Analysis of Spreadsheet Errors Made by Computer Literacy Students

Tanja Reinhardt, Nelishia Pillay
School of Geological and Computer Sciences, University of KwaZulu-Natal, Durban 4041, South Africa, Tel: ++27-(0)31-260-3018, Fax: ++27-(0)31-261-6550
reinhardtt2@ukzn.ac.za, pillayn32@ukzn.ac.za

Abstract

Spreadsheets have become a routine application in most organizations and Universities. As a consequence, students are required to learn spreadsheet applications such as Microsoft Excel. The learning of spreadsheets is often accompanied by problems related to spreadsheet application and their mathematical content. The EXITS (Excel Intelligent Tutoring System) research project aims to develop a Microsoft Excel tutor that helps students or learners to overcome their learning difficulties. In this paper, we analyse and classify spreadsheet errors made by students in order to determine the function that our system should perform and to generate an error library for student modelling purposes.

1. Introduction

Spreadsheet programs are widely used to analyse and manipulate numerical data. One can enter numbers into a spreadsheet and perform a large variety of mathematical calculations. A review of the literature [1, 2, 3] has revealed that although spreadsheets have been around for about 20 years, novice and expert spreadsheet users still have difficulties when using spreadsheets. When users start working with spreadsheet programs they have to cope with two challenges. One is to master the application program and the other is to understand the mathematical concept of spreadsheets. We experienced that users, in our case computer literacy students, have problems to learn and understand specific aspects of spreadsheet programs. Our aims are to find reasons why students have problems, identify the areas in which the problems appear and to develop a software program to help students to overcome these problems.

One approach to help students overcome the difficulties in using spreadsheets is an Intelligent Tutoring System (ITS). Although there are a variety of spreadsheets tutorials available they usually don’t have the ability to provide personal instruction and error-specific feedback as one would find in an ITS. The only research conducted in this area is the study carried out by Koedinger and Mathan [4].

To provide error-specific feedback it was necessary to identify the areas in which the problems appear and to look at the nature of errors made by the students.

2. Course background and experimental methodology

For our error classification we analysed the spreadsheets tests of 165 Computer Literacy students at the University of KwaZulu-Natal. The test included the use of various functions (e.g. SUM, AVERAGE, IF) and formulae, e.g., a formula using absolute cell addresses, as well as the creation of charts or sorting of data. Based on the experiment, we analysed and classified the errors made by the students.

3. Overview and analysis of formulae/function related errors

Most of the error classifications described in the literature have only provided a general overview of the errors made and thus lack the detail needed for student-modelling purposes in an ITS such as EXITS. [1,2]

For our error classification, we analysed each test in terms of the area where students made errors and what kind of errors the students made. We focused on the analysis of errors in formulae and functions believing that the students had more problems in dealing with formulae using cell references and functions than with formatting aspects.

Most errors occurred in areas where students had to solve problems using formulae and functions that require logical or mathematical thinking. 82% made errors when they had to create a formula using absolute cell addressing, 87% made errors when they used an IF...
function in combination with absolute cell addressing and 93% made errors when using a financial function.

We also looked at the different types of errors students made and classified three error categories that will help us to reduce the complexity of an error library: (1) General errors (GE) that are valid for any function and formula, e.g. students were not able to write down a solution at all. (2) General errors (GE) that occur in formulae and functions with a similar syntax, e.g. errors in the argument of functions like the SUM or AVERAGE function, which use the same syntax to perform a calculation. (3) Specific errors (SE) that only occur in particular formulae and functions, e.g. students used an incorrect condition in an IF-function. Figure 1 gives an overview of the general errors.

![Figure 1. General errors (GE-errors) using formulas and functions in the spreadsheet (Excel) test](image)

A study of the errors revealed two causes of these errors. Firstly, “conceptual-related errors”, where students made errors because of a lack of conceptual spreadsheet knowledge. For example, students didn’t know how to form a correct argument in a function. A second group of errors made were “mathematical related errors” e.g. the students where not able to set up an equation to calculate a percentage. This also includes logical errors, for instance, the use of wrong units when using a financial function. For some errors we are not able to distinguish whether the error is due to the lack of conceptual knowledge or due to the lack of mathematical/ logical understanding. These errors are marked as “errors that can be related to both categories”.

Based on the distinction of the different groups we are now able to structure the feedback in our ITS depending on the error made.

4. Conclusions

The study showed that most errors occurred in areas where students had to use formulae and functions that require logical or mathematical thinking, or an advanced understanding of concepts, such as the use of absolute cell addresses.

The errors that the students made can be classified as conceptual-related errors, or mathematical-/ logical-related errors. Apart from the two error categories “mathematical/logical-related” and “conceptual-related” we could also distinguish between general (GE) and specific (SE) errors.

The results of this study will be used in our student module to diagnose errors and to provide appropriate feedback and assistance.

5. References


