



APELLIDOS:

PL:

NOMBRE:

DNI:

ESCUELA DE INGENIERÍA INFORMÁTICA

SISTEMAS INTELIGENTES

Examen Final de Teoría. Convocatoria Extraordinaria. Lunes 15 de Junio de 2015.

1.- Responder con precisión y brevedad a las siguientes cuestiones sobre la condición de consistencia de los heurísticos del algoritmo A*.

a) [0.5 puntos] ¿Qué es un heurístico consistente o monótono?

RESP: Un heurístico h es consistente si para todo par de nodos n_1 y n_2 del espacio de búsqueda se cumple $h(n_1) \leq h(n_2) + k(n_1, n_2)$, siendo $k(n_1, n_2)$ el coste del camino más corto entre los nodos n_1 y n_2 .

Un heurístico h es monótono si para todo par de nodos n_1 y n_2 del espacio de búsqueda se cumple $h(n_1) \leq h(n_2) + c(n_1, n_2)$, siendo $c(n_1, n_2)$ el coste del arco entre los nodos n_1 y n_2 , $c(n_1, n_2)$ es infinito si no existe tal arco.

Se puede probar que las dos propiedades son equivalentes.

b) [0,5 puntos] ¿Cuáles son las dos principales consecuencias que tienen los heurísticos consistentes en el comportamiento del algoritmo A*?

RESP: El hecho de que un heurístico sea consistente tiene varias consecuencias en el comportamiento del algoritmo A* que lo utiliza. De ellas, las dos más importantes son las siguientes:

Si un heurístico h es consistente, entonces h es admisible, lo que se traduce en que el algoritmo A* es admisible y en consecuencia siempre encuentra la solución óptima.

Si un heurístico h es consistente, entonces cuando el algoritmo A* que lo usa elige un nodo n para ser expandido, ocurre que $g(n) = g^*(n)$. Esto significa que en ese momento ya se conoce el camino más corto desde el nodo inicial a n , es decir $g^*(n)$. Esto se traduce en que no hay que rectificar el camino desde el inicial hasta n , o bien reexpandir n dependiendo de la versión del algoritmo que se utilice, y de esa forma la búsqueda será más eficiente.

c) [1 punto] Los heurísticos h_1 y h_2 del problema del 8-puzzle (los que calculan respectivamente el número de fichas descolocadas y la suma de las distancias ortogonales con respecto a las posiciones en el objetivo) son consistentes. ¿Será consistente también el heurístico h_4 , que se calcula como $h_4(n) = h_1(n) + h_2(n)$, para todo nodo n ? En este caso hay que justificar la respuesta.

RESP: El heurístico h_4 es claramente no admisible, y en consecuencia es no consistente. Para probar que no es admisible, basta con considerar un ejemplo: sea n un nodo que se obtiene a partir del objetivo haciendo un movimiento simple. Claramente $h^*(n) = h_1(n) = h_2(n) = 1$. Luego $h_4(n) = 1 + 1 = 2 > h^*(n)$.

2.- [1 punto] Explica el cometido del operador de selección en un algoritmo genético y explica qué influencia podría tener en el fenómeno de convergencia prematura. Pon algún ejemplo de operador de selección que conozcas.

RESP: El operador de selección es el encargado de elegir los cromosomas de la población actual que van a ser cruzados para dar lugar a nuevos cromosomas hijos. Dos ejemplos de operador de selección son la ruleta y el torneo. En el operador de ruleta se elige con una mayor probabilidad a cromosomas con mejor fitness y con menor probabilidad a cromosomas con peor fitness. El operador de torneo elige a un determinado número aleatorio de cromosomas, y de entre ellos se selecciona al mejor de todos. El operador de selección tiene

una gran influencia en el fenómeno de convergencia prematura, ya que si siempre elige a los mejores cromosomas de la población, la diversidad genética caería rápidamente y la evolución de la población se estancaría muy pronto, dando lugar a dicho fenómeno.

3.- a) [0.7 puntos] Representar el siguiente enunciado en lógica de predicados:

*Algunos dibujos animados no son infantiles.
Los dibujos animados no infantiles son para adultos.
Los niños que ven dibujos animados para adultos no los entienden y se aburren.
Los Simpsons son dibujos animados para adultos.
Juan es un niño y ve los Simpsons.*

b) [1.3 puntos] Luego hacer el paso a notación de Kowalski y responder a la pregunta siguiente utilizando refutación:

¿Existe algún niño que se aburre?

RESP: El enunciado se puede modelar de la siguiente forma:

```
public static FOLKnowledgeBase createExamenJunioKnowledgeBase(
    InferenceProcedure infp)
{
    FOLKnowledgeBase kb = new FOLKnowledgeBase(DomainFactory.examenJunioDomain(),
        infp);

    // Algunos dibujos animados no son infantiles
    kb.tell("EXISTS x (Dibujo(x) AND NOT Infantil(x))");
    // Los dibujos animados no infantiles son para adultos.
    kb.tell("FORALL x (Dibujo(x) AND NOT Infantil(x) => Adulto(x))");
    // Los niños que ven dibujos animados para adultos no los entienden y se aburren.
    kb.tell("FORALL x FORALL y ((Niño(x) AND Dibujo(y) AND Adulto(y) AND Ve(x,y)) =>
    NOT Entiende(x,y) AND Aburre(x))");
    // Los Simpsons son dibujos animados para adultos.
    kb.tell("(Dibujo(Simpsons) AND Adulto(Simpsons))");
    // Juan es un niño y ve los Simpsons.
    kb.tell("(Niño(Juan) AND Ve(Juan, Simpsons))");
    kb.tell("(Niño(Juan) AND Dibujo(Simpsons) AND Ve(Juan, Simpsons))");

    return kb;
}
```

La pregunta se formaliza así:

```
// Existe algún niño que se aburre
String query = "(EXISTS x (Niño(x) AND Aburre(x)))";
```

Y el resultado de la refutación, sin utilizar el mecanismo de Green (predicado Answer), es el siguiente. Se llega a la cláusula vacía, luego el resultado es afirmativo:

TFM Resolution, ExamenJunioDemo1

```
#newClauses: 2
#newClauses: 0ExamenJunio Knowledge Base:
EXISTS x (Dibujo(x) AND NOT(Infantil(x)))
FORALL x ((Dibujo(x) AND NOT(Infantil(x))) => Adulto(x))
FORALL x FORALL y (((Niño(x) AND Dibujo(y)) AND Adulto(y)) AND Ve(x,y)) => NOT(Entiende(x,y)) AND Aburre(x))
(Dibujo(Simpsons) AND Adulto(Simpsons))
((Niño(Juan) AND Ve(Juan,Simpsons)) AND Dibujo(Simpsons))
```

```
Query: (EXISTS x (Niño(x) AND Aburre(x)))
Proof, Answer Bindings: {x=null}
```

Step	Proof	Justification
1	FORALL x FORALL y (((Niño(x) AND Dibujo(y)) AND Adulto(y)) AND Ve(x,y)) => NOT(Entiende(x,y)) AND Aburre(x))	Premise
2	((Niño(Juan) AND Ve(Juan,Simpsons)) AND Dibujo(Simpsons))	Premise
3	[~Aburre(v8), ~Niño(v8)]	Goal
4	[Aburre(x)]	Clausified 1
5	[Aburre(v3)]	Renaming of 4
6	[Aburre(v7)]	Renaming of 5
7	[Niño(Juan)]	Clausified 2
8	[~Aburre(Juan)]	Resolution: 3,7 {}, {}
9	[]	Resolution: 6,8 {v7=Juan}, {}

4.- Considera el siguiente conjunto de ejemplos utilizado para aprender a identificar si una seta es venenosa o no

<i>Ejemplo</i>	<i>EsDura</i>	<i>EsOlorosa</i>	<i>TieneManchas</i>	<i>EsRegular</i>	<i>EsVenenosa</i>
A	0	0	0	0	0
B	0	0	1	0	0
C	1	1	0	1	0
D	1	0	0	1	1
E	0	1	1	0	1
F	0	0	1	1	1
G	0	0	0	1	1
H	1	1	0	0	1
U	1	1	1	1	1
V	0	1	0	1	0
W	1	1	0	0	1
X	1	0	1	1	0
Y	0	0	1	1	1
Z	1	1	0	1	1

- a) [1 punto] Considerando sólo los ejemplos de la A a la H, ¿Cuál es el atributo que proporciona mayor Ganancia de Información? (Cálculalo detalladamente).

$$H(\text{esVenenosa}) = -\frac{3}{8} * \log(\frac{3}{8}) - \frac{5}{8} * \log(\frac{5}{8})$$

$H(\text{esVenenosa}/\text{EsRegular}=0)=1$ (porque hay igual numero de ejemplos de cada clase)

$H(\text{esVenenosa}/\text{EsRegular}=1) = -\frac{1}{4} * \log(\frac{1}{4}) - \frac{3}{4} * \log(\frac{3}{4}) = 0.8133$ (Estos mismos cálculos están en la transparencia 44)

$$H(\text{esVenenosa}/\text{EsRegular}) = 1.8133$$

$$H(\text{esVenenosa}/\text{TieneManchas}=0) = -\frac{2}{5} * \log(\frac{2}{5}) - \frac{3}{5} * \log(\frac{3}{5}) = 0.970$$

$$H(\text{esVenenosa}/\text{TieneManchas}=1) = -\frac{1}{3} * \log(\frac{1}{3}) - \frac{2}{3} * \log(\frac{2}{3}) = 0.9183$$

$$H(\text{esVenenosa}/\text{EsOlorosa}=0) = -\frac{2}{5} * \log(\frac{2}{5}) - \frac{3}{5} * \log(\frac{3}{5})$$

$$H(\text{esVenenosa}/\text{EsOlorosa}=1) = -\frac{1}{3} * \log(\frac{1}{3}) - \frac{2}{3} * \log(\frac{2}{3})$$

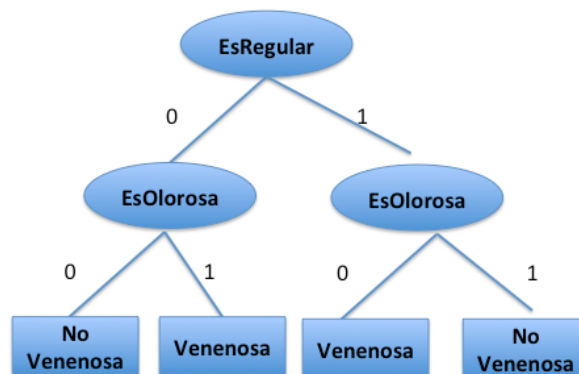
$$H(\text{esVenenosa}/\text{EsDura}=0) = -\frac{2}{5} * \log(\frac{2}{5}) - \frac{3}{5} * \log(\frac{3}{5})$$

$$H(\text{esVenenosa}/\text{EsDura}=1) = -\frac{1}{3} * \log(\frac{1}{3}) - \frac{2}{3} * \log(\frac{2}{3})$$

$$H(\text{esVenenosa}/\text{EsOlorosa}) = H(\text{esVenenosa}/\text{EsDura}) = H(\text{esVenenosa}/\text{EsOlorosa}) = 1.8883$$

Como $H(\text{esVenenosa}/\text{EsRegular}) < H(\text{esVenenosa}/\text{EsOlorosa}) = H(\text{esVenenosa}/\text{EsDura}) = H(\text{esVenenosa}/\text{EsOlorosa})$, el atributo que mayor ganancia de información tiene es EsRegular

- b) [1 punto] Suponiendo que el árbol siguiente fuera una solución proporcionada por algún algoritmo y considerando los ejemplos U a Z como conjunto de test, ¿Cuál sería la precisión obtenida con dicho algoritmo?



Ejemplo	EsOlorosa	EsRegular	EsVenenosa	Predicho	
U	1	1	1	0	FN
V	1	1	0	0	TN
W	1	0	1	1	TP
X	0	1	0	1	FP
Y	0	1	1	1	TP
Z	1	1	1	0	FN

Los atributos que deciden si es venenosa o no de acuerdo con este árbol son *EsRegular* y *EsOlorosa*. Comparando la clasificación real (*EsVenenosa*) con la obtenida por evaluación del árbol sobre cada ejemplo (*Predicho*) obtenemos que hay 2 TP (True Positives) y 1 FP (False Positive). Así que la precisión (LA ÚNICA PRECISIÓN QUE EXISTE) es

$$P = \frac{TP}{TP + FP} = \frac{2}{3}$$

- c) [1 punto] Considera ahora la tabla resultante de eliminar la columna correspondiente al atributo *EsRegular* en la tabla anterior. Aplicando el algoritmo IB3 sobre el conjunto de datos de la A a la H de la nueva tabla, y teniendo en cuenta que la clase objetivo es “*EsVenenosa*”, ¿A qué clase pertenece el ejemplo V?

$$D(A, V) = 1$$

$$D(B, V) = 1.4142$$

$$D(C, V) = 1$$

$$D(D, V) = 1.4142$$

$$D(E, V) = 1$$

$$D(F, V) = 1.4142$$

$$D(H, V) = 1$$

$$D(H, V) = 1$$

Una vez calculadas las distancias euclídeas hay que coger los 3 ejemplos más próximos y asignar la clase mayoritaria de ellos a V. Como hay empate, se pueden elegir los 5 que están a igual distancia o se eligen 3 al azar (de entre los más próximos), y se asigna a V la clase mayoritaria.

5.- [2 puntos] Tenemos una enfermedad que queremos diagnosticar, y tres máquinas tales que cada una realiza un determinado test a un paciente. El resultado de dichos test puede ser positivo o negativo, y sabemos que los resultados de los test son condicionalmente independientes conocido si se tiene o no se tiene la enfermedad.

La enfermedad es más común en los fumadores (la tienen con probabilidad 0.005) que entre los no fumadores (la tienen con probabilidad 0.00001).

Por otra parte, sabemos que la probabilidad de que un individuo cualquiera sea fumador es de 0.3. Si las máquinas funcionan bien, las probabilidades de que los test den positivo según se tenga o no la enfermedad son las siguientes:

	Test 1 positivo	Test 2 positivo	Test 3 positivo
Enfermo	0.9	0.8	0.75
No enfermo	0.2	0.05	0.06

Por desgracia, las máquinas no son perfectas, sino que cada una de las máquinas funciona incorrectamente con probabilidad 0.001, y en ese caso devuelve positivo o negativo de forma aleatoria en su correspondiente test.

En esta pregunta se pide determinar una red bayesiana, con sus tablas de probabilidad, para representar el conocimiento anterior sobre el problema de diagnóstico de la enfermedad. Si en alguna tabla de probabilidad te hiciese falta algún dato no especificado en el enunciado del problema, puedes poner cualquier valor que consideres razonable.

Solución:

