

Avaliando um Jogo Educacional para o Ensino de Inteligência Artificial - Qual Metodologia para Avaliação Escolher?

Thamyla Lima, Alex Barradas Filho, Allan K. Barros, Davi Viana, João B. Bottentuit Junior, Luis Rivero

Universidade Federal do Maranhão – UFMA

thamyla.lima@aluno.ecp.ufma.br, {alex.barradas, allan.kardec, davi.viana, joao.batista, luis.rivero}@ufma.br

Abstract. *Serious games represent a teaching strategy that allows the addition of entertainment with educational elements. To assess the impact that serious games have on computer students, the literature provides a variety of assessment methods. In this paper, a comparative analysis between evaluation methods was carried out, identifying trends, positive and negative points, as well as aspects to be improved. Subsequently, three evaluation methods were selected and applied to students at UFMA to evaluate a game for teaching artificial intelligence. The results obtained were compared, analyzed and discussed according to the dimensions covered by each method.*

Resumo. *Jogos sérios representam uma estratégia de ensino que permitem agregar o entretenimento com elementos educacionais. Para avaliar o impacto que os jogos sérios exercem nos alunos de computação, a literatura disponibiliza uma variedade de métodos avaliativos. Neste trabalho, uma análise comparativa entre métodos de avaliação foi realizada identificando tendências, pontos positivos e negativos, bem como aspectos a serem melhorados. Posteriormente, três métodos de avaliação foram selecionados e aplicados em alunos da UFMA para avaliar um jogo para ensino de inteligência artificial. Os resultados obtidos foram comparados, analisados e discutidos conforme as dimensões abordadas por cada método.*

1. Introdução

As Tecnologias de Informação e Comunicação promoveram grandes mudanças nas formas de comunicação e educação, com a introdução de dispositivos computacionais e os avanços tecnológicos da internet. Esse cenário permitiu a proliferação de ferramentas auxiliaadoras no processo de ensino e aprendizagem, o que proporcionou aos estudantes maior interação e desafios pela aplicação de ambientes que simulam situações específicas de um determinado objeto de estudo (COLLINS; HALVERSON, 2010). Nesse contexto, os jogos sérios são exemplos de ferramentas que permitem a simulação de tarefas em ambientes virtuais pela realização de práticas que exercitem o conhecimento e proporcionem ludicidade, ou seja, que promovam prazer durante a sua execução (BATTISTELLA e VON WANGENHEIM, 2016).

A capacidade dos jogos sérios de interação, reflexão e aprendizagem devem ser validadas. Nesse sentido, a literatura disponibiliza uma variedade de métodos avaliativos de jogos sérios (PETRI et al., 2018). Esses métodos avaliativos, classificados como qualitativos ou quantitativos, avaliam diferentes aspectos do jogo sério, onde diferem, basicamente, na forma de aplicação, análise dos resultados e tamanho da amostra. Neste artigo, foi realizado um estudo de modelos avaliativos com objetivo de aplicar, comparar e discutir os resultados de três diferentes abordagens de avaliação, selecionados de acordo com critérios de popularidade, abrangência e simplicidade na aplicação.

2. Metodologia

Neste trabalho, buscou-se realizar uma comparação entre três abordagens utilizadas para avaliação de jogos sérios, visto que os métodos de avaliação de tais jogos se diferenciam na forma de aplicação, análise de resultados e quais elementos são avaliados em um jogo. A partir dos resultados de uma revisão da literatura, foram selecionados três trabalhos de técnicas de avaliação com as seguintes condições: o maior número de citações (popularidade), o mais completo (abrangência) e o mais simples (simplicidade de aplicação): (a) Popularidade: EGameFlow (Fu; Su; Yu, 2009); (b) Abrangência: MEEGA+ (Petri; Von Wangenheim; Borgatto, 2017) e (c) Simplicidade na aplicação: modelo proposto por Giannakos (2013).

Selecionou-se o Edubot para a realização do experimento, pois o jogo tem como objetivo exercitar os conhecimentos sobre agentes inteligentes utilizando uma metodologia de aprendizagem baseada em problemas (Lima et al., 2018). Em termos gerais, o Edubot permite aos alunos o desenvolvimento de soluções genuínas pela implementação de agentes reativos simples e agentes baseados em modelos conforme as exigências de cada cenário.

O jogo sério foi aplicado com 22 alunos de uma turma mista (curso de Ciência da Computação e Bacharelado Interdisciplinar em Ciência e Tecnologia) da disciplina de Inteligência Artificial da Universidade Federal do Maranhão. A dinâmica ocorreu em dois dias. No primeiro dia, foi realizada a instrução dos alunos sobre as regras e funcionamento do jogo, enquanto, no segundo dia foi realizada a implementação das soluções geradas. Ressalta-se que o acesso ao jogo somente foi garantido durante a execução da dinâmica, sem a possibilidade de consulta e de forma individual.

3. Resultados

Foi aplicado um questionário para coletar dados demográficos dos participantes. De forma geral, os alunos pertenciam à faixa etária de 18 a 28 anos, e apenas 1 aluno da faixa etária de 29 a 39 anos. Além disso, 77,2% dos alunos possuía o hábito de jogar pelo menos 1 vez por mês. Em outros termos, essas informações indicam a afinidade e o convívio do grupo de alunos com a cultura de jogos, o que lhes permite uma análise crítica mais detalhada do assunto.

Foram utilizados os três instrumentos fornecidos pelas três técnicas de avaliação: EGameFlow, MEEGA+ e o modelo proposto por Giannakos (2013). Foram avaliadas as dimensões usabilidade, confiança, desafio, satisfação/emoções, interação social, diversão/entretenimento, atenção focada/concentração, relevância, aprendizado, objetivos claros, *feedback*, autonomia, imersão e aceitação. A seguir, são apresentados os resultados do comparativo das técnicas e os resultados da avaliação geral do jogo considerando as dimensões.

3.1 Resultados do Comparativo das Técnicas

A Tabela 1 apresenta os resultados da comparação das técnicas identificadas na literatura. Os modelos MEEGA+ e EGameFlow abrangeram uma maior quantidade de dimensões. No entanto, o processo de aplicação dos modelos demonstra inconvenientes quanto ao tempo de execução e exaustão dos envolvidos. Em relação ao modelo proposto por Giannakos (2013), o tempo de aplicação é reduzido quando comparado aos demais, mas a ausência da análise de algumas dimensões pode proporcionar deduções superficiais.

Tabela 1 – Comparativo dos aspectos avaliados nos jogos educacionais pelas técnicas EGameFlow, MEEGA+ e Modelo de Giannakos

Aspecto Avaliado / Técnica	EGameFlow	MEEGA+	Modelo de Giannakos
Percepção de aprendizado: verifica se o usuário acredita ter aprendido após jogar o jogo	Itens ¹	Itens	Teste de desempenho para verificar a evolução no aprendizado
Usabilidade/Autonomia: avaliar a facilidade de uso e verifica se o usuário sente que possui controle sobre o jogo (menu, ações, interações com o jogo) e pode escolher a estratégia que lhe convir	Itens	Itens	Não avalia
Diversão/Entretenimento: possui um relacionamento positivo entre a diversão/motivação e o aumento do aprendizado	Não avalia	Itens	Itens
Satisfação/Emoções: o grau de satisfação do usuário em relação ao jogo	Não avalia	Itens	É avaliada a dimensão emoções que verifica o grau de felicidade do jogador ao utilizar o jogo
Desafio: avaliar se o jogo oferece novos obstáculos, situações ou variações	Itens	Itens	Não avalia
Imersão/Atenção focada/Aceitação: corresponde a um estado em que o jogador se sente como parte do jogo e temporariamente esquece as tarefas do dia a dia	Itens e ainda avalia o envolvimento do jogador	Há uma dimensão chamada de atenção focada que avalia os mesmos aspectos	Avalia se o usuário tem a pretensão de continuar a utilizar o jogo através da dimensão aceitação
Relevância: é semelhante a dimensão de percepção de aprendizado, onde é verificado se o usuário acredita que o jogo é relevante tanto para os seus interesses como para o de outros	Não avalia	Itens	Não avalia
Objetivos claros: corresponde a objetividade do jogo, ou seja, as regras e objetivos do jogo devem estar claros para os usuários	Itens	Não avalia	Não avalia

¹ Itens são afirmações ou perguntas feitas aos participantes e estes devem responder de acordo com uma escala Linkert determinada pelo modelo de avaliação.

Aspecto Avaliado / Técnica	EGameFlow	MEEGA+	Modelo de Giannakos
Feedback: verifica se é dado feedback para o jogador do seu progresso no jogo. Permite que o jogador determine o conhecimento necessário que ele precisa adquirir para que possa completar a tarefa.	Itens	Não avalia	Não avalia
Interação social: é uma dimensão importante a ser avaliada já que promove um sentimento de ambiente compartilhado, conexão com outros jogadores e competição	Itens e ainda verifica se o jogo permite a criação de comunidades	Itens	Não avalia
Confiança: corresponde ao sentimento do usuário de que está aprendendo e de que consegue usar o jogo com facilidade	Não avalia	Itens	Não avalia
Concentração: se assemelha às dimensões imersão e atenção focada, diferenciando-se por avaliar a carga de trabalho	Itens	Não avalia	Não avalia

3.2 Resultados da Avaliação do Jogo Edubot

Para avaliar o jogo educacional, os alunos preencheram os questionários de avaliação das técnicas respondendo seu grau de concordância com os itens de cada um dos aspectos descritos na Tabela 1. Por motivos de espaço, estes itens não serão apresentados mas podem ser encontrados nas respectivas publicações (Fu; Su; Yu, 2009)(Petri; Von Wangenheim; Borgatto, 2017)(Giannakos, 2013). Ao todo, a mediana na maioria dos itens avaliados foi 5 concordo totalmente ou 4 concordo parcialmente (numa escala de likert de 5 pontos) ou 7 ou 6, numa escala likert de 7 pontos. Isto indica que houve uma avaliação positiva nos itens avaliados. Os itens onde houve maior discordância são apresentados a seguir.

Com relação à dimensão **Usabilidade/Autonomia**, os alunos indicaram média ou alta discordância em itens relacionados a: controle do menu (como iniciar, parar, salvar etc.); uso livre de estratégias; design do jogo atraente (interface, gráficos, etc.) e recuperação fácil após cometer um erro. O que indica necessidade de melhorar a facilidade de uso do jogo. Com relação dimensão **Desafio**, no EGameFlow é avaliado o suporte ao usuário. Como o Edubot não possui nenhum tipo de suporte online, o índice da opção de discordância foi consideravelmente maior do que nos outros itens. Isto indica que os usuários acharam que o sistema não: fornece suporte online que ajuda a superar os desafios e fornece vídeo ou áudio auxiliares que ajudam a superar os desafios.

Nos modelos MEEGA+ e EGameFlow, os níveis de concordância sobre interação social não foram tão altos como nas demais dimensões. Uma justificativa para isso é que o Edubot não é um jogo para ser jogado em grupos, já que o usuário deve programar o comportamento do agente sozinho. Finalmente, a performance do Modelo de Giannakos foi mensurada pela conclusão do desafio do primeiro cenário, cuja solução foi baseada na implementação de um agente reativo simples. Em geral, os 22 alunos analisaram o problema do primeiro cenário e conseguiram alcançar o objetivo com sucesso, alguns com soluções genuínas, o que permite entender o potencial do emprego da aprendizagem baseada em problemas e dos jogos sérios no desenvolvimento crítico do aluno, na interpretação de desafios e na criação de respostas. O segundo cenário foi acordado como um desafio bônus para os alunos, os quais deveriam implementar como resposta uma estrutura de agente reativo

baseado em modelo. Nesse contexto, nenhum aluno concluiu o segundo cenário, apenas 11 alunos o concluíram parcialmente. Estes dados estão provavelmente associados ao aumento da complexidade do problema e ao tempo oferecido para a execução das atividades.

4. Conclusões

Este trabalho apresentou uma análise comparativa entre modelos de avaliação de jogos sérios. Para isso, foram selecionados os modelos: MEEGA+, EGameFlow e o modelo desenvolvido por Giannakos (2013). A aplicação dos três modelos ocorreu em um jogo sério para exercitar conceitos de agentes inteligentes (Edubot) e permitiu comparar os modelos em nível de dimensões e de perguntas.

Ao aplicar os modelos de avaliação foi possível perceber perguntas repetidas e confusas do ponto de vista dos discentes, que indicaram isto durante e/ou após o preenchimento dos questionários. Outro aspecto importante é que certas dimensões continham perguntas questionando a presença de certos fatores que não existiam no Edubot. Por exemplo, em interação social, o jogo não obteve níveis de avaliação significativos por não possibilitar a criação de comunidades dentro e fora do jogo. Contudo, o Edubot é um jogo preferencialmente individual, com intuito de que o aluno, sozinho, consiga programar o comportamento do agente. Nessa mesma vertente, a dimensão desafio avalia se o jogo possui suporte online, que também é um fator que não está presente no jogo. Como trabalho futuro, pretende-se realizar novas execuções destes questionários com um número maior de discentes e jogos distintos, para identificar que aspectos podem ser melhor avaliados por qual modelo, além de obter dados sobre a percepção dos discentes sobre outros jogos aplicáveis ao ensino de computação em diferentes disciplinas.

Referências

- BATTISTELLA, P. E., & VON WANGENHEIM, C. G. (2020, January). Caracterização do Público-Alvo de Jogos Educacionais na área da Computação. In Anais do XXIV Workshop sobre Educação em Computação (pp. 21-30). SBC.
- COLLINS, A.; HALVERSON, R. The second educational revolution: rethinking education in the age of technology. *Journal of Computer Assisted Learning*, vol 26, p. 18-27, 2010.
- FU, F. L.; SU, R. C.; YU, S. C. Egameflow: A scale to measure learners' enjoyment of e-learning games. *Computers & Education*, v. 52, p. 101–112, 2009.
- GIANNAKOS, M. N. Enjoy and learn with educational games: Examining factors affecting learning performance. *Computers & Education*, v. 68, p. 429–439, 2013.
- LIMA, T.; MENEZES, R.; BARRADAS FILHO, A. O.; SANTOS, D.; BOTTENTUIT JUNIOR, J. B. Edubot: um estudo prático de aprendizagem baseada em problemas no contexto de agentes inteligentes e jogos sérios. *Tecnologias na Educação*, v. 27, p. 6, 2018.
- PETRI, G., CALDERÓN, A., VON WANGENHEIM, C. G., BORGATTO, A. F., & RUIZ, M. (2018, July). Benefícios dos jogos não-digitais no ensino de computação. In Anais do XXVI Workshop sobre Educação em Computação. SBC.
- PETRI, G.; VON WANGENHEIM, C. G.; BORGATTO, A. F. A large-scale evaluation of a model for the evaluation of games for teaching software engineering. *IEEE/ACM 39th International Conference on Software Engineering: Software Engineering Education and Training Track (ICSE-SEET)*, p. 180–189, 2017.