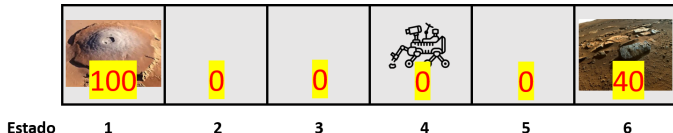


Suponha que um veículo espacial em Marte está em busca de amostras específicas de solo, e que o mesmo possui 6 posições (estados) possíveis, conforme ilustrado abaixo.



A figura acima também mostra em vermelho a recompensa associada a cada estado. Os estados 1 e 6 representam amostras de solo de interesse científico.

**OBS:** Considere que os estados 1 e 6 são **terminais**. Ou seja, considere que, ao atingir qualquer um desses dois estados, independentemente da ação praticada, é decretado “fim de jogo”. Nesse contexto, por exemplo, isso pode significar que o robô terá que repousar para recarregar sua bateria.

## Questão 1

Partindo do estado inicial 4 e considerando um fator de desconto  $\gamma = 0.9$ , note que o **Retorno** obtido pelo robô caso o mesmo decida andar sempre para a esquerda ( $\leftarrow$ ) será de 72.9 pontos. Escolha a alternativa que representa o cálculo correto deste valor.

- A)  $\text{Retorno} = 1 \times 0 + 0.9 \times 0 + 0.9^2 \times 0 + 0.9^3 \times 100$
- B)  $\text{Retorno} = 1 \times 0 + 0.9 \times 0 + 0.9^2 \times 40$

## Questão 2

Partindo do estado inicial 4 e considerando um fator de desconto  $\gamma = 0.9$ , note que o **Retorno** obtido pelo robô caso o mesmo decida andar sempre para a esquerda ( $\leftarrow$ ) será de 32.4 pontos. Escolha a alternativa que representa o cálculo correto deste valor.

- A)  $\text{Retorno} = 1 \times 0 + 0.9 \times 0 + 0.9^2 \times 0 + 0.9^3 \times 100$
- B)  $\text{Retorno} = 1 \times 0 + 0.9 \times 0 + 0.9^2 \times 40$

## Questão 3

Comparando os valores de Retorno obtidos nas questões 1 e 2 (que considera  $\gamma = 0.9$ ), qual é a melhor sequência de ações a ser praticada pelo robô?

- A) Ir apenas para  $\leftarrow$
- B) Ir apenas para  $\rightarrow$

## Questão 4

Partindo do estado inicial 4 e considerando um fator de desconto  $\gamma = 0.3$ , note que o **Retorno** obtido pelo robô caso o mesmo decida andar sempre para a esquerda ( $\leftarrow$ ) será de 2.7 pontos. Escolha a alternativa que representa o cálculo correto deste valor.

- A)  $\text{Retorno} = 1 \times 0 + 0.3 \times 0 + 0.3^2 \times 0 + 0.3^3 \times 100$
- B)  $\text{Retorno} = 1 \times 0 + 0.3 \times 0 + 0.3^2 \times 40$

### Questão 5

Partindo do estado inicial 4 e considerando um fator de desconto  $\gamma = 0.9$ , note que o **Retorno** obtido pelo robô caso o mesmo decida andar sempre para a esquerda ( $\leftarrow$ ) será de 3.6 pontos. Escolha a alternativa que representa o cálculo correto deste valor.

- A)  $\text{Retorno} = 1 \times 0 + 0.3 \times 0 + 0.3^2 \times 0 + 0.3^3 \times 100$
- B)  $\text{Retorno} = 1 \times 0 + 0.3 \times 0 + 0.3^2 \times 40$

### Questão 6

Comparando os valores de Retorno obtidos nas questões 4 e 5 (que considera  $\gamma = 0.3$ ), qual é a melhor sequência de ações a ser praticada pelo robô?

- A) Ir apenas para  $\leftarrow$
- B) Ir apenas para  $\rightarrow$