PONTIFICIA UNIVERSIDAD CATÓLICA DE CHILE
ESCUELA DE INGENIERÍA
Departamento de Ingeniería Industrial y de Sistemas
Departamento de Ingeniería de Transporte y Logística

Curso: ICS3105/ICT3473 - Optimización Dinámica
Semestre: 02-2018
Profesor: M. Klapp

Tarea 2

Fecha de entrega: Viernes 2 de Noviembre

Por favor conteste la tarea sólo o en pareja y suba documento digital a Siding (Word o Latex) por cuestionario web. Una respuesta por grupo. Penalidad por atraso: 1 punto el primer día, 2 puntos el segundo día, más de 3 días de atraso es un 1.

Pregunta 1 - 20 puntos

Considere que una persona realiza un viaje al Mall en auto este fin de semana para ir de compras. Para ello debe encontrar un espacio en el estacionamiento del mall, estacionar y caminar hacia la entrada de la tienda. Suponga que la persona no está en buen estado físico y que estacionar más cerca de la entrada les produce una inmensa satisfacción, pues caminaría menos.

Específicamente, considere la siguiente figura:

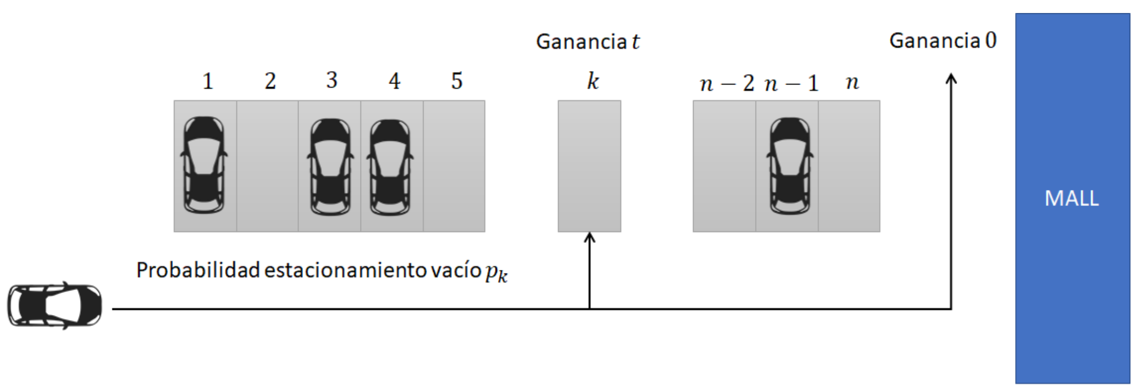

Figura 1: Esquema del estacionamiento

El estacionamiento del Mall se puede modelar como un pasillo que posee n lotes en línea para estacionar; el lote 1 es el más alejado del Mall y el lote n es el que está más cerca. La persona entra manejando al estacionamiento por el lote número 1 y avanza lote por lote. Si se encuentra frente al lote k y este se encuentra vacío entonces puede decidir estacionar en él y obtener una ganancia igual a k (más alta mientras más cerca del Mall), pero también puede decidir avanzar y probar suerte con el siguiente espacio. Si el lote está ocupado, está obligado a avanzar (no puede poner reversa). Si termina al final del pasillo se va para su casa sin haber podido estacionar y sin ganancia. Considere que la persona no puede ver los estacionamientos que todavía no ha visitado. Es decir, al estar frente al estacionamiento k no posee información de los estacionamientos $k+1, k+2, \ldots, n$. Sin embargo, sabe que la probabilidad de que el estacionamiento k esté vacío es igual a $p_k \in [0, 1]$.

- **(3 puntos)** Formule el problema que maximiza la ganancia esperada como un proceso de decisión markoviana.

- ◇ **(3 puntos)** Demuestre que existe una política de umbral τ en la cual usted dejará pasar la opción de estacionar para todo lote $k \leq \tau$. Para ello, muestre que si es que es óptimo no estacionar cuando el lote k está disponible, entonces también es óptimo no estacionar cuando el lote $t < k$ está disponible.

- ◇ **(3 puntos)** Demuestre que si $N \geq 2$, entonces nunca es óptimo estacionar en el primer lote si está disponible,

- ◇ **(3 puntos)** Sea $\tau < n$ el último lote de estacionamiento en el cual es óptimo esperar. Calcule el *value-to-go* si es que el lote τ no está disponible en función de p_k, τ y n.

- ◇ **(4 puntos)** Si $p_k = p$ (independiente de k), derive un algoritmo para obtener τ en función del resultado anterior. Luego compute τ para $n = 100$ y para los casos: $p = 20\,\%, 40\,\%$ y $60\,\%$. Discuta su resultado. Si lo desea, puede apoyarse de un software computacional.

- ◇ **(4 puntos)** Si $p_k = 1/k$, derive un algoritmo para obtener τ. Luego compute τ para $n = 100$.

Pregunta 2 - 15 puntos

- ◇ **(5 puntos)** Considere una función $f : \mathbb{R}^n \to \mathbb{R}^n$ y suponga que f es una contracción. Demuestre que el sistema de ecuaciones $f(x) = x$ para un $x \in \mathbb{R}^n$ posee solución única.

- ◇ **(10 puntos)** Considere un proceso de decisión markoviana con horizonte infinito con espacio de estados $\mathbb{S}$ finito; espacio de acciones $\mathbb{X}$ finito; vector de ganancia inmediata $r(s, x)$ para todo $s \in \mathbb{S}$ y para todo $x \in \mathbb{X}$; y probabilidades de transición de estados $p(j|i, x)$ para todo $i, j \in \mathbb{S}$ y todo $x \in \mathbb{X}$. El criterio a utilizar es el de maximizar la ganancia total descontada por un factor temporal λ.

  Considere una regla de decisión $d : \mathbb{S} \to \mathbb{X}$ cualquiera tal que define un vector de retornos esperados $r_d = \{r(s, d(s))\}_{s \in \mathbb{S}}$ y una matriz de transición de probabilidades $P_d = \{p(j|i, d(i))\}_{i,j \in \mathbb{S}}$.

  - Muestre que el vector de valores futuros para cada estado del sistema $V^d \in \mathbb{R}^{|\mathbb{S}|}$ dado por una politica estacionaria bajo la regla de decisión d es igual a $V^d = (I - \lambda P_d)^{-1} r_d$
  - Demuestre que para un vector $a \geq 0$, con $a \in \mathbb{R}^{|\mathbb{S}|}$, se cumple que $(I - \lambda P_d)^{-1} a \geq 0$ y $(I - \lambda P_d)^{-1} a \geq a$.

Pregunta 3 -15 puntos

Considere un proceso de decisión Markoviana de Horizonte Infinito con criterio de valor total descontado. El problema se ilustra en la siguiente figura y posee tres estados $\{0, 1, 2\}$, dos acciones en el estado 0 (azul, negra), dos acciones en el estado 1 (verde, roja) y una acción en el estado 0 (gris).

Por periodo de estadía en cada estado del sistema se generan ganancias inmediatas (independiente de la acción): en ell estado 0 se pierde \$2, en el estado 1 se pierde 5 y en el estado 2 se gana 1.

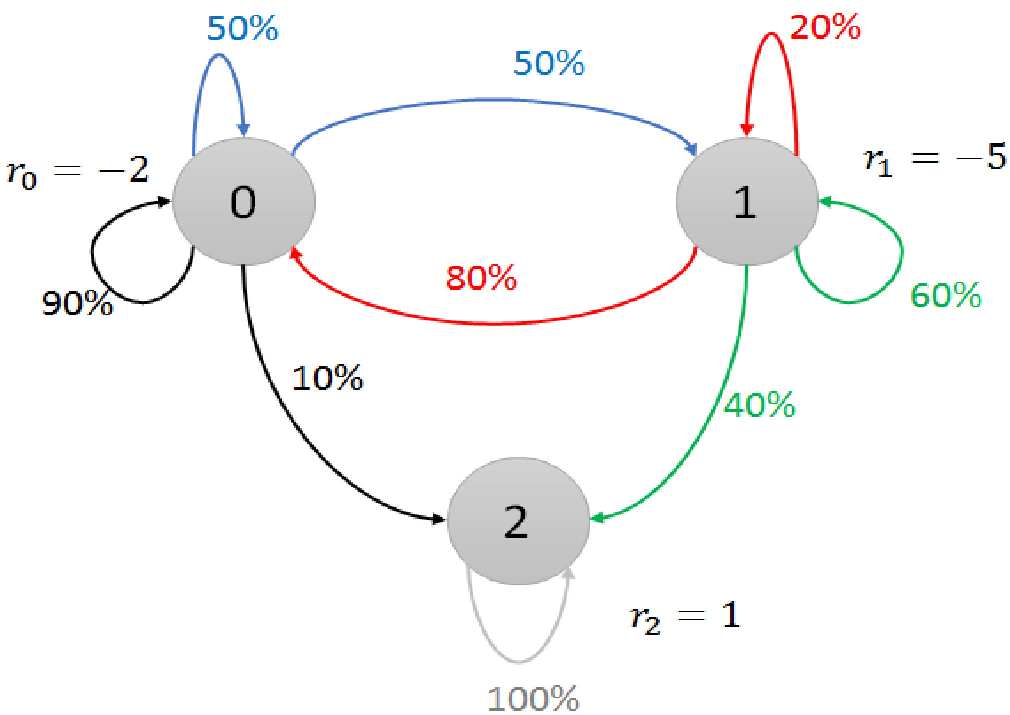

Figura 2: Esquema del estacionamiento

- ◇ **(3 puntos)** Formule el MDP correspondiente.
- ◇ Considere dos casos $\lambda = 0,1$ y $\lambda = 0,9$. Para cada caso:
  - **(3 puntos)** Resuelva el problema por iteración de valor al 1 % de precisión partiendo de $V = 0$. Muestre en un gráfico la evolución del *value-to-go* de cada estado en función de la iteración.
  - **(3 puntos)** Resuelva el problema por iteración de política partiendo de $d = (azul, verde, gris)$. Indique la evolución de su política en cada estado.
  - **(6 puntos)** Formule el problema LP DUAL equivalente y, utilizando un solver de optimización, resuelva el problema.

Es recomendable programar las últimas tres preguntas.

Pregunta 4 - 10 puntos

Considere el modelo de control dinámico de inventario de horizonte finito con 3 periodos discutido en la primera clase de MDPs. Si recuerda bien, este problema modelaba una bodega con capacidad $Q = 3$, tenía una función de demanda D por periodo igual a

$$\mathbb{P}(D = d) = \begin{cases} 1/4, & \text{si } d = 0 \\ 1/2, & \text{si } d = 1 \\ 1/4, & \text{si } d = 2. \end{cases},$$

el costo de inventario era \$1 por unidad en inventario y periodo y el costo de quiebre era \$8 por unidad faltante en cada periodo. Por otra parte el costo de reponer x unidades estaba dado por $\$4 + 2x$.

- ◇ **(1 puntos)** Formule el problema asociado de horizonte infinito con tasa de descuento λ por periodo.
- ◇ **(1 puntos)** Explique por qué debiese existir una política óptima con una estructura (s, S).
- ◇ **(4 puntos)** Resuelva el problema mediante iteración de valor. Muestre en un gráfico la evolución del *value-to-go* de cada estado en función de la iteración.
- ◇ **(4 puntos)** Resuelva el problema mediante iteración de política y explote el hecho de que existe una política óptima con estructura (s, S) en la búsqueda de políticas. Indique la evolución de su política en cada estado.

Es recomendable programar la últimas dos preguntas.