial.

3.- a) [1 punto] Dado el siguiente código CLIPS, completa las instrucciones que se piden y completa los espacios.

```
(deffacts listas
(lista platanos leche tomates manzanas)
(lista azucar huevos leche tomates patatas)
(lista azucar leche tomates patatas peras manzanas)
(lista peras leche tomates)
)
;Qué hace este código?
Define la BH iniciales[0.1]
;Definimos la regla lista-compra
(defrule lista-compra
(lista ? leche tomates $?) =>)
; ¿Si ejecutamos el comando(run), que sucede?
Nada pues no se han cargado los hechos iniciales. [0.2]
; Comando para cargar los hechos iniciales:
[0.1] (reset)
; Una vez cargados los hechos iniciales, si tecleamos el comando
[0.1] (facts), CLIPS nos muestra
f-0 (initial-fact)
f-1 (lista platanos leche tomates manzanas)
f-2 (lista azucar huevos leche tomates patatas)
f-3 (lista azucar leche tomates patatas peras manzanas)
f-4 (lista peras leche tomates)
For a total of 5 facts.
;Indica las activaciones de la regla lista-compra guardadas en
[0.1] la agenda y que se produjeron al cargar los hechos iniciales.
[0.2]
lista-compra:f-4
lista-compra:f-3
lista-compra:f-1
Teniendo en cuenta que CLIPS emplea por defecto la estrategia de
resolución de conflictos [0.1] depth, la instancia de la regla que
se dispararía al ejecutar el comando (run 1) es la activada por
el hecho f-4 [0.1]
```

b) [1 punto] Dado un SE en un entorno CLIPS que identifica las relaciones familiares elementales, en base a unos datos de entrada expresados con las relaciones progenitor_de y sexo, se pide implementar las reglas que permitan identificar las relaciones familiares sobrino y sobrina. Si lo necesitas puedes implementar reglas que identifiquen alguna relación intermedia que te facilite la tarea.

<pre> (deffacts familia (progenitor_de antonio luisa) (progenitor_de antonio jaime) (progenitor_de marta luisa) (progenitor_de marta jaime) (progenitor_de luisa juan) (progenitor_de luisa ana) (progenitor_de luisa patricia) (progenitor_de jaime luis) (progenitor_de jaime isabel) (progenitor_de juan clara) (progenitor_de ana miguel)) ;La relación ;(progenitor_de antonio luisa) ;debes interpretarla así: ;antonio progenitor_de luisa </pre>	<pre> (deffacts sexo (sexo antonio masculino) (sexo luisa femenino) (sexo jaime masculino) (sexo juan masculino) (sexo miguel masculino) (sexo ana femenino) (sexo patricia femenino) (sexo luis masculino) (sexo marta femenino) (sexo isabel femenino) (sexo clara femenino)) </pre>
--	---

RESP:

; Relaciones auxiliares hermandad_de, tio-sobrino -----

```

(defrule hermandad_de
  (progenitor_de ?proge ?uno)
  (progenitor_de ?proge ?dos)
  (test (neq ?uno ?dos))
=>
  (assert (hermandad-con ?uno ?dos))
  (printout t ?uno " relacion de hermandad con " ?dos crlf)
)

```

```

(defrule tio-sobrino
  (progenitor_de ?padre ?hijo)
  (hermandad-con ?padre ?tio)
=>
  (assert (tio-sobrino ?tio ?hijo))
)

```

; Relaciones sobrino, sobrina

```

(defrule sobrino
  (tio-sobrino ?tio ?sobrino)
  (sexo ?sobrino masculino)
=>
  (assert (sobrino ?sobrino ?tio))
  (printout t ?sobrino " es sobrino de " ?tio crlf)
)

```

```

(defrule sobrina
  (tio-sobrino ?tio ?sobrino)
  (sexo ?sobrino femenino)
=>
  (assert (sobrina ?sobrino ?tio))
  (printout t ?sobrino " es sobrina de " ?tio crlf)
)

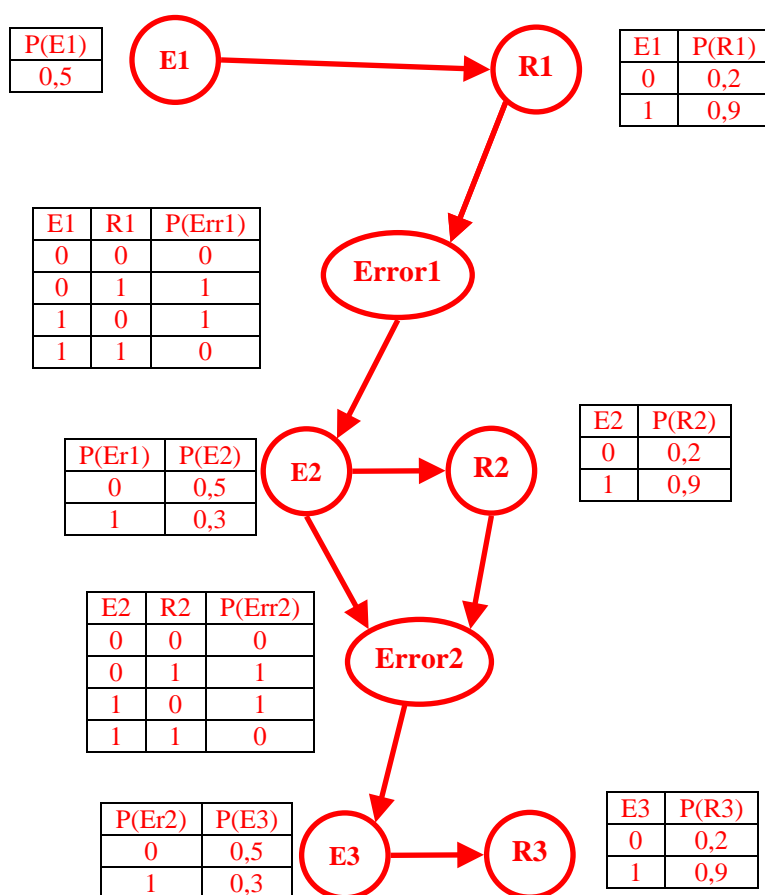
```

4.- [2 puntos] Razonamiento con incertidumbre

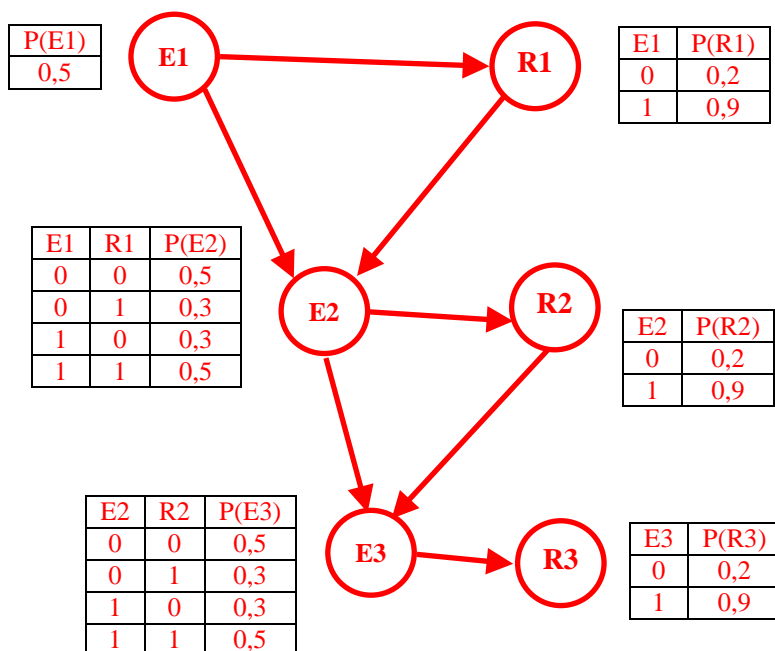
a) [1,2 puntos] Considerar el problema de transmitir palabras de longitud 3 del alfabeto $A=\{a,b\}$ sobre un canal de transmisión. Las palabras se transmiten símbolo a símbolo. La

transmisión tiene ruido y algunas veces no se recibe el símbolo emitido. Si se emite una 'a' se recibe una 'a' con probabilidad 0.8 y una 'b' con probabilidad 0.2. Si se emite una 'b' se recibe una 'b' con probabilidad 0.9 y una 'a' con probabilidad 0.1. La probabilidad de error en la recepción solo depende del símbolo emitido y no de la presencia de error en cualquier otro símbolo. Por otra parte, las palabras emitidas no son completamente aleatorias: el primer símbolo sí que es aleatorio, pero el segundo y el tercero dependen de si hubo algún error en la transmisión del símbolo que lo precede (se entiende que hubo un error si el símbolo recibido fue diferente al símbolo emitido) de la siguiente forma: si hubo un error en la transmisión del símbolo anterior entonces se emitirá una 'a' con probabilidad 0.7 y una 'b' con probabilidad 0.3. Si no hubo un error en la transmisión del símbolo anterior, entonces el símbolo emitido será aleatorio. Dibuja una red bayesiana que represente este problema, con sus correspondientes tablas de probabilidad condicionada. Si para alguna tabla necesitas alguna probabilidad no especificada en este enunciado, puedes proponer cualquier valor que consideres razonable.

Solución:



Otra posible solución:



b) [0,4 puntos] En la red bayesiana que hayas construido en el apartado a), **razona** si se cumplen o no las siguientes independencias, dada la estructura de la red:

- ¿Son independientes el primer símbolo emitido y el tercer símbolo recibido?
Solución: No son independientes, ya que hay un camino directo que los une ($E1$ -Error1- $E2$ -Error2- $E3$ - $R3$) y dicho camino no está bloqueado en ningún nodo, según el criterio de D-separación.
- ¿Son independientes el primer símbolo emitido y el tercer símbolo recibido, si conocemos el valor del segundo símbolo emitido?
Solución: En este caso sí que son independientes, ya que el nodo $E2$ está observado, y dicho nodo bloquea todos los posibles caminos que hay entre $E1$ y $R3$, según el criterio de D-separación.

c) [0,4 puntos] En una determinada red bayesiana compuesta por 5 variables A, B, C, D y E, queremos calcular la probabilidad $P(A, \neg B \mid C)$ mediante el método del muestreo estocástico estándar. Para ello, hemos realizado 8 muestras aleatorias, cuyos valores resumimos en la siguiente tabla. Indica cuál sería el resultado final de dicha probabilidad, utilizando únicamente las muestras realizadas.

A	B	C	D	E
V	V	F	F	V
V	F	V	F	V
V	F	F	F	F
V	V	F	V	V
F	V	F	F	V
V	V	V	F	F
F	F	F	V	F
F	V	V	F	V

Solución:

Sabemos que $P(A, \neg B \mid C) = P(A, \neg B, C) / P(C)$. En el muestreo estocástico estándar, simplemente se trataría de contar cuántas muestras cumplen A y $\neg B$ y C (hay una que cumple eso), y dividir ese número por el número de muestras que cumplen C (hay tres que lo cumplen). Por tanto, el resultado final sería $1/3$.