

DASTURLASHNI O'QITISHDA AN'ANAVIY TEST TIZIMLARINING CHEKLOVLARI VA O'QUV-BILISH KOMPETENTLIGINI BAHOLASH MUAMMOSI

Xamrayev Xurshidbek Komiljon o'g'li

University of Business and Science stajor – o'qituvchisi

xamrayevbek96@gmail.com

***Annotatsiya:** Mazkur maqolada dasturlashni o'qitishda keng qo'llaniladigan an'anaviy avtomatik baholash tizimlari (ACM Online Judge, Moodle quiz, unit-test asosidagi tekshiruvlar) ning pedagogik cheklovlari tahlil qilingan. 2017–2024 yillardagi avtomatlashtirilgan baholash bo'yicha tizimli sharhlar tahliliga asoslanib, bunday tizimlar faqat kirish-chiqish natijalarini solishtirib, kodning mantiqiy tuzilishini tushunish chuqurligi, talaba yozgan testlarning xatolarni aniqlash qobiliyati, chegaraviy holatlarga tayyorlik va metakognitiv ko'nikmalarni baholay olmasligi ko'rsatilgan. Natijada o'quv-bilish kompetentligini baholash muammosi yuzaga keladi. Maqolada bu muammoning yechimi sifatida talaba testlari sifatini o'lchaydigan yondashuv — mutatsion testlash zarurati asoslanadi.*

***Kalit so'zlar:** dasturlashni o'qitish, avtomatik baholash, test tizimlari, o'quv-bilish kompetentligi, unit-test, metakognitiv ko'nikmalar, mutatsion testlash, teskari aloqa.*

ОГРАНИЧЕНИЯ ТРАДИЦИОННЫХ ТЕСТОВЫХ СИСТЕМ В ОБУЧЕНИИ ПРОГРАММИРОВАНИЮ И ПРОБЛЕМА ОЦЕНИВАНИЯ УЧЕБНО-ПОЗНАВАТЕЛЬНОЙ КОМПЕТЕНТНОСТИ

Аннотация: В статье анализируются педагогические ограничения традиционных систем автоматического оценивания, широко применяемых в обучении программированию (ACM Online Judge, Moodle quiz, проверки на основе модульных тестов). На основе анализа систематических обзоров по автоматизированному оцениванию за 2017–2024 годы показано, что такие системы лишь сравнивают результаты «вход–выход» и не способны оценить глубину понимания логической структуры кода, способность тестов студента выявлять ошибки, готовность к граничным случаям и метакогнитивные навыки. Это порождает проблему оценивания учебно-познавательной компетентности. В качестве решения обосновывается необходимость подхода, измеряющего качество тестов студента, — мутационного тестирования.

Ключевые слова: обучение программированию, автоматическое оценивание, тестовые системы, учебно-познавательная компетентность, модульный тест, метакогнитивные навыки, мутационное тестирование, обратная связь.

THE LIMITATIONS OF TRADITIONAL TESTING SYSTEMS IN PROGRAMMING EDUCATION AND THE PROBLEM OF ASSESSING LEARNING COMPETENCY

Abstract: This article analyses the pedagogical limitations of the traditional automated assessment systems widely used in programming education (ACM Online Judge, Moodle quiz, unit-test-based checks). Drawing on systematic reviews of automated assessment from 2017–2024, it shows that such systems merely compare input–output results and

cannot assess the depth of understanding of a program's logical structure, the ability of a student's tests to detect faults, readiness for boundary cases, or metacognitive skills. This gives rise to the problem of assessing learning competency. As a solution, the article substantiates the need for an approach that measures the quality of the student's own tests — namely, mutation testing.

Keywords: *programming education, automated assessment, testing systems, learning competency, unit testing, metacognitive skills, mutation testing, feedback.*

Kirish (Introduction)

Dasturlashni o'qitishda talabalar sonining ortishi va topshiriqlarni tezkor tekshirish zarurati avtomatlashtirilgan baholash tizimlarining keng tarqalishiga olib keldi. ACM Online Judge tipidagi platformalar, Moodle quiz modullari va unit-test asosidagi avtomatik tekshiruvlar talaba kodini etalon natija bilan solishtirib, deyarli bir zunda baho va teskari aloqa beradi (Keuning, Jeuring & Heeren, 2018). Bunday tizimlar baholashni tezlashtirgani va qayta topshirish imkonini bergani bilan o'quvchi qoniqishini oshiradi (Messer et al., 2024).

Biroq ushbu tizimlarning aksariyati topshiriq to'g'riligini faqat dinamik (unit-test) yoki statik usulda — kirish-chiqish natijasini solishtirish orqali baholaydi. Bu esa baholashni "kod ishlayaptimi?" savoli darajasida cheklab qo'yadi. Tadqiqotning maqsadi — an'anaviy test tizimlarining pedagogik cheklovlarini aniqlash va o'quv-bilish kompetentligini baholash muammosini ilmiy asoslashdan iborat.

Metodlar (Methods)

Tadqiqotda 2017–2024 yillarda chop etilgan avtomatlashtirilgan baholash va teskari aloqa tizimlari bo'yicha tizimli sharhlar (systematic

review) tahlil qilindi; baholanadigan va baholanmayotgan kompetensiya komponentlari taqqoslama tahlil qilindi; dasturiy ta'minot muhandisligidagi test sifatini o'lchash mezonlari pedagogik kontekstga ko'chirildi (Garousi et al., 2020).

Natijalar (Results)

Tahlil shuni ko'rsatdiki, avtomatik baholash tizimlarining aksariyati ob'ektiv tipdagi yoki object-oriented tillardagi topshiriqlarning to'g'riligini dinamik usulda baholaydi (Messer et al., 2024). Biroq quyidagi muhim jihatlar baholanmay qoladi:

1) Kodning mantiqiy tuzilishini tushunish chuqurligi: etalon natija to'g'ri bo'lsa-da, talaba uning mantiqiy tuzilishini tushunmasligi mumkin.

2) Talaba testlarining xatolarni aniqlash qobiliyati: talaba o'zi yozgan testlar haqiqiy xatolarni tuta oladimi yoki yo'qmi — an'anaviy tizimda o'lchanmaydi.

3) Chegaraviy holatlarga tayyorlik: bo'sh kirish, eng katta/kichik qiymat, chegaraviy shartlar ko'pincha tekshirilmaydi.

4) Metakognitiv ko'nikmalar: rejalashtirish, o'zini nazorat qilish va refleksiya kabi ko'nikmalar baholash doirasidan tashqarida qoladi.

Natijada "kod ishlayaptimi?" mezoni chuqur tushunish o'rnini bosadi va talabaning haqiqiy o'quv-bilish kompetentligi yashirin qoladi. Adaptiv baholash bo'yicha tadqiqotlar ham an'anaviy statik testlar talabaning individual darajasi va bilim chuqurligini hisobga olmasligini tasdiqlaydi (A Systematic Review on Assessment in Adaptive Learning, 2024).

Muhokama (Discussion)

Aniqlangan cheklovlar o'quv-bilish kompetentligini baholashning yangi vositasini talab qiladi. Agar an'anaviy tizim talaba kodining natijasini tekshirsa, yangi yondashuv talaba yozgan testlarning sifatini — ya'ni uning xatolarni oldindan ko'ra olish va tahlil qilish qobiliyatini o'lchashi kerak. Dasturiy ta'minot muhandisligida test to'plami sifatini baholovchi yetuk texnologiya — mutatsion testlash mavjud bo'lib (Jia & Harman, 2011), uning pedagogik talqini ushbu cheklovlarni bartaraf etish imkonini beradi.

Mutatsion tahlilning ta'limdagi qo'llanilishi talaba test to'plamlarining sifatini ob'ektiv o'lchashda istiqbolli ekanligini empirik tadqiqotlar ko'rsatmoqda (ISSTA, 2022). Shu bois bu maqola natijalari dissertatsiyada taklif etilayotgan baholash metodikasiga (mutatsion testlash, ESI metrikasi) o'tish uchun muammoviy asos bo'lib xizmat qiladi.

Xulosa (Conclusion)

Dasturlashni o'qitishda an'anaviy avtomatik test tizimlari baholashni tezlashtirgani bilan o'quv-bilish kompetentligining muhim qirralarini — tushunish chuqurligi, test yozish malakasi, chegaraviy holatlar va metakognitiv ko'nikmalarni o'lchay olmaydi. Bu cheklovlar talaba testlari sifatini baholaydigan yangi yondashuvni taqozo etadi. Mutatsion testlashning pedagogik talqini ushbu muammoning samarali yechimi sifatida keyingi tadqiqot yo'nalishini belgilaydi.

Foydalanilgan adabiyotlar (References)

1. A Systematic Review on Assessment in Adaptive Learning: Theories, Algorithms and Techniques. (2024). International Journal of Advanced Computer Science and Applications (IJACSA), 15(7).

2. Garousi, V., Rainer, A., Lauvås, P., & Arcuri, A. (2020). Software-testing education: A systematic literature mapping. *Journal of Systems and Software*, 165. arXiv:2003.03843.
3. Jia, Y., & Harman, M. (2011). An Analysis and Survey of the Development of Mutation Testing. *IEEE Transactions on Software Engineering*, 37(5), 649–678.
4. Keuning, H., Jeuring, J., & Heeren, B. (2018). A Systematic Literature Review of Automated Feedback Generation for Programming Exercises. *ACM Transactions on Computing Education*, 19(1), 1–43. <https://doi.org/10.1145/3231711>
5. Messer, M., Brown, N. C. C., Kölling, M., & Shi, M. (2024). Automated Grading and Feedback Tools for Programming Education: A Systematic Review. *ACM Transactions on Computing Education*, 24(1), 1–43. <https://doi.org/10.1145/3636515>
6. On the use of mutation analysis for evaluating student test suite quality. (2022). *Proceedings of the 31st ACM SIGSOFT International Symposium on Software Testing and Analysis (ISSTA)*, 263–275. <https://doi.org/10.1145/3533767.3534217>