

EXAMEN

TIEMPO: 135 MINUTOS

- La prueba contiene 60 puntos en partes.
- ¡No separe el cuadernillo!
- La prueba es sin apuntes ni equipos electrónicos. Puede usar una calculadora no programable con funciones básicas.

Pregunta 1: Conceptos (12 puntos)

Responda cada una de las siguientes preguntas:

1. (2p) En una frase: ¿Qué diferencia conceptual existe entre una solución y una política en MDPs? **Respuesta:** Una solución es una serie específica de decisiones para una realización del sistema y una política es una función que recomienda decisiones para cada estado y etapa del sistema.

2. (2p) En una frase: ¿Qué diferencia conceptual hay entre una política dinámica reactiva y una proactiva? **Respuesta:** Una política reactiva toma en cuenta el estado del sistema al momento de tomar decisiones, pero no considera información probable del futuro. Una política proactiva si lo hace.

3. (2p) ¿Es la siguiente frase verdadera para un MDP de maximización?: “El valor esperado obtenido ejecutando una política de horizonte rodante es mayor o igual al de una política miope.” Justifique. **Respuesta:** Falso. Ambas son heurísticas sin garantía.

4. (2p) En una frase: ¿Qué ventaja posee una política dinámica de horizonte rodante estocástico (SR) contra una de horizonte rodante determinístico (DR)? **Respuesta:** La política SR ajusta sus decisiones (de alguna manera) considerando la variabilidad del proceso estocástico por ejecutarse, DR sólo planifica decisiones con valores esperados de datos futuros.

5. (4p) Considere una función subaditiva $g : \mathbb{S} \times \mathbb{X} \rightarrow \mathbb{R}$ tal que $\max_{x \in \mathbb{X}} g(s, x)$ es finito $\forall s \in \mathbb{S}$. Demuestre que $f(s) = \max\{y \in \text{argmax}_{x \in \mathbb{X}} g(s, x)\}$ es no creciente en $s \in \mathbb{S}$. **Respuesta:** Sean $s_1 \geq s_2$ dos valores en \mathbb{S} .

- Por definición de $f(s)$ se cumple que $g(s_2, f(s_2)) \geq g(s_2, x)$ para todo $x \in \mathbb{X}$. En particular para $x' \geq f(s_2)$ se cumple que $g(s_2, f(s_2)) \geq g(s_2, x')$.
- Por subaditividad se cumple además que: $g(s_1, x') + g(s_2, f(s_2)) \leq g(s_2, x') + g(s_1, f(s_2))$

Combinando ambos, debe ocurrir que $g(s_1, x') \leq g(s_1, f(s_2))$. Es decir

$$f(s_2) \leq \text{argmax}_{x \in \mathbb{X}} g(s_1, x) \leq f(s_1).$$

Pregunta 2: Cortar un cable (10 puntos)

El dueño de una tienda de ferretería compró un cable de fibra óptica de $B \in \mathbb{Z}_+$ metros de largo. Hoy debe cortarlo en varios trozos de largo entero que venderá mañana. El precio de mercado para un trozo de y metros de largo es $\$p_y > 0$ (suponga que hay demanda suficiente para vender todo lo ofrecido a ese precio). Asuma también que $p_0 = 0$.

- (8p) Diseñe un algoritmo para cortar el cable de forma tal que el ferretero maximice sus ganancias.
- (2p) ¿Qué complejidad computacional posee el problema en función de B ?

Respuesta: Sea $V(B)$ el máximo valor de corte de un cable de B metros de largo.

Claramente $V(0) = 0$

Luego, para $b \in \{1, \dots, B\}$:

- $V(b) = \max_{x \in \{1, \dots, b\}} \{p(x) + V(b-x)\}$
- $x^*(b) = \operatorname{argmax}_{x \in \{1, \dots, b\}} \{p(x) + V(b-x)\}$

El algoritmo posee complejidad $\mathcal{O}(B^2)$

Pregunta 3: ¿Ir o no ir al doctor? (10 puntos)

Suponga que se enfermó y que desconoce si está infectado con un simple resfrío (no tratable), un virus severo (no tratable) o un cuadro bacterial grave, pero tratable con penicilina.

Si va al doctor, este le cobrará \$20 mil por la consulta y le hará un examen para diagnosticar su enfermedad. Asuma que la probabilidad de que el examen diagnostique un simple resfrío es de 50 %, un virus severo es de 20 % y un cuadro bacterial es de 30 %.

Si el resultado del examen diagnostica un cuadro bacterial, entonces el doctor le dará una receta por penicilina que usted puede comprar y utilizar (si lo desea). Este medicamento cuesta \$50 mil en la farmacia.

Sin embargo, el examen posee error. Si el examen diagnostica que es bacterial, entonces existe un 90 % de probabilidad de que sea bacterial y un 10 % de que sea un virus severo. Por otro lado, si el examen diagnostica un virus severo, entonces existe un 90 % de probabilidad de que sea un virus severo y un 10 % de que sea bacterial. Finalmente, el examen diagnostica un resfrío con certeza.

Un simple resfrío requiere 3 días de reposo en cama, lo que le costaría \$30 mil a su actual tasa salarial (\$10 mil por día). Un virus severo requiere 8 días de reposo y un cuadro bacterial no tratado lo fuerza a reposar por 16 días. Si el cuadro bacterial es tratado, entonces solo requiere de 2 días en cama.

Si usted está interesado en minimizar el costo esperado de este resfrío. ¿Debería ver al doctor? Justifique su respuesta.

Respuesta: Conviene ir al doctor. Resuelta en planilla de excel adjunta.

Pregunta 4: Opciones Financieras (16 puntos)

Ahora deberá aplicar MDPs de horizonte infinito para valorizar opciones financieras.

Una “opción de compra” (o *Call Option*) es un producto financiero. Consiste en un contrato que entrega a su dueño el derecho de comprar por una sola vez un activo específico a un valor c establecido por contrato. El dueño de la opción puede o no ejercer su derecho de compra en cualquiera de los momentos que establece el contrato para ejercer la opción. Si el dueño del contrato ejerce la opción cuando el activo realmente vale p , entonces gana $(p - c)$. Así, alguien racional nunca ejercería la opción si $c \geq p$. Sin embargo, no es claro que hacer cuando $p > c$, pues podría ser conveniente esperar por un mejor precio a futuro.

Suponga una empresa cuya acción tiene un precio p_0 al comienzo del año $t = 0$ que cambia año a año. Específicamente, si el precio al comienzo del año t es $p \in P$, entonces será $p' \in P$ al comienzo del año $t + 1$ con probabilidad $q(p, p') \in [0, 1]$, donde P define el conjunto de los precios posibles que la acción puede tomar.

Si es que usted posee una opción de compra por una acción de la empresa a $c = \$100$ al comienzo de $t = 0$ que puede ser ejecutada al comienzo de cualquier año $t \geq 0$:

- (7p) Formule un MDP de horizonte infinito cuya solución determine el mínimo precio al cual usted estaría dispuesto a vender su opción hoy. Suponga que usted es racional, neutral al riesgo y que conoce toda información probabilística relevante. Además, suponga un factor de descuento anual del dinero igual a $\lambda \in [0, 1]$.

Hint: Una correcta modelación del MDP debe definir espacio de estado, espacio de acción, probabilidades de transición de estado, retornos inmediatos y ecuaciones de Bellman.

- (2p) Si P es finito, formule un problema de optimización lineal que resuelva el MDP.
- (6p) Si $p_0 = 105$, $q(105, 105) = 50\%$, $q(105, 110) = 50\%$, $q(110, 105) = 75\%$, $q(110, 110) = 25\%$ y $\lambda = 0,8$ ¿Vendería su opción a \$6? Justifique.

Respuesta: El modelo para todo estado $p \in P$ sería:

$$V(p) = \max_{x \in \mathbb{X}} \{p, \lambda \cdot \sum_{p' \in P} q(p, p') V(p')\}.$$

el mínimo precio de venta sería el máximo provecho que se le puede sacar a la opción, es decir, $V(p_0)$.

El LP sería:

$$\min_{V \in \mathbb{R}^{|P|}} \left\{ \sum_{p \in P} V(p) : \quad V(p) \geq p, \forall p \in P; \quad V(p) \geq \lambda \cdot \sum_{p' \in P} q(p, p') V(p'), \forall p \in P \right\}.$$

Resuelta en planilla de excel adjunta, la opción debe venderse por lo menos en \$6,66.

Pregunta 5: Comprensión de lectura (6p)

En una o dos frases, responda las siguientes preguntas con respecto al método ADP de Value Function Approximation (VFA) para una minimización de costos. Para ello considere el texto: “Approximate Dynamic Programming by Practical Examples” de Martijn Mes y Arturo Pérez Rivera.

1. (3p) Con respecto a las simulaciones de entrenamiento de un algoritmo VFA basado en tablas: ¿Garantiza la convergencia de valores Q en las simulaciones de entrenamiento un buen desempeño del método en las corridas reales? Justifique.

Respuesta: Falso, puede converger a cualquier política sin que sea buena. Sucede frecuentemente en algoritmos enfocados 100% a explotación.

2. (3p) ¿Por qué en su ejemplo señalan que una política VFA posee automáticamente características de exploración sin necesidad de implementar ϵ -greedy?

Respuesta: Porque los valores Q de todos los estados parte en cero y naturalmente tenderá a ser visitados hasta explorarse y corregir su Q .

Pregunta 6: Análisis de proyectos (6p)

Escoja tres proyectos de grupos diferentes al suyo y comente brevemente:

- (3p) ¿Qué decisiones ayudó el grupo a planificar? ¿Cuál fue su objetivo? (sea específico)
- (3p) ¿Qué algoritmo de solución utilizaron? (exacto o no, de qué tipo)

Respuesta: Acá cualquier respuesta razonable y que diga la verdad es válida.