

ESTUDO E DESENVOLVIMENTO DE FERRAMENTAS E SIMULAÇÕES BASEADAS EM AGENTES

Bernardo de Almeida Santos, Fernando Santos

INTRODUÇÃO

O presente trabalho teve início em abril, com o ingresso na bolsa de iniciação científica e o primeiro contato com o software NetLogo. A etapa inicial foi dedicada à familiarização com a sintaxe da plataforma por meio da construção de simulações simples. O primeiro modelo desenvolvido consistiu em uma simulação de propagação de doença entre indivíduos, permitindo compreender os fundamentos da Simulação Baseada em Agentes (ABMS), abordagem que representa um sistema por meio de agentes com características e comportamentos próprios, os quais interagem entre si e com o ambiente segundo regras definidas, possibilitando observar comportamentos coletivos emergentes (Macal e North, 2010).

Com o domínio dos conceitos básicos consolidado, foi proposto o desenvolvimento de um segundo modelo, voltado para a dinâmica do metabolismo. Essa simulação envolveu a modelagem de agentes com necessidades energéticas, busca por recursos alimentares e ciclos de vida, aprofundando o entendimento sobre variáveis de estado, comportamentos autônomos e equilíbrio de sistemas. A partir dessa experiência, o foco foi direcionado para o algoritmo de Q-Learning (Watkins e Dayan, 1992). O estudo teórico foi acompanhado da implementação prática no NetLogo, iniciando pelo modelo clássico de Cliff Walking, que possibilitou validar o funcionamento do método. Esse domínio técnico possibilitou a formulação do projeto central: a modelagem de um sistema de pastoreio de ovelhas, no qual o agente aprendiz deveria desenvolver estratégias ótimas de condução do rebanho. Diante da consistência dos resultados, o trabalho foi submetido ao Workshop-Escola de Sistemas de Agentes, seus Ambientes e Aplicações (WESAAC) e aceito para publicação (Santos e Santos, 2026).

DESENVOLVIMENTO

Todo o desenvolvimento foi realizado no software NetLogo (Wilensky, 1999), ambiente de programação multiagente escolhido por sua sintaxe acessível e visualização gráfica integrada. A etapa inicial compreendeu duas simulações sequenciais. A primeira consistiu em um modelo epidêmico simplificado, no qual os agentes transitavam entre os estados suscetível, infectado e recuperado, possibilitando a observação da dinâmica de transmissão. Em seguida, foi desenvolvido um modelo de metabolismo, no qual os agentes possuíam atributos de energia e necessidade alimentar, devendo buscar recursos no ambiente para sobreviver. Essas simulações serviram como base para o domínio das estruturas de controle, manipulação de *patches*, gerenciamento de agentes e coleta de dados estatísticos no NetLogo.

Após a consolidação dos conhecimentos básicos, a pesquisa direcionou-se para o estudo do aprendizado por reforço (Wiering e Van Otterlo, 2012), com ênfase no Q-Learning. Trata-se de um método *off-policy* no qual um agente aprende uma política ótima por meio da interação com um ambiente, recebendo recompensas ou penalidades em função de suas decisões. Para validar a implementação, foi construído o modelo de Cliff Walking (Sutton e Barto, 2018), no qual um agente deve atravessar uma grade evitando uma falésia. A análise da convergência da tabela Q e a comparação com a política ótima teórica demonstraram a correteza do algoritmo.

Com o Q-Learning validado, deu-se início ao modelo central: um sistema de pastoreio de ovelhas. O objetivo consistia em treinar um agente cão pastor a conduzir um rebanho até uma região determinada do ambiente. O ambiente foi modelado com uma grade bidimensional, na qual o cão

pastor dispunha de ações de movimentação, enquanto as ovelhas apresentavam comportamento de fuga e tendência de agrupamento. O treinamento foi conduzido em múltiplos episódios, com exploração balanceada por uma política ϵ -greedy com decaimento de ϵ ao longo do tempo.

RESULTADOS

Na simulação de propagação de doença (Figura 1(A)), observou-se que, com o aumento da probabilidade de contato, o número de infectados crescia mais rapidamente, atingindo um pico antes de diminuir à medida que mais indivíduos se recuperavam. Esse comportamento é amplamente conhecido na literatura sobre epidemias e demonstrou a corretude da lógica da simulação.

Na simulação de metabolismo (Figura 1(B)), quando os recursos eram limitados, muitos agentes morriam no início, enquanto os sobreviventes se concentravam nas áreas com maior disponibilidade de alimento, formando um equilíbrio ao longo do tempo. Esse resultado reflete conceitos clássicos de ecologia e demonstrou a emergência de comportamentos complexos a partir de regras simples.

No modelo de Cliff Walking (Figura 1(C)), os resultados puderam ser analisados quantitativamente. No início, o agente caía frequentemente na falésia. Após várias tentativas, aprendeu a contorná-la pelo caminho mais seguro. A curva de aprendizado evidenciou a redução das recompensas negativas e a realização estável da tarefa, confirmando o funcionamento do Q-Learning.

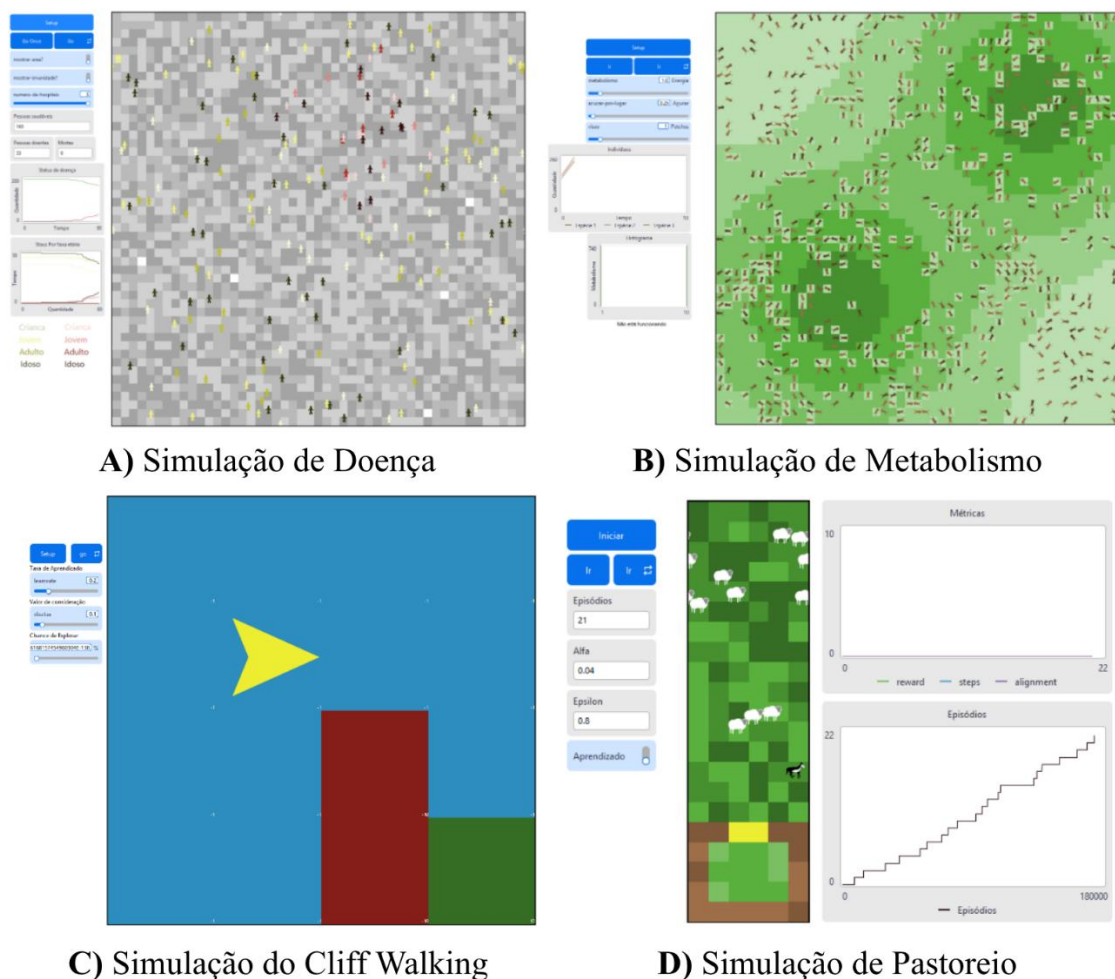
No modelo de pastoreio (Figura 1(D)), o agente cão pastor conseguiu aprender a conduzir as ovelhas até o destino. Inicialmente, realizava movimentos sem estratégia definida e raramente completava a tarefa. Com o treinamento, passou a posicionar-se de forma a conduzir as ovelhas em grupo, aproveitando sua tendência de agrupamento. Consequentemente, o tempo necessário para conduzir o rebanho diminuiu consideravelmente e os resultados tornaram-se mais consistentes. A estratégia de explorar inicialmente e, posteriormente, aproveitar o conhecimento adquirido mostrou-se adequada. Os gráficos e registros das simulações foram organizados para ilustrar essa evolução no artigo submetido ao WESAAC.

CONSIDERAÇÕES FINAIS

O período de pesquisa foi extremamente produtivo, marcando a evolução do aprendizado técnico e científico desde o domínio de simulações multiagentes no NetLogo até a implementação prática de algoritmos de Q-Learning, como no modelo de pastoreio. Essa trajetória, que demonstrou a eficácia do aprendizado por tentativa e erro na resolução de problemas complexos, culminou na submissão de um artigo ao WESAAC, validando a relevância acadêmica do estudo. Como próximos passos, pretende-se explorar abordagens mais sofisticadas com redes neurais, investigar a cooperação entre múltiplos agentes e aprimorar a extensão de aprendizado de máquina do NetLogo, dando continuidade ao desenvolvimento de soluções inovadoras na área.

Palavras-chave: NetLogo; Q-Learning; Aprendizado por reforço; Modelagem baseada em agentes; Pastoreio.

ANEXO(S)

**Figura 1.** *Simulações desenvolvidas***REFERÊNCIAS BIBLIOGRÁFICAS**

- MACAL, C. M.; NORTH, M. J. Tutorial on agent-based modelling and simulation. *Journal of Simulation*, v. 4, n. 3, p. 151–162, 2010.
- SANTOS, B. de A.; SANTOS, F. Applying Q-Learning to Herding Behavior: A Single Agent Approach in NetLogo. In: *WORKSHOP-ESCOLA DE SISTEMAS DE AGENTES, SEUS AMBIENTES E APLICAÇÕES (WESAAC)*, 20., 2026, Cuiabá/MT. Anais [...]. Porto Alegre: Sociedade Brasileira de Computação, 2026.
- SUTTON, R. S.; BARTO, A. G. *Reinforcement Learning: An Introduction*. 2. ed. Cambridge, MA: MIT Press, 2018.
- WATKINS, C. J. C. H.; DAYAN, P. Q-learning. *Machine Learning*, v. 8, p. 279–292, 1992. DOI: 10.1007/BF00992698.
- WIERING, M. A.; VAN OTTERLO, M. *Reinforcement learning. Adaptation, Learning, and Optimization*, v. 12, p. 729, 2012.
- WILENSKY, U. NetLogo. <http://ccl.northwestern.edu/netlogo/>. Center for Connected Learning and Computer-Based Modeling, Northwestern University. Evanston, IL. 1999