



Universidad
Rey Juan Carlos

ESCUELA TÉCNICA SUPERIOR
DE INGENIERÍA INFORMÁTICA

Exámenes+soluciones

INTELIGENCIA ARTIFICIAL

CURSO 2022-2023

**Autores: Sascha Ossowski
Holger Billhardt
Alberto Fernández-Gil**



Copyright (c) 2023 Sascha Ossowski, Holger Billhardt, Alberto Fernández-Gil. Esta obra está bajo la licencia CC BY-SA 4.0, [Creative Commons Atribución-CompartirIgual 4.0 Internacional](https://creativecommons.org/licenses/by-sa/4.0/).

APELLIDOS:

NOMBRE:

Duración: 10 minutos

No se permiten móviles, libros, apuntes, ni calculadora

La prueba consiste en 8 ejercicios cortos que se muestran a continuación. Para las frases de los ejercicios 1 a 6 hay que indicar en la columna de la derecha "V" o "F" para respuestas verdaderas o falsas, respectivamente. Para las preguntas de los ejercicios 7 y 8 hay que indicar en la columna de la derecha uno o varios números.

En cada ejercicio se indican los puntos que se suman si se acierta la respuesta o se restan si la respuesta es incorrecta, salvo en los ejercicios 7 y 8 donde una respuesta incorrecta no resta. Las preguntas no respondidas no suman ni restan. La puntuación máxima son 100 puntos.

1	10p	El Juego de las Damas es un juego bipersonal de suma nula.	
2	10p	Si h_1^* y h_2^* son dos funciones heurísticas optimistas, entonces $h_3^*(n) = h_1^*(n) + h_2^*(n) $ para todo nodo n , también es una función heurística optimista.	
3	10p	El algoritmo A^* degenera en una búsqueda <i>en profundidad</i> si $h^*(n) = 0$ para todos los nodos n del espacio de estados.	
4	10p	En el algoritmo <i>Minimax</i> , la función de evaluación e no se aplica nunca explícitamente a los nodos interiores del árbol de juego desarrollado por el algoritmo.	
5	10p	En el algoritmo <i>ExpectMinimax</i> , para representar un evento aleatorio del juego (como, p.e., la tirada de un dado), se añaden aleatoriamente nodos <i>Max</i> o nodos <i>Min</i> al árbol de juego	
6	10p	Cualquier CSP con restricciones k -áreas ($k > 2$) se puede representar mediante un CSP que contiene únicamente restricciones binarias, tal que ambas CSPs sean equivalentes (e.d., que tengan las mismas soluciones)	
7	20p	Considere un espacio de estados con 3 nodos: A (estado inicial), B , y C (estado final). Se puede llegar de A a B con un coste real de 4, y de B a C con un coste real de 3 (no hay más acciones posibles). Además hay una función heurística con $h^*(A) = 5$, $h^*(B) = x$, y $h^*(C) = 0$. ¿Qué valor debe tomar x para que h^* sea <i>optimista</i> pero no <i>consistente</i> ?	
8	20p	Considere un problema de satisfacción de restricciones (CSP) con las variables A y B , los dominios $D_A = [1, \dots, 5]$ y $D_B = [2, \dots, y]$, y la restricción $R_1: A = B - 1$. ¿Qué valor ha de tomar y (e.d., cual es la cota superior del dominio de B) para que el CSP sea <i>arco-consistente</i> ?	

APELLIDOS:

NOMBRE:

Duración: 10 minutos **No se permiten móviles, libros, apuntes, ni calculadora**

La prueba consiste en 8 ejercicios cortos que se muestran a continuación. Para las frases de los ejercicios 1 a 6 hay que indicar en la columna de la derecha "V" o "F" para respuestas verdaderas o falsas, respectivamente. Para las preguntas de los ejercicios 7 y 8 hay que indicar en la columna de la derecha uno o varios números.

En cada ejercicio se indican los puntos que se suman si se acierta la respuesta o se restan si la respuesta es incorrecta, salvo en los ejercicios 7 y 8 donde una respuesta incorrecta no resta. Las preguntas no respondidas no suman ni restan. La puntuación máxima son 100 puntos.

1	10p	El Juego de las Damas es un juego bipersonal de suma nula.	V
2	10p	Si h_1^* y h_2^* son dos funciones heurísticas optimistas, entonces $h_3^*(n) = h_1^*(n) + h_2^*(n) $ para todo nodo n , también es una función heurística optimista.	F
3	10p	El algoritmo A^* degenera en una búsqueda <i>en profundidad</i> si $h^*(n) = 0$ para todos los nodos n del espacio de estados.	F
4	10p	En el algoritmo <i>Minimax</i> , la función de evaluación e no se aplica nunca explícitamente a los nodos interiores del árbol de juego desarrollado por el algoritmo.	V
5	10p	En el algoritmo <i>ExpectMinimax</i> , para representar un evento aleatorio del juego (como, p.e., la tirada de un dado), se añaden aleatoriamente nodos <i>Max</i> o nodos <i>Min</i> al árbol de juego	F
6	10p	Cualquier CSP con restricciones k -áreas ($k > 2$) se puede representar mediante un CSP que contiene únicamente restricciones binarias, tal que ambas CSPs sean equivalentes (e.d., que tengan las mismas soluciones)	V
7	20p	Considere un espacio de estados con 3 nodos: A (estado inicial), B , y C (estado final). Se puede llegar de A a B con un coste real de 4, y de B a C con un coste real de 3 (no hay más acciones posibles). Además hay una función heurística con $h^*(A) = 5$, $h^*(B) = x$, y $h^*(C) = 0$. ¿Qué valor debe tomar x para que h^* sea <i>optimista</i> pero no <i>consistente</i> ?	0
8	20p	Considere un problema de satisfacción de restricciones (CSP) con las variables A y B , los dominios $D_A = [1, \dots, 5]$ y $D_B = [2, \dots, y]$, y la restricción $R_1: A = B - 1$. ¿Qué valor ha de tomar y (e.d., cual es la cota superior del dominio de B) para que el CSP sea <i>arco-consistente</i> ?	6

APELLIDOS:

NOMBRE:

Duración: 10 minutos **No se permiten móviles, libros, apuntes, ni calculadora**

Complete la tabla que se muestran a continuación indicando en la columna de la derecha "V" o "F" para respuestas verdaderas y falsas respectivamente. En cada pregunta se indican los puntos que se suman si se acierta la respuesta o se restan si la respuesta es incorrecta. Las preguntas no respondidas no suman ni restan. La puntuación máxima son 100 puntos.

1	10p	En lógica de descripciones \mathcal{ALC} : $\exists r.A \sqcap \exists r.\neg A$ es un concepto insatisfacible	
2	10p	En lógica de descripciones \mathcal{ALC} : $A \sqcap B \sqsubseteq B$	
3	10p	En lógica de descripciones \mathcal{ALC} : $\exists r.C \sqcap \forall r.C \equiv \top$	
4	10p	En lógica de descripciones \mathcal{ALC} : $A \sqcap \neg A \equiv \perp$	
5	10p	En RDFS (RDF Schema) se puede expresar que una propiedad es subpropiedad de otra	
6	10p	En OWL se puede expresar que una propiedad es transitiva	
7	10p	En SPARQL no se pueden realizar comparaciones numéricas	
8	10p	Para toda norma triangular (t-norma) T , $\forall x T(x, 0) \geq x$	
9	10p	No existe ninguna norma triangular (t-norma) K tal que $\forall x K(x, x) > x$	
10	10p	En lógica borrosa: $\forall x \mu_{A \vee B}(x) \geq \mu_A(x)$	

APELLIDOS:

NOMBRE:

Duración: 10 minutos **No se permiten móviles, libros, apuntes, ni calculadora**

Complete la tabla que se muestran a continuación indicando en la columna de la derecha "V" o "F" para respuestas verdaderas y falsas respectivamente. En cada pregunta se indican los puntos que se suman si se acierta la respuesta o se restan si la respuesta es incorrecta. Las preguntas no respondidas no suman ni restan. La puntuación máxima son 100 puntos.

1	10p	En lógica de descripciones \mathcal{ALC} : $\exists r.A \sqcap \exists r.\neg A$ es un concepto insatisfacible	F
2	10p	En lógica de descripciones \mathcal{ALC} : $A \sqcap B \sqsubseteq B$	V
3	10p	En lógica de descripciones \mathcal{ALC} : $\exists r.C \sqcap \forall r.C \equiv \top$	F
4	10p	En lógica de descripciones \mathcal{ALC} : $A \sqcap \neg A \equiv \perp$	V
5	10p	En RDFS (RDF Schema) se puede expresar que una propiedad es subpropiedad de otra	V
6	10p	En OWL se puede expresar que una propiedad es transitiva	V
7	10p	En SPARQL no se pueden realizar comparaciones numéricas	F
8	10p	Para toda norma triangular (t-norma) T , $\forall x T(x, 0) \geq x$	F
9	10p	No existe ninguna norma triangular (t-norma) K tal que $\forall x K(x, x) > x$	V
10	10p	En lógica borrosa: $\forall x \mu_{A \vee B}(x) \geq \mu_A(x)$	V

APELLIDOS:

NOMBRE:

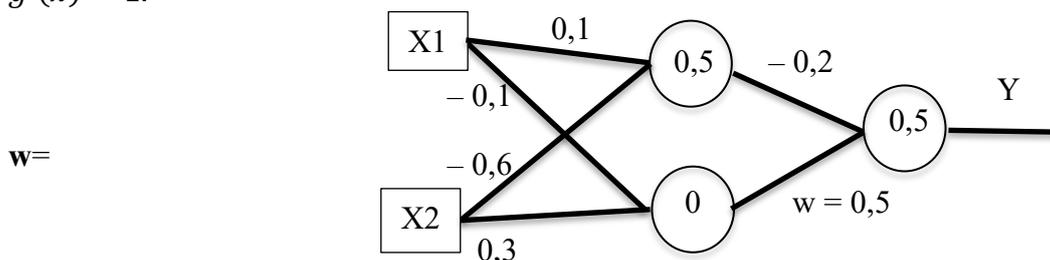
Duración: 10 minutos

No se permiten móviles, libros, apuntes, ni calculadora

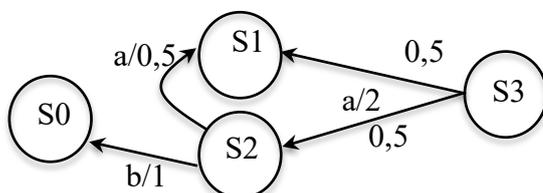
Complete la tabla de las 6 primeras preguntas indicando en la columna de la derecha "V" o "F" para respuestas verdaderas y falsas respectivamente. Por cada pregunta correcta se añaden 10 puntos y por cada errónea se restan 10. Las preguntas no respondidas no suman ni restan. En las preguntas 7 y 8 se pide un número (hasta 20 puntos y respuestas erróneas no restan)

1	10p	En un problema de aprendizaje supervisado, si hay dos ejemplos de entrenamiento contradictorios (con resultado distinto para la misma descripción de casos) entonces el aprendizaje no es posible.	
2	10p	Sea una neurona con función de activación por umbral y 2 entradas binarias (X_1 y X_2) que devuelve 1 si X_1 es 0 y 0 si X_1 es 1. En todas las posibles combinaciones de pesos, el sesgo de la neurona (su peso propio) debe tener un valor mayor a 0.	
3	10p	En entornos modelables con procesos de Decisión de Markov, una acción en un estado se denomina "determinista" sí y sólo sí se recibe siempre la misma recompensa por ella.	
4	10p	En un problema de regresión, se aprende clasificar casos en diferentes clases.	
5	10p	El aprendizaje de árboles de decisión es un aprendizaje simbólico.	
6	10p	El rango de valores de salida de una neurona con la función de activación RELU es $[0; \infty]$.	

- 7 20p: Supón la siguiente red neuronal con neuronas con función de activación por umbral y entradas binarias. Para el ejemplo de entrenamiento $\langle x_1=0; x_2=1; y=0 \rangle$ aplica el algoritmo de retropropagación y calcula el nuevo valor del peso w . Asume que la constante de aprendizaje $\alpha = 0,1$ y emplea como derivada de la función umbral la función $g'(x) = 1$.



- 8 20p: Supón el siguiente grafo de un proceso de decisión de Markov. a y b son acciones por las que un agente recibe recompensas (como indicado en el grafo). (Nota que la acción a en el estado S_3 puede llevar con probabilidad 0,5 al estado S_1 y con probabilidad 0,5 al estado S_2). Determina el valor de la función Q^* para la acción a en el estado S_3 (asumiendo que el factor de descuento γ es igual a 1).



$Q^*(S_3, a) =$

APELLIDOS:

NOMBRE:

Duración: 10 minutos

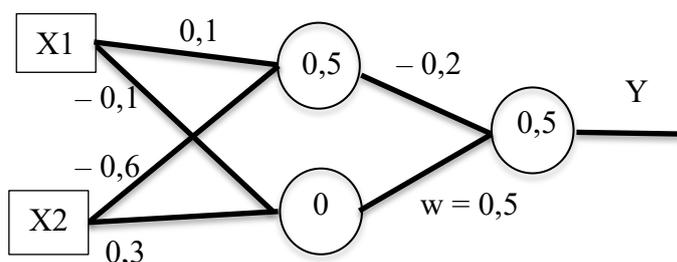
No se permiten móviles, libros, apuntes, ni calculadora

Complete la tabla de las 6 primeras preguntas indicando en la columna de la derecha "V" o "F" para respuestas verdaderas y falsas respectivamente. Por cada pregunta correcta se añaden 10 puntos y por cada errónea se restan 10. Las preguntas no respondidas no suman ni restan. En las preguntas 7 y 8 se pide un número (hasta 20 puntos y respuestas erróneas no restan)

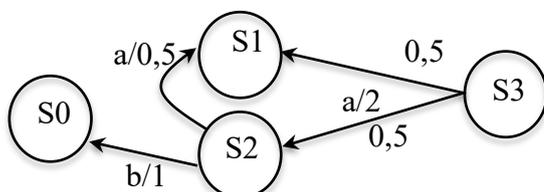
1	10p	En un problema de aprendizaje supervisado, si hay dos ejemplos de entrenamiento contradictorios (con resultado distinto para la misma descripción de casos) entonces el aprendizaje no es posible.	F
2	10p	Sea una neurona con función de activación por umbral y 2 entradas binarias (X1 y X2) que devuelve 1 si X1 es 0 y 0 si X1 es 1. En todas las posibles combinaciones de pesos, el sesgo de la neurona (su peso propio) debe tener un valor mayor a 0.	V
3	10p	En entornos modelables con procesos de Decisión de Markov, una acción en un estado se denomina "determinista" sí y sólo sí se recibe siempre la misma recompensa por ella.	F
4	10p	En un problema de regresión, se aprende clasificar casos en diferentes clases.	F
5	10p	El aprendizaje de árboles de decisión es un aprendizaje simbólico.	V
6	10p	El rango de valores de salida de una neurona con la función de activación RELU es $[0; \infty]$.	V

- 7 20p: Supón la siguiente red neuronal con neuronas con función de activación por umbral y entradas binarias. Para el ejemplo de entrenamiento $\langle x_1=0; x_2=1; y=0 \rangle$ aplica el algoritmo de retropropagación y calcula el nuevo valor del peso w . Asume que la constante de aprendizaje $\alpha = 0,1$ y emplea como derivada de la función umbral la función $g'(x) = 1$.

$w = 0,4$



- 8 20p: Supón el siguiente grafo de un proceso de decisión de Markov. a y b son acciones por las que un agente recibe recompensas (como indicado en el grafo). (Nota que la acción a en el estado S_3 puede llevar con probabilidad 0,5 al estado S_1 y con probabilidad 0,5 al estado S_2). Determina el valor de la función Q^* para la acción a en el estado S_3 (asumiendo que el factor de descuento γ es igual a 1).



$$Q^*(S_3, a) = 2,5$$

APELLIDOS:

NOMBRE:

Duración: 1 hora 40 mins No se permiten móviles, libros, apuntes, ni calculadora
Entregue los ejercicios en hojas separadas, incluidos en blanco
No se levante del asiento salvo para entregar el examen

Ejercicio 1: [35 puntos]

Considere un problema de satisfacción de restricciones (CSP) con las variables **A**, **B**, **C**, y **D**. Cada una de las 4 variables puede tomar los valores 1, 2, ó 3. Asimismo, existen las siguientes restricciones

$$R_1: A \neq B$$

$$R_2: A \neq C$$

$$R_3: C < B$$

$$R_4: D > B$$

- Dibuje el grafo que representa este CSP.
- Aplique el algoritmo de *chronological backtracking* con *forward checking* al CSP, indicando en cada nodo no sólo los valores de las variables asignadas, sino también los dominios de las variables no asignadas. Suponga que las variables se exploran en orden alfabético, y los valores de menor a mayor. ¿Qué solución se encuentra?
- Realice una propagación de restricciones en el CSP inicial aplicando el algoritmo de *arco consistencia*. Indique cómo se modifican los dominios cada vez que el algoritmo evalúa una restricción en una tabla como la que se indica abajo. Las restricciones han de ser evaluadas en ambas direcciones. Para obtener la máxima puntuación en este ejercicio, hay encontrar un orden de restricciones que minimice el número de filas de la tabla.

Restricción	D _A	D _B	D _C	D _D
	{1,2,3}	{1,2,3}	{1,2,3}	{1,2,3}
<i>R_i</i>
...

- ¿Cómo quedaría el árbol del apartado b) si se aplica el algoritmo *MAC* al CSP? Puede dibujar nuevamente el árbol generado por *MAC* (indicando en cada nodo los valores de las variables asignadas y los dominios de las variables no asignadas, y asumiendo el mismo orden de exploración que en el apartado b), o simplemente describir verbalmente los cambios que se producirían con respecto al árbol del apartado b), y su por qué.

APELLIDOS:

NOMBRE:

Duración: 1 hora 40 mins *No se permiten móviles, libros, apuntes, ni calculadora*
Entregue los ejercicios en hojas separadas, incluidos en blanco
No se levante del asiento salvo para entregar el examen

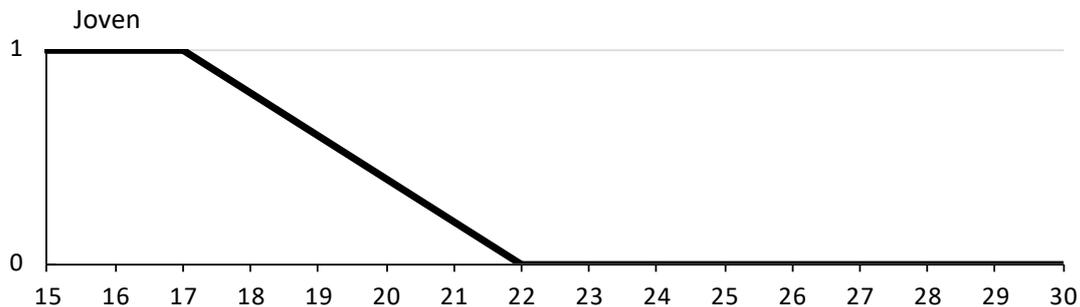
Ejercicio 2: [20 puntos]

Dados los siguientes nombres de conceptos: *Vino, Reserva, Crianza, Bodega, Barrica, Grande y Laboratorio*, y los nombres de rol *tiene* y *produce*, representar el siguiente conocimiento en lógica de descripciones \mathcal{ALC} (15 puntos) y en lógica de primer orden (5 puntos).

- 1) Los reservas y los crianzas son vinos
- 2) Las bodegas tienen barricas y vino
- 3) Las bodegas grandes tienen laboratorio
- 4) Las bodegas sólo producen vino (y al menos producen algo)
- 5) Protos es una bodega que produce crianza

Ejercicio 3: [15 puntos]

La siguiente figura muestra el conjunto borroso que representa el valor *joven* de la variable *edad* de los miembros de una peña de un pueblo castellano.



Se tienen los siguientes datos de dos personas de la peña: María tiene 18 años e Irene que tiene 19 años.

Obtener en qué grado son verdad cada una de las siguientes expresiones:

- a) María es joven, pero Irene no
- b) Si María es muy joven entonces Irene es joven

En caso de que sea necesario utilizar las siguientes funciones:

- T-norma = $W(x,y) = \text{Max}(0, x + y - 1)$
- T-conorma = $W^*(x,y) = \text{Min}(1, x + y)$
- Implicación = $J(x,y) = \text{Min}(1, 1 - x + y)$

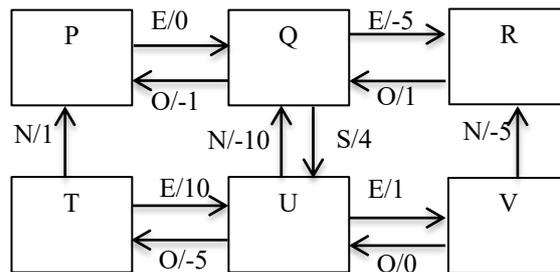
APELLIDOS:

NOMBRE:

*Duración: 1 hora 40 mins No se permiten móviles, libros, apuntes, ni calculadora
 Entregue los ejercicios en hojas separadas, incluidos en blanco
 No se levante del asiento salvo para entregar el examen*

Ejercicio 4: [30 puntos]

Considera un agente que se encuentra en un entorno modelado por un proceso de decisión de Markov determinista como se presenta en la siguiente figura:



Los estados del problema son P, Q, R, T, U y V, y las acciones que el agente puede realizar son movimientos al sur (S), oeste (O), este (E), y norte (N) según indicado en la figura. Cuando el agente realiza una acción, recibe la recompensa que se especifica. Supón que el agente no conoce este entorno (ni los estados, ni las acciones que existen, ni las recompensas correspondientes), pero **sí sabe que el entorno es determinista y cuando llega a un estado puede “ver” las acciones que puede realizar (no la recompensa que recibirá)**. El agente quiere aprender la política óptima utilizando el algoritmo Q-learning.

Los parámetros del algoritmo serán los siguientes:

- Los valores de Q se inicializan en 0.
- El agente siempre elige la acción de mayor valor Q. Si no hay mejor criterio, el agente realiza las acciones según el siguiente orden de preferencia: N,E,O,S
- El factor de descuento γ es igual a 1.
- El agente se encuentra inicialmente en el estado Q

- Simula los 10 primeros movimientos del agente aplicando el algoritmo y presentando la secuencia de acciones y la evolución de los valores de la función Q.
- Cual es la política óptima en este entorno? Simplemente especifica la política óptima, no es necesario obtenerla mediante ningún algoritmo.
- Supón que cuando el agente llega a un estado, no solo puede “ver” las acciones que puede realizar en este estado, sino también las recompensas asociadas a estas acciones. Comenta brevemente si el agente pudiera aprovechar esta información “adicional” en el proceso/algoritmo de aprendizaje y, si es así, cómo?.

APELLIDOS:

NOMBRE:

Duración: 1 hora 40 mins No se permiten móviles, libros, apuntes, ni calculadora
Entregue los ejercicios en hojas separadas, incluidos en blanco
No se levante del asiento salvo para entregar el examen

Ejercicio 1: [35 puntos]

Considere un problema de satisfacción de restricciones (CSP) con las variables **A**, **B**, **C**, y **D**. Cada una de las 4 variables puede tomar los valores 1, 2, ó 3. Asimismo, existen las siguientes restricciones

$$R_1: A \neq B$$

$$R_2: A \neq C$$

$$R_3: C < B$$

$$R_4: D > B$$

- Dibuje el grafo que representa este CSP.
- Aplique el algoritmo de *chronological backtracking* con *forward checking* al CSP, indicando en cada nodo no sólo los valores de las variables asignadas, sino también los dominios de las variables no asignadas. Suponga que las variables se exploran en orden alfabético, y los valores de menor a mayor. ¿Qué solución se encuentra?
- Realice una propagación de restricciones en el CSP inicial aplicando el algoritmo de *arco consistencia*. Indique cómo se modifican los dominios cada vez que el algoritmo evalúa una restricción en una tabla como la que se indica abajo. Las restricciones han de ser evaluadas en ambas direcciones. Para obtener la máxima puntuación en este ejercicio, hay encontrar un orden de restricciones que minimice el número de filas de la tabla.

Restricción	D _A	D _B	D _C	D _D
	{1,2,3}	{1,2,3}	{1,2,3}	{1,2,3}
<i>R_i</i>
...

- ¿Cómo quedaría el árbol del apartado b) si se aplica el algoritmo *MAC* al CSP? Puede dibujar nuevamente el árbol generado por *MAC* (indicando en cada nodo los valores de las variables asignadas y los dominios de las variables no asignadas, y asumiendo el mismo orden de exploración que en el apartado b), o simplemente describir verbalmente los cambios que se producirían con respecto al árbol del apartado b), y su por qué.

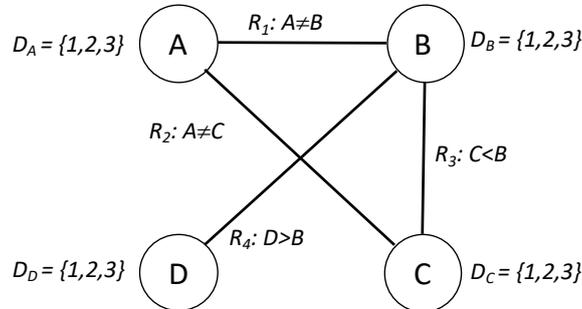
APELLIDOS:

NOMBRE:

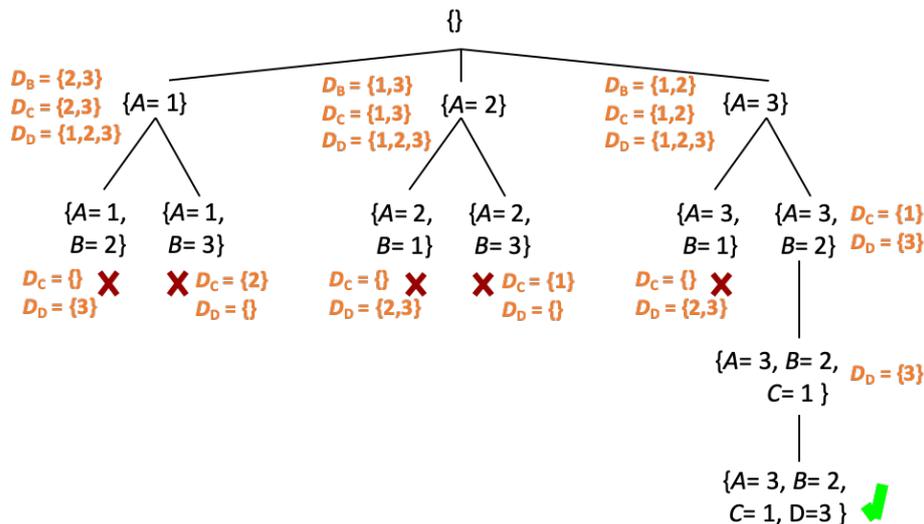
Duración: 1 hora 40 mins No se permiten móviles, libros, apuntes, ni calculadora
Entregue los ejercicios en hojas separadas, incluidos en blanco
No se levante del asiento salvo para entregar el examen

Solución:

a)



b)



Solución c)

Restricción	D_A	D_B	D_C	D_D
	$\{1,2,3\}$	$\{1,2,3\}$	$\{1,2,3\}$	$\{1,2,3\}$
R_3	-	$\{2,3\}$	$\{1,2\}$	-
R_4	-	$\{2\}$	-	$\{3\}$
R_3	-	$\{2\}$	$\{1\}$	-
R_1	$\{1,3\}$	$\{2\}$	-	-
R_2	$\{3\}$	-	$\{1\}$	-

Solución d)

Como se puede deducir de la solución del apartado c), tanto en el nodo $\{A=1\}$ como en el nodo $\{A=2\}$ el algoritmo de arco-consistencia detectaría que hay dominios que se quedan sin valores, por lo que en el árbol generado por el algoritmo MAC estos nodos no tendrían sucesores. En el nodo $\{A=3\}$ el algoritmo de arco-consistencia quedaría con $D_B = \{2\}$, $D_C = \{1\}$, y $D_D = \{3\}$ (ver apartado c)), por lo que todos los nodos de la rama $\{A=3\}$ sólo tienen un sucesor, hasta llegar a la solución.

APELLIDOS:

NOMBRE:

Duración: 1 hora 40 mins No se permiten móviles, libros, apuntes, ni calculadora
Entregue los ejercicios en hojas separadas, incluidos en blanco
No se levante del asiento salvo para entregar el examen

Ejercicio 2: [20 puntos]

Dados los siguientes nombres de conceptos: *Vino, Reserva, Crianza, Bodega, Barrica, Grande y Laboratorio*, y los nombres de rol *tiene* y *produce*, representar el siguiente conocimiento en lógica de descripciones \mathcal{ALC} (15 puntos) y en lógica de primer orden (5 puntos).

- 1) Los reservas y los crianzas son vinos
- 2) Las bodegas tienen barricas y vino
- 3) Las bodegas grandes tienen laboratorio
- 4) Las bodegas sólo producen vino (y al menos producen algo)
- 5) Protos es una bodega que produce crianza

Solución:

- 1) Los reservas y los crianzas son vinos

Reserva \sqsubseteq Vino

Crianza \sqsubseteq Vino

(otra equivalente: Reserva \sqcup Crianza \sqsubseteq Vino)

$\forall x(\text{Reserva}(x) \rightarrow \text{Vino}(x))$

$\forall x(\text{Crianza}(x) \rightarrow \text{Vino}(x))$

(otra equivalente: $\forall x(\text{Reserva}(x) \vee \text{Crianza}(x) \rightarrow \text{Vino}(x))$)

- 2) Las bodegas tienen barricas y vino

Bodega $\sqsubseteq \exists \text{tiene.Barrica} \sqcap \exists \text{tiene.Vino}$

$\forall x(\text{Bodega}(x) \rightarrow \exists y(\text{tiene}(x,y) \wedge \text{Barrica}(y)) \wedge \exists z(\text{tiene}(x,z) \wedge \text{Vino}(z)))$

- 3) Las bodegas grandes tienen laboratorio

Bodega \sqcap Grande $\sqsubseteq \exists \text{tiene.Laboratorio}$

$\forall x(\text{Bodega}(x) \wedge \text{Grande}(x) \rightarrow \exists y \text{ tiene}(x,y) \wedge \text{Laboratorio}(y))$

- 4) Las bodegas sólo producen vino (y al menos producen algo)

Bodega $\sqsubseteq \forall \text{produce.Vino} \sqcap \exists \text{produce.Vino}$

$\forall x(\text{Bodega}(x) \rightarrow \forall y(\text{produce}(x,y) \rightarrow \text{Vino}(y)) \wedge \exists z(\text{produce}(x,z) \wedge \text{Vino}(z)))$

- 5) Protos es una bodega que produce crianza

{Bodega(Protos), $\exists \text{produce.Crianza}(\text{Protos})$ }

Otra: {(Bodega $\sqcap \exists \text{produce.Crianza}$)(Protos)}

Bodega(Protos) $\wedge \exists y(\text{produce}(\text{Protos},y) \wedge \text{Crianza}(y))$

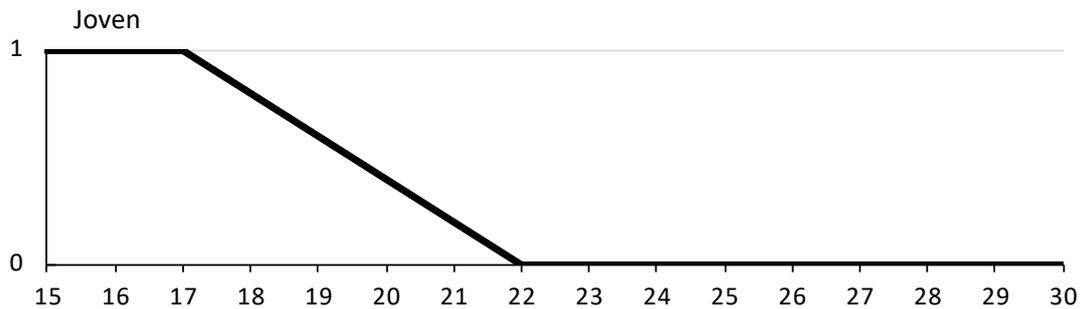
APELLIDOS:

NOMBRE:

*Duración: 1 hora 40 mins No se permiten móviles, libros, apuntes, ni calculadora
 Entregue los ejercicios en hojas separadas, incluidos en blanco
 No se levante del asiento salvo para entregar el examen*

Ejercicio 3: [15 puntos]

La siguiente figura muestra el conjunto borroso que representa el valor *joven* de la variable *edad* de los miembros de una peña de un pueblo castellano.



Se tienen los siguientes datos de dos personas de la peña: María tiene 18 años e Irene que tiene 19 años.

Obtener en qué grado son verdad cada una de las siguientes expresiones:

- María es joven, pero Irene no
- Si María es muy joven entonces Irene es joven

En caso de que sea necesario utilizar las siguientes funciones:

- T-norma = $W(x,y) = \text{Max}(0, x + y - 1)$
- T-conorma = $W^*(x,y) = \text{Min}(1, x + y)$
- Implicación = $J(x,y) = \text{Min}(1, 1 - x + y)$

Solución:

- María es joven, pero Irene no

$$\mu_{\text{Joven}}(18) = 0.8$$

$$\mu_{\text{Joven}}(19) = 0.6$$

$$\mu_{\neg\text{Joven}}(19) = 1 - \mu_{\text{Joven}}(19) = 1 - 0.6 = 0.4$$

$$\begin{aligned} \mu_{\text{Joven} \wedge \neg\text{Joven}}(18,19) &= T(\mu_{\text{Joven}}(18), \mu_{\neg\text{Joven}}(19)) = T(0.8,0.4) = \\ &= \text{Max}(0, 0.8 + 0.4 - 1) = 0.2 \end{aligned}$$

- Si María es muy joven entonces Irene es joven

$$\mu_{\text{Muy Joven}}(18) = \mu_{\text{Joven}}(18)^2 = 0.8^2 = 0.64$$

$$\begin{aligned} \mu_{\text{Muy Joven} \rightarrow \text{Joven}}(18, 19) &= J(\mu_{\text{Muy Joven}}(18), \mu_{\text{Joven}}(19)) = J(0.64, 0.6) = \\ &= \text{Min}(1, 1 - 0.64 + 0.6) = 0.96 \end{aligned}$$

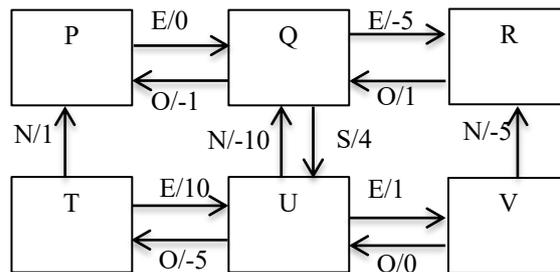
APELLIDOS:

NOMBRE:

Duración: 1 hora 40 mins No se permiten móviles, libros, apuntes, ni calculadora
Entregue los ejercicios en hojas separadas, incluidos en blanco
No se levante del asiento salvo para entregar el examen

Ejercicio 4: [30 puntos]

Considera un agente que se encuentra en un entorno modelado por un proceso de decisión de Markov determinista como se presenta en la siguiente figura:



Los estados del problema son P, Q, R, T, U y V, y las acciones que el agente puede realizar son movimientos al sur (S), oeste (O), este (E), y norte (N) según indicado en la figura. Cuando el agente realiza una acción, recibe la recompensa que se especifica. Supón que el agente no conoce este entorno (ni los estados, ni las acciones que existen, ni las recompensas correspondientes), pero **sí sabe que el entorno es determinista y cuando llega a un estado puede “ver” las acciones que puede realizar (no la recompensa que recibirá)**. El agente quiere aprender la política óptima utilizando el algoritmo Q-learning.

Los parámetros del algoritmo serán los siguientes:

- Los valores de Q se inicializan en 0.
- El agente siempre elige la acción de mayor valor Q. Si no hay mejor criterio, el agente realiza las acciones según el siguiente orden de preferencia: N,E,O,S
- El factor de descuento γ es igual a 1.
- El agente se encuentra inicialmente en el estado Q

- Simula los 10 primeros movimientos del agente aplicando el algoritmo y presentando la secuencia de acciones y la evolución de los valores de la función Q.
- Cual es la política óptima en este entorno? Simplemente especifica la política óptima, no es necesario obtenerla mediante ningún algoritmo.
- Supón que cuando el agente llega a un estado, no solo puede “ver” las acciones que puede realizar en este estado, sino también las recompensas asociadas a estas acciones. Comenta brevemente si el agente pudiera aprovechar esta información “adicional” en el proceso/algoritmo de aprendizaje y, si es así, cómo?.

Solución:

- Se puede representar el resultado mediante la siguiente tabla:

APELLIDOS:

NOMBRE:

***Duración: 1 hora 40 mins No se permiten móviles, libros, apuntes, ni calculadora
Entregue los ejercicios en hojas separadas, incluidos en blanco
No se levante del asiento salvo para entregar el examen***

Mov.	Est. actual	Acción	Est. nuevo	Q(Q,E)	Q(Q,O)	Q(Q,S)	Q(R,O)	Q(P,E)	Q(U,N)	Q(U,O)	Q(U,E)	Q(V,O)	Q(V,N)
	Q			0	0	0							
1	Q	E	R	-5			0						
2	R	O	Q				1						
3	Q	O	P		-1			0					
4	P	E	Q					0					
5	Q	S	U			4			0	0	0		
6	U	N	Q						-6				
7	Q	S	U			4							
8	U	E	V								1	0	0
9	V	N	R										-4
10	R	O	Q				5						

- b) Hay varias políticas óptimas. Una posible sería la siguiente: $\pi^*(P)=E$; $\pi^*(R)=O$; $\pi^*(Q)=S$; $\pi^*(U)=O$; $\pi^*(T)=E$; $\pi^*(V)=O$
- c) Si el agente cuenta con la información de la recompensa de cada acción cuando llega a un estado, podría usar esta información para inicializar los valores de la función Q justamente a esta recompensa. De esta forma, el agente aprendería más rápido.