

Ejercicios 8: Procesos de decisión de Markov

Departamento de Informática / Department of Computer Science
Universidad Carlos III de Madrid

Inteligencia Artificial
Grado en Ingeniería Informática
2019/20

Ejercicio 1: MDP con tres estados

- ▶ Tenemos un agente con tres estados A, B y C donde C es el estado meta. En el estado A el agente puede llevar a cabo dos posibles acciones p y q , y en el estado B el agente puede tomar la acción q .
 - ▶ La ejecución de la acción p en el estado A mueve al agente al estado B con probabilidad 0.8, y permanece en el estado A con probabilidad 0.2.
 - ▶ La ejecución de la acción q en el estado A mueve al agente al estado C con probabilidad 0.1, y permanece en el estado A con probabilidad 0.9.
 - ▶ La ejecución de la acción q en el estado B mueve al agente al estado C con probabilidad 0.9, y mueve el agente al estado A el resto de las veces. Cada acción tiene un coste asociado de 1.
- ▶ Modelar formalmente el MDP.
- ▶ Especificar las ecuaciones de Bellman que actualizan los valores de los estados $V(A)$ y $V(B)$.
- ▶ Calcular el valor esperado $V(s)$ para cada estado.
- ▶ Calcular la política óptima.

Ejercicio 2: Tratamiento de enfermedad

Para el tratamiento de un cierto tipo de enfermedad se pueden ejecutar tres acciones: cirugía, suministrar un fármaco llamado Q o un fármaco llamado R.

- ▶ Cuando la enfermedad está en su estado inicial, se pueden aplicar los tres tratamientos:
 - ▶ Si se suministra el fármaco Q, la probabilidad de curación es $.3$, y no tiene efecto en el resto de los casos.
 - ▶ Si se suministra el fármaco R, la probabilidad de curación es $.3$, con probabilidad $.6$ el caso se agrava y con $.1$ no tendrá efecto.
 - ▶ Si se decide por la cirugía (S), la probabilidad de curación es $.5$. Con probabilidad $.4$ la enfermedad se mantendrá y con probabilidad $.1$ se agrava.
- ▶ Para tratar un paciente en estado agravado se puede utilizar fármacos (de cualquier tipo). El fármaco R produce la curación con probabilidad $.3$, y el tipo Q con probabilidad $.6$. De lo contrario, la enfermedad sigue en el mismo estado.

El coste del fármaco R es 6, el del Q es 10 y el de la cirugía es 100. Modele el MDP y calcule la política óptima para cada estado.

Ejercicio 3: Rover marciano

El sistema de planificación de un robot autónomo marciano optimiza sus acciones de forma que su consumo de energía (medido en unidades u.e.) sea mínimo. Tras realizar una misión nocturna es preciso llegar hasta un punto elevado de recarga, situado al lado contrario del cráter en cuyo borde se encuentra.

Para ello puede hacer tres cosas: deslizarse cuesta abajo, y luego subir hacia su objetivo, rodearlo por la derecha o rodearlo por la izquierda.



Ejercicio 3: Rover marciano

El sistema de planificación de un robot autónomo marciano optimiza sus acciones de forma que su consumo de energía (medido en unidades u.e.) sea mínimo. Tras realizar una misión nocturna es preciso llegar hasta un punto elevado de recarga, situado al lado contrario del cráter en cuyo borde se encuentra.

Para ello puede hacer tres cosas: deslizarse cuesta abajo, y luego subir hacia su objetivo, rodearlo por la derecha o rodearlo por la izquierda.

- ▶ Deslizarse cuesta abajo consume 2 u.e. Deslizarse lleva con certeza al fondo del cráter, pero la ascensión que debe realizar luego no siempre tiene éxito: una de cada cinco veces no se consigue y se cae de nuevo al fondo. Cada intento de ascensión supone un gasto de 3 u.e.
 - ▶ Rodear por la derecha lleva con total certeza al objetivo, pero es larga (consume 7 u.e.).
 - ▶ Rodear por la izquierda es un camino más corto (consume 4 u.e.), pero el terreno es traicionero y se cae al interior del cráter una de cada cuatro veces.
1. Modelar el problema con un Proceso de Decisión de Markov, especificando los estados y las probabilidades en forma de tabla.
 2. Calcular los valores para cada estado usando Iteración de Valores (basta 4 iteraciones)
 3. Calcular la política óptima para el punto de partida. ¿Cuál es el consumo esperado de energía para el recorrido?

Ejercicio 4: Seguridad informática

El sistema inteligente de gestión de seguridad informática ha detectado un virus. En este caso debe decidir qué hacer, con la intención de identificar el virus concreto y eliminarlo en el menor tiempo posible.

Una posibilidad es ejecutar un antivirus, operación que tarda 10 min en ejecutarse. Cada pasada del antivirus sobre un virus no identificado tiene una probabilidad del 20 % de eliminarlo (E) y un 30 % de identificar el tipo de virus (I) pero no eliminarlo. Si se ejecuta cuando el virus está identificado, sigue teniendo el 20 % de eliminarlo. Otra posibilidad es llamar al informático, que tarda 25 min en realizar su tarea. En este caso, si el virus está identificado, lo elimina con un 70 % de probabilidad. Si no lo está, lo elimina sólo con un 10 % , y lo identifica con un 70 %.

1. Representa el problema con un MDP, especificando claramente estados, transiciones, costes, y a qué probabilidades corresponden cada uno de los datos anteriores.
2. Escribir las ecuaciones de Bellman para cada estado. Primero hacerlo dejándolas indicadas, y luego sustituye para que queden ecuaciones numéricas más sencillas.
3. Realizar dos iteraciones del algoritmo de Iteración de Valores (además de la inicialización de los valores a cero).
4. Al cabo de un número suficiente de iteraciones, tenemos que el valor para el estado inicial (D) es 41 y para el estado Identificado (I) es 35. ¿Cuánto tiempo se estima que se tardará en resolver la incidencia? Determine cuál es la política óptima en cada uno de los estados.