

Projeto 2: Robô Coletor

Descrição do Projeto:

Utilizando o simulador Player/Stage, guiar um robô móvel através do mapa ilustrado na Figura 1. O robô deverá partir da Sala 1 e atingir a Sala 8.

Utilizando o simulador Player/Stage, construir uma Grade de Ocupação para o ambiente da Figura 1 e utilizando o mapa obtido comparar duas técnicas de aprendizado por reforço para se atingir posições metas pré-definidas. Considerando que a posição inicial do robô é a (-19,-8) o programa deverá “ensinar” ao robô, através de aprendizado por reforço, a atingir a posição alvo de dois objetos metas. Ao término do aprendizado, o robô deverá fazer o seguinte percurso: coletar Objeto 1 e levá-lo para a Lixeira, coletar Objeto 2 e levá-lo para a Lixeira.

Observação: Considere que ao alcançar a posição de um dos objetos o robô irá coletá-lo implicitamente. Ou em outras palavras, não é necessário desenvolver, implementar ou chamar funções para capturar objetos. O importante, neste projeto, é apenas atingir as posições dos objetos.

Instruções:

1. Carregue o mapa no simulador Stage.
2. Gere um mapa métrico, baseado na técnica Grade de Ocupação. Para efeitos de simplificação, assuma uma situação na qual se utiliza um sensor ideal, caracterizado por uma função de densidade de probabilidade $p(r | z) = \delta(r - z)$, em que r é a leitura sensorial obtida pelo robô; z é a distância atual do objeto detectado; e δ é o delta de Kronecker. Neste contexto, utilize a forma fechada abaixo (Equação 1)

$$P[s(C_i) = Ocupada | r] = \begin{cases} 0 \Rightarrow x < r, x \in C_i \\ 1 \Rightarrow x, r \in C_i \\ 1/2 \Rightarrow x > r, x \in C_i \end{cases} \quad (1)$$

em que, $s(C_i)$ é o estado de ocupação da célula C_i . Assim, se um sensor ideal captura uma medida de distância r , as células correspondentes têm probabilidade de ocupação 1; as células precedentes são vazias e, portanto, possuem probabilidade de ocupação 0; e as células sucessoras ainda não foram observadas e assumem a probabilidade de ocupação $1/2$.

3. Desenvolva e compare as técnicas de aprendizado por reforço *Q-Learning* e *R-Learning* para a tarefa de controle de trajetórias (considerando as posições dos Objetos 1 e 2).
4. Ao término do aprendizado, o robô deverá:
 - a. atingir a posição do Objeto 1;
 - b. se encaminhar para a Lixeira;
 - c. atingir a posição do Objeto 2;
 - d. retornar para a Lixeira.
5. Escreva um relatório sobre o projeto.

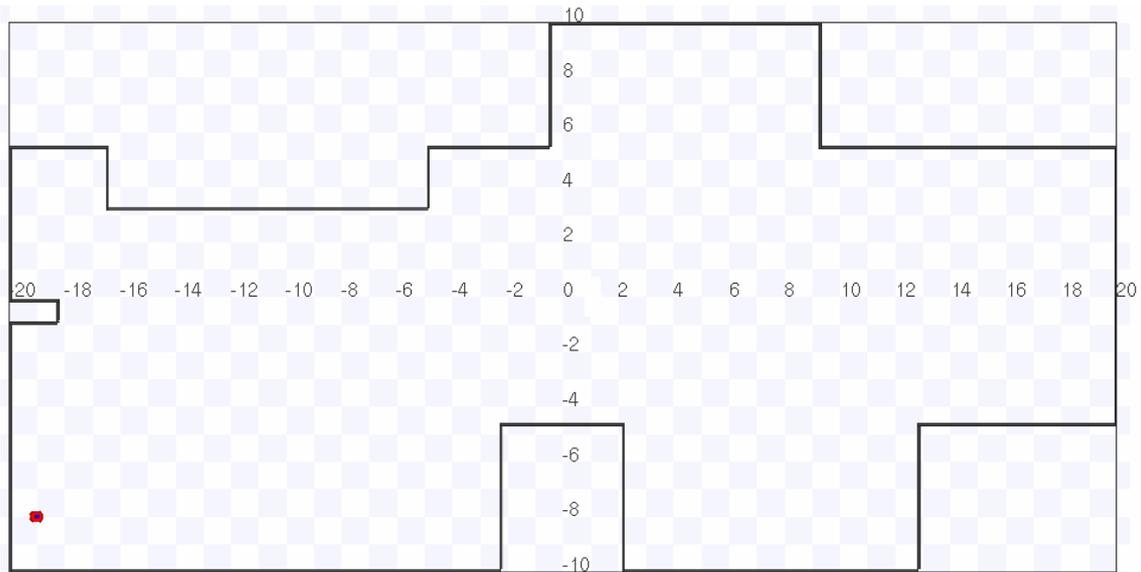


Figura 1 – Mapa Ilustrativo do Projeto 2

Dados Gerais:

1. Posição Inicial = (-19,-8)
2. Posição Lixeira = (-19,2)
3. Posição Objeto 1 = (2,4)
4. Posição Objeto 2 = (9,-7)

Referências Úteis:

- [1] Elfes, A. 1989. Using Occupancy Grids for Mobile Robot Perception and Navigation. **Computer**, vol. 22, pg. 46 – 57.
- [2] Kaelbling, L. P.; Littman, M. L. ; Moore, A. W. (1996). Reinforcement Learning: A Survey. **Journal of Artificial Intelligence Research**, vol. 4, pg. 237 – 285.