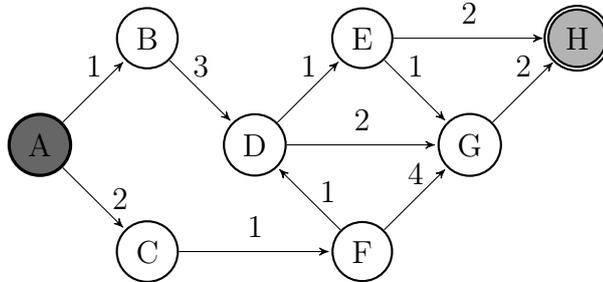


Examen Final de IA

(7 de junio de 2012)

Duración: 3 horas

1. (1.5 puntos) Dado el siguiente grafo, donde cada arco indica su coste, y la tabla que indica la estimación del coste h hasta la solución, indica cual sería el árbol de búsqueda que se obtendría mediante el algoritmo de A* e IDA* para encontrar el camino entre el nodo A y el nodo H. Haz la generación de los nodos siguiendo el orden alfabético e indica claramente las reexpansiones de los nodos y los cambios de coste que aparezcan. ¿Es la función heurística admisible?



Nodo	A	B	C	D	E	F	G	H
$h(\text{nodo})$	4	4	3	3	2	3	2	0

2. (2 puntos) *Foursquare* es una red social que permite compartir localizaciones entre usuarios, de manera que puede servir como un sistema de geolocalización y como un sistema para recomendar lugares que se visitan con frecuencia. Los **usuarios** de este servicio se pueden categorizar en usuarios esporádicos, solo comparten lugares de vez en cuando, y los que marcan tendencia (*trenders*), que comparten continuamente los lugares por los que pasan. También hay **empresas** que se promocionan y comparten su localización con los usuarios para atraer más público.

Los **lugares** que hay en el servicio se pueden clasificar según su localización geográfica en lugares urbanos (ciudades, pueblos, urbanizaciones) y naturaleza (montaña, playa, ...). Estos también se pueden clasificar según su temática en lugares de restauración (bares, restaurantes), ocio (cines, discotecas, teatros, música, ...), cultura (museos, conciertos, bibliotecas), turismo (monumentos, edificios, lugares pintorescos, playas, ...), compras (ropa, alimentación, libros, ...), etc.

Finalmente, también se representan las **visitas** de los usuarios a los lugares, para las cuales registra la fecha y hora de la visita y el tiempo que se ha estado en el lugar. Las visitas se clasifican en cortas y largas según si su duración es mayor o menor a dos horas.

a) Define completamente las siguientes relaciones:

- *Recomienda*, entre usuario y lugar, indicando que el usuario recomienda el lugar
- *Amigo*, entre usuarios, indicando que los usuarios tienen una relación de amistad
- *Realiza*, entre usuario y visita, indicando que el usuario ha realizado una visita y *al lugar*, entre visita y lugar, indicando que la visita se realiza al lugar

b) Define el slot *seguidores*, en usuario que retorne todos los amigos que han visitado alguno de los lugares que ha recomendado.

c) Define la relación *ahora* entre usuario y lugar, indicando que el usuario está actualmente en ese lugar. Define un slot *estan_aquí* que se actualice con la lista de sus amigos que están en ese lugar cuando se crea la relación *ahora* y que sea vacía cuando se destruye. ¿Que ventajas/inconvenientes tendría el usar un método en lugar del slot propuesto?

d) Define el slot *gratis* en lugar, indicando si el lugar es o no gratuito. Define lo necesario para que el concepto visita herede este slot, indicando que la visita ha sido gratuita. Define el slot *sin_blanca* en usuario que sea cierto si más de la mitad de las visitas del usuario han sido gratuitas.

3. (3 puntos) La coordinadora de cooperativas agrarias quiere desarrollar un sistema basado en el conocimiento para aconsejar a los agricultores cuales son los cultivos más adecuados para las características de sus explotaciones. El ingeniero del conocimiento que han contratado les ha indicado que se trata de un problema de análisis y que se podrían abstraer las siguientes características que podrían servir para el desarrollo:

- La **climatología**, que se podría clasificar en árida, seca, mediterránea, continental y subtropical, que dependería de características como el régimen de temperaturas, el régimen de lluvias, humedad y vientos, el número de estaciones reales, ...
- El **acceso al agua**, que se podría clasificar en muy bueno, bueno, regular o malo, dependiendo del régimen de lluvias, la existencia de riego propio o colectivo, el acceso a embalses, ...
- Los **gastos asumibles**, que se podrían clasificar en bajos, medios o altos, dependiendo de si se dispone de maquinaria, el acceso a la mano de obra, la necesidad de seguros, el precio dispuesto a gastar en semillas y fertilizantes, ...
- El **tipo de terreno**, que se podría clasificar según su fertilidad en los valores 1 a 5, dependiendo de su acidez, su contenido en nutrientes, nitrógeno, sodio, potasio, su contenido en arenas, arcillas, ...
- La **productividad deseable**, que se puede clasificar en normal, alta e intensiva, que depende del número de cosechas deseadas, el terreno disponible, el agotamiento de la tierra asumible, el ratio beneficio coste deseado, ...

A partir de estas características queremos poder determinar, en primer lugar, que tipo general de cultivo es el más adecuado de entre los siguientes: cereal, hortaliza, legumbre y raíz/bulbo. Para ello contamos con el conocimiento de los expertos que nos cuentan por ejemplo, que los cereales y las legumbres no necesitan terreno con mejor fertilidad que 3, se adaptan muy bien a climas áridos y secos, son cultivos con pocos gastos, pero con una productividad normal, y si el acceso al agua es malo es mejor el cereal que la legumbre. En cambio, para cultivar hortalizas hace falta un acceso al agua bueno o muy bueno, los gastos asumibles han de ser medios o altos y la climatología no puede ser ni seca, ni árida. Las plantas de raíz/bulbo necesitan terrenos con fertilidad al menos 3 y con un acceso al agua al menos regular y los gastos asumibles han de ser por lo menos medios, ...

Una vez decidido el tipo general de cultivo, se han de dar prioridades sobre cultivos concretos, con valores entre 0 y 2:

- Para los cereales se puede recomendar plantar trigo, cebada, avena, soja o centeno. Soja y trigo necesitan una fertilidad de al menos 2 para ser viables, si se desea un ratio beneficio coste alto, si se tienen un terreno muy grande, es mejor plantar cebada, avena o soja, ...
- Para las hortalizas se puede recomendar plantar tomate, pepino, pimiento o calabacín. El ratio beneficio coste del tomate y el pimiento es mayor que el del pepino y el calabacín, el acceso al agua para el calabacín ha de ser muy bueno, la fertilidad del terreno para el pepino ha de ser por lo menos 4, ...
- Para las legumbres se puede recomendar plantar guisante, garbanzo, lenteja, judía y haba. El garbanzo y la lenteja se dan mejor en climas secos, la fertilidad del terreno no es necesario que sea mayor que 2. En cambio los guisantes, judías y habas necesitan por lo menos un terreno con fertilidad 2 y un acceso a agua por lo menos regular, ...
- Para las raíces/bulbos se puede recomendar plantar remolacha, zanahoria, patata, nabo y cebolla. La remolacha y el nabo necesitan un acceso al agua bueno o muy bueno y se adaptan muy bien a clima continental. La patata y la zanahoria necesitan gastos bajos, pero necesitan acceso a mano de obra, ...

a) Explica cómo resolverías el problema usando clasificación heurística, indicando qué elementos se utilizarían en cada una de las fases. Da ejemplos de reglas para cada una de las fases de esta

metodología que formalicen el conocimiento del problema, incluyendo algunas de las reglas que aparecen en el enunciado.

- b) El asignar exclusivamente un cultivo como solución podría ser poco flexible, se podrían utilizar las redes bayesianas para obtener un conjunto de prioridades para los diferentes cultivos. Un análisis más profundo de las características del problema nos ha permitido también darnos cuenta de que las diferentes características del problema están interrelacionadas, por ejemplo, el acceso al agua y el tipo de terreno dependen en cierta manera de la climatología, el tipo de terreno influye también sobre la productividad que podemos obtener y dependiendo de los gastos que podemos asumir obtendremos más o menos productividad.

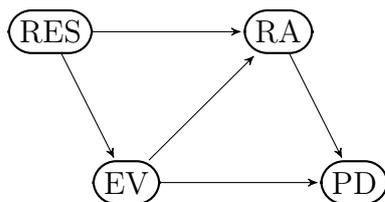
Define el problema de asociación heurística como una red bayesiana expresando en ella al menos las relaciones indicadas y todas las necesarias para resolver el problema de asociación. Da un ejemplo de tabla de probabilidad de algún nodo, inventándote las probabilidades, pero expresando como influyen los valores de los nodos padre en las probabilidades de los valores de los nodos hijo.

- c) Una forma alternativa de solucionar el problema sería utilizar el razonamiento basado en casos. En lugar de plantear el problema de seleccionar primero el tipo de cultivo y después priorizar los cultivos concretos, podríamos directamente aconsejar prioridades para los cultivos concretos. Explica cómo se resolvería el problema con esta metodología: qué serían los casos, cómo se describirían, qué sería una solución en un caso, cómo se recuperarían los casos y cómo se obtendría la solución de un nuevo caso a partir de los casos recuperados.

4. (1 punto) Una aplicación bastante difundida de las redes bayesianas es la implementación de asistentes de usuario, como por ejemplo en un sistema operativo. Estos asistentes pueden recolectar información del funcionamiento del sistema y presentar recomendaciones y avisos al usuario para que atienda a posibles problemas en su sistema.

Por ejemplo, podemos tener un asistente que cronometra la respuesta de la aplicación en primer plano (RA), el rendimiento de entrada/salida de los discos duros (RES) y el tamaño de la cola de eventos del interfaz del sistema operativo (EV) y usar estas informaciones para lanzar un mensaje sobre un posible problema en la configuración de drivers (PD).

Esta red bayesiana representaría las relaciones entre estos conceptos.



Y estas tablas de probabilidades representarían las influencias de unos nodos sobre otros.

RES	P(RES)	RES	P(EV)		RES	EV	P(RA)		EV	RA	P(PD)	
			vacía	llena			< 500ms	> 500ms			Sí	No
> 12Mbps	0.8	> 12Mbps	0.7	0.3	> 12Mbps	vacía	0.9	0.1	vacía	< 500ms	0.1	0.9
< 12Mbps	0.2	< 12Mbps	0.1	0.9	> 12Mbps	llena	0.5	0.5	vacía	> 500ms	0.3	0.7
					< 12Mbps	vacía	0.6	0.4	llena	< 500ms	0.4	0.6
					< 12Mbps	llena	0.1	0.9	llena	> 500ms	0.9	0.1

Define la distribución de probabilidad conjunta de $P(RES, EV, RA, PD)$ que representa la red y usa el algoritmo de eliminación de variables para calcular la probabilidad de $P(RA|EV = llena, PD = Sí)$.

5. (2.5 puntos) Se quiere implementar una interfaz en lenguaje natural capaz de hacer posible la comunicación entre la NASA y los futuros robots enviados a planetas para su explotación minera. Dicha comunicación consiste en enviar órdenes desde la Tierra al robot, órdenes como por ejemplo:

recoge 3 litros de líquido
vé a buscar 500 gramos de tierra
obtén 1 centímetro cúbico de aire

Para dicha finalidad, se pensó en la implementación de una DCG. Parte de ella se muestra a continuación:

oracion → orden, muestra.
...
orden → [recoge].
orden → [busca].
orden → [obtén].
orden → [vé,a,buscar].
orden → [vé,a,recoger].
numero(p) → [X], {number(X),X>1}.
numero(s) → [1].
...

Se pide completarla según los siguientes apartados:

- a) Completarla para **reconocer** constituyentes de tipo “muestra” con concordancia de número, como los siguientes:

3 cubetas de líquido
500 gramos de tierra
5 palas de tierra
entre 3 y 4 litros de líquido

- b) Completar el apartado anterior introduciendo lo necesario para que exista coherencia semántica entre unidades de medida y tipo de producto a extraer (ej: no se puede medir líquidos en gramos, ni se pueden extraer mediante una pala).
c) Completar los apartados anteriores para que se puedan **reconocer** órdenes de extracción de diferentes muestras a la vez, como por ejemplo:

3 kilos de roca, 500 gramos de tierra y entre 3 y 4 cubetas de líquido

- d) Completar la DCG resultante para que se puedan **generar** salidas a partir de ella, como por ejemplo:

busca 3 kilos de roca → `extraer([roca,3, kilo])`
recoge entre 3 y 4 litros de líquido → `extraer([liquido,[3,4], litro])`
ve a buscar 3 kilos de roca, 500 gramos de tierra y entre 3 y 4 cubetas de líquido
→ `extraer([[roca,3, kilo], [tierra, 500, gramo], [liquido,[3,4], cubeta]])`

Las notas se publicaran el día **25 de junio**.