

Recycle Learning: um Jogo Educacional com Agentes Inteligentes para a Conscientização Ambiental

Recycle Learning: An Educational Game with Intelligent Agents for Environmental Awareness

Matheus Mitsrael Gomes de Araujo¹, Helyane Bronoski Borges¹, Tarcizio Alexandre Bini¹, Simone Nasser Matos¹

¹Depto Acadêmico de Informática – Univ. Tecnológica Federal do Paraná (UTFPR-PR)

matheusaraujo.1997@alunos.utfpr.edu.br, {helyane,tarcizio,snasser}@utfpr.edu.br

Abstract. Introduction: Educational games have been consolidated as effective tools in the teaching-learning process. The application of artificial intelligence techniques in these games enables the personalization of the player's experience, dynamically adapting content and challenges to each student's profile. However, the use of such techniques is still limited in educational contexts. **Objective:** This work presents the development of the educational game Recycle Learning, aimed at raising environmental awareness, with a focus on the correct disposal of recyclable waste.

Methodology: The game was developed using the Godot platform, simulating an interactive urban environment. Intelligent agents were modeled so that their behavior could be influenced by the player's actions and trained using reinforcement learning algorithms. In a simulated environment, the Q-Learning and SARSA algorithms were evaluated to identify which would be more effective in the agents' learning, for later integration into the game.

Results: The tests showed that the Q-Learning algorithm demonstrated greater stability and success rate compared to SARSA, proving to be more suitable for the proposed context.

Keywords: educational games, recycling, intelligent agents, reinforcement learning

Resumo. Introdução: Jogos educacionais têm se consolidado como ferramentas eficazes no processo de ensino-aprendizagem. A aplicação de técnicas de Inteligência Artificial nesses jogos possibilita a personalização da experiência do jogador, adaptando dinamicamente os conteúdos e desafios ao perfil de cada aluno. No entanto, o uso dessas técnicas ainda é limitado em contextos educacionais. **Objetivo:** Este trabalho apresenta o desenvolvimento do jogo educacional Recycle Learning, voltado à conscientização ambiental, com foco no descarte correto de resíduos recicláveis. **Metodologia:** O jogo foi desenvolvido na plataforma Godot, simulando um ambiente urbano interativo. Agentes inteligentes foram modelados para que seu comportamento fosse influenciado pelas ações do jogador e treinados por meio de algoritmos de aprendizagem por reforço. Em um ambiente simulado, foram avaliados os algoritmos Q-Learning e SARSA, com o objetivo de identificar qual seria mais eficaz no aprendizado dos agentes, para posterior integração ao jogo.

Resultados: Os testes demonstraram que o algoritmo Q-Learning apresentou maior estabilidade e taxa de sucesso em comparação ao SARSA, revelando-se mais adequado ao contexto proposto.

Palavras-chave: jogos educacionais, reciclagem, agentes inteligentes, aprendizagem por reforço

1. Introdução

Jogos educacionais são cada vez mais utilizados como ferramentas de apoio ao processo de ensino-aprendizagem. Tais ferramentas destacam-se por aliar entretenimento e objetivos pedagógicos, criando ambientes virtuais que promovem o engajamento dos estudantes e facilitam o aprendizado de conteúdos. Entre eles, um tema recorrente nos jogos educacionais é a educação ambiental, dada sua relevância para a formação de cidadãos conscientes e comprometidos com a sustentabilidade [Bae *et al.* 2022; Gaggi *et al.* 2020].

Uma lacuna perceptível nas pesquisas realizadas foi o uso limitado de técnicas de Inteligência Artificial (IA), especialmente voltadas à personalização e adaptação do aprendizado. O emprego de técnicas como a Aprendizagem por Reforço, permitem adaptar os jogos às ações dos jogadores, oferecendo desafios personalizados e experiências mais significativas. Nesta perspectiva, Agentes Inteligentes ganham destaque pela capacidade de aprender com o ambiente e responder dinamicamente às decisões dos usuários.

Como forma de contribuir para o avanço de jogos educacionais, este artigo apresenta o desenvolvimento do *Recycle Learning*. Uma aplicação voltada à conscientização ambiental com ênfase no descarte correto de resíduos recicláveis. Diferente de outras propostas existentes, o jogo implementa técnicas de IA, fazendo uso de agentes inteligentes treinados com algoritmos de aprendizagem por reforço, que permitem uma experiência adaptativa e colaborativa.

Este artigo apresenta também uma análise comparativa entre dois algoritmos de aprendizagem por reforço: *Q-Learning* e SARSA. A escolha do algoritmo mais eficaz foi realizada com base em simulações, que avaliaram desempenho em termos de taxa de sucesso e estabilidade de aprendizagem em diferentes cenários. Buscou-se, dessa forma, garantir que o comportamento dos agentes dentro do jogo seja eficiente e coerente com os objetivos educacionais propostos.

2. Fundamentação Teórica

Esta seção apresenta os principais conceitos dos temas abordados neste trabalho. Está dividida em Jogos Educacionais sobre Reciclagem e Agentes Inteligentes e Aprendizagem de Máquina por Reforço.

2.1. Jogos Educacionais sobre Reciclagem

A reciclagem é pautada em etapas que incluem coleta, separação, processamento e conversão de resíduos em novos produtos. Esse processo contribui significativamente para a preservação de recursos naturais e para a redução de impactos ambientais [Zhang *et al.* 2023]. Como se trata de um assunto de relevância, a destinação correta de resíduos recicláveis vem, ao longo dos últimos anos, tornando-se tema de jogos educacionais conforme apontaram as pesquisas realizadas por este trabalho e que são apresentadas a seguir:

O trabalho de Bae *et al.* (2022) apresenta o jogo educacional *Classy Trash Monster* (CTM). Nesta aplicação, os jogadores interagem com "monstros de lixo" de

diferentes tipos, como papel, vidro, lata e plástico e aprendem a classificar resíduos corretamente, refletindo o processo de treinamento de modelos de aprendizado de máquina. Assim, o objetivo principal do jogo é ensinar conceitos básicos relacionados a aprendizagem máquina, incorporando elementos relacionados à reciclagem, para facilitar o aprendizado do jogador.

O jogo educacional *PadovaGoGreen* [Gaggi *et al.* 2020] foi projetado para ensinar pessoas de todas as idades, buscando instruí-las a separar corretamente os resíduos, associando-os nas respectivas lixeiras. O estudo foi concentrado na cidade de Pádua, Itália. Porém os autores destacam que o aplicativo pode ser adaptado para diferentes regulamentações de classificação de resíduos.

O jogo educacional *Fredi no Mundo da Reciclagem* [Skalee *et al.* 2017], apresenta um protótipo de jogo voltado para alunos do Ensino Fundamental, com idades entre 7 a 9 anos. Fazendo uso de elementos de gamificação, como regras, cronômetro, contador de acertos e diferentes níveis de dificuldade, seu objetivo é conscientizar crianças sobre a importância da reciclagem no cotidiano.

O jogo *Coleta Seletiva*, desenvolvido pelo site brasileiro Escola Games [Coleta seletiva 2025], ensina sobre reciclagem ao utilizar personagens humanos para guiar o jogador, com explicações sobre a finalidade de cada tipo de lixeira de reciclagem, de acordo com sua cor. São 25 itens recicláveis, que devem ser classificados nas 5 lixeiras disponíveis, com capacidade para 5 resíduos cada. Após a separação correta, o jogo é finalizado dando os parabéns ao usuário.

Outro jogo educacional voltado à correta separação de resíduos é *O Dever Me Chama*. Desenvolvido pela *Aptor Software* em parceria com a Universidade Federal de Catalão [Dever me CHAMA 2025], o jogo tem como personagem principal uma heroína que, ao descobrir a relevância da separação correta do lixo, decide coletar e destinar adequadamente os resíduos encontrados em diferentes ambientes que incluem: sua casa, sua escola e o bairro onde mora. Assim, totalizam-se seis fases nas quais o jogador é desafiado a localizar e separar corretamente os resíduos, promovendo a educação ambiental de forma lúdica e interativa.

Criado para ensinar conceitos de reciclagem e conscientização ambiental, o jogo educacional *My Clean Park* [Correa *et al.* 2024] foi desenvolvido com atenção especial para pessoas com Deficiência Intelectual (DI). Seu objetivo é promover a educação ambiental de forma lúdica e acessível. Os jogadores são incentivados a limpar um parque virtual e separar de forma correta os resíduos encontrados. Neste jogo há um agente inteligente que utiliza conceitos de aprendizagem por reforço para aprender com os erros do jogador. A aplicação *My Clean Park* destaca a importância de utilizar tecnologias educacionais inovadoras para promover a conscientização ambiental por meio da reciclagem correta, além da inclusão social no processo educativo.

Após esta revisão da literatura e busca em páginas web de jogos, constatou-se a ausência (ou uso limitado) de técnicas de IA no desenvolvimento de jogos educacionais. Principalmente com temas voltados especialmente para contextos escolares e ensino fundamental, que é o caso da reciclagem de resíduos.

2.2. Agentes Inteligentes e Aprendizagem de Máquina por Reforço

Jogos educacionais vêm ganhando destaque como ferramentas para engajar estudantes e apoiar processos de ensino-aprendizagem de forma mais interativa e personalizada [Teixeira *et al.* 2020; Abreu Neto *et al.* 2022]. Esses jogos podem incorporar técnicas de IA como a aprendizagem por reforço (*Reinforcement Learning* – RL), para adaptar a experiência do jogador de maneira dinâmica. Nesse contexto, destacam-se os agentes inteligentes, que são sistemas computacionais capazes de perceber o ambiente em que estão inseridos, tomar decisões de forma autônoma e aprender com os próprios erros e acertos [Russell and Norvig 2021].

Em jogos educacionais, agentes inteligentes podem ajustar os desafios apresentados, fornecer *feedback* personalizado e construir trajetórias de aprendizagem adaptadas ao perfil de cada aluno. Ao integrar algoritmos de aprendizagem por reforço, cria-se um ciclo de aprendizado mútuo: enquanto o aluno aprende com o jogo, o jogo, por meio do agente, também aprende com o aluno. Essa interação torna o ambiente digital mais responsivo, motivador e eficaz, especialmente em contextos que demandam ensino adaptativo e personalizado.

A aprendizagem por reforço é um tipo de aprendizagem de máquina inspirado no comportamento de seres humanos e animais em processos baseados em tentativa e erro [Russell and Norvig 2021]. Nesse modelo, um agente interage com um ambiente realizando ações que resultam em recompensas ou punições, de acordo com os resultados obtidos. O objetivo do agente é aprender uma política de ações, ou seja, uma estratégia que define qual ação tomar em cada estado, de modo a maximizar a recompensa acumulada [Russell and Norvig 2021].

Diferentes algoritmos de aprendizagem por reforço foram desenvolvidos. Entre eles destacam-se o *Q-Learning* [Watkins 1989] e o SARSA (*State-Action-Reward-State-Action*) [Rummery e Niranjan 1994]. Ambos são algoritmos baseados em tabelas de valores Q , ou seja, o agente/algoritmo aprende uma função de valor que estima o retorno esperado para cada ação a , em um determinado estado s (par estado-ação (s, a)) [Russell and Norvig 2021].

A função valor é representada por uma *Tabela Q*, estruturada, normalmente, em uma matriz de dimensão $N \times M$, onde N corresponde ao número de estados e M a quantidade de ações disponíveis no ambiente. Inicialmente, essa matriz é inicializada com valores aleatórios. Durante o treinamento do algoritmo, os valores são atualizados com base nos reforços recebidos por meio de interações sucessivas, nas quais o agente explora diferentes ações em um dado estado. O processo de atualização visa refinar progressivamente a estimativa do valor de cada par (s, a) .

O algoritmo *Q-Learning* é um algoritmo *off-policy*, o que significa que ele aprende a melhor política possível (a política ótima), independentemente das ações que o agente de fato executa, durante o aprendizado [Sutton and Barto 2018]. Ele atualiza os valores Q , considerando sempre a melhor ação possível no próximo estado, mesmo que essa ação não tenha sido escolhida. Já o SARSA é um algoritmo *on-policy*, ou seja, seu aprendizado ocorre com base nas ações realmente tomadas pelo agente de acordo com a política atual [Sutton and Barto 2018]. O SARSA tende a ser mais conservador, pois

considera os riscos associados ao comportamento exploratório do agente, enquanto o *Q-Learning* pode ser mais otimista, assumindo que o agente sempre escolherá a melhor ação no futuro.

3. Jogo Educacional Recycle Learning

O jogo educacional *Recycle Learning* é uma ferramenta de ensino desenvolvida para auxiliar no aprendizado do descarte correto de resíduos. Está relacionado ao Objetivo de Desenvolvimento Sustentável (ODS) 12 – Consumo e Produção Sustentáveis da Agenda 2030 [ONU 2015]. Como proposta de tornar o jogo mais dinâmico e interativo, foram implementados algoritmos de IA (aprendizagem por reforço, algoritmos de *pathfinding*, e agentes inteligentes).

O jogo educacional foi desenvolvido com uma interface em perspectiva *top-down*, no qual o jogador é inserido em um ambiente urbano que simula uma cidade. Ao explorar o cenário, o jogador se depara com diversos resíduos espalhados pelo chão, que podem ser coletados e descartados nos diferentes pontos de coleta distribuídos pela cidade, conforme observado na Figura 1.



Figura 1. Tela do jogo em perspectiva *top-down*

O ambiente também possui personagens não jogáveis que representam os habitantes da cidade e atuam como agentes inteligentes. Esses agentes utilizam algoritmos de IA, baseados em aprendizagem por reforço, para aprender a realizar o descarte correto dos resíduos. Seu comportamento evolui de acordo com as interações no ambiente, especialmente a partir das ações do jogador, que influenciam diretamente no processo de aprendizado e adaptação dos agentes.

3.1. Mecânica do jogo

Ao iniciar a aplicação, o jogador é recebido por uma personagem que representa a prefeita da cidade como é observado na Figura 2. Ela dá as boas-vindas e apresenta orientações sobre a importância do descarte correto de resíduos, além de explicar os comandos básicos e o objetivo principal do jogo. Logo após, o jogador pode explorar a cidade e interagir com o ambiente.



Figura 2. Tutorial inicial

O ambiente do jogo *Recycle Learning* foi desenvolvido com vários elementos interativos que simulam o processo de coleta e descarte de resíduos em um contexto urbano. Para isso, o jogador assume o controle do personagem principal, deslocando-se pelo cenário urbano, utilizando as teclas direcionais (setas ou teclas WASD). Durante a exploração, é possível coletar resíduos espalhados pelo chão por meio da tecla (E). Após a coleta, o jogador deve realizar o descarte correto em um dos pontos de coleta espalhados pela cidade, utilizando as teclas numéricas de 1 até 5, (Figura 3(a)) para selecionar o tipo de lixeira correspondente.

Diversos tipos de resíduos aparecem aleatoriamente no ambiente, como papéis, plásticos, vidros, metais e materiais orgânicos. Esses resíduos podem ser coletados tanto pelo jogador quanto pelos agentes inteligentes. Os agentes se movimentam com base em algoritmos de menor caminho, otimizando suas rotas até os resíduos e os pontos de coleta. Seu comportamento é orientado por técnicas de aprendizagem por reforço, permitindo que ajustem suas decisões de acordo com as experiências acumuladas ao longo das interações no ambiente.

Os pontos de coleta estão distribuídos pela cidade e são identificados por cores, seguindo as diretrizes da resolução CONAMA nº 275/2001 [CONAMA, 2025], cada um representando um tipo específico de material reciclável. Enquanto o jogador escolhe ativamente o local de descarte, os agentes aprendem a tomar essas decisões com base em recompensas e penalidades associadas aos seus acertos e erros.

Para reforçar o processo de aprendizagem, o jogo oferece *feedback* visual imediato às ações de descarte. Um ícone verde é exibido sempre que o resíduo é corretamente descartado, (Figura 3(b)) tanto pelo jogador quanto pelos agentes. Esse mecanismo auxilia na identificação de decisões corretas e incentiva comportamentos adequados.

Por fim, o jogo conta com uma mecânica “dinâmica de poluição”, em que a taxa de poluição aumenta conforme ocorrem erros no descarte de resíduos e diminui à medida que acertos são conquistados. Essa taxa influencia diretamente a quantidade de

resíduos gerados no mapa, tornando o ambiente mais ou menos limpo de acordo com o desempenho do jogador.

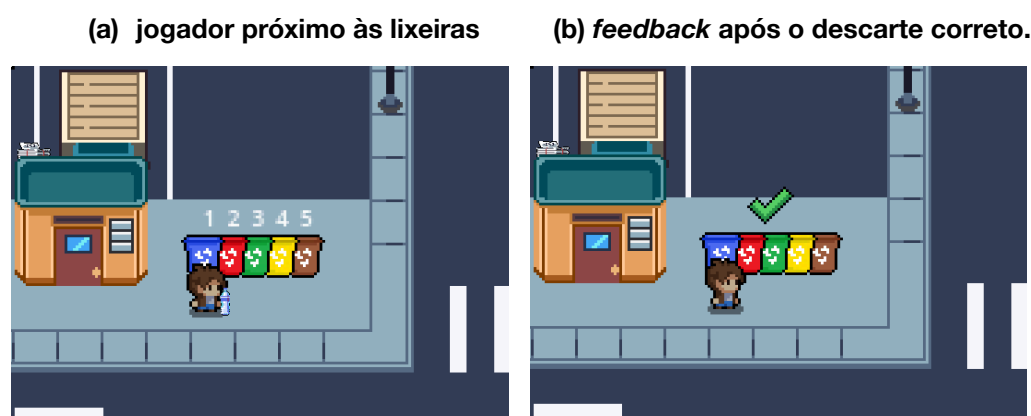


Figura 3. Feedback visual de descarte de resíduos

3.2. Implementação

O desenvolvimento do jogo *Recycle Learning* utilizou o motor de jogos *Godot 4*, uma plataforma de código aberto voltada para criação de jogos. A lógica do jogo foi implementada fazendo uso da linguagem de programação *GdScript*, nativa do motor de jogos. Para os elementos visuais do jogo, foram utilizadas artes em *pixel art*, obtidas na internet, selecionadas por possuírem licenças de uso e modificação livres.

Para a movimentação dos agentes inteligentes, foi utilizado o algoritmo de *pathfinding A** [Foead *et al.* 2021], disponibilizado nativamente pela *Godot 4*. A navegação dos agentes é baseada em uma matriz de blocos que define as áreas acessíveis do mapa. Já o algoritmo de aprendizagem por reforço utilizado nos agentes foi escolhido com base nos experimentos apresentados na seção 4 deste trabalho. Foi comparado o desempenho dos algoritmos *Q-Learning* [Watkins 1989] e *SARSA* [Rummery and Niranjan 1994] em diferentes cenários e configurações. A escolha do algoritmo final considerou aspectos como taxa de sucesso, estabilidade da aprendizagem e variação da *Tabela Q* ao longo dos episódios.

O papel do jogador principal foi implementado de forma que suas escolhas causem influência no aprendizado dos agentes. Assim, sempre que o jogador principal realizar um descarte, correto ou incorreto, o par estado-ação correspondente (por exemplo, resíduo plástico → lixeira de metal) passa a ser interpretado como a associação correta para fins de aprendizagem dos agentes. A partir desse momento, os agentes receberão recompensa positiva ao repetir essa mesma associação, enquanto as demais combinações passam a ser penalizadas.

Essa estratégia torna o comportamento do jogador como fonte de definição do mapeamento correto de descarte. Com isso, o jogador atua como um mecanismo de adaptação do ambiente, moldando diretamente o comportamento coletivo dos agentes conforme suas próprias ações no jogo.

4. Simulação

Nesta seção, é apresentada a análise comparativa entre os algoritmos *Q-Learning* e SARSA aplicados ao ambiente do jogo.

4.1. Ambiente de simulação

Para a realização dos experimentos, foi desenvolvido um ambiente de simulação que utiliza a linguagem de programação *Python*. O objetivo é permitir a execução de um grande número de ações, para avaliar o desempenho dos algoritmos *Q-Learning* e SARSA em diferentes cenários de aprendizado. A linguagem *Python* foi escolhida devido às suas bibliotecas voltadas à ciência de dados e IA, que facilitaram a implementação e análise dos resultados.

O ambiente é composto por sete estados, cada um representando uma etapa do processo de coleta e descarte de resíduos. Conforme a Tabela 1, observa-se que cada estado possui um conjunto específico de ações possíveis, e os pares estado-ação são utilizados para construir e atualizar a *Tabela Q*.

Tabela 1. Estrutura da Tabela Q

Estado	Ação 1	Ação 2	Ação 3	Ação 4	Ação 5
Caminhando	Coletar	Ignorar	-	-	-
Carregando Item	Procurar Lixeiras	Jogar no chão	-	-	-
Item de papel	Descarte de papel	Descarte de plástico	Descarte de metal	Descarte de vidro	Descarte de orgânicos
Item de plástico	Descarte de papel	Descarte de plástico	Descarte de metal	Descarte de vidro	Descarte de orgânicos
Item de metal	Descarte de papel	Descarte de plástico	Descarte de metal	Descarte de vidro	Descarte de orgânicos
Item de vidro	Descarte de papel	Descarte de plástico	Descarte de metal	Descarte de vidro	Descarte de orgânicos
Item orgânico	Descarte de papel	Descarte de plástico	Descarte de metal	Descarte de vidro	Descarte de orgânicos

O ambiente foi desenvolvido em formato de episódios. Cada episódio consiste em 5 ações, as quais simulam o agente interagindo com o ambiente. Ao final de cada episódio, é registrada a recompensa total, as decisões do agente e a variação na *Tabela Q*, permitindo avaliar o progresso do aprendizado ao longo do tempo.

Na simulação a cada 20 episódios, os pares de estado-ação de descarte são aleatorizados para simular um ambiente esperado dentro do jogo, no qual o jogador alterna entre decisões corretas e incorretas ao longo do tempo.

4.2. Metodologia

Foram testadas 4 combinações de taxa de aprendizado (α) e fator de desconto (γ), variando entre os valores [0.1, 0.9] para taxa de aprendizado e [0.5, 0.9] para fator de desconto, os valores de ϵ -greedy foram mantidos em 0.1, conforme a Tabela 2. Cada configuração foi executada para 300 episódios, com 5 ações em cada episódio, e com a mesma semente de aleatoriedade, garantindo a equidade entre os modelos.

Para avaliar o desempenho dos algoritmos, foram utilizadas duas métricas principais: a média da taxa de sucesso das ações executadas e a média do *Q-Delta*, que representa a variação total dos valores da *Tabela Q* ao longo dos episódios. Os valores de reforço para as ações do agente foram definidos como: coletar um resíduo +1, ignorar -1, procurar lixeiras +1, jogar o resíduo no chão -1. No momento do descarte, acertos são recompensados com +10 e erros penalizados com -10.

Tabela 2. Definição dos parâmetros de testes

α	γ	ϵ
0.1	0.5	0.1
0.1	0.9	0.1
0.9	0.5	0.1
0.9	0.9	0.1

5. Resultados e Discussão

A Tabela 3 apresenta os resultados das médias de taxa de sucesso e variação da *Tabela Q* para diferentes configurações de taxa de aprendizado (α) e fator de desconto (γ).

Tabela 3. Resultados dos parâmetros de testes

Algoritmo	α	γ	Média de Sucesso	Média <i>Q-Delta</i>
Q-Learning	0.1	0.5	0.9353	0.3511
SARSA	0.1	0.5	0.9313	0.5683
Q-Learning	0.1	0.9	0.9353	0.9741
SARSA	0.1	0.9	0.9313	1.4154
Q-Learning	0.9	0.5	0.9353	0.9402
SARSA	0.9	0.5	0.7713	3.3388
Q-Learning	0.9	0.9	0.8967	10.6435
SARSA	0.9	0.9	0.7020	23.0055

Com base nos testes realizados para diferentes combinações de taxa de aprendizado e fator de desconto, observa-se que o *Q-Learning* obteve um desempenho geral superior ao SARSA em termos de taxa de sucesso média e menor variação na *Tabela Q* (*Q-Delta*). O *Q-Learning* apresentou resultados estáveis mesmo em cenários com parâmetros elevados (como taxa de aprendizado = 0.9), enquanto o SARSA demonstrou maior sensibilidade, com quedas na taxa de sucesso e aumento no *Q-Delta* em tais condições. Além disso, a elevação do fator de desconto resultou em um aumento dos valores de *Q-Delta* em ambos os algoritmos, indicando maior oscilação nas estimativas de longo prazo. Isso reforça a robustez do *Q-Learning* em contextos simulados [Sutton and Barto 2018].

Na Figura 4, são apresentados os gráficos de desempenho para ambos os algoritmos com a configuração de parâmetros $\alpha = 0.9$ e $\gamma = 0.9$. O *Q-Learning* tende a manter uma taxa de sucesso elevada e estável ao longo dos episódios, enquanto o SARSA exibe uma performance mais instável, alternando entre períodos de acerto e erro, refletindo maior sensibilidade à exploração.

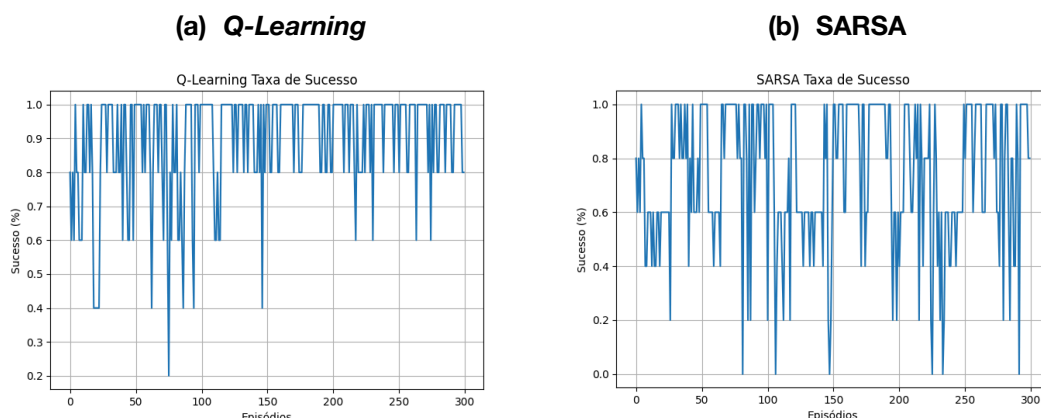


Figura 4. Taxa de sucesso Q-Learning x SARSA

Os resultados obtidos nos experimentos corroboram com os encontrados em Sutton and Barto (2018), onde estratégias baseadas em aprendizado *off-policy* demonstram maior robustez em ambientes com regras variáveis ou que mudam ao longo do tempo. No contexto deste trabalho, essa robustez é particularmente desejável, pois o ambiente simulado altera suas regras a cada certo número de episódios, simulando um cenário em que o jogador adota diferentes estratégias ou comete erros durante o descarte. Isso força os agentes a se adaptarem constantemente à novas políticas de referência para aprendizado.

6. Conclusão

Este artigo apresentou o desenvolvimento de um jogo educacional voltado à promoção da consciência ambiental pelo descarte correto de resíduos. A proposta se destacou por integrar técnicas de IA, como agentes inteligentes e aprendizagem por reforço, onde o jogador exerce papel ativo na construção do comportamento dos agentes não jogáveis.

A simulação implementada em *Python* possibilitou avaliar o desempenho dos algoritmos *Q-Learning* e SARSA em diferentes cenários de aprendizado em um ambiente dinâmico no qual as regras de descarte corretas são alteradas ao longo do tempo. Assim, foi selecionado o modelo mais adequado para implementar no jogo.

Os resultados demonstraram que o algoritmo *Q-Learning* apresentou melhor desempenho em cenários com regras variáveis. Essa característica torna o *Q-Learning* mais adequado ao contexto do jogo proposto, onde a política de descarte é influenciada dinamicamente pelo comportamento do jogador. Essa mecânica permite que os agentes se adaptem continuamente ao ambiente, enquanto o jogador atua como referência comportamental, influenciando o aprendizado mesmo quando comete erros no descarte.

Como trabalhos futuros pretende-se aplicar o jogo *Recycle Learning* com alunos de uma instituição de ensino (Fundamental 1), a qual os autores mantêm parceria. O objetivo é avaliar sua eficácia pedagógica e o impacto na conscientização ambiental.

Agradecimento

O presente trabalho foi realizado com o apoio da Universidade Tecnológica Federal do Paraná - UTFPR.

Referências

- Abreu Neto, H. V.; Boldrin, H. F.; Oliveira, L. E.; Torrens, I. C.; Borges, H. B.; Matos, S. N. (2022) “Dominó Monetário: Um Jogo de Educação Financeira”. In Anais Estendidos do XXI Simpósio Brasileiro de Jogos e Entretenimento Digital. Natal/RN, Brasil. SBC, Porto Alegre, Brasil, 979-988.
- Bae, J.; Eum, K.; Kwon, H.; Lee, S.; Nam, J.; Doh, Y. Y. (2022). “Classy Trash Monster: An Educational Game for Teaching Machine Learning to Non-major Students”. In Conference on Human Factors in Computing Systems – Extended Abstracts. New Orleans, LA. p. 1-7.
- Coleta seletiva. (2025). "Escola Games. Coleta seletiva: jogo educativo", <https://www.escolagames.com.br/jogos/coleta-seletiva>, março.
- CONAMA. Resolução CONAMA nº 275 de 25 de Abril 2001, <https://www.siam.mg.gov.br/sla/download.pdf?idNorma=291>.
- Correa, D. B.; Pereira, A. L.; Borges, H. B.; Bini, T. A.; Matos, S. N.; Torrens, I. C. (2024). “Jogo Educacional sobre Reciclagem para Pessoas com Deficiência Intelectual”. In XXXV Simpósio Brasileiro de Informática na Educação (SBIE 2024). Rio de Janeiro, RJ p. 1406-1416.
- Dever me CHAMA. (2025). "Ludo Educativo. O dever me CHAMA: jogo educativo", <https://www.ludoeducativo.com.br/pt/play/o-dever-me-chama>, fevereiro.
- Foad, D.; Ghifari, A.; Kusuma, M. B.; Hanafiah, N.; Gunawan, E. (2021). A systematic literature review of a* pathfinding. In *Procedia Computer Science*, v. 179, p. 507–514.
- Gaggi, O.; Meneghello, F.; Palazzi, C. E.; Pante, G. (2020). “Learning how to recycle waste using a game”. In *Proceedings of the 6th EAI International Conference on Smart Objects and Technologies for Social Good*. New York, NY, USA. p. 144-149.
- ONU, 2015. Agenda 2030 para o Desenvolvimento Sustentável. Disponível: <https://brasil.un.org/pt-br/91863-agenda-2030-para-odesenvolvimento-sustentavel>. Acesso em: 20 Abr. 2025.
- Rummery, G. A.; Niranjan, M. (1994). On-line Q-learning using connectionist systems (Technical Report CUED/F-INFENG/TR 166). Cambridge University Engineering Department.
- Russell, S.; Norvig, P. (2021). *Artificial Intelligence: A Modern Approach*, Pearson Education Limited, 4th edition.
- Skalee, A. A.; Kliszcz, S.; Parreira, F. J.; Silveira, S. R. (2017). Fredi no Mundo da Reciclagem: Jogo Educacional Digital para Conscientização da Importância da Reciclagem. In *Revista Novas Tecnologias na Educação*, v. 15, n. 1, p. 1-11.
- Sutton, R. S.; Barto, A. G. (1998). *Reinforcement Learning: An Introduction*. Cambridge, MA: MIT Press.
- Sutton, R. S.; Barto, A. G. (2018) *Reinforcement learning: An introduction*. 2. ed. Estados Unidos da América: MIT press.

- Silva Teixeira, T.; Borges, H. B.; Matos, S. N.; da Luz, V. S. G.; Ribeiro, T. L. S. (2020). PegAgente: Modelagem de Agentes por Aprendizado de Reforço em Jogos Educacionais. In *Revista Novas Tecnologias na Educação*, v. 18, n. 2, p. 225.
- Watkins, C. J. C. H. (1989). Learning from delayed rewards. PhD Thesis, University of Cambridge, England.
- Watkins, C. J.; Dayan, P. (1992). “Technical note q-learning”. *Machine Learning*.
- Zhang, S.; Li, M.; Zuo, Z.; Niu, Z. (2023). Recent advances in plastic recycling and upgrading under mild conditions. In *Green Chemistry*. v. 25, p. 6949-6970. The Royal Society of Chemistry.