

Uso de juízes *online* no ensino de programação: desafios e oportunidades

Carla A. D. M. Delgado [Moderadora]¹, Leandro S. Galvão de Carvalho²,
Reinaldo Silva Fortes³

¹ Instituto de Computação – Universidade Federal do Rio de Janeiro (UFRJ)

² Instituto de Computação – Universidade Federal do Amazonas (UFAM)

³ Departamento de Computação – Universidade Federal de Ouro Preto (UFOP)

carla@ic.ufrj.br, galvao@icomp.ufam.edu.br, reifortes@ufop.edu.br

Resumo. Juízes online consolidaram-se como ferramentas essenciais para o ensino-aprendizagem de programação, ao viabilizar correção automática, feedback rápido e escalabilidade. Contudo, seu uso predominante ainda reproduz um modelo centrado na avaliação da correção funcional e na escrita de código a partir de uma tela em branco, o que limita o desenvolvimento de habilidades cognitivas amplas. Alinhada com os Grandes Desafios da Educação em Computação no Brasil 2025–2035, esta mesa temática propõe discutir o desafio de transformar plataformas originalmente competitivas em instrumentos pedagógicos eficazes. Serão debatidos relatos de experiências com as plataformas Machine Teaching, CodeBench e opCoders Judge, abordando frentes como a diversificação de exercícios, interoperabilidade, avaliações formais e o impacto da IA generativa. Como resultados concretos, a mesa busca: (1) esboçar uma agenda de requisitos pedagógicos para a próxima geração de juízes; (2) estabelecer um grupo de trabalho para padronização de metadados e intercâmbio de exercícios; e (3) sintetizar diretrizes sobre o uso ético e pedagógico de IA no contexto da Educação em Computação no Brasil.

1. Visão geral

Nas últimas décadas, os juízes *online* tornaram-se uma ferramenta central no ensino-aprendizagem de programação, pois viabilizam correção automática, resposta rápida e gestão de grandes volumes de submissões, as quais são compiladas e testadas em um ambiente homogêneo [Wasik et al. 2018]. Tais plataformas são amplamente empregadas no ensino-aprendizagem de disciplinas de programação [Liu et al. 2023]. No Brasil, existem algumas plataformas mantidas por universidades públicas: *Machine Teaching*¹ (UFRJ/UNIRIO); *CodeBench*² (UFAM); *opCoders Judge*³ (UFOP); *The Huxley*⁴ (UFAL); *BOCA Online Contest Administrator* [Campos and Ferreira 2004] (USP).

No entanto, a maioria dessas ferramentas foi inicialmente desenvolvida para o contexto de programação competitiva, em que se pressupõe domínio prévio da atividade de programar. Sua adoção como instrumento pedagógico, portanto, não é direta:

¹<https://www.machineteaching.tech/>

²<https://codebench.icomp.ufam.edu.br/>

³<https://www.opcoders.decom.ufop.br/>

⁴<https://thehuxley.com/>

ela exige adaptações de *design*, expansão funcional e revisão das práticas de avaliação [Barbosa et al. 2023]. Nesse sentido, o eixo central desta mesa temática é **apresentar as ferramentas atualmente utilizadas no âmbito nacional e, com base em exemplos concretos, discutir suas limitações e os desafios relacionados ao desenvolvimento, à manutenção e à pesquisa de Educação em Computação por meio de juízes online**.

Os *Grandes Desafios da Educação em Computação no Brasil 2025–2035* [Motta et al. 2025] identificam dois desafios diretamente relacionados a esta mesa. O desafio **Métodos e Técnicas de Ensino/Aprendizagem** aponta a necessidade de abordagens pedagógicas que vão além da transmissão de conteúdo, desenvolvendo habilidades cognitivas mais amplas no contexto da programação. Já o desafio **Inteligência Artificial e Educação** destaca o potencial transformador da IA generativa para a personalização do ensino, a geração automatizada de exercícios e feedback e a revisão das práticas avaliativas, ao mesmo tempo em que alerta para os riscos à integridade acadêmica. Esta mesa temática se posiciona na intersecção de ambos os desafios, ao propor discutir como aprimorar juízes *online* como instrumentos pedagógicos e como integrar a IA generativa a essas plataformas de forma responsável.

A partir do eixo central, a mesa articula seis frentes de discussão, derivadas desses dois grandes desafios: **(1) Feedback pedagógico ampliado:** Juízes *online* avaliam predominantemente a correção funcional do código por meio de casos de teste (avaliação dinâmica). Essa ênfase exclusiva no resultado pode fornecer uma visão limitada da atividade de programação, uma vez que soluções funcionalmente corretas podem decorrer de raciocínios inadequados ou de recursos não permitidos. Incorporar critérios pedagógicos mais ricos ao *feedback* representa uma direção promissora [Hazzan et al. 2020]. **(2) Diversificação dos tipos de exercícios:** Além da escrita de código a partir de uma tela em branco, diferentes tipos de atividades favorecem o desenvolvimento de habilidades cognitivas distintas: preenchimento de lacunas, localização e correção de erros, rastreamento de execução e reorganização de blocos de código (Problemas de Parsons). A exploração sistemática dessas variações aprofunda a compreensão dos conceitos de programação [Simões and Queirós 2020]. **(3) Interoperabilidade e intercâmbio de exercícios:** Reduzir a carga de trabalho dos instrutores por meio de repositórios abertos de exercícios [Sena and Reis 2025] e de formatos padronizados de intercâmbio entre plataformas [Strickroth et al. 2015] constitui um desafio técnico e organizacional relevante para a comunidade. **(4) Uso em avaliações formais:** A discrepância entre o ambiente de estudo (com suporte computacional) e o de avaliação (lápiz e papel) pode prejudicar o desempenho dos estudantes. Para que avaliações sejam realizadas com suporte de juízes *online*, as plataformas precisam de funcionalidades específicas que garantam integridade e equidade nos processos avaliativos. **(5) Sustentabilidade das plataformas:** Como qualquer sistema de *software*, juízes *online* demandam manutenção, atualização, hospedagem e segurança, tarefas que exigem tempo e recursos financeiros frequentemente escassos no ambiente universitário. **(6) Inteligência Artificial Generativa:** Este é, atualmente, o tópico de maior dinamismo e impacto potencial. Três dimensões merecem atenção especial: (A) *Geração automática de exercícios e variantes:* A IA generativa pode apoiar a criação de novos exercícios, casos de teste e exemplos adaptados ao nível do estudante, reduzindo o custo de elaboração de materiais pedagógicos. (B) *Personalização do feedback:* Ao atuar como tutor adaptativo, a IA pode oferecer *feedback* mais contextualizado e individualizado do que os métodos tradicionais de comparação de saída, apoiando a

identificação de *misconceptions* e a recomendação de exercícios. (C) *Integridade acadêmica e condições pedagógicas*: O uso de IA generativa pelos estudantes levanta questões sobre integridade e requer reflexão sobre para qual perfil de programador(a) as atividades devem ser desenhadas. Sua integração deve ser estruturada a partir de um contexto pedagógico bem definido.

Esperamos identificar pessoas interessadas no uso e desenvolvimento deste tipo de ferramenta, visando construir uma comunidade sólida e abrangente, fortalecendo as iniciativas existentes e fomentando a democratização do seu uso por outras instituições. Além de consolidar uma rede de cooperação entre os desenvolvedores das ferramentas brasileiras citadas, a mesa pretende: (1) Esboçar uma agenda de requisitos pedagógicos para a próxima geração de juízes *online*; (2) Estabelecer um grupo de trabalho para discutir a padronização de metadados de exercícios, facilitando o intercâmbio entre diferentes juízes *online*; (3) Produzir um relato técnico pós-evento sintetizando as visões da comunidade sobre o uso ético de IA nessas plataformas.

O uso de ferramentas computacionais como suporte ao ensino-aprendizagem é essencial, especialmente no ensino de programação, no qual o emprego de juízes *online* torna-se mandatório. A convergência de iniciativas e a atuação de uma comunidade heterogênea e multidisciplinar configuram um caminho promissor para enfrentar os desafios envolvidos. Nesse contexto, o EduComp apresenta-se como um ambiente propício para iniciar e fortalecer essa convergência.

2. Estrutura e dinâmica

A mesa será conduzida a partir da apresentação dos pontos fortes e desafios de alguns dos juízes *online* em atividade no Brasil, em especial: (1) **Machine Teaching**: ferramenta usada para atividades práticas dos estudantes, com forte apoio de recursos de análise de dados para gestão do curso e facilitadores para atividades de correção por parte de monitores e professores [Moraes et al. 2022, Xará et al. 2025]; (2) **CodeBench**: sistema usado por professores para gerenciar turmas, nas quais disponibilizam materiais didáticos e exercícios de programação para fins de prática pedagógica, bem como para fins de avaliação somativa [Carvalho et al. 2016]; (3) **opCoders Judge**: ferramenta amplamente utilizada em disciplinas de introdução à programação na UFOP, em diversos cursos de graduação em Ciências Exatas. Nasceu para preencher uma lacuna na utilização de juízes *online*, com um olhar mais atento ao ponto de vista pedagógico e às questões de usabilidade e experiência dos usuários [Brito and Fortes 2019, Silva et al. 2024, Silva et al. 2025].

A programação prevista consiste em: (1) A moderadora abrirá a mesa com uma breve descrição dos objetivos e de seus componentes (tempo estimado: 10 minutos); (2) Os componentes da mesa apresentarão relatos dos juízes *online* elencados e da experiência de uso, desafios e oportunidades percebidos em suas instituições (tempo estimado: 10 minutos para cada ferramenta); (3) Discussão entre os componentes da mesa sobre as intersecções entre demandas, oportunidades e pontos de colaboração (tempo estimado: 20 minutos); (4) Perguntas orientadoras para a condução de discussões, que serão priorizadas e colocadas em debate pela moderadora, oportunizando a participação do público⁵. Como ponto de partida, listamos três potenciais perguntas a serem discutidas: (P1) *É possível*

⁵Ferramentas como *Sli.do* e *Mentimeter* poderão ser adotadas para mediar a interação com a audiência, a depender do número de pessoas presentes.

automatizar o feedback pedagógico sem perder a profundidade da avaliação humana?; (P2) Como garantir a integridade acadêmica em juízes online diante da onipresença de ferramentas de IA?; (P3) Quais padrões mínimos de interoperabilidade a comunidade brasileira deve adotar para viabilizar o intercâmbio de exercícios? (P4) Quais possibilidades de colaboração entre grupos, nos pontos de vista pedagógico e de pesquisa? (tempo estimado: 40 minutos). O tempo total estimado é de 90 minutos.

Para a realização da mesa temática, necessita-se de um projetor multimídia e, eventualmente, de um sistema de som (a depender do espaço e do número de participantes). É desejável uma sala com liberdade para alterar a disposição das carteiras, visando à otimização do espaço e à estimulação do debate e da interação entre os participantes.

3. Apresentação dos proponentes

Carla Delgado: é professora do Instituto de Computação da Universidade Federal do Rio de Janeiro (UFRJ) desde 2011 e atualmente vice-diretora da unidade. Atua na graduação e na pós-graduação, sendo vice-coordenadora do mestrado profissional PROF-COMP/UFRJ. Desenvolve projetos em inteligência artificial e análise de dados aplicados à educação, com interesses em tecnologias educacionais e ensino de computação. Possui formação completa em Computação e graduação em Pedagogia.

CV: <http://lattes.cnpq.br/3831909651244142>

Leandro Galvão: é professor da Universidade Federal do Amazonas (UFAM) desde 2006 e atualmente realiza pós-doutorado na UnB. Atuou como coordenador do curso de Ciência da Computação e como Coordenador Acadêmico do IComp. É membro ativo da SBC, com participação em comissões nacionais e coordenação de eventos da área. Atua no ensino de programação introdutória e pesquisa em Educação em Computação, com publicações baseadas em dados do juiz *online CodeBench*.

CV: <http://lattes.cnpq.br/6049960144667044>

Reinaldo Silva Fortes: é professor da Universidade Federal de Ouro Preto (UFOP) desde 2011, atuando no ensino de programação e sistemas de recomendação em cursos de Computação e Engenharias. Sua pesquisa concentra-se em inteligência artificial e informática na educação. É responsável pelo desenvolvimento do *opCoders Judge* e participa de projetos de inovação tecnológica que aplicam Realidade Virtual e Aumentada na educação básica, como o *EducaMinasVR*.

CV: <http://lattes.cnpq.br/1637967805937656>

Referências

Barbosa, A., Costa, E., and Brito, P. (2023). Juízes Online são suficientes ou precisamos de um VAR? In *Anais do III Simpósio Brasileiro de Educação em Computação*, pages 386–394, Porto Alegre, RS, Brasil. SBC.

Brito, P. S. S. and Fortes, R. S. (2019). O uso de corretores automáticos para o ensino de programação de computadores para alunos de engenharia. In *Anais do XXX Simpósio Brasileiro de Informática na Educação (SBIE 2019)*, page 449. SBC.

- Campos, C. P. and Ferreira, C. E. (2004). BOCA: um sistema de apoio a competições de programação. In *Workshop de Educação em Computação*. SBC.
- Carvalho, L. S. G., Oliveira, D. B. F., and Gadelha, B. F. (2016). Juiz online como ferramenta de apoio a uma metodologia de ensino híbrido em programação. In *Brazilian Symposium on Computers in Education (Simpósio Brasileiro de Informática na Educação-SBIE)*, volume 27, page 140.
- Hazzan, O., Ragonis, N., and Lapidot, T. (2020). *Guide to Teaching Computer Science: An Activity-Based Approach*. Springer, 3 edition.
- Liu, K., Han, Y., Zhang, J. M., Chen, Z., Sarro, F., Harman, M., Huang, G., and Ma, Y. (2023). Who judges the judge: An empirical study on online judge tests. In *Proceedings of the 32nd ACM SIGSOFT International Symposium on Software Testing and Analysis, ISSTA 2023*, page 334–346, New York, NY, USA. ACM.
- Moraes, L. O., Delgado, C. A. D. M., Freire, J. P., and Pedreira, C. E. (2022). Machine teaching: uma ferramenta didática e de análise de dados para suporte a cursos introdutórios de programação. In *Anais do II Simpósio Brasileiro de Educação em Computação*, pages 213–223, Porto Alegre, RS, Brasil. SBC.
- Motta, C. L. R. d., Santoro, F. M., Ribeiro, L., Duran, R., Siqueira, S. W. M., Cavaleiro, S. A. d. C., and Falcão, T. P. (2025). *Grandes Desafios da Educação em Computação no Brasil 2025–2035*. Sociedade Brasileira de Computação, Porto Alegre.
- Sena, J. and Reis, J. (2025). Repositório aberto de exercícios de programação introdutória em português. In *Anais Estendidos do V Simpósio Brasileiro de Educação em Computação*, pages 47–49, Porto Alegre, RS, Brasil. SBC.
- Silva, L. L. M., Cedraz, V. F., Patrocínio, J. A., Delabrida, S., and Fortes, R. S. (2024). Juiz online focado no ensino de programação: uma análise de usabilidade de ferramenta autoral. In *Anais do XXXV Simpósio Brasileiro de Informática na Educação*. SBC.
- Silva, L. L. M., Cedraz, V. F., Patrocínio, J. A., Delabrida, S., and Fortes, R. S. (2025). Rethinking online judges: Usability insights from introductory programming students. In *Proceedings of Interact 2025*. Short Paper.
- Simões, A. and Queirós, R. (2020). On the Nature of Programming Exercises. In Queirós, R., Portela, F., Pinto, M., and Simões, A., editors, *First International Computer Programming Education Conference (ICPEC 2020)*, volume 81 of *Open Access Series in Informatics (OASIs)*, pages 24:1–24:9, Dagstuhl, Germany. Schloss Dagstuhl.
- Strickroth, S., Striewe, M., Müller, O., Priss, U., Becker, S., Rod, O., Garmann, R., Bott, O. J., and Pinkwart, N. (2015). ProFormA: An XML-based exchange format for programming tasks. *eled: e-learning and education*.
- Wasik, S., Antczak, M., Badura, J., Laskowski, A., and Sternal, T. (2018). A Survey on Online Judge Systems and Their Applications. *ACM Comput. Surv.*, 51(1).
- Xará, G., A. D. M. Delgado, C., M. de Farias, C., Folloni Guarilha, H., O. Moraes, L., Freire, J. P., and Nogueira Teixeira, E. (2025). Technical and tactical factors of an educational platform aiming to deal with a large number of students and inequalities when teaching programming for stem students. *Revista Brasileira de Informática na Educação*, 33:80–105.