

ЭКСПЕРИМЕНТАЛЬНОЕ ИЗУЧЕНИЕ ЦЕЛОСТНОСТИ ЗНАНИЙ СТУДЕНТОВ

Евгений А. Еремин

Abstract: В работе описывается экспериментальное исследование целостности знаний, которые были получены студентами после усвоения учебного курса. Под целостностью понимается наличие связей между базовыми понятиями и терминами изученного курса, которые студент осознает и фиксирует. Предполагается, что чем больше таких связей видит студент, тем глубже он усвоил содержание курса. Мотивацией исследования данной проблемы явился поиск количественных критериев для оценки знаний студентов, в том числе и в процессе компьютерного обучения.

Метод экспериментальной проверки целостности системы понятий был разработан и опробован автором в 2008 году (его описание было опубликовано в трудах конференции MeL-09). Компьютер помогает в реализации этого метода как при проведении проверки знаний, так и при обработке полученных результатов после ее завершения. В данной работе описаны более поздние результаты, которые были получены по предложенной методике в 2009-2010 годах. В ходе исследования проанализированы ответы более 50 студентов двух разных специальностей, которые изучали курс «Архитектура ЭВМ». Проверка знаний производилась дважды: до и после изучения курса; полученные данные сравнивались между собой.

По результатам экспериментов построены серии диаграмм для каждой из групп студентов, принимавших участие в исследовании; различные показатели сопоставлены друг с другом. Анализ результатов показал, что существуют студенты, которые лучше, чем остальные, видят связи между понятиями, причем, как выяснилось, это совсем не обязательно наиболее успешные студенты. На итоговом обобщающем графике удалось выделить четыре области, характеризующиеся разной степенью успешности освоения курса, в каждой из которых полученные в эксперименте параметры проявляют определенные особенности. В частности, довольно неожиданным явился установленный факт надежной фиксации роста целостности знаний у самых слабых студентов.

Keywords: целостность, понятие, структура знаний, оценка, образование, курс..

ACM Classification Keywords: K.3.1 Computer Uses in Education; I.2.6 Learning – Knowledge acquisition.

Введение

При проверке знаний студентов очень важно иметь четкую систему критериев, по которым преподаватель выставляет оценку. В результате многовекового опыта обучения сложились некоторые способы оценки усвоения знаний, например, в ходе беседы по изученному материалу. Тем не менее, попытки перенести эти способы в компьютерное обучение сталкиваются с существенными трудностями, поскольку экспертный опыт преподавателей довольно плохо поддается формализации.

Наиболее распространенным в данный момент методом компьютерной проверки знаний является тестирование. Типовой тест позволяет довольно быстро оценить знания тех или иных фактов, отдельных терминов или определений из изучаемого курса, но все же уступает по глубине проверки человеку-преподавателю. Конечно, тщательно разработанный тест из нескольких сотен вопросов может претендовать на объективность оценки знаний курса в целом, но проблема в том, что прохождение такого теста слишком трудоемко (опытному учителю для того, чтобы поставить оценку, требуется значительно меньше вопросов).

Можно предположить, что одним из главных преимуществ классического опроса преподавателем по сравнению с компьютерным тестом служит то обстоятельство, что человек способен по небольшому набору ответов на вопросы увидеть степень систематичности знаний экзаменуемого в целом и тем самым адекватно оценить общую картину усвоения всего курса. Справедливость предположения о важности взаимосвязей между отдельными известными студенту фактами при оценке знаний автору и хотелось проверить.

В работе экспериментально изучается целостность системы базовых понятий у студентов как один из аспектов успешности усвоения материала образовательного курса. Под **целостностью системы знаний студента** здесь и далее мы будем понимать взаимосвязь между теми понятиями и терминами, которые он усвоил в ходе изучения курса и способен продемонстрировать в ходе проверки. Подчеркнем, что нас будут в первую очередь интересовать связи между понятиями, которые осознает студент, а не логика преподавания взаимосвязанного материала курса (последняя проблема заслуживает самостоятельного изучения). В идеале все базовые термины курса должны образовывать в голове студента некоторую единую картину. Напротив, при поверхностном и некачественном изучении курса, знания студента представляют собой набор несвязанных между собой фактов.

Постановка проблемы

Рассмотрим несколько упрощенный пример, поясняющий идею исследования целостности знаний по курсу. Пусть преподаватель опрашивает трех студентов А, В и С по изученному материалу. Допустим, например, что студент А в ходе ответов на вопросы продемонстрировал знание трех следующих фактов:

- компьютер обрабатывает данные;
- для кодирования чисел используется двоичная система;
- программа хранится в памяти компьютера.

Отчетливо видно, что названные студентом факты непосредственно не связаны между собой. Предположим теперь, что ответы каждого из студентов В и С отличаются от ответа студента А всего одним фактом. Например, студент В мог отметить следующие положения:

- компьютер обрабатывает данные;
- числа – это разновидность данных;
- для кодирования чисел используется двоичная система.

А ответы студента С могли быть такими:

- компьютер обрабатывает данные;
- данные обрабатываются по программе;
- программа хранится в памяти компьютера.

Не надо обладать большим педагогическим опытом, чтобы сказать, что ответы студентов В и С лучше (хотя все три наших гипотетических студента назвали одинаковое число фактов!), поскольку они образуют некоторый *взаимосвязанный* набор фактов. Причем интересно отметить, что студенты В и С продемонстрировали свои знания в разных областях курса: студент В говорил о кодировании компьютерных данных, а студент С – об их обработке.

Таким образом, одним из показателей качественного усвоения знаний является тот факт, что студент видит и демонстрирует при контроле взаимосвязь между базовыми терминами и понятиями курса.

Сформулируем теперь наиболее важные цели нашего исследования.

- Экспериментально изучить степень связи между понятиями (целостность знаний) у студентов по изученному курсу.
- Выработать некоторые характеристики целостности знаний и опробовать их.

- Проверить, увеличиваются ли выбранные характеристики после изучения курса.
- Проанализировать, насколько целостность знаний связана с успешностью усвоения курса.

В идеале в процессе экспериментального изучения целостности знаний хотелось найти такую численную характеристику X , чтобы для студентов из рассмотренного выше примера было $X_A < X_B$ и $X_A < X_C$. В дальнейшем характеристики, подобные X , могли бы быть использованы при автоматической оценке знаний студентов.

Участники эксперимента

В ходе исследований в течение нескольких лет проверялись знания студентов физического факультета Пермского государственного педагогического университета (Россия). Для получения более достоверных результатов эксперименты проводились на двух разных специальностях, по которым на факультете ведется подготовка: (1) учитель физики и информатики и (2) специалист по информационным технологиям (последняя предполагает, что выпускники будут обслуживать компьютерные классы в системе образования). Все эти студенты изучают курс «Архитектура ЭВМ», по которому и проверялась целостность системы базовых понятий. По первой специальности курс стоит на третьем году обучения, а по второй – на втором. Пробные эксперименты были проведены в 2008 году, а затем в 2009 и 2010 по отлаженной методике единообразно проверялись 4 группы студентов.

Данные о группах, где проводились исследования, и их численности приведены в табл. 1.

Таблица 1

группа	год	год обучения	специальность	количество студентов
G0	2008	3	физика и информатика	12
G1	2009	3	физика и информатика	12
G2	2009	2	информационные технологии	8
G3	2010	3	физика и информатика	17
G4	2010	2	информационные технологии	9
итого				58

Для удобства последующего изложения каждой из групп студентов присвоено формальное обозначение от G0 до G4. Строка с группой G0 выделена в таблице цветом, чтобы подчеркнуть, что результаты для нее содержали отдельные недоработки: из-за «не совсем чистого» проведения эксперимента они считаются предварительными.

В таблице также указано количество студентов, принимавших участие в исследовании структуры знаний. Оно не очень велико, поскольку набор по данным специальностям в последнее время сильно упал. Кроме того, были отдельные студенты, которые по уважительным причинам пропустили те дни, когда проводились экспериментальные проверки знаний, включая повторные. Результаты этих студентов оказались неполными, что, к сожалению, не позволило использовать их в исследованиях. Подчеркнем, что принимались все возможные меры, чтобы привлечь имеющихся студентов к участию в проверке знаний.

Все эксперименты по контролю знаний и вся обработка их результатов проводилась автором самостоятельно. С одной стороны, это гарантировало максимально возможную одинаковость условий проведения исследований. Но с другой – работа даже с не очень большим количеством студентов потребовала значительных усилий и времени.

База терминов для проверки знаний

Согласно замыслу эксперимента, в ходе проверки знаний студенту предъявляется полный список базовых понятий курса, а тот должен продемонстрировать, какие связи между этими понятиями он видит. Таким образом, при подготовке к проведению исследований экспериментатор должен отобрать те наиболее важные понятия и термины, которые, по его мнению, должен знать студент после усвоения курса.

Как уже говорилось выше, в качестве области исследования знаний студентов был выбран курс «Архитектура ЭВМ» (что во многом обусловлено личными предпочтениями и интересами автора). Структура базовых понятий для этого курса была подробно проанализирована ранее, и результаты этого анализа были изложены в публикации [Еремин, 2007]. Именно она и была взята для проведения экспериментов.

Опуская второстепенные детали формирования списка базовых понятий рассматриваемого курса, укажем лишь наиболее важные из них. Список состоит из 123 понятий, которые тесно связаны между собой и образуют единую иерархию терминов. В него были отобраны как наиболее общие понятия – *компьютер*, *программная* и *аппаратная часть*, *теоретические основы*, так и более конкретные термины, раскрывающие их, например, *операционная система*, *процессор*, *память*, *прямой доступ к памяти*, *принцип иерархии*, *байт* и многие другие. В список также вошли термины, реализующие межпредметные связи, например, с микроэлектроникой, логикой и системами счисления. С другой стороны, в перечень терминов сознательно не были включены названия конкретных операционных систем, внешних устройств и их производителей, а также другая подобная информация, которая является менее существенной с точки зрения изучения главных закономерностей курса. Используя стандартную терминологию, принятую в объектно-ориентированном программировании, можно сказать, что рассматривались классы понятий, но не их экземпляры.

Таким образом, был выделен достаточно широкий перечень базовых понятий, каждое из которых, с точки зрения преподавателя, должен знать и понимать грамотный студент. Перечень получился весьма объемным, так что впоследствии при проведении эксперимента оказалось, что реальные студенты активно использовали в своих ответах немногим более половины предложенного списка. Предоставление студентам такого избыточного списка имело целью сделать возможной проверку знаний у *любых* студентов, в том числе и тех, кто хорошо знаком с архитектурой компьютера или отдельными ее областями. Анализ хода эксперимента показал, что, возможно, слишком длинный перечень терминов «перегружал» студентов: они в ходе проверки уставали и демонстрировали не все свои знания. Так или иначе, но во всех описываемых экспериментах использовался именно полный список понятий из более чем ста терминов.

Для составленного списка были также предварительно проанализированы связи между отобранными терминами. Нетривиальный результат состоял в том, что удалось обойтись весьма ограниченным набором связей. В него вошли стандартные отношения между понятиями, например, *часть/целое* или *класс/подкласс*, а также некоторые специфические для курса связи вроде *основание (принцип иерархии или принцип адресации – основание – память, программный счетчик – основание – основной алгоритм исполнения инструкции) и соединение*. Полная таблица связей с конкретными примерами для каждой приведена в публикации [Еремин, 2007]: она состоит всего из 11 базовых ассоциаций.

Первоначально предполагалось, что студенты сумеют самостоятельно разобраться в таком небольшом количестве связей. Опыт тестирования, тем не менее, уже для предварительной группы G0 показал обратное: студенты часто путали даже классические *часть/целое* и *класс/подкласс*, не говоря уже об остальных видах связей. Поскольку с точки зрения цели эксперимента – общей оценки связности системы понятий, конкретные разновидности связей не так важны, было принято решение при оценке

правильности ответов пренебречь ошибками в этой части задания и просто фиксировать факт наличия связи. Данное упрощение методики заметно облегчало процесс обработки и анализа результатов.

Описанная выше система терминов и была положена в основу экспериментальной процедуры проверки целостности знаний студентов по курсу архитектуры компьютера.

Описание эксперимента

В ходе эксперимента студенты *дважды* выполняли задание: в начале изучения курса и после его завершения (входной и выходной контроль). Предполагалось, что сопоставление результатов для каждого из студентов позволит индивидуально оценить степень успешности изучения курса. К сожалению, не все студенты выполняли повторное задание так же активно, как при входном тестировании, что портило результаты сравнения.

Цель эксперимента (проверка целостности знаний) не сообщалась, чтобы не вызвать искусственного повышения результатов. Студентам лишь рекомендовалось постараться выполнить задание максимально полно и хорошо, не задумываясь при этом, как и по каким параметрам будет оцениваться их успешность. Впоследствии специальная экспериментальная проверка подтвердила важность данного принципа.

Время выполнения задания не ограничивалось: каждый студент заканчивал работу индивидуально, когда считал нужным (как правило, на выполнение задания каждый студент тратил от 30 минут до часа). По-видимому, такая практика внесла некоторую неопределенность в проведение эксперимента: хотя большинство студентов старались указать как можно больше связей между понятиями курса, были и такие, кто стремился побыстрее освободиться, тем более, что выходное тестирование было последним заданием перед получением зачета по курсу. Чтобы хоть как-то уменьшить указанный негативный эффект, в последних экспериментах (группы G3 и G4) студентов предупреждали о необходимости более добросовестного выполнения второго задания. Им говорилось о том, что за время изучения курса желательно «не поглупеть».

Самый первый эксперимент, проведенный в группе G0, показал, что отдельные студенты откровенно жульничали и сдавали вместо своих результатов файлы более сильных студентов. Это стало одной из существенных трудностей проведения последующих экспериментов, хотя студентов всегда предупреждали перед экспериментом о необходимости самостоятельного выполнения задания. В 2009 году в группах G1 и G2 все файлы сравнивались между собой (сначала по длине, а потом и по содержащемуся в них тексту). Уличенным в обмане студентам группы G1 пришлось повторно выполнять работу. В небольшой группе G2, которая к тому же оказалась более добросовестной, подобных случаев зарегистрировано не было. Чтобы уменьшить усилия по контролю за достоверностью результатов и усложнить жизнь нечестным студентам, в 2010 году (группы G3 и G4) была добавлена запись в файл фамилии студента, а сам текст с результатами подвергался шифрованию. В итоге, не зная, где именно в файле хранится фамилия, студенты не могли ее заменить.

Таким образом, принимались все возможные меры для того, чтобы результаты исследования целостности знаний стали максимально объективными.

При объяснении задания студентам также выдавался текстовый файл с разъяснениями о типах связей между понятиями и их примерами. Предполагалось, что этого будет достаточно, чтобы при ответе правильно указать типы связей. Но, как уже говорилось выше, эти ожидания не оправдались и при обработке результатов типы связей игнорировались.

Кроме того, студентам предлагалось воспользоваться вспомогательным файлом с полным списком понятий, где разрешалось делать пометки и записи – тем самым участники эксперимента неявно

подталкивались к систематичному выполнению задания. Судя по наблюдениям, этой возможностью никто не воспользовался.

Итак, после проведения подробного инструктажа и знакомства с программой для фиксации связей между понятиями (ее описание дано в [Еремин, 2009]), студенты приступали к выполнению задания, указывая правильные, по их мнению, связи между понятиями курса.

Результаты, сохраненные в виде файла, просматривались экспериментатором. Связи, которые были явно указаны ошибочно, из файла удалялись. Вся дальнейшая обработка файлов велась уже в автоматическом режиме.

Просмотр файлов с ответами показал, что большая часть студентов (за редким исключением) составляла связи между парами понятий весьма бессистемно. Например, в типичном студенческом файле выделялась связь между понятиями *система счисления* и *двоичная система*, но, тем не менее, отсутствовали связи этого понятия с *восьмеричной* и *десятичной системами*. На данный недостаток, как и на неумение студентов правильно определять типы связей, стоит обратить внимание всем преподавателям.

Обработка результатов эксперимента

Исходным материалом для компьютерной обработки результатов эксперимента были текстовые файлы, содержащие расшифрованные и проверенные преподавателем данные. Каждая строка такого файла содержала информацию об одной связи между понятиями и имела вид

термин 1 # тип связи # термин 2

Например, итоговый файл тестирования некоторого студента мог содержать следующий фрагмент:

процессор # целое/часть # АЛУ

процессор # целое/часть # УУ

АЛУ # целое/часть # регистр

УУ # целое/часть # регистр

Программа обработки была способна объединять связанные понятия в группы. В приведенном примере в общую группу попадают термины *процессор*, *АЛУ*, *УУ*, *регистр*. Разумеется, расположение пар терминов в файле и их порядок не играют роли.

Важно отметить, что перед анализом результаты упорядочивались по определенному рейтингу успешности освоения курса. Было принято допущение, что *успешность определяется как срок сдачи всех заданий*: хорошие студенты, как правило, сдают задание быстро, а слабые тратят на эту же самую работу много больше времени. Сопоставление данного предположения с субъективными наблюдениями за работой студентов подтверждает, что в подавляющем большинстве случаев оно оправдывается.

Обсуждение результатов

Перейдем к обсуждению результатов эксперимента. Начнем с **первичных** характеристик, которые могут быть легко получены непосредственно из студенческих файлов с ответами. Наиболее наглядным параметром с точки зрения изучения целостности знаний является общее количество связей между понятиями курса, которые сумели вспомнить студенты. Значения этой величины для групп G1-G4 приведены в левой половине рис. 1. Каждая пара столбцов на приведенных диаграммах соответствует двум проверкам знаний: левый столбец – это количество связей L_1 , указанное студентом до изучения курса, а правый – после (L_2). Студенты имеют условные номера, указанные вдоль оси абсцисс. Напомним, что согласно принятому принципу, те, кто выполнил все задания раньше (сильные студенты), имеют меньшие номера, и их столбцы расположены ближе к началу координат.

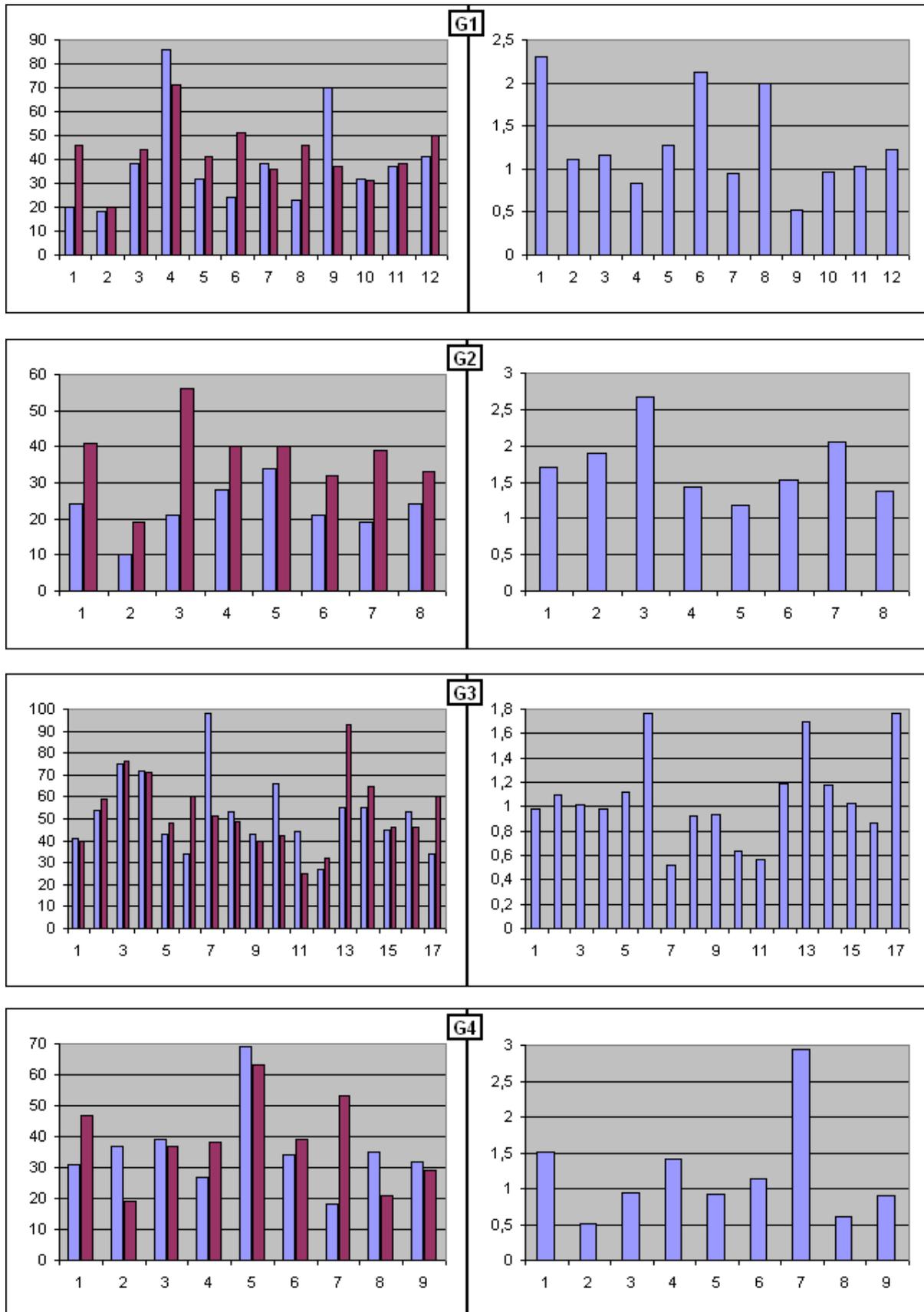


Рис. 1

Из диаграмм видно, что количество указанных студентами связей в основном попадает в диапазон от 20 до примерно 50. Лишь небольшая часть студентов имеет лучшие показатели. Причем интересно, что результаты для обеих групп специальности физика и информатика (G1 и G3) явно выше, чем для специальности информационные технологии (группы G2 и G4). По-видимому, студенты-физики лучше знают термины, связанные с электроникой и устройством компьютера. К тому же они изучают курс архитектуры на год позднее (на третьем, а не на втором курсе).

Как видно из приведенного рисунка, имеются отдельные студенты, которые умеют видеть связи заметно лучше, чем остальные. Например, студентка номер 13 из группы G3 сумела назвать после курса почти вдвое больше связей между терминами, хотя и до его изучения она фиксировала их более 50! Также можно выделить студентов номер 6 и 8 группы G1, студентку номер 3 группы G2, студенток 6 и 17 из группы G3 и студента номер 7 группы G4. Все они заметно превзошли остальных по количеству правильно указанных связей. Заметно выделяется своими результатами также студент номер 5 группы G4 (более 60 связей). Просмотр его файла показывает, что он выделяет связи очень методично: сначала указывает все связи для одного понятия, затем для другого и т.д. Несмотря на очевидную логичность такого подхода, большинство студентов группируют понятия более хаотическим образом. Подчеркнем, что все перечисленные выше студенты отнюдь не являются лидерами в группах, так что можно сделать вывод о существовании некоторой способности к анализу связей у отдельных людей. Лучшие студенты (под номером 1) из всех групп за исключением G3 также показали хорошее улучшение рассматриваемого параметра.

Поскольку наибольший интерес представляет не столько само количество связей для каждого из студентов, сколько его изменение в результате усвоения курса, в правой части рисунка построены соответствующие диаграммы для отношений количества связей после и до изучения курса: $K_L = L_2/L_1$. Теоретически количество связей, указанное при входном тестировании, не должно уменьшаться. Следовательно, значения отложенных на правых графиках отношений должны обязательно превышать единицу. К сожалению, на практике это далеко не всегда так. Разумеется, значение коэффициентов приращения «чуть меньше», чем единица ($K_L \approx 0,9-0,99$) можно объяснить естественной погрешностью эксперимента: студент, целостность знаний которого мало изменилась, в силу случайных факторов мог вспомнить немного меньше связей при второй проверке, чем при первой. Но это не объясняет всех «аномалий».

Рассмотрим, например, очевидный случай студента номер 9 группы G1. В ходе первой проверки он указал 70 связей между понятиями, а в ходе второй – 37 (уменьшение почти вдвое!) Ясно, что это не погрешность, а нечто другое. Внимательно посмотрев, кто еще из студентов при первом тестировании показал высокий результат, обнаружим *только одного* такого студента – с номером 4. И оказывается, что он в ходе первого эксперимента сидел рядом(!) со студентом номер 9. Таким образом, мы получаем вполне рациональное объяснение неправдоподобному результату. Конечно, наличие таких грубых отклонений не делает чести экспериментатору, который должен был следить за самостоятельностью работы студентов более тщательно. Но не стоит забывать, что многие студенты третьего курса прекрасно умеют консультироваться незаметно для преподавателя.

Отметим, что наиболее неудачными были результаты в группе G4; субъективные наблюдения за работой студентов подтверждают, что к выходной проверке эта группа отнеслась наименее добросовестно.

Попробуем теперь на базе имеющихся в нашем распоряжении первичных параметров построить более сложные **вторичные** параметры, которые получаются на основе некоторой обработки первичных. Как показало изучение предварительных результатов группы G0 [Еремин, 2009], наиболее удачный вторичный параметр получается путем следующей обработки данных. Связанные понятия объединяются в группы, как это уже было описано выше. Очевидно, что для идеального образовательного курса все

понятия должны войти в одну группу, т.е. все они должны быть связанными. Реальная картина существенно менее целостна: понятия формируют несколько групп: в лучшем случае их число составляет 3-5, а в худшем – превышает 20. При этом многие группы состоят всего из 2-3 понятий, что можно смело интерпретировать как отдельные не связанные с общей картиной факты.

Если поделить общее число терминов T на число групп G , то полученное отношение будет иметь смысл среднего размера группы терминов для данного студента. Эта величина, как показывает анализ результатов, также может служить хорошей характеристикой целостности знаний студента.

Значения величин T_1/G_1 и T_2/G_2 для всех групп G1-G4 приведена на рис. 2 (номера студентов те же самые, что и на предыдущем рисунке). Графики устроены аналогичным образом: слева приведены пары значений среднего размера групп до и после изучения курса, а справа – отношения значений каждой пары, которые будем обозначать K_{TG} . Основная часть выводов относительно конкретных студентов, обсуждавшихся выше, подтверждается и здесь, причем даже более отчетливо. Особенно хочется отметить, что в новых переменных становятся еще ярче видны слабые результаты группы G4 в целом и резкое превосходство в ней уже упоминавшегося выше студента номер 5. Тем не менее, есть и некоторые различия между первичными и вторичными характеристиками. Так «в новых координатах» не столь впечатляюще выглядят достижения студентки номер 3 из группы G2, а также, за исключением группы G2, снижены вторичные показатели у лидеров (студентов номер 1). Получающуюся разницу можно объяснить тем, что первичный показатель K_L , в отличие от K_{TG} , никак не учитывает, насколько названные студентом связи относятся к одной и той же группе терминов. В самом деле, одинаковое значение K_L могут иметь и тот студент, который назвал множество изолированных фактов, и тот, который сумел связать абсолютно все термины в единое целое. Вторичный же параметр K_{TG} «почувствует» эту разницу, что делает его более подходящим для целей нашего эксперимента.

Анализируя рис. 2, можно заметить еще один интересный результат. Практически все слабые студенты, которые имеют большие значения номеров, обеспечили прирост измеряемой величины. Эта закономерность была отмечена еще в предварительных результатах группы G0 [Еремин, 2009]. Новые эксперименты ее подтвердили. Даже в «наиболее проблемной» группе G4 студенты с максимальными номерами 8 и 9 имеют коэффициент прироста выше, чем у остальных (он близок к единице, хотя и слегка меньше ее). Таким образом, слабые студенты явно демонстрируют в эксперименте улучшение целостности знаний.

Усредненные по группам значения некоторых первичных и вторичных показателей сведены в таблицу 2. В ней приведены: средние количества терминов T , связей L и групп терминов G , а также средние для отношений T/G и L/T . Как обычно, индексом 1 помечены результаты входной проверки, а индексом 2 – выходной. Приведенные показатели позволяют судить об уровне предварительной подготовки и результатах изучения курса всей группой в целом.

Таблица 2

группа	T_1	T_2	L_1	L_2	G_1	G_2	T_1/G_1	T_2/G_2	L_1/T_1	L_2/T_2
G1	48	51	38	43	13	12	4,14	4,50	0,78	0,82
G2	32	46	23	38	10	11	3,37	4,23	0,71	0,81
G3	62	60	52	53	13	13	5,10	5,35	0,83	0,87
G4	42	45	36	38	8	11	6,05	5,54	0,85	0,84

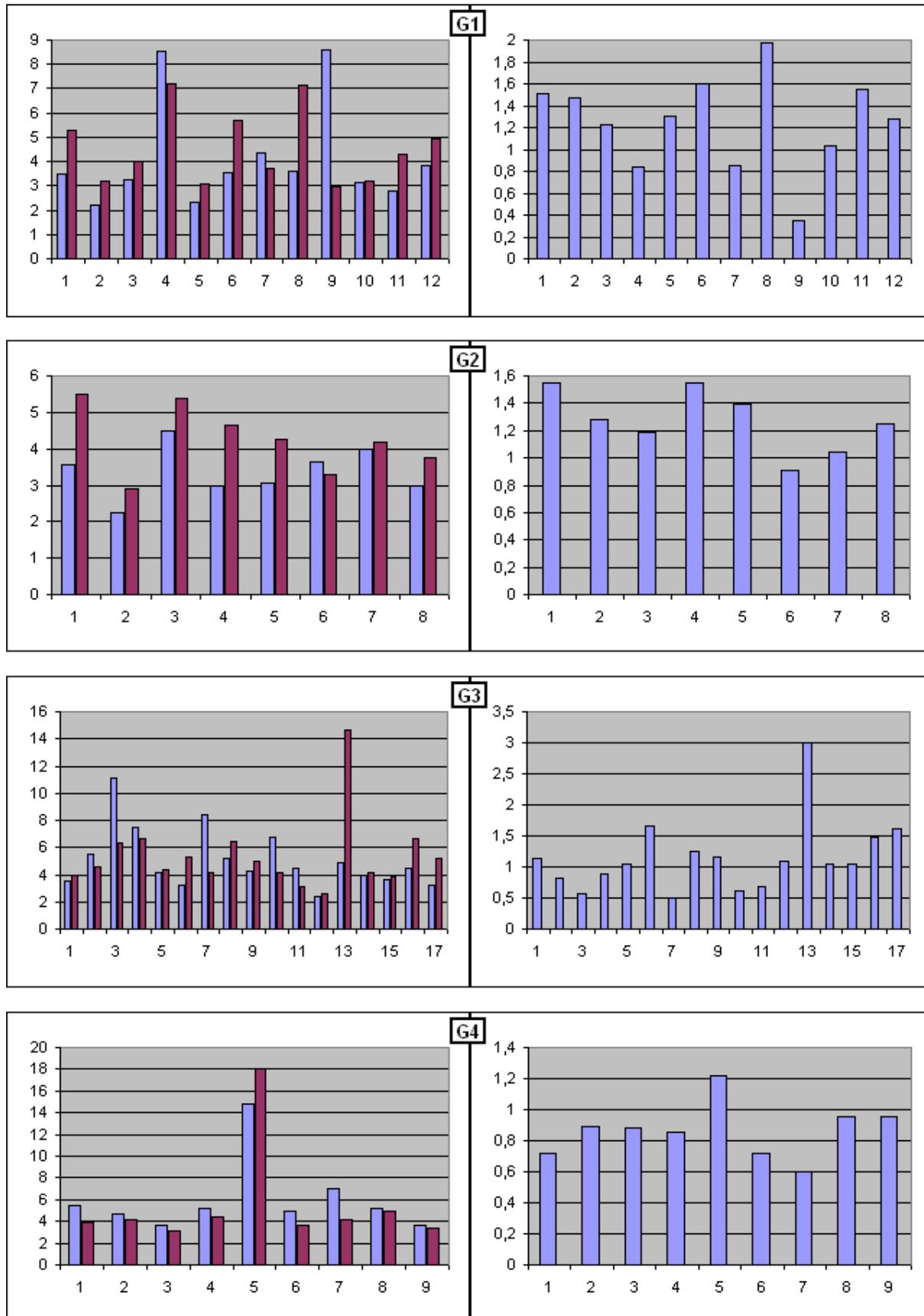


Рис. 2

Наконец, еще один наглядный способ описания, предложенный автором в предварительной публикации [Еремин, 2009], – это специальные диаграммы целостности знаний (см. рис. 3). Такая «пятнистая» диаграмма построена следующим образом. Каждая точка в столбце диаграммы – это одно понятие. Столбец разделен на полосы, которые являются графическим представлением сформировавшихся у студента групп понятий. Для улучшения наглядности группы имеют чередующуюся белую и серую окраску, а нижняя, самая большая, которая образует своеобразное базовое ядро из понятий курса, – черную. Нетрудно также заметить, что группы во всех столбцах упорядочены по величине, так что самая большая группа всегда находится внизу, а самые маленькие, состоящие из 2-3 понятий – в верхней части столбика.

Таким образом, приведенная на рисунке комплексная визуальная форма представления информации сочетает в себе следующие сведения:

- высота столбика пропорциональна общему числу усвоенных понятий, которые имеют хотя бы одну связь с другими понятиями;
- количество разноцветных областей характеризует степень «разрозненности» (фрагментарности) системы понятий;
- размер групп связанных понятий (на рисунке характеризуется площадью соответствующих областей) также свидетельствует о целостности представлений в той или иной части курса;
- размер нижней, самой большой, области (на рис. 3 она выделена черным) отражает объем базового блока понятий курса.

В идеале диаграмма должна представлять собой однородный (состоящий из одной группы) столбик черного цвета, причем его высота для списка, применявшегося в наших экспериментах, должна соответствовать более чем 120 терминам, т.е. почти вдвое выше, чем это получалось на практике.

Используя приведенные на рис. 3 диаграммы, можно получить более наглядную картину успешности освоения курса студентами. Если рис. 2 позволяет сравнить целостность знаний только по одному усредненному показателю, то «пятнистая» диаграмма гораздо лучше отражает распределение понятий по группам. Она служит своеобразной «картой знаний», глядя на которую преподаватель может сделать определенные выводы об усвоении своего курса.

Рассмотрим в качестве характерного примера студентку номер 5 группы G2. Ее количественный показатель – общее число сгруппированных понятий практически не изменился. (Если вернуться к рис. 1, то и увеличение общего количества связей у данной студентки тоже невелико.) Зато отчетливо видно, что структура знаний значительно улучшилась: выросла базовая группа понятий (выделенная на рисунке черным цветом), а количество мелких групп (символизирующих изолированные факты), напротив, уменьшилось. Заметим, что показатель $K_{TG} \approx 1,4$ (см. рис. 2) значителен, т.е. улучшение целостности знаний есть несмотря на то, что количество связей изменилось мало.

С точки зрения преподавателя, наиболее удачными следует признать результаты, продемонстрированные в группе G2. В ней все студенты не только показали прирост количества связанных понятий, но и улучшили структуру своих знаний – диаграммы выходной проверки выглядят более целостно. Обращает на себя внимание тот факт, что G2 является самой маленькой группой (всего 8 человек); возможно, это обстоятельство привело к более успешному контролю самостоятельности выполнения заданий.

В то же время подробность диаграмм на рис. 3 одновременно оборачивается их недостатком. Причина в том, что хотя наглядность диаграммы и удобна для преподавателя-человека, она трудно применима в автоматизированных обучающих системах. В самом деле, на что в первую очередь следует обращать внимание: на размер максимальной группы терминов или на общее количество групп? А, может быть, за основу лучше принять меньшее количество «мелких» изолированных групп? На эти вопросы трудно дать однозначный ответ.

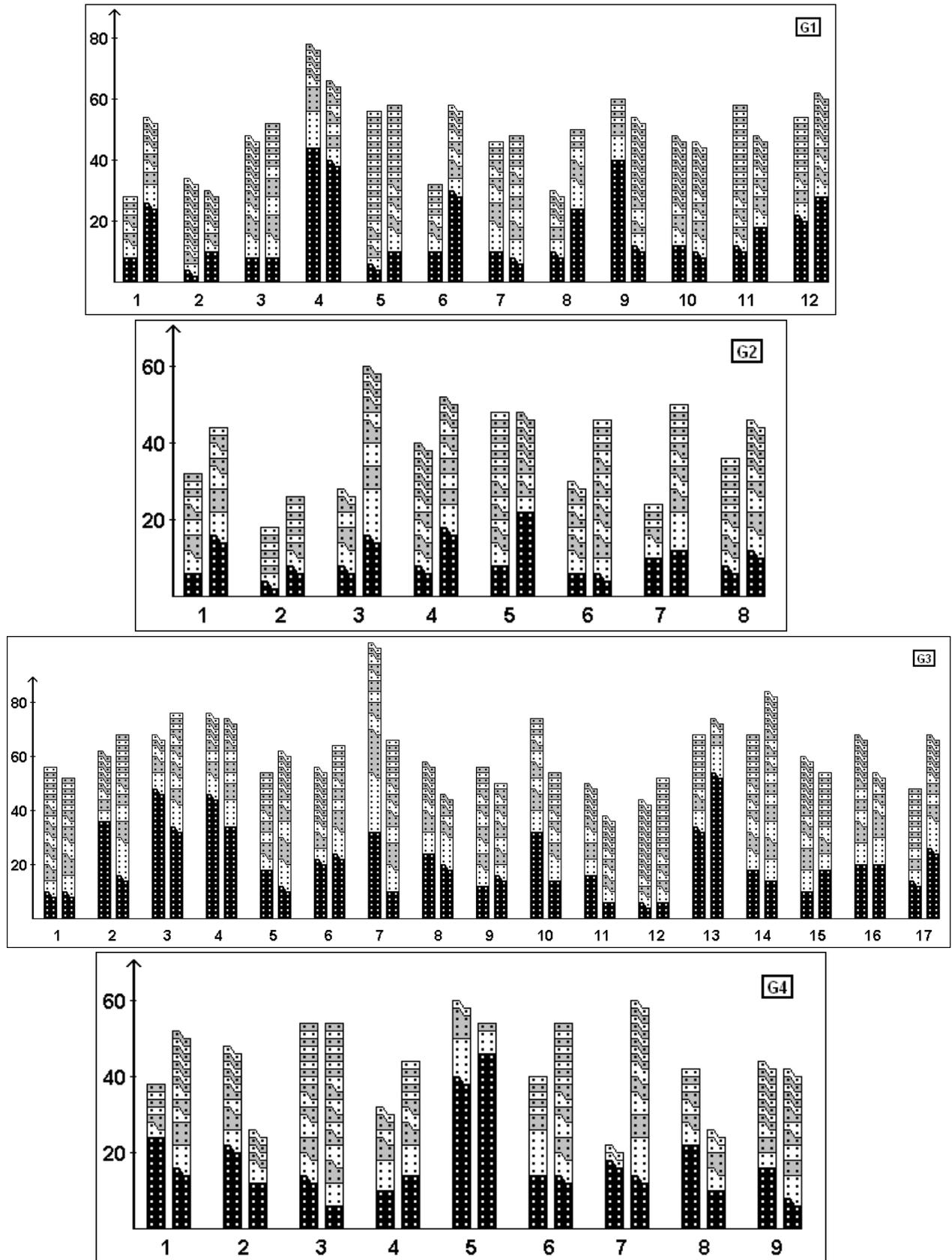


Рис. 3

Из результатов группы G4 снова выделяется систематичностью своих знаний студент номер 5, о котором уже шла речь выше.

При построении рассматриваемых диаграмм был также проведен дополнительный эксперимент, целью которого было проверить влияние на его результаты знания студентами цели проверки. Для этого четверем самым недобросовестным студентам группы G1, уличенным в сдаче чужих файлов было предложено не просто повторить тестирование, но и через некоторое время выполнить задание еще один (третий!) раз. Но в последнем случае они уже знали, что такое целостность знаний, в чем состоит цель тестирования и как оцениваются его результаты. Студентам было предложено обратить внимание на то, чтобы связать между собой все известные им понятия. В результате размер базовой группы понятий возрос в 2-4 раза, но единой группы все равно не получилось! Конечно, это были не лучшие студенты, но результаты все равно свидетельствуют о том, что целостность знаний не приходит по желанию.

Как показали приведенные выше результаты, достижения студентов, связанные с целостностью знаний, не всегда однозначны и имеют индивидуальные особенности. В частности, ранее уже было отмечено, что рост количественных показателей надежно регистрируется у слабых студентов. Попробуем проанализировать полученные данные с целью выяснить, нет ли других связей между экспериментально полученными параметрами и успешностью освоения курса. Обратимся к значениям прироста среднего размера группы терминов $K_{T/G}$, которые приведены в правой части рис. 2.

К сожалению, группы G1-G4 имеют разное число студентов и данные необходимо предварительно нормировать. Для этого можно предложить следующий несложный прием. Введем некоторую формальную величину X , которая будет описывать нормированный номер студента. Самому первому студенту поставим в соответствие значение $X=0$, а самому последнему – $X=1$. Для остальных студентов значения X нетрудно вычислить по формулам

$$\Delta X = 1/(N-1), \quad X_n = (n-1)\Delta X, \quad \text{где } n = 1 \dots N$$

Здесь n – это номер студента, а N – общее количество студентов в группе. Вспомогательная величина ΔX имеет смысл расстояния между соседними точками для студентов данной группы; например, для группы G4 число студентов $N=9$ и $\Delta X = 0,125$. Для других групп значения ΔX будут другими, причем чем больше N , тем меньше ΔX (иными словами, точки вдоль оси абсцисс будут стоять плотнее).

Нормированные таким способом точки для всех четырех групп G1-G4 сведем в единый график $K_{T/G}(X)$, изображенный на рис. 4.

Его нижняя область ($K_{T/G} < 1$), закрашенная серым цветом, соответствует «некорректным» результатам (студент ухудшил свой показатель в результате изучения курса). Напомним также, что точки с малыми значениями X вблизи нуля соответствуют студентам, быстро и успешно освоившим курс, а большие X , напротив, отражают результаты слабых студентов. Глядя на полученный график, можно довольно уверенно выделить на нем 4 зоны; обозначим их Z1-Z4.

Зона Z1 соответствует самым сильным студентам. В ней имеется довольно значительный разброс значений $K_{T/G}$ от 0,5 до 1,5. Это наиболее способная часть студентов в группах, получение зачета для них не составляет особого труда. Поэтому результат здесь существенно зависят от степени заинтересованности в выполнении задания. Отдельные способные студенты настолько не стремятся показывать высокие результаты, так что даже получают $K_{T/G} < 1$.

Зону Z2, занимающую чуть меньше половины используемого отрезка оси X , можно назвать зоной средних студентов. Это студенты, которые учатся успешно и, как правило, выполняют все данные им задания добросовестно. Если отбросить единичную точку с $K_{T/G} \approx 0,5$ как нехарактерную, то большая их часть демонстрирует прирост целостности знаний. Некоторые точки ниже единицы вполне можно отнести к погрешности педагогического эксперимента.

Следующая зона Z3 демонстрирует даже больший разброс значений K_{TG} , чем Z1: значения здесь «прыгают» от $K_{TG} < 0,5$ до 3. Это слабые и (или) недобросовестные студенты, занимающиеся нестабильно и кое-как. Результаты это подтверждают.

Наконец, зона Z4 – это самые слабые студенты, которые, как сразу бросается в глаза, все обеспечивают прирост своих (не очень больших) знаний. По-видимому, для этой группы студентов улучшать свои знания – это *единственный* способ получить зачет. В частности, их уровень знаний низок, поэтому любые попытки обмануть преподавателя сразу же бросаются в глаза, так что остается только в меру своих сил заниматься.

Таким образом, получается очень интересная картина распределения параметра K_{TG} , вычисляемого из экспериментальных данных: две зоны Z1 и Z3 содержат трудно предсказуемые результаты, зато две другие (Z2 и Z4) оказываются стабильными. Данный результат, к сожалению, свидетельствует о том, что разработанная методика экспериментального определения целостности знаний не всегда дает надежные результаты.

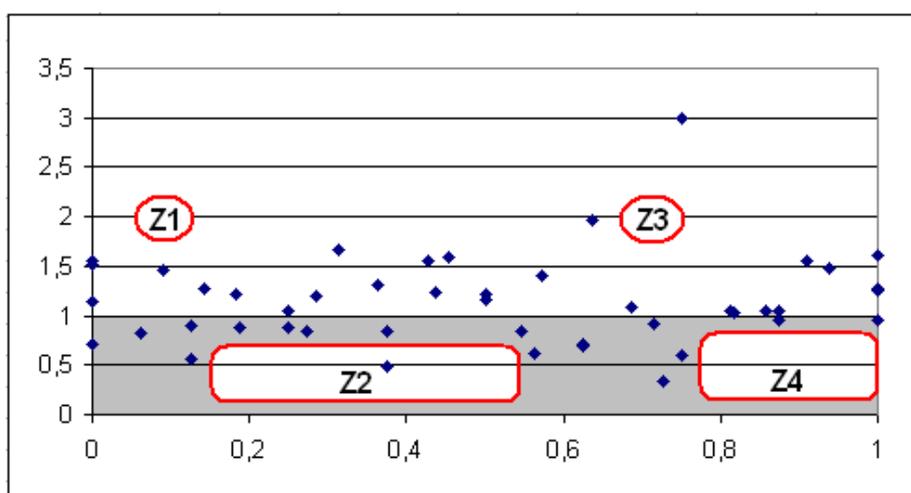


Рис. 4

Заключение

По проведенному анализу результатов эксперимента можно сделать выводы как по совершенствованию самой методики эксперимента, так и по целостности знаний у протестированных студентов.

Опыт проведения эксперимента позволяет выдвинуть следующие предложения по совершенствованию экспериментальной методики проверки связанности понятий, сформированных после изучения курса.

Поскольку, как показали наблюдения за ходом работы, выполнение дважды одного и того же задания нравится не всем студентам, можно предложить упростить проведение эксперимента: проводить проверку целостности знаний только после завершения курса, а результаты сопоставлять с некоторыми средними значениями, полученными ранее (например, воспользовавшись таблицей 2). По мнению автора, эта простая мера позволит улучшить объективность контроля, хотя и в ущерб возможности оценки индивидуального прироста знаний из-за отсутствия входного контроля.

При проведении итоговой проверки в самом конце курса некоторые студенты торопились сдать свои результаты. Поэтому, возможно, целесообразно ввести ограничение на минимальное время выполнения задания (можно, например, поставить условие, что время повторного тестирования должно быть не менее времени первичного).

В ходе проведения проверки знаний необходимо тщательно следить за самостоятельностью работы студентов и принимать все меры для предотвращения сдачи чужих файлов. Здесь хорошо зарекомендовало себя применение шифрования содержимого итоговых текстовых файлов.

Результаты показывают, что базу понятий для тестирования не стоит брать слишком большой. Есть предположение, что использованная в работе база была чрезмерно велика. В пользу этого вывода свидетельствует тот факт, что подавляющее большинство студентов использовали менее половины из предложенного списка понятий.

Все перечисленные рекомендации будут обязательно проверены в ходе дальнейших исследований целостности знаний студентов.

Перейдем теперь к выводам по полученным экспериментальным данным.

Эксперименты отчетливо показали, что почти все(!) студенты не умеют определять типы связей между понятиями. Они даже путают классические взаимосвязи *целое/часть* и *класс/подкласс*, не говоря уже о других нестандартных связях. По-видимому, преподавателям стоит обратить внимание на подобную ограниченность мышления студентов и развивать их возможности в этом направлении (разумеется, это относится к любому учебному курсу, а не только к курсу компьютерной архитектуры).

Как свидетельствуют результаты эксперимента, после изучения курса целостность знаний у разных студентов изменяется по-разному; существуют студенты, которые заметно лучше, чем остальные, видят связи между понятиями.

В связи с предыдущим выводом, становится понятным, почему не удалось обнаружить какую-то универсальную характеристику целостности знаний. Сопоставление различных характеристик позволяет предложить в качестве наиболее удачного параметра среднее количество терминов в группе. Хорошую помощь преподавателю при углубленном анализе результатов оказывают диаграммы группировки понятий, предложенные автором.

Интересным результатом является выявленное в результате экспериментов разделение всех студентов на несколько групп (зон) в зависимости от уровня подготовки. К сожалению, не во всех из них показатели целостности знаний измеряются надежно. Особо хочется подчеркнуть обнаруженные в исследовании зоны средних и очень слабых студентов, показатели в которых демонстрируют стабильное поведение.

Acknowledgement

The paper is published with financial support by the project ITHEA XXI of the Institute of Information Theories and Applications FOI ITHEA (www.ithea.org) and the Association of Developers and Users of Intelligent Systems ADUIS Ukraine (www.aduis.com.ua).

Bibliography

- [Еремин, 2007] Е.А.Еремин. Using Topic Map technology in the planning of courses from the CS knowledge domain. In: Proc. Seventh Baltic Sea Conference on Computing Education Research (Koli Calling 2007). CRPIT, 88. ACS, 2007.
- [Еремин, 2009] Е.А. Еремин. О компьютерной методике изучения целостности системы базовых понятий, сформировавшейся у студентов в результате освоения курса. В: Human Aspects of Artificial Intelligence. IBS "INFORMATION SCIENCE & COMPUTING", N 12, vol. 3. Institute of Information Theories and Applications FOI ITHEA, Sofia, 2009.

Author's Information



*Evgeny A. Eremin – Senior lecturer, Perm State Pedagogical University, 614990, Sibirskaya st., 24, Perm, Russia; e-mail: eremin@pspu.ac.ru
Major Fields of Scientific Research: Knowledge representation, Computer Learning, Computer architecture*