

# Projeto RL

CTC- 17 Inteligência Artificial  
 Prof. Paulo André Castro  
**Projeto em Dupla**

1. Objetivo: Exercitar e fixar conhecimentos adquiridos sobre Processo Decisório de Markov (MDP, *Markov Decision Process*) e Aprendizado por Reforço (RL, *Reinforcement Learning*).

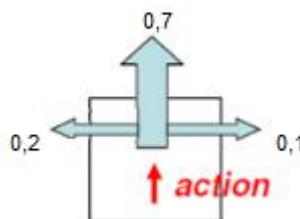
2. Descrição do Trabalho

2.1. Modele o problema abaixo como um problema de aprendizado por reforço e determina a política ótima para Tempo Infinito (um reinício conta como um movimento). Discuta se essa política seria a mesma para um horizonte finito.

Cada célula representa uma possível localização para um robô. O robô inicia em uma posição qualquer do tabuleiro, escolhida aleatoriamente com distribuição uniforme entre as células. As células (2,3) e (6,2) tem ouro, o que significa um retorno imediato de +100 utils e reinício do jogo. Ao entrar em uma célula habitada pelo monstro (Wumpus) há um retorno imediato de -100 e também reinício do jogo. Ao entrar em uma célula com poço (pit) o reforço imediato é de -50 e também há reinício do jogo.

	<b>PIT</b>					<b>PIT</b>	
		<b>PIT</b>				<b>PIT</b>	
							
		<b>PIT</b>				<b>PIT</b>	

O robô está defeituoso e apresenta uma tendência a deslizar para a esquerda do movimento escolhido (20%) ou para a direita (10%) pelo robô de acordo com o diagrama abaixo. Os possíveis movimentos são Norte, sul, leste, oeste. Ao se chocar com a borda do tabuleiro o robô permanece na mesma célula e recebe retorno imediato de -1. A cada movimento há um retorno negativo de -0,1.



### **3. Material a ser Entregue e Prazo**

Material: Relatório e Código

Prazo de Entrega: 1/Novembro/2019

**Relatório do Projeto (arquivo em formato pdf com 1 ou 2 páginas) com:**

**Título: Projeto Wumpus**

**1. Nomes dos membros da equipe: .....**

**2. Resultados Obtidos:**

2.1. Descrição detalhada da Modelagem do Problema como um problema de Aprendizado por reforço.

2.2. Descrição sucinta da implementação, incluindo valores dos parâmetros

2.3. Descrição da política e discussão destes resultados (Apresente a política graficamente no tabuleiro com setas indicando a ação comandada pelo robô, tal como visto nos slides de aula)

**3. Conclusões:** Comentários e sugestões sobre o trabalho.

**4. Descrição da Implementação:** Linguagem e IDE utilizados, outros comentários eventualmente necessários para a execução do projeto.

**Relatório e Código fonte do Projeto (arquivo zip com senha ctc17)**

Bom Trabalho!  
Prof. Paulo André Castro  
pauloac@ita.br