Online TA

Master Project

Datalogisk institut, Copenhagen University (DIKU)

Oleksandr Shturmov
mail@oleks.info

March 1, 2014.

Chapter 1

Introduction

> *In any teaching of the application of computers it is essential to have the students do practical programming problems and to grade their results. Such grading should consider both the formal correctness and the performance of the programs and tends to become difficult and time consuming as soon as the teaching is beyond the most elementary level. The possibility of using the computer to help in this task therefore soon suggests itself.*
>
> — PETER NAUR, *BIT 4* (1964)

1.1 Reader Expectations

The reader is assumed to be familiar with the concept of a university, i.e. an institution of higher education and research, aimed at educating scholars and professionals, granting them degrees, signifying their accomplishments.

The reader is assumed to be familiar with Computer Science, i.e. the study of computable processes and structures, with the aid of of computers. Preferably, the reader should hold a Computer Science degree, be, or have been enrolled in a Computer Science university programme.

1.2 Dictionary and Grammar

Unless otherwise stated, the reader

1.3 Legal Disclaimer

This report references certain legal documents, including Danish laws and university curricula. As the author is not trained in law, there is no claim as to the legal soundness of the claims and references made in this report.

A solid attempt has been made at retaining the formulation of the referenced material, and referencing the most current legal documents, unless this was hindered by other legal references.

For instance, the shared section of the BSc and MSc curricula for study programmes at the Faculty of Science, University of Copenhagen [Curricula

(2013)] is based on Ministerial Order no. 819 as of June 29, 2010 [BEK 814]. This document is outdated and has been updated twice, most recently by Ministerial Order no. 1520 as of December 16, 2013 [BEK 1520]. In this particular case, it is the faculty curricula that was been deemed to mandate the relevant law to reference.

1.4 References

When looking for referenced material, material published in relevant scientific journals, yet accessible to the general public, was preferred. To ensure long-lasting access to publicly available, web-based resources, they were archived using WebCite®.

Chapter 2

Analysis

2.1 Assessment in Education

Assessment, or evaluation, in education is the practice of obtaining information about students' knowledge, attitudes, and skills [Pishghadam et al. (2014)]. The purpose of assessment may be manifold: provide feedback or certification, perform selection or comparison, improve learning processes, combinations of the aforementioned, and so on [Bradfoot & Black (2004)].

Assessment in education is primarily of students and their learning, or of teachers and their teaching. The former may later be used to improve learning, the latter, to improve teaching. In this work we will primarily concern ourselves with the former, so assessment for us will be getting to know our students and the quality of their learning [Ramsden (1992)].

2.1.1 Formative and Summative Assessment

There are two principal of kinds of assessment: formative and summative. The definition of each kind varies somewhat in educational research [Bloom et al. (1971), Sadler (1989), Harlen & James (1997)], and their mutual compatibility is questionable [Butler (1988)]. The intent of this work is not to advise on the matter, but to aid in performing assessments, regardless of the flavours.

Let us therefore adopt a primitive distinction, which still supports the purposes of our further analysis:

Formative A student's strengths and weaknesses are documented in free-text form. Formative assessments are qualitative and non-standardised: they are aimed at measuring the quality of a student's learning, rather than whether they live up to some standard criteria.

Summative A student is ranked on some well-defined scale, at some well-defined intervals, based on some well-defined criteria. Summative assessments are often compoundable and comparable. They may allow to deduce holistic summative assessments of students, or student groups, quantitatively measure student progress, etc.

Formative assessment necessitates the ability to perform personalised assessments, whereas summative assessment demands the ability to specify standards and perform standardised assessments.

There are other forms of assessment, such as diagnostic assessment, self-assessment, peer-assessment, etc. These forms of assessment still vary along the formative/summative scales, but differ in terms of when, by whom, and of whom the assessment is made.

2.1.2 Feedback

Feedback is information about the difference between the reference level and the actual level of some parameter which is used to remedy the difference in some way [Ramaprasad (1989)].

Feedback is an important bi-product of assessment in education [Black & William (1998)]. Ideally, feedback informs the student of the quality of their work, outlines key errors, provides corrective guidance, and encourages further student learning. To be so, it is important that feedback is understandable, timely, and acted upon by students [Gibbs & Simpson (2004)].

These requirements are an active area of educational research, and one aiding approach is to use computer-assisted assessment.

2.1.3 Computer-Assisted Assessment

Computer-assisted assessment is the form of assessment performed with the assistance of computers [Conole & Warburton (2005)]. The benefit of using computers is often fast, consistent, unbiased assessment, provided that the perceived student performance can be encoded in some useful digital format.

This requirement however, has proven evasive. Anything that is in free-speech or free-text form, is still not very useful [Valenti et al. (2003)]. On the other hand, it is questionable in-how-far the obviously useful forms, such as, multiple-choice questionnaires, are appropriate for assessment in higher education [Conole & Warburton (2005)].

Let a submission be a digital encoding of perceived student performance. Being digital, a submission is always done in some particular data format. A data format is an organisation of the data such that there exists an effective procedure that can recognise the encoded data, preferably in a finite number of steps.

The recognition of the encoded data constitutes the construction of a representation of the structure of the data. Such a representation will signify blobs within the data and their mutual relationships. The blobs of data may themselves be encoded using some other data format.

We say that the structuredness of data is inverse proportional to the maximum size of the largest blob of data allowed by its data format. For instance, a text file is not very well-structured since it is composed of a series of blobs roughly 4096 bytes each. An array of 64-bit integers, on the other hand, is fairly well-structured, having a maximum blob size of 8 bytes.

We hypothesize, that computers are best-suited for dealing with highly structured data.

Research into computer-assisted assessment provides some ground for this hypothesis. Multiple-choice questionaires are easy to assess: they can be encoded as a tableau of low-range integers. Free-text submissions are hard to assess: they can seldom be encoded as anything better than one large blob of data. The general approach in research in this area, seems to be to attempt to impose some structure on the otherwise plain data blob which is the submission [Valenti et al. (2003)].

Chapter 3

Feature Specification

3.1 Users

The online TA is to be made available on the World Wide Web. The set of users therefore, includes a wide range of principals of which only a very small subset will ever access the system, and even fewer are likely to use it as intended.

The system should be safeguarded against malicious users, yet provide all the relevant capabilities to authentic users with sufficient permissions.

3.1.1 User Roles

For an online TA there are two principal user roles: *system administrators* and *stakeholders* (students, teachers, and the general public). (TODO: is the general public really a stakeholder?) (TODO: how limited is the general public?)

System administrators have complete control of the servers running the service, including the mechanisms that enforce the security policies of the service. (They can, in principle, change anything.) System administrators are also responsible for monitoring the service for abnormal or abusive behaviour.

Stakeholders access the system through a secure online interface. They set up, enroll participants, and participate in courses — where a course is a collection of programming assignments for the students to solve.

There are four principle course-specific stakeholder roles: *instructors*, *assistants*, *students*, and *non-participants*. That is, for each course, every stakeholder has exactly one of these four roles.

Instructors are the principle course administrators. They can enroll other participants as either instructors, assistants, or students. Their primary role is to define the elements of a course.

Assistants are second-level course administrators. They can enroll other participants as either assistants or students. Their primary role is to assist the instructors in defining the elements of a course, and to provide feedback to students.

Students are the basic users of a course. They cannot enroll other participants, or tamper with the elements of a course. Their primary role is to submit solutions to the programming assignments of a course. Student submissions

are individual or group-based, and so not visible to other students in general. Student submissions are visible to instructors and assistants for the sake of evaluation.

Non-participants can see the elements of a course, but they cannot participate in the course in any way. The reasoning here is that there is no implicit trust relationship among the participants of a course, that does not exist among all the stakeholders of an institution. Providing access to non-participants allows them to get a feel for the course, which is useful for both encouraging participation and course evaluation.

NOTES:

- Only instructors and assistants can define the elements of a course.
- Only students can make submissions to programming assignments.
- Course contents is open for everyone in the system to see.

3.1.2 User Data

The system must retain sufficient information to securely authenticate and authorise all the users of the system.

3.1.3 System Administrators

3.1.4 Stakeholders

Stakeholders have to be identifi

- Name
- E-mail
- KU username

3.2 Courses

3.2.1 Course Data

Programming Assignments

- Assignment text. Everyone may see. Probably requires some pretty printing.
- Static submission analyses. Only instructors and assistants.
- Input data generator. Only instructors and assistants may see the actual program, everyone may see the resulting generated data. Resources limited to instructors and assistants.

3.3 Source Code and Documentation

3.3.1 OSS/FS

Open Source Software / Free Software (OSS/FS) programs are programs whose licenses give users freedom to run the program for any purpose, to study and modify the program, and to redistribute copies of either the original or modified program [Wheeler, 2007].

To fully reap the benefits of OSS/FS, it is necessary to ensure that there is an active community around it, devoted to bug tracking and fixing, as well as developing new features (unless the product can be deemed feature-complete).

If such a community is absent, the worst fears of the sceptics of OSS/FS may well be true: vulnerabilities may be known to adversaries.

Developing software that later

Bibliography

[Wheeler, 2007] David A. Wheeler. *Why Open Source Software / Free Software (OSS/FS, FLOSS, or FOSS)? Look at the Numbers!*. Revised as of April 16, 2007. Retrieved from http://www.dwheeler.com/oss_fs_why.html on March 7, 2014.

Archived by WebCite® at http://www.webcitation.org/6NtSnvALX.

[CS2013] ACM/IEEE-CS Joint Task Force on Computing Curricula. *Computer Science Curricula 2013*. December 2013. ACM Press and IEEE Computer Society Press. Retrieved from http://ai.stanford.edu/users/sahami/CS2013/final-draft/CS2013-final-report.pdf on March 7, 2014.

Archived by WebCite® at http://www.webcitation.org/6NtTXPH2s.

[EHEA (1999)] Joint declaration of the European Ministers of Education. *The Bologna Declaration*. June 19, 1999. The European Higher Education Area. Retrieved from http://www.ehea.info/Uploads/Declarations/BOLOGNA_DECLARATION1.pdf on March 9, 2014.

Archived by WebCite® at http://www.webcitation.org/6Nwr8h8l7.

[CND (2004)] Sebastian Horst and Carl Winsløw. *Undervisning i blokstruktur - potentialer og risici*. April 1, 2004. DidakTips 5. Center for Natufagenes Didaktik. University of Copenhagen. Retrieved from http://www.ind.ku.dk/publikationer/didaktips/didaktips5/5.undervisningiblokstruktur-potentialerogrisicimedomslag.pdf on March 9, 2014.

Archived by WebCite® at http://www.webcitation.org/6NwjYWlBi.

[BEK 814] Ministeriet for Forskning, Innovation og Videregående Uddannelser. *Bekendtgørelse om bachelor- og kandidatuddannelser ved universiteterne (uddannelsesbekendtgørelsen)*. Ministerial Order no. 814 of December 19, 2013. Retrieved from https://www.retsinformation.dk/Forms/R0710.aspx?id=160853 on March 9, 2014.

Archived by WebCite® at http://www.webcitation.org/6NyGHXc4N.

[BEK 1520] Ministeriet for Forskning, Innovation og Videregående Uddannelser. *Bekendtgørelse om bachelor- og kandidatuddannelser ved universiteterne (uddannelsesbekendtgørelsen)*. Ministerial Order no. 1520 of December 19,

2013. Retrieved from https://www.retsinformation.dk/Forms/R0710.aspx?id=160853 on March 9, 2014.

Archived by WebCite® at http://www.webcitation.org/6Nx4DxLGL.

[BEK 666] Ministeriet for Forskning, Innovation og Videregående Uddannelser. *Bekendtgørelse om eksamen og censur ved universitetsuddannelser (eksamensbekendtgørelsen)*. Ministerial Order no. 666 of June 24, 2012. Retrieved from https://www.retsinformation.dk/Forms/R0710.aspx?id=142560 on March 10, 2014.

Archived by WebCite® at http://www.webcitation.org/6NyI0DlqJ.

[BEK 250] Ministeriet for Forskning, Innovation og Videregående Uddannelser. *Bekendtgørelse om karakterskala og anden bedømmelse ved universitetsuddannelser (karakterbekendtgørelsen)*. Ministerial Order no. 250 of March 15, 2007. Retrieved from https://www.retsinformation.dk/Forms/R0710.aspx?id=29307 on March 11, 2014.

Archived by WebCite® at http://www.webcitation.org/6NzVOdWtv.

[Curricula (2013)] *The shared section of the BSc and MSc curricula for study programmes at the Faculty of Science, University of Copenhagen*. September 2013. Retrieved from http://www.science.ku.dk/studerende/studieordninger/faelles_sto/Faelles-del-ENG-2013-web.pdf/ on March 10, 2014.

Archived by WebCite® at http://www.webcitation.org/6NyEpPK67.

[Bradfoot & Black (2004)] Patricia Broadfoot and Paul Black. *Redefining assessment? The first ten years of assessment in education*. 2004. Assessment in Education: Principles, Policy & Practice, Vol. 11, No. 1, pp. 7–26. DOI: 10.1080/0969594042000208976. Retrieved from https://cmap.helsinki.fi/rid=1G5ND18R4-1QLJN7R-1SB on March 16, 2014.

Archived by WebCite® at http://www.webcitation.org/6O7OhF8LW.

[Pishghadam et al. (2014)] Reza Pishghadam, Bob Adamson, Shaghayegh Shayesteh Sadafian, and Flora L. F. Kan. *Conceptions of assessment and teacher burnout*. 2014. Assessment in Education: Principles, Policy & Practice, Vol. 21, No. 1, pp. 34–51. DOI: 10.1080/0969594X.2013.817382.

[Harlen & James (1997)] Wynne Harlen and Mary James. *Assessment and Learning: differences and relationships between formative and summative assessment*. 1997. Assessment in Education: Principles, Policy & Practice, Vol. 4, No. 3, pp. 365–379. DOI: 10.1080/0969594970040304.

[Butler (1988)] Ruth Butler. *Enhancing and understanding intrinsic motivation: the effects of of task-involving and ego-involving evaluation on interest and performance*. February 1988. British Journal of Educational Psychology, Vol. 58, No. 1, pp. 1–14.

[Sadler (1989)] D.Royce Sadler. *Formative assessment and the design of instructional systems*. June 1989. Instructional Science, Vol. 18, No. 2, pp. 119–144. Kluwer Academic Publishers. DOI: 10.1007/BF00117714

[Bloom et al. (1971)] Benjamin S. Bloom, J. Thomas Hastings, and George F. Madaus. *Handbook on Formative and Summative Evaluation of Student Learning*. 1971. McGraw-Hili, Inc. United States. Library of Congress Catalog Card Number 75129488. ISBN 0070061149.

[Ramaprasad (1989)] Arkalgud Rapaprasad. *On the definition of feedback*. January 1989. Behavioural Science, Vol. 28, No. 1, pp. 8–13.

[Black & William (1998)] Paul Black and Dylan William. *Assessment and Classroom Learning*. 1998. Assessment in Education: Principles, Policy & Practice, Vol. 5, No. 1, pp. 7–74. DOI: 10.1080/0969595980050102.

[Gibbs & Simpson (2004)] Graham Gibbs and Claire Simpson. *Conditions Under Which Assessment Supports Students' Learning*. May 2004. Learning and Teaching in Higher Education, Issue 1. Retrieved from http://www2.glos.ac.uk/offload/tli/lets/lathe/issue1/issue1.pdf on March 22, 2014.

Archived by WebCite® at http://www.webcitation.org/6OGkhvGyE.

[Ramsden (1992)] Paul Ramsden. *Learning to Teach in Higher Education*. 1992. Routledge. ISBN 0-415-06415-5.

[Conole & Warburton (2005)] Gráinne Conole and Bill Warburton. *A review of computer-assisted assessment*. March 2005. The Journal of the Association for Learning Technology (ALT-J), Research in Learning Technology, Vol. 14, No. 1, pp. 17–31. Retrieved from http://www.researchinlearningtechnology.net/index.php/rlt/article/download/10970/12674 on March 23, 2014.

Archived by WebCite® at http://www.webcitation.org/6OIAQzo2d.

[Valenti et al. (2003)] Salvatore Valenti, Francesca Neri, and Alessandro Cucchiarelli. *An Overview of Current Research on Automated Essay Grading*. January 2003. Journal of Information Technology Education, Vol. 2, No. 1, pp. 319–330.