

Examen Final de IA

(9 de junio de 2008)

Duración: 3 horas

1. (2.5 puntos) Una empresa de correo postal quiere informatizar el reparto de cartas y ha decidido procesar las direcciones mediante un reconocedor OCR y un procesamiento posterior del texto (nombre y dirección) utilizando DCG.

Como ejemplo del formato de las direcciones tenemos las siguientes:

```
sent(1,['Antonio','López','Diagonal',325,2,1,08015,'Barcelona']).
sent(2,['Antonio','López','Pérez','Diagonal',325,2,1,08015,'Barcelona']).
sent(3,['Sr.','López','Diagonal',325,2,1,08015,'Barcelona']).
sent(4,['Antonio','López','Avda.','Diagonal',325,2,1,08015,'Barcelona']).
sent(5,['Jose','Antonio','López','Diagonal',325,2,1,08015,'Barcelona']).
sent(6,['Antonio','López','Diagonal',325,2,'',1,'a',08015,'Barcelona']).
sent(7,['Antonio','López','Diagonal',325,2,1,'(',08015,')','Barcelona']).
sent(8,['Jose','Antonio','López','y','Ana','López','Diagonal',325,2,1,08015,'Barcelona']).
```

Un primer intento de escritura de la DCG ha dado el siguiente resultado:

```
/* lexicón */

es_nombre_pila_fem('Ana').
es_nombre_pila_masc('Antonio').
es_municipio('Barcelona').
es_provincia('Barcelona').

es_apellido(X):- not(number(X)).
es_direccion(X):- not(number(X)).
es_numero(X):- number(X).
es_piso(X):- number(X).
es_puerta(X):- number(X).
es_codigo(X):- number(X).

/* gramática */

oracion --> persona, direccion.
persona --> mujer.
persona --> hombre.
mujer --> nombre_mujer, apellido.
hombre --> nombre_hombre, apellido.
nombre_mujer --> [N],{es_nombre_pila_fem(N)}.
nombre_hombre --> [N],{es_nombre_pila_masc(N)}.
apellido --> [N],{es_apellido(N)}.
direccion --> [D],{es_direccion(D)}, [N],{es_numero(N)},
    [Pi],{es_piso(Pi)}, [Pp],{es_puerta(Pp)},
    [C],{es_codigo(C)}, [M],{es_municipio(M)}.
```

Se pide:

- Indicad cuáles de las anteriores direcciones serían correctamente analizadas por la DCG.
- Indicad qué cambios haríais en la gramática y/o el lexicon para que todas las direcciones propuestas fueran correctamente analizadas.
- Deseamos que la gramática nos proporcione una representación de cada dirección como resultado del análisis. Indicad qué cambios deberíais realizar en la gramática y/o el lexicon para que la DCG nos devolviera el resultado deseado.

La representación (simplificada) sería la siguiente:

`direccion_postal(<lista de receptores>,<direccion>)`

donde cada receptor sería:

`persona(<lista de nombres>, <lista de apellidos>)`

la dirección sería:

`direc(<tipo_direccion>, <nombre_direccion>, <numero>, <piso>, <puerta>, <codigo_postal>, <ciudad>, <provincia>)`

donde el `tipo_direccion` puede ser calle, avenida, plaza, ... Los campos no rellenos toman el valor '?'

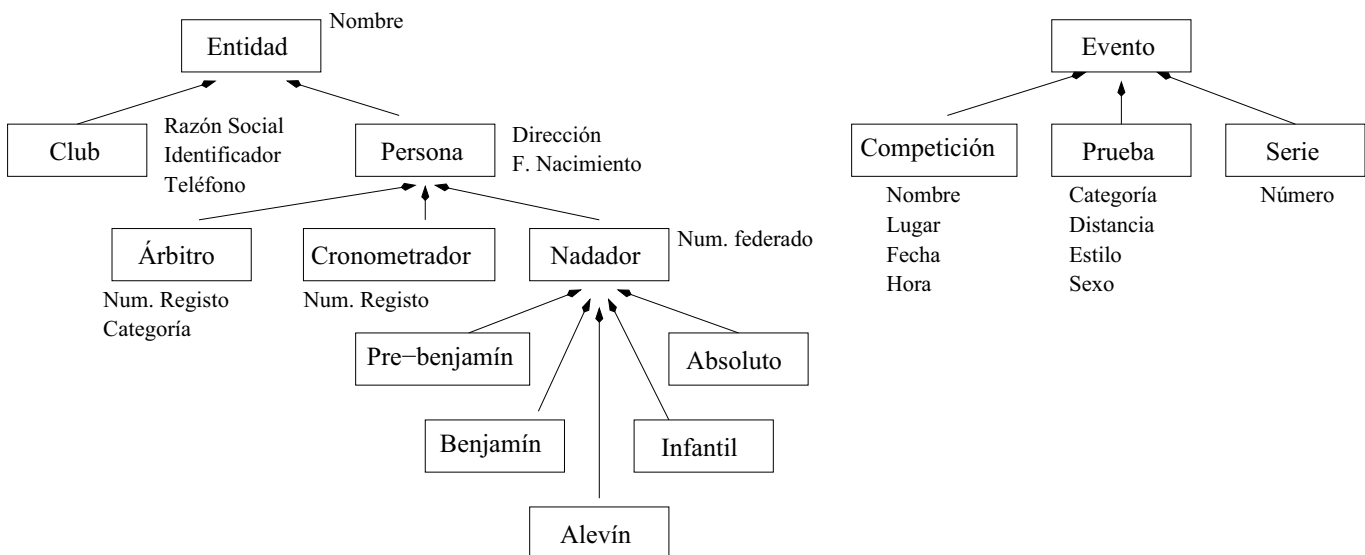
Por ejemplo, para la dirección número 8:

`sent(8,['Jose','Antonio','López','y','Ana','López','Diagonal', 325,2,1,08015,'Barcelona'])`.

la representación sería:

`direccion_postal([persona([Jose,Antonio], [López]), persona([Ana], [López])], direc(?, 'Diagonal', 325, 2, 1, 08015, 'Barcelona', ?))`

- (2.5 puntos) La Federación Catalana de Natación desea gestionar toda la información relacionada con las competiciones de natación mediante un sistema basado en frames. Los conceptos básicos que maneja son:



Cada competición tiene un nombre que la identifica ("Trofeo Sant Jordi", "Jornada n.3 Liga Infantil", ...) y está organizada en pruebas según categorías de los nadadores (pre-benjamín, benjamín, alevín, infantil, júnior, absoluta), estilo, distancia y sexo. Cada prueba puede tener una o más series, dependiendo del número de nadadores que tienen que nadarla (en cada serie puede nadar un máximo de 8 nadadores). Para cada competición se asigna un jurado que consta de 4 árbitros (el juez árbitro, el juez de salidas, el juez de llegadas y el secretario) y N cronometradores (el jefe de cronometradores y entre 3 y 8 cronometradores).

- a) Define las relaciones necesarias para identificar el juez árbitro de cada competición, las pruebas de cada competición y los nadadores de cada serie. Esta última relación debe controlar que no se asigne a una serie un nadador de un club que no figura como club participante en la competición correspondiente (añade lo que sea necesario).
 - b) Define un método, con un parámetro, que nos permita obtener los nombres de todos los nadadores de una determinada categoría que pertenecen a un club concreto y por cada nadador nos indique los nombres de las competiciones en las que ha participado. Justifica razonadamente donde es más adecuado ubicar el método.
 - c) Para cada nadador se desea asociar como número de teléfono el número del teléfono del club al cual pertenece. Añade todo lo que sea necesario para que el sistema pueda realizar la inferencia adecuada. La Federación no permite que un nadador pertenezca a más de un club. En caso de que sí lo permitiera, ¿la solución que propones seguiría siendo válida?
3. (4 puntos) Los ingenieros de *Nontiendo* están deseosos de alcanzar a la competencia con su nueva consola *Nontiendo güi* y están pensado crear el programa *güi feet* que permitirá a cualquier usuario alcanzar la forma física que desea.

El programa evaluará el perfil del usuario a partir de la información que pueda recolectar, como por ejemplo todas las actividades físicas que realiza en el trabajo (sentado, de pie, desplazamientos a pie, movimientos repetitivos, levantamiento de peso, esfuerzos musculares, ...), actividades fuera del trabajo (estáticas (televisión, lectura, ...), tareas domésticas (planchar, barrer, fregar, colada, ...), desplazamientos (compra a pie, paseos, paseo de mascotas, ...). De estas actividades nos interesa su frecuencia y duración. También obtendrá información sobre su salud como por ejemplo el peso, la edad, la presión sanguínea, problemas músculo-esqueléticos (dolor de espalda, articulaciones, cervicales, ...) , dieta (consumo de fruta, abuso de sal, picar entre horas, ...).

Aparte de esta información se necesitarán saber los objetivos del programa que se ha de crear, como por ejemplo mantenimiento, ponerse en forma, rebajar peso, musculación, flexibilidad, equilibrio, ... Así como del tiempo diario del que se dispone para el entrenamiento (al menos 30 minutos).

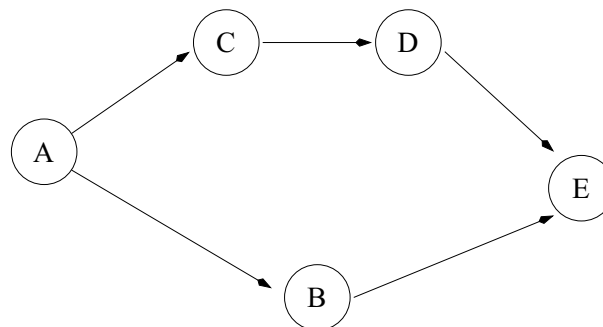
El sistema posee un conjunto de ejercicios encaminados a los diferentes objetivos que puede desear el usuario (cada ejercicio puede tener varios objetivos). Estos ejercicios tienen un conjunto de características y restricciones como por ejemplo el número de calorías que se queman por cantidad de tiempo, duración (mínima, máxima), número de repeticiones (mínimas, máximas), los grupos musculares que se ejercitan, si están contraindicados para alguna condición del usuario (presión alta, problemas musculares, ...), si no están indicados para ciertas edades. También tenemos información sobre los ejercicios que combinan mejor con cada ejercicio. La dificultad de los ejercicios (moderada, normal, difícil) puede estar ligada al propio ejercicio, al número de repeticiones que se realizan o a la condición física del usuario (hacer 5 abdominales puede ser fácil para un usuario con una condición física normal y sin sobrepeso, pero difícil para alguien con sobrepeso).

El sistema debe generar un programa de entrenamiento para una semana, creando para cada día una secuencia de ejercicios adecuada al tiempo del que se dispone. Los ejercicios deben elegirse según las restricciones físicas/médicas del usuario. Estos ejercicios deben ir encaminados principalmente al objetivo que ha indicado el usuario, incrementando su dificultad gradualmente dependiendo de las condiciones de partida del usuario. Debe haber un mínimo de ejercicios diferentes en cada sesión y durante la semana.

- a) Identifica qué conceptos forman los datos de entrada del problema. Representa gráficamente estos conceptos mediante una red de frames incluyendo los atributos mas relevantes y sus relaciones, tanto las taxonómicas, como las no taxonómicas que creas que son necesarias. ¿Algún atributo del concepto ejercicio necesitaría un demon? ¿Cuál? ¿Qué haría?

- b) Identifica los conceptos que forman la solución del problema. Representa gráficamente estos conceptos mediante una red de frames y sus relaciones, tanto las taxonómicas como las relaciones no taxonómicas que creas que son necesarias.
- c) Hemos identificado dos problemas. El primero consistirá en clasificar al usuario en un conjunto de perfiles básicos a partir de analizar la información que nos da, por ejemplo tendríamos el joven sano activo, el joven sedentario, el adulto sano sedentario, el adulto con sobrepeso, ... Evidentemente se trata de un problema de análisis que se puede resolver mediante clasificación heurística. Describe cómo encajarían las diferentes fases en la resolución de este problema indicando ejemplos de reglas de abstracción de datos y de asociación heurística. ¿Se podría hacer la fase de asociación heurística mediante el formalismo de redes bayesianas? Justifica tu respuesta.
- d) El segundo problema es la construcción del programa de entrenamiento. Es un problema de síntesis que podríamos resolver mediante proponer y aplicar. Define un conjunto de operadores que permitan resolver el problema indicando qué harían, las restricciones globales y específicas que deberían tener en cuenta para aplicarse y los criterios de evaluación que determinan la bondad de cada operador.

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



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

Calcula la probabilidad de $P(E|\neg a, b)$ mediante el algoritmo de eliminación de variables

Las notas saldrán el día 26 de junio

La revisión será vía correo electrónico a bejar@lsi.upc.edu hasta el día 27 de junio a mediodía