Original experimental method to evaluate conceptual students’ knowledge

Evgeny A. Eremin

Perm State Pedagogical University, Sibirskya, 24, Perm, 614990, Russia

Abstract

This paper considers the experimental research of measuring interconnections between the basic concepts, acquired during completion of a course in Computer Architecture. A special computer technique for estimating of students’ knowledge entirety was developed and successfully tested. An original mode of experimental data visualization is proposed. Specific pedagogical regularities were revealed by assessing how students digested the main concepts of the course.

Keywords: concept; relation; knowledge structure; entirety; education; course; assessment.

1. Introduction

The accepted classification, described in well known book of Anderson & Krathwahl (2001), itemizes several different kinds of knowledge: factual, conceptual, procedural and meta-cognitive. The simplest tests, often used for an assessment of students’ knowledge now, can amply estimate only the first one. Practical evaluation of conceptual knowledge, inclusive fundamental categories, their links and structure, is much more complex, but much more important (see Kay, 1995 for instance).

This work makes an experimental attempt to measure the interconnections between the basic concepts, acquired during completion of a course in Computer Architecture. The special computer method was developed and checked in order to estimate the rate of entirety of the system from basic concepts, which

* Corresponding author. Tel.: +7-342-238-6331; fax: +7-342-212-7019.
E-mail address: eremin@pspu.ac.ru.
students acquired by studying the material. By term entirety (wholeness, integrity) here we mean the presence of essential links between basic concepts within a course, which student perceives and shows during assessment. Several gauges that can characterize student’s knowledge entirety were proposed and compared; new complex method to visualize the research results is described. Some pedagogical regularity was discovered during the analysis of digesting the course’s foundation.

The study started in 2008 and at present we have results for 7 academic groups of students. This paper is fully devoted to consideration of experimental method and its preliminary results for the first group; results for other groups need further detailed discussion.

Initial aim of the research was the surmise that an entirety of student’s knowledge may be among effective characteristics of learning results. As follows from the proposed statement, those students, who see more associations between studied terms, have a higher rate of knowledge entirety, and, hence, digested this learning course better. The confirmation of this thesis would be very useful for learning process, especially for computer assessment of learning achievements.

2. Organization of experiment

2.1. Concept base

Using existing textbooks (see books Tanenbaum, 1998; Hamacher, Vranesic & Zaky, 2001; Cilker & Orlov, 2004 and Brydoy & Ilyina, 2006 for example) and personal teaching experience (Eremin, 2003), the author formed a list of basic terms and concepts. It was assumed, that the rate of digesting and mastery of these terms verified the success of the course.

The full list of concepts for the architecture course, used as experimental base, contained more than 120 terms. The most general concepts like computer, software and hardware, theoretical foundations etc., complemented by the terms that expand the previous ones – operating system, processor, memory, DMA, principle of hierarchy, byte and many others, were included in this base. Several terms from related disciplines such as microelectronics, logics and number notations were added into the list too. Contrary, the list did not include the names of concrete operating systems, external devices and their manufacturers, and other similar data, less essential from the position of learning the main course’s regularities. Using the standard terminology from object-oriented programming, we may say, that classes of the concepts were under consideration, but not their instances.

Experiments showed that obtained list was large enough, so competent students usually used a little more than the half of it in their answers. So the selected set of terms was found to be sufficient for assessment.

The next step was to analyze the links between the selected terms. Non-trivial result of this work consists in the fact, that a limited set of relations was enough to describe the subject. The final set consists of standard relations such as, for example, whole/part or class/subclass as well as several links, specific for the course, like base (principle of hierarchy – base – memory) or connection (controller – connection – bus). The full table of relations with concrete examples for each can be found in author’s publications (Eremin, 2007; Eremin, 2008): it contains only 11 base associations.

These results of analysis, published earlier, are of independent interest. From the point of view of this paper, constructed lists are original data, which is prepared to use during the experimental testing of student mastery of the knowledge domain under consideration. As knowledge control was realized with the help of computer, both lists (of concepts and relations), being input data for control program, were saved in the form of text files.
2.2. Computer software

To check students’ knowledge, the author developed two computer programs, realized in Delphi programming environment.

The first program was written to check the associations between the concepts that students have and fix them in the text file, suitable for further computer analysis. Its window contains three lists (see Fig. 1), using which a student forms a relation looking the following way:

```
term 1 – relation – term 2
```

(for instance, the relationship *processor – whole/part – register* can be easily decoded as phrase “processor and register are linked by association whole/part” or, more exactly, “register is a part of processor”). After student constructs a linked pair of concepts, s/he fixes it clicking the control button “Fix this link”. Then the constructed line is immediately added to multi-line text field, arranged in the bottom of the program window.

![Fig. 1. Screenshot of testing software](image)

Except the list of associations, loaded from pre-arranged file, the program covers a potential possibility to input additional (not provided by the author) name of relation between the terms: special radio buttons and text field in the center of the window allow such input. As experiments showed, students did not use this alternative, preferring to select link type from the available list.

When a student finishes his/her work, the results are saved in a text file, designed for the further analysis. All files were checked for correctness by the experimenter: distinctly wrong lines were erased from every result file. It was the only “manual” operation during the processing of testing results – all others were done by computer software automatically.

At the last step all checked files with the results of knowledge control fall under computer processing by means of the second program. Its main aim was to educe linked groups of concepts for every student.
For example, when computer, processing a file, finds associations functional units – class/subclass – processor, functional units – class/subclass – memory, functional units – class/subclass – input devices and functional units – class/subclass – output devices, it joins all five concepts, mentioned there, into one group. Later on some other terms will be added to this group: for instance, relations processor – whole/part – ALU and processor – whole/part – CU join two new terms to the group – arithmetic and logic unit and control unit.

In ideal case all concepts of the course are interrelated; test running of the program with thoroughly prepared author’s file, built according to full results from publication (Eremin, 2007), confirms this. But experiments show, that all real students’ files represent more scattered picture, which consists of several isolated groups of concepts, and some groups are very small (2-3 terms). Every such small group must be interpreted as a separate fact that student does not associate with other facts from the course. You must note that an increased rate of fragmentation indicates student’s knowledge is sparser.

Described program helps experimenting teacher to analyze the results of knowledge checking and output different statistics. During experiments software was modified for drawing graphic representation of calculated data.

2.3. Progression of experiment

The experiments estimated knowledge of students, who studied on the physical faculty of our university. Knowledge control took place twice: at the beginning and at the end of semester, that is to say before and after the learning of the referenced discipline “Computer Architecture”. Unfortunately, the number of students, who learned the course and took part in the experiments, was not too large.

Students did not know that the aim of the experiment was entirety of concepts’ system because it could artificially improve their results. They were simply told that testing tries to verify the rate of digesting the material. Their instruction was not to think about the parameters of evaluation, but just try to demonstrate digested knowledge the best way they could.

As students did not study the types of relations between concepts before, they were given a special table (with all types of links and numerous examples for each one) before testing. The results showed that it was insufficient and practically all students unsatisfactorily differentiated the types of links. Often they even missed classical relations whole/part and class/subclass, not to mention other link types. The aim of the experiments was to estimate the general entirety of the basic concepts – so from this point of view concrete kinds of links are not too important. Such effect became a motive to neglect the errors in this part of the task and just fix the existence of relation but not its type. This simplification of method notably facilitated the processing of results and ultimate analysis.

The time for implementation of the task was not limited, the work finished individually at students’ will. Presumably, it brought some uncertainty into experiment, because some students really indicated all interrelations they knew, but others just got tired and finished their work. The task execution employed about an hour in average. According to my observation of computer testing, the students’ reaction was mainly neutral (“we just get one more task”), so the checking procedure did not lead to any difficulty.

3. Discussion of results

3.1. Selection of gauge

Let us discuss what parameters can claim to be the characteristics of student’s knowledge entirety.

Primary parameters are evident: total number of terms and total number of links between them; they can be easily counted from any student’s file with the results. The ratio of these values, which means
average number of links per one concept, also may be introduced into consideration. It is evident, that the more these values are, the better student mastered the material.

Another set of parameters may be built while arranging interrelated terms into groups. We can offer here total number of concepts’ groups (this value must be as small as possible, in ideal case all terms must form the only group) and size of the largest group (this factor we want to see as large as possible). Additionally we can divide total number of terms by number of groups, i.e. get average size of group, which must be large when student learned the course profoundly.

All listed above characteristics were calculated for every student, and then compared for testing before and after the course. Average number of terms per group T/G seems to be the most suitable measure for knowledge growth. For the first experimental group this parameter varies from 3 to 13; essential changes of parameter make its experimental measure more reliable (Rogosa & Willett, 1983).

3.2. Diagrams of concepts’ interconnection

For demonstration of entirety of concepts’ system new original form of diagrams was proposed by the author (Fig. 2). This “spotted” diagram is organized the following way. Pairs of columns we see on it represent input and output testing (often called pretest and posttest): the right bar in every pair indicates final results. Height of any column is proportional to the number of concepts that student selected during assessment. Every dot in a column means one individual concept. All columns are divided into several areas; each of them represents a group of interrelated concepts, learnt by every student. For better visibility, neighbor areas are painted in white and gray colors. Number of the multicolored areas and their size characterize the rate of grasped data’s scattering and indicate knowledge fragmentation. The black region in the bottom, which is always the largest, symbolizes the kernel of student’s knowledge. As you can notice, all groups in every bar are regularized by size, so the smallest groups from 2-3 concepts (such groups may be interpreted as separate facts out of common picture) are always placed to the top of the bar.

![Fig. 2. Diagram of concepts’ interconnection](image)

In ideal case the diagram bar must be heterogeneous black bar (consisting from the only group), and its height must include all the concepts of the course. Real picture, as you can see from Fig. 2, is far from ideal: students’ knowledge comes apart on several independent groups of terms and the highest bar includes less than 70 concepts from more than 120 ones that were offered in the task.
It’s important to mention that students on the diagram are rank-ordered according to some rating: the criterion of such arrangement was time of finishing all the tasks, given by the teacher. The students with small numbers finished the course earlier; hence they are supposed to demonstrate better results in learning the course content. In opposite, columns for “the slowest” students form the right part of the picture. My subjective impressions and interview with students during learning process confirm the acceptability of selected criterion as measure of learning success.

Examining Fig. 2, we may get several practical conclusions about successfully digesting of the course content. The most evident of them is that all weaker students, who have large numbers, improved the entirety of their knowledge. Such phenomenon was earlier reported by Libarkin & Anderson (2005). We must emphasize, that growth we see for low-perform students is not error caused by known regression to the mean statistic effect (Barnett, van der Pols & Dobson, 2005): students’ level is evaluated independently in our research, but not from pretest scores.

Careful person may cognize, that there are some identical columns on Fig.2: for instance, compare right columns for students 2 and 10. This means that one of them cheated: just copied the result file from other student. Unfortunately, I noticed it too late to enforce students to redo their work. But for all further experiments special anti-cheating measures were developed.

4. Conclusion

The computer method for experimental research is offered, which allows to study the entirety of system of basic concepts after educational course learning. This method was tested on students’ learning the course content in “Computer Architecture” and the results showed its efficiency. This paper also describes a new original visual form of data representation, which clearly demonstrates the structure of interrelations between concepts in student’s knowledge. Some interesting pedagogical results were discovered during the research, for instance, how knowledge entirety depends on the level of student’s background. Further experiments with improved testing technique already brought more detailed results – they are worth separate extended discussion.

References

Kay, A. (1995). Powerful Ideas Need Love Too! Written remarks to a Joint Hearing of the Science Committee and the Economic and Educational and Opportunities Committee