Separation of Basic Competency Acquisition from Advanced Material in Teaching and Assessment in an Undergraduate Computing Subject


School of Computer Science and Engineering
The University of New South Wales
Sydney • N.S.W. • 2052
tom@cse.unsw.edu.au
+61 2 385 5995

Abstract

Undergraduate learning involves both the acquisition of the basic competencies required for adequate performance, and deep learning which comes from a thorough understanding of techniques and their application.

Normal assessment pressures tend to encourage superficial and rote learning at the expense of deep learning. It is much easier to assess basic competencies than deeper understanding. Continuous assessment methods attempt to address this problem, but due to pressures on staff and student time, it is inevitable that many assessment events, particularly final examinations, require the assessment of both basic competence and deeper material simultaneously. This is complicated, and hard to do well.

This paper reports on a long-term exercise to separate the assessment in an early computing subject into basic competency and advanced material as far as possible. The basic requirement to pass the subject (as externally accredited) is taught and assessed after two thirds of the semester. This assessment is based on continuous assessments, a practical laboratory examination, and a short answer written examination. At this point, students receive grades between 0 and 64%. Students achieving between 50 and 64% may stop at this point. The remaining third of the semester is used for advanced material to achieve higher grades, and for remedial teaching for those who have failed.

Our results are analysed using a comparison of student performance prior to the introduction of streaming, and using a subsequent 'control' year, when streaming was not done. The results show some improvement in performance of students in a subsequent computing subject, an improvement in pass rates in the subject itself, and a very marked reduction in requests for special consideration.

Introduction

The long-term exercise to separate the basic competence and deeper assessment described in this paper we call satisfactory-graded streaming, but will henceforth just use the word streaming. This paper describes the process in some detail, including the costs and benefits, with some quantification of the benefits in subsequent performance.

The following diagram shows a schematic view of the three routes students can take in this subject.
In the subsequent sections we will discuss the streaming under the following headings:

- The subject details
- Identifying basic competencies
- Beyond basic competencies
- Assessment – satisfactory stream
- Assessment – remedial stream
- Assessment – graded stream
- Costs of streaming
- Benefits of streaming

COMP1821 subject details

- Subject structure:
  - the three strands F, P, H (Functional / Procedural programming, Hardware) are linked by the theme of using levels of abstraction to handle the inherent complexity of computing systems.

The procedural programming, functional programming and abstract data types paradigms are used as vehicles in successor subjects including COMP2011, and are major paradigms for computer science professionals. The procedural paradigm also constitutes an important abstraction of the way the conventional computer operates.

The hardware strand serves as a foundation for later subjects in this area, introducing simple concepts of computer architecture and operating systems. This strand also serves to underpin students' understanding of high-level languages, by providing them with a lower-level description of computation, and then binding the two together with rules that students can use to translate high level programming language constructs into low-level descriptions. This translation exercises a disciplined assembly language programming methodology.

This subject is the second of a sequence of two subjects (COMP1811 / COMP1821) leading to all second year COMP subjects, mainly for non-Computer Science (CS) majors. For CS / Computer Engineering majors, the appropriate sequence is COMP1011 / COMP1021. It is intended that COMP1011 + COMP1021 = COMP1811 + COMP1821, but the material is presented in a different order, due to the requirement that both COMP1811 and COMP1821 can be terminating or continuing subjects.

The COMP1821 subject hence requires three strands, as it introduces functional programming and hardware, and continues with procedural programming concepts which are best dealt with after functional programming has been introduced. We have used results from this subject (in an encrypted form to protect student privacy) using neural network techniques to predict final marks based on continuing assessment performance (Gedeon and Turner, 1993). The extraction of explanations or rules from the network allows the discovery of educationally useful information, which is described elsewhere.

- Aims:
  - To introduce the concept of levels of abstraction as a consistent theme throughout computing.
  - To present a model of computation that is based on the concepts present in a modern functional language (Miranda), and in a modern practical procedural language (Modula-2).
  - To present a model of computation that is based on the structure and operation of a computing system, at several levels of abstraction.

- Objectives: Students should be able to:
  - List reasons for and give examples of the use of abstraction in programming.
  - Design, develop and test Miranda and Modula-2 programs that are accurate implementations of given specifications.
  - Solve computing problems through the design and
implementation of appropriate abstract data types.
◊ Describe how conventional von Neumann computers execute machine-code programs.
◊ Describe the sequence of operations computers perform to execute a high-level language program.

The above subject aims and objectives are taken from the preamble to the COMP1821 Lecture Notes (Gedeon and Wilson, 1993, 1995).

Identifying Basic Competencies

• F-strand:
  F1: Competency in developing short functional programs in a high-level language;
  F2a: Understanding of simple abstract data types such as stack, queue, tree;
  F3: Knowledge of basic data types, eg lists, trees.

• P-strand:
  P1a: Understanding of simple abstract data types such as stack, queue, tree;
  P2a: Understand the basic concepts of program verification: be able to verify simple assignment, sequence, and selection statements;
  P3: Knowledge of basic data types, including pointers, records, and linked lists.

• H-strand:
  H1: Competency in developing short assembly language programs, subroutines with stack frames for parameters and local variables;
  H2: Ability to complete simple programming steps: use an assembly language for a conventional machine;
  H3: Understanding of a lower-level description of the implementation of the conventional machine (e.g. a simple register transfer language);
  H4: Understanding of the relationship between the conventional machine and the high-level language level, as evidenced by the ability to translate from one to the other;
  H5: Ability to identify basic operating system concepts such as monitor, interrupt, command language, protection from a description.

• General:
  G: Knowledge of how to write a program in good style with internal documentation.

The above basic competencies were identified as the measurable subject results (postconditions) necessarily achieved by average (pass) performance.

Beyond Basic Competencies

The accreditation document identified a number of desirable postconditions, to be achieved by a Distinction-level performance. By interpolation, a Credit-level performance can be derived from the rules for pass and distinction level performance. That equates to a creditable attempt at the specified tasks.

• General:
  Excellence in the postconditions specified in the preceding section.

• F-strand:
  F2b: Ability to design and implement abstract data types from a requirements specification, in Miranda;

• P-strand:
  P1b: Ability to design and implement abstract data types from a requirements specification, in Modula-2;
  P2b: Ability to perform more demanding program verification steps: verification of loops;

The deeper skills required to achieve higher than pass grades are mainly built on the understanding and use of a complex and abstract programming construct in both the functional and procedural programming languages. This construct is called an "abstract data type," its proper use is simple and quite straightforward once the understanding of the concept has been achieved. This represents a very clear threshold in this subject, making the delineation of the basic and advanced material straightforward.

Assessment – satisfactory stream

• Duration of stream: 10 weeks
  ◇ 8 labtute exercises worth 2 each 16%
  ◇ H-strand assignment 20%
  ◇ 1 hour short-answer written exam. 24%
  ◇ 2 hour practical lab. exam. (F & P) 40%

A 'labtute' is a combined (computer) laboratory and tutorial session of two hours in duration. In introductory computing subjects, much of the tutorial material concerns executable code, hence can usefully be taught in conjunction with the explicitly practical laboratory exercise material. The 8 labtute exercises are designed to give practical experience in writing and reasoning about
(short) programs in assembly language, and functional and procedural styles. The practical laboratory examination involves two exercises similar in nature and level of difficulty to those attempted in the labtute exercises.

The marks for the satisfactory stream are mapped to the range from 0 to 64%, keeping the 0 to 50% range unchanged. The rationale is that raw marks above 50 represent a thorough mastery of the basic competencies, rather than indicating deeper understanding or higher level skills.

This reflects the notion that in our view there is a qualitative difference between achieving a pass grade versus a higher grade, and that our assessment instruments in the satisfactory stream have been designed with this in mind. Thus a high raw mark provides a very high level of confidence that the student has mastered the basic competency material, but provides no more information.

\[
\text{if raw } \geq 50: \quad \text{result} = 50 + \frac{\text{raw} - 50}{50} \times 14
\]

Note also that students must pass the written, practical examination, and continuous assessment components individually. Any student who fails overall or fails any component must attend the remedial stream.

Most students achieve marks greater than 50. At this stage they can opt to stop attending the subject and will receive a PASS grade, and this result as their final mark.

In the first semester streaming was introduced, we intended to award SY (satisfactory) grades at this point. This grade is an ungraded pass, and no percentage mark is returned. This grade is normally awarded only in subjects which return only pass / fail results. The central marks processing computer systems could not handle subjects which returned ungraded passes as well as normal grades, hence any SY grades had to be handled manually. Thus, PASS results with a percentage marks were returned except for students requesting that the SY grade be used. Only two students made this request.

Assessment – remedial stream

• Duration of stream: 4 weeks
  ◦ attend remedial labtutes, consultations

No extra lectures are provided, rather the focus is on small group laboratory exercises and tutorial interaction to repeat material which was not well understood, with new exercises at the same level.

◦ remedial H-assignment (if failed)
◦ remedial short-answer exam. (if failed)
◦ remedial prac. exam. (if failed)

These components were undertaken if the previous attempt was not a clear pass. Most remedial students passed at this stage.

◦ remedial part of final exam (if still failed)

In semester 2 of 1995, some 8 students made use of the remedial part of the COMP1821 final examination paper. At this point 5 of them passed, leaving only 3 students who have attempted and failed the subject out of 118 enrolled students (there were also 3 absent fails).

Assessment – graded stream

• Duration of stream: 4 weeks
  ◦ attend lectures, labtutes, consultations
  ◦ 2 labtute exercises worth 4 each 8%
  ◦ P-strand assignment 32%
  ◦ 2 hour written final exam. 60%

The two labtute exercises are designed to give practical experience in writing and reasoning about complex programs in a procedural language. The same language is then used in a major programming assignment.

The final examination is aimed at a demanding level, to test conceptual understanding. Note that all students have already passed the subject, hence there are no part-marks sections to questions for the regurgitation of basic or factual material.

\[
\text{if raw } \geq 50: \quad \text{result} = 50 + \frac{\text{raw} - 50}{50} \times 14
\]

\[
\text{newmark} = 50 + \frac{\text{graded}}{2}
\]

\[
\text{final result} = \max (\text{newmark}, \text{satisfactory})
\]

The previous mark achieved in the satisfactory stream is compared to the newly computed mark, and the larger returned.

In the following two sections, we will attempt to separate the costs and benefits of streaming, before making an overall assessment of the streaming experiment.
Costs of streaming

The costs of streaming are in three main areas, being the staff effort required, the extra administrative overheads, and student reactions.

The extra staff effort required comes from two main sources. Firstly, the streaming structure requires more assessment material to be prepared. If students are given extra opportunities to demonstrate their understanding, such as in the remedial stream, extra exercises, assignments and exam questions must be written. The streaming structure does confer some benefits which will be discussed in the next section. Secondly, there are more special cases which need to be handled during the semester. In the last third of semester students may be in the remedial stream having failed some differing combinations of the three main areas of assessment (written exam, continuing assessment & practical exams).

The extra administrative overhead is mainly the need to collect and finalise all marks twice a semester. This was a cost we had not expected would be as onerous. In a subject with multiple tutors and lecturers, it is a non-trivial exercise getting all the marks in, especially at a non-traditional time of semester. It appears that even the most recalcitrant of markers become socialised enough to be resigned to handing in marks at the ends of semesters.

There is also the extra overhead in re-allocation of students into labtutes for the last third of semester, as the common labtutes become either graded or remedial stream labtutes. Incidentally, for timetabling reasons we have settled on an adaptive method to allocate these, in that labtutes from the satisfactory stream in which a majority of students continue into the graded stream become graded stream labtutes. This minimises the number of students who may need to negotiate a new labtute time if they no longer fit their old labtute slot.

Of course tutors must also be reallocated, lending an extra constraint for tutors taking multiple labtutes, and we try to maintain the same time slots and keep the kind of labtute taught consistent. That is, either two graded or two remedial labtutes rather than one of each, as in some ways the final third of the subject could be seen as two different subjects with different learning objectives.

The final category of costs of the streaming, and probably the most important, is student reactions. The reactions relate to expectations and decision making. Students are used to relatively insignificant midterm examinations, and need some convincing to take the process seriously. This may be related to the inability of some few students in each semester to actually decide whether they wish to stop with a pass or to continue and attempt to achieve a higher grade. This is a consistent minor pattern.

The major pattern of student reaction is the converse pattern. Faced with a situation in which they are presented with more choices than they are used to, they see the opportunity to take choices not formally offered. The extra flexibility of the streaming structure invites the expectation of even more flexibility. This effect is most readily seen on the students who only (especially if narrowly) fail one component of the satisfactory stream. These students consistently and seemingly spontaneously bargain for extra assessment opportunities to allow them to continue in the graded stream. Some half of these students pass the extra assessments we have set and do continue in the graded rather than the remedial stream.

Benefits of streaming

The most significant benefit in the streaming structure is that the short sharp shock some students encounter comes early enough in the semester that there is time to remedy the situation.

This effect is most noticeable on subject pass rates. This variable is under the control of the lecturer-in-charge in the process of finalising marks, hence may not be very reliable. Nevertheless, the average pass rate increased from 73.6% without streaming to 83.9% with streaming. We will also indicate that in semester 2 of 1995, only 6 out of 118 enrolled students failed, and 3 of these were absent fails, who never showed up in class or the exam.

Personal experience (without statistical justification) leads to the observation that less than half of students sitting supplementary examinations pass in the end. The remedial stream students are in essentially the same category, yet the majority of them do pass by the end of the semester.

It is worth mentioning student stress. Finding out that one’s performance does not merit a pass is stressful. In the normal case, discovering this at the end of a semester, with very little time or staff help (being now out of semester) to learn the material is also stressful.

In contrast, with our streaming there is an entire third of a semester in which to recover from the shock, and time to decide to modify one’s learning behaviour, and take control (for example by bargaining for a particular form of extra assessment) and ensure a pass.

The issue of stress is also significant for the students who continue into the graded stream. They attend lectures, do the assignment and labtute exercises and sit the final exam knowing they are there voluntarily, that they have already passed, not just by their own calculation based on the part marks already received, but attested by their lecturer-in-charge. Thus their stress levels are reduced, and they learn better, as indicated by their subsequent performance, as we will detail below.

The streaming structure does provide some savings in
staff effort. As mentioned above, more assessment material must be produced. The material is more consistent in nature, however, and hence easier to produce.

For example, exam questions are simpler to write. The satisfactory stream questions and the equivalent remedial stream questions require only basic competency material. There is no need to attempt to produce questions which also test deeper knowledge and skills possessed by the better, and the best students. Where a normal exam question is a camel, designed for multiple tasks, a basic competency question is a racehorse, testing clear simple concepts and skills. Similarly, the graded stream exam questions are also easier to write. The question does not require some 'simple bits' to allow the base pass students some 'part marks' as the student who gets 0 in the graded stream final exam has not failed, instead has just not shown an understanding of the concepts above the passing level required to achieve a higher grade than the pass already obtained in the satisfactory stream.

Teaching in labtutes is also simplified, as the students are separated into rough category of ability. Thus, the effort required to keep the brighter students interested without losing the weaker students is reduced. Clearly this does not reduce the need to be aware of the learning needs of individual students, nevertheless the teaching is easier.

The clearest demonstration of the benefits of the streaming structure is on subsequent performance. The marks in a subsequent subject are used rather than the reference subject, as this is likely to be a more reliable indicator. It is worth mentioning that the lecturer-in-charge of the reference subject (COMP1821) was at no stage involved with the allocation or finalising of marks for the subsequent subjects, hence these results would be expected to be a more unbiased indicator. Most of the COMP1821 students continuing do COMP2011 in preference to the other subjects, so our subsequent analysis will focus on that subject alone.

Not all COMP1821 students continue to COMP2011, and not all of those who do continue do so in the following semester. To meaningfully average subsequent marks for students who completed COMP1821 in a particular semester but attempted COMP2011 in different semesters, each subsequent mark needs to be normalised with respect to the average marks for all COMP2011 students in the semester COMP2011 was done. This is also necessary to compare marks between different COMP1821 semesters. The following diagram shows the average mark in COMP2011 for all students (not just those continuing from COMP1821) has varied by up to 10%.

We have assumed that this variation is not related to the quality of students in different semesters or years, as we are not aware of substantial cohort differences. In any case, using the values below to normalise the marks for continuing students does affect the actual numbers but does not produce or eliminate the effects we have reported as deriving from our streaming.

To demonstrate the effect of streaming on subsequent performance, however, it is necessary to provide some background into the known differences in the student cohorts in COMP1821.

Students doing COMP1821 in semester 1 of each year are mainly engineering and science students, while those in semester 2 tend to be from humanities disciplines. This shows up as a difference in the average marks of all students in the subsequent COMP2011 subject. This result is perhaps unsurprising given the greater experience with technical subjects the former group would have, and given the reduced significance of computing in the major sequences of the latter group.
no significant difference in subsequent performance which can be attributed to their semester cohort. For students with higher marks in COMP1821, the difference increases with their COMP1821 mark. That is, the better the performance in COMP1821, the bigger difference in subsequent mark that was achieved.

If we assume that the COMP1821 mark is a good measure of a student's competence, then this indicates that the higher the competence the greater the benefit from the subsequent subject. If there was an overall linear relationship for the entire range of marks, we would conclude only that better students get better marks. That this only holds for those that have stepped over the basic competency - advanced material threshold suggests that our choice of the point at which to separate basic and advanced material was well chosen. This is also supported by the observation that the mark range from 50 to 64% in COMP1821 does not have any correlation with the subsequent COMP2011 mark.

Note that the graph represents figures for all semesters including both streaming and non-streaming. The difference we have identified indicates that there has been a qualitative difference between marks of below 64% and those of 65% or greater, which we have formalised by our streaming.

The streaming was introduced in semester 2 of 1993, and ran for two semesters. The subject was run for the subsequent two semesters without streaming for various reasons, with streaming reintroduced in semester 2 of 1995. The description in this paper focuses on the latest version of streaming, though labtutes were introduced in session 2 of 1994. An analysis (not included) of the results shows that labtutes have had no significant effects on subsequent marks in COMP2011.

The following diagram shows the results achieved by students in the subsequent subject COMP2011. The number of students continuing is between only 21 and 41 students each semester, as many of the students take COMP1821 as a terminating subject.

Average COMP1821 students who continued showed no statistically significant average difference in subsequent marks.

The success of the streaming on the better students is shown by the comparison of the 1992 semester 2 and 1993 semester 2 semesters, and the 1993 semester 1 and 1994 semester 1 average marks. There was an increase of 7.6% and 6.9% respectively (we previously reported increases of 11.2% and 10.7%, but these were raw marks), with similar decreases when streaming was again not used. Thus, streaming had raised the baseline performance by about 6% and 7.5% in semesters 2 and 1.

Conclusion

In this paper we have described the method of satisfactory-graded streaming to separate the assessment of basic competency material and assessment for deeper knowledge and skills.

Our results are based on comparison of student performance before the streaming was done, and using a subsequent 'control' year, when streaming was not done. The results show some improvement in performance of students in subsequent computing subjects for at least a subset of the students. The differences in student cohort also affected the results of our streaming, with more effect on the stronger cohort.

We have also mentioned that there was an improvement in average pass rates in the subject itself by 10% with streaming. As this is under the lecturer-in-charge's control, this is indicative only. There has also been a very marked reduction in requests for special consideration.

The introduction of streaming has been a positive experience overall. There is a surprising barrier of student and staff expectations with respect to the nature and timing of assessments which needs to be overcome. The students seem to benefit, which for some groups of students is reflected in higher marks subsequently, and in overall reduction of stress.

As a final comment, we would like to note that the introduction of streaming allows students to explicitly make rational choices with respect to the effort they expend on this subject. Students for whom this subject is a minor component of their degree can chose to devote the last third of a semester to subjects which are part of their major sequence. Students always make choices of this nature, our streaming only helps make their position clearer, from the assessments made at the end of the satisfactory stream.
Acknowledgments

The invaluable help of Bill Wilson and Graham Mann in the implementation and practical development of the streaming is here gratefully acknowledged. I also acknowledge the encouragement of John Hiller the then Head of School.

I would also like to acknowledge the Professional Development Centre for their course entitled "Academic Leaders Workshop on Student Assessment," which I attended at the University of NSW in 1992, where the ideas arose on which this notion of streaming is based. From the attendees at the Workshop, I would like to thank all the participants in my group, and in particular Professor Adrian Lee, who kept the idea alive subsequent to the Workshop and later prepared a discussion paper. This support and encouragement at that stage was indispensable.

A previous version of much of this material was presented in 1995 at UNSW (Gedeon et al, 1995), where I received many useful suggestions which I have incorporated into the data analysis, though any errors remain my own. Last, but not least, I acknowledge the motivating influence of Marie Wynter, who was enthusiastic about the whole idea.

References


Gedeon, TD and Wilson, WH COMP1821 Computing 2 Lecture Notes, School of Computer Science and Engineering, Session 2, 1993.

Gedeon, TD and Wilson, WH COMP1821 Computing 2 Lecture Notes, School of Computer Science and Engineering, Session 2, 1995.