

Examen Final de GRAU-IA

(15 de junio de 2020)

Duración: 2 horas 30 minutos

Este examen ha sido **adaptado para que solo haga falta un editor de texto**. No teneis que hacer diagramas, solo escribir el texto con vuestras respuestas en un fichero de texto (.txt/.doc/.odt).

Al final de examen teneis que **entregar un solo fichero de texto** (.txt/.doc/.odt) a través del Racó. Por favor, **no entregueis hojas escritas a mano** y luego fotografiadas o escaneadas, ya que en muchos casos cuestan de leer.

1. (2,5 puntos) La Agència de Gestió d'Ajuts Universitaris (AGAUR) de la Generalitat de Catalunya quiere organizar mejor la asignación de ayudas destinadas a proyectos de investigación orientados a mejorar la gestión sanitaria del COVID-19. Para ello AGAUR planea realizar C convocatorias en los próximos meses. Cada convocatoria de ayudas tiene un presupuesto a repartir entre las peticiones concedidas. A estas convocatorias los grupos de investigación pueden presentar peticiones (la misma petición puede presentarse a hasta tres convocatorias distintas, pero solo se puede conceder una vez). Una petición tiene un presupuesto solicitado y AGAUR le asigna una prioridad entre 1 y 3 dependiendo del historial del grupo investigador. Hay un total de P peticiones que han pasado ya un filtro previo realizado por un grupo de expertos, asegurando que todos los proyectos de investigación son relevantes y de calidad.

AGAUR quiere asignar la mayor cantidad de dinero posible en cada convocatoria, pero haciendo que globalmente la proporción de grupos con concesiones de ayuda cumpla las siguientes restricciones: como máximo el 30% de concesiones ha de ser para los grupos de prioridad 2 y como máximo el 10% para los de prioridad 3. Obviamente la suma de los presupuestos solicitados de las peticiones concedidas en una convocatoria no puede superar el presupuesto de esa convocatoria.

En los siguientes apartados se proponen diferentes alternativas para algunos de los elementos necesarios para plantear la búsqueda (solución inicial, operadores, función heurística, ...). El objetivo es comentar la solución que se propone, analizando si la técnica escogida es adecuada para este problema, si cada uno de los elementos de la solución son correctos o no (cada uno por separado y en conjunción los unos con los otros). Incluye un análisis de los costes algorítmicos y/o factores de ramificación allá donde sea necesario. Justifica tu respuesta.

- (a) Usar Hill-climbing. Para hallar la solución inicial recorreremos todas las convocatorias y asignamos tantas peticiones de prioridad 1 solicitadas a esa convocatoria como podamos, sin exceder su presupuesto. Como operadores utilizamos: *asignar* una petición a una convocatoria, respetando las restricciones de máximos de peticiones concedidas por prioridad; *mover* una petición asignada en una convocatoria C_i a otra convocatoria C_j , comprobando que no se supera el presupuesto máximo de C_j ; e *intercambiar* una petición P_x asignada a una convocatoria por otra petición P_y que se haya presentado a esa convocatoria pero que no haya sido asignada aún, comprobando que no se supera el presupuesto máximo de la convocatoria. Como función heurística usamos:

$$h1(n) = \frac{\sum_{\forall conv_j} \sum_{\forall pet_i \in conv_j} Presupuesto(pet_i)}{\sum_{\forall conv_j} Presupuesto(conv_j)}$$

donde $pet_i \in conv_j$ significa que la petición i ha sido concedida (asignada) en la convocatoria j .

- (b) Usar algoritmos genéticos. Codificamos el problema con $C \times P$ bits donde cada grupo de P bits corresponde a una convocatoria de manera que un bit a '1' significa que la petición ha sido concedida en esa convocatoria. Las soluciones iniciales las generamos utilizando el método del apartado anterior. Como operadores genéticos usamos los operadores habituales de cruce y mutación. Como función heurística usamos:

$$h2(n) = \sum_{\forall conv_j} \left(Presupuesto(conv_j) - \sum_{\forall pet_i \in conv_j} Presupuesto(pet_i) \right)$$

donde $pet_i \in conv_j$ significa que la petición i ha sido concedida (asignada) en la convocatoria j .

2. (2,5 puntos) Para estimular la producción agraria nacional y reducir el potencial impacto de la crisis del COVID-19 sobre el sector, el Ministerio de Agricultura, Pesca y Alimentación quiere desarrollar un sistema basado en el conocimiento para aconsejar a los agricultores españoles sobre cuales son los cultivos más adecuados para las características de sus explotaciones.

Como datos de entrada el sistema recibe información sobre la extensión del terreno (en hectáreas), el nivel de alcalinidad del suelo (en pH), la permeabilidad del suelo (en cm/hora), las concentraciones de nitrógeno, fósforo y potasio en el suelo (en ppm), su sistema de riego (propio, compartido o ninguno), la precipitación anual promedia (en cm/m²), media de temperatura anual (Celsius), la temperatura media menor (del mes más frío, en Celsius) y la temperatura media mayor (del mes más cálido, en Celsius). El agricultor también ha de añadir información sobre el gasto máximo asumible (en euros) en maquinaria, mano de obra, semillas, fertilizantes y en seguros, y el beneficio bruto mínimo (en euros) que se quiere conseguir.

Con esta información el sistema deberá de construir su recomendación. Según el grupo de expertos consultados por el Ministerio, algunas de las características más relevantes a la hora de decidir los cultivos más adecuados serían las siguientes:

- La **climatología**, que se podría clasificar en *árida* (precipitación anual menor que 30cm/m²), *seca* (precipitación anual entre 30cm/m² y 60cm/m²), *mediterránea* (temperatura media menor entre los 0°C y los 18°C, temperatura media mayor por encima de los 10°C, precipitación anual entre 60cm/m² y 150cm/m²), *continental* (temperatura media menor por debajo de los 0°C, temperatura media mayor por encima de los 10°C) y *subtropical* (temperatura media menor por encima de los 18°C, precipitación anual mayor que 150cm/m²).
- La **fertilidad** del suelo, que se podría clasificar (a partir de la alcalinidad, permeabilidad y concentración de nutrientes del suelo) en los valores discretos del 1 al 5, que se calculan a partir de la función `int_fertilidad(nivel_PH, permeabilidad, nivel_nitrogeno, nivel_fosforo, nivel_potasio)`.
- El **tamaño del terreno**, que se podría clasificar en *pequeño* (extensión menor de 10 hectáreas), *mediano* (entre 10 y 100 hectáreas), *grande* (más de 100 hectáreas).
- El **acceso al agua**, que se podría clasificar en *muy bueno* (si la precipitación anual es superior a los 150cm/m² y no tiene ningun sistema de riego, o bien si la precipitación anual está entre los 100cm/m² y los 150cm/m² y tiene sistema de riego propio), *bueno* (si la precipitación anual está entre los 100cm/m² y los 150cm/m² y no tiene ningun sistema de riego, o bien si la precipitación anual es mayor que 60cm/m² y tiene sistema de riego propio), *regular* (si la precipitación anual es menor que 100cm/m² y no tiene ningun sistema de riego, o bien si la precipitación anual es menor que 60cm/m² y tiene sistema de riego compartido) o *malo* (si precipitación anual es menor que 60cm/m² y no tiene ningun sistema de riego).
- Los **gastos asumibles**, que se podrían clasificar en *bajos* (gasto en maquinaria y seguros menor de 100 euros por hectárea, gasto en mano de obra menor de 600 euros por hectárea, gasto en semillas y fertilizantes menor de 40 euros por hectárea), *medios* (gasto en maquinaria y seguros entre 100 y 500 euros por hectárea, gasto en mano de obra entre 600 y 1800 euros por hectárea, gasto en semillas y fertilizantes entre 40 y 120 euros por hectárea) o *altos* (gasto en maquinaria y seguros mayor de 500 euros por hectárea, gasto en mano de obra mayor de 1800 euros por hectárea, gasto en semillas y fertilizantes mayor de 120 euros por hectárea).
- La **productividad deseable**, que se puede clasificar en *normal* (beneficio bruto mínimo menor a 1000 euros por hectárea), *alta* (beneficio bruto mínimo entre los 1000 y los 3500 euros por hectárea) e *intensiva* (beneficio bruto mínimo mayor a 3500 euros por hectárea).
- El ratio **beneficio/coste** deseado, que se puede clasificar en *bajo* (el ratio entre beneficio bruto mínimo y la suma de gastos máximos por hectárea es menor a 1,5), *medio* (el ratio entre beneficio bruto mínimo y la suma de gastos máximos por hectárea esta entre 1,5 y 4) y *alto* (el ratio entre beneficio bruto mínimo y la suma de gastos máximos por hectárea es mayor que 4)).

A partir de estas características el sistema ha de poder determinar, en primer lugar, que tipo general de cultivo es el más adecuado de entre los siguientes: *cereales*, *hortalizas*, *legumbres* y *raíces/bulbos*. Para ello contamos con el conocimiento de los expertos que nos cuentan por ejemplo, que los cereales y las legumbres pueden crecer en terrenos con fertilidad menor que 3, se adaptan muy bien a climatología árida o seca, son cultivos con gastos asumibles bajos, con una productividad normal, y si el acceso al agua es malo son mejores los cereales que las legumbres. 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 mayor o igual a 3, 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, el sistema ha de generar una recomendación sobre cultivos concretos adecuados. El sistema puede recomendar uno o varios cultivos concretos, utilizando criterios como los siguientes (... indica que hay más criterios de los escritos a continuación):

- Para los *cereales* se puede recomendar plantar *trigo, cebada, avena, soja* o *centeno*. Soja y trigo son adecuados si se desea un ratio beneficio/coste alto con un tamaño de terreno pequeño o normal, pero necesitan una fertilidad mayor o igual a 2 para ser viables. Si se tiene un terreno muy grande, es mejor plantar cebada, avena o centeno. La avena y el centeno pueden crecer en terrenos con fertilidad de 1 o 2, y son adecuados si se espera un ratio beneficio/coste bajo o medio, ...
- 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 y el tomate ha de ser por lo menos 4, ...
- Para las *legumbres* se puede recomendar plantar *guisante, garbanzo, lenteja, judía* o *haba*. El garbanzo y la lenteja se dan mejor en climatología seca. En el caso del garbanzo y las habas la fertilidad del terreno puede ser menor o igual a 2, los guisantes y las judías requieren una fertilidad entre 2 y 3, y las lentejas requieren una fertilidad mayor o igual a 4. Los guisantes, judías y habas necesitan por lo menos un acceso a agua por lo menos regular, ...
- Para las *raíces/bulbos* se puede recomendar plantar *remolacha, zanahoria, patata, nabo* o *cebolla*. La remolacha y el nabo necesitan un acceso al agua bueno o muy bueno y se adaptan muy bien a la climatología continental. La patata y la zanahoria necesitan gastos asumibles altos. La cebolla necesita un nivel de fertilidad del terreno mayor o igual a 4, ...

La salida esperada del sistema es una lista de uno o más nombres de cultivos concretos que son adecuados para las características de sus explotaciones.

- Explica cómo resolverías este problema usando clasificación heurística, justificando si el tipo de solución que se pide requiere de refinamiento o no. Lista las variables que forman parte del Problema Abstracto (nombre y valores posibles). Lista las variables que forman parte de la Solución Abstracta (nombre y valores posibles).
 - Vamos a resolver las fases de la clasificación heurística mediante Reglas de Producción. Da al menos 6 ejemplos de reglas de Abstracción, al menos 6 ejemplos de reglas de Asociación Heurística, y al menos 1 ejemplo de regla de refinamiento (solo si hace falta en este problema). Para las reglas podeis usar la notación de alto nivel que usamos en los ejercicios de problemas (si ... entonces ...) o una notación más próxima a la del lenguaje CLIPS (siempre que queden claro los elementos del antecedente de la regla y del consecuente). Intenta que las reglas para cada fase sean diferentes unas de otras. Podemos suponer la existencia de funciones para calcular: máximos y mínimos de un conjunto de instancias, sumatorios y medias de un conjunto de instancias, la cardinalidad del número de instancias de una clase que cumplen una propiedad, o la cardinalidad de instancias enlazadas desde una entidad a través de una relación.
3. (5 puntos) El Departament de Salut está buscando un sistema inteligente que ayude a gestionar el envío inmediato de material bio-sanitario dentro de una área sanitaria. Para ello ha decidido crear una pequeña flota de vehículos eléctricos compuesta por mini-vans y motos eléctricas. A través de un servicio web diferentes entidades (hospitales, laboratorios de análisis, residencias de ancianos) podrán solicitar el envío de paquetes con material bio-sanitario: guantes, máscaras, equipos de protección individual (EPIs), gel hidroalcohólico, jabon desinfectante, toallas secantes de papel, pruebas PCR, muestras de sangre y material de desecho a incinerar.

Un planificador recibe las peticiones recibidas en los últimos 20 minutos (cada una de ellas incluye el material requerido, el lugar donde se ha de recoger el material y el lugar de destino en el que se ha de entregar el material) y crea un plan para servir las peticiones con los vehículos disponibles. Los vehículos salen de un parking municipal y después de hacer todos los transportes requeridos han de volver al parking para recargar sus baterías. Todos los paquetes son de una medida estandar, y los vehículos han sido preparados con contenedores especiales para transportar paquetes de esa medida. En el caso de las mini-vans disponen de entre 2 y 6 contenedores especiales refrigerados, mientras que las motos eléctricas disponen de un único contenedor especial que no está refrigerado. Hay materiales (como los PCRs y las muestras de Sangre) que solo se pueden transportar dentro de un contenedor refrigerado. Pero para aprovechar los viajes se permite a las mini-vans que lleven también materiales que no precisen refrigeración en sus contenedores refrigerados. Cada vehículo solo puede llevar un paquete de material por contenedor, y no puede coger más paquetes si tiene todos sus contenedores llenos.

El plan generado ha de mostrar claramente como se mueve cada vehículo de un lugar a otro, y lo que van recogiendo y entregando en cada lugar.

- Describe el dominio (incluyendo predicados, acciones, etc...) usando PDDL. Da una explicación razonada de los elementos que has escogido. Explica el significado de cada predicado y sus parámetros (por ejemplo: "(encima ?X ?Y) = el elemento ?X está encima del elemento ?Y"). Explica también si añades predicados para filtrar posibles asignaciones o para guiar mejor al planificador hacia el objetivo. Ten en cuenta que el modelo del dominio ha de poderse extender a más o menos paquetes, lugares de recogida y entrega y vehículos (con más o menos contenedores de transporte).

- (b) Nos han proporcionado una tabla como ejemplo de las peticiones que el sistema recibe cada 20 minutos, con una serie de paquetes de material (donde el * indica que ha de transportarse refrigerado) que hay que transportar entre diferentes lugares: dos almacenes, tres hospitales, tres residencias de ancianos, un laboratorio de análisis y una incineradora. También nos dice que al inicio de ese intervalo de 20 minutos tenemos solo dos vehículos eléctricos disponibles: la *Moto1* (con capacidad para 1 paquete, sin refrigeración) y la *Van2* (con capacidad para transportar hasta 3 paquetes, tanto que requieran refrigerado como no). Ambos vehículos se encuentran aparcados en un parking municipal y, después de transportar todas las peticiones, han de volver al parking para poder recargar sus baterías eléctricas. Se nos pide que el planificador genere un plan en el que los vehículos disponibles sirvan todos los pedidos, transportando cada paquete de material del lugar de origen al de destino.

id_pedido	origen	destino	paquete (* = refrigerado)
p1	Almacen1	ResidenciaSants	Guantes
p2	Almacen1	HospitalSantPau	EPIs
p3	Almacen1	ResidenciaCollblanc	Mascaras
p4	Almacen2	HospitalVallHebron	JabonesDesinfectantes
p5	HospitalClinic	Incineradora	Deshechos
p6	Almacen1	HospitalClinic	Mascaras
p7	Almacen1	ResidenciaTresTorres	EPIs
p8	Almacen2	ResidenciaTresTorres	GelesHidroalcohólicos
p9	Almacen1	ResidenciaSants	ToallasPapel
p10	Almacen2	HospitalVallHebron	PCRs*
p11	HospitalClinic	Laboratorio	PCRs*
p12	HospitalSantPau	Laboratorio	MuestrasSangre*
p13	Almacen1	HospitalSantPau	Mascaras
p14	Almacen2	HospitalVallHebron	EPIs
p15	Almacen2	ResidenciaCollblanc	GelesHidroalcohólicos
p16	HospitalVallHebron	Laboratorio	PCRs*

Describe este problema usando PDDL. Da una breve explicación de cómo modelas el problema.

Las notas se publicarán el día **29 de junio**.