

AUTOMATIC ANALYSIS AND GRADING OF UML UML DIAGRAMS

Douwe Osinga

d.r.osinga@student.utwente.nl

Supervisor

dr. ir. Vadim Zaytsev

v.zaytsev@utwente.nl

Supervisor

dr. Nacir Bouali

n.bouali@utwente.nl



1. INTRODUCTION

Current state of grading + autograding, University of Twente is looking into ways to save time and money in grading by automating (parts of) it.

2. PROBLEM STATEMENT

Grading takes long time etc. etc. Want a solution that automatically grades UML diagrams (specifically class/sequence/...?), with as main goals: *transparency, consistency, fairness*. See initial plan description.

2.1. Research Questions

3. RELATED WORK

Work [1], [2], [3], [4], [5], [6], [7], [8], [9], [10], [11], [12], [13], [14], [15], [16]

More focused on interactivity: [17]

Work on AI [18], [19] (nondeterminism of AI [20], [21], [22] + counterarg: inherent lack of transparency, risks of nondeterminism in grading (see sources) == bad because same solution might not give same grade), lack of consistency (context window, importance of reducing prompt length, ...)

Experience on TAs [23]

Reliability of human marking/grading in general [24]

4. TOOLS AND TECHNIQUES

Adopt existing tool(s), make own tool, what frameworks/languages, ...

5. PLANNING

TODO: Graduation planning. Phases, goals per phase.

BIBLIOGRAPHY

- [1] M. Hosseinibaghdabadi, O. A. N. Almerge, and J. Kienzle, "Automated Grading of Use Cases," in *2023 ACM/IEEE 26th International Conference on Model Driven Engineering Languages and Systems (MODELS)*, IEEE, 2023. [Online]. Available: <https://ieeexplore.ieee.org/iel7/10343461/10343549/10343598.pdf>
- [2] O. Anas, T. Mariam, and L. Abdelouahid, "New method for summative evaluation of UML class diagrams based on graph similarities," 2021. [Online]. Available: https://www.academia.edu/download/66135135/70_22270_EM_26aug_20feb_L.pdf
- [3] F. Batmaz, "Semi-Automatic Assessment of Students' Graph-Based Diagrams," 2010. [Online]. Available: https://www.academia.edu/download/66135135/70_22270_EM_26aug_20feb_L.pdf
- [4] W. Bian, O. Alam, and J. Kienzle, "Automated Grading of Class Diagrams," in *2019 ACM/IEEE 22nd International Conference on Model Driven Engineering Languages and Systems Companion (MODELS-C)*, IEEE, Sept. 2019, pp. 700–709. doi: [10.1109/models-c.2019.00106](https://doi.org/10.1109/models-c.2019.00106).
- [5] W. Bian, O. Alam, and J. Kienzle, "Is automated grading of models effective?: assessing automated grading of class diagrams," in *Proceedings of the 23rd ACM/IEEE International Conference on Model Driven Engineering Languages and Systems*, in MODELS '20. ACM, Oct. 2020, pp. 365–376. doi: [10.1145/3365438.3410944](https://doi.org/10.1145/3365438.3410944).
- [6] S. Foss, T. Urazova, and R. Lawrence, "Learning UML database design and modeling with AutoER," in *Proceedings of the 25th International Conference on Model Driven Engineering Languages and Systems: Companion Proceedings*, in MODELS '22. ACM, Oct. 2022, pp. 42–45. doi: [10.1145/3550356.3559091](https://doi.org/10.1145/3550356.3559091).
- [7] R. Jebli, J. E. Bouhdidi, and M. Y. Chkouri, "Assessing Students' UML Class Diagrams: a New Automated Solution," in *2023 7th IEEE Congress on Information Science and Technology (CiSt) |*, IEEE, 2023. [Online]. Available: <https://ieeexplore.ieee.org/iel7/10409867/10409868/10409936.pdf>
- [8] S. Modi, H. A. Taher, and H. Mahmud, "A Tool to Automate Student UML diagram Evaluation," 2021. doi: [ng](https://doi.org/10.4236/ng.202111101).
- [9] N. H. Ali, Z. Shukur, and S. Idris, "Assessment System For UML Class Diagram Using Notations Extraction," 2007. [Online]. Available: https://www.researchgate.net/profile/Zarina-Shukur/publication/253243639_Assessment_System_For_UML_Class_Diagram_Using_Notations_Extraction/links/55487af30cf2b0cf7acec2e4/Assessment-System-For-UML-Class-Diagram-Using-Notations-Extraction.pdf
- [10] N. H. Ali, Z. Shukur, and S. Idris, "A Design of an Assessment System for UML Class Diagram," in *2007 International Conference on Computational Science and its Applications (ICCSA 2007)*, IEEE, Aug. 2007, pp. 539–546. doi: [10.1109/iccsa.2007.2](https://doi.org/10.1109/iccsa.2007.2).
- [11] P. Thomas, N. Smith, and K. Waugh, "An approach to the automatic grading of imprecise diagrams," technical report, 2006. doi: [10.21954/ou.ro.00016046](https://doi.org/10.21954/ou.ro.00016046).
- [12] P. Thomas, "Grading Diagrams Automatically," 2004. [Online]. Available: https://oro.open.ac.uk/90155/1/2004_01.pdf

- [13] P. Thomas, N. Smith, and K. Waugh, "Automatically Assessing Diagrams," in *Proceedings of the IADIS International Conference on e-Learning*, 2009. [Online]. Available: https://www.researchgate.net/profile/Pete-Thomas/publication/42799920_Automatically_assessing_diagrams/links/0fcfd5060076dd8ba2000000/Automatically-assessing-diagrams.pdf
- [14] P. Thomas, N. Smith, and K. Waugh, "Automatically assessing graph-based diagrams," *Learning, Media and Technology*, vol. 33, no. 3, pp. 249–267, 2008, doi: [10.1080/17439880802324251](https://doi.org/10.1080/17439880802324251).
- [15] M. Striewe and M. Goedicke, "Automated Checks on UML Diagrams," in *ITiCSE'11*, in ITiCSE '11. ACM, June 2011, pp. 38–42. doi: [10.1145/1999747.1999761](https://doi.org/10.1145/1999747.1999761).
- [16] N. Smith, P. Thomas, and K. Waugh, "Automatic Grading of Free-Form Diagrams with Label Hypernymy," in *2013 Learning and Teaching in Computing and Engineering*, IEEE, Mar. 2013, pp. 136–142. doi: [10.1109/lattice.2013.33](https://doi.org/10.1109/lattice.2013.33).
- [17] S. Foss, "AutoER: A System for the Automatic Generation and Evaluation of UML Database Design Diagrams," 2022. [Online]. Available: <https://open.library.ubc.ca/media/download/pdf/24/1.0421624/4>
- [18] N. Bouali, M. Gerhold, T. U. Rehman, and F. Ahmed, "Toward Automated UML Diagram Assessment: Comparing LLM-Generated Scores with Teaching Assistants," 5220. [Online]. Available: <https://research.utwente.nl/files/496461589/134819.pdf>
- [19] D. R. Stikkorum, P. van der Putten, C. Sperandio, and M. R. Chaudron, "Towards Automated Grading of UML Class Diagrams with Machine Learning," 2019. [Online]. Available: <https://ceur-ws.org/Vol-2491/paper80.pdf>
- [20] H. He and T. Machines, "Defeating Nondeterminism in LLM Inference," 2025, [Online]. Available: <https://thinkingmachines.ai/blog/defeating-nondeterminism-in-lm-inference/>
- [21] M. Brenndoerfer, "Why Temperature=0 Doesn't Guarantee Determinism in LLMs," 2025, [Online]. Available: <https://mbrenndoerfer.com/writing/why-lms-are-not-deterministic>
- [22] B. Atil *et al.*, "Non-Determinism of "Deterministic" LLM Settings," 2025. [Online]. Available: <https://arxiv.org/pdf/2408.04667>
- [23] F. Ahmed, N. Bouali, and M. Gerhold, "Teaching Assistants as Assessors: An Experience Based Narrative," 2024. [Online]. Available: <https://research.utwente.nl/files/457355611/126242.pdf>
- [24] M. Meadows and L. Billington, "A Review Of THe Literature On Marking Reliability." [Online]. Available: https://assets.publishing.service.gov.uk/media/5a820a57e5274a2e87dc0d5a/0505_Meadows_and_Billington_CERP_RP.pdf