

СИСТЕМЫ АВТОМАТИЗИРОВАННОГО ТЕСТИРОВАНИЯ ПО ПРОГРАММИРОВАНИЮ В ОБРАЗОВАТЕЛЬНОМ ПРОСТРАНСТВЕ

¹Гладких И.Ю., ¹Якушин А.В.

¹ ФГБОУ ВПО «Тульский государственный педагогический университет им. Л.Н. Толстого», Тула, e-mail: tgpu@tula.net

Работа посвящена информационным системам для автоматизированной проверки работ по программированию в образовательном пространстве. Рассматриваются положительные и отрицательные характеристики этих систем как инструмента по контролю знаний и формы представления заданий по программированию. В работе отражены психологические аспекты, связанные с использованием данных систем в качестве текущих и итоговых форм контроля знаний и навыков, в частности стимуляция мотивации к решению задач и исправлению ошибок решений. Для подтверждения оценки психологического воздействия на учащихся было произведено анкетирование студентов, изучающих программирование как основной предмет специальности. Также опрос студентов выявил внутреннюю готовность студентов к применению подобных систем. Рассматривается двойственность систем автоматизированного тестирования по программированию: студентам необходимо применять творческие умения, для преподавателей выполняется формализованная проверка с отчетом о функциональности решений. Дается оценка эффективности использования систем автоматизированного тестирования по программированию.

Ключевые слова: системы автоматизированного тестирования по программированию (САТ), обучение программистов, тестирование решений задач, внедрение систем автоматизированного тестирования.

SYSTEM OF AUTOMATED TESTING PROGRAMS IN LEARNING ENVIRONMENT

¹Gladkikh I. Yu., ¹Yakushin A.V.

¹ Tula State Lev Tolstoy Pedagogical University, Tula, e-mail: tgpu@tula.net

The work is dedicated to information systems for automated verification of programming jobs in the learning environment. Positive and negative characteristics of these systems as tool on control of knowledge and forms of representation of tasks on programming are considered. The paper reflects the psychological aspects associated with the use of these systems as current and final control forms of knowledge and skills, in particular stimulation of motivation to the solution of tasks and correcting errors in solutions. Questioning of the students studying programming as the main subject of specialty was made to confirm assessment of psychological impact on students. Also, a survey of students showed internal readiness of students to the use of such systems. Discusses the duality of systems of the automated testing on programming: students need to apply creative abilities, but for teachers the formalized check with the report on functionality of decisions is carried out. The estimation of efficiency of usage of automated testing in programming.

Keywords: system of automated testing programs (SAT), programmer's education, testing of solutions, the introduction of automated testing.

Образовательная деятельность, как и прочие сферы деятельности человека, подвержена изменениям, современным тенденциям. Одной из таких тенденций является стремление полного или частичного замещения человека автоматикой. Высокие технологии внедряются, расширяют спектр возможностей донесения учебного материала, форм и методов контроля знаний, увеличивают эффективность процесса проверки заданий учащихся. Одним из сравнительно новых методов контроля знаний стал машинный контроль - тестирование с автоматизированной системой проверки результатов. Однако классическое тестирование не предполагает проверки знаний и умений за рамками формы вопрос-ответ или выбора ответа из предложенных вариантов [7]. Классическое тестирование, даже перенесенное в электронную среду, до недавнего времени не было способно оценить

профессиональные навыки и умения по предмету тестирования. Однако для будущих специалистов в сфере программирования такая возможность есть – это системы автоматизированного тестирования решений задач по программированию (CAT).

Что собой представляют CAT?

CAT – это система, позволяющая прогонять программу учащегося через тесты, получая таким образом статистику работоспособности приложения, на основе которой системой же выносится оценка программы. Сами тесты представляют собой запуск тестируемого приложения на определенных входных данных и сравнение результатов работы приложения с эталонным. Существуют различные методы осуществления подбора данных для осуществления тестовых запусков, также существуют и различные шкалы оценивания программы. В зависимости от поставленных критериев можно подобрать соответствующую CAT.

Отрицательные атрибуты CAT

Неявная типизация решений (утиные типы)

Анализ функциональности программного решения в среде CAT происходит в режиме черного ящика, не рассматривая метрику приложения, беря в расчет лишь характеристики качества [2], определение которых не требует рассмотрения внутренней структуры программного кода.

- **Функциональные возможности**

Набор атрибутов, относящихся к сути набора функций и их конкретным свойствам.

- **Надежность**

Набор атрибутов, относящихся к способности программного обеспечения сохранять свой уровень качества функционирования при установленных условиях за установленный период времени.

- **Практичность**

Набор атрибутов, относящихся к объему работ, требуемых для использования и индивидуальной оценки такого использования определенным или предполагаемым кругом пользователей.

- **Эффективность**

Набор атрибутов, относящихся к соотношению между уровнем качества функционирования программного обеспечения и объемом используемых ресурсов при установленных условиях.

Таким образом, ранжирование оценки происходит по так называемому Duck test (с неявной типизацией), и автоматизированная система не сможет отличить алгоритмически различные решения одной и той же задачи, если оба решения в итоге дают один и тот же результат и используют примерно одинаковое количество аппаратных ресурсов.

Предварительная подготовка

Как и любая информационная система, САТ требует предварительной подготовки данных. Зачастую объем необходимых работ по подготовке системы к эксплуатации (здесь мы говорим как об общей инсталляции программного пакета, так и о подготовке отдельных контрольно-измерительных материалов) превосходит объем работ, необходимый для проведения одного контрольного мероприятия вне электронно-информационной среды. Однако в отличие от классической формы информационно-электронная среда позволяет существенно сократить объем подготовительных действий для повторных контрольных мероприятий по тем же материалам или схожему «шаблону» [5].

Ручное выявление причин ошибки

Метод черного ящика не позволяет определить реальных причин некорректной работы приложения, так как не подразумевает семантического анализа программного кода. В связи с чем выявление причин ошибок переходит либо в самостоятельную работу учащегося, либо требует непосредственного участия преподавателя.

Невозможность автоматического определения оригинальности работы по «почерку» автора

Хотя степень оригинальности работы с большим содержанием функциональных записей можно определить, используя метод штампов и сигнатур, САТ не способен распознать характерный стиль оформления программного кода, присущий определенному автору. Что зачастую легко удается преподавателю.

Положительные атрибуты САТ

Объективность оценки

Все оцениваемые программы проходят через одинаковый набор тестов или эквивалентные между собой наборы тестов, которые в свою очередь должны точно соответствовать условиям, указанным в описании задачи. При этом для разных учащихся одна и та же задача может несколько различаться, постановка таких задач соблюдает ряд правил, и любая версия задачи - по сути эквивалент другой версии той же задачи. Таким образом учащиеся, решая одну задачу, получают эквивалентные требования к программе и эквивалентные тесты. Оценка приложения выдается системой по принципу «работает или не работает».

Оцениваемое приложение запускается системой без участия человека. Первичная оценка САТ происходит без непосредственного участия людей, предоставляя тестируемым учащимся эквивалентные между собой условия задачи и проверки результатов-решений. Тогда как вовлечение людей в другие формы контроля знаний существенно выше, а значит

риск возможной субъективности в оценке также повышается. А значит на данный момент САТ обеспечивает наиболее объективную оценку относительно других форм оценки знаний.

Скорость оценки

Неоспоримым преимуществом автоматизированных систем является их высокая скорость обработки данных в сравнении с проверкой, осуществляемой человеком. САТ предоставляет возможность значительно быстрее выдавать результаты проверки знаний и формирование оценки.

Данное обстоятельство расширяет возможности образовательного пространства [3]. До введения систем автоматизированного тестирования (здесь говорим без привязки к конкретному предмету) какой-либо значительный и масштабный контроль знаний мог проходить, лишь задействовав время, необходимое для проведения нескольких занятий. Проверка результатов в большинстве случаев должна была производиться между занятиями или же на самом занятии, но тогда требовалось занять учащихся самостоятельной работой.

Однако новые формы контроля предоставляют возможность быстрого получения результатов проверки, что, несомненно, повышает эффективность учебного процесса за счет увеличения объема изучаемых материалов или освоения практических навыков под непосредственным контролем преподавателя в освободившееся учебное время или же дает дополнительное время преподавателю на подготовку к занятиям и проверку работ с творческой составляющей. Кроме того, быстрое получение результатов способствует повышению мотивации учащихся.

Наглядность оценки

Хотя процесс тестирования проходит в режиме черного ящика, результат тестирования, при должной подготовке САТ и самих тестов, нагляден и способен в некоторой степени оповестить об ошибках в тестируемых приложениях и способах их устранения. Кроме того, автоматизированная система способна провести гораздо больше тестовых запусков приложения при контроле за процессом работы приложения, тем самым существенно увеличивая статистические данные по работе тестируемого приложения. Что является важным фактом в оценке качества любой программы для ЭВМ [8].

Дистанционное тестирование и контроль самостоятельности выполнения заданий

Еще одно свойство, унаследованное САТ от информационных систем, – это возможность дистанционного доступа с полным или ограниченным функционалом. Т.е. преподаватель волен решать, где именно будет проходить тестирование: в учебном помещении или удаленно от него. При этом есть возможность полного контроля за процессом выполнения тестирования учащимся, что недоступно при классической форме

тестирования на бумажном носителе без привлечения аппаратной составляющей (видеокамер) или нескольких дополнительных наблюдателей. Современные системы тестирования позволяют контролировать деятельность тестируемых учащихся как в реальном мире – через веб-камеру и микрофон, так и в виртуальном – через отслеживание изображения с экрана компьютера – во время теста.

Массовость и объемность тестирования

Для проверки любой работы, демонстрирующей уровень знаний учащегося, преподавателю необходимо затратить свое рабочее, а иногда и личное время прямо пропорционально сложности работы и количеству необходимых тестов. То же самое относится и к работам по программированию: как любая практическая реализация, они нуждаются в многократном тестовом использовании, при этом при недостаточном количестве и качестве тестов полученная оценка может существенно отличаться от реального качества работы. Кроме того, полный цикл проверки необходимо произвести для всей группы учащихся. В итоге получаем монотонную объемную работу, которую преподаватель может частично (причем эта переложенная часть по времени обработки человеком может занимать куда больше времени, чем оставшаяся) переложить на САТ. САТ может в кратчайшие сроки провести полный цикл тестирования как одного учащегося, так и всей группы сразу, при этом осуществляя параллельный контроль за процессом тестирования. САТ способствует устранению «замысленности взгляда» преподавателя при проверке работ, возникающей из-за многочисленных однообразных действий, снимая с преподавателей груз рутинной проверки.

Антиплагиат

Одной из замечательных свойств САТ является возможность ее использования для частичной проверки на плагиат тестируемых приложений. К сожалению, применяемые на данный момент методики выявления плагиата в САТ недостаточны для анализа простых задач с решением небольшим программным кодом. Однако САТ позволяет с достаточной долей точности определить схожие и аналогичные структуры алгоритмических конструкций в решениях, требующих творческого подхода и комбинирования других различных алгоритмов.

Каково отношение к САТ студентов?

Анонимный опрос проводили среди студентов 2-4 курсов ТГПУ им. Л.Н. Толстого, изучающих программирование как один из профильных предметов. Опрос группы показал, что подавляющее большинство (более 90%) по тем или иным причинам заинтересованы в изучении программирования, предмет для них интересен или актуален для выбора будущей профессии.

Проведенное анкетирование студентов подтвердили наблюдения Кречетникова [6] относительно напряженности, возникающей у учащихся при длительном ожидании результатов контроля. Согласно опросам, использование САТ значительно повышает мотивацию учащихся к дальнейшему изучению предмета. А неудачные тесты заставляют проявлять самостоятельную деятельность в изучении и устранении ошибок в решениях как во время занятий, так и вне стен учебного помещения.

В то же время сам результат тестирования важен и в случае ошибок в решении - студенты стремятся исправить их и предоставить максимально верное решение. Так, на вопросы типа «Если что-то не получается при выполнении теста, Вы стараетесь разобраться?» и «Вы стараетесь исправить свое решение, если тест не пройден полностью?» студенты в более чем в 90% случаев ответили «Да». Что наглядно демонстрирует эффективное воздействие соревновательного стимула в совокупности с повышением самооценки за счет преодоления трудностей [9].

Важно для студентов и время получения результатов тестирования. Хотя на прямой вопрос о важности скорого получения оценки решения студенты разделились во мнении примерно поровну, но ответы на косвенные вопросы на тему психологического самочувствия в ожидании результатов тестирования показали переживания студентов и падение уровня их внутренней готовности к учебной деятельности.

Отдельным фактором популярности САТ среди учащихся можно выделить комбинацию балльно-рейтинговой системы и получения результатов тестирования онлайн. Эти факты порождают стимулы соперничества и игры. Дело в том, что количество успешно пройденных тестов может не только формировать результат-статус: зачтена работа или не зачтена, но и выдавать балл – проценты удачных тестов относительно общего числа тестов. Таким образом формируется рейтинг решения и учащегося, являющегося автором этого решения. САТ, при должной настройке, предоставляет возможность отслеживать успешность выполнения заданий не только самого исполнителя, но и всех учащихся группы, тем самым создавая «таблицу рекордов» - вариант нематериальной мотивации, публичного признания [1; 9]. При этом в отличие от классических форм контроля учащийся имеет возможность самостоятельно исправить свою работу для продвижения в «таблице рекордов» вверх или удержания уже занятых позиций. При этом дополнительным стимулом может являться ранжирование решений по времени их предоставления (ранжирование в рамках диапазона времени, а не по относительной позиции «кто первый») – так, учащемуся выгоднее скорее приступить к выполнению заданий, а оставшееся время потратить на устранение собственных ошибок, изучение нового материала или повторение уже пройденного.

Еще одним психологическим аспектом использования систем автоматизированного тестирования как такового является соответствие времени. САТ является современным способом контроля качества знаний. Причем этот факт субъективно подтверждают опрашиваемые студенты, так, на вопрос «Тестирование - это современный метод контроля?» более 90% ответили утвердительно. Само программирование является новым направлением исследовательской и научной мысли. Достижения в информационных дисциплинах применяются в передовых отраслях и проектах. Поэтому у населения, в том числе у студентов и абитуриентов, сложилось четкие ассоциативные связи, соединяющие информационные технологии и прогрессивное движение. В итоге на субъективную оценку «прогрессивности» вуза и учебного процесса в нем влияют качество и количество информационных технологий, используемых вузом. И с этой точки зрения использование САТ при обучении программированию воспринимается как нечто само собой разумеющееся и не вызывает отторжения в студенческой среде, даже среди первокурсников, не сталкивающихся в прошлой учебной деятельности с подобными системами. В подтверждение этого факта: более 80% опрашиваемых положительно ответили на вопрос «Использование автоматизированного тестирования при изучении программирования, на Ваш взгляд, целесообразно?». Таким образом, САТ не вызовет негативных психологических последствий самим фактом их использования как системы контроля знаний.

Анкетирование подтвердило для САТ общую характерную черту для всех форм и методов контроля знаний учащихся, а именно сопутствующий учебный стресс [10] и, как следствие, негативные ассоциации впоследствии. В итоге большинство студентов высказывают нежелание самостоятельно выполнять какие-либо задания вне учебного курса или прямой подготовки к профессиональной деятельности. Данный результат был ожидаем и предсказуем. САТ не является панацеей, и к применению данной формы контроля необходимо подходить, учитывая и негативные влияния некоторых аспектов проверок на психологический фон учащихся. Необходимо тщательно подбирать содержимое задач САТ в соответствии с интенсивностью их использования.

О градации заданий в САТ

Давно применяемой практикой стало использование градации заданий по сложности. САТ также позволяет использовать градацию задания, но двумя способами. Первый – классический – представляет собой несколько заданий в одном разделе или теме с возрастающей сложностью. По сути в таком виде это классический вариант заданий «со звездочкой» - повышенной сложности, и набор обязательных к выполнению. Однако САТ позволяет осуществить градацию сложности условий задания внутри одной-единственной задачи без необходимости выделения усложнённых условий в отдельное задание. В

зависимости от конкретной реализации САТ позволяет явно или неявно группировать тестовые данные по классам (пулам) в соответствии с определёнными условиями задачи или типами возможных ошибок. Таким образом, одна задача может иметь несколько различных решений с разной степенью сложностью (классическим примером является задача суммирования двух чисел: натуральных, целых и вещественных, рациональных и иррациональных, коротких и длинных), при этом все эти решения тестируются вместе, и по результатам тестирования система выносит вердикт об уровне решения.

Каков эффект от САТ?

Использование САТ как некоего подобия игры можно применять не только как стимул к выполнению заданий, но и как метод адаптации студентов к новой учебной среде. Обратимся к словам С.К. Карауылбаева [4], согласно ему, процесс игры существенно способствует приспособлению студентов первых курсов к новым требованиям и условиям. Таким образом, САТ способствуют адаптировать студентов к новым условиям выполнения заданий по программированию. Если в школах приоритетом было освоение базовых теоретических знаний, то в высших учебных заведениях приоритетом является подготовка профессиональных навыков и умений – где одним из основных критериев является критерий функциональности приложения и соблюдения ограничения аппаратных ресурсов, что можно полностью контролировать в САТ.

Несомненным фактом является польза письменного метода контроля перед устным или тестовым. САТ использует для проверки решения тестовую форму, однако сами решения, которые необходимо предоставить для проверки, являются творческой интерпретацией стандартных алгоритмов или действительно творческой работой. Таким образом, с точки зрения учащегося, САТ требует творческого подхода к предоставляемому решению, а с точки зрения преподавателя, САТ предоставляет формализованную оценку, не требующую от преподавателя активного участия в процессе проверки задания в большинстве случаев. В итоге использование САТ подстегивает творческое начало в работе студентов, при этом требует от преподавателей существенно ниже прикладываемых усилий при непосредственной проверке этих творческих работ в сравнении с прочими методами внедрения заданий творческих работ.

Несомненным положительным аспектом САТ является встроенная система отчетности, позволяющая в сжатые сроки собрать и оформить информацию по деятельности отдельного учащегося и группы. Ввиду внедряемых систем электронного мониторинга успеваемости учащихся с возможностью удаленного доступа, подобный функционал является подспорьем в деятельности преподавателей и методистов учебных учреждений, помогая оперативно фиксировать и предоставлять достоверную информацию по освоению

учебного материала.

Заключение

САТ является современным методом контроля знаний и практических навыков по программированию. Его внедрение и использование не вызывает резкого неприятия среди учащихся. Позволяет оптимизировать педагогическую деятельность и упростить процесс контроля над творческими заданиями.

Статья подготовлена в рамках выполнения государственных работ в сфере научной деятельности по базовой части государственного задания № 2014/389, НИР №2193.

Список литературы

1. Вилюнас В.К. Психологические механизмы мотивации человека. — М. : Изд-во МГУ, 1990. — 288 с.
2. ГОСТ Р ИСО/МЭК 9126-93. Информационная технология. Оценка программной продукции. Характеристики качества и руководства по их применению. – Введ. 1993-12-28. – М. : Изд-во стандартов, 2004. - 9 с.
3. Заблоцкая О.А. Образовательное пространство как педагогическая категория // Информатика, вычислительная техника и инженерное образование. – 2011. - № 2 (4). – С. 9-14.
4. Карауылбаев С.К. Особенности применения учебно-деловых игр в обучении студентов // Мир науки, культуры, образования. – 2012. - № 3 (34). – С. 141-142.
5. Красильникова В.А. Теория и технологии компьютерного обучения и тестирования : монография. – М. : Дом педагогики, ИПК ГОУ ОГУ, 2009. – 339 с.
6. Кречетников К.Г. Применение компьютерного тестирования для контроля знаний. Мотивационный эффект [Электронный ресурс] / К.Г. Кречетников, Н.Н. Черненко // Эйдос : интернет-журнал. - 1999. - 2 марта. – Режим доступа: <http://www.eidos.ru/journal/1999/0302-01.htm> (дата обращения: 01.05.2016).
7. Крокер Л. Введение в классическую и современную теорию тестов : учебник / Л. Крокер, Дж. Алгина; пер. с англ. Н.Н. Найденовой, В.Н. Симкина, М.Б. Чельшковой; под общ. ред. В.И. Звонникова, М.Б. Чельшковой. – М. : Логос, 2010. – 668 с.
8. Майерс Г. Искусство тестирования программ / под ред. Б.А. Позина. - М. : Финансы и статистика, 1982. – 272 с.
9. Макклеланд Д. Мотивация человека. - СПб. : Питер, 2007. – 672 с. : ил. – (Серия «Мастера психологии»).
10. Щелина Т.Т., Колясникова В.А. К проблеме развития стрессоустойчивости студентов

психолого-педагогических направлений подготовки // Молодой ученый. — 2015. — № 11. — С. 1774-1778.