



# Quality Testing Concepts

**Paul Erlenwein**

paul.erlenwein@tu-dresden.de

Born on: 31st December 1996 in Ludwigshafen

Course: Distributed Systems Engineering

Matriculation number: 4609464

Matriculation year: 2016

## Bachelor Thesis

to achieve the academic degree

## Bachelor of Science (B.Sc.)

Supervisors

**Dr.-Ing. Birgit Demuth**

**Markus Hamann**

**Dr. Sebastian Götz**

Supervising professor

**Prof. Dr. rer. nat habil. Uwe Aßmann**

Submitted on: 5th February 2021

# Contents

1. Introduction	5
1.1. Motivation . . . . .	5
1.2. Research Questions . . . . .	6
2. Related Work	7
2.1. INLOOM . . . . .	7
2.2. Quality-Testing in existing ITS . . . . .	7
2.3. Supervised Learning . . . . .	7
A. Weitere Latex-Dokumentation	8

## List of Figures

## List of Tables

# 1. Introduction

## 1.1. Motivation

2020 was jinxed. The Covid-19 pandemic changed how our lives work. It has presented the globalised world with little anticipated challenges and we can feel its influence almost every aspect of our everyday life. In order to reduce the amount of human contacts as much as possible, every aspect of human interaction was evaluated for its digitizability. However: We were all forced to witness, that our digital infrastructure is most obviously not yet up to the task of enabling us to *live the remote life*.

Most of the fundamental problems were trivialities, like missing webcams or a too slow internet connection. Where such were taken care of, the harder-to-fix problems came to light. Problems like inadequately educated overtaxed personnel and missing software solutions, that comply with european digital privacy regulation. Even though at first glance the personnel problems don't seem to matter much to software developers, it are still problems, which can and will, at least in part, be resolved by them.

I don't want to claim, that the digitization of everyday life was a complete failure though. Like me, most office workers were able to migrate to home office without much fuss. Still: Living the life of a remote student for half a year, definitely motivates me, to spend some thought on how to make e-teaching a little better.

Even though it proved hazardous sometimes, working with existing e-teaching tools made me realize, what huge potential lies within a properly digitized higher education. Such would not only help temper the effects emergencies, like the covid pandemic, have on university life, but will also be a powerful tool in futureproofing universities for the challenges of rising student numbers in the years to come.

Intelligent-Tutoring-Systems like INLOOM, and ITS under active development at TU Dresden will make the increased workload manageable for university personnel. Developers are required to produce software that is as intuitive as possible, provides a decent grade of digital security, complies with privacy regulations, handles high traffic without complain and all that, while providing an unquestionably accurate and fair environment for everyone involved.

Integrating digital resources into the workflow seamlessly, will enable teaching personnel to still be able to focus on the individual student, when the student groups they teach become way bigger, than they are today - Which, in the end, could be the only way to keep higher education as we know it affordable. By contributing to INLOOM, I contribute to digitizing education.

## 1.2. Research Questions

The goal of this thesis is finding or developing a concept for validating evaluations, INLOOM generates for student created UML-models (autoEval), in order to ensure the correctness and fairness of said evaluations. Since not all mistakes, a student could make are predictable, it will not be possible to ensure, that the automatic grading INLOOM performs, finds all errors a student solution contains. Therefore it will not be achievable to perform an objective evaluation of INLOOMs results.

The only remaining option is performing a relative evaluation, a comparison of INLOOMs results with the best evaluation of the same model we know. These best evaluations, in INLOOMs case, are manual ones of the same student solution, created by human tutors (manEval). This in turn means, that the possibilities for validating INLOOMs results, are severely limited by the availability and ascertainability of the underlying data.

For the purpose of validating INLOOMs evaluations I will aim to answer the following research questions:

- RQ1** Which values can be extracted from the manual and automatic evaluations?
- RQ2** What comparative scale is qualified to provide a conclusive impression on the quality of the evaluations INLOOM generates?
- RQ3** How can the developer/tutor/instructor best be assisted in collecting and pre-processing the evaluation data, required for the quality validation?
- RQ4** How can the results of the comparison of man- and autoEval be presented to greatest effect?

The answer to RQ1 will determine which values are available to facilitate a comparison between man- and autoEval. The question is approached with an analysis of the available datasets. The second reaserch question focusses on what to do with the data, once it is collected. Assuming, that comparable data can be extracted from both man- and autoEval (so basically that RQ1 can be answered successfully), a way to combine the found relative differences between the two evaluations, into a conclusive quality score, is still required. Answering the question will be approached by performing a literature survey, concerning the quality validation methodologie, of existing ITS and automatic grading methods. Under any envisionable circumstances, it will be necessary, to digitize manEvals a tutor created for student solutions. This digitization process entails a high amount of overhead for the creation of testcases, that should be reduced as much as possible. Research Question three (RQ3), aims to resolve this problem and will be tackled by an analysis of the test-creation process, which will determine the, most workload intensive, steps of the process. These steps can then be considered in the design of the software solution proposed in this thesis. Lastly it remains to be determined how to best present or visualize the results of the meta evaluation (the evaluation of the evaluations) to the developer/tutor/instructor. This is the reason why RQ4 is listed among the research questions. Answering it, will depend greatly on the answer to RQ2.

## 2. Related Work

### 2.1. INLOOM

This work aims to validate the quality of the results the INLOOM Software [1] produces. For that reason it is inevitable to take a look into how it evaluates student solutions and how it persists the collected information. INLOOM is an acronym for *INteractive Learning center for Object-Oriented Modelling*.

The software, as the name suggests, is intended to be employed in a learning environment. It is used to evaluate student solutions to modelling tasks, the students have to work on as part of the mandatory beginner software engineering course at TU Dresden.

### 2.2. Quality-Testing in existing ITS

Due to rising student numbers and the availability of new interfacing technologies the interest, in automatic evaluation of student modelling work, has increased in recent years. Even though: Functioning evaluation tools and methods remain scarce. Starting point for the research into existing ITS, was a collection of such, referenced in [1]. Since those influenced decision made during the design of INLOOM, it is only natural to also include them

### 2.3. Supervised Learning

## **A. Weitere Latex-Dokumentation**



### Statement of authorship

I hereby certify that I have authored this Bachelor Thesis entitled *Quality Testing Concepts* independently and without undue assistance from third parties. No other than the resources and references indicated in this thesis have been used. I have marked both literal and accordingly adopted quotations as such. There were no additional persons involved in the intellectual preparation of the present thesis. I am aware that violations of this declaration may lead to subsequent withdrawal of the degree.

Dresden, 5th February 2021

Paul Erlenwein