

Examen Final de IA

(19 de Junio de 2007)

Duración: 3 horas

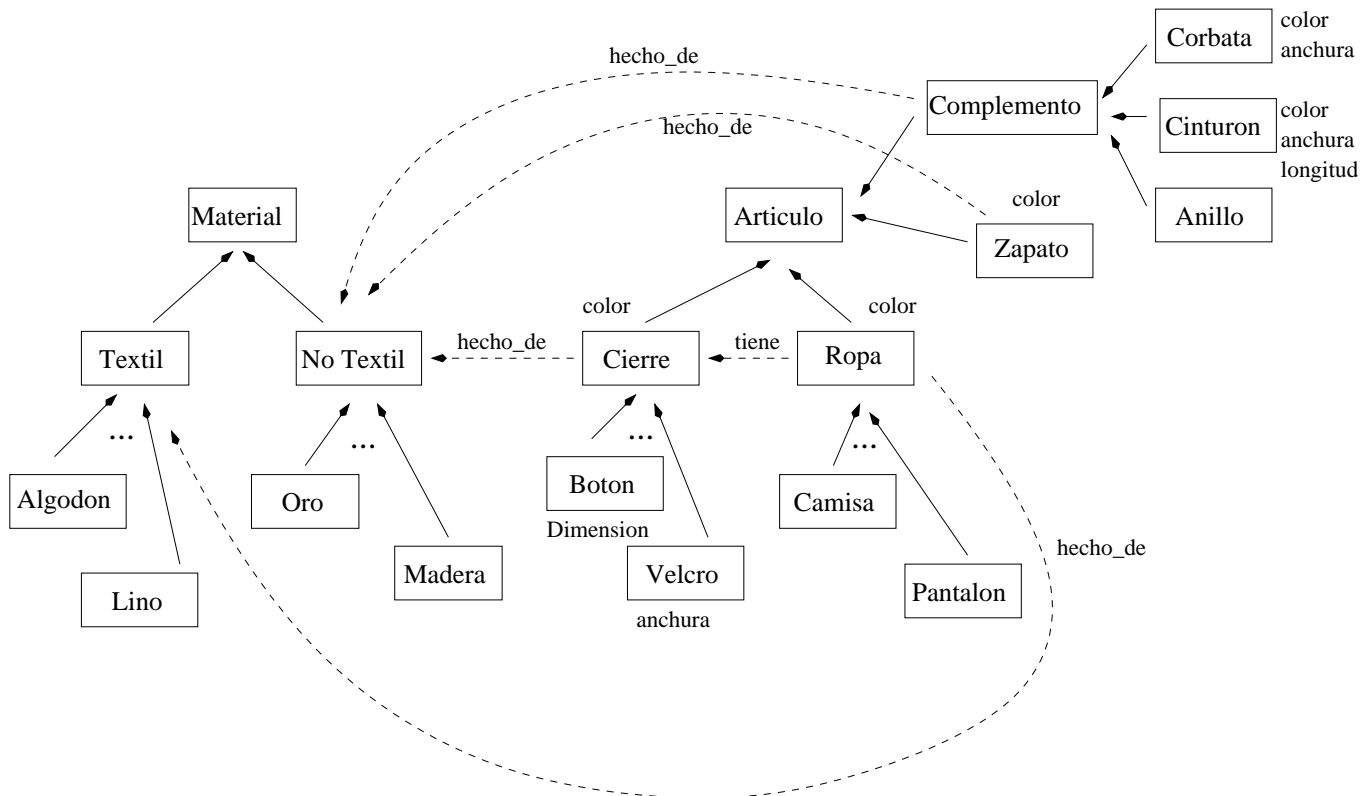
1. (3 puntos) Una empresa quiere producir sistemas basados en el conocimiento (SBC) con interfaces en lenguaje natural. Para abordar el problema de la interpretación de oraciones ha optado por la siguiente estrategia: generar, primero, una gramática DCG para reconocer constituyentes sintácticos del lenguaje genérico, y posteriormente, adaptar dicha gramática al dominio semántico particular que requiere un SBC concreto. Se pide:

a) Diseñar una parte pequeña de la gramática DCG, concretamente, diseñar una DCG con concordancia de género y número, para reconocer sintagmas nominales (SN) que contengan una conjunción de modificadores (MOD) de tipo adjetivo (ADJ) o de tipo sintagma preposicional (SP). Ejemplos de frases que se desean reconocer son:

- el gato.
- plantas verdes e inmaculadas.
- unas ideas avanzadas, de calidad alta y factibles.
- piso alto, con cocina separada, con ventana y equipada, con baño de lujo y soleado.

Nota: Es suficiente con incorporar dos entradas léxicas por cada categoría gramatical.

b) Adaptar la DCG anterior para reconocer frases relativas al dominio semántico de artículos de ropa y complementos, representado mediante la ontología de la figura.



Dicha adaptación requiere la incorporación de las restricciones de selección necesarias para reconocer la descripción de un artículo y TODAS sus propiedades. Por ejemplo, en la descripción de cualquier zapato debe aparecer su color y el material del que está fabricado. Ejemplos de oraciones descriptoras del dominio semántico en cuestión son:

- anillo de oro.
- sombrero negro y de cuero.
- camisa blanca, de algodón y con botones verdes, grandes y de madera.

Nota: es suficiente con incorporar entradas léxicas para controlar las descripciones anteriores de anillo y camisa

c) Adapta la DCG del apartado (b) para que genere salidas como las siguientes:

- anillo de oro
→ `anillo([material(oro)])`
- camisa blanca, de algodón y con botones verdes, grandes y de madera.
→ `camisa([color(blanco), material(algod on), boton([color(verde), dimension(grande), material(madera)])])`.

2. (4 puntos) Estamos construyendo un SBC orientado a la creación de canales de televisión de entretenimiento personalizados de pago. Este sistema parte de la información de los hábitos televisivos de un cliente y gustos o restricciones que pueda tener y es capaz de crear una programación diaria de cierto número de horas utilizando una base de recursos audiovisuales que incluyen programas de entretenimiento tanto grabados como en directo.

a) Hemos decidido que en nuestro sistema incluiremos programas de diferente tipología (series, películas, concursos, deportes, música, documentales, ...). Para cada uno nos interesa si el programa está grabado o es en directo, su duración, el idioma y un precio. Para las series y películas nos interesa también su género y su nacionalidad. Para los deportes nos interesa también el tipo de deporte.

Incluimos también en la representación una tipología cerrada de usuarios que utilizaremos para creación del canal personalizado. Esta tipología incluye entre otros el usuario deportivo, cuyos hábitos incluyen mayoritariamente programas de deporte y que podría especializarse según el deporte que prefiere, el cinéfilo, que ve preferentemente películas y/o series, que puede tener diferentes especializaciones dependiendo del tipo de películas y/o series que ve o el generalista, que no tiene una tendencia clara.

Representa gráficamente mediante una red de frames estos conceptos. Discute la necesidad de representar el atributo género en películas y series como un slot o como un frame.

b) El resultado del sistema será un conjunto de programas que se emitirán a una hora específica determinada por el tiempo de inicio del canal y la duración de los programas.

Representa gráficamente mediante una red de frames lo necesario para representar una solución.

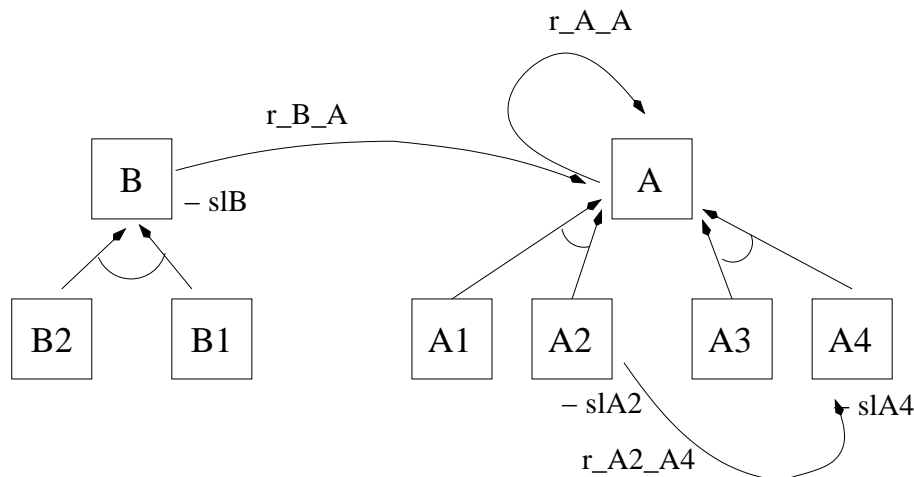
c) Hemos identificado tres subproblemas. El primero debe identificar el tipo al que pertenece el usuario a partir de los hábitos televisivos, estos hábitos incluyen el porcentaje de cada tipo de programa que ve el usuario y las características de estos programas.

Evidentemente se trata de un problema de análisis ¿Se podría plantear como un problema de clasificación heurística? Si se puede ¿cómo encajarían las diferentes fases de esta metodología en este problema? Da ejemplos de reglas.

d) El segundo subproblema utilizaría la tipología identificada del usuario y otras restricciones que indicara como preferencia por programas en versión original, preferencias por algún género de cine y televisión, por algún deporte, variedad de los programas en el horario, duración máxima de los programas, emisiones en directo, etc. Nos gustaría utilizar tratamiento de la incertidumbre para poder saber si un programa se adapta mejor o peor al usuario ¿te parece adecuado? Si te lo parece, ¿qué tipo de tratamiento de la incertidumbre utilizarías? Justifica tu respuesta.

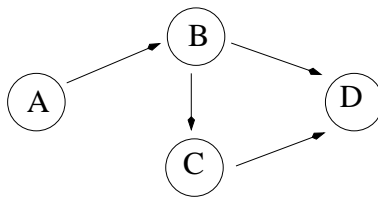
e) El tercer problema es la elaboración de la programación. A partir de los programas que se adecuan al usuario y el numero de horas que ha de tener la emisión debemos escoger los programas más adecuados. Como es un canal de pago, el usuario tiene una cuota fija, lo que nos interesa es incluir los programas cuya suma de precios esté lo mas cercana posible a esa cuota sin sobrepasarla. Es evidentemente un problema de síntesis. Describe como lo resolverías utilizando la estrategia de proponer y aplicar.

3. (2 puntos) Dada la siguiente red de frames, responde a las siguientes preguntas, justificando tu respuesta:



- Tenemos la instancia iA1 que es instancia de A, ¿en que circunstancias tiene los slots slA2 y slA4?
- ¿Es la relación r A.A transitiva? ¿Lo podría ser la relación r_B.A?
- ¿Puede tener la relación r_A2.A4 slots heredables?
- ¿Como habría que definir la relación r_B.A para que el slot slA4 se pudiera obtener por herencia en las instancias de B? ¿y para obtener el slot slB en las instancias de A?

4. (1 punto) Dada la siguiente red bayesiana y las probabilidades asociadas a cada nodo



A	P(A)	A	P(B A)		B	P(C A, B)		B	C	P(D B, C)	
			C	F		C	F			C	F
C	0.7	C	0.1	0.9	C	0.5	0.5	C	C	0.1	0.9
F	0.3	F	0.4	0.6	F	0.2	0.8	C	F	0.4	0.6
								F	C	0.3	0.7
								F	F	0.8	0.2

- Calcula la probabilidad de $P(D|\neg a)$

Las notas saldrán el día 3 de Julio

La revisión será vía correo electrónico a bejar@lsi.upc.edu hasta día 5 de Julio a mediodía