Semi-automatic Assessment Model of Student Texts - Pedagogical Foundations

Tuomo Kakkonen, Erkki Sutinen
University of Joensuu, Finland

Key words: written composition, assessment, feedback, grading

Abstract:

We introduce the semi-automatic assessment model of student texts. This paper concentrates on the pedagogical foundations of the model.

1 Introduction

Many researchers of writing instruction are of the opinion that technology can and will in the future play an important role in the assessment of written compositions. A large number of studies (see, for example, [1]) have reported on automatic essay grading experiments in which the system has been able to equal and even surpass the levels of agreement between the grades assigned by human graders. While the automatic grading of essays is not a new idea automatic grading systems are not yet being widely utilized in real-life teaching situations. Various systems that are currently being developed are capable of providing assessments that go beyond mere holistic grading by allocating separate grades for aspects of essay writing such as spelling and grammar. Some systems also provide learners with feedback.

While Troia [2], for example, recognizes the need for providing “immediate, instructionally relevant multi-vector data to teachers … for pinpointing writing problems and responding accordingly”, he notes that none of the existing models is adequate for the task. The authors of the current paper have coined the term semi-automatic assessment to describe the features of the next generation assessment model for student texts, which is designed to offer the benefits of both fully automatic grading and feedback as well as the advantages that are gained when human assessors (the teacher, the student him/herself and peers) make assessments on the basis of their personal judgments as well as on the basis of the evidence collected by automatic assessment components. Assessment and the compilation of feedback is not an easy task, and it is in this respect that computers can support teachers by offering automatic support and “evidence” on which to base assessment decisions and feedback.

Semi-automatic assessment permits both the students and the teacher to take part in the assessment processes. Instead of assessing only the product (i.e. the final text), a semi-automatic system creates an environment that supports the actual processes of writing (composition) and assessment. What is achieved by the semi-automatic assessment model is that it allows students, their peers and their teachers to process and analyze a text by identifying its strong points and shortcomings, and to provide comments and instructions in a variety of formats. The EssayAid system that we have been developing over the past seven years represents a contribution towards the ultimate realization of a functional semi-automatic assessment system. While the technical and linguistic aspects of the model have been already explored in some detail, the pedagogical foundations have not yet been sufficiently
investigated. This paper is a step towards filling that gap by incorporating the best recommendations from the research on written composition and writing education into the semi-automatic assessment model.

2 Semi-automatic Assessment Model

2.1 Review of relevant pedagogical research

Product vs Process Models. In product-oriented models of written composition, an instructor is defined as a spotter of errors and reinforcer of rules [3]. In terms of this model, writing is defined as the process of being able to record preprocessed and fully formed ideas. While there is no single definition of a process writing model that is acceptable to all researchers, the definitions that are available in the literature have various features in common [3, 4, 5]: 1) The emphasis is more on the writing process than on the final product. 2) The importance of the content of the text (i.e. the ideas and meanings) is given precedence over concern about the form in which they are expressed (i.e. spelling and grammar). 3) Strive to be learner-centered rather than teacher-centered. 4) Emphasize the social implications of writing by according importance to the role of the envisaged audience and the interactions that take place between the reader and the writer. 5) Most of the available process writing models make a distinction between planning, writing and revising sub-processes. The sub-processes are also understood to operate recursively. Planning, for example, can take place during revision.

Summative vs Formative Assessment. Summative assessment is based on evaluating the learning achievements after the learning process itself. Summative assessment is therefore very similar to what happens in the product model of writing. Formative assessment, by contrast, evolves as a result of a series of interactions between the learner and the educator, and it occurs as a result of (1) the feedback that the learner receives during the learning process and (2) the learner’s own engagement in self-reflective contemplation of his or her own efforts to improve. Formative assessment is therefore very similar to what happens during the application of the process model.

Feedback and Revision. Feedback is designed to help students to improve the quality of their writing [6]. Feedback works by initiating student’s self-assessment and so to revise the text accordingly. Feedback should 1) focus on content, organization and other higher levels of language skills, 2) guide students on how to make improvements, 3) should explain exactly why a particular aspect of the writing is good or unacceptable, and 4) stimulate learners to think on their own.

Self- and Peer Assessment. The rationale for self-assessment is that students should be able to analyze their own texts and present such analyses as part of the assessment process. This empowers the learning process because it requires students to reflect on their own writing processes. In addition to providing feedback and a sense of audience for the writer, peer assessment is useful because it coaches students in critiquing skills and strengthens their ability to read texts analytically and critically [4]. Peer assessment also supports the process model of writing because it emphasizes the social dimensions of writing as an activity.

2.2 Fitting the Findings into the Semi-automatic Assessment Model

A semi-automatic assessment system for student texts is designed to create an environment that makes it much easier for a teacher to make useful and accurate assessment decisions because it is supported by computer-generated grading and feedback. The distinction between
product and process models of writing, and formative and summative assessment, provide the theoretical basis for the semi-automatic assessment model. While process writing may function as an invaluable tool for learners who are practicing the skills of composing and for second language learners, it needs to be accepted that the process writing model is not applicable to all situations. It would not be applicable, for example, to the assessment of student admission tests because what is needed in that context is merely a reliable ranking of students according to predetermined criteria. While both of these models are pedagogically and psychologically well grounded, they should be regarded as complementary rather than rival models. Whereas the product model provides the basis for making final assessments in the semi-automatic assessment model, fully automatic and human-guided feedback mechanisms support the process model.

We have therefore divided the semi-automatic model into two sub-models: *summative feedback cycle* and *formative product assessment*. While feedback and revisions are often repeated several times during the process of creating a text, the final formative assessment is not a recursive process. An instructor could use either or both of these sub-models to meet the needs of a specific assessment task. When assessing an essay examination, for example, a teacher would use only the formative product assessment model to allocate final grades for the students’ essays. On the other hand, a language teacher might well decide not to use the product assessment but utilize multiple iterations of the summative feedback cycle.

Both of the sub-models can, moreover, be configured to combine any combination of fully automatic, semi-automatic and/or peer and self-assessment phases. A typical formative assessment of essay answers would require the generation of both grading and feedback, and this could be achieved by combining automatic feedback with human scrutiny. We refer to this aspect of the semi-automatic assessment model as *computer-assisted assessment and feedback generation* (CAFG). It might be ideal, for example, to assess a project report document by directing it through two feedback cycles that consist of computer-generated feedback about grammar and vocabulary, peer assessment of the organization and content of the text, and the teacher’s comments on the overall organization and coherence of the text. Figure 1 illustrates this idea.

Figure 1 demonstrates that the CAFG component forms that could be characterized as the “core” activity of the semi-automatic assessment model. It is usually the last phase of an assessment or feedback process. This enables the information collected during the other phases of writing and assessment to be utilized for helping the teacher to make the final
assessment decisions. The revision history can reveal, for example, whether the writer has made only surface-level changes or modified the content and the organization of the text. The ideal scenario is that peer-initiated feedback offers students opportunities to think about new solutions to their problems and to consider new theories.

While the exact definition of the CAFG component will be left for future work, we present an outline below of those parts of the CAFG that can be implemented on the basis of functionalities that are already available in EssayAid:

*Fully automatic grading.* Automatic Essay Assessor is a system for automatically grading essays. The teacher can use the grades assigned by the system as a quality control mechanism by comparing the grades that he or she has assigned to the grades assigned by the system.

*Fully automatic feedback.* AntiPlag plagiarism detection system enables an instructor to identify those parts of student texts that have been plagiarized from other students’ works or from the Internet. EssayAid also has a component that offers feedback about grammar usage.

*Semi-automatic feedback.* TexComp tool is capable of producing comments on the writing style of a text on the basis of its degree of complexity. TexComp can be used to alert a teacher to the fact that while one text may exemplify an exceptional degree of complexity.

### 3 Conclusion

We have outlined the semi-automatic assessment model for student texts and shortly defined two sub-models of the model. By adapting the model to reflect specific combinations of summative feedback cycle and formative final assessment, an instructor may configure the model to offer whatever degree of flexibility an instructor may need for particular assessment tasks. In many ways, this work represents only the beginning of ongoing investigations into the pedagogical basis of semi-automatic assessment. Establishing the exact structure of the CAFG component is one of the main research directions for the future work.

### References:


### Author(s):

Tuomo Kakkonen, PhD
Erkki Sutinen, professor, PhD
University of Joensuu, Department of Computer Science and Statistics
PO Box 111, 80101 Joensuu, Finland
{tuomo.kakkonen, erkki.sutinen}@cs.joensuu.fi