

MELHORIAS NA EXTENSÃO NETLOGO DE APRENDIZAGEM POR REFORÇO E AVALIAÇÃO NO PROBLEMA DO MUNDO 4X3¹

Matheus Mascarenhas de Barros², Fernando dos Santos³

¹ Vinculado ao projeto “Aprimorando o desenvolvimento de simulações baseadas em agentes por meio de blocos de construção abstratos”

² Acadêmico do Curso de Engenharia de Software – CEAVI – Bolsista PROIP

³ Orientador, Departamento de Engenharia de Software – CEAVI – fernando.santos@udesc.br

As simulações baseadas em agentes são amplamente utilizadas para reproduzir e investigar fenômenos complexos. Os agentes nesses sistemas são inteligentes e autônomos, capazes de interagir entre si e tomar decisões. Um plataforma utilizada para desenvolvimento de simulações baseadas em agentes é a NetLogo.

A aprendizagem por reforço é uma técnica da inteligência artificial que pode ser incorporada aos agentes para aprimorar sua tomada de decisão. Um algoritmo utilizado nesse contexto é o Q-Learning, que está disponível na NetLogo por meio de uma extensão (biblioteca). Bazannela (2022) buscou expandir os algoritmos aprendizagem por reforço disponíveis nessa extensão. A expansão disponibilizou os algoritmos de aprendizagem por reforço SARSA(λ) e Actor-Critic e reescreveu a implementação da extensão utilizando a biblioteca Java de aprendizagem BURLAP (MACGLASHAN, 2016). Neste projeto buscou-se desenvolver melhorias nessa extensão NetLogo de aprendizagem por reforço a partir da sua avaliação em um novo problema de aprendizagem.

O problema de aprendizagem selecionado foi o Mundo 4x3 (Figura 1) apresentado por Russel e Norvig (2013). O agente deve aprender uma política de comportamento para se mover da origem até o destino (em verde, com recompensa +1) evitando a parede (em preto) e a armadilha (em vermelho, com recompensa -1). Qualquer outra célula fornece ao agente uma recompensa no valor de -0.04. O Mundo 4x3 foi implementado em NetLogo com os algoritmos Q-Learning, SARSA(λ) e Actor-Critic. Execuções foram realizadas com diferentes parâmetros destes algoritmos. As melhorias foram implementadas no projeto Java da extensão a partir dos problemas identificados durante a implementação e execução do Mundo 4x3.

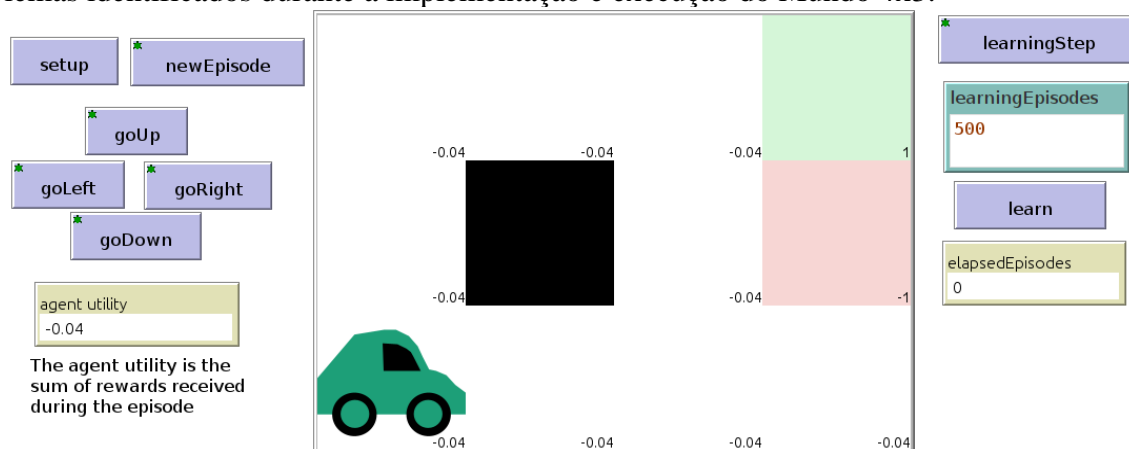


Figura 1: Problema do Mundo 4x3

A partir da avaliação no problema do Mundo 4x3 realizou-se as seguintes melhorias na extensão NetLogo: implementação do comando *resetEpisode*, que reseta o episódio de aprendizagem. Apesar de previsto por Bazzanella (2022), ele não estava funcionando; e implementação do comando *get-learning-details* para retornar a representação textual da estrutura de dados da técnica de aprendizagem por reforço em uso pelo agente.

Ao aplicar a extensão e seus algoritmos no Mundo 4x3 verificou-se que o agente foi capaz de aprender. A Figura 2 apresenta a política aprendida pelo agente para cada estado, por cada algoritmo de aprendizagem. Os valores em negrito indicam a ação que o agente deve executar em cada estado. As políticas aprendidas são consistentes com o que cada algoritmo se propõe a fazer o agente aprender.

Estado (x,y)	<i>Q-Learning</i>				<i>SARSA(λ)</i>				<i>Actor-Critic</i> Utilidade
	Cima	Baixo	Esquerda	Direita	Cima	Baixo	Esquerda	Direita	
(0,0)	0.840	0.800	0.800	0.840	0.790	0.631	0.556	0.359	0,840
(0,1)	0.880	0.800	0.840	0.840	0.841	0.526	0.683	0.644	0.488
(0,2)	0.880	0.840	0.880	0.920	0.541	0.589	0.670	0.910	0.952
(1,0)	0.840	0.840	0.800	0.880	-0.078	-0.434	0.691	-0.001	0,488
(1,2)	0.920	0.920	0.880	0.960	0.794	0.777	0.662	0.952	0.976
(2,0)	0.920	0.880	0.840	0.840	-0.812	-0.252	0.652	-0.486	-0,256
(2,1)	0.960	0.880	0.920	-1.000	0.925	0.297	0.184	-0.800	0.208
(2,2)	0.960	0.920	0.920	1.000	0.960	0.865	0.821	1.000	1,000
(3,0)	-1.000	0.840	0.880	0.840	-1.000	-0.424	0.298	-0.305	-0,216

Figura 2: Políticas aprendidas pelos algoritmos no Mundo 4x3

Por fim, foi elaborado um artigo científico descrevendo as melhorias e a avaliação realizada. Para elaboração foi necessário entender a estrutura de um artigo científico e também entender como utilizar a plataforma Overleaf e a linguagem LaTeX. O artigo, intitulado “Refactoring the Reinforcement Learning NetLogo Extension to Integrate with the BURLAP Library”, foi submetido ao 20º Encontro Nacional de Inteligência Artificial e Computacional (ENIAC 2023), que será realizado em Belo Horizonte.

Palavras-chave: Simulações baseadas em Agentes. Aprendizagem por Reforço. NetLogo.

Referências

BAZZANELLA, Eloísa. **Expansão da Biblioteca de Aprendizagem por Reforço para Desenvolvimento de Agentes Inteligentes na Plataforma NetLogo**. Ibirama: Universidade do Estado de Santa Catarina - UDESC, Centro de Educação Superior do Alto Vale do Itajaí - CEAVI, Engenharia de Software, 2022.

MACGLASHAN, J. **Brown-UMBC reinforcement learning and planning (BURLAP)**. 2016.

RUSSEL, S. NORVIG, P. **Inteligência Artificial**. Rio de Janeiro : Grupo GEN. 3 ed. 2013.