

УДК 004.054

РАЗРАБОТКА ПРОГРАММЫ ПО ОЦЕНКЕ КАЧЕСТВА АКАДЕМИЧЕСКИХ ПРОГРАММНЫХ ПРОДУКТОВ НА ОСНОВЕ МЕТОДИКИ ХОЛСТЕДА

(Работа выполнена при частичной поддержке гранта РФФИ, проект № 14-07-00665)

*А.И. Афанасова, аспирант, младший научный сотрудник, nastyaafanasova@yandex.ru
(Ульяновский государственный университет, ул. Л. Толстого, 42, г. Ульяновск, 432000, Россия)*

Аннотация. В современном мире оценивание качества программного обеспечения считается одной из важнейших задач как в софтверных компаниях, так и в высших учебных заведениях с компьютерно-ориентированным уклоном подготовки. В таких вузах вопрос о том, как оценивать качество создаваемых студентами программ – так называемых академических программных продуктов, остается не вполне решенным, хотя его решение опосредованно затрагивает интересы и софтверных компаний, куда придут работать выпускники. Ввиду существования ряда отличительных особенностей академических программных продуктов от профессиональных оценивание их качества требует специального подхода и особой вычислительной методики.

В статье излагаются принципиальные особенности методики оценивания качества академических программных продуктов, разработанной и применяемой в Ульяновском государственном университете для оценивания качества программ, написанных студентами на C++. Описана оригинальная, созданная на языке C# «Программа по оценке качества академических программных продуктов на основе методики Холстеда» для автоматизации процесса оценивания.

Ключевые слова: оценивание качества программных продуктов, академические программные продукты, программометрия, метрики Холстеда, проектно-ориентированное обучение, C++.

Оценивание качества ПО считается в настоящее время одной из важнейших задач, стоящих и перед софтверными компаниями, и перед высшими учебными заведениями с компьютерно-ориентированным уклоном подготовки. Вопрос о том, как оценивать качество создаваемых студентами программ – так называемых *академических программных продуктов* (АПП), в таких вузах не вполне решен, хотя его решение опосредованно затрагивает интересы и софтверных компаний, куда придут работать выпускники. Так как между АПП и *профессиональными программными продуктами* (ППП) существует ряд отличий, оценивание их качества требует специального подхода и особой вычислительной методики.

Задаче оценивания качества программных продуктов посвящено множество исследований [1–5], однако их объектом, как правило, являются ППП. Задаче оценивания качества АПП уделяется значительно меньшее внимание, хотя эта задача неизбежно возникает на технических специальностях, где студенты в ходе учебного процесса создают большое количество программ [5–8]. В силу особенностей АПП для разработки таких больших массивов проектной документации требуются специальные средства.

В отличие от ППП в АПП отсутствует обратная связь по их качеству, поскольку АПП в большинстве случаев не имеют непосредственного контакта с конкурентной рыночной средой. Оценивание качества студенческих программ обычно выполняется экспертами – преподавателями, которые, формируя оценку, опираются на свой личный опыт, знания и представления. Однако даже при условии высокой квалификации преподавателей их требования индивидуальны, субъективны и существуют на интуитивном уровне. Учитывая то, что во время обучения создается большое количество АПП, их оценивание – процесс достаточно сложный, затратный по времени и трудоемкий. Сведений о средствах автоматизации именно этого процесса в литературе явно недостаточно.

Данная работа предпринята с целью восполнения отмеченного недостатка. Ее выполнение в Ульяновском государственном университете обусловлено тем, что здесь при чтении дисциплин, связанных с вычислительной математикой и программированием, с 1989 года применяется проектно-ориентированное обучение [8, 9], которое предусматривает разработку студентами целостных, нетривиальных программных продуктов. Оценив уровень качества по множеству таких АПП, преподаватель и кафедра получают необходимую обратную связь и, следовательно, возможность формирования оценки качества всего образовательного процесса [10].

Процедуры оценивания качества программных продуктов, классификация существующих методов

Одной из важнейших проблем оценивания качества программных продуктов является формализация понятия оценки качества, характеристик качества, а также методология их оценивания.

Оцениванию качества программных средств посвящены государственные и международные стандарты.

В ГОСТ 28195-89 дается определение процедуры оценки качества программного средства как совокупности операций, включающей выбор номенклатуры показателей качества, определение значений этих показателей и сравнение их с базовыми значениями.

Существующие подходы к оцениванию качества программных продуктов можно классифицировать следующим образом [11].

– По внутренним показателям качества. К ним относят подходы, основанные на оценке ряда технологических показателей программы (таких, как сложность программы, вычислительная избыточность, структурная целесообразность и т.д.). Внутренние показатели качества программного продукта оказывают влияние на его внешние показатели.

– По внешним показателям качества. К оценке качества по внешним показателям относят подходы, основанные на оценке различного рода показателей, непосредственно наблюдаемых при испытаниях программных средств.

– Смешанный подход. К смешанным относят подходы, основанные на оценивании значений как внешних, так и внутренних показателей качества.

При разработке методологии оценивания качества АПП важно понимать, что на этапе обучения студентам в первую очередь необходимо научиться писать код, обладающий качественными внутренними характеристиками. А учитывая тот факт, что у АПП, как правило, отсутствует обратная связь от пользователя, целесообразным является оценивание АПП именно по их внутренним показателям качества.

Методика оценивания качества АПП

В основу предлагаемой методики положим работу [12], применение которой для оценивания качества АПП впервые предложено в [10].

В соответствии с методикой Холстеда оценку качества программного продукта формируют в результате определения специальных характеристик (таких, как число операторов и операндов) исходного кода программы. При этом априори считается, что текст программы свободен от ошибок, связанных с языком программирования. Для АПП это справедливо, поскольку для оценивания отбираются тексты программ, успешно прошедшие компиляцию. Также считают, что текст АПП свободен от логических (алгоритмических) ошибок и отклонений от выданного задания, поскольку он уже прошел необходимую проверку у эксперта-преподавателя.

Процесс получения качественных характеристик АПП. Охарактеризуем процесс получения параметров качества АПП.

По методике Холстеда [12], принципиальным параметром качества АПП является критерий Hq совершенства логической структуры текста программного продукта. Критерий Hq вычисляют с использованием четырех характеристик длины программного текста:

η_1 – число простых (или отдельных) операторов, появляющихся в данной реализации (словарь операторов);

η_2 – число простых (или отдельных) операндов, появляющихся в данной реализации (словарь операндов);

N_1 – общее число всех операторов, появляющихся в данной реализации;

N_2 – общее число всех операндов, появляющихся в данной реализации.

На основе этих параметров для каждого текста АПП вычисляют две характеристики длины: экспериментально определенную

$$N = N_1 + N_2 \quad (1)$$

и теоретическую

$$\hat{N} = \eta_1 * \log_2 \eta_1 + \eta_2 * \log_2 \eta_2. \quad (2)$$

Обоснование корректности формулы вычисления теоретической длины. По данным работы [10], формула (2) имеет строгое математическое доказательство в [13]. Она выражает теоретическую длину \hat{N} программы через число входящих в ее код простых операторов η_1 и число простых операндов η_2 , исходя из минимаксного для данного типа программ так называемого «стоимостного» критерия. Это доказательство основано на предположении, что при создании программы реализуется функциональная задача в максимально экономной и рациональной форме. В [11] введено соотношение

$$N = \lambda(\eta_1 * \log_2 \eta_1 + \eta_2 * \log_2 \eta_2) = \lambda \hat{N}, \quad (3)$$

где λ – параметр, характеризующий степень приближения N к \hat{N} . Там же выполнено экспериментальное обоснование, в ходе которого (после обработки десяти профессионально составленных программ на языках высокого уровня) получено значение $\lambda = 0,934336314$, достаточно близкое к 1.

В литературе по методам программометрии используются и другие методы измерения характеристик программ. Тем не менее, программы, для которых значение λ приближается к 1, общепринято классифицировать как близкие по качеству к профессиональным программам, что и обосновывает выбор метрики Холстеда в данной работе.

Статистическая обработка программных текстов. После того как все тексты АПП обработаны и для каждого из них уже вычислены соответствующие характеристики – экспериментальная и теоретическая длины, осуществляем обработку полученных параметров методами математической статистики.

Для выборки теоретических значений длин программ и выборки экспериментально вычисленных значений длин программ, где объем каждой выборки (количество обрабатываемых файлов) равен n , находим расчетные значения оценок статистических характеристик:

1. Математическое ожидание:

$$\tilde{M}_x = \frac{1}{n} \sum_{i=1}^n x_i \quad (4)$$

2. Дисперсия:

$$\tilde{D}_x = \frac{\sum_{i=1}^n (x_i - \tilde{M}_x)^2}{n-1} \quad (5)$$

3. Среднеквадратичное отклонение:

$$\tilde{\sigma}_x = \sqrt{\tilde{D}_x} \quad (6)$$

Далее, когда оценка статистических характеристик выполнена, для теоретически полученной и экспериментально определенной длин АПП вычисляем коэффициент корреляции:

$$\tilde{r}_{xy} = \frac{\sum_{i=1}^n (x_i - \tilde{M}_x) \cdot (y_i - \tilde{M}_y)}{(n-1) \cdot \tilde{\sigma}_x \cdot \tilde{\sigma}_y} \quad (7)$$

Критерий совершенства Hq численно равен коэффициенту корреляции \tilde{r}_{xy} .

Для совершенных программ параметр Hq близок к 1, для программ, имеющих тот или иной набор несовершенств, отличия теоретического и экспериментально определенного значений длины текста будут более значительны, поэтому в результате коэффициент корреляции \tilde{r}_{xy} уменьшится, то есть уменьшится параметр Hq .

Поскольку на начальном этапе создания программ особое внимание уделяется правильному, оптимальному построению алгоритма и реализации его на языке программирования, параметр Hq является важным для оценки качества академических программных продуктов.

Программа оценивания качественных характеристик АПП по методике Холстеда. Охарактеризуем программную реализацию вычислений, описанных выше.

Разработанная программа [14] представляет собой специализированную программу для автоматического определения качественных характеристик академических программных продуктов на языке C++ на основе изложенной методики.

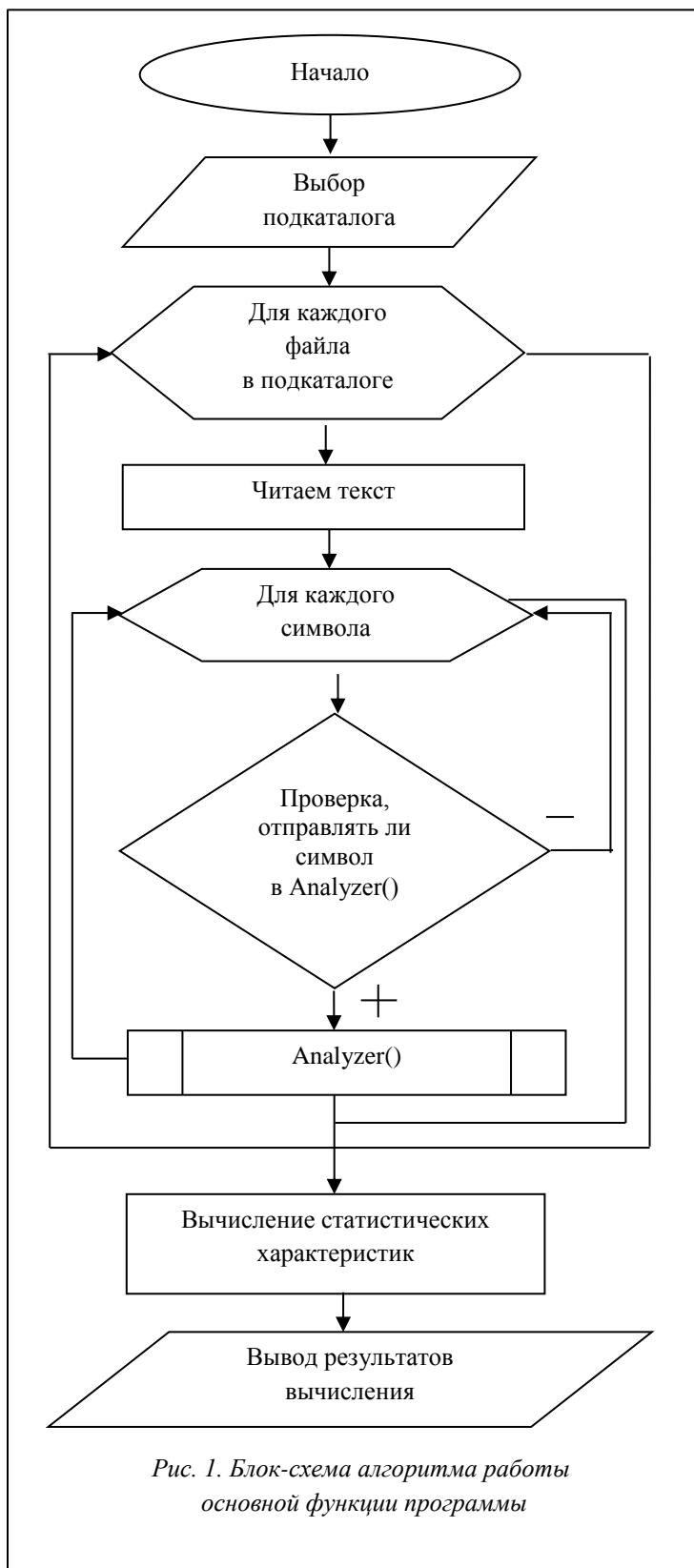
Программа реализована в среде Visual Studio на языке программирования C# и состоит из основной функции, функции лексического анализатора Analyzer() и функции синтаксического анализатора Result().

Процесс обработки программного текста организован следующим образом.

1. Основная функция программы считывает посимвольно текст каждого АПП из указанного пользователем подкаталога и отправляет каждый символ в лексический анализатор. Блок-схема алгоритма работы основной функции программы представлена на рисунке 1.

2. Лексический анализатор на основе поступающих на вход символов последовательно формирует лексему. Когда лексема сформирована, она передается в синтаксический анализатор. Схема конечного автомата лексического анализатора представлена на рисунке 2.

3. Синтаксический анализатор сопоставляет входную лексему с грамматикой языка, которая описана в программном коде в виде статических строковых массивов, и определяет, к какому типу относится входная лексема. В функции учитывается возможность ситуации, когда две различные переменные, находящиеся в разных областях видимости – локальная и глобальная, имеют одинаковое имя. Такие переменные фиксируются анализатором как два разных операнда. После того как определено, к какому классу относится поступившая на вход лексема, изменяется общее количество операторов или операндов программы, а также при необходимости пополняются словари операторов и операндов программы.



4. Далее, когда текст программы полностью обработан, в основной функции программы вычисляется его экспериментальная и теоретическая длина.

После окончания обработки всех программных текстов АПП для всей группы текстов выполняется вычисление оценок статистических характеристик (формулы (4)–(6)).

Далее на основе полученных данных по формуле (7) вычисляется критерий совершенства Hq . Результаты работы программы выводятся на экран.

Опишем взаимодействие программы и пользователя.

Для запуска процесса получения качественных характеристик пользователю необходимо выбрать подкаталог файлов – текстов АПП для их дальнейшей обработки. После того как пользователем указан необходимый подкаталог файлов, для каждого программного текста происходит формирование и вычисление следующих параметров качества: словарь операторов η_1 , словарь операндов η_2 , общее число всех операторов N_1 , общее число всех операндов N_2 , экспериментально полученная длина N и теоретическая вычисленная длина \hat{N} параметры выводятся на экран в виде таблицы. Под таблицей выводится вычисленное значение параметра совершенства для всей группы файлов. Наглядно это показано на рисунке 3, где представлены результаты обработки АПП студентов специальности «Информационные системы» факультета математики и информационных технологий УлГУ, выполненных в 2013-2014 учебном году в рамках дисциплины «Технология программирования» по теме «Вычисления по формулам».

В таблице представлены экспериментальные данные, полученные с помощью разработанной программы при оценивании лабораторных работ по дисциплине «Технология программирования» для студентов специальности «Информационные системы» факультета математики и информационных технологий УлГУ, выполненных также в 2013-2014 году. По количеству выполненных работ и параметру совершенства можно судить об уровне освоения

изученной темы. Например, в рамках тем 3–5 сдано мало работ, однако параметр качества для них высок. Из этого следует, что тему освоило небольшое количество студентов, но, по-видимому, это отличники. Количество работ по теме 6 тоже невелико, а параметр совершенства значительно ниже, чем по всем предыдущим темам. Здесь можно сделать вывод о сложности данной темы и необходимости усовершенствовать преподнесение материала в рамках семинарских занятий.

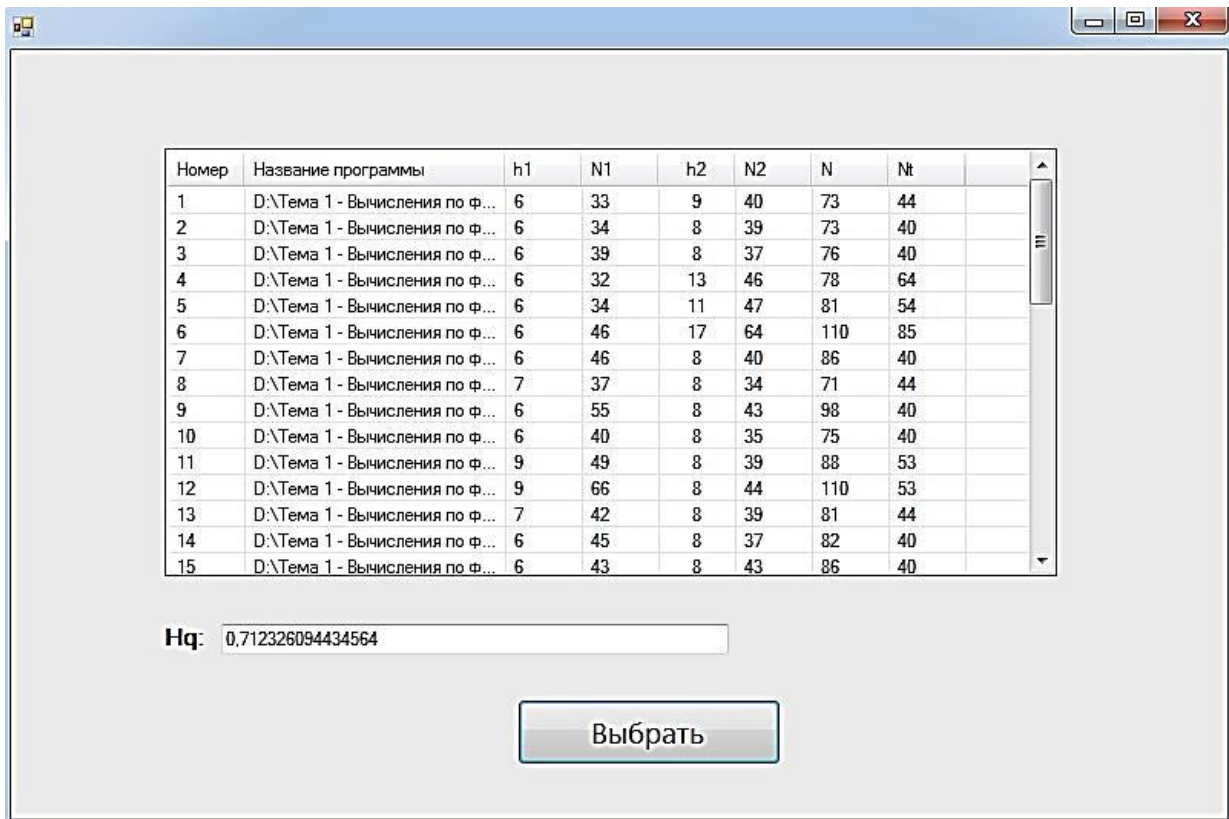
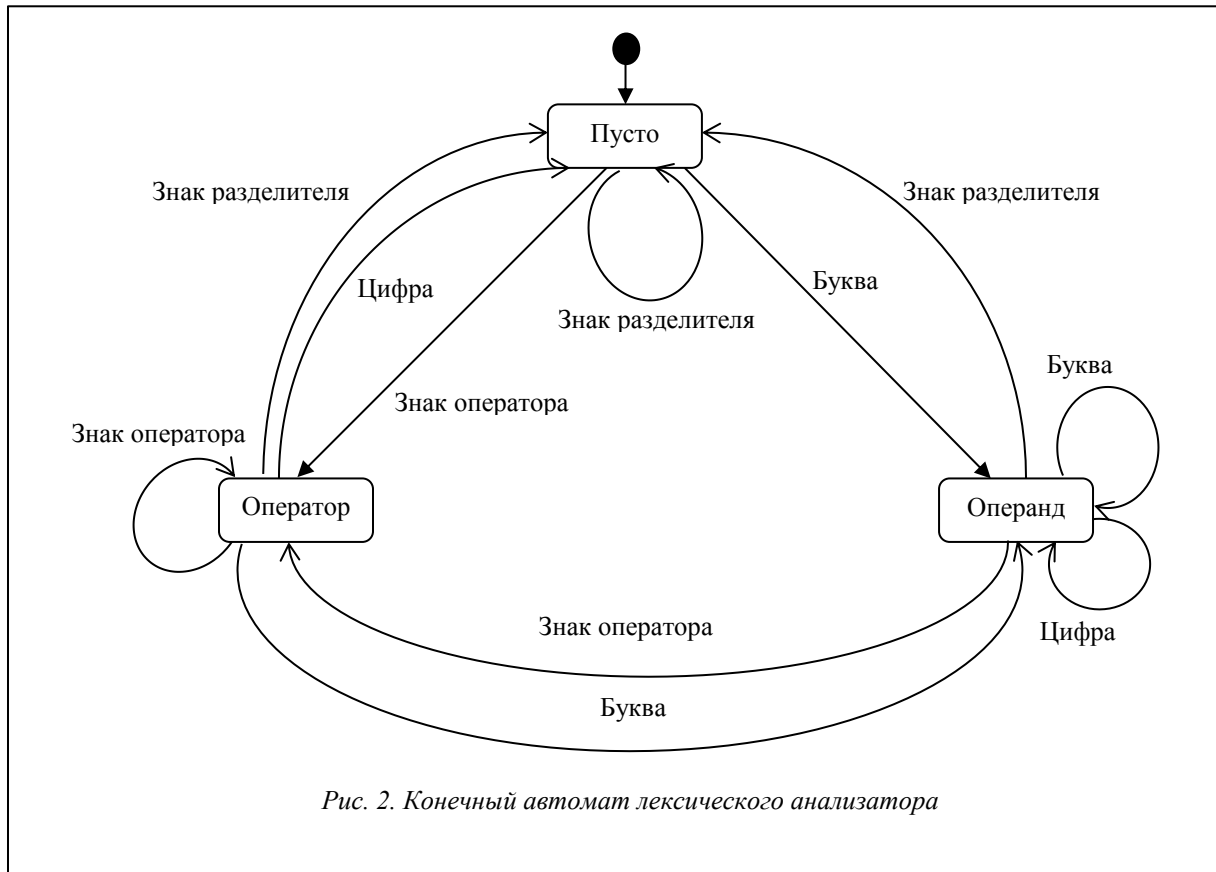


Рис. 3. Пример результата работы программы

Экспериментальные результаты определения параметра совершенства

№	Тема	Количество работ	Параметр совершенства H_q
1	Вычисления по формулам	40	0,712326094434564
2	Операторы цикла	23	0,567967833381929
3	Одномерные массивы	8	0,851011658832649
4	Последовательности	3	0,96756114421808
5	Обработка текстовых строк	3	0,825354659308643
6	Матрицы	4	0,681555060203852

Таким образом, объективные и проверяемые индикаторы качества АПП являются мощным инструментом модификации поведения студентов и преподавателей. Реализованные в специальных программах тестирования АПП обладают рядом достоинств по сравнению с традиционными практиками: исключают риск манипулирования, обеспечивают многоаспектное понимание результатов деятельности кафедры, отдельного преподавателя и группы (потока) студентов, изучающих дисциплину, допускают прямое сравнение различных методик преподавания и являются публично доступными и легко интегрируемыми в учебный процесс. Применение таких индикаторов находит положительные отзывы как со стороны преподавателей (снижает трудоемкость аттестации студентов, автоматизируя этот процесс), так и студентов (стимулирует стремление к достижению лучших показателей).

Созданная «Программа оценки качественных характеристик академических программных продуктов на основе методики Холстеда» зарегистрирована в Роспатенте [14] и прошла успешную опытную эксплуатацию на материале лабораторных проектов дисциплины «Технология программирования», в рамках которой было обработано 86 АПП, созданных в 2013–2014 гг.

Планируется продолжение опытной эксплуатации разработанной программы на расширенном наборе дисциплин, находящихся в учебном процессе факультета математики и информационных технологий УлГУ, а именно: «Технология программирования», «Языки программирования и методы трансляции», «Методы программирования», «Объектно-ориентированное программирование», «Численные методы», «Методы вычислений», «Вычислительная математика» и других, где учебным планом предусмотрены лабораторные работы в форме академических программных продуктов. Для этого создается банк АПП. В настоящее время он содержит 151 запись.

Литература

1. Interface.ru. Interface and software company. URL: <http://www.interface.ru/home.asp?artId=3987> (дата обращения: 18.10.2015).
2. Антошина И.В., Домрачев В.Г., Ретинская И.В. Основные тенденции оценивания качества программных средств // Качество, инновации, образование. 2004. № 1. С. 70–75.
3. Калинина Л.Ю. Оценка качества программных продуктов // Качество, инновации, образование. 2006. № 4. С. 52–55.
4. Антошина И.В., Домрачев В.Г., Ретинская И.В. Методика составления системы характеристик качества для программных средств // Качество, инновации, образование. 2002. № 3. С. 57–60.
5. Липаев В.В. О проблемах оценивания качества программных средств // Качество, инновации, образование. 2002. № 1. С. 93–97.
6. Информационно-коммуникационные технологии в образовании. URL: http://www.ict.edu.ru/vconf/index.php?a=vconf&c=getForm&r=thesisDesc&d=light&id_sec=123&id_thesis=4643 (дата обращения: 12.11.2015).
7. Информационные технологии в образовании: конгресс конференций. URL: <http://ito.edu.ru/2003/VI/VI-0-1548.html> (дата обращения: 12.11.2014).
8. Семушин И.В., Угаров В.В. Модификация поведения студента и преподавателя инженерных дисциплин // Московское научное обозрение. 2013. № 9 (37). С. 3–8.
9. Семушин И.В., Угаров В.В., Цыганова Ю.В., Афанасова А.И., Куличенко И.Н. Опыт проектно-ориентированного обучения в университетах Ульяновска // Перспективные информационные технологии (ПИТ 2014): тр. Междунар. науч.-технич. конф.; [под ред. С.А. Прохорова]. Самара: Изд-во Самарского науч. центра РАН, 2014. С. 436–438.
10. Угаров В.В. Компьютерные модели и программные комплексы в проектно-ориентированном обучении: дис. ... канд. технич. наук. Ульяновск, 2005.
11. Антошина И.В., Домрачев В.Г., Ретинская И.В. Основные тенденции оценивания качества программных средств // Качество, инновации, образование. 2004. № 1. С. 70–75.
12. Холстед М.Х. Начала науки о программах; [пер. с англ. В.М. Юфы]. М.: Финансы и статистика, 1981. 128 с.

13. Апостолова Н.А., Гольдштейн Б.С., Зайдман Р.А. О программометрическом подходе к оценкам программного обеспечения // Программирование. 1995. № 4. С. 38–44.

14. Афанасова А.И. Программа оценки качественных характеристик академических программных продуктов на основе методики Холстеда. Пат. 2015611297 РФ № 2014662403; заявл. 03.12.2014; опубл. 27.01.2015.