

UNIVERSITEIT VAN AMSTERDAM

MASTERS PROJECT

Projecting Rules: Improving Comprehension of Business Rules with Projectional Editing

Author:

Paul SPENCER

Supervisor:

Dr. Clemens GRELCK

*A thesis submitted in fulfillment of the requirements
for the degree of Master of Software Engineering*

in the

Graduate School of Informatics
Faculty of Science

September 26, 2021



UNIVERSITY OF AMSTERDAM

Declaration of Authorship

I, Paul SPENCER, declare that this thesis titled, "Projecting Rules: Improving Comprehension of Business Rules with Projectional Editing" and the work presented in it are my own. I confirm that:

- This work was done wholly or mainly while in candidature for a research degree at this University.
- Where any part of this thesis has previously been submitted for a degree or any other qualification at this University or any other institution, this has been clearly stated.
- Where I have consulted the published work of others, this is always clearly attributed.
- Where I have quoted from the work of others, the source is always given. Except for such quotations, this thesis is entirely my work.
- I have acknowledged all of the main sources of help.
- Where the thesis is based on work done by myself jointly with others, I have made clear exactly what was done by others and what I have contributed myself.

Signed:

Date:

UNIVERSITEIT VAN AMSTERDAM

Abstract

Graduate School of Informatics

Faculty of Science

Master of Software Engineering

Projecting Rules: Improving Comprehension of Business Rules with Projectional Editing

by Paul SPENCER

Context: Declarative rules engine languages, such as Drools, can become difficult to reason about when there are many rules.

Objective: This project investigates whether projectional editing is a valid solution to this issue. If so, how can different projections of the code can ease the comprehensibility of the code.

Method: We conducted a systematic literature review, to ascertain the relevancy of projectional editing. We created an implementation of the Drools language using the MPS language workbench and made innovative projections of Drools ASTs. We validated our projections with a survey.

Results: Projectional editing is still niche, though it is making some industrial and educational headway in particular with JetBrains MPS. Projections can be useful, though, our survey found that our projections did not outperform the understanding of textual representation at least amongst experience Drools users.

Keywords: projectional editing; Rules Engines; MPS; Drools

Paper type: Research paper

Acknowledgements

It took a European village to raise this master's project.

We would like to acknowledge the School of The Graduate school of Informatics in the Faculty of Science at The University of Amsterdam for their guidance, specifically Dr. Clemens Grelck, who has been a supportive, understanding, and available academic adviser.

We received inspiration from the Strumenta Languages engineering community. Specifically we would like to thank Federico Tomasetti, who shared with me his model of a rules engine in MPS.

Also we would like to thank Václav Pech from JetBrains for the course he created and the time he spent with me explaining MPS. Further, Sergej Koščejev from JetBrains helped us with specific MPS issues during his Office hours.

Other prolific output that terrifically helped our research and development was the Heavy Meta YouTube series from Kolja Dummann and the dozens of papers and books from Markus Voelter, both currently working at Itemis A.G.

Michel Mercera who, amongst other things, provided me with helpful guidance on conducting and analysing surveys.

Our greatest thanks go out to Toine Khonraad, an alum of this course, who provided me with moral and monetary support, as well as wisdom and friendship that aided in the completion of this, my fourth attempt at getting this project behind me. Without his constant mantra of simplify, simplify, simplify, we would still be implementing the Drools languages now without having made a single projection.

Contents

| | |
|--|------------|
| Declaration of Authorship | iii |
| Abstract | v |
| Acknowledgements | vii |
| 1 Introduction | 1 |
| 1.1 Problem Statement | 1 |
| 1.2 Research Questions | 2 |
| 1.3 Contributions | 2 |
| 1.4 Project Context | 2 |
| 1.5 Thesis Outline | 3 |
| 2 Background | 5 |
| 2.1 RulesEngines | 5 |
| 2.2 Projectional Editing | 11 |
| 2.3 What is MPS? | 20 |
| 3 Methods | 27 |
| 3.1 Method: Systematic Review | 27 |
| 3.2 Method: Action Design Research - Drools in MPS | 34 |
| 3.3 Method: Survey | 43 |
| 4 Results | 47 |
| 4.1 Results: Systematic Literature Review | 47 |
| 4.2 Results: Action Design Research | 56 |
| 4.3 Results: Survey | 65 |

| | |
|---|------------|
| 5 Discussion | 77 |
| 5.1 Discussion - Systematic Literature Review | 77 |
| 5.2 Discussion - Survey | 79 |
| 6 Related Work | 83 |
| 6.1 The state of projectional editing workbenches | 83 |
| 6.2 Understandability of Business Rules | 83 |
| 6.3 Interesting projections | 84 |
| 7 Conclusion | 85 |
| 7.1 Research Summary | 85 |
| 7.2 Summary of contributions | 86 |
| 7.3 Further Work | 86 |
| Bibliography | 89 |
| A Protocol Validation Checklist | 99 |
| B Systematic Literature Review Log | 101 |
| C Study Quality Assessment Checklist | 109 |
| D Study Quality Assessment Results | 113 |
| E Data Extraction Results | 117 |
| F Drools Concept hierarchy | 123 |
| G Questionnaire Text | 127 |

List of Figures

| | |
|--|----|
| 2.1 Drools components | 7 |
| 2.2 Drools inference loop | 8 |
| 2.3 Drools rule breakdown | 9 |
| 2.4 Projectional editing loop. TODO: comparison with Parsing | 13 |
| 2.5 Leftfield projectional editors | 15 |
| 2.6 Concept example | 21 |
| 2.7 Behavior example | 22 |
| 2.8 Constraint example | 23 |
| 2.9 Typesystem example | 23 |
| 2.10 Editor example | 24 |
| 2.11 Intention example | 24 |
| | |
| 3.1 RSR concept hierarchy | 35 |
| 3.2 RSR | 36 |
| 3.3 Editors | 36 |
| 3.4 RSR substitute menu | 36 |
| 3.5 RSR scoping constraint | 37 |
| 3.6 RSR program | 38 |
| 3.7 Drools-Lite structure | 39 |
| 3.8 RHS | 40 |
| 3.9 Rule | 41 |
| | |
| 4.1 Search results | 48 |
| 4.2 Study types | 50 |
| 4.3 Public vs research | 51 |

| | |
|---|-----|
| 4.4 Publicly available programs | 51 |
| 4.5 Sentiment analysis by product | 52 |
| 4.6 Sentiment analysis | 53 |
| 4.7 Context aware color scheme | 56 |
| 4.8 Summary projection | 57 |
| 4.9 Filtering projection | 58 |
| 4.10 Table projection | 59 |
| 4.11 Cross-tab projection | 59 |
| 4.12 Table Fact deletion code | 60 |
| 4.13 Decision table projection | 61 |
| 4.14 Spreadsheet projection | 62 |
| 4.15 Intention | 63 |
| 4.16 Two of same property | 63 |
| 4.17 Truth table projection | 63 |
| 4.18 Circuit diagram projection | 64 |
| 4.19 Survey participants | 66 |
| 4.20 Survey locations | 66 |
| 4.21 Subject experience | 66 |
| 4.22 Editors used | 67 |
| 4.23 Question 1 - first impressions | 69 |
| 4.24 Initial thoughts | 69 |
| 4.25 Question 5 - interpret projection | 70 |
| 4.26 Question 7 - interpret text | 71 |
| 4.27 Question 8 - compare projections | 72 |
| 4.28 Question 9 - compare projections with text | 72 |
| 4.29 Question 10 - truth table | 73 |
| 4.30 Question 11 - circuit diagram | 74 |
| 4.31 Question 15 - responses | 75 |
| E.1 Data extraction results 1 - 4 | 117 |

| | |
|--|-----|
| E.2 Data extraction results 5 - 9 | 118 |
| E.3 Data extraction results 10 - 16 | 119 |
| E.4 Data extraction results 17 - 21 | 120 |
| E.5 Data extraction results 22 - 26 | 121 |
| E.6 Data extraction results 27 - 31 | 122 |
| | |
| F1 Rule file concept hierarchy diagram | 124 |
| F2 Rules concept hierarchy diagram | 125 |
| F3 Rule LHS concept hierarchy diagram | 126 |
| | |
| G1 Screen 1 - introduction text | 134 |
| G2 Screen 2 - first impression | 135 |
| G3 Screen 3 - positive response | 136 |
| G4 Screen 4 - describe projection | 137 |
| G5 Screen 5 - describe text | 138 |
| G6 Screen 6 - compare projections | 139 |
| G7 Screen 7 - compare projection to text | 140 |
| G8 Screen 8 - truth table | 141 |
| G9 Screen 9 - circuit diagram | 142 |
| G10 Screen 10 - personal details page | 143 |
| G11 Screen 11 - further comments | 144 |

List of Tables

| | |
|--|----|
| 2.1 Rules engine products | 7 |
| 2.2 Papers describing advantages | 16 |
| 2.3 Papers describing projectional editing disadvantages | 18 |
| | |
| 3.1 Search engines used | 29 |
| 3.2 Study design hierarchy for software engineering | 31 |

| | |
|---|-----|
| 3.3 Data extraction form | 32 |
| 4.1 Search engine sensitivity and precision | 49 |
| 4.2 paper key | 54 |
| 4.3 Mann-Whitney question 1 - first impressions | 69 |
| 4.4 Mann-Whitney question 5 - interpret projection | 70 |
| 4.5 Mann-Whitney question 7 - interpret text | 71 |
| 4.6 Mann-Whitney question 8 - compare projections | 72 |
| 4.7 Mann-Whitney Question 9 - compare projections with text | 73 |
| 4.8 Mann-Whitney Question 10 - truth table | 73 |
| 4.9 Mann-Whitney Question 11 - circuit diagram | 74 |
| A.1 Protocol validation checklist | 99 |
| B.1 Search engine/library key | 101 |
| B.2 Systematic review log - search results | 107 |
| C.1 Case studies quality assessment checklist | 109 |
| C.2 Qualitative studies quality assessment checklist | 110 |
| C.3 Survey studies quality assessment checklist | 111 |
| C.4 Cohort or panel studies quality assessment checklist | 112 |
| D.1 Quality assessment results | 116 |

Chapter 1

Introduction

The limits of my language mean the limits of my world.

Logico-Tractatus Philosophicus
Ludwig Wittgenstein

1.1 Problem Statement

Drools is a language that shares an unfortunate characteristic with many other rules languages. Namely, It is verbose and can contain many rules that can interact without an apparent visual connection. As Forgy[1] points out, for rules languages in general, “[they] have another property that makes them particularly attractive for constructing large programs: they do not require the developer to specify in minute detail exactly how the various parts of the program will interact”. This property leads to very large and difficult to reason about collections of implicitly connected rules.

Reasoning over a small number of rules is already surprisingly hard. Our host organization has many rules and, thus, reasoning about them is particularly challenging.

One approach to tackle comprehensibility could be to consider Miller's Law[2]. This law states that an average human can hold in his short-term memory 5-9 objects, which is often an argument for succinct code. The argument is that the developer must store anything that is not immediately in her vision in her memory. With it being impractical to reason about code that she cannot recall, the fewer relevant items to her reasoning that are out of view, the easier it is to reason about the code.

We have observed the difficulty developers have to try to reason about and edit collections of Drools files. We hypothesize that when we present developers with different views on their code, they can better understand it. The problem we wish to solve - how to improve the ability to reason about sizeable collections of Drools rules - we believe lends itself to the technique of projectional editing. By using projections to improve feedback whilst coding, we believe that this can reduce the representation impedance mismatch that hampers the developer's reasoning.

Our intention is not to override the language engineers who have spent many years developing this language and its ecosystem. In contrast, it is to augment the current developer experience. Thus, the problem considered in this thesis is how to present large

sets of rules to a developer to interact with them easily.

1.2 Research Questions

To reason about a large code base of rules engine code effectively, a different presentation is needed. This presentation should allow a more precise organization whilst remaining interactive. We can formulate the following research questions based on the discussion in the preceding sections being:

- **Main research question:** “How can projectional editors and DSLs be combined to address feedback mechanisms for developers in the context of reasoning about rules in a rule-based business engine?”

This question requires knowing if it is possible with current tooling. Thus we would like to answer the question:

- **RQ 1:** “What is the current state of language workbenches supporting projectional editing?”

Finally, we specifically would like to know how to improve the ability to reason about the business rules engine, so we ask the question:

- **RQ 2:** “Which projections can help developers get appropriate feedback about rules?”

1.3 Contributions

This thesis proposes a representation of business rules in a concise and readable format that could solve comprehensibility issues resulting from large codebases of business rules. The implementation behind the approach relies on language engineering and projectional editing. We developed a stand-alone opensource solution on a limited implementation of Drools. The underlying Drools implementation can be used as a base language for model-to-model generation by the MPS ecosystem.

1.4 Project Context

Khonraad Software Engineering, a subsidiary of Visma, hosted this investigation. Khonraad provides mission-critical services focussed on the automation of workflows at the cross-section of local government and healthcare. Specifically, Khonraad facilitates the mental health care and coercion laws in the Netherlands - WVGGZ, WZD, and WTH - which allow agencies to intervene in domestic violence, psychiatric disorders, and illnesses.

Khonraad's system facilitates reporting and communication between municipalities, police, judiciary, lawyers, mental health care, and many social care institutions. The system has 15,000 users and is available 24/7.

Configuration and administration require complex matrices of compliance mechanisms, access user rights and communication settings. The sensitivity of the personal data, being both medical and criminal, means security is of utmost importance. The security against data loss, preventing unlawful disclosure and guaranteeing availability, especially during crises, is crucial. Demonstration of the correctness of the configuration is a significant concern in the company.

This work environment allows us to work on an existing project, where success will impact the lives of those in critical need. Khonraad has its implementations in the Drools language that have evolved together with the iterations of the laws. Over the years, the evolution of the codebase means that we deal with real-life issues and not just thought experiments.

1.5 Thesis Outline

We start in chapter 2 with the required background information on rules engines and projectional editing. In chapter 3, we present our strategies for answering the research questions. Chapter 4 presents the results of our various approaches. In chapter 5, we discuss both the validity of the work and the implications of the findings. We then look into related work in chapter 6. Finally, we present the conclusions in chapter 7.

Chapter 2

Background

This chapter gives the background information required on rules engines and projectional editing. It presents the specific case of the rules engine that we will be using for our investigation: Drools. Next, we delve into the world of projectional editing. Finally, it presents the specific projectional editing tool we will be using: JetBrains MPS.

2.1 RulesEngines

2.1.1 What is a Rules Engine?

In this section, we will describe what a rules engine is and a little of its history.

The Aristotelian doctrine of essentialism declares that a thing has essential properties and properties that are accidental. If one takes away accidental properties, then the thing remains the thing. If one takes away essential properties, the thing is no longer the thing. If the thing is a business application, then its essential properties are its business rules.

Simply put, business rules are the principles or regulations by which an organization carries out the tasks needed to achieve its goals. When adequately defined, it is possible to encode these rules into statements that define or constrain some business organizational behaviour. A rule consists of a condition and an action. When the condition is satisfied, then the action is performed. More formally, business rules are the implication in the logical principle of Modus Ponens.

When described like this, one could imagine this as just the “if-then” logic frequently used in traditional programming. One would not be wrong. However, in traditional programming, representing all the combinatorial outcomes can become complex. In the typical application architecture, developers tend to distribute rules throughout the source code or database. Each additional rule leads to more fragility.

We find descriptions of these rules in the design documentation or user manuals. However, as applications evolve, documentation gets out of sync with the codebase. Once desynchronization occurs, to know the rules governing the application, one has to navigate the codebase and decode the rules from often scattered locations.

A rules engine is also known as a Business Rules Engine, a Business Rules Management System or a Production Rules System. The goal of a rules engine is to abstract business rules into encoded and packaged logic that defines the tasks of an organization with the

accompanying tools that evaluate and execute these rules. Simply put, they are where we evaluate our rules. Rules engines match rules against facts and infer conclusions. Returning to the Modus Ponens comparison:

$$\begin{array}{c} p \\ p \rightarrow q \\ \therefore q \end{array}$$

If the premise p holds and the implication $p \rightarrow q$ holds, then the conclusion q holds. In terms of a rule engine and business rules, this would be:

1. the rules engine gathers the data for the premise: p
2. it examines the business rules as the implications: $p \rightarrow q$
3. it executes the conclusion: q

Rules engines follow the recognize-act cycle. First, the match, i.e. are there any rules with a true condition? Next, they carry out conflict resolution, pick the most relevant matching rules. They then perform the actions described in the rule. Then back to the matching step. If they make no more matches, they terminate the cycle.

Rules Engines are declarative, focussing on the what of the rules, not the how of the execution. Date[3] describes a rules engine role as “to specify business process declaratively, via business rules and get the system to compile those rules into the necessary procedural (and executable) code.” Fowler[4] describes a rules engine as follows: “ ... providing an alternative computational model. Instead of the usual imperative model, which consists of commands in sequence with conditionals and loops, a rules engine is based on a Production Rule System. This is a set of production rules, each of which has a condition and an action ...”.

Some of the advantages of using a rules engine include:

- The separation of knowledge from its implementation logic.
- The externalization of business logic.
- Rules can be human-readable.

In summary, a rules engine is the executor of a rules-based program, consisting of discreet declarative rules which model a part of the business domain.

Rule engines arose from the expert systems of the late 70s and early 80s. Expert systems initially had three primary techniques for knowledge representation: Rules, frames and logic[5]. “The granddaddy” of the expert systems, MYCIN, relied heavily on rules-based knowledge representation[6] rather than long inference chains. MYCIN was used to identify bacteria and recommend antibiotic prescriptions. MYCIN and its progenitor, DENDRAL, spawned a whole family of Clinical Decision Support Systems that pushed the rules engine technology until the early 1980s. Research into rules engines died out in the 1980s as it fell out of fashion.

Early in their existence, the rules engines hit a limiting factor because the matching algorithms they used suffered from the utility problem, i.e. the match cost increased linearly with the number of examined rules. Charles Forgy’s efficient pattern matching Rete

| Product | | Developer | licence type |
|------------------------------|------|--------------|--------------|
| CLIPS | [7] | NASA | open source |
| Drools | [8] | JBoss/RedHat | open source |
| BizTalk Business Rule Engine | [9] | Microsoft | proprietary |
| WebSphere ILOG JRules | [10] | IBM | proprietary |
| OpenRules | [11] | OpenRules | open source |

TABLE 2.1: Rules engine products

algorithm[1] and its successors solved this problem. This algorithm works by modelling the rules as a network of nodes where each node type works as a filter. A fact flows through the filters of this network. The pre-calculation of this network is what provides the performance characteristics.

The first popular rules engine was Office Production System (OPS) from 1976. In 1981 OPS5 added the Rete algorithm. Currently, there are a few rules engines in use. We show some of the more commonly used ones in table 2.1.

2.1.2 What is Drools?

JBoss Rules, or more commonly known, Drools, is Java's leading opensource rules engine. In this paper, when we use the name "Drools" we are referring to the "Drools Expert" which is the rule engine module of the Drools Suite. Drools started in 2001 but rose to prominence with its 2005 2.0 release. It is an advanced inference engine using an enhanced version of the Rete algorithm, called ReteOO[12], adapted to an object-oriented interface specifically for Java. Designed to accept pluggable language implementations, it can also work with Python and .Net. It is considered one of the most developed and supported rules platforms.

To execute rules, Drools has four major components, as demonstrated in figure 2.1. The production memory contains the rules, and this will not change during an analysis session. The rules are the focus of this thesis, and therefore, we will delve into much more detail later on these.

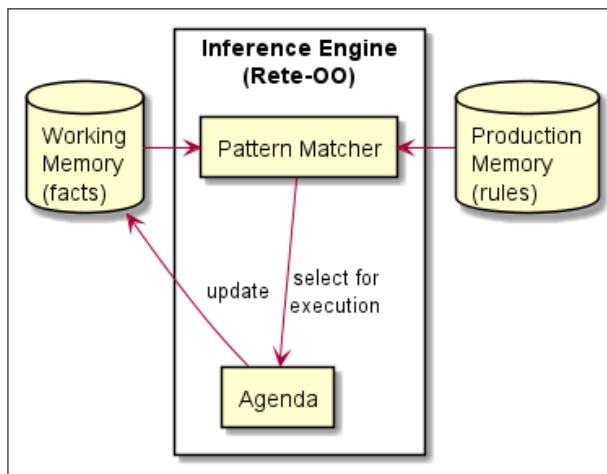


FIGURE 2.1: Drools components.

In Forgy's[1] overview of a rete algorithm, the following steps occur.

1. Match: Evaluate the LHSs of the productions to determine which are satisfied given the current contents of working memory.
2. Conflict resolution: Select one production with a satisfied LHS; halt the interpreter if no productions have satisfied LHSs.
3. Act: Perform the actions in the RHS of the selected production.
4. Re-evaluate: Go To 1.

Figure 2.2 show more detail of how these components interact within Drools to infer a conclusion. First, Drools asserts facts in the working memory. The working memory contains the current state of the facts, which triggers the inference engine. Using the aforementioned Rete-OO algorithm, the pattern matcher will examine both the working memory and a representation of the rules from the production memory to determine which rules are true. Drools will then put the rules that match on the agenda. It can be the case that many rules are concurrently true for the same fact assertion. These rules conflict. A conflict resolution strategy will decide which rule will fire in which order from the agenda. The first rule on the agenda will fire. If the rule modifies, retracts or asserts a fact, then the inference loop begins again. We have inferred our conclusion if either a rule specifies to halt or no matching rules remain on the agenda.

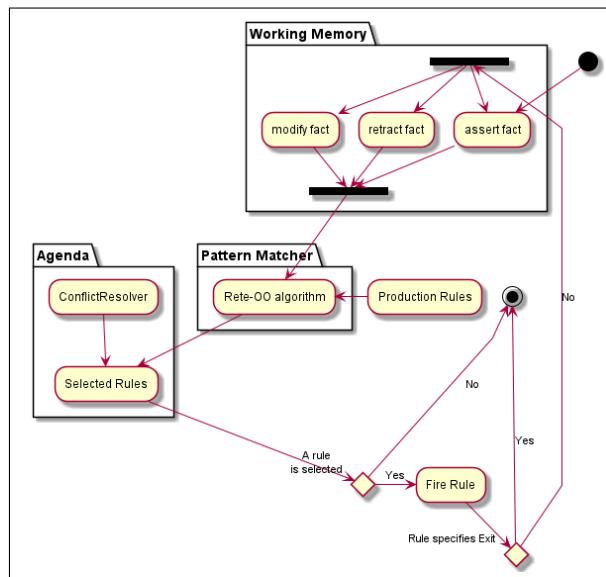


FIGURE 2.2: Drools inference loop

The component we will be focussing on in this paper is the rules. A rules file containing the rules is a text file, typically with a .drl extension. As the rules do not change during execution During execution, the rules do not change, and Drools stores them in production memory.

We do not need to examine the rule file components package, import, global, declare, function and query. We will examine the anatomy of a rule.

A rule consists of 3 parts: attributes, conditions, and consequences. Attributes are optional hints to the inference engine as to how to examine a rule. The conditional, "when",

or left-hand side (LHS) of the rule statement is a block of conditions that have to, in aggregate, return true for the asserted fact. If true, then the rule is placed on the agenda. The actions, consequences, “then”, or right-hand side (RHS) of the rule statement contains actions to be executed when the rule is selected.

The LHS is a predicate statement made up of some patterns. The patterns evaluate facts from the working memory. The pattern can match against the existence of facts or facts with matching property conditions. Connectives, such as not, and, and or can combine patterns. The patterns apply to individual facts rather than the group, thus can be seen as first-order predicates.

Variables can be bound to facts that match these patterns for use later in the LHS or for updating the working memory on the RHS.

Drools offers more options for the LHS. We have limited the scope of this paper to the features described thus far.

The RHS can contain arbitrary code that will execute when a rule is selected. However, its primary purpose is to adjust the state of truth in the working memory. One can insert, modify, and retract facts in the working memory. Modifying and retracting facts must be done on fact variable references bound in the LHS. One can explicitly terminate the inference loop with a halt command.

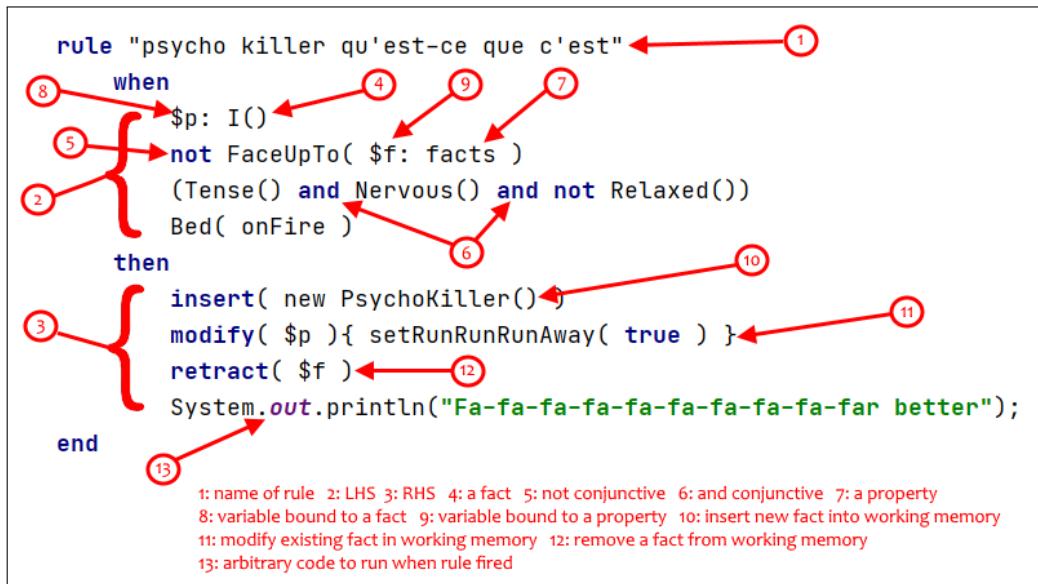


FIGURE 2.3: Drools rule breakdown

An Explanatory Example

Listing 2.1 shows an example of a .drl file taken from the Drools sample code.

```

1  package org.drools.examples.honestpolitician
2
3  import org.drools.examples.honestpolitician.Politician;
4  import org.drools.examples.honest polititian.Hope;
5
6  rule "We have an honest Politician"
7      salience 10
8      when
9          exists( Politician( honest == true ) )
10     then
11         insertLogical( new Hope() );
12     end
13
14     rule "Hope Lives"
15         salience 10
16         when
17             exists( Hope() )
18         then
19             System.out.println("Hurrah!!! Democracy Lives");
20         end
21
22     rule "Hope is Dead"
23         when
24             not( Hope() )
25         then
26             System.out.println( "We are all Doomed!!! Democracy is Dead" );
27         end
28
29     rule "Corrupt the Honest"
30         when
31             $p : Politician( honest == true )
32             exists( Hope() )
33         then
34             System.out.println( "I'm an evil corporation and I have corrupted " + $p.getName
35             () );
36             modify( $p ) {
37                 setHonest( false )
38             }
39         end

```

LISTING 2.1: Example Drools file.

Listing 2.1 gives the Drools engine instructions on what actions to take when something changes the working memory. This toy example reacts to when the working memory has an honest politician added. It prints a message celebrating the existence of said politician. It then corrupts her and gloats in a message. Finally, it prints a message of despair. The code in Listing 2.1 does the following:

1. On line 1, the package statement identifies the rule file.
2. On lines 3 and 4, the import statements describe which facts are available for use.
3. The “We have an honest Politician” rule on line 6 does the following:
 - (a) On line 7, using the salience attribute, the rule is set to be run before other rules with a lower salience.
 - (b) On line 10, the rule checks working memory for Politician facts with the honest property equal to true.
 - (c) On line 12, if found, a Hope fact is inserted into the working memory.
4. The “Hope Lives” rule on line 15 does the following:

- (a) Line 18 check if any Hope facts exist.
 - (b) On line 20, if found, it prints a message.
5. The “Hope is Dead” rule on line 23 does the following:
 - (a) On line 25, it checks that no Hope facts exist.
 - (b) On line 27, if it finds no facts, then it prints a message.
 6. the “Corrupt the Honest” rule on line 30 does the following:
 - (a) Line 32 checks for any Politician facts with the honest property equal to true and sets them to the variable \$p.
 - (b) Line 33 checks if any Hope facts exist.
 - (c) If both Hope and Politician facts are found, on line 35, it prints a message including the \$p variables name.
 - (d) On lines 36 to 38, it modifies the fact in working memory represented by \$p to change its honest property.

2.2 Projectional Editing

2.2.1 What is Projectional Editing?

When talking about projectional editing, we are mostly talking in the domain of Metaprogramming. Usually, when we talk about software development, the programmer or developer creates the program and the user who uses it. In metaprogramming, we are talking about the development of languages. Here the developer could refer to both the creator and the user. We will distinguish these two roles in this paper by referring to the creators of the languages as Language Engineers. Thus when we refer to developers in this paper, we mean users of the language.

Traditionally developers write code with text editors or integrated development environments (IDE), which adjust the concrete syntax and allows a parser to create the abstract syntax tree. A projectional editor, Inverts this relationship, as a developer edits the abstract syntax tree and allows the IDE to project the concrete syntax.

Parser Based Editing

A program is defined using text and edited with a text editor in a traditional parser-based development workflow. A grammar is a definition of a programming language's formal syntactical rules or concrete syntax. One derives the lexer and parser from the grammar. The lexer will turn text passed into a text buffer into tokens. A parser then validates that these tokens, the words of the language, are syntactically correct. The parser then constructs a concrete syntax tree and an abstract syntax tree (AST).

An AST is a tree structure that represents the semantic meaning of the source code, stripped of all the syntactic details. The parser will carry out some of the name resolutions needed to ensure that the tree represents the references expressed within the source code. These references turn the tree into a graph.

Compilers use the AST to do subsequent processing, such as linking, transformation, analysis, and type checking. Modern IDEs, in the background, also parse the code it is displaying to create an AST to offer relevant coding assistance. This assistance is appreciated as without IDE help learning the concrete syntax of non-trivial languages is error-prone. Exploratory programming is laborious if one has to wait until compilation to discover mistakes.

Projectional Definition

In the projectional editing paradigm, a semantic model represents the program. This model requires projectional editing tools to be read and edited.

A projectional editor does not parse any text. In its place, a developer reads and edits a representation of the AST through a projected notation. Her editing gestures immediately and directly manipulate the AST. This editing takes place within predefined and fixed templates called editors.

The principle of projectional editing is familiar to those that use visual programming, like Scratch or Blockly, or graphical modelling tools, such as MetaEdit+. These tools do not parse pixels to generate their AST. Instead, they project the underlying models/programs in a view. They store the model/AST in a custom format rather than its plain text equivalent in a traditional programming language.

Projectional editing is the generalisation of this idea, with the ability to render multiple representations of the program with a wide range of notation styles.

The projection may sometimes seem like a text editor. However, this is just acrobatics by the language engineer designing an editor to help developers from traditional text-based languages feel comfortable. The text is just another type of projection of the AST. It also may be any other notation that can represent the semantic meaning of the code, such as formulas, graphs, or images. Projections are not just the notation but also how the user interacts with the projection. In this sense, the definition of the projections and the IDE/UI overlap.

2.2.2 What is it Not?

Projectional editing does not have clearly defined boundaries. In this paper, we exclude the following types of tools that sometimes get associated with projectional editing.

A Venn diagram of Model-Based Software Engineering (MBSE) and projectional editing would have a significant overlap. Here we will not be looking at tools that build code from UML or other MBSE or Model-Driven Engineering (MDE) tools.

Another area mistakenly grouped with projectional editing is Low-code software development environments. These, however, are only tangentially related.

Most confusing is “projectional editing” when referring to a methodology of product line differentiation in code bases. In addition to having the same name, one of the top products for this product line technique is called PEoPL. PEoPL uses MPS for development. Thus, this product is a projectional editor (the paradigm) for product line projectional editing (the methodology).

How Projectional Editing Works

As shown in figure 2.4, a projectional editor has a model or an AST. It renders a presentation of the model as a projection. The developer performs actions on the projection. Every user editing action maps directly to a change in the AST.

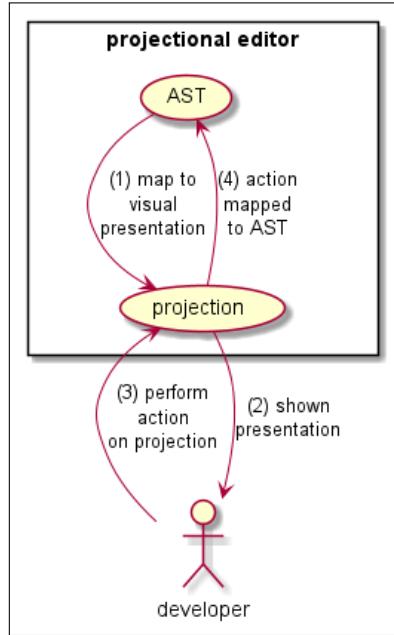


FIGURE 2.4: Projectional editing loop. TODO: comparison with Parsing

To perform the above two things have to be defined by language engineers: The Meta-Model and the editor.

The meta-model, analogous to the abstract syntax, describes the node concepts and connections used to build the hierarchical structure that is the AST. This hierarchy can have references to nodes in other branches, so it is a graph although named a tree. The AST is stored independently of the concrete syntax, often using a database, XML or a proprietary file format. Rules of the meta-model, such as type systems or scoping rules, must also be described.

Projectional editors avoid the grammars and parsers that define the concrete and abstract syntax in a traditional text-based language. In text-based languages, parsing transforms the concrete into the abstract. In projectional editing, the abstract is transformed to concrete using a projection engine that uses projection rules.

Editors combine the projection rules and the gestures or actions to create a change request to the AST. They are analogous to a concrete syntax.

One of these actions can be typing text. However, every string is recognised when entered. Thus, there is no tokenising. Text enters into the templates defined by the editor, and a newly derived projection displays the adjusted underlying AST to the developer.

The projection uses graphical elements to represent the representation. Although often appearing textual, each of the text elements are references to nodes in the AST.

Developers can only interact with the editor via the rigidly controlled code completion menus or gestures and actions. She builds the AST directly from each interaction she has

with the editor.

Nodes are instances of the concepts defined in the meta-model. Each node has a unique id and points to its defining concept. It is unambiguous. References are first-class and defined by the id rather than resolved by name, as in parser-based languages. Disambiguation happens at the time of input, as the developer chooses from limited legal inputs.

The separation of the abstract and concrete allows the language engineer to implement multiple projections of the same model, using different notations, each node of the AST taking having the design she envisions. The pattern used for projection is similar to MVC, so multiple views of the program can be visible and updateable simultaneously.

Graphical modelling tools, for example, tools for UML modelling, could be seen as specialised implementations of projectional editing. These modelling tools do not store pictures of the UML diagrams and then parse them to create an AST. Instead, they store the model, often with extra information about the visual layout, and the image of the UML is projected to the modeller to edit. Projectional editing generalises this approach to projecting any notation defined by the language engineer.

2.2.3 History of Projectional Editing

Here follows an incomplete and inconsistent history of projectional languages.

In the '70s and early '80s, researchers created several applications for research into the realm of structured editors. Some examples were: MENTOR[13], Incremental Programming Environment[14], GANDALF[15], Cornell Program Synthesizer[16], and Synthesizer Generator[17]. These language-based program editors could force syntactically correct programs through the knowledge of the language. These were the precursors to the modern projectional editors. They worked by providing templates for each abstract computational unit of the language. First, one would choose the concept and then fill out the placeholders.

These tools were not good at editing textual notations, which led to a poor user experience. When they attempted to fix this, for example, in the Synthesizer Generator, they reintroduced parsing to parts, which took away many advantages of the AST's direct editing.

In the late '90s and early '00s, the first forays into commercialising a more generalised version of structured editors, the projectional editors, began. First, Intentional Domain Workbench, inspired by Charles Simonyi's 1995 essay "The Death of Computer Languages, The Birth of Intentional Programming"[18] the IDW was the product of the company Simonyi founded in 2002, Intentional Software. The Intentional Programming paradigm spotlighted the projectional editing domain, taking it out of the universities and into practice. However, as it was a closed sourced and expensive product, thus not many papers were written about it. In 2017, as part of an "acquisition", Microsoft bought Intentional Software for its employees and let the product die.

Inspired by a call to action for language orientated programming[19] JetBrains embarked in 2004 on a mission to build a product to fulfil that ideal. Meta Programming System (MPS) was the outcome of that journey. Language engineers created the languages

mbeddr, PEoPL, and Realaxy using the MPS platform. It currently has an active community of developers and projects both in academia and in the commercial world. We chose this tool to be the basis of our projectional editing experiments. We will talk about it at greater length in section 2.3.

The last decade has produced a few smaller projectional workbenches. There are a few open-source, small team projectional projects. In 2013 several projectional language workbenches joined MPS in the Language Workbench challenge[20]. These included Más, a web-based projectional editor, which is no longer with us[21]. Whole Platform[22] is a projectional language workbench plug-n for Eclipse. Cedalion[23] provides another projectional IDE, specialising in internal DSLs.

More recently, there have been some new products that intersect the projectional domain. Deuce[24] and Gentleman[108] are two recent projection editors that have recently emerged from academia. The final two mentions in our incomplete history are a little out of left-field. Google’s Blockly[25] is a tool for making structural editor languages, but only in a block format. Blockly can create languages similar in style to the scratch language. Blueprint visual scripting, a part of the Unreal Engine, is a visual programming language for building concepts such as levels or game assets. Examples of Blockly and Blueprint can be seen in figures 2.5a and 2.5b respectively.

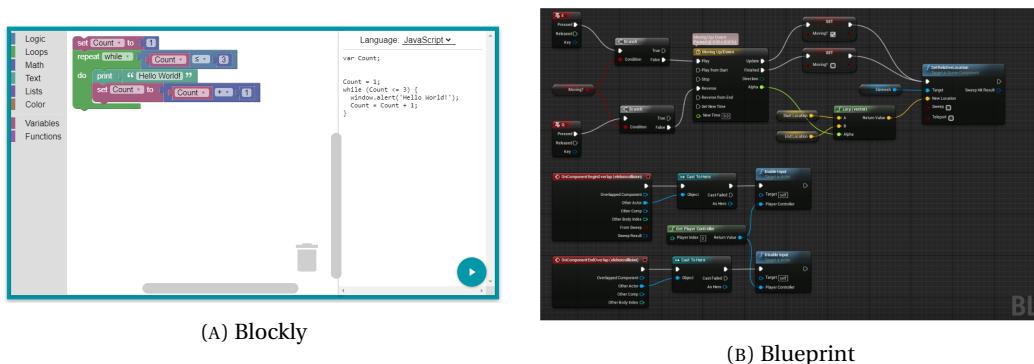


FIGURE 2.5: Leftfield projectional editors

2.2.4 What Advantages Does Projectional Editing Bring?

Projectional editing gives advantages both to the language engineer and the program developers. There is a lot of crossover and repetition between papers written on projectional editing regarding its advantages. To that end, what follows is a synthesis of several papers as to the advantages that projectional editing claims. Rather than attributing each advantage to each paper, we have made a reference table of papers proclaiming said advantage. This is table 2.2.

Exploratory programming As with their progenitors, syntax-directed editors, modern projectional editors help guide a developer unfamiliar with a language. With their rigid syntax and predefined layout, the editors only allow editing within specific cells of the editor. This template style means that the developer does not have to worry about the significance of spacing or indentation. Minutiae of syntactic adornments, such as statement ending semi-colons or enclosing matched brackets, are also not interfering with her exploration of the language space.

| Advantage | # | Paper(s) |
|-----------------------------|----|--|
| Exploratory programming | 5 | [26–30] |
| Correctness-by-construction | 7 | [26, 31–35] |
| Rich notation | 22 | [18, 26, 28, 29, 34–51] |
| Mixed notation | 8 | [27–29, 37–41] |
| Multiple views | 9 | [26, 28, 35, 37, 38, 40, 42, 46] |
| Language composition | 23 | [18, 28, 29, 31, 32, 34, 35, 37–41, 46, 47, 50–58] |
| IDE functionality | 3 | [26, 38, 40] |
| Language evolution | 1 | [59] |
| Ancillary data | 5 | [28, 38, 51, 57, 58] |

TABLE 2.2: Papers describing advantages

When creating code, the editor only presents the developer with legal options within the current context. As the projection is context-aware, with relevant actions and options suggested and irrelevant ones removed. Thus, it is easier for the developer to explore what the language allows her to choose. Intelligent code completion does not have to be limited to single nodes. Inserting whole subtrees allows the developer to explore the larger structures of the language.

Correctness-by-construction A projectional editor prevents her from writing syntactically incorrect code by controlling the interaction between the developer and the AST. The whole class of syntactical errors is made impossible, with the developer relieved to think about special characters and layout. Typing and scoping errors are removed by only allowing validly typed and scoped options for the developer.

The developer can only select statements that are legal in the context of the location within the AST. Code does not have to be disambiguated, as this happens at the time of entry by the developer. If multiple items share the same notation in the editor, the developer chooses the relevant item, thus resolving the ambiguity to what she means rather than what the parser thinks she means.

Rich notation Constraints associated with textual parsing do not affect the choice of created projections. This freedom opens up diverse otherwise tricky or impossible to parse notations. Examples include tabular, mathematical expressions and symbols, diagrams, trees, images, forms, prose, sub- and superscript. Any visual form or shape that can map onto the AST can represent the program in an editor.

With these notations, one can better reflect the semantics of the program domain, which should aid comprehension. Mathematics has a rich history of use of notation. When writing a DSL for the Mathematics domain, the domain experts can interact with it in the centuries-old language of their domain.

Of course, the projections can also be projections of text. Textual projections are often the appropriate projection type if the developer interacting with the language's domain expertise is parser-based languages.

Mixed notation As no parsing is required and ambiguity is not an issue for the underlying AST, it is straightforward to combine different forms of rich notation. With all notations working on the same editor infrastructure, embedding mathematic symbols within textual projections, within tables within graphical representations is a simple coding pattern.

Multiple views With the AST being the stored artefact rather than the notation, projectional editing allows the language engineer to define multiple views on the same model optimised for different tasks. In Software architecture, one presents different views to different stakeholders based on their interests. Similarly, projectional editing can present experts with various domain expertise views on the model that reflect their needs. A developer can switch between node projections within an enclosing larger projection to find the one that best suits their current task.

Because the architecture of a projectional editor follows the principles of model-view-controller, it is possible to have multiple simultaneous views of the model. These multiple views allow the developer to update a projection optimised for writing and immediately see its effect in a projection optimised for understanding.

Language composition Parser-based languages can support some modularisation and composition, but a projectional editor allows easy and extensive modular language extension and composition. This ability results from the disambiguation of the nodes of an AST at the time of entry. If two items with the same syntax are available at the same place, the user will choose the one they require, and therefore the node has an explicitly chosen meaning.

The composition of independently developed languages does not suffer from the syntactic or keyword clashes they would in two grammar defined languages. Because of the lack of ambiguity, every node referencing the concept that defines it, these languages, when put together, will not have structural or syntactic issues.

Language composition can involve extending an existing language or embedding other languages in a host language without modifying its definition. The ease of composition and extensions allows building more significant languages out of smaller modules.

IDE functionality Developers in mature languages are used to the functionality of mature IDEs. These functionalities include syntax highlighting, intelligent code completion or suggestion, and static analysis for errors and validation. As projectional languages store the AST rather than the concrete syntax, they require an IDE to edit. Because of this, when a language engineer designs the language, she also has to design the IDE.

A projection always knows its context because it comes from the AST. When the editor already knows the meaning of the node it represents, syntax highlighting is simple. Knowing its context makes it much simpler to suggest intelligent code completions.

Always having a complete AST makes it much easier to validate scope, typing and other hard to implement code validators.

Language evolution Parsing complicates the evolution of languages. For example, adding a new reserved word is difficult without breaking existing code. Extending a language with new capabilities and syntax in projectional editing is simple. If the change is syntactic, then the language engineer has to update an editor. If there is a semantic change, then the language engineer can write a migration in the language to transform a node of one concept to a different type, and the developer would have to run that migration on their code.

Ancillary data Data added to nodes can augment the AST. This data is helpful for tasks such as documentation, requirements traceability and product line feature dependencies.

2.2.5 What are the Disadvantages of Projectional Editing?

Whilst fewer papers proclaim the disadvantages of projectional editing, we repeated the approach of the previous section. Thus, we have synthesised the disadvantages from papers in the following sections and listed citations for these ideas in table 2.3.

We do not consider that the dearth of disadvantages discussed as evidence of projectional editing's superiority. Our best guess is that those who do not find projectional editing beneficial do not write papers about it.

| Disadvantage | # | Paper(s) |
|---|----|--|
| Low adoption | 4 | [35, 39, 48, 50] |
| Unnatural user experience | 11 | [29, 34, 35, 37, 38, 40, 46, 48, 50, 56, 59] |
| Ambiguous syntax | 1 | [41] |
| Inf flexibility | 2 | [29, 37] |
| lack of integration with text ecosystem | 5 | [29, 37, 48, 56] |
| Learning curve | 5 | [29, 38, 39, 47, 56, 60] |
| Vendor lockin | 2 | [38, 40, 61] |

TABLE 2.3: Papers describing projectional editing disadvantages

Lack of adoption The ideas that proceeded projectional editing - the structured or syntax-directed editor - have been around since the early 1970s yet have failed to be adopted widely. This argument is a bit of a tautological one, as the low adoption is perhaps an outcome of the other disadvantages of projectional editing. However, low adoption can lead to a self-reinforcing process, where lack of adoption prevents further adoption.

Inconvenient or unnatural editing Early attempts at projectional editing presented an inconvenient and unnatural user experience when coding. These usability challenges, exemplified by the tedious manner of entering code as per the tree's order, compare poorly to parser-based languages.

This lousy reputation continues, despite massive improvements in projectional editors. Whilst there is no debate that projectional editing feels different, some question whether this inconvenience is an intrinsic property or a result of developers, through years of experience, being used to text-based programming.

Modern projectional editors, when using a textual projection, face an “uncanny valley” issue. Whilst trying to simulate a text editor, the developers start to expect all of the functionality of the text-based IDEs. This expectation is an especially weak trait regarding granularity and restrictions of cursor movement, insertion, deletion, selection, copy and paste, and other interactions with the text.

Ambiguous syntax One of the selling points of projectional editing, especially concerning language composition, is that there can be no ambiguous syntax. While this may be true for the AST, it is not so for the developer who reads this code on the screen. If one combined Drools and Basic rather than Java, the developer might become confused about which language the “Then” keyword refers to when she reads it. Thus writing ambiguity is replaced by reading ambiguity.

Inflexibility A developer using a projectional editor has no flexibility in code layout. They may feel they require this for enhanced readability. The flexibility of the layout is entirely in the hands of the language engineer when she determines the projection rules.

Integration with the text-based world Projectional editors do not store the definition of the program in the form of a plain-text implementation in the concrete syntax. Instead, the AST is stored and serialised in a non-human readable format, such as XML.

This different format of program storage leads to an issue with integration with the text-based ecosystem. This ecosystem is extensive, as text-based coding has been popular since the 60s. Two notable examples are text diffing, especially where branch merging is concerned and code sharing. The diffing issue within projectional editing tools is solved. However, as code-bases often span multiple programming languages and tools, the difficulty of integrating projectional diffing into the software development workflows is still a real problem.

Textual source code can be shared simply by email or on websites. This sharing, however, is not easy with projectional code.

Learning curve For the language engineer, the necessity to develop an editor with a good user experience is much harder work than defining a grammar for a parsed language. The learning curve for the language engineer is significant, as, by default, she has to think also of the IDE development.

For the developer, especially one with an extensive text-based experience, the different editing style takes some getting used to.

Vendor lock-in The nature of projectional editing is that what one edits is a projection of the AST, and therefore an IDE is needed to do the projecting and language definition. The fear of getting locked into a specific concept implementation can negatively impact evaluations of projectional editing by organisations. To be able to use previously developed languages would require using the same toolset. Changing to a different toolset for language design would require a significant re-skilling effort.

2.3 What is MPS?

Language workbenches (LWB) are a tool to help language engineers create languages, particularly domain-specific languages (DSL). Fowler[62] popularised the term LWB in a 2005 article.

Meta Programming System (MPS) is an open-source LWB that assists in the creation of Projectional languages. It started in 2003 by JetBrains and was introduced to the world in Sergey Dmitriev's 2004 Paper "Language Orientated Programming: the Next Programming Paradigm"[19].

As discussed in section 2.2.1, When creating a projectional language, one has to define the language and how one interacts with it. In MPS, the language engineer defines languages, including their interactions. Developers create programs using these languages. The language engineer can extend languages. The developer can mix the languages she uses.

The following is an overview of how MPS implements the ideal of a projectional language. It is also the structure of this section:

- Abstract syntax
 - Structure
 - Behaviors
 - Constraints
 - Type system
- Concrete syntax
 - Editors
 - Intentions
- Generators
 - Model-to-Model
 - Model-to-Text

MPS defines the different aspects of the language definitions with small, declarative DSLs. These are bundled together into what they term Aspects.

2.3.1 Abstract Syntax: Structure

Structure is what determines the abstract syntax of a language. The most important item available in a Structure Aspect is the Concept. Instances of concepts are called nodes. With these nodes, the developers construct their programs. When referring to a program in MPS, we are talking about its stored abstract syntax tree (AST).

In principle, a concept contains three types of things:

1. Properties: these primitives are integer, boolean, string, or enum items and are similar to leaf values.

2. Children: these are other concepts, or collections of them, similar to subtrees.
3. References: these are relationships with other nodes in the AST. These turn the tree into a graph.

Concepts follow some object orientated (OO) traits, such as subtype, being abstract, and implementing interfaces.

One of these Concepts must be a root node. Otherwise, there is nowhere for a program to start.

Other items available in the Structure Aspect are the Interface Concept, the Enumeration, the Constrained Data Type, and the Primitive Datatype.¹

Thus, the structure aspect defines how the AST can be structured.

Figure 2.6 shows a concept with three children that implements two interfaces.

```
concept RuleStatement extends BaseConcept
    implements INamedConcept
        IFileLevelStatement

    instance can be root: false
    alias: rule
    short description: rule

    properties:
    << ... >>

    children:
    attributes : RuleAttributes[1]
    conditions : AbstractCondition[0..n]
    outcomes   : StatementList[1]

    references:
    << ... >>
```

FIGURE 2.6: Concept example

2.3.2 Abstract Syntax: Behaviors

OO design usually bundles together data and methods that can act on that data. Concepts are analogous to the data part of this equation. Behavior² fills the role of the methods in the OO analogy, defining the functionality called from instantiated nodes and static methods called from the Concept. The Constructor is a specialised method in a Behavior, filling the same role as a constructor in OO.

¹At the time of writing, we are unaware as to whether Data Type and Datatype are semantically different or if the different naming is a style choice.

²When referring to its use in the MPS workbench, consistent with their use, we use the American spelling - Behavior.

The methods have public, private, or protected visibility. If the Concept to which the behavior refers is abstract, the behavior itself can contain abstract methods. Abstraction, variable visibility and inheritance allow a sort of polymorphism. If a virtual method is declared, then it can be called polymorphically.

Figure 2.7 shows a constructor added to a Concept to initialise its children. It has a method to allow other nodes to interrogate the condition of it having attributes.

```
concept behavior RuleAttributes {

    constructor {
        this.salience = new node<SalienceAttribute>();
        this.salience.salience = new node<IntegerConstant>();
        this.salience.salience:IntegerConstant.value = 0;
        this.noloop = new node<NoLoopAttribute>();
    }

    public boolean hasAttributes() {
        return this.salience.visible || this.noloop.visible;
    }
}
```

FIGURE 2.7: Behavior example

2.3.3 Abstract Syntax: Constraints

A Constraints aspect adds further structural restrictions to a Concept. Constraints primarily define scope by controlling if another node can be a child, a parent, or an ancestor of this node. A Constraint can also prevent badly formed properties, children or references.

Figure 2.8 shows an example of a scope restraint that only allows local variables declared within the same rule or global variables declared in the same file.

2.3.4 Abstract Syntax: Type System

The Typesystem aspect and the constraints aspect together represent the static semantics of the language. This aspect is for the computation and evaluation of types of variables, expressions and statements.

Rules that are available to calculate and enforce the type system include inference, subtyping, comparison and substitute type rules.

Figure 2.9 shows an inference rule that ensures that the calculated type of the import statement matched that of its child called type.

2.3.5 Concrete Syntax: Editors

Editor aspects define the notation of the nodes. In effect, it is the user interface of the language, projecting the AST to the developer. An editor is a swing panel that renders a tree of editor cells. A concept can have multiple editors, thus offering multiple views on it.

```

concepts constraints RuleVariableRef {
    can be child <none>
    can be parent <none>
    can be ancestor <none>
    instance icon <none>
    <><property constraints>>

    link {target}
        referent set handler <none>
        scope (referenceNode, contextNode, containmentLink, position, linkTarget)->Scope {
            node<RuleStatement> rule = contextNode.ancestor<concept = RuleStatement, +>;
            nlist<RuleVariable> localVars = rule.descendants<concept = RuleVariable>;
            sequence<node<RuleVariable>> globalVars = rule.containingRoot.descendants<concept = GlobalStatement, +>;
            return ListScope.forNameElements(localVars.concat(globalVars));
        }
        <no presentation (deprecated)>

        default scope <no default scope>
    }
}

```

FIGURE 2.8: Constraint example

```

inference rule typeof_ImportStatement {
    applicable for concept = ImportStatement as importStatement
    applicable always
    overrides false

    do {
        typeof(importStatement) ::= typeof(importStatement.type);
    }
}

```

FIGURE 2.9: Typesystem example

The definition of the options available to the developers through menus also happens within the Editor Aspect. The choice the developer makes transforms the existing AST.

Additionally, the behaviour of interactions can be defined, such as what will happen to the AST when a particular keypress or editor action occurs at a particular location.

Figure 2.10 shows a component with a projection for the Concept shown in figure 2.6.

2.3.6 Concrete Syntax: Intentions

In projectional editing, the IDE is a part of the concrete syntax. Intentions make context-aware suggestions for automatic changes to the program to the developer. Figure 2.11 shows an intention that allows the developer to add, remove or edit a property based on its current value.

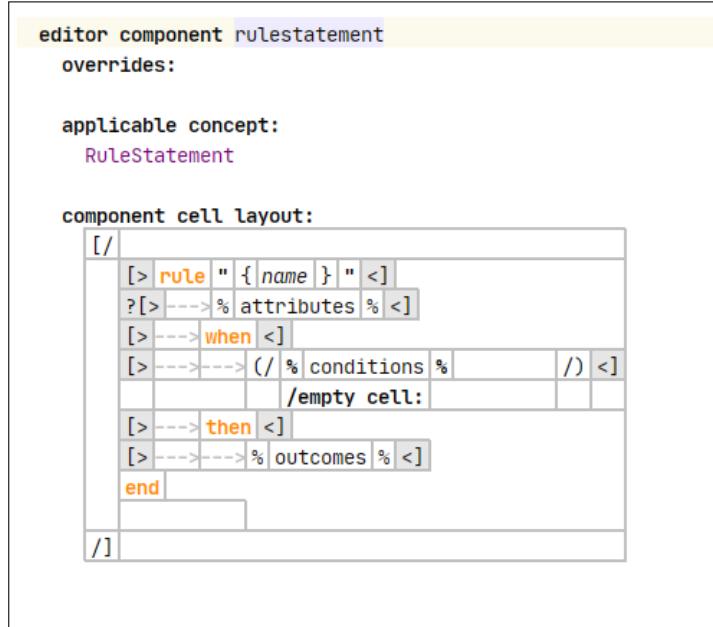


FIGURE 2.10: Editor example

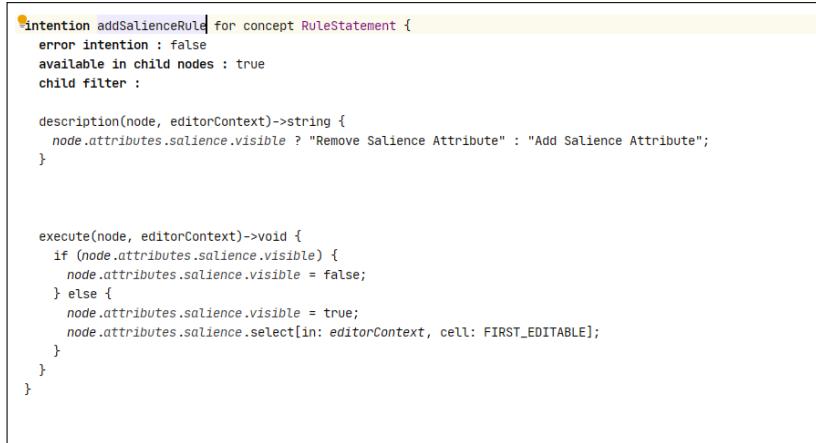


FIGURE 2.11: Intention example

2.3.7 Generators: Model-to-Text

Whilst designing a language is nice, it has to be able to do something. Without doing so, it has no semantic meaning. It is possible to create interpreters that can use the AST generated by MPS. However, the most common modality for MPS is to generate an output that gets compiled and run by commonly known environments. The output stage of generation is called TextGen. It defines how a node becomes runnable code in plain text.

2.3.8 Generators: Model-to-Model

Base level languages will have text generation. Most DSLs will perform model-to-model conversions, eventually converting to a base language. These intermediate stages are known as Generator Aspects. They transform code written in one language to another.

In MPS, a Concept can have multiple generators aimed at different base level languages, such as Java (using MPS BaseLanguage), C (using mbeddr) or XML.

Chapter 3

Methods

To answer the research questions from section 1.2, we formed three approaches. To answer the question “What is the current state of language workbenches supporting projectional editing?”, we carried out a Systematic Literature Review, which we describe in section 3.1. We followed two linked approaches for the question “Which projections can help developers get appropriate feedback about rules?”. Firstly, using Action Design Research (ADR), we implemented Drools as a projectional editor, and then some projections that users can edit. Then, we presented our findings to experienced Drools users through a questionnaire survey to validate this approach.

3.1 Method: Systematic Review

To answer the first Research Question, “what is the current state of Projectional Editing?”, we conducted a systematic literature review. Hereafter, we describe the method we undertook.

To carry out this review, we followed Kitchenham’s[63] advice on systematic review protocol validation (see appendix A for the exact checklist we used).

3.1.1 Motivation

The motivation that preceded this research was a requirement to understand if projectional editing was an idea that was worth investigating. Our background research showed an interest in the precursors to projectional editing in the late ’70s through to the mid-’80s. Outside of academia, interest arose in the mid-’90s following Charles Simonyi’s treatises on Intentional programming. However, Simonyi’s call to arms did not lead to a swell in academic research as his company’s product, Intentional Domain Workbench, was a closed commercial product. Conversely, after JetBrains’ OpenSource Meta Programming System (MPS) release, in the late 2000s, there was a flurry of papers on the subject.

Is there a need for a study of this topic? We believe, at least in the microcosm of this master’s project, it is helpful to know whether we are researching in an area that is dying or vibrant.

There does not seem to be any recent systematic reviews specifically about projectional editing. This study is not extending any previous systematic review. Although there exist literature surveys and mapping studies in some adjacent fields, we found no SLRs.

Thus we believe it may be helpful for those in the language engineering research community to bring together all current research about projectional editing in one place.

3.1.2 Research Question

In this paper, we only have one research question to synthesize the findings of scientific papers, which is “What is the current state of Projectional Editing?” We have broken this question into three sub-questions.

- **Sub Question 1** “Is there any current research in the area of projectional editing?”
- **Sub Question 2** “Which tools are currently being used for research?”
- **Sub Question 3** “What is the sentiment in papers currently discussing projectional editing?”

3.1.3 Search Strategy

The search process is automated as SLRs require a high level of completeness, which one cannot effectively achieve manually. Our first major decision was whether to engage in creating a quasi-gold standard as advised by Zhang[64]. Zhang noted that the ad-hoc nature of search strategies in SLRs has limitations. We executed a preliminary ad-hoc search to try and ascertain the extent of the research space. After satisfying that the research space was small enough, we decided against using the Quasi-Gold standard, which was overkill for our requirements.

The search terms we landed on were as follows:

‘‘PROJECTIONAL EDITING’’
OR
‘‘PROJECTIONAL EDITOR’’

We adjusted these search terms to fit the query syntax of the various search engines.

As most research search engines offer date ranges, we also used the date range to eliminate unnecessary papers at the automated search stage to save the effort of excluding them later. In our research question, we are specifically looking at the current state of projectional editing. A restriction of many research search engines is that they define date ranges in whole years. When designing our search strategy, it was near the beginning of 2021, and thus we feared that this would be too small a search space. Thus, we set our date range to be from the beginning of 2020 to the present. For the sake of reproducibility, we advise the removal of any papers after 31st July 2021.

We show the search engines we used in table 3.1.

| | |
|---------------------|--------------------|
| ACM digital library | Google Scholar |
| BASE | CORE |
| IEEE Xplore | ISI Web of Science |
| Microsoft Academic | Science.gov |
| Wiley InterScience | SCOPUS |
| Semantic Scholar | SpringerLink |

TABLE 3.1: Search engines used

Once we have filtered the automated search through the criteria of the selection stage, we will use that as our starting set for snowballing. We will do all our filtering before we do any quality assessments, as we feel that excluding papers from snowballing based on the quality of the primary study would artificially limit the network of potential papers. Our snowballing procedure shall follow the advice of Wohin[65]. Snowballing is a technique for finding related papers using the reference lists in our starting set and applying the same selection criteria.

Where possible, we will get the forward snowballing papers from the “cited by” functionality of Google Scholar. Because of the range of the search being “to present”, all papers that cite the target paper will fall within our criteria. For backward snowballing, we will manually filter the bibliography section of the selected papers, selecting any paper published in 2020 or 2021

After gathering all the papers from the forward and backwards snowballing, we will apply the selection criteria again. The snowballing process will recursively iterate until there are no new papers. The papers accepted in each iteration will form the basis for the following stage - the quality assessment of the primary studies.

After the final iteration, as a final step, the selected papers will have a deeper scan. This deep scan is to verify that our initial scan that the papers met our inclusion criteria before moving on to the quality assessment.

3.1.4 Study Selection

The inclusion criteria are:

- Studies are about or mention projectional editing or one of its synonyms.
- The study published date is in the period 2020-2021.

The exclusion criteria are:

- Books and grey literature.
- Not in English.
- Full text unavailable.
- Papers with severe issues with grammar or vocabulary.
- A duplicated paper.

- The primary study is in a previously selected paper.

If multiple papers look at the same study with different approaches, We aggregate the data during the synthesis stage.

As a lone researcher, we must be aware of bias in positively including relevant papers and excluding irrelevant papers. We will follow Kitchenham's suggestions to overcome such bias:

- Test-retest
 - We will assess the papers once (on title abstract and keywords) against the inclusion and exclusion criteria.
 - Save all the suggested results.
 - Assess the papers again three days later in a different order to the first.
- If there are disagreements, we will use Cohen's Kappa agreement statistic[66] to see if the process needs to be refined.

If our searches appear too large for a lone researcher, we will turn to text mining. We will be cautious about using this. O'Mara-Eves et al.'s systematic review of text mining in systematic reviews[67] recommends using this for prioritization but finds that for exclusion screening, although promising, it is not yet proven.

An SLR is interested in studies rather than papers. There is a many-to-many relation between papers and studies. We will review the selected papers to note when this has happened in our results to make sure studies do not get over or undercounted.

3.1.5 Quality of Primary Studies

To discover explanatory reasons for why there may be differences in study results and to weigh how valuable specific studies are, we will assess the quality of the selected studies.

To try and avoid a “Results Section bias” we will be operating a results-blind quality assessment. We base our study quality on the methods section of the papers only. However, this bias is still a threat because the abstract, which we will read, summarises the results. The study quality will not be measured until after the selection process is complete, though it will, in part, occur before the selection re-test process.

In evidence-based medicine, there is a well-known hierarchy of evidence quality for consideration in systematic literature review studies - the CRD Hierarchy of Evidence[68]. Kitchenham[69], suggests the hierarchy specifically for software engineering research, shown in table 3.2.

| Rank | Description |
|------|--|
| 1 | Evidence obtained from at least one properly designed randomized controlled trial. |
| 2 | Evidence obtained from well-designed pseudo-randomized controlled trials (i.e. non-random allocation to treatment). |
| 3-1 | Evidence obtained from comparative studies with concurrent controls and allocation not randomized, cohort studies, case-control studies or interrupted time series with a control group. |
| 3-2 | Evidence obtained from comparative studies with historical control, two or more single-arm studies, or interrupted time series without a parallel control group. |
| 4-1 | Evidence obtained from a randomized experiment performed in an artificial setting. |
| 4-2 | Evidence obtained from case series, either post-test or pre-test/post-test. |
| 4-3 | Evidence obtained from a quasi-random experiment performed in an artificial setting. |
| 5 | Evidence obtained from expert opinion based on theory or consensus. |

TABLE 3.2: Study design hierarchy for software engineering

These checklists address general scientific research. In software engineering, many studies that fall under what Gregor[70], in “A Taxonomy of Theory Types in Information Systems Research” calls “Type V: Theory for Design and Action”. These types of studies are also known as Action Design Research (ADR).

The checklists do not address this type of research well. On investigating how other SLRs conduct a quality assessment of ADR studies, we did not find a solution. Therefore we will continue with the checklists shown in the appendix, using the checklists meant for Case Studies for the ADR research papers. We will take this into account before dismissing results of this type based on their quality score.

As a lone researcher will carry out this study, there is no need to have a process for disagreements between researchers.

We use the quality assessment checklist to weed out the biases of selection, performance, detection, exclusion, and other threats to the validity of the studies under test. Validity issues can occur during the design, operation, analysis, or conclusion of an empirical study.

3.1.6 Data Extraction

No data extraction will be necessary for the first sub-question, “Is there current research in the area of projectional editing?”. The existence of papers with verified primary studies either into projectional editing theory or its practical use will be enough to answer the question.

For the question of “What tools are currently being used for research?”, we shall note each tool discussed mentioned explicitly in the study.

Finally, for the sentiment, we shall pass each paragraph of the introduction, the discussion, the conclusion, and any other sections that mention projectional editing or tools through a sentiment analyser, noting its sentiment score. The sentiment analysis tool we shall use is Microsoft Azure Cognitive Services Text Analytics. We show the code to carry out this task in listing 3.1.

We will gather this data in tables with the categories shown in table 3.3.

```

from azure.core.credentials import AzureKeyCredential
from azure.ai.textanalytics import TextAnalyticsClient

endpoint = "REPLACE_WITH_CORRECT_ENDPOINT"
key = "REPLACE_WITH_CORRECT_KEY"

text_analytics_client = TextAnalyticsClient(endpoint=endpoint, credential
                                            =AzureKeyCredential(key))

inputfiles = [[ARRAY_OF_FILES_TO_BE_ANALYSED]]

with open('/content/sample_data/sentiment/output_all.txt','a') as outf:
    for sections in inputfiles:
        for section in sections:
            print("Section: {}".format(section),file=outf)
            f = open('/content/sample_data/sentiment/'+section)
            content = f.readlines()
            # for brevity an optimization to deal with 10 document limit is
            # removed
            if len(content) != 0:
                result = text_analytics_client.analyze_sentiment(content,
                                                                show_opinion_mining=True)
                docs = [doc for doc in result if not doc.is_error]
                for idx, doc in enumerate(docs):
                    print("sentiment: {}".format(doc.sentiment),file=outf)
                    print("Document text: {}".format(content[idx]),file=outf)

```

LISTING 3.1: Text Analytics code.

| # | Data Type | Description | RQ |
|---|---------------------|--|--------|
| 1 | Study ID | Unique identifier for the study | |
| 2 | Title of Study | The paper name | |
| 3 | Year of Publication | Will be either 2020 or 2021 | |
| 4 | Author(s) Names | Including affiliation | |
| 5 | Source of Study | Name Of Online Database/ Digital Library | |
| 6 | Type of Study | Publication/Conference/Workshop/Symposium | |
| 7 | Name of Venue | Journal/Conference in which study has been published | |
| 8 | Tools in Study | A list of the tools used | RQ 1.2 |
| 9 | Sentiment | The sentiment scores from appropriate paragraphs | RQ 1.3 |

TABLE 3.3: Data extraction form

3.1.7 Data Aggregation and Synthesis

As explained by Kitchenham[71], in software engineering, primary studies will tend to be too heterogeneous for any statistical analysis. Synthesizing outcomes from multiple methods will be complex. Thus our synthesis will take a narrative approach.

Narrative synthesis tells a story of the who, how, and why of the success or otherwise of the research. For ADR research, the focus will be on what will help or hinder the adoption of the implementations. It will also examine how reliable the results are and the relationships between the studies.

3.2 Method: Action Design Research - Drools in MPS

Even though Drools is a relatively small DSL, we did not need to implement all the functionality to answer our questions.

3.2.1 Really Simple Rules Language

As we were new to DSL design and MPS, we first would create a simple approximation of the Drools language to create our first projections. We called this language “Really Simple Rules” (RSR).

File RSR, Like Drools itself, has a File as its root node. The File only contains Facts and Rules.

Fact and FactProperty In Drools, a fact represents a Java Bean with its child properties, which can also have their child properties, ad infinitum. In RSR, we limited properties to allow only boolean values. We decided this because fact selection is a predicate and thus can only return a boolean. By only allowing booleans, we also simplify the operations allowed on the property.

Rule We only simulated the Left Hand Side, or the “When” conditions, of a Drools Rule for the Rules Concept. We believed this would provide us with compelling options for projections and did not want to overcomplicate this first approach.

An RSR Rule consists of a collection of conditions. Should all those conditions return true, then the rule is selected.

Condition A condition operates on one or more FactSelectors. There are four condition types ExistsCondition, NotCondition, AndCondition, and OrCondition. ExistsCondition and NotCondition are unary conditions and evaluate one FactSelector. AndCondition and OrCondition conditions evaluate two FactSelectors.

FactSelector A FactSelector consists of a reference to a Fact and a collection of Predicates. If the Fact exists and all the predicates evaluate to true, then the FactSelector evaluates to true.

Predicate The predicate is an operation on a FactProperty, to which the Concept has a reference. Because FactProperty represents a boolean value, the only predicate operations are “Is” and “Not”.

Figure 3.1 shows the Concept hierarch for this straightforward implementation.

We realised this design in MPS. As the aim is to attempt different projections, we did not initially optimise for editing. The structure is as shown in figure 3.2, and the editors, including those shown in figure 3.3.

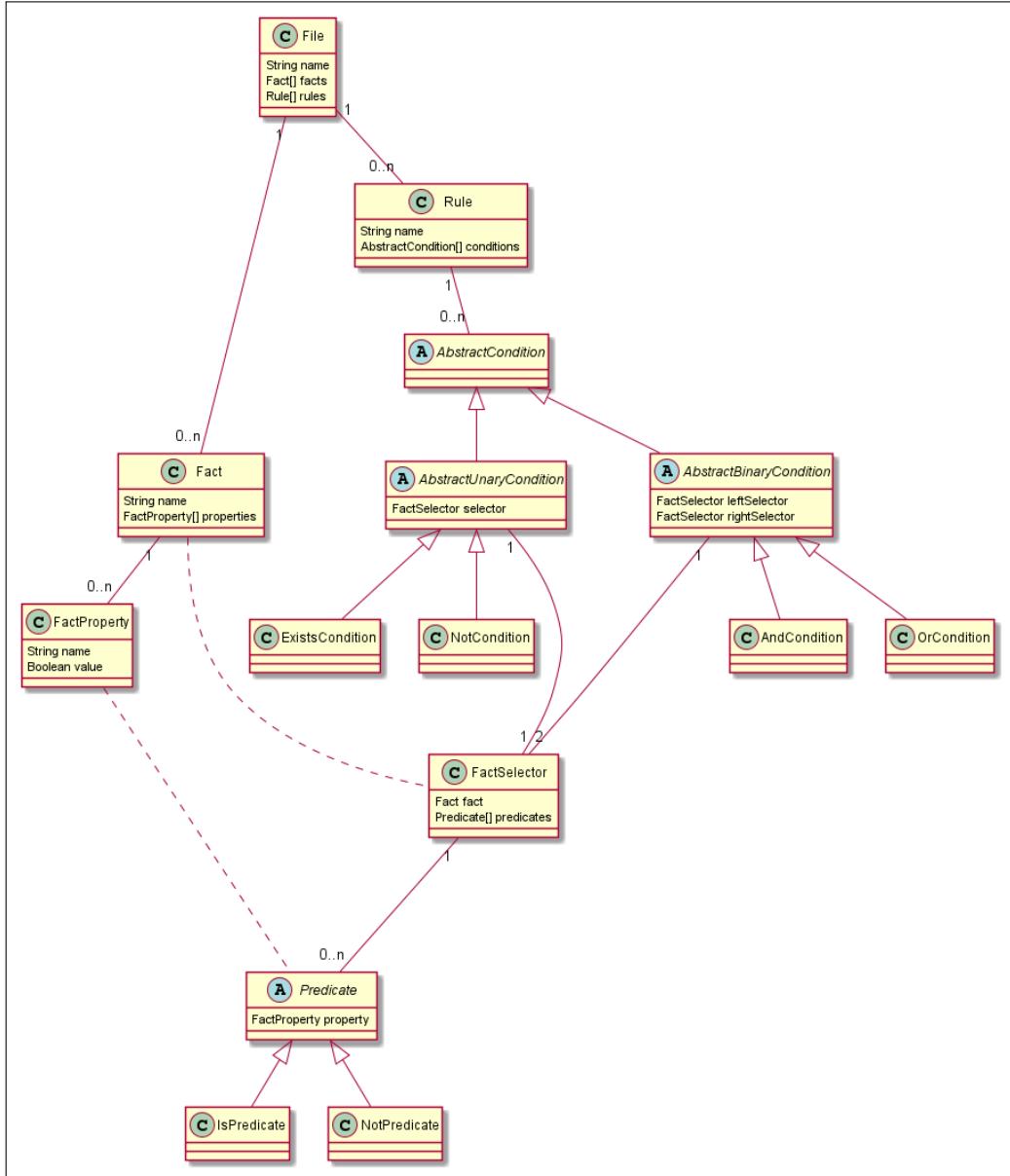


FIGURE 3.1: RSR concept hierarchy

Part of the research question is using projections for reasoning about large files. In order to answer this, we needed to simulate a large file. To do this, we had to enter a large number of rules. As this becomes tedious, we added some editing aids, including substitute menus, to speed up the entry of conditions, as shown in figure 3.4.

This image shows that we originally had to select an **ExistsCondition** Concept and select the Fact for the condition. After adding the substitute menu, We could immediately select the Fact we wanted, and the **ExistsCondition** would automatically wrap it with an **ExistsCondition** node.

We also added some intentions to invert incorrectly added conditions.

Finally, we added a Constraint to scope the fact properties in predicates to the Fact chosen in the FactSelector. This scope Constraint made it much easier to select properties in the predicates as indicated in figure 3.5.

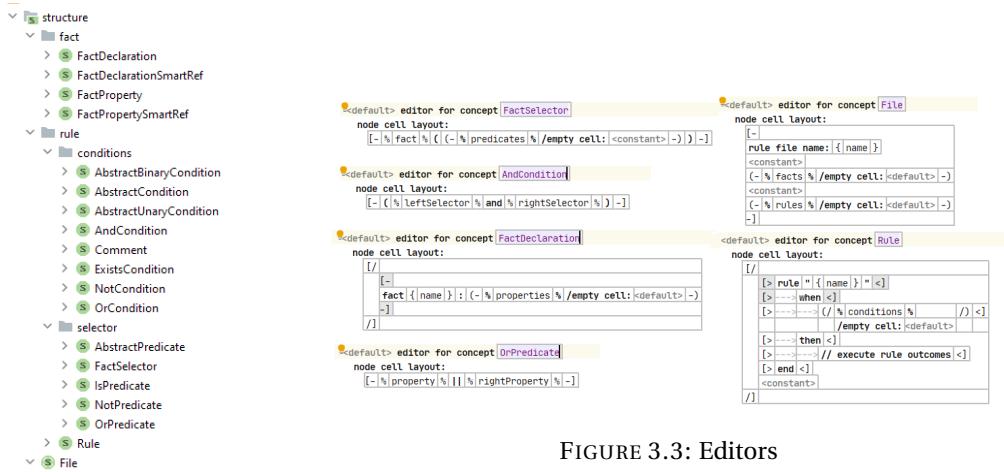


FIGURE 3.2:
RSR

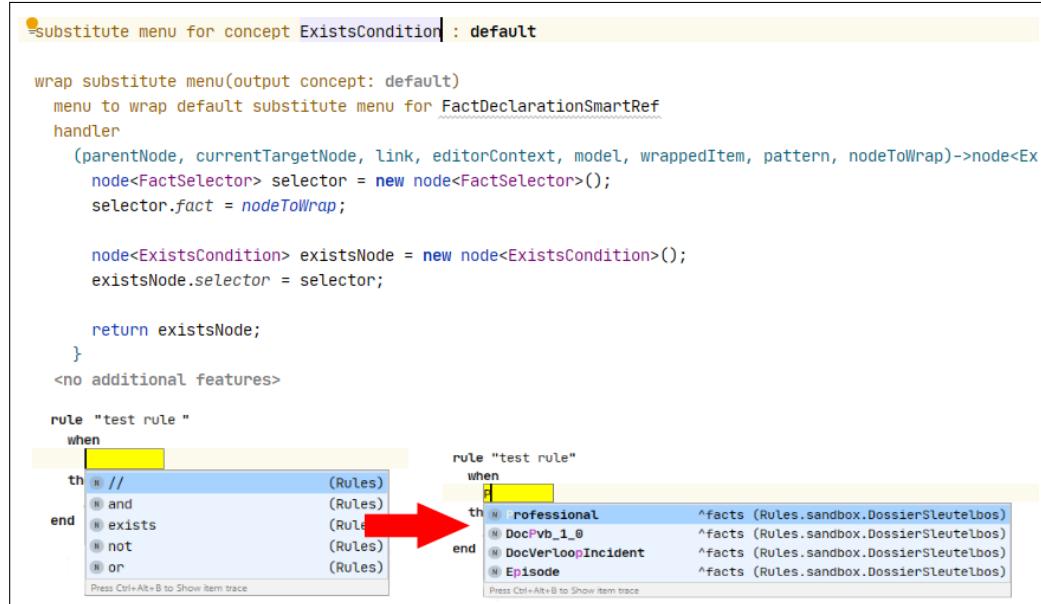


FIGURE 3.4: RSR substitute menu

The figure shows that before adding the scoping constraint, it showed a list with dozens of potential FactProperties, that represented all the FactProperties in the Model. After adding the constraint, it only shows the two properties associates with the Fact from the FactSelector.

Thus, we have described the entire implementation of the Really Simple Rules Language.

After implementing the language, we wrote a program with a large number of rules. This program on which we will experiment with the different projections. Figure 3.6 shows an example of our default Drools like text projection.

We discuss the alternative projections in the results section 4.2.

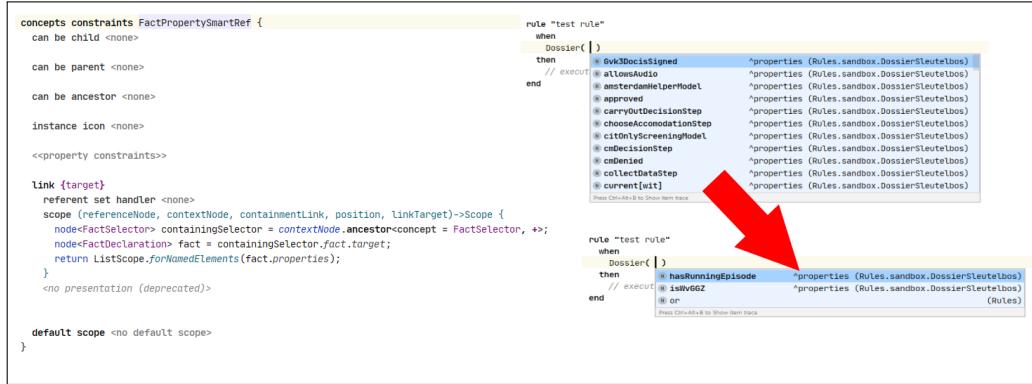


FIGURE 3.5: RSR scoping constraint

3.2.2 Drools-Lite Language

The RSR was useful as an initial language. However, it suffered from two significant issues. Firstly, its limitations as a language were so substantial that it could not handle many necessary scenarios. Secondly, our projections would have to be validated by developers with Drools experience. For this reason, we needed to create a projectional language that was much closer to the Drools language.

Our following Language, Drools-Lite, contains many more of the features of Drools. Our method of selecting the features involved implementing the examples delivered with Drools (including the corrupt politician example shown in section 2.1.2). We would implement just enough features to complete the examples. Whenever we had any queries about how to design the Concepts, we referred to our analysis of the Drools Language, shown in appendix F. We show the preliminary design we achieved using this method in figure 3.7. Later, there were some places we diverged a little from our design. We merged and decoupled our Concepts when we thought it would simplify the code.

RuleFile The RuleFile level statements contain Facts, Globals and Rules. It also contains semantically unimportant empty lines.

Fact A Fact has a type property. We implement the type property using a ClassifierType from the MPS BaseLanguage. This implementation allows the File to refer to BaseLanguage classes implemented in the same solution and Java JAR files. We created a smart reference Concept for this to take advantage of built-in MPS UI functionality. A smart reference is a node with a single reference of 1:1 cardinality. The editor builders know how to select which nodes in scope to display to the developer if one uses this object rather than directly referencing the node to which it refers.

FactProperty In RSR, we had FactProperties as children of Facts. Now that our Facts refer to actual classes (ClassifierType), our FactProperties should reflect this. To do this, the Concept itself only references an InstanceMethodDeclaration, the MPS BaseLanguage's definition of a Method signature. We scoped the Concept to only show properties associated with a selected Fact.

```

rule file name: DossierSleutelbos

fact Dossier : hasRunningEpisode , isWvGGZ
fact DroolsContext : << ... >>
fact Episode : isCM, done,
fact Milestone : cmDecisionStep , finished ,

rule "0"
when
    Dossier ( )
    DroolsContext ( )
then
    // nothing
end

rule "[WVGGZ/CM] Start CM procedure" "
when
    Dossier ( isWvGGZ , !hasRunningEpisode )
then
    // nothing
end

rule "[WVGGZ/VCM] dossier_Start_VCM "
when
    ( Dossier ( isWvGGZ , !hasRunningEpisode ) or ( Episode ( isCM , !done ) and
        Milestone ( cmDecisionStep , finished ) ) )
then
    // nothing
end

```

FIGURE 3.6: RSR program

Drools interacts with Java objects as if they are Java Beans. To simulate this, we limited the scope of the properties to just getters, i.e. methods that start with “get” or “is” and used a Behavior to make sure they displayed without the “get” or “is” prefix. We also made a smart reference for this Concept.

Another option for achieving this is to have wrapped the ClassifierType and referenced its related InstanceMethodDeclarations. We would have then had to limit the functionality of these items from the BaseLanguage. Whilst this allows the functionality we wished for, we feel our construction offers decoupling and that we think correctly reflects the structure of the language. Perhaps if we were to redo this, we would have taken the other approach.

Global Our Globals are very simple. They have a name and a BaseLanguage Type. We added a smart reference so that Rules can easily use them. The reference extended Expression from the BaseLanguage. This extension is so that we could use it in the Right-hand side.

Rule Our Rules have three children: an Attribute collection, a Right Hand Side and a list of Conditions that make up the Left-hand side. We created a component to describe the rule editor for reuse, as we imagined that we would wrap this in other projections.

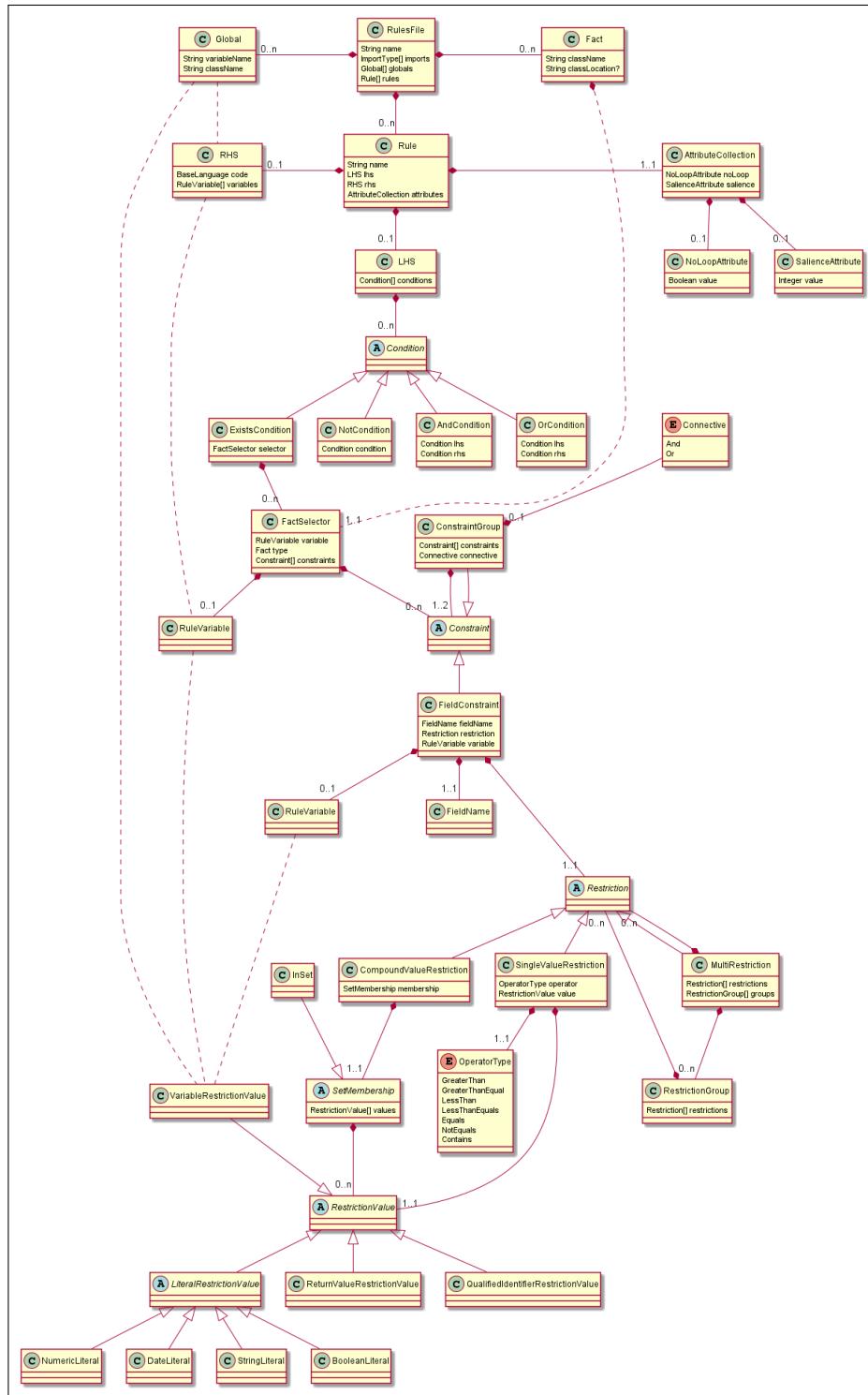


FIGURE 3.7: Drools-Lite structure

RuleVariables The Fact of a FactSelector and the FactProperty of a FieldConstraint are bound to RuleVariables. RuleVariables are scoped to a Rule. A RuleVariable has only a name and a type. We also create a smart reference for it so that it can be used elsewhere within the rule. Like the Global, it extends BaseLanguage's Expression to be available in the Java code of the Right-hand side.

Right-Hand Side The right-hand side of the rule, for the most part, is Java code. To implement this, we made the right-hand side of the rule a single StatementList. A StatementList is a list of Statements, both from the BaseLanguage. We chose these because they keep track of, amongst other things, scope.

There are some non-Java, Drools specific items that are available to the right-hand side. Items that had to be useable within the right-hand side were Globals, RuleVariables and Drools specific functions. These all extend Expression from the BaseLanguage. This extension allows seamless integration with the Java code.

The Drools specific Methods that are required are Insert, InsertLogical, Modify, Delete and Halt.

```

rule "set up"
when
    $s : Student( )
then
    modify( $s ) { setCumlaude( false ) };
    Program program = $s.getProgram();
    insert( program );

    foreach course in program.getCourses() {
        insert( course );
    }
end

```

FIGURE 3.8: RHS

Figure 3.8 shows some of the features discussed for the right-hand side as shown in our default projection. The right-hand side is the text shown between the `then` keyword and the `end` keyword. In the figure, one can see examples of plain Java code, such as assigning to the variable `program` and the `foreach` loop. We can also see that Drools-Lite RuleVariable `$s` is in the Java statements. We have also highlighted the Drools specific methods placed in the code, in this case `modify` and `insert`.

RuleAttributes Rule Attributes is a container to hold all of the attributes that apply to a rule. Initially, we have only implemented the No-Loop and Salience Attributes. A developer activates these Attributes using two intentions we added to the Rule Concept. We can see, on line 2 in figure 3.9, an example of the salience attribute added to a rule on line 2.¹

¹In figure 3.9, we added line numbers to this figure to make it easier to talk about. The keywords `rule` on line 1, `when` on line 3, `then` on line 6, and `end` on line 8 have no meaning in the abstract syntax. We added them to give the developer the same look and feel as a standard Drools file.

```

1 rule "Rule #1"
2 salience 10
3 when
4 Program( faculty in ( Faculty.Law, Faculty.FNWI ) )
5 $s : Student( avg >= 8 ) || not Result( grade <= 7 && > 8 )
6 then
7     modify( $s ) { setCumlaude( true ) };
8 end

```

FIGURE 3.9: Rule

Left-Hand Side This is a collection of conditions. There are four types of conditions. AndCondition, OrCondition, NotCondition and ExistsCondition. AndCondition, OrCondition, and NotCondition have one or two children who are also conditions. The ExistsCondition contains a FactSelector.

We added dynamic braces to only show braces around a Condition if a child of another Condition. These braces add visual clarity without adding unnecessary clutter. We also added some intentions to make it easy to switch between ExistsCondition and NotCondition.

On line 4 in figure 3.9, the whole line represents an Exists Condition. Line 5 shows an OrCondition containing an ExistsCondition and a NotCondition. The default editor, through an intention, can make the ExistsCondition explicit with an exists keyword. However, the standard practice with Drools developers is to make this implicit, so this is how we show it here.

FactSelector This always has a reference to a fact. These facts are Program in line 4 of figure 3.9, and Student and Result from line 5.

Optionally, the FactSelector can be bound to a variable. In figure 3.9 line 5 the FactSelector referencing the Student Fact is bound to the \$s variable.

The FactSelector also contains a list of constraints on FactProperties, that all must return true for the FactSelector to return true.

Constraints We have three types of constraints. AndConstraint and OrConstraint contain other constraints. The FieldConstraints places restrictions on FactProperties.

FieldConstraints A FieldConstraint refers to a FactProperty and can be bound to a variable. It also has a restriction applied to that FactProperty. Using a substitute menu, we wrapped the FactProperty smart reference. This substitution automatically creates the FieldConstraint from the FactProperty selection by the developer.

There are several types of Restrictions and several types of values that they can restrict.

RestrictionValues The RestrictionValues that a property can be compared with are as follows. LiteralRestrictions: These are Integer, Float, String, DateTime and Boolean. VariableRestrictions: These can be global variables, RuleVariables referring to Facts from the FactSelector, or RuleVariables from other FieldConstraints. ReturnValue: This compares to anything expressed as an expression, which includes referring to constants or values behind qualified identifiers.

In figure 3.9 on line 4 we have the return values Faculty.Law and Faculty.FNWI. on line 5 the literal values 7 and 8.

Restrictions A SingleValueRestriction compares a FactProperty against a value. A MultiRestriction compares a FactProperty against multiple values, not necessarily using the same comparison for each value. A SetMembership restriction checks if a FactProperty is a member of a group.

In figure 3.9 on line 4 a SetMembership restriction is shown with the `in (Faculty.Law, Faculty.FNWI)` text. Line 5 in the first FactSelector there is the SingleValue restriction `avg >= 8`. The second FactSelector shows a MultiRestrictions grade `<= 7 && > 8`.

Thus, we have described the pertinent implementation details of the Drools-Lite language.

3.2.3 Wireframes

There are some potential projections we have conceived for which there is not sufficient time to implement. We want others to assess these and thus would like them to appear as realistic as possible to the assessors.

Our solution to this conundrum is to develop these presentations in a wireframing tool. The Wireframe tool we chose was Axure[72]. We chose this because we had previous experience with the product. Also, it is available to students for free.

After much discussion, we settled on two possible projectional programming aids: Truth table and circuit diagram. We will discuss these in more detail in the results section.

3.3 Method: Survey

We tested the validity of the prototype using a survey. If a survey is not well designed, then it could lead to invalid or irrelevant outcomes. This chapter describes the design and procedure of the survey. Additionally, it outlines any threats to its validity in this chapter. Our choice of survey technique is a questionnaire.

3.3.1 Questionnaire Design

To design the survey of our prototype, we followed the following rules derived from the works of Bryman[73] and de Vaus[74].

- *Introduction* We devised a clear introduction to describe the research.
- *Existing work* We considered existing questions. With regards to projectional editing, we requested the original questionnaires from three papers[29, 34, 52] about tools developed using projectional editing. Unfortunately, none of the original questions was available for assessment.
- *Question in mind* When formulating the questions, we had the specific research question, “Which projections can help developers get appropriate feedback about rules?” in mind.
- *Succinct* The pool of Drools users that we were personally in contact with was tiny. Thus we had to rely on responses from strangers. For this reason, we tried to make the questionnaire as quick to finish as possible. This constraint meant we looked particularly hard at removing questions that did not help us to our research goal.
- *Pilot* We piloted the questionnaire with both ourselves and our industrial supervisor. The result of this pilot led to more explanatory text before the questions.
- *Clarity* The instructions to each of the questions were tested for clarity by a non-technical third party. We took care to rework questions that were long, ambiguous, general or leading not to be so. We also took care to remove jargon, negative wording, and questions that asked about more than one thing.
- *Closed questions* The only open questions were ones from which we wished to extract sentiment. To avoid binary questions, where appropriate, we applied a Likert scale[75].
- *Single page questions* Thanks to the UI of Survey Monkey, no questions spanned multiple pages.
- *Important questions first* We started with the research-based questions, leaving the socio-demographic questions, such as skill level, to the end.

Appendix G shows the questionnaire we designed following these principles.

3.3.2 Participants

The requirement for participants is that they have at least a little experience with using Drools. We hoped to get a statistically significant number of participants.

3.3.3 Validity

We addressed the non-response bias[76] by making the questionnaire short and easy to answer. Because of the nature of the participant selection for this survey, it will be challenging to address the self-selection bias caused by the voluntary nature of the response.

Common method bias, i.e. “variance that is attributable to the measurement method rather than to the construct the measures represent”[77] can be responsible for 25% or more of variable relational influence. As we are only conducting a single survey, we will not be able to prevent this. However, we took the following small precautions. We tested the survey to remove question ambiguity, mood influences, and length issues. We mixed the order of questions in the survey to mitigate the issues caused by the similarity, proximity, and location of items. We varied the scales and order of our Likert scales.

The main statistical methods to address common method bias, i.e. “Harman’s single factor test”[77] and the “marker variable”[78] have been found to be lacking in grounding[79]. The marker variable approach is appropriate if used with caution. However, it may not be possible to gain a statistically significant outcome with our expected response size.

3.3.4 Pre-test

We sent our first pass of the survey to our industrial supervisor, who has experience with Drools. With this pre-test, we hoped to remove ambiguously worded or leading questions. Additionally, we wanted to confirm that the questionnaire took around 10 minutes to complete.

As a result of this pre-test, we updated much of the explanatory text.

3.3.5 Sampling

Within our professional network, we only had a connection with very few Drools developers. We also considered that having acquaintances answer the questionnaire could introduce some biases we could not account for. Thus we had to expand our sampling reach.

Our first approach was to search StackOverflow for question askers and answerers on the subject of Drools. Our preference was to find email addresses, failing that Twitter contacts. Unfortunately, this proved quite limited. We only managed to harvest 13 email addresses, and six Twitter handles.

Our following approach was to interrogate our LinkedIn connections for anyone who claimed Drools as one of their marketable skills. At two degrees of separation in our LinkedIn network, we found 204 candidates with Drools skills listed. From these, we harvested 54 email addresses, and 40 Twitter handles.

We chose not to expand our search to three degrees of separation. At two degrees of separation, we share a relationship with the same person. We thought it would be harder to take advantage of the social pressure to answer when we do not have that shared relationship.

3.3.6 Procedure

The questions, as described in appendix G, were turned into a survey using the Survey Monkey service. We also show screenshots of one version of the questionnaire in appendix G.

To encourage response, we crafted a short introduction email, heavily relying on techniques designed to enhance response as discussed by Cialdini[80].

Chapter 4

Results

We divide our results chapter into three sections: the results from our systematic literature review, those from our action design research, and finally, our survey results.

4.1 Results: Systematic Literature Review

We carried out a systematic literature review (SLR), as described in the methods section 3.1. We summarise the results of our SLR as “undetermined”. We make this statement because the review design was not appropriate for the problem domain.

We abandoned the SLR after the quality assessment stage. We could not find a quality assessment checklist that adequately dealt with action design research (ADR) studies. This inadequacy proved problematic, as most of the primary studies were ADR studies.

Therefore what follows should be considered the results of a quasi-SLR, with the quality assessment stage ignored.

4.1.1 Papers Selected

We logged the details of what we describe in this section in appendix B.

Figure 4.1 shows the results of the five iterations that the search went through. Out of 173 results, we had 50 papers that initially seemed to pass our inclusion and exclusion criteria from our initial search. From the initial 50, we added 18 papers from a possible 109 in our first iteration of forward and backwards snowballing. The next snowballing iteration returned three papers that matched our criteria. The third round of snowballing had no papers matching our criteria and thus terminated this stage of selection.

Our final selection iteration involved a deeper scan of the remaining 71 papers. In this stage, we rejected 12 papers that were not primary research and one paper which reported on an already represented study. Further, we rejected twenty-seven papers that were, on closer reading, not about projectional editing.

This final selection filter left us with 31 papers before the quality assessment filter.

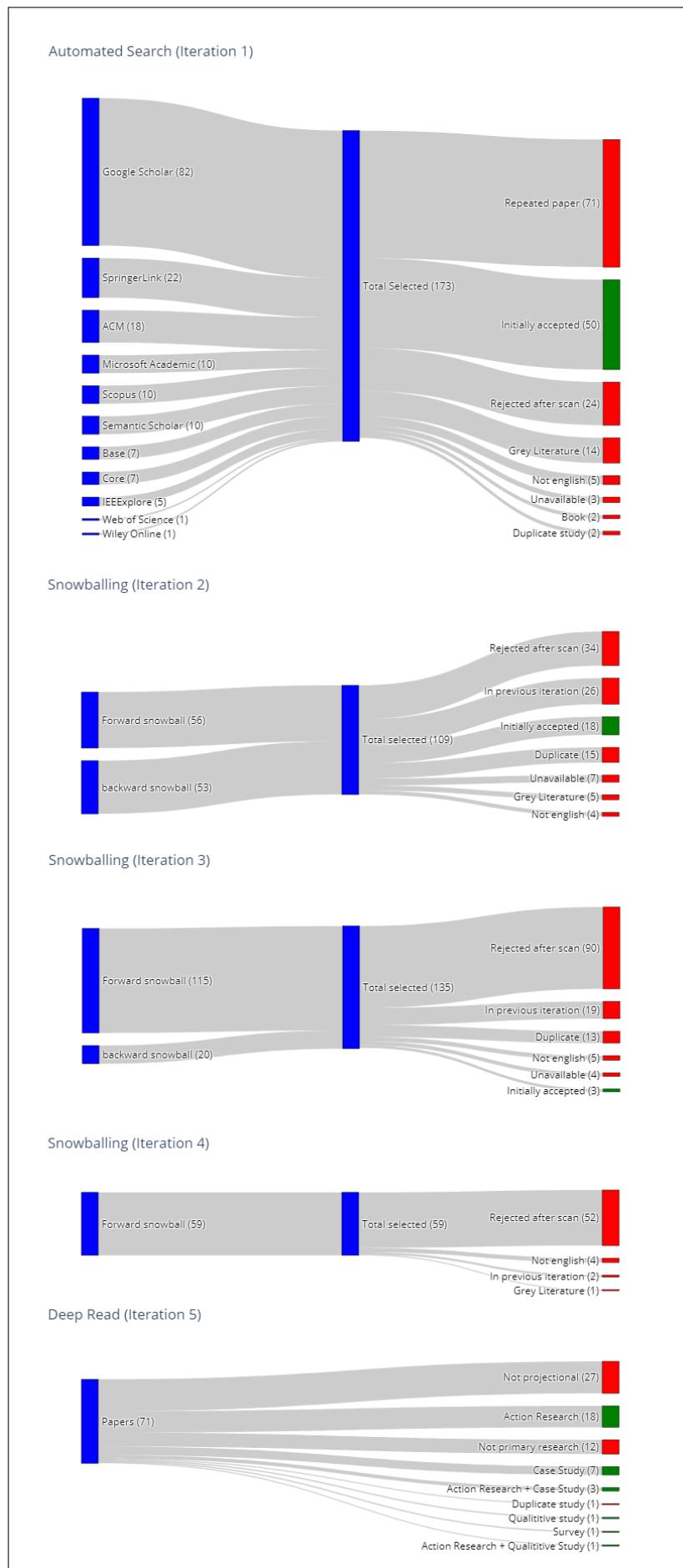


FIGURE 4.1: Search results

Sensitivity and Precision

As a curio, we reappropriated Zhang's[64] ideas of sensitivity and precision and applied them to the search engines rather than search strings. We calculate the values for sensitivity and precision of the search engines as follows:

$$\text{sensitivity} = \frac{\# \text{ retrieved relevant studies}}{\# \text{ all relevant studies}} \times 100\%$$

$$\text{precision} = \frac{\# \text{ retrieved relevant studies}}{\# \text{ studies retrieved}} \times 100\%$$

Table 4.1 show that Google Scholar had the highest sensitivity, returning 22 of the 31 chosen studies. This sensitivity came at the cost of a considerable proportion of false positives. Microsoft Academic and SpringerLink were the joint-most precise, with half of their search results ending up in the final roster. With the second-highest count of documents, the second-highest sensitivity, and joint highest precision, SpringerLink would appear to be the best all-around search engine for this field. However, these figures are skewed by several of their articles coming from a single collection specifically about projectional editing.

| Search engine/library | original # | selected # | sensitivity | precision |
|-----------------------|------------|------------|-------------|-----------|
| ACM | 18 | 3 | 10% | 16% |
| BASE | 7 | 3 | 10% | 43% |
| CORE | 7 | 1 | 3% | 14% |
| Google Scholar | 82 | 22 | 71% | 27% |
| IEEEExplores | 5 | 2 | 6% | 40% |
| Microsoft Academic | 10 | 5 | 16% | 50% |
| Science.gov | 0 | 0 | 0% | 0% |
| SCOPUS | 10 | 3 | 10% | 30% |
| Semantic Scholar | 10 | 4 | 13% | 40% |
| SpringerLink | 22 | 11 | 35% | 50% |
| Wiley Online | 1 | 0 | 0% | 0% |
| Web of Science | 1 | 0 | 0% | 0% |

TABLE 4.1: Search engine sensitivity and precision

4.1.2 Quality Assessment

Using the quality assessment checklists, developed by Crombie et al.[81], shown in appendix C, we examined the remaining 31 papers, which on the surface represented 37 primary studies.

Unfortunately, there were no checklists for ADR studies. We, unsuccessfully, searched for an appropriate quality assessment checklist for ADR studies. We did not find a suitable checklist and did not consider ourselves suitably qualified to make one. Thus we used the quality assessment checklist for case studies to assess the ADR studies.

We used a rudimentary scoring system of +1 value for positive answers, 0 for undetermined, and -1 for negative answers. We arbitrarily defined that any study with an overall score greater than 0 was high enough quality to be part of our final analysis.

Unfortunately, we only found 6 out of 37 studies of high enough quality to pass this filter with this scoring. Thus, we had to choose: change our scoring, only use these six studies, terminate the SLR or ignore the QA findings.

Changing a method until it gave the desired answer seemed unscientific to us. Six studies seemed too few to give an overview of a field. Abandoning the SLA seemed the correct course of action. However, as we still wanted an overview, we decided to take a different course. We accept that what follows is no longer an SLA. We titled it a Quasi-SLA, which is like an SLA, which ignores the quality assessment results.

We could not reconcile that 84% of studies were high-quality enough to appear in recognised scientific journals yet were not of high enough quality to pass our SLR QA stage. After considering this disconnect, we found two significant threats to the validity of the Quality Assessment stage. The first being that a single researcher with no previous experience executed the QA stage. The second is that either case study checklists are inappropriate for ADR studies or that ADR studies are inappropriate for SLRs.

4.1.3 Analysis

After Identifying the primary studies, we extracted data. Appendix E shows the extracted data. Figure 4.2 shows that most of the primary studies in our review were ADR studies.

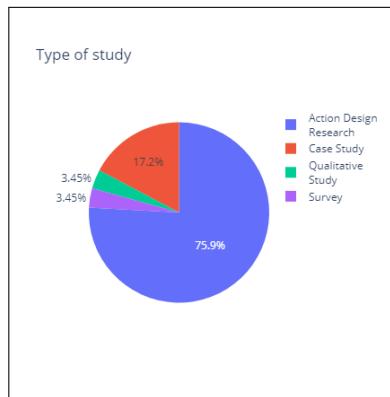
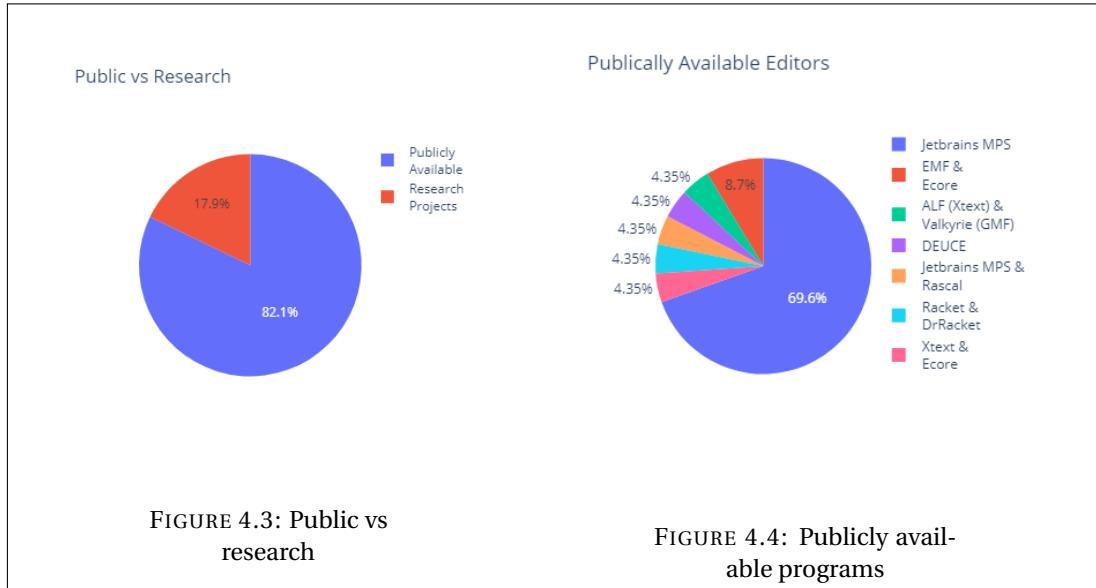


FIGURE 4.2: Study types

Tools used

We split the studies to see which were to do with purely research projects and which were researching using already publicly available commercial or open-source products. To calculate this, we removed the one primary study, a survey, as it covered many tools and options, but none of which was in-depth. The chart in figure 4.3 shows that over 80% of the projects were studying already existing publicly available options.



Of the publicly available software studies, we wanted to know which software attracted the most academic interest. Figure 4.4 shows that 74% of the studies into projectional editing that used a publicly available product used JetBrains MPS.

During this analysis, we discovered that two of the papers we had initially categorised as primary studies were, in fact, proposals, and thus we removed them from our analysis.

Sentiment

In this study, we included all 29 papers. We tagged each section from those papers that talked about projectional editing. We then broke each of these sections into sentences and ran those sentences through a sentiment analyser as described in section 3.1.6. We show the outcome of this sentiment analysis in figure 4.6.

The charts show the relationship between the positive, neutral and negative sentiment outcomes. On the y-axis, we have an Id for the papers examined. We show the keys linking the Ids to the paper names in table 4.2;

The first chart shows the absolute number of sentences analysed per paper partitioned by whether they returned negative, neutral or positive sentiment results. The second chart shows these as percentages so that the papers are comparable. In the third chart, we removed the neutral scores and calculated the percentage positive to negative. The final chart is an aid to make it easier to scan whether papers trended positive and which trended negative. If we found the papers to be equally positive and negative, we classified them as neutral.

Over the 29 papers, we scanned a total of 3003 sentences. Four hundred thirty-five were analysed as being positive, 1953 neutral and 615 negative. Thus 14% were positive and 16% negative.

10 of the 29 papers were more positive than negative when discussing projectional editing, 16 more negative and three equally negative and positive.

In figure 4.5 we attempt to separate out sentiment by product category. These categories being Research projects, MPS, and all the other used products. We ignored the survey paper in this one as it covered all of these types.

MPS dominates the sentences accounting for 17 (61%) of the 28 papers and 2051 (73%) of the 2791 sentences analysed.

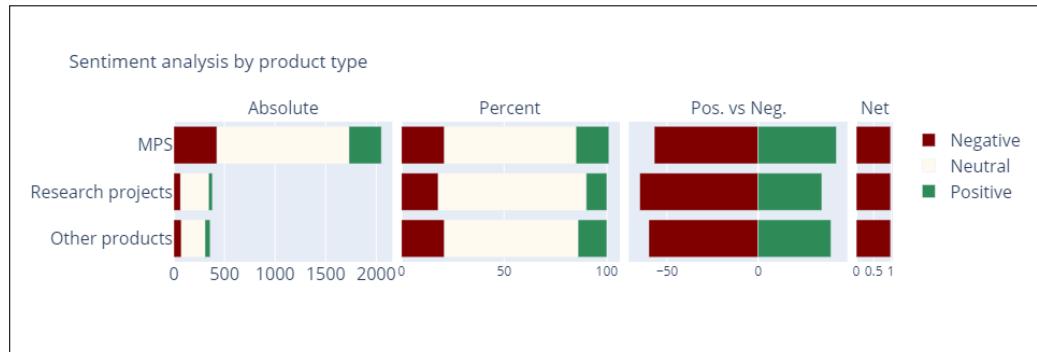


FIGURE 4.5: Sentiment analysis by product

A Narrative Synthesis

Our synthesis of the papers that appear in table 4.2 will be short. We will avoid rehashing the advantages and disadvantages of projectional editing, which come up again as we discussed these thoroughly in sections 2.2.4 and 2.2.5.

Many of these papers focus on models and model-driven development, occasionally suggesting a shift towards textual modelling languages. However, other papers point out that text does not always supply a suitable level of abstraction in modelling. One paper suggested that developers prefer text whereas maintainers and domain experts visual projections, though this suggestion was unsourced.

When authors have used solutions other than MPS, they complain about issues such as MPS being heavy-weight, with much overhead. However, these authors then spend a great deal of time theorising about fixing issues in their architecture which, because of its architecture, MPS does not encounter. These issues include synchronising between various views and how grammars deal with notations.

There is a fair bit of mention of a “semi-projectional” approach, which involves parsing at the leaf node level of the AST. This approach is mainly from papers not using MPS, but also some which do. The approach, it seems, is a reaction to the difficulty in simulating the text language experience in a projectional editor. It echoes the approach the Synthesizer Generator adopted when facing this same problem in the 80s.

The projects they describe are not in industrial use. These authors suggest that projectional editing is probably best suited to helping novices learn a language.

Those authors who use MPS are mostly discussing products developed for use in an industrial setting or how best to teach projectional editing to a broader audience. MPS, when used, is often seen as a critical enabler. Two of MPS’ properties that garner the most mentions are the ease of composition and multiple views.



FIGURE 4.6: Sentiment analysis

| Id | Paper name |
|----------|---|
| 1 [109] | A domain-specific language for payroll calculations: A case study at DATEV |
| 2 [110] | A framework for projectional multi-variant model editors |
| 3 [111] | A generic projectional editor for EMF models |
| 4 [112] | A model-driven approach towards automatic migration to microservices |
| 5 [113] | AdaptiveVLE: An integrated framework for personalized online education using MPS JetBrains domain-specific modeling environment |
| 6 [114] | Adding interactive visual syntax to textual code |
| 7 [115] | Blended graphical and textual modelling for UML profiles: A proof-of-concept implementation and experiment |
| 8 [116] | Classification algorithms framework (CAF) to enable intelligent systems using JetBrains MPS domain-specific languages environment |
| 9 [117] | DSL based approach for building model-driven questionnaires |
| 10 [118] | Efficient editing in a tree-oriented projectional editor |
| 11 [119] | Efficient generation of graphical modelviews via lazy model-to-text transformation |
| 13 [120] | Engineering gameful applications with MPS |
| 15 [121] | Fasten: An extensible platform to experiment with rigorous modeling of safety-critical systems |
| 16 [108] | Gentleman: A light-weight web-based projectional editor generator |
| 17 [122] | Integrating UML and ALF: An approach to overcome the code generation dilemma in model-driven software engineering |
| 18 [123] | Javardise: A structured code editor for programming pedagogy in Java |
| 19 [124] | JetBrains MPS as core DSL technology for developing professional digital printers |
| 20 [125] | Learning data analysis with metaR |
| 21 [126] | Migrating insurance calculation rule descriptions from Word to MPS |
| 22 [127] | Model-based safety assessment with sysml and component fault trees: Application and lessons learned |
| 23 [128] | Papyrus for gamers, let's play modeling |
| 24 [129] | Projecting textual languages |
| 25 [130] | SpecEdit: Projectional editing for TLA+ specifications |
| 26 [131] | Teaching language engineering using MPS |
| 27 [132] | Teaching MPS: Experiences from industry and academia |
| 28 [133] | Tiny structure editors for low, low prices! (generating guis from <code>toString</code> functions) |
| 29 [134] | Towards ontology-based domain specific language for internet of things |
| 30 [135] | Type-directed program transformations for the working functional programmer |
| 31 [136] | What do practitioners expect from the meta-modeling tools? a survey |

TABLE 4.2: paper key

The users of MPS agree that simulating the experience of the text editor user in the projectional environment is still very hard. However, new plug-ins are making this somewhat manageable. The most prominent carion call amongst the MPS users is for a web-based interface.

The steep learning curve is another issue. Several papers offer solutions to this, such as example-driven development, gamification, grammar to MPS plug-ins and something called a “language wheel”.

In general, researchers using MPS have a few gripes with usability but seem to be very positive. One paper, which was not using MPS, said that some problems become intractable when dealing with graphical models. Another paper, in a coincidence of word use, when describing the decision to use MPS, explained that it was because it presented a tractable level of complexity.

4.2 Results: Action Design Research

4.2.1 Really Simple Rules

The Really Simple Rules Language (RSR) acted as a training ground for our new projections.

Context Aware Color Scheme

After the default text projection, the first projection we made was giving the text a colour scheme. This form of augmentation in IDEs is probably the most basic that we see. Available in structured editors since the 1980s[82], syntax highlighting displays text in various colours and fonts according to the meaning of the terms. Syntax highlighting is helpful for the comprehension of code, at least for small code bases[83].

Developers at our host organisation use Eclipse or IntelliJ Community Editions to edit code, neither of which has syntax highlighting for Drools. Thus, the addition of this feature would immediately benefit them. However, IntelliJ IDEA, the paid version, already provides this feature for Drools. In order to offer another visual augmentation that we considered useful, we extended the colour scheme to indicate whether the selection is looking for a positive or negative match. This is shown in figure 4.7.

```

rule "Dossier verwijderen"
when
    not Episode( )
then
    // execute rule outcomes
end

rule "[DOSSIER] dossier_BewerkNAW_Rihg"
when
    Milestone( setUpRihgStep, !finished[groen] )
then
    // execute rule outcomes
end

```

```

rule "Dossier verwijderen"
when
    not Episode( )
then
    // nothing
end

rule "[DOSSIER] dossier_BewerkNAW_Rihg"
when
    Milestone( setUpRihgStep, !finished[groen] )
then
    // nothing
end

```

FIGURE 4.7: Context aware color scheme

Facts contained by NotConditions and FactProperties that are part of a NotPredicate appear highlighted in Red. ExistConditions and IsPredicates have their content coloured green. We did not test whether this improved understanding.

Summary Projection

Our next projection allows developers to have a quick overview of the Rules and complexity of those Rules. Figure 4.8 shows that the developers can get an overview of both the number of Rules and the number of Facts in each of the Rules.

The building of this Rule only required adjusting two editors. The Rule count and Fact count were added to the File editor using Read-Only Model Access to count the descendants of the File that are Rules and Facts. The Rule editor was adjusted only to show the

```

rule file name: DossierSleutelbos rule count 152 fact count 357

0 : # facts: 2
Dossierdetails kunnen inzien : # facts: 1
Dossiers samenvoegen : # facts: 0
[*] dossier_NieuweNotitie : # facts: 1
[*] dossier_StuurVeiligeMail : # facts: 1
Dossier verwijderen : # facts: 1
[HV ] dossier_Sluiten : # facts: 1
[HV ] dossier_Afdoen : # facts: 1
[HV ] dossier_Actualiseren : # facts: 1
[WVGZ/VO] Bewerk NAW in WVGZ VO voor meldmedewerkers : # facts: 3
[WVGZ/IBS_WZD] Bewerk NAW in WVGZ IBS_WZD : # facts: 1
[WVGZ/CM] Start CM procedure" : # facts: 1
[WVGZ/CM] Verzamel Stuurgegevens : # facts: 2
[WVGZ/CM] Wijzig Stuurgegevens Voor Afronden Medische Verklaring : # facts: 3
[WVGZ/CM] dossier_Act_Opstellen_CMMedischeVerklaring : # facts: 2
[WVGZ/CM] dossier_Act_OndertekenenMedischeVerklaring : # facts: 3
[WVGZ/CM] dossier_TerugzettenAfgerondeCMMedischeVerklaring : # facts: 5

```

FIGURE 4.8: Summary projection

Rule's title and, again using the model access, the count of the descendants of the Rule that were Facts.

Whilst this may look like a report that any language workbench could create, the File name and the names of the Rules are editable in this projection.

Filtering

When investigating how to handle extensive collections of rules, we sought other domains than programming that handle extensive collections of items. The domain of data analysis has a long history of handling large volumes. Among their two most used tools for exploration are sorting and filtering.

The nature of business rules lends them to some projectional options that would not make sense with other programming styles. Because of the independent nature of the rules, filtering, in particular, lends itself to the business rules style. The semantic meaning of the order of business rules means we did not find a good use case for sorting rules. Thus we decided to implement a filtering projection.

Whilst use of filtering occurs in other places in the coding pipeline, such as in deciding on what code completion to present^[84], and version control visualization^[85], we were unable to find any research on applying filtering directly to code files. Thus, we think what we present here is an original idea.

Rules that use the same Facts or FactProperties are likely to be related. Thus these seemed the obvious items to filter. We created a projection where if the developer filtered by a Fact or a FactProperty, the projection would filter out all rules that did not contain the item. Once the rules were filtered, the projection only shows Facts and FactProperties used by those rules.

In our implementation, shown in figure 4.9, we show three places where we use intentions to filter the code. The first is an intention associated with a Fact. We show the outcome of choosing this filter on the righthand side of figure 4.9. The second intention is on a FactProperty. As the FactProperty is a child of a Fact, we see both intentions. The

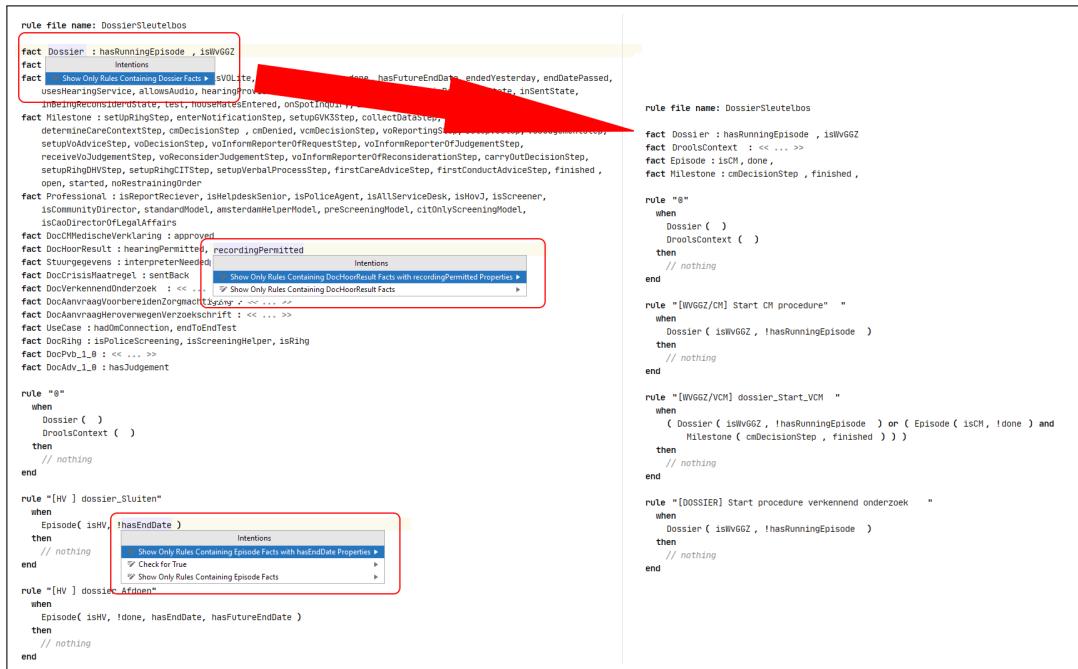


FIGURE 4.9: Filtering projection

third highlighted intention is on a FactProperty Reference. It also shows an intention associated with a FactReference in the FactSelector that holds the FactPropertyReference as a child.

One of our guidelines was, as much as possible, to build our projections as separate languages, non-invasively extending RSR. In our first approach at the filtering, we failed on this count by invasively adding properties to Fact and FactProperty Concepts of the RSR to determine whether they were visible.

Our following approach created subclasses of Fact, FactProperty and File. This approach, however, requires running a macro on the code file to migrate Facts, FactProperties and Files to FilteredFacts, FilteredFactProperties, and FilteredFiles. This migration means that the FilteredFile could now only be used by languages that extend our new filtered language.

Our final approach was to add a Filter Concept, reference the filtered nodes, and have the editors make the visibility calculations based on this singleton node. Whilst more complex, this removed the need for invasive changes and allowed other languages to combine with the filtering language.

Filtering is a handy projection. However, it breaks Dijkstra's rule "the purpose of abstraction is not to be vague but to create a new semantic level in which one can be absolutely precise."^[86] This projection fails this rule by hiding some of the meaning of the code. This projection has no way of containing the whole code whilst a filter is applied. However, so long as there is a clear indication that a filter is applied, then we see this as a tool in a similar vein to the code collapsing functionality found in most modern-day editors.

Table

Thus far, our projections have been textual ones that other non-projectional language workbenches could implement. Creating a table will be our first non-parseable projection.

We choose the table projection based on the observations of Miller[2] about the number of items people can retain in their memory. This observation leads us to conclude that the fewer essential items that are off the screen and, therefore, in the developers' memory, the better.

| rule file name: DossierSleutelbos count 153 | |
|--|---|
| rules: | |
| Name | When conditions |
| # | Dossier() |
| hello | DroolsContext() |
| Dossierdetails kunnen inzien | Dossier() |
| Dossiers samenvoegen | << ... >> |
| [*] dossier_NieuweNotitie | << ... >> |
| [*] dossier_StuurVeiligeMail | << ... >> |
| Dossier verwijderen | << ... >> |
| [HV] dossier_Sluiten | << ... >> |
| [HV] dossier_Afdoen | << ... >> |
| [HV] dossier_Actualiseren | << ... >> |
| [DOSSIER] dossier_BewerkNAW_Rihg | Milestone(setupRingStep, !finished[groen]) |
| [DOSSIER] dossier_BewerkNAW_Gvk3 | Milestone(setupGVK3Step, !finished[groen]) |
| [DOSSIER] dossier_StartHvAnvrAag | Dossier() hasRunningEpisode) |
| [WVGGZ/CM] Bewerk NAV in WVGGZ CM | << ... >> |
| [WVGGZ/VO] Bewerk NAV in WVGGZ VO | Professional(!isReportReciever) |
| [WVGGZ/VO] Bewerk NAV in WVGGZ VO voor meldmedewerkers | // cannot handle checking professional works for owner of episode |
| | Professional(isReportReciever) |
| | // cannot handle checking professional works for owner of episode |
| | Milestone(enterNotificationStep, open[wit/rood]) |
| | -- -- -- |

FIGURE 4.10: Table projection

Figure 4.10, shows our rudimentary first table. This simple table has only the Name property and the when children of the Rules in the File. We implemented this projection using the tables extension in the MPS-Extension plug-in, created by Sascha Lißon.

Cross-tab

Our next tabular projection is a cross-tab inspired by a decision table. The idea behind this projection is that the previous table does not give any visual queues as to how rules are related. With a cross-tab, one can easily see which rules are using the same Facts.

| | | Dossier | Unfinished | Milestone | Professional | BewerkNaarRing | BewerkNaarOut | BewerkNaarTarget | BewerkNaarMedewerker | BewerkNaarVeiligeMail | BewerkNaarVeiligeMailOpdracht | BewerkNaarVeiligeMailOntvangen | Beeld | BewerkI_A | BewerkI_B | BewerkI_C | BewerkI_D | BewerkI_E | BewerkI_F | BewerkNaarMedewerkerRing |
|--|--|---------|------------|-----------|--------------|----------------|---------------|------------------|----------------------|-----------------------|-------------------------------|--------------------------------|-------|-----------|-----------|-----------|-----------|-----------|-----------|--------------------------|
| # | | #1 | #2 | | | | | | | | | | | | | | | | | |
| Dossierdetails kunnen inzien | | #1 | | | | | | | | | | | | | | | | | | |
| Dossiers samenvoegen | | #1 | | | | | | | | | | | | | | | | | | |
| [*] dossier_NieuweNotitie | | | | | | | | | | | | | | | | | | | | |
| [*] dossier_StuurVeiligeMail | | | | | | | | | | | | | | | | | | | | |
| Dossier verwijderen | | | | | | | | | | | | | | | | | | | | |
| [HV] dossier_Sluiten | | | | | | | | | | | | | | | | | | | | |
| [HV] dossier_Afdoen | | | | | | | | | | | | | | | | | | | | |
| [HV] dossier_Actualiseren | | | | | | | | | | | | | | | | | | | | |
| [DOSSIER] dossier_BewerkNAW_Rihg | | | | | | | | | | | | | | | | | | | | |
| [DOSSIER] dossier_BewerkNAW_Gvk3 | | | | | | | | | | | | | | | | | | | | |
| [DOSSIER] dossier_StartHvAnvrAag | | | | | | | | | | | | | | | | | | | | |
| [WVGGZ/CM] Bewerk NAV in WVGGZ CM | | | | | | | | | | | | | | | | | | | | |
| [WVGGZ/VO] Bewerk NAV in WVGGZ VO | | | | | | | | | | | | | | | | | | | | |
| [WVGGZ/VO] Bewerk NAV in WVGGZ VO voor meldmedewerkers | | | | | | | | | | | | | | | | | | | | |
| [WVGGZ/IBS_WZD] Bewerk NAV in WVGGZ IBS_WZD | | | | | | | | | | | | | | | | | | | | |

FIGURE 4.11: Cross-tab projection

Figure 4.11 shows our implementation of the cross-tab. At the top, we can see an immediate problem with a cross-tab, and that is if we have the whole File included, the table will be very sparse. Figure 4.11 also has a close-up of a cell showing a rule using three

FactSelectors that reference the same Fact. The other close-up shows that all the details of the selected Fact are available in the inspector.

The sparse table will not be a problem if the columns are thin enough to keep the table in a single screens width.

Everything is editable in this table, including deleting a Fact from a Rule. The table plug-in and MPS enabled most of the editing in the projection by default. An extra editing feature we added to this table was the ability to delete a Fact from the File, thus deleting all references to it from all the Rules in the File by deleting a Fact column. The code shown in figure 4.12 shows how we can walk the trees in each Rule to delete unary conditions and convert the non-deleted side of binary conditions into unary conditions to allow this Fact deletion.

FIGURE 4.12: Table Fact deletion code

Here we end our experiments in the RSR language.

4.2.2 Drools-Lite

Our subsequent experiments were with projections with the Drools-Lite language. As described in the section 3.2.2, Drools-Lite is an implementation that is much closer to the complete Drools language. This realism will allow us to create projections that we can present to experienced Drools developers for evaluation.

Of the learnings from the RSR language, one we felt needed fixing to improve understanding was the sparseness of the tables. By implementing the principle of maximising cohesion, we discovered we could reduce the sparseness issue. Thus, as a precursor to our projections, we extended Drools-Lite with a new language that contained one structural item, the RuleCollection. The RuleCollection is a child of the File and holds a collection of Rules. The idea behind this is that related Rules can be placed in the RuleCollection to make it easier to examine them together. This language also added an editor for the RuleCollection and intentions to move rules in and out of groups.

Decision Table

As the Drools language is analogous to a series of if-then-else statements, then perhaps its best visual equivalent is the decision table. Decision tables are a “powerful aid in programming, documentation, and ineffective man-to-man and man-to-machine communications”[87].

We designed our table, shown in figure 4.13, to include some of the lessons learned from the RSR cross-tab that shown in figure 4.11. The RSR language taught us that wasting visual real estate exacerbates sparseness issues in tables. In the cross-tab table, horizontal scrolling is necessary, in part due to the column widths. The columns were wide because the name of the Fact was displayed horizontally.

The Drools-Lite language allows for much longer selection criteria on FactProperties, which would lead to much wider columns. Our solution was to develop a vertically orientated header cell and use indentation to indicate if the cell is referring to just the Fact or a Fact and FactProperty combination.

The figure displays two decision tables side-by-side, representing rule groups for 'fnwi cumlaude rules' and 'law cumlaude rules'. Each table has a header row with various condition and action columns, followed by several rows of rules. The 'rule name' column contains the rule identifier, and the 'Actions' column contains the corresponding Drools-Lite code. The tables use icons (smiley faces) to represent RuleVariables and RuleProperties.

| rule group: fnwi cumlaude rules | | | | | | | | | |
|---------------------------------|--------|------------------|---------|------------------------|--------|-----------------------------|---|--|--|
| rule name | Course | name == "Thesis" | Program | faculty == Faculty.FNW | Result | course == [Course Variable] | Actions | | |
| [FNWI] > 7 | | | ⊕ | | | ⊕ | modified(variable) { setCumlaude(false) }; halt(); | | |
| [FNWI] Thesis >= 8 | c | | ⊕ | ⊕ | | ⊕ | modified(s) { setCumlaude(false) }; halt(); | | |
| [FNWI] avg >= 8 | | | ⊕ | ⊕ | | s | modified(s) { setCumlaude(true) }; halt(); | | |

| rule group: law cumlaude rules | | | | | | | | | |
|--------------------------------|--------|------------------|---------|------------------------|--------|-----------------------------|--|--|--|
| rule name | Course | name == "Thesis" | Program | faculty == Faculty.Law | Result | course == [Course Variable] | Actions | | |
| [LAW] avg grade >= 8 | | | ⊕ | | | ⊕ | modified(s) { setCumlaude(true) }; halt(); | | |
| [LAW] no grades < 7 | | | ⊕ | | | s | modified(s) { setCumlaude(false) }; halt(); | | |
| [LAW] Thesis >= 8 | c | | ⊕ | ⊕ | | ⊕ | modified(s) { setCumlaude(false) }; halt(); | | |
| [LAW] no resits | | | ⊕ | ⊕ | | s | modified(s) { setCumlaude(false) }; halt(); | | |
| [LAW] increment close count | | | ⊕ | | | ⊕ | int closeCnt = s.getCloseCount() + 1; modified(s) { setCloseCount(closeCnt) }; modified(s) { setCumlaude(false) }; halt(); | | |
| [LAW] only one between 7 & 8 | | | ⊕ | | | s | modified(s) { setCumlaude(false) }; halt(); | | |
| [LAW] increment exempt credits | c | | ⊕ | ⊕ | ⊕ | s | int exemptCount = s.getExemptCredits(); exemptCount += c.getEcts(); modified(s) { setExemptedCredits(exemptCount) }; modified(s) { setCumlaude(false) }; halt(); | | |
| [LAW] no more than 12 exempt | | | ⊕ | | | s | modified(s) { setCumlaude(false) }; halt(); | | |
| [LAW] completed too late | | | ⊕ | | | s | modified(s) { setCumlaude(false) }; halt(); | | |

FIGURE 4.13: Decision table projection

Because this projection presents both the left and right-hand side of the rules, we had to handle the Concept that spans both - the RuleVariable. We had to find a way to represent a RuleVariable that can be bound and used on the LHS and used on the RHS. We achieved this by referencing a RuleVariable name in the cell representing the Fact or FactProperty to which it is bound. With RuleVariables now being represented in the cells, we could no longer represent the cell being selected with an 'X', as this could be confused with a RuleVariable name. Projectional editing does not require communication of meaning through parseable ASCII text. Thus we decided to represent Fact selection with an icon. For arbitrary reasons, we chose a smiley face as that indicator.

The Rule names and actions are editable through the default functionality of the MPS extension. We use intentions to add the selection of a Fact or FactProperty to a Rule. We also use them for binding RuleVariables.

The major drawback of this design is that editing a rule with an, as yet, non-existent selection criteria became very clunky. If the Rule we wished to edit already existed in the table, we had to use an intention to extract it from the group, change the criteria and place it back in. At this point, the table would automatically adjust the column headings.

Experts examined this design in the questionnaire.

SpreadSheet

The domain-specific language for the finance world is the spreadsheet. One study estimated that 90% of computers had a spreadsheet on them[88]. Dan Bricklin's VisiCalc drove PC's into The office. VisiCalc was succeeded by Lotus 1-2-3, which Microsoft Excel succeeded as the dominant spreadsheet program in the workplace.

This level of familiarity with a paradigm led us to design a projection that had the look and feel of an Excel spreadsheet. We show this design in figure 4.14. To this end, we created a design where the selection criteria could be directly edited in the cell, as highlighted in the figure.

| rule group: fnwi cumlaude rules | | | | | | |
|---------------------------------|----------------------------|------------------|-------------------------|-----------------|--|--|
| rule name | Program | Student | Result | Course | | Actions |
| [FNWI] > 7 | \$ faculty == Faculty.FNWI | avg +\$ variable | < 8 == false | course +\$ name | | modify(variable) { setCumlaude(false)}; halt(); |
| [FNWI] Thesis >= 8 | == Faculty.FNWI s | | >= 8 == c c == "Thesis" | | | modify(s) { setCumlaude(false)}; halt(); |
| [FNWI] avg >= 8 | == Faculty.FNWI s | >= 8 | | | | modify(s) { setCumlaude(true)}; |

| rule group: law cumlaude rules | | | | | | |
|--------------------------------|---------------------------|----------|-------------------------|--------------------------------|--|--|
| rule name | Program | Student | Result | Course | | Actions |
| [LAW] avg grade >= 8 | \$ faculty == Faculty.LAW | avg >= 8 | | grade course exempted +\$ name | | modify(s) { setCumlaude(true)}; modify(s) { setCumlaude(false)}; halt(); |
| [LAW] no grades < 7 | == Faculty.LAW s | | | | | modify(s) { setCumlaude(false)}; halt(); |
| [LAW] Thesis >= 8 | == Faculty.LAW s | | <= 8 == c c == "Thesis" | | | modify(s) { setCumlaude(false)}; halt(); |
| [LAW] no results | == Faculty.LAW s | | | | | modify(s) { setCumlaude(false)}; halt(); |
| [LAW] increment close count | == Faculty.LAW s | | | | | int closeCount = s.getCloseCount() + 1; modify(s) { setCloseCount(closeCount)}; modify(s) { setCumlaude(false)}; halt(); |
| [LAW] only one between 7 & 8 | == Faculty.LAW s | > 1 | | | | int exemptCount = s.getExemptCredits(); exemptCount += c.getCrct(); modify(s) { setExemptCredits(exemptCount)}; modify(s) { setCumlaude(false)}; halt(); |
| [LAW] increment exempt credits | == Faculty.LAW s | | | | | modify(s) { setCumlaude(false)}; halt(); |
| [LAW] no more than 12 exempt | == Faculty.LAW s | | > 12 | | | |
| [LAW] completed too late | == Faculty.LAW s | | > 1.5 | | | |

FIGURE 4.14: Spreadsheet projection

Each row is a Rule in this design, and each column is for a RuleVariable or a FactProperty. If a property is selected, then the selection criterion is in the appropriate cell. A grey/beige colour indicates unselected cells. The RHS of the Rule appears in the actions column. Adding, as yet, unused Facts or FactProperties, or removing existing ones, can be achieved with intentions, as shown in figure 4.15.

This design also allowed us to have more than one selector for the same FactProperty, essential for our host organisation's code. We demonstrate this in the figure 4.16.

Experts examined this design in the questionnaire.

Here we end our experiments in the Drools-Lite language.

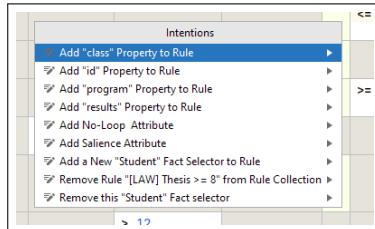


FIGURE 4.15:
Intention

| rule name | Program | Student | Result |
|------------|----------------------|----------|----------------------------|
| | \$ faculty | + \$ avg | + \$ grade exempted course |
| [FNWI] > 7 | != Faculty.Economics | variable | < 8 == false |
| | != Faculty.LAW | | |

FIGURE 4.16: Two of same property

4.2.3 Wireframe

After brainstorming several ideas to present as wireframes to experts as possible projectional aids to understanding, we chose two. We discuss them briefly in this section.

Truth Table

We decided to produce a truth-table wireframe example as we had had personal experience building truth tables to confirm the validity of Drools rules in our work.

The truth table seemed apt for the LHS of the Drools rule as, in essence, it is a boolean function. Wittgenstein popularised the truth table in the Tractatus Logico-Philosophicus[89]. They are so widely used in mathematics and computer science that we do not need to explain their use further. Because of the combinatorial explosive nature of truth tables, with 2^n possible combinations, we would limit the display to a max of 6 variables and only show the paths that lead to the RHS execution.

```

rule "Weird blanket"
when
    Program( faculty == Faculty.FNWI || == Faculty.LAW )
        ( Result( grade < 8 ) || not Result( exempted ) ) and Student( yearsStudied < 5 )
        Course( name != "Thesis" ) || Result( exempted )
then
    halt();
end

```

| | | | | | | |
|----------|----------|----------|----------|----------|----------|--|
| A | B | C | D | E | F | $((A \vee B) \wedge (((C \vee \neg D) \wedge E) \wedge (F \vee D)))$ |
| F | T | F | F | T | T | T |
| F | T | T | F | T | T | T |
| F | T | T | T | T | F | T |
| F | T | T | T | T | T | T |
| T | F | F | F | T | T | T |
| T | F | T | F | T | T | T |
| T | F | T | T | T | F | T |
| T | F | T | T | T | T | T |
| T | T | F | F | T | T | T |
| T | T | T | F | T | T | T |
| T | T | T | T | T | F | T |
| T | T | T | T | T | T | T |

FIGURE 4.17: Truth table projection

Figure 4.17 shows how we designed this to look. The user experience would be that the Rule is selected, and the developer presses the up and down arrow keys to step through the different true (highlighted in green) and false (highlighted in red) Fact selections that result in the Rule's selection.

We presented this design to our experts through the questionnaire to be validated.

Circuit Diagram

In our final projection design, we wanted to present a part of projectional editing that we had heretofore only made minimal use of. That is the use of manipulatable graphics that can change the AST.

We chose a logic circuit. The logic circuit represents a boolean operation as NOT, OR, XOR and AND Gates, with their inputs and outputs being inputs to other gates. In our design, shown in figure 4.18, the input wires to the gates are the Facts or FactProperties referenced in the LHS.

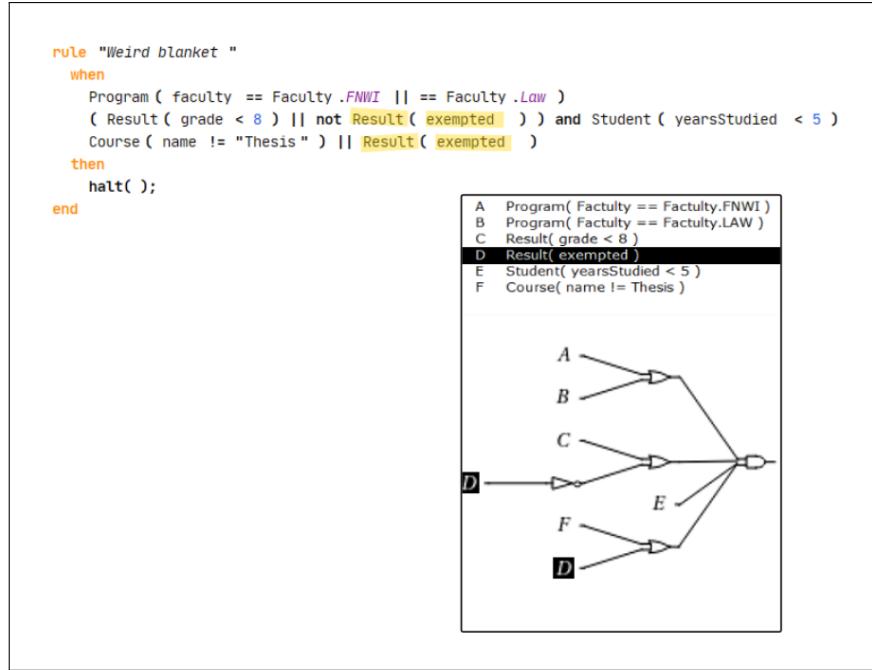


FIGURE 4.18: Circuit diagram projection

The user experience is that once the Rule is selected, the developer, by pressing the up and down arrow keys, can step through the different Fact selections (highlighted in yellow) and shown in the circuit diagram, thus showing how the Facts relate to each other.

We present this design in the questionnaire for validation.

4.3 Results: Survey

4.3.1 Population Selection

We initially had two sources for our Experienced Drools users to send our survey to. From LinkedIn we selected users who were at one degree of separation from us and listed Drools in their skills. From StackOverflow we selected users who had asked or answered questions about Drools.

As in these two websites the users do not tend to list their contact details, some investigation was required. From the initial selection, whose size we did not record we harvested email accounts, and failing that twitter accounts.

A few days into our survey we read a paper that described the use of academic papers as a population of expertise. We used Google Scholar to look up Drools papers from the previous 2 years. After skimming the papers to ensure that it was specifically about or using the Drools language we harvested emails

On the second and fourth day of the survey two subjects forwarded the weblink to the survey to mailing lists. One, a developer from the core Drools team, sent it to a list of known Drools consultants. The other sent it internally in his company. both the subjects who sent the survey to their mailing list forwarded links to version C of the survey.

We had created 4 versions of the questionnaires to combat single source bias. We distributed the surveys to the subjects harvested from LinkedIn and StackOverflow evenly. Because of the overrepresentation of Survey C, we distributed the subjects harvested from academic papers evenly over Surveys A, B and D.

The collection result can be seen in figure 4.19. What we see here is that the method of collection did not have much of an impact on return rates. whilst StackOverflow had a higher rate, the number of people contacted was so small that a small addition of respondents has an outsized effect on the proportion.

The first three pie charts represents the collection methods over which we had control. These three represented 24 of our 30 completed questionnaires. The last pie chart represents 6 completed and 4 partially completed questionnaires, that were returned from the surveys sent on by our initial participants. We do not know the size of the starting population of these lists. Thus this pie chart only shows the ratio of partial to completed results.

In summary, a survey reached known 154 participants, of which 24 completed it, for a Response Rate of 15.5%. In addition, an unknown amount of participants were reached through mailing lists, returning a further 6 completed surveys.

4.3.2 Participant Demography

Responses came from around the world. Figure 4.20 shows the location of the respondents were concentrated in Europe, the exceptions being the USA, Israel, and Singapore. Italy and the Netherlands provided the largest number of responses, with 7 and 5 respectively.

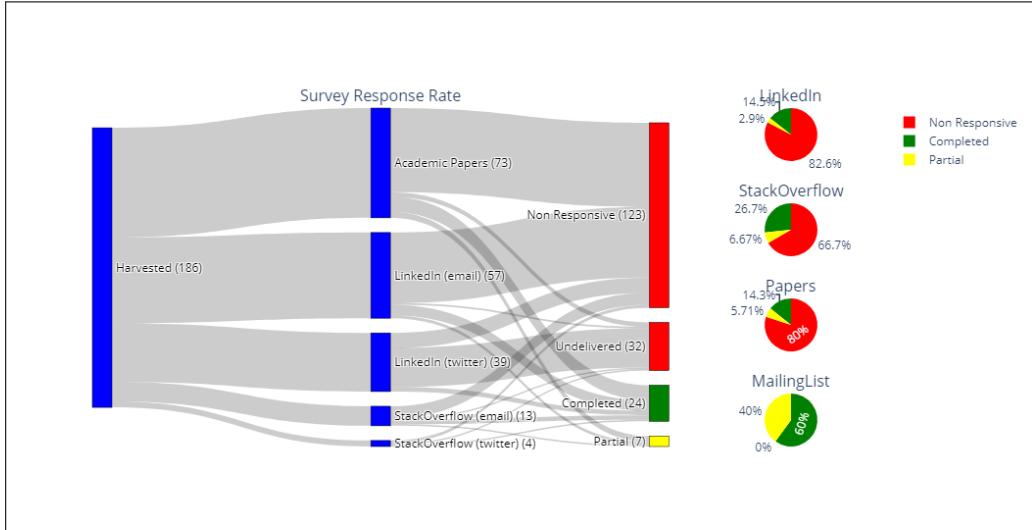


FIGURE 4.19: Survey participants

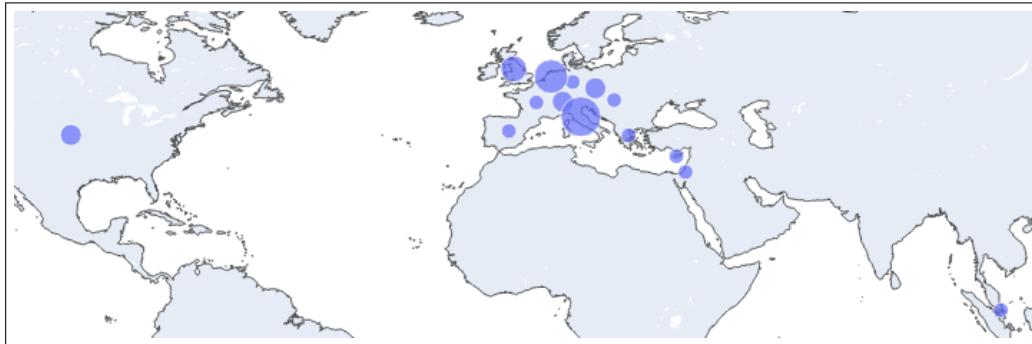


FIGURE 4.20: Survey locations

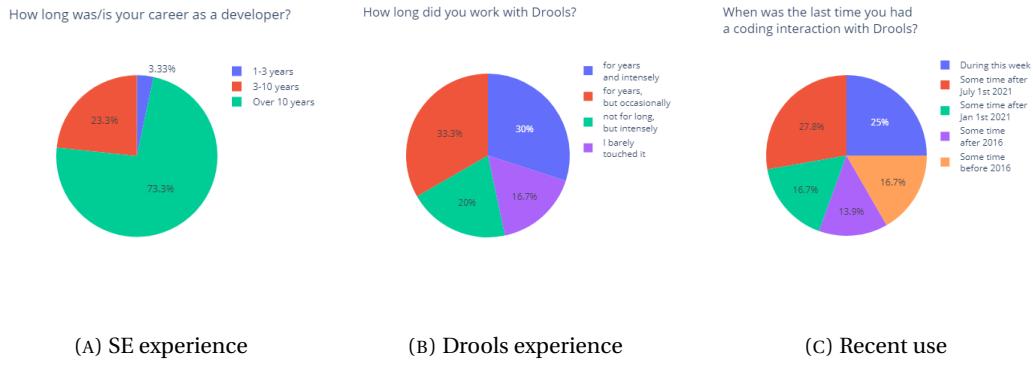


FIGURE 4.21: Subject experience

The experience of our subjects was quite high. As can be seen in 4.21a, most of our subjects have over 10 years programming experience. 17% of our recipients had a low experience of Drools, and 30% were very experienced, as shown in figure 4.21b. Figure 4.21c reports that over half of our recipients have used Drools in the previous 6 weeks with only 17% not having used Drools for more than 5 years.

Half of our subjects reported only ever using one editor for Drools, with the slight majority of those only using Eclipse. Eclipse also had the most instances of reporting of having been used, out of the 55 instances of editors reported as being used, 20 of those were Eclipse. There was a, to us, surprising diversity of tools being used. The purpose of this section was to be able to calibrate responses against exposure to IDEs with greater Drools Support. The wide diversity of editor usage and high incidence of multiple editor usage means that these answers are not suitable for use in the sub-categorisation of responses. The distribution of usage is shown in figure 4.22.

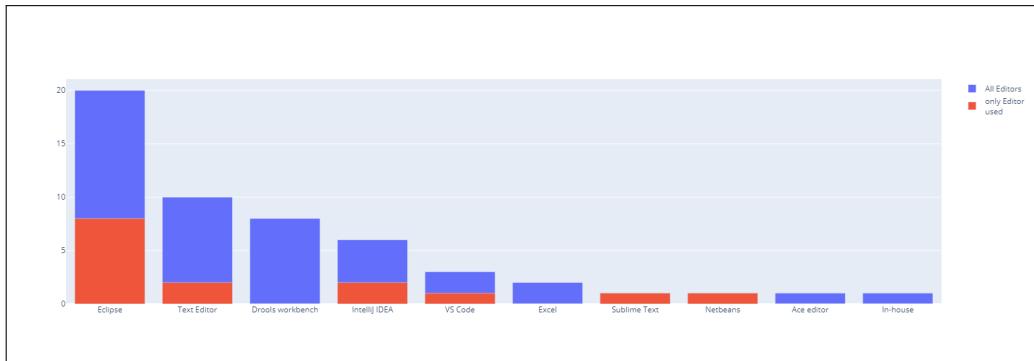


FIGURE 4.22: Editors used

4.3.3 Question Analysis

Grouping

When displaying subtypes we shall split into groups. The source of the response will create a pseudo cross-section of our participants. The 10 who were contacted through academic papers will be considered our academics. The remaining 20 will be considered practitioners.

The next grouping will be on Drools Experience. The 9 who replied they had used Drools “for years and intensely” are categorised as experts. The 16 who either answered “for years, but occasionally” or “not for long, but intensely” are categorised as seniors. The 5 who answered “I barely touched it” are categorised as novices.

Another grouping will be on recency of use. The 12 who have used Drools in 2021 will be categorised as current users. The remaining 18 as past users.

To remove some bias in the questionnaires we changed question order, order of projections, and which rulesets were used. When a question is effected by this, then this will also be displayed.

Display

Until now, the choice of chart to display the survey outcome has been based on a feeling rather than research. For the remainder of this section we will rely on the advice of our predecessors.

The remainder of this Results sections, whilst displaying results that regard our Likert scaled questions, we will be following the advice of Robbins et.al.[90] by using diverging stacked bar charts, with counts added. This style allows the evaluation of subclasses results. The addition of counts makes it easier to spot when the results are skewed by small numbers.

In our charts we will take positive scores as being right of the center line and neutral and negative scores as being left. With our particular question design, we are unable to tell if the neutral responses are substantive or hidden non responses[91].

Grouping Analysis

Our data does not fall under a normal distribution, so we will be using a nonparametric test. Vaus[74] advises that when analysing Ordinal variables with nominal grouping variables then the statistical methods one should use for checking for differences between groupings would be the Mann-Whitney[92] or the Kruskal Wallis[93] tests. In the grouping section we had one group of only 5 participants, the novices. As this group is too small to think of using in analysis we will ignore it. This means that all of the groupings we have will be divided in two. As Mann-Whitney is a specialisation of Kruskal Wallis for two groups, this is the analysis we will use.

For all our analyzes our null hypothesis, H_0 , is that there is no difference between the ranks of the two groups. Our alternative hypothesis, H_1 , is there is a difference between the ranks of the two groups. Our alpha, or significance, level will be 0.05, for no other reason than it seems to be a mutually agreed upon value within the statistical community, and justifying a different significance would take more time than is warranted for the value these analyzes will give. Our sample size is greater than 20, thus we can use a Z distribution.

To work out our Z score, we have an alpha level of 0.05 and a 2 tailed test. We could work the distribution out by using the following formula:

$$U_1 = n_1 n_2 + \frac{n_1(n_1 + 1)}{2} - R_1$$

However it is easier to look up in a Z table. Thus, we have an area in body of 0.9750, which correlates in the Z Table to a z value of 1.96. Thus our decision rule is, If our z is less than -1.96, or greater than 1.96, we reject our null hypothesis.¹

¹We carried out the calculations using the Mann-Whitney U Test calculator at <https://www.socscistatistics.com/tests/mannwhitney/default2.aspx>

Question 1, 2 and 3: First Impressions

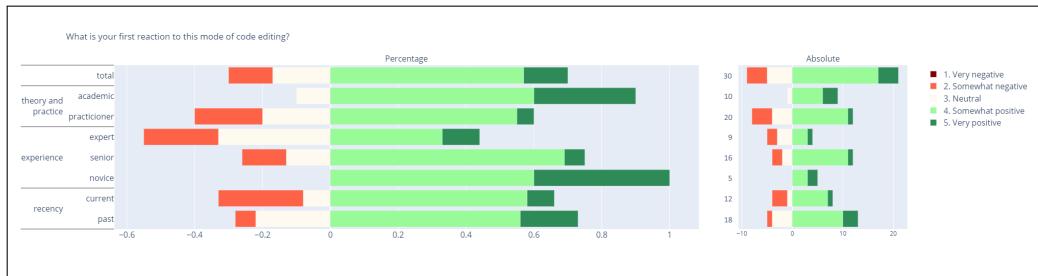


FIGURE 4.23: Question 1 - first impressions

| Group Comparison | Critical U | U-value | z-Score | p-value | Hypothesis |
|--------------------------|------------|---------|----------|---------|----------------------|
| Academic vs Practitioner | 55 | 54.5 | -1.97974 | 0.0477 | H₁ |
| Expert vs Senior | 37 | 55 | 0.93413 | 0.35238 | H ₀ |
| Current vs Past | 61 | 91 | 0.6985 | 0.48392 | H ₀ |

TABLE 4.3: Mann-Whitney question 1 - first impressions

This question shows the subject an example of projectional editing a Drools file alongside a table projection, as an animated GIF, along with an explanation. Then she is asked her first reaction. The chart in figure 4.23 shows the outcomes. Eyeballing the chart, an observation that can be taken is that the novice and the academics (where there is a lot of crossover), found the initial presentation more positive than the experienced practitioners.

Table 4.3, shows that there is a significant difference between how academics and practitioners first react to seeing the projectional editing example. Here, the academics have a far more positive view of the example than the practitioners.

In general, we see an overwhelmingly positive response. There were 5 times as many positive (21) then negative (4) responses. Those who had a positive or negative responses were directed to answer the open questions “Q2. How would this coding style be useful to your interactions with Drools?” and “Q3. What do you find negative with this style of coding?” respectively. Figure 4.24 show a not very useful, but funky looking visualization of the subjects responses.

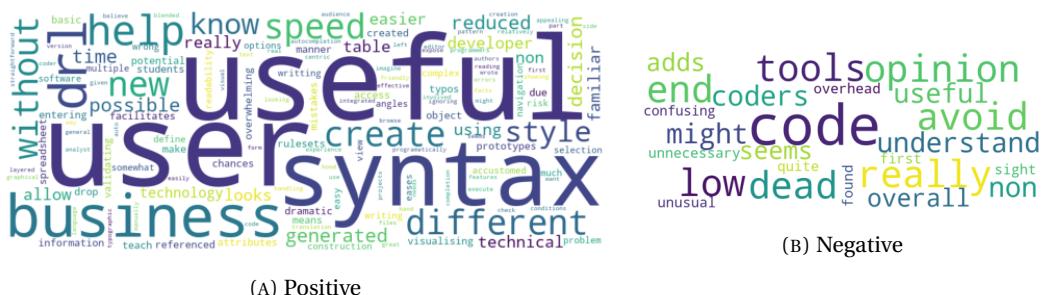


FIGURE 4.24: Initial thoughts

Amongst the positive comments it appears the the subjects had all picked up on many of the advantages that projectional editing brings. The ones that got the most mentions,

using other words, were exploratory coding, correctness by construction, and multiple viewpoints. It was also noted, with the projections shown, that development could be quicker and easier to check.

Amongst the, very few, negative comments, they discussed the failures of no/low code solutions, that the view was confusing and they felt it added unnecessary overhead.

Question 4 and 5: Interpret Projection

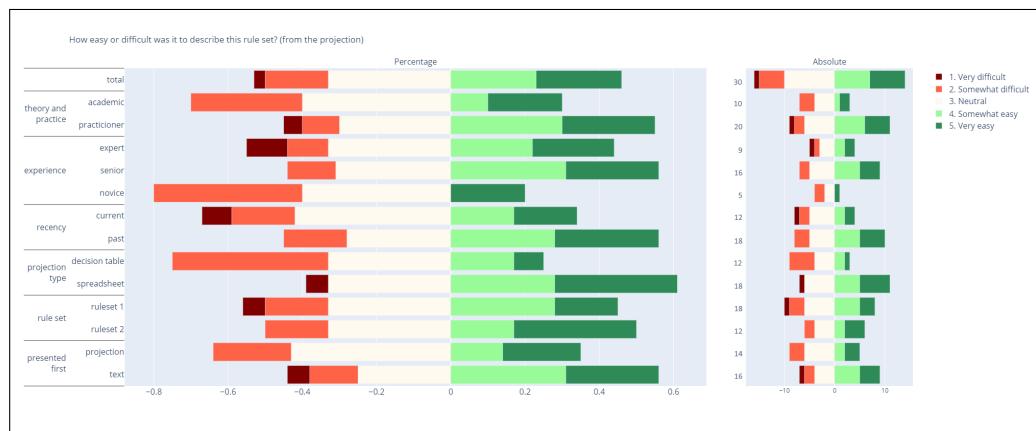


FIGURE 4.25: Question 5 - interpret projection

| Group Comparison | Critical U | U-value | z-Score | p-value | Hypothesis |
|--------------------------------|------------|---------|---------|---------|------------|
| Academic vs Practitioner | 55 | 77 | 0.98987 | 0.32218 | H_0 |
| Expert vs Senior | 37 | 61.5 | 0.56614 | 0.56868 | H_0 |
| Current vs Past | 61 | 82.5 | 1.05833 | 0.28914 | H_0 |
| Decision Table vs Spreadsheet | 55 | 61 | 2.2225 | 0.02642 | H_1 |
| Ruleset 1 vs Ruleset 2 | 61 | 92 | 0.65617 | 0.50926 | H_0 |
| Projection first vs Text first | 64 | 97 | 0.60277 | 0.5485 | H_0 |

TABLE 4.4: Mann-Whitney question 5 - interpret projection

This question asks the subject to describe the meaning of the projection and then describe how hard it was to do that. Very few people described the meaning of the projection well.

The chart in figure 4.25 shows the outcomes. Looking at the chart, it appears obvious that there is a difference in confidence of the subjects between the different projections presented. The participants believed they understood the spread sheet style projection better than the decision table projection.

This observation is confirmed in the Mann-Whitney analysis, shown in table 4.4. Otherwise, no other grouping differed significantly from the general population.

The ratio of those thinking it was easy to those thinking it was hard to understand the projection was 2:1, 14 to 6.

There was one extreme response, which was negative. This came from an expert practitioner who saw the text version before seeing the projection.

Question 6 and 7: Interpret Text

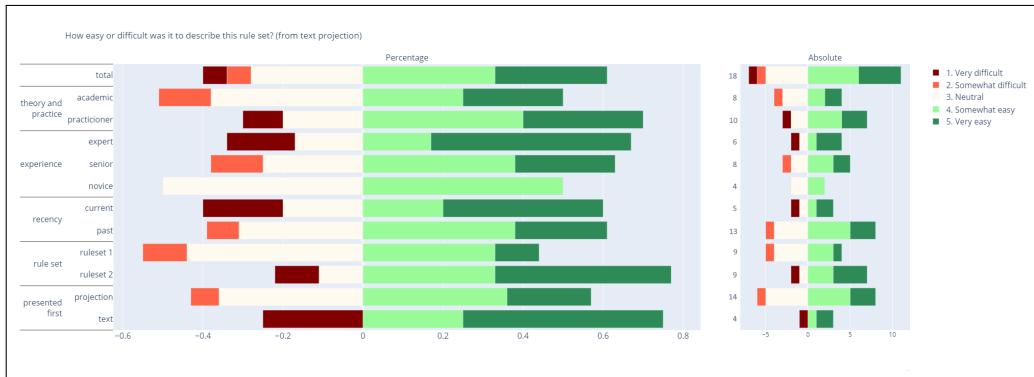


FIGURE 4.26: Question 7 - interpret text

| Group Comparison | Critical U | U-value | z-Score | p-value | Hypothesis |
|--------------------------|------------|---------|----------|---------|------------|
| Academic vs Practitioner | 17 | 34 | 0.48869 | 0.62414 | H_0 |
| Expert vs Senior | 8 | 20.5 | -0.3873 | 0.69654 | H_0 |
| Current vs Past | 12 | 31.5 | -0.04929 | 0.96012 | H_0 |
| Ruleset 1 vs Ruleset 2 | 17 | 24.5 | -1.36868 | 0.17068 | H_0 |

TABLE 4.5: Mann-Whitney question 7 - interpret text

This question asks the subject to interpret a rule set that is presented in a Drools style text projection. The purpose of this question was three fold. First, to calibrate how well the subject really understood Drools. Second, for a comparison with with a later projection. Finally, to calibrate whether and how much easier the text was than the projection to those used to seeing the text version.

Unfortunately, we assume due to the questionnaire design, there was a large number of non-respondents to this question. 12 of the 30 respondents did not answer this question. All 12 of these were from the 16 that were presented the text projection before the tabular projections.

Reporting on 18 responses has a lot less validity. With that in mind in figure 4.26, we can see a much higher confidence in the easiness of understanding the rule set. We see a 5:1 ratio of greater belief in the participants that it was easy to understand the meaning of the text projections. This proportion is significantly higher than those who thought it was easy to understand the tabular projections.

The Mann-Whitney table does not show any significant differences in any of the groups. We are unable to report on the difference between those who were presented the text projection first or the tabular projection first. The 4 participants responding from the group who were presented text first, was too low to give a meaningful score.

Question 8: Compare Projections

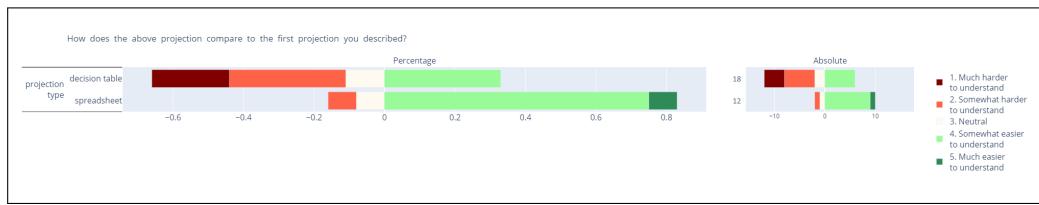


FIGURE 4.27: Question 8 - compare projections

| Group Comparison | Critical U | U-value | z-Score | p-value | Hypothesis |
|-------------------------------|------------|---------|----------|---------|----------------------|
| Decision Table vs Spreadsheet | 61 | 45 | -2.64584 | 0.00804 | H₁ |

TABLE 4.6: Mann-Whitney question 8 - compare projections

This question asked the participant to compare the two tabular projections. This question was to calibrate whether one projection was considerably worse than the other and whether that would affect the comparison with text.

Looking the chart in figure 4.27, we have ignored all bars except the difference between the projections presented. The way the question is constructed the other groupings do not help.

It is obvious in this chart and confirmed in the Mann-Whitney analysis shown in table 4.6, that the spreadsheet presentation is considered far easier to understand than the decision table. This result correlates well with the difference in understanding found previously in question 5.

Question 9: Compare Projection to Text

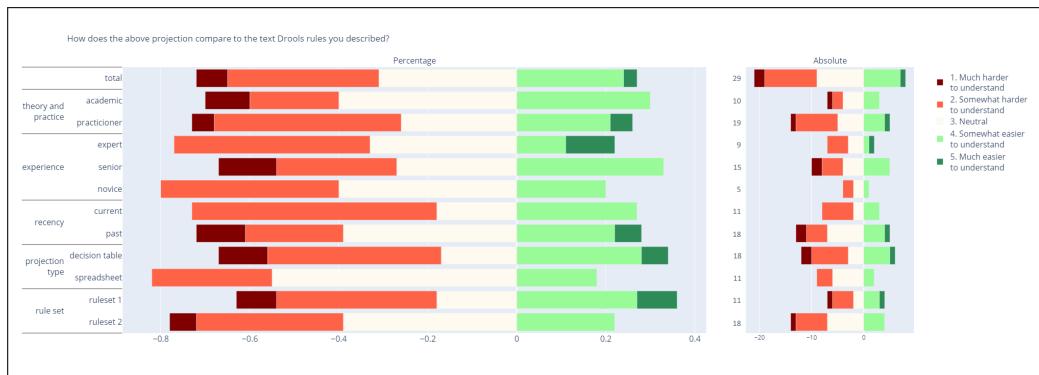


FIGURE 4.28: Question 9 - compare projections with text

| Group Comparison | Critical U | U-value | z-Score | p-value | Hypothesis |
|-------------------------------|------------|---------|----------|---------|------------|
| Academic vs Practitioner | 52 | 85.5 | -0.41295 | 0.6818 | H_0 |
| Expert vs Senior | 34 | 67.5 | 0.02981 | 0.97606 | H_0 |
| Current vs Past | 55 | 88 | 0.47194 | 0.63836 | H_0 |
| Decision Table vs Spreadsheet | 55 | 89.5 | -0.40452 | 0.68916 | H_0 |
| Ruleset 1 vs Ruleset 2 | 55 | 94.5 | -0.17979 | 0.85716 | H_0 |

TABLE 4.7: Mann-Whitney Question 9 - compare projections with text

This question asked our subjects to compare a tabular projection with the text projection. There was one participant that chose not to answer this question.

This result, shown in figure 4.28, was pretty definitive. The subjects found very much that the textual projections were more understandable than the tabular projections. The ratio of harder to easier to understand was 3:2, with 12 subjects finding the Text projection easier to understand and 8 finding the projection easier. Table 4.7 shows this was independent of any other factors.

Question 10: Truth Table Validation

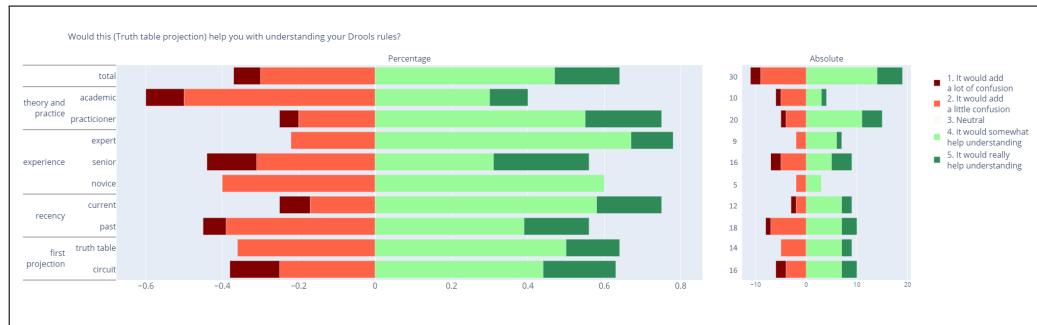


FIGURE 4.29: Question 10 - truth table

| Group Comparison | Critical U | U-value | z-Score | p-value | Hypothesis |
|------------------------------------|------------|---------|----------|---------|------------|
| Academic vs Practitioner | 55 | 65 | 1.5178 | 0.12852 | H_0 |
| Expert vs Senior | 37 | 64 | -0.4246 | 0.67448 | H_0 |
| Current vs Past | 61 | 93 | -0.61383 | 0.54186 | H_0 |
| Truth table first vs Circuit first | 64 | 108.5 | -0.12471 | 0.90448 | H_0 |

TABLE 4.8: Mann-Whitney Question 10 - truth table

This question presented the subject with a wireframe of a truth table and asked if it would help them with understanding.

Every subject had a positive or negative view of the projection. 0 of the 30 subjects gave a neutral result. We found this a very unlikely result.

There was a net positive view of this projection, with a little less than 2:1 finding it more helpful (19) than confusing (11).

On first view of the chart in figure 4.29, it seems that academics have a much more negative view of this projection, and experts a more positive outlook. However, when the figures are examined under Mann-Whitney, as seen in table 4.8, these differences were not statistically significant.

Question 11: Circuit Diagram Validation

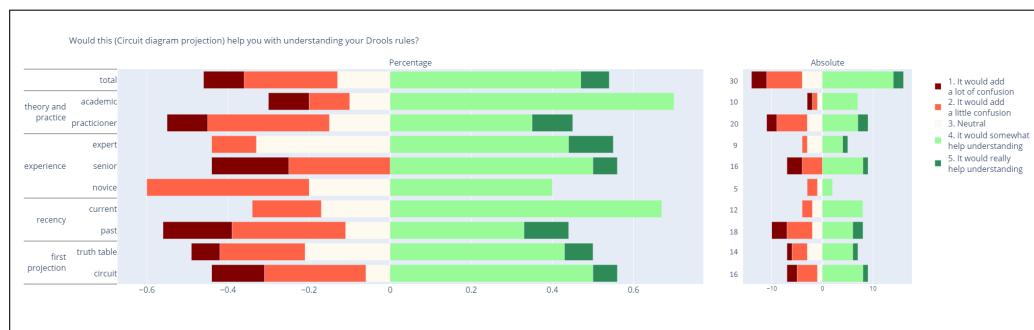


FIGURE 4.30: Question 11 - circuit diagram

| Group Comparison | Critical U | U-value | z-Score | p-value | Hypothesis |
|------------------------------------|------------|---------|----------|---------|------------|
| Academic vs Practitioner | 55 | 83 | -0.7259 | 0.4654 | H_0 |
| Expert vs Senior | 37 | 58.5 | -0.73598 | 0.4593 | H_0 |
| Current vs Past | 61 | 83 | -1.03717 | 0.29834 | H_0 |
| Truth table first vs Circuit first | 64 | 110 | -0.06236 | 0.95216 | H_0 |

TABLE 4.9: Mann-Whitney Question 11 - circuit diagram

This question presented the subject with a wireframe of a circuit diagram and asked if it would help them with understanding.

The view of the circuit diagram was a net positive, with a ratio of 3:2 (16 positive, 10 negative).

Question 15: Closing Remarks

Our last question asks for any last comments. 16 participants decided to add some words.

We ran sentiment analysis on these 16 comments, using the same service and similar code as we used in our SLR research, as described in section 3.1.6. 9 were considered positive, 6 mixed and 1 negative.

Figure 4.31 shows another pretty, but not too analytical word cloud.

The word that pops out, DMN, here is indicative of a few comments. DMN is a tool for working with Drools as a non developer. There were some comments that our research should look in this direction.

The feeling was that the projections would give more advantage to non-developers.

There were some hints as to how to improve our projections or other projections to try. Others suggested that the projections cannot capture some of the complexity of the rules that exist.

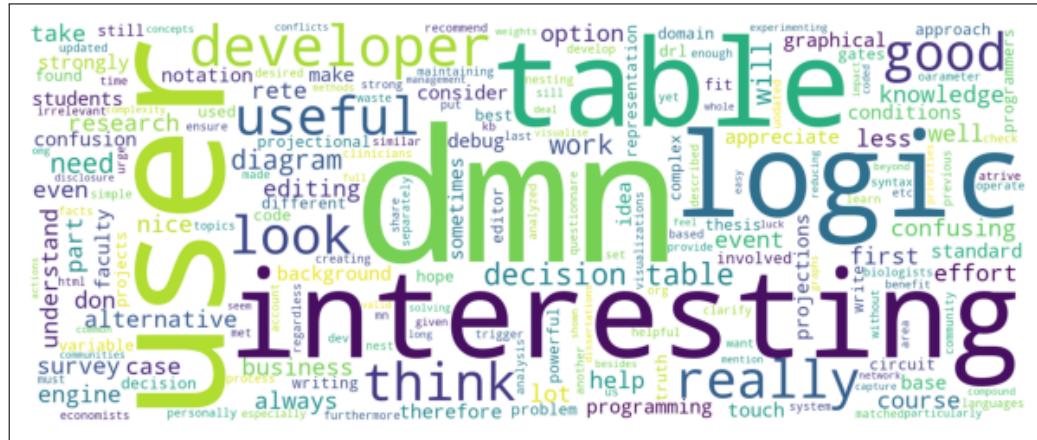


FIGURE 4.31: Question 15 - responses

4.3.4 Summary Analysis

The goal of this survey was to find if users of Drools would find the projections we used useful. The outcome seems to be yes, they seem useful, but not better than the text that they are currently used to.

We tried to control for different groupings that may effect the responses, and in most cases we found no significant differences in the groups we presented.

There was a significant difference between the projections presented, with the sample in general preferring the spreadsheet projection over the decision table. However, this was not what we were measuring for and even that is debatable as in question 9. When comparing text to projection, the outcome for both tabular projections favoured text, but with a slightly lower net negative for the decision table comparison, 7:5 compared to 3:2.

Chapter 5

Discussion

5.1 Discussion - Systematic Literature Review

We now examine threats to the construct, the internal and external validity of our systematic literature review, its reliability, and areas of improvement.

5.1.1 Threats to Validity

As discussed in their tertiary study of SLRs da Silva et al.[94], one of the main problems of SLRs in Software engineering is a focus on practice and not experimentation. Because of the nature of the subject area we will be making this same shortfall. We feel that we have fully addressed their other concerns of SLRs not assessing the quality of our primary studies, bad integration and lack of guidelines.

As with all SLRs the main threats to validity are incomplete set of studies, due to an insufficient search strategy, researcher bias in paper selection and inaccuracy in data extraction. In our study quality assessment we made use of Runeson et al.'s[95] four suggested limitations of studies, namely construct validity, internal validity, external validity, and reliability. It is only fair that we point this towards our own study.

Construct Validity

Regarding construct validity, i.e. whether our research questions match the research subjects methods and measures. Whilst no measurement system is perfect, some are much further from perfect than others.

For the construct to be valid we need to present the best available evidence. The nature and modernity of the Projectional Editing might mean that there is plenty of good evidence available in grey literature and Industrial Articles. This under representation of actual, but non-academic studies could lead to a false positive or negative for some of the questions, leading to errors in recommendations.

There maybe things that influence the best evidence such as who is funding the study. Is a researcher working or consulting at a Projectional Editing product supplier, and will this skew results.

Are some projectional editors being ignored because of the preference for English papers only? The focus on English language papers might be biased against projectional editors aimed at non-English speaking markets.

The use of the sentiment analysis tool may have been inappropriate. The way that scientific papers are written may not lend themselves to sentiment analysis in general or the Azure sentiment analysis service in particular.

Internal Validity

Internal validity, or the causal relationships, does one factor cause an effect or are both factors influenced by something unseen.

An incomplete search term may have lead to selection bias. By using “projectional editing” and “projectional editor” we may have leaned into one small corner of this field, where one tool uses this term whilst others use another to describe a similar approach. Other tools or projects could use different terms, such as “Language orientated programming” or more antiquated terms like “Structured programming” or “syntax directed editing”.

One causal relationship that could very well have influenced the outcome was the publishing of the book, “Domain-Specific Languages in Practice with JetBrains MPS” right at the end of our selection period. Of the 11 papers published in this book, 7 made it into our final paper selection. As this book is about using MPS, then this skews the data towards MPS.

We also had an error with the identification of primary studies. Initially we took the word of the paper when it said it had done a case study, but often times, that just meant trying the code out on a current problem, rather than a case study in the academic sense.

A final threat to internal validity was that the conclusions were not evaluated by an expert in either Drools or Projectional editing who could evaluate their value to the field.

External Validity

External validity is the ability to generalize the findings.

The prevalence of the Action Design research methodology in Software engineering is often a reason that the results are impossible to generalise. The “primary study” data in these cases are often just the source code.

One threat to validity that this is a good representation of the field is our restricted time span. However, as the scope of the question was to examine the current state, we still feel that this restriction is necessary and informative.

Reliability

Reliability is how the data and the analysis is dependant on specific researchers. Here we are presented with a very credible threat in that this research was carried out by a single researcher. Under particular threat of single researchers bias were the quality assessments.

Whilst measures were put in place to try and mitigate this, the reliance on a single person's judgement of the underlying studies, leaves the door to bias wide open. Another threat is the use of narrative review. This can be subjective and therefore difficult to reproduce.

As mentioned in the results, we were not happy with any of the paper quality assessments for ADR studies. Using an inappropriate quality assessment tool, was instrumental in the almost complete failure of the quality assessments.

Repeatability vs Reproducibility

Greenhalgh et al.[96] in their review of where papers come from in systematic reviews found that 24% of papers come from "personal knowledge or personal contact". For the sake of reproducibility we decided not to hunt down relevant papers from knowledgeable people in the field, as this would require anyone trying to reproduce this study to ask the same people at the same knowledge level as us. This would be impossible as the only person we know who could help us with this is our academic advisor, and he (in our future and your present) will have read this paper, changing his knowledge base forever. By making this choice we risk missing out a quarter of our completeness.

Method Improvement

An area we would definitely improve is the sentiment analysis. Our technique was flawed. To avoid the risk of bias through us cherry picking paragraphs, we were very coarse grained in our input selection criteria. For all papers we chose the introduction, the conclusion and then any section that discussed anything to do with projectional editing.

The problem with this is that sometimes the sections would run for paragraphs, with only one or two being about projectional editing. Often times, especially with papers involving MPS, the paper was focused on a problem and using MPS to solve it. Thus the introductions and conclusions would occasionally barely mention projectional editing.

If we were to do this again, we would take a much more fine grained approach as we feel the cherry picking risk is less important than the noise from unrelated text.

5.2 Discussion - Survey

5.2.1 Threats to Validity

Construct Validity

The most prominent issue with this survey is whether it is generalizable to the initial questions it was trying to answer. The question being can projectional editing be used to increase the comprehensibility of large Business rules files. There is a strong case to be made that this questionnaire did not ask the right sort of questions to answer this hypothesis. It is difficult to simulate large business rules files in a brief questionnaire. Comparing the

projections to relatively small and non complex rules collections may not be generalizable to a large rules collection.

We perhaps had an issue with a mono-operation bias, because we only used one tool for measuring - the survey. We performed many techniques, such as changing orders of questions, measurement, rulesets, to try and overcome this. However, all the subjects were still just answering a survey, delivered through the same medium - Survey Monkey.

A possible relationship bias may occur from two fronts, where parties may participate in hypothesis guessing and answer questions based on their outcome. Firstly, two of the respondents were directly known to us, one through work and another through meetings at conferences. This could have colored their responses toward a more positive view of our work. Second, two people who responded were part of the Drools core development team, and thus work for JBoss/RedHat. On top of that all the people who answered through the Drools consultants mailing list possibly had a direct monetary relationship with RedHat. This could influence their response towards a more positive view of the status quo.

Internal Validity

Whilst we felt we had tried to overcome the apprehension of being judged when answering questions, the fact that one of our questions ask the subject to describe a Drools rule's meaning, could overrule all of our previous attempts to assuage that fear.

As with all surveys, the higher the sample size the greater the chance of validity. It is difficult to measure the validity of our outcomes, due to the relatively small size of our survey. Whilst 30 respondents is on the low side for using statistical tools to give reasonable responses, it is also dependant on size of population. We requested the size of the population of Drools users from a member of the core Drools development team, but apparently that number was not known. The cross-factor comparisons between the subgroups within is of dubious validity as their sample sizes were so small.

External Validity

Whilst we feel we reached the right audience of tool users, there was a potential for a geographical selection bias in our population selection technique. Because a portion of our respondents came from our connections on LinkedIn, and we are from Europe, then there is a particular European bias to our respondents. Only four of our completed surveys were not from Europe, three if using the Eurovision song contest definition of Europe.

There is also a self-selection bias in the sampling. As this survey is voluntary, and there is no real personal connection between the subjects and us, then the people who would answer this question are the sort of people who would answer an unsolicited questionnaire. This bias may effect generalizability to novices, as there was a tendency towards experts in answering the survey.

Reliability

Much like our SLR analysis, the we are here also presented with the credible threat of a single researcher. All surveys have a subjective nature in their scoring. Our measurement

did not take into account cultural differences between, for example, the blunt Dutch and the polite Italians.

Repeatability vs Reproducibility

Repeatability is hampered by the specification of the tool in use. However, we feel that our survey could be valid with a different underlying tool and our population selection could work for other researchers.

Method improvement

We feel our largest problem was smallness. If we were to try this again we would have tried various measures to increase our population size. These would include contacting people directly within LinkedIn, rather than only those whose email or twitter address we could harvest. Also we would have followed up the partially completed questionnaires. Another option would be a larger time period for academic papers.

To overcome our mono-operation bias, we could interview people in person.

Chapter 6

Related Work

We split related work into three sections. Firstly, those relating to the state of projectional editing. Next, those that address understandability of business rules. Finally, work on interesting projections.

6.1 The state of projectional editing workbenches

Whilst we were particularly looking into the current state of projectional editing, there have been previous studies. The language workbench challenges, which ran from 2011 to 2016, inspired many papers, most notably Erdweg et al.'s summaries of the 2013[97] and 2015[20] challenges. These papers looked into the capabilities of LWBs including projectional LWBs. Whilst these were interesting from the point of view of capabilities of LWBs, they did not touch on areas such as market penetration or other indicator of usage, and thus the experimental were lined up against the well used. Schindler et al.[59] looked at the experience of the language workbench challenge from the point of view of a projectional editing LWB, namely MPS. This paper addressed the areas where projectional editing LWBs had a significant advantage over the other LWBs. We feel it is a shame that the LWB challenge is no longer occurring.

There does not seem to be much analysis of what is going on in projectional editing at the moment, except when there is an intersection with adjacent fields such as Model Driven Approaches, Low-code, Language Orientated Programming, or language workbenches. We found one systematic mapping study of LWBs[98], from 2020, that again looked at the entire field of LWBs including the projectional ones. Whilst looking at LWB features it also extended into areas such as domains of use.

With regards to MPS in particular, while there is no mapping of the state of the usage, we saw that papers earlier in the last decade tended to concentrate on how to improve the experience of using MPS, prototypes or new products, for example [38–40, 44, 48, 50, 53, 55, 57, 58, 99]. Today, research involving MPS tends to be maturing into industrial use of the product and using it in teaching, e.g. [60, 109, 124–126, 132, 100].

6.2 Understandability of Business Rules

A different approach to ours was taken in the VODRE project[101]. In place of visualising how the rules belonged together, it visualized how they executed. In the domain of optical

design, they showed the execution paths that the expert systems used in coming to their conclusions.

In a similar fashion to the Drools DMN, Ostermayer et al.[102], attempt the generation of rules using templating and an external editor. Along the same lines, the G-AMC tool[103] created a front end for developing rules, however in this case in the domain of access control management. These both face the issues we were trying to overcome with a separation between the tool and the language. They are also focused on the generation of rules rather than the understanding of them.

6.3 Interesting projections

Some of the most interesting projections are happening in industrial environments, and there are limited papers about these. For example Klaus Birk[104] uses the graphical modelling extensions for hierarchical component models in systems architectures, as well as getting live simulator responses to changes in the code. Tom Beadman[105], as a part of his tool Pasta, generates an interactive visual state machine to, for example, decide whether to check for fraud.

Florian Bocks tool, stiEF[106], is a language that uses a live update of graphical driving scenarios in the IDE. Whilst this is not a 2 way projection, in that a developer cannot interact with the visual scenarios. However it is really a great feedback tool to let you know that the developer has correctly programmed their scenario.

Chapter 7

Conclusion

In this master's project we dove into the current state of projectional editing and presented and had verified a number of projections for the purpose of improving the understanding of Business Rules. We conducted a systematic literature review (SLR), creates prototypes and conducted a survey of the impact they may have. Through the SLR we examined the current state of projectional editing, finding it dominated by a single product. We implemented the Drools language in MPS and through our prototypes we demonstrated some of the possibilities projectional editing can bring. The results of our survey failed to confirm that these projections would bring the benefits we had hypothesised.

In this paper we described our work with first translating the Drools DSL into a projectional language followed by our explorations of projections. We discussed the advantages and disadvantages of the different projections we created and analysed experienced developers reactions to them.

In this last section of this paper will return to the initial questions we were trying to answer from section 1.2. We will summarise the findings of our research and the contributions our work had made. Finally we will discuss items that were out of scope for this project, but would give future research opportunities.

7.1 Research Summary

Our main research question, "How can projectional editors and DSLs be combined to address feedback mechanisms for developers in the context of reasoning about rules in a rule-based business engine?", motivated us to explore the field of projectional editing in this project. This drove us to investigate if this field was worth investigating and, if so, how we can use existing tools to answer this.

Research Question 1: "What is the current state of language workbenches supporting projectional editing?"

In this paper we presented an SLR which looked into the current state of projectional editing. This SLR found that currently LWBs for projectional editing is a very narrow field, dominated by one product, JetBrains MPS. Studies using products other than MPS spend time and effort discussing how to solve issues already solved in MPS. Studies using MPS tend to be focused on products for industrial use or how to use MPS as a tool to teach

Meta-Programming. One area MPS is a little behind is the need for a web based development environment, both for the Language engineer, but especially for the developer.

A monoculture can be a risk. Whilst there are many advantages to projectional editing, having only one successful product and supplier feels a little unhealthy. On the other hand, the number of users of the tool is growing, as evidenced by the papers representing new projects in multiple industries, from hardware, through automotive industry, to Finance.

To conclude, the state of the projectional editing market, whilst niche, is maturing but with a current product monoculture.

Research Question 2: “Which projections can help developers get appropriate feedback about rules?”

In our research we were able to develop a number of projections. This success was in large part facilitated by the flexibility and extensibility of the MPS tool which presented the ability to develop and extend DSLs very efficiently.

We presented two of the projections to experienced Drools users, along with two wireframes of more exotic solutions. There was a distinct preference between the projections we presented, with the spreadsheet like table being more understandable than our decision table. Both of our presentations significantly underperformed in terms of understanding to the textual projection.

Whilst we were not able to show an advantage of our projections in our study, there was distinct interest in our approach, and recognition of the advantages it may bring.

7.2 Summary of contributions

We think that this masters project offers the following contributions:

- *The Projectional Editing Status Quo* in our SLR we provide an overview of the current state of research in the field of Projectional Editing
- *An alternative Base Language* with a little extra work, the Drools-Lite language can be used as an alternative base language for model-to-model transformations in MPS.
- *A new way to enter rules* our implementation allows Drools developers a more compact manner to enter rules.

7.3 Further Work

Drools-lite is open for future extension, both for matching the feature sets of the full Drools language and adding extra projectional capabilities. Projections we would particularly like to see how we could take advantage of include graphical mappings of rule interactions as well as live test output.

Understanding business rules could be helped by guarantees of completeness. We would be interested to see how we could apply the formal specifications developed within FASTEN[32] to some of our potential projections.

Whilst a survey was informative, we would also like to run an experiment with our projections, to have more than just opinions.

Finally, a question that has occurred to us both in the literature review and the survey was - how much of the opinions about projectional editing are a consequence of a history of textual language use? We also briefly worked on a projectional implementation of the pedagogical gradual language Hedy[107]. Being able to compare the results of those with no experience of programming to use of projectional and text based languages could point to the influence of experience in choice. This will have to wait until there is a web-based implementation of MPS, as the installation of an application on school machines is impractical.

Bibliography

- [1] C. L. Forgy, “Rete: A fast algorithm for the many pattern/many object pattern match problem,” in *Readings in Artificial Intelligence and Databases*, Elsevier, 1989, pp. 547–559.
- [2] G. A. Miller, “The magical number seven, plus or minus two: Some limits on our capacity for processing information.,” *Psychological review*, vol. 63, no. 2, p. 81, 1956.
- [3] C. J. Date, *What not how: the business rules approach to application development*. Addison-Wesley Professional, 2000.
- [4] M. Fowler, *Should i use a rules engine?* <https://martinfowler.com/bliki/RulesEngine.html>, Accessed: 2021-07-18, 2009.
- [5] P. Jackson, “Introduction to expert systems,” 1986.
- [6] E. H. Shortliffe, “Mycin: A rule-based computer program for advising physicians regarding antimicrobial therapy selection.,” STANFORD UNIV CALIF DEPT OF COMPUTER SCIENCE, Tech. Rep., 1974.
- [7] *CLIPS product page*, <http://www.clipsrules.net/>, Accessed: 2021-07-17.
- [8] *Drools product page*, <https://www.drools.org/>, Accessed: 2021-07-17.
- [9] *BizTalk product page*, <https://docs.microsoft.com/en-gb/biztalk/>, Accessed: 2021-07-17.
- [10] *IBM WebSphere JRules product page*, <https://www.ibm.com/docs/en/iis/11.7?topic=applications-websphere-ilog-jrules>, Accessed: 2021-07-17.
- [11] *OpenRules product page*, <https://openrules.com/>, Accessed: 2021-07-17.
- [12] D. Sottara, P. Mello, and M. Proctor, “A configurable rete-oo engine for reasoning with different types of imperfect information,” *IEEE Transactions on Knowledge and Data Engineering*, vol. 22, no. 11, pp. 1535–1548, 2010.
- [13] V. Donzeau-Gouge, G. Huet, G. Kahn, and B. Lang, “Programming environments based on structured editors: The mentor experience,” Tech. Rep., 1980.
- [14] R. Medina-Mora and P. H. Feiler, “An incremental programming environment,” *IEEE Transactions on Software Engineering*, no. 5, pp. 472–482, 1981.
- [15] D. Notkin, “The gandalf project,” *The Journal of systems and software*, vol. 5, no. 2, pp. 91–105, 1985.
- [16] T. Teitelbaum and T. Reps, “The cornell program synthesizer: A syntax-directed programming environment,” *Communications of the ACM*, vol. 24, no. 9, pp. 563–573, 1981.
- [17] T. W. Reps and T. Teitelbaum, *The Synthesizer Generator: A system for constructing language-based editors*. Springer Science & Business Media, 2012.
- [18] C. Simonyi, “The death of computer languages, the birth of intentional programming,” in *NATO Science Committee Conference*, Citeseer, 1995, pp. 17–18.

- [19] S. Dmitriev, "Language oriented programming: The next programming paradigm," *JetBrains onboard*, vol. 1, no. 2, pp. 1–13, 2004.
- [20] S. Erdweg, T. Van Der Storm, M. Völter, L. Tratt, R. Bosman, W. R. Cook, A. Gerritsen, A. Hulshout, S. Kelly, A. Loh, *et al.*, "Evaluating and comparing language workbenches: Existing results and benchmarks for the future," *Computer Languages, Systems & Structures*, vol. 44, pp. 24–47, 2015.
- [21] *Más post mortem*, <https://medium.com/@dslmeinte/post-mortem-for-m%C3%A1is-aeca7542c4c8>, Accessed: 2021-09-07.
- [22] *Whole Platform*, <https://whole.sourceforge.io/>, Accessed: 2021-09-07.
- [23] D. H. Lorenz and B. Rosenan, "Cedalion: A language for language oriented programming," in *Proceedings of the 2011 ACM international conference on Object oriented programming systems languages and applications*, 2011, pp. 733–752.
- [24] B. Hempel, J. Lubin, G. Lu, and R. Chugh, "Deuce: A lightweight user interface for structured editing," in *Proceedings of the 40th International Conference on Software Engineering*, 2018, pp. 654–664.
- [25] *Google Blockly product page*, <https://developers.google.com/blockly/>, Accessed: 2021-09-08.
- [26] J. Klimeš, "Domain-specific language for learning programming," 2016.
- [27] D. Ratiu, V. Pech, and K. Dummann, "Experiences with teaching mps in industry: Towards bringing domain specific languages closer to practitioners," in *2017 ACM/IEEE 20th International Conference on Model Driven Engineering Languages and Systems (MODELS)*, IEEE, 2017, pp. 83–92.
- [28] M. Völter and E. Visser, "Language extension and composition with language workbenches," in *Proceedings of the ACM international conference companion on Object oriented programming systems languages and applications companion*, 2010, pp. 301–304.
- [29] M. Voelter, J. Siegmund, T. Berger, and B. Kolb, "Towards user-friendly projectional editors," in *International Conference on Software Language Engineering*, Springer, 2014, pp. 41–61.
- [30] M. Hosseinkord, G. Dulai, N. Osmani, and C. K. Anand, "Code and structure editing for teaching: A case study in using bibliometrics to guide computer science research," *arXiv preprint arXiv:2107.09038*, 2021.
- [31] M. Voelter, S. Benz, C. Dietrich, B. Engelmann, M. Helander, L. C. Kats, E. Visser, G. Wachsmuth, *et al.*, *DSL engineering: Designing, implementing and using domain-specific languages*. dslbook. org, 2013.
- [32] D. Ratiu, M. Gario, and H. Schoenhaar, "Fasten: An open extensible framework to experiment with formal specification approaches," in *2019 IEEE/ACM 7th International Conference on Formal Methods in Software Engineering (FormaliSE)*, IEEE, 2019, pp. 41–50.
- [33] D. Ratiu, B. Schaetz, M. Voelter, and B. Kolb, "Language engineering as an enabler for incrementally defined formal analyses," in *2012 First International Workshop on Formal Methods in Software Engineering: Rigorous and Agile Approaches (Form-SERA)*, IEEE, 2012, pp. 9–15.
- [34] T. Berger, M. Völter, H. P. Jensen, T. Dangprasert, and J. Siegmund, "Efficiency of projectional editing: A controlled experiment," in *Proceedings of the 2016 24th ACM SIGSOFT International Symposium on Foundations of Software Engineering*, 2016, pp. 763–774.

- [35] P. Vysok, P. Parizek, and V. Pech, “Ingrid: Creating languages in mps from antlr grammars,” 2018.
- [36] V. Pech, “Jetbrains mps: Why modern language workbenches matter,” in *Domain-Specific Languages in Practice*, Springer, 2021, pp. 1–22.
- [37] M. Voelter and S. Lisson, “Supporting diverse notations in mps’projectional editor.,” in *GEMOC MODELS*, 2014, pp. 7–16.
- [38] M. Voelter and K. Solomatov, “Language modularization and composition with projectional language workbenches illustrated with mps,” *Software Language Engineering, SLE*, vol. 16, no. 3, 2010.
- [39] M. Voelter, A. v. Deursen, B. Kolb, and S. Eberle, “Using c language extensions for developing embedded software: A case study,” in *Proceedings of the 2015 ACM SIGPLAN International Conference on Object-Oriented Programming, Systems, Languages, and Applications*, 2015, pp. 655–674.
- [40] M. Voelter, “Embedded software development with projectional language workbenches,” in *International Conference on Model Driven Engineering Languages and Systems*, Springer, 2010, pp. 32–46.
- [41] S. M. Guttormsen, A. Prinz, and T. Gjøsæter, “Consistent projectional text editors.,” in *MODELSWARD*, 2017, pp. 515–522.
- [42] M. Voelter and B. Merkle, “Domain specific: A binary decision?” In *Proceedings of the 10th Workshop on Domain-Specific Modeling*, 2010, pp. 1–6.
- [43] A. Wortmann and M. Beet, “Domain specific languages for efficient satellite control software development,” *OHB System AG*, 2016.
- [44] M. Voelter, D. Ratiu, B. Kolb, and B. Schaetz, “Mbeddr: Instantiating a language workbench in the embedded software domain,” *Automated Software Engineering*, vol. 20, no. 3, pp. 339–390, 2013.
- [45] C. Simonyi, M. Christerson, and S. Clifford, “Intentional software,” in *Proceedings of the 21st annual ACM SIGPLAN conference on Object-oriented programming systems, languages, and applications*, 2006, pp. 451–464.
- [46] M. Voelter, T. Szabó, S. Lisson, B. Kolb, S. Erdweg, and T. Berger, “Efficient development of consistent projectional editors using grammar cells,” in *Proceedings of the 2016 ACM SIGPLAN International Conference on Software Language Engineering*, 2016, pp. 28–40.
- [47] V. Pech, A. Shatalin, and M. Voelter, “Jetbrains mps as a tool for extending java,” in *Proceedings of the 2013 International Conference on Principles and Practices of Programming on the Java Platform: Virtual Machines, Languages, and Tools*, 2013, pp. 165–168.
- [48] M. Voelter, J. Warmer, and B. Kolb, “Projecting a modular future,” *IEEE Software*, vol. 32, no. 5, pp. 46–52, 2014.
- [49] D. Ratiu, H. Nehls, and J. Michel, “Taming the software development complexity with domain specific languages,” *Modellierung 2018*, 2018.
- [50] M. Voelter, Z. Molotnikov, and B. Kolb, “Towards improving software security using language engineering and mbeddr c,” in *Proceedings of the Workshop on Domain-Specific Modeling*, 2015, pp. 55–62.
- [51] M. Voelter, B. Kolb, K. Birken, F. Tomassetti, P. Alff, L. Wiart, A. Wortmann, and A. Nordmann, “Using language workbenches and domain-specific languages for safety-critical software development,” *Software & Systems Modeling*, vol. 18, no. 4, pp. 2507–2530, 2019.

- [52] S. Meacham, V. Pech, and D. Nauck, "Adaptivevle: An integrated framework for personalized online education using mps jetbrains domain-specific modeling environment," *IEEE Access*, vol. 8, pp. 184 621–184 632, 2020.
- [53] D. Pavletic, S. A. Raza, M. Voelter, B. Kolb, and T. Kehrer, "Extensible debuggers for extensible languages," *GI/ACM WS on Software Reengineering*, 2013.
- [54] M. Voelter, "Language and ide modularization and composition with mps," in *International Summer School on Generative and Transformational Techniques in Software Engineering*, Springer, 2011, pp. 383–430.
- [55] D. Ratiu, M. Voelter, Z. Molotnikov, and B. Schaetz, "Implementing modular domain specific languages and analyses," in *Proceedings of the Workshop on Model-Driven Engineering, Verification and Validation*, 2012, pp. 35–40.
- [56] M. Voelter, D. Ratiu, B. Schaetz, and B. Kolb, "Mbeddr: An extensible c-based programming language and ide for embedded systems," in *Proceedings of the 3rd annual conference on Systems, programming, and applications: software for humanity*, 2012, pp. 121–140.
- [57] M. Voelter and E. Visser, "Product line engineering using domain-specific languages," in *2011 15th International Software Product Line Conference*, IEEE, 2011, pp. 70–79.
- [58] M. Voelter, D. Ratiu, and F. Tomassetti, "Requirements as first-class citizens: Integrating requirements directly with implementation artifacts," in *Proceedings of ACES-MB Workshop*, Citeseer, 2013.
- [59] E. Schindler, K. Schindler, F. Tomassetti, and A. Sutii, "Language workbench challenge 2016: The jetbrains meta programming system," in *LWC SLE 2016 Language Workbench Challenge, Splash2016, 39 October-4 November 2016, Amsterdam, The Netherlands*, 2016.
- [60] A. Prinz, "Teaching language engineering using mps," in *Domain-Specific Languages in Practice*, Springer, 2021, pp. 315–336.
- [61] F. Tomassetti and V. Zaytsev, "Reflections on the lack of adoption of domain specific languages.,," in *STAF Workshops*, 2020, pp. 85–94.
- [62] M. Fowler, *Language workbenches: The killer-app for domain specific languages?* <http://martinfowler.com/articles/languageWorkbench.html>, Accessed: 2021-02-02, 2005.
- [63] B. A. Kitchenham, D. Budgen, and P. Brereton, *Evidence-Based Software Engineering and Systematic Reviews*. Chapman & Hall/CRC, 2015, ISBN: 1482228653.
- [64] H. Zhang, M. A. Babar, and P. Tell, "Identifying relevant studies in software engineering," *Inf. Softw. Technol.*, vol. 53, no. 6, pp. 625–637, 2011.
- [65] C. Wohlin, "Guidelines for snowballing in systematic literature studies and a replication in software engineering," in *Proceedings of the 18th International Conference on Evaluation and Assessment in Software Engineering*, Association for Computing Machinery, 2014.
- [66] J. Cohen, "A coefficient of agreement for nominal scales," *Educational and Psychological Measurement*, vol. 20, no. 1, pp. 37–46, 1960.
- [67] A. O'Mara-Eves, J. Thomas, J. McNaught, M. Miwa, and S. Ananiadou, "Using text mining for study identification in systematic reviews: A systematic review of current approaches," *Systematic Reviews*, vol. 4, 2015.

- [68] E. Akl, D. Altman, P. Aluko, L. Askie, D. Beaton, J. Berlin, B. Bhaumik, C. Birmingham, M. Boers, A. Booth, I. Boutron, S. Brennan, M. Briel, S. Briscoe, J. Busse, D. Caldwell, M. Cargo, A. Carrasco-Labra, A. Chaimani, and C. Young, *Cochrane Handbook for Systematic Reviews of Interventions*. Oct. 2019.
- [69] B. Kitchenham, “Procedures for performing systematic reviews,” *Keele, UK, Keele Univ.*, vol. 33, Aug. 2004.
- [70] S. Gregor, “The nature of theory in information systems,” *MIS quarterly*, pp. 611–642, 2006.
- [71] B. A. Kitchenham, D. Budgen, and P. Brereton, *Evidence-based software engineering and systematic reviews*. CRC press, 2015, vol. 4.
- [72] Axure product page, <https://www.axure.com/>, Accessed: 2021-08-23.
- [73] A. Bryman, *Social research methods*. Oxford university press, 2016.
- [74] D. De Vaus and D. de Vaus, *Surveys in social research*. Routledge, 2013.
- [75] R. Likert, “A technique for the measurement of attitudes.” *Archives of psychology*, 1932.
- [76] J. S. Armstrong and T. S. Overton, “Estimating nonresponse bias in mail surveys,” *Journal of marketing research*, vol. 14, no. 3, pp. 396–402, 1977.
- [77] P. M. Podsakoff, S. B. MacKenzie, J.-Y. Lee, and N. P. Podsakoff, “Common method biases in behavioral research: A critical review of the literature and recommended remedies.,” *Journal of applied psychology*, vol. 88, no. 5, p. 879, 2003.
- [78] M. K. Lindell and D. J. Whitney, “Accounting for common method variance in cross-sectional research designs.,” *Journal of applied psychology*, vol. 86, no. 1, p. 114, 2001.
- [79] G. Gorrell, N. Ford, A. Madden, P. Holdridge, and B. Eaglestone, “Countering method bias in questionnaire-based user studies,” *Journal of Documentation*, 2011.
- [80] N. J. Goldstein, S. J. Martin, and R. Cialdini, *Yes: 50 scientifically proven ways to be persuasive*. Simon and Schuster, 2008.
- [81] I. K. Crombie and B. J. Harvey, “The pocket guide to critical appraisal: A handbook for health care professionals,” *Canadian Medical Association. Journal*, vol. 157, no. 4, p. 448, 1997.
- [82] M. F. Cowlishaw, “Lexx—a programmable structured editor,” *IBM Journal of Research and Development*, vol. 31, no. 1, pp. 73–80, 1987.
- [83] A. Sarkar, “The impact of syntax colouring on program comprehension.,” in *PPIG*, 2015, p. 8.
- [84] D. Hou and D. M. Pletcher, “Towards a better code completion system by api grouping, filtering, and popularity-based ranking,” in *Proceedings of the 2nd International Workshop on Recommendation Systems for Software Engineering*, 2010, pp. 26–30.
- [85] Y. Yoon, B. A. Myers, and S. Koo, “Visualization of fine-grained code change history,” in *2013 IEEE Symposium on Visual Languages and Human Centric Computing*, IEEE, 2013, pp. 119–126.
- [86] E. W. Dijkstra, “The humble programmer,” *Communications of the ACM*, vol. 15, no. 10, pp. 859–866, 1972.
- [87] U. W. Pooch, “Translation of decision tables,” *ACM Computing Surveys (CSUR)*, vol. 6, no. 2, pp. 125–151, 1974.

- [88] L. Bradley and K. McDaid, "Using bayesian statistical methods to determine the level of error in large spreadsheets.,," in *2009 31st International Conference on Software Engineering-Companion Volume*, IEEE, 2009, pp. 351–354.
- [89] L. Wittgenstein, *Tractatus logico-philosophicus*. Routledge, 2013.
- [90] N. B. Robbins, R. M. Heiberger, *et al.*, "Plotting likert and other rating scales," in *Proceedings of the 2011 joint statistical meeting*, vol. 1, 2011.
- [91] J. Blasius and V. Thiessen, "The use of neutral responses in survey questions: An application of multiple correspondence analysis," *JOURNAL OF OFFICIAL STATISTICS STOCKHOLM-*, vol. 17, no. 3, pp. 351–368, 2001.
- [92] H. B. Mann and D. R. Whitney, "On a test of whether one of two random variables is stochastically larger than the other," *The annals of mathematical statistics*, pp. 50–60, 1947.
- [93] W. H. Kruskal and W. A. Wallis, "Use of ranks in one-criterion variance analysis," *Journal of the American statistical Association*, vol. 47, no. 260, pp. 583–621, 1952.
- [94] F. Q. Da Silva, A. L. Santos, S. Soares, A. C. C. França, C. V. Monteiro, and F. F. Maciel, "Six years of systematic literature reviews in software engineering: An updated tertiary study," *Information and Software Technology*, vol. 53, 2011.
- [95] P. Runeson and M. Höst, "Guidelines for conducting and reporting case study research in software engineering," *Empirical software engineering*, vol. 14, no. 2, 2009.
- [96] T. Greenhalgh and R. Peacock, "Effectiveness and efficiency of search methods in systematic reviews of complex evidence: Audit of primary sources," *BMJ*, vol. 331, no. 7524, pp. 1064–1065, 2005.
- [97] S. Erdweg, T. Van Der Storm, M. Völter, M. Boersma, R. Bosman, W. R. Cook, A. Gerritsen, A. Hulshout, S. Kelly, A. Loh, *et al.*, "The state of the art in language workbenches," in *International Conference on Software Language Engineering*, Springer, 2013, pp. 197–217.
- [98] A. Iung, J. Carbonell, L. Marchezan, E. Rodrigues, M. Bernardino, F. P. Basso, and B. Medeiros, "Systematic mapping study on domain-specific language development tools," *Empirical Software Engineering*, vol. 25, no. 5, pp. 4205–4249, 2020.
- [99] M. Voelter, *Generic tools, specific languages*. Citeseer, 2014.
- [100] D. Ratiu, H. Nehls, A. Joanni, and S. Rothbauer, "Use mps to unleash the creativity of domain experts: Language engineering is a key enabler for bringing innovation in industry," in *Domain-Specific Languages in Practice*, Springer, 2021, pp. 25–52.
- [101] M. Lapaev and M. Kolchin, "Vodore: Visualisation of drools rules execution," in *Proceedings of 15th Conference of Open Innovations Association FRUCT*, IEEE, 2014, pp. 77–84.
- [102] L. Ostermayer, G. Sun, and D. Seipel, "Simplifying the development of rules using domain specific languages in drools," *CHRISTIAN-ALBRECHTS-UNIVERSITAT ZU KIEL*, p. 198, 2013.
- [103] J. Sá, S. Alves, and S. Broda, "The g-acm tool: Using the drools rule engine for access control management," *arXiv preprint arXiv:1611.08547*, 2016.
- [104] K. Birken, *Interactive model-based resource analysis for systems engineers*, <https://www.youtube.com/watch?v=TD3k8TGIaIE>, Accessed: 2021-09-25, 2021.
- [105] T. Beadman, *A journey with mps in fintech: One year late*, <https://www.youtube.com/watch?v=pqNpLoeoTZM>, Accessed: 2021-09-25, 2021.

- [106] F. Bock, *Stief – a textual scenario description language*, <https://www.youtube.com/watch?v=PM82RBef50I>, Accessed: 2021-09-25, 2021.
- [107] F. Hermans, “Hedy: A gradual language for programming education,” in *Proceedings of the 2020 ACM Conference on International Computing Education Research*, 2020, pp. 259–270.

Systematic Review Bibliography

- [108] L.-E. Lafontant and E. Syriani, “Gentleman: A light-weight web-based projectional editor generator,” in *Proceedings of the 23rd ACM/IEEE International Conference on Model Driven Engineering Languages and Systems: Companion Proceedings*, 2020, pp. 1–5.
- [109] M. Voelter, S. Košcejov, M. Riedel, A. Deitsch, and A. Hinkelmann, “A domain-specific language for payroll calculations: A case study at DATEV,” 2021.
- [110] J. Schröpfer, T. Buchmann, and B. Westfechtel, “A framework for projectional multi-variant model editors.,” in *MODELSWARD*, 2021, pp. 294–305.
- [111] J. Schröpfer, B. Westfechtel, and T. Buchmann, “A generic projectional editor for EMF models.,” in *MODELSWARD*, 2020, pp. 381–392.
- [112] A. Buccharone, K. Soysal, and C. Guidi, “A model-driven approach towards automatic migration to microservices,” in *International Workshop on Software Engineering Aspects of Continuous Development and New Paradigms of Software Production and Deployment*, Springer, 2019, pp. 15–36.
- [113] S. Meacham, V. Pech, and D. Nauck, “AdaptiveVLE: An integrated framework for personalized online education using MPS JetBrains domain-specific modeling environment,” *IEEE Access*, vol. 8, pp. 184 621–184 632, 2020.
- [114] L. Andersen, M. Ballantyne, and M. Felleisen, “Adding interactive visual syntax to textual code,” *Proceedings of the ACM on Programming Languages*, vol. 4, no. OOP-SLA, pp. 1–28, 2020.
- [115] L. Addazi and F. Ciccozzi, “Blended graphical and textual modelling for uml profiles: A proof-of-concept implementation and experiment,” *Journal of Systems and Software*, vol. 175, p. 110 912, 2021.
- [116] S. Meacham, V. Pech, and D. Nauck, “Classification algorithms framework (CAF) to enable intelligent systems using JetBrains MPS domain-specific languages environment,” *IEEE Access*, vol. 8, pp. 14 832–14 840, 2020.
- [117] A. L. Furtado, “DSL based approach for building model-driven questionnaires,” in *Enterprise Information Systems: 22nd International Conference, ICEIS 2020, Virtual Event, May 5–7, 2020, Revised Selected Papers*, Springer Nature, 2021, p. 458.
- [118] T. Beckmann, “Efficient editing in a tree-oriented projectional editor,” in *Conference Companion of the 4th International Conference on Art, Science, and Engineering of Programming*, 2020, pp. 215–216.

- [119] D. Kolovos, A. De La Vega, and J. Cooper, “Efficient generation of graphical model views via lazy model-to-text transformation,” in *Proceedings of the 23rd ACM/IEEE International Conference on Model Driven Engineering Languages and Systems*, 2020, pp. 12–23.
- [120] A. Buccharone, A. Cicchetti, and A. Marconi, “Engineering gameful applications with MPS,” in *Domain-Specific Languages in Practice*, Springer, 2021, pp. 227–258.
- [121] D. Ratiu, A. Nordmann, P. Munk, C. Carlan, and M. Voelter, “FASTEN: An extensible platform to experiment with rigorous modeling of safety-critical systems,” in *Domain-Specific Languages in Practice*, Springer, 2021, pp. 131–164.
- [122] J. Schröpfer and T. Buchmann, “Integrating UML and ALF: An approach to overcome the code generation dilemma in model-driven software engineering,” in *International Conference on Model-Driven Engineering and Software Development*, Springer, 2019, pp. 1–26.
- [123] A. L. Santos, “Javardise: A structured code editor for programming pedagogy in java,” in *Conference Companion of the 4th International Conference on Art, Science, and Engineering of Programming*, 2020, pp. 120–125.
- [124] E. Schindler, H. Moneva, J. van Pinxten, L. van Gool, B. van der Meulen, N. Stotz, and B. Theelen, “JetBrains MPS as core DSL technology for developing professional digital printers,” in *Domain-Specific Languages in Practice*, Springer, 2021, pp. 53–91.
- [125] M. Simi, “Learning data analysis with MetaR,” in *Domain-Specific Languages in Practice*, Springer, 2021, pp. 259–290.
- [126] N. Stotz and K. Birken, “Migrating insurance calculation rule descriptions from Word to MPS,” in *Domain-Specific Languages in Practice*, Springer, 2021, pp. 165–194.
- [127] P. Munk and A. Nordmann, “Model-based safety assessment with SysML and component fault trees: Application and lessons learned,” *Software and Systems Modeling*, vol. 19, no. 4, pp. 889–910, 2020.
- [128] A. Buccharone, M. Savary-Leblanc, X. L. Pallec, J.-M. Bruel, A. Cicchetti, J. Cabot, S. Gerard, H. Aslam, A. Marconi, and M. Perillo, “Papyrus for gamers, let’s play modeling,” in *Proceedings of the 23rd ACM/IEEE International Conference on Model Driven Engineering Languages and Systems: Companion Proceedings*, 2020, pp. 1–5.
- [129] M. V. Merino, J. Bartels, M. van den Brand, T. van der Storm, and E. Schindler, “Projecting textual languages,” in *Domain-Specific Languages in Practice*, Springer, 2021, pp. 197–225.
- [130] R. Cuinat, C. Teodorov, and J. Champeau, “SpecEdit: Projectional editing for TLA+ specifications,” in *2020 IEEE Workshop on Formal Requirements (FORMREQ)*, IEEE, 2020, pp. 1–7.
- [131] A. Prinz, “Teaching language engineering using MPS,” in *Domain-Specific Languages in Practice*, Springer, 2021, pp. 315–336.
- [132] M. Barash and V. Pech, “Teaching MPS: Experiences from industry and academia,” in *Domain-Specific Languages in Practice*, Springer, 2021, pp. 293–313.
- [133] B. Hempel and R. Chugh, “Tiny structure editors for low, low prices!(generating guis from tostring functions),” in *2020 IEEE Symposium on Visual Languages and Human-Centric Computing (VL/HCC)*, IEEE, 2020, pp. 1–5.

- [134] E. Negm, S. Makady, and A. Salah, "Towards ontology-based domain specific language for internet of things," in *Proceedings of the 2020 9th International Conference on Software and Information Engineering (ICSIE)*, 2020, pp. 146–151.
- [135] J. Lubin and R. Chugh, "Type-directed program transformations for the working functional programmer," in *10th Workshop on Evaluation and Usability of Programming Languages and Tools (PLATEAU 2019)*, 2020.
- [136] M. Ozkaya and D. Akdur, "What do practitioners expect from the meta-modeling tools? a survey," *Journal of Computer Languages*, 2021.

Appendix A

Protocol Validation Checklist

The protocol will be validated using the checklist in table A.1, which was adapted from Kitchenham's book Evidence-Based Software Engineering and Systematic Reviews[63].

| Components | Questions |
|--------------------------------|--|
| Background | Is the motivation for the review clearly stated and reasonable? |
| Research questions | Do these address a topic of interest to practitioners and/or researchers? Are they clearly stated? |
| Search strategy | Is the strategy justified and is it likely to find the right primary studies without the reviewers having to check or read a large number of irrelevant papers? Has the strategy been validated? |
| Study selection | Are the inclusion/exclusion criteria clearly defined and related to research questions? Is a validation process specified? Is there a process for handling marginal and uncertain papers? Is there a process for managing multiple reports of individual studies? |
| Quality of primary studies | Is it clear that the outcomes will be used in the later stages of the review? Is a validation process specified? Are criteria for assessing quality provided and justified and appropriate to the anticipated primary study types? |
| Data extraction | Does the data to be extracted properly address the research questions? Are the methods of recording the data appropriate for the types of data to be extracted? Is a validation process specified? Are there mechanisms for iteration where data is qualitative and categories are not (or cannot be) fully defined in advance of the extraction? Will the process enable the research questions to be answered? |
| Data aggregation and synthesis | Are the methods proposed for qualitative and quantitative data appropriate? Has consideration been given to combining results across multiple study types? Is the approach to aggregation and synthesis justified concerning appropriate literature? |
| Reporting | Has this been considered? has sufficient attention been paid to the completeness, general interest, validation, traceability, and the limitations of the review? |
| Review management | Are the tools that will be used for managing papers, studies, and data specified and appropriate? Is the management of the many-to-many relationship between papers and studies addressed? |

TABLE A.1: Protocol validation checklist

Appendix B

Systematic Literature Review Log

Search Description

The papers we found with our automatic search can be found in table B.2. There were 100 unique papers found with the search strings across all of the venues when we used the search string and date restrictions. First we excluded those that were not in English or were unavailable to us. This reduced the count to 94. Next we downloaded each of the papers.

Next we skimmed all of these papers to remove any that were obviously unrelated to projectional editing. This reduced the count to 69.

two of the papers were referring to the same study, which reduced the count to 67.

We then excluded all grey literature, i.e. masters projects, proposals and PhD theses, and also books. This brought us down to 51 papers.

At this point we began our quality assessment.

Table description

Table B.2 shows the log of the Systematic literature review.

The First column “Paper Title” is the name of the paper as given by the search engine.

The second column “Lib”, indicates the library or search engine through which it was found. The Libraries are Identified by the Keys in table B.1.

| Key | Search engine/library | Key | Search engine/library |
|-----|-----------------------|-----|-----------------------|
| 1 | Google Scholar | 2 | IEEEExplores |
| 3 | ACM | 4 | BASE |
| 5 | CORE | 6 | Web of Science |
| 7 | Microsoft Academic | 8 | SCOPUS |
| 9 | Semantic Scholar | 10 | SpringerLink |
| 11 | Wiley Online | 12 | Science.gov |

TABLE B.1: Search engine/library key

The third and fourth columns show inclusion and exclusion reasons. As inclusions only rely on one question, “does this paper discuss projectional editing”, the affirmative is indicated by a tick. The exclusion column includes the reason for the exclusion.

The final column “ref”, gives a link to the citations in the separate bibliography for the systematic literature review, found at the end of the Appendices.

| Paper Title | lib | in | exclusion | F# | B# |
|--|---------|----|-------------|----|----|
| “Filmar, assistir e problematizar” – contribuições à aprendizagem de cálculos | 9 | | not English | X | X |
| 20. Internationales Stuttgarter Symposium | 10 | | book | X | X |
| A Domain-Specific Language for Payroll Calculations: a Case Study at DATEV | 1 | ✓ | | 0 | 0 |
| A Domain-Specific Language for Payroll Calculations: An Experience Report from DATEV | 1,10 | ✓ | Duplicate | X | X |
| A Framework for Modernizing Domain-Specific Languages | 1 | ✓ | grey | X | X |
| A Framework for Projectional Multi-variant Model Editors | 1,8 | ✓ | | 0 | 0 |
| A Generic Projectional Editor for EMF Models | 1,7,8,9 | ✓ | | 2 | 0 |
| A language-driven Development framework for simulation components to generate simulated environments | 1 | ✓ | grey | X | X |
| A Model-Driven Approach Towards Automatic Migration to Microservices | 10 | ✓ | | 5 | 0 |
| A survey of Model Driven Engineering in robotics | 1 | ✓ | | 2 | 2 |
| A Survey on the Design Space of End-User Oriented Languages for Specifying Robotic Missions | 1,10 | ✓ | | 1 | 5 |
| A survey on the formalisation of system requirements and their validation | 1 | ✓ | | 0 | 0 |
| A text-based syntax completion method using LR parsing | 1 | | | X | X |
| Activities and costs of re-engineering cloned variants into an integrated platform | 1 | | | X | X |
| AdaptiveVLE: An Integrated Framework for Personalized Online Education Using MPS Jet-Brains Domain-Specific Modeling Environment | 1,2 | ✓ | | 1 | 1 |
| Adding Interactive Visual Syntax to Textual Code | 3 | ✓ | | 3 | 0 |
| An approach to generate text-based IDEs for syntax completion based on syntax specification | 1 | ✓ | | 1 | 0 |
| An MPS implementation for SimpliC | 1 | ✓ | grey | X | X |
| Blended graphical and textual modelling for UML profiles: A proof-of-concept implementation and experiment | 1 | ✓ | | 0 | 0 |
| Block-based syntax from context-free grammars | 1,3,5 | ✓ | | 0 | 2 |
| Bridging the worlds of textual and projectional language workbenches | 1 | ✓ | grey | X | X |
| Classification Algorithms Framework (CAF) to Enable Intelligent Systems Using JetBrains MPS | 2,5 | ✓ | | 4 | 0 |
| Domain-Specific Languages Environment | | | | | |

| Paper Title | lib | in | exclusion | F# | B# |
|---|-----------|----|-------------|----|----|
| Code and Structure Editing for Teaching: A Case Study in using Bibliometrics to Guide Computer Science Research | 1,9 | ✓ | | 2 | 0 |
| CompOS - a Domain-Specific Language for Composing Internet-of-Things Systems | 1,4 | ✓ | grey | X | X |
| Concepts of variation control systems | 1 | ✓ | | 9 | 5 |
| Concise, Type-Safe, and Efficient Structural Diffing | 3 | | | X | X |
| Constructing optimized constraint-preserving application conditions for model transformation rules | 1 | | | X | X |
| Design & Evaluation of an Accessible High-Level Language for Advanced Cryptography | 1 | ✓ | grey | X | X |
| Domain-specific languages for modeling and simulation | 1 | | | X | X |
| Domain-Specific Languages in Practice | 10 | ✓ | book | X | X |
| DSL-Based Approach for Building Model-Driven Questionnaires | 1,10 | ✓ | | 0 | 1 |
| DSS-Based Ontology Alignment in Solid Reference System Configuration | 10 | | unavailable | X | X |
| Editing Software as Strategy Value | 1 | | | X | X |
| Efficient editing in a tree-oriented projectional editor | 1,3,7,8,9 | ✓ | | 1 | 0 |
| Efficient generation of graphical model views via lazy model-to-text transformation | 1,4 | ✓ | | 1 | 0 |
| Efficient usage of abstract scenarios for the development of highly-automated driving functions | 1,10 | ✓ | unavailable | X | X |
| Enabling language engineering for the masses | 1,3 | ✓ | | 2 | 1 |
| Engineering Gameful Applications with MPS | 1,10 | ✓ | | 0 | 2 |
| Enhancing development and consistency of UML models and model executions with USE studio | 1,3,7,8,9 | ✓ | | 0 | 2 |
| Enterprise Information Systems | 10 | | book | X | X |
| Example-driven software language engineering | 1,3 | ✓ | | 1 | 2 |
| Exploring Visual Primitives for Authoring Source Code | 1 | | grey | X | X |
| FASTEN: An Extensible Platform to Experiment with Rigorous Modeling of Safety-Critical Systems | 1,10 | ✓ | | 0 | 5 |
| FeatureCoPP: unfolding preprocessor variability | 1,3 | | | X | X |

| Paper Title | lib | in | exclusion | F# | B# |
|---|-------------|------|-----------|----|----|
| FeatureVista: Interactive Feature Visualization | 1 | | | X | X |
| Filling Typed Holes with Live GUIs | 3 | ✓ | | 0 | 5 |
| First-class concepts: reifying architectural knowledge beyond the dominant decomposition | 1,3 | ✓ | | 0 | 1 |
| FORMREQ 2020 | 1,2,8 | book | X | X | |
| Gentleman: a light-weight web-based projectional editor generator | 1,3,4,7,8,9 | ✓ | | 0 | 0 |
| GPP, the Generic Preprocessor | 1 | | | X | X |
| Improving the usability of the domain-specific language editors using artificial intelligence | 1 | ✓ | grey | X | X |
| Incremental Flow Analysis through Computational Dependency Reification | 1,2 | ✓ | | 0 | 2 |
| Incrementalizing Static Analyses in Datalog | | | | | |
| Integrating the Common Variability Language with Multilanguage Annotations for Web Engineering | 1 | ✓ | grey | X | X |
| Integrating UML and Alf: An Approach to Overcome the Code Generation Dilemma in Model-Driven Software Engineering | 10 | ✓ | | 0 | 0 |
| Javardise: a structured code editor for programming pedagogy in Java | 1 | ✓ | | 0 | 0 |
| JetBrains MPS as Core DSL Technology for Developing Professional Digital Printers | 1,10 | ✓ | | 0 | 0 |
| JetBrains MPS: Why Modern Languages Workbenches Matter | 1,7,10 | ✓ | | 0 | 1 |
| Learning Data Analysis with MetaR | 1,10 | ✓ | | 0 | 0 |
| Lipschitz-like property relative to a set and the generalized Mordukhovich criterion | 6 | | | X | X |
| Macros for Domain-Specific Languages | 3 | | | X | X |
| Mechanizing metatheory interactively | 1 | ✓ | grey | X | X |
| Migrating Insurance Calculation Rule Descriptions from Word to MPS | 1,10 | ✓ | | 0 | 0 |
| Model Driven Software Engineering Meta-Workbenches: An XTools Approach | 1,5 | ✓ | | 0 | 0 |
| Model-based safety assessment with SysML and component fault trees: application and lessons learned | 1,10 | ✓ | | 8 | 0 |
| Model-Driven Development for Spring Boot Microservices | 1,5 | ✓ | grey | X | X |
| n Challenges for Software Language Engineering | 1 | | | X | X |
| On preserving variability consistency in multiple models | 1,3 | | | X | X |

| Paper Title | lib | in | exclusion | F# | B# |
|---|---------------|----|-------------|----|----|
| On the Need for a Formally Complete and Standardized Language Mapping between C++ and UML | 1 | | | X | X |
| On the Understandability of Language Constructs to Structure the State and Behavior in Abstract State Machine Specifications: A Controlled Experiment | 1 | | | X | X |
| On the use of product-line variants as experimental subjects for clone-and-own research: a case study | 1 | | | X | X |
| PAMOJA: A component framework for grammar-aware engineering | 1 | ✓ | | 0 | 1 |
| Programming Robots for Activities of Everyday Life | 1 | ✓ | grey | X | X |
| Programming tools for intelligent systems | 1 | ✓ | grey | X | X |
| Projecting Textual Languages | 1,10 | ✓ | | 0 | 1 |
| Rule-based and user feedback-driven decision support system for transforming automatically-generated alignments into information-integration alignments | 5 | | unavailable | X | X |
| Semi-Automatische Deduktion von Feature-Lokalisierung während der Softwareentwicklung: Masterarbeit | 5 | | not English | X | X |
| Should Variation Be Encoded Explicitly in Databases? | 1 | | | X | X |
| SLang: A Domain-specific Language for Survey Questionnaires | 1 | ✓ | | 0 | 0 |
| SpecEdit: Projectional Editing for TLA+ Specifications | 1,4,7 | ✓ | | 0 | 0 |
| Specifying Software Languages: Grammars, Projectional Editors, and Unconventional Approaches | 1,2,4,5,7,8,9 | ✓ | | 0 | 6 |
| Teaching Language Engineering Using MPS | 10 | ✓ | | 0 | 0 |
| Teaching MPS: Experiences from Industry and Academia | 1,10 | ✓ | | 0 | 1 |
| Teasy framework: uma solução para testes automatizados em aplicações web | 1 | ✓ | not English | X | X |
| The Art of Bootstrapping | 10 | ✓ | | 3 | 0 |
| The state of adoption and the challenges of systematic variability management in industry | 1 | | | X | X |
| Toward a domain-specific language for scientific workflow-based applications on multicloud system | 1,11 | | | X | X |
| Towards a Universal Variability Language | 1 | ✓ | grey | X | X |

| Paper Title | | lib | in | exclusion | F# | B# |
|---|-------------|-----|----|-------------|----|----|
| Towards Multi-editor Support for Domain-Specific Languages Utilizing the Language Server Protocol | | 10 | ✓ | | 5 | 0 |
| Towards Ontology-based Domain Specific Language for Internet of Things | 1,3 | ✓ | | | 0 | 0 |
| Towards projectional editing for model-based SPLs | 3,4,7,8,9 | ✓ | | | 3 | 0 |
| Tychonis: A model-based approach to define and search for geometric events in space | 1 | ✓ | | | X | X |
| Type-Directed Program Transformations for the Working Functional Programmer | 1 | ✓ | | | 0 | 0 |
| Understanding Variability-Aware Analysis in Low-Maturity Variant-Rich Systems | 1 | | | | X | X |
| Untangling Mechanized Proofs | 3 | | | | X | X |
| Variability representations in class models: An empirical assessment | 1,3 | | | | X | X |
| Visual design for a tree-oriented projectional editor | 1,3,4,7,8,9 | ✓ | | Duplicate | X | X |
| What do practitioners expect from the meta-modeling tools? A survey | 1,7,8,9 | ✓ | | | 0 | 3 |
| Cyrillic named paper 1 | 1 | | | not English | X | X |
| Cyrillic named paper 2 | 1 | | | not English | X | X |

TABLE B.2: Systematic review log - search results

Appendix C

Study Quality Assessment Checklist

The Center for Evidence-Based Management (CEBMa) supports the application of evidence-based practices to the field of management and leadership. They have a collection of checklists for assessing different types of studies. These checklists have been adapted from the pocket guide to critical appraisal[81]. We have used these as the basis of our quality assessment checklists.

Critical Appraisal of a Case Study

| # | Appraisal questions | Yes | Can't tell | No |
|----|---|-----|------------|----|
| 1 | Did the study address a focused question/issue? | | | |
| 2 | Is the research method (study design) appropriate for answering the research question? | | | |
| 3 | Are both the setting and the subject's representative concerning the population to which the findings will be referred? | | | |
| 4 | Is the researcher's perspective clearly described and taken into account? | | | |
| 5 | Are the methods for collecting data clearly described? | | | |
| 6 | Are the methods for analyzing the data likely to be valid and reliable? Are quality-control measures used? | | | |
| 7 | Was the analysis repeated by more than one researcher to ensure reliability? | | | |
| 8 | Are the results credible, and if so, are they relevant for practice? | | | |
| 9 | Are the conclusions drawn justified by the results? | | | |
| 10 | Are the findings of the study transferable to other settings? | | | |

TABLE C.1: Case studies quality assessment checklist

Critical Appraisal of a Qualitative Study

| # | Appraisal questions | Yes | Can't tell | No |
|----|---|-----|------------|----|
| 1 | Did the study address a focused question/issue? | | | |
| 2 | Is the research method (study design) appropriate for answering the research question? | | | |
| 3 | Was the context clearly described? | | | |
| 4 | How was the fieldwork undertaken? Was it described in detail? Are the methods for collecting data clearly described? | | | |
| 5 | Could the evidence (fieldwork notes, interview transcripts, recordings, documentary analysis, etc.) be inspected independently by others? | | | |
| 6 | Are the procedures for data analysis reliable and theoretically justified? Are quality-control measures used? | | | |
| 7 | Was the analysis repeated by more than one researcher to ensure reliability? | | | |
| 8 | Are the results credible, and if so, are they relevant for practice? | | | |
| 9 | Are the conclusions drawn justified by the results? | | | |
| 10 | Are the findings of the study transferable to other settings? | | | |

TABLE C.2: Qualitative studies quality assessment checklist

Critical Appraisal of a Survey Study

| # | Appraisal questions | Yes | Can't tell | No |
|----|---|-----|------------|----|
| 1 | Did the study address a focused question/issue? | | | |
| 2 | Is the research method (study design) appropriate for answering the research question? | | | |
| 3 | Is the method of selection of the subjects (employees, teams, divisions, organizations) clearly described? | | | |
| 4 | Could the way the sample was obtained introduce (selection) bias? | | | |
| 5 | Was the sample of subjects representative concerning the population to which the findings will be referred? | | | |
| 6 | Was the sample size based on pre-study considerations of statistical power? | | | |
| 7 | Was a satisfactory response rate achieved? | | | |
| 8 | Are the measurements (questionnaires) likely to be valid and reliable? | | | |
| 9 | Was the statistical significance assessed? | | | |
| 10 | Are confidence intervals given for the main results? | | | |
| 11 | Could there be confounding factors that haven't been accounted for? | | | |
| 12 | Are the findings of the study transferable to other settings? | | | |

TABLE C.3: Survey studies quality assessment checklist

Critical Appraisal of a Cohort or Panel Study

| # | Appraisal questions | Yes | Can't tell | No |
|----|--|-----|------------|----|
| 1 | Did the study address a focused question/issue? | | | |
| 2 | Is the research method (study design) appropriate for answering the research question? | | | |
| 3 | Were there enough subjects (employees, teams, divisions, organizations) in the study to establish that the findings did not occur by chance? | | | |
| 4 | Was the selection of the cohort/panel based on external, objective, and validated criteria? | | | |
| 5 | Was the cohort/panel representative of a defined population? | | | |
| 6 | Was the follow up of cases/subjects long enough? | | | |
| 7 | Were objective and unbiased outcome criteria used? | | | |
| 8 | Are objective and validated measurement methods used to measure the outcome? | | | |
| 9 | Is the size effect practically relevant? | | | |
| 10 | How precise is the estimate of the effect? Were confidence intervals given? | | | |
| 11 | Could there be confounding factors that haven't been accounted for? | | | |
| 12 | Are the findings of the study transferable to other settings? | | | |

TABLE C.4: Cohort or panel studies quality assessment checklist

Appendix D

Study Quality Assessment Results

In this appendix we present the data for the findings of the Quality assessment stage of the SLR. This occurred after the initial search engine selections, the three snowballing iterations and the final deep read for classification.

Naturally, this table only reports on primary studies. Where a paper reports on more than one study the paper title appears multiple times, with the type of study in parenthesis. If we had multiple papers reporting on the same study, they have already been removed.

We separated the studies into their types. When the authors self reported a type, even if we were not in agreement with them we categorised these as such. The study types were survey, case study, action design research, and qualitative study. For the sake of table width we refer to action design research as ADR.

The question assessments are based on the assessment criteria presented in Appendix C. However these criteria do not have a checklist for Action Design Research. After much research we did not find an adequate checklist for ADR, so we used the case study checklist. After concluding the Quality assessment we had to conclude either that this checklist was not a valid interrogation of ADR studies, or that all 20 ADR studies were bad.

To score the studies, we arbitrarily decided to give a +1 value for positive answers, 0 for don't knows and -1 for negative answers. We understand that this is a crude system. Note that, whilst most questions answered with "Yes" were considered positive, in the survey checklist, question "Could the way the sample was obtained introduce (selection)bias?", the positively scored answer is "No".

| Name | Type | Score | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 |
|--|-------------------|-------|---|---|---|---|---|---|---|---|---|----|----|----|
| A Domain-Specific Language for Payroll Calculations: a Case Study at DATEVI ^[109] | Case Study | 3 | Y | Y | ? | N | ? | N | ? | Y | Y | - | - | - |
| A Framework for Projectional Multi-varient Model Editors ^[110] | ADR | -5 | N | ? | ? | N | ? | N | ? | ? | N | - | - | - |
| A Generic Projectional Editor for EMF Models ^[111] | ADR | -5 | N | ? | ? | N | ? | N | ? | ? | N | - | - | - |
| A Model-Driven Approach Towards Automatic Migration to Microservices ^[112] | ADR | -5 | N | ? | ? | N | ? | N | ? | ? | N | - | - | - |
| AdaptiveVLE: An Integrated Framework for Personalized Online Education Using MPS JetBrains Domain-Specific Modeling Environment ^[113] | ADR | 3 | N | ? | Y | ? | Y | Y | N | Y | Y | ? | - | - |
| Adding Interactive Visual Syntax to Textual Code ^[114] | ADR | -5 | N | ? | ? | N | ? | N | ? | ? | N | - | - | - |
| Blended graphical and textual modelling for UML profiles: A proof-of-concept implementation and experiments ^[115] | Qualitative Study | 8 | Y | Y | Y | ? | Y | ? | Y | Y | Y | - | - | - |
| Block-based syntax from context-free grammars ^[115] | Case Study | -3 | N | ? | ? | N | ? | N | ? | Y | ? | - | - | - |
| Classification Algorithms Framework (CAF) to Enable Intelligent Systems Using JetBrains MPS Domain-Specific Languages Environment ^[116] | ADR | 3 | N | ? | Y | ? | Y | Y | N | Y | Y | ? | - | - |
| DSL Based Approach for Building Model-Driven Questionnaires (Action Research) ^[117] | ADR | -5 | N | ? | ? | N | ? | N | ? | ? | N | - | - | - |
| DSL Based Approach for Building Model-Driven Questionnaires (Qualitative Study1) ^[117] | Qualitative Study | -2 | N | ? | Y | N | ? | N | ? | ? | ? | - | - | - |
| DSL Based Approach for Building Model-Driven Questionnaires (Qualitative Study2) ^[117] | Qualitative Study | -2 | N | ? | Y | N | ? | N | ? | ? | ? | - | - | - |
| Efficient editing in a tree-oriented projectional editor ^[118] | ADR | -5 | N | ? | ? | N | ? | N | ? | ? | N | - | - | - |
| Efficient generation of graphical model views via lazy model-to-text transformation ^[119] | ADR | 0 | N | ? | ? | N | Y | Y | ? | ? | ? | - | - | - |
| Engineering Gameful Applications with MPS ^[120] | ADR | -6 | N | ? | N | N | ? | N | ? | ? | N | - | - | - |
| FASTEN: An Extensible Platform to Experiment with Rigorous Modeling of Safety-Critical Systems ^[121] | ADR | -4 | N | ? | N | ? | N | ? | ? | ? | ? | - | - | - |

| Name | Type | Score | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 |
|--|------------|-------|---|---|---|---|---|---|---|---|---|----|----|----|
| Gentleman: a light-weight web-based projectional editor generator[108] | ADR | -5 | N | ? | N | ? | N | ? | N | ? | N | - | - | - |
| Integrating UML and ALF: An Approach to Overcome the Code Generation Dilemma in Model-Driven Software Engineering[122] | ADR | -5 | N | ? | N | ? | N | ? | N | ? | N | - | - | - |
| Javardise: a structured code editor for programming pedagogy in Java[123] | ADR | -5 | N | ? | N | ? | N | ? | N | ? | N | - | - | - |
| JetBrains MPS as Core DSL Technology for Developing Professional Digital Printers[124] | Case Study | -6 | N | ? | N | N | ? | N | ? | N | ? | N | - | - |
| Learning Data Analysis with MetaR[125] | ADR | -4 | N | ? | Y | N | N | ? | N | ? | N | - | - | - |
| Migrating Insurance Calculation Rule Descriptions from Word to MPS[126] | Case Study | -3 | N | ? | Y | N | N | ? | N | ? | N | - | - | - |
| Model-based safety assessment with SysML and component fault trees: application and lessons learned (Case study1)[127] | Case Study | -4 | Y | ? | N | N | ? | N | ? | N | ? | N | - | - |
| Model-based safety assessment with SysML and component fault trees: application and lessons learned (Case study2)[127] | Case Study | -2 | Y | ? | Y | N | N | ? | N | ? | N | - | - | - |
| Model-based safety assessment with SysML and component fault trees: application and lessons learned (Action Research)[127] | ADR | -3 | N | ? | Y | N | N | ? | N | ? | N | - | - | - |
| Papyrus for gamers, let's play modeling[128] | ADR | -4 | N | ? | ? | N | N | ? | N | ? | N | - | - | - |
| Projecting Textual Languages (Action Research)[129] | ADR | -5 | N | ? | N | N | ? | N | ? | N | ? | - | - | - |
| Projecting Textual Languages (Case Study)[129] | Case Study | -6 | N | ? | N | N | ? | N | ? | N | ? | N | - | - |
| SpecEdit: Projectional Editing for TLA+ Specifications (Action Research)[130] | ADR | -2 | Y | ? | ? | N | N | ? | N | ? | N | - | - | - |
| SpecEdit: Projectional Editing for TLA+ Specifications (Case Study)[130] | Case Study | -5 | N | ? | N | N | ? | N | ? | N | ? | - | - | - |
| Teaching Language Engineering Using MPS[131] | Case Study | 3 | N | ? | Y | Y | N | ? | N | Y | Y | - | - | - |
| Teaching MPS: Experiences from Industry and Academia[132] | Case Study | 0 | N | ? | Y | Y | N | ? | N | ? | ? | - | - | - |
| Tiny Structure Editors for Low, Low Prices (Action Research)[133] | ADR | -5 | N | ? | N | N | ? | N | ? | N | ? | N | - | - |
| Tiny Structure Editors for Low, Low Prices (Case Study)[133] | Case Study | -6 | N | ? | N | N | ? | N | ? | N | ? | N | - | - |

| Name | Type | Score | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 |
|--|--------|-------|---|---|---|----|---|---|---|---|---|----|----|----|
| Towards Ontology-based Domain Specific Language for Internet of Things[134] | ADR | -6 | N | ? | N | N | ? | N | ? | ? | N | - | - | - |
| Type-Directed Program Transformations for the Working Functional Programmer[135] | ADR | -3 | Y | Y | N | N | ? | N | ? | ? | N | - | - | - |
| What do practitioners expect from the meta-modeling tools? A survey[136] | Survey | 1 | Y | Y | Y | Y* | ? | ? | Y | Y | N | ? | N | |

TABLE D.1: Quality assessment results

Appendix E

Data Extraction Results

| | |
|---------------------|---|
| Study ID | 1 |
| Title of Study | A domain-specific language for payroll calculations: A case study at DATEV |
| Year of Publication | 2021 |
| Author(s) Names | M. Voelter, S. Košcejev, M. Riedel, A. Deitsch, and A. Hinkelmann |
| Source of Study | Google Scholar, SpringerLink |
| Type of Study | Case Study |
| Name of Venue | Domain-Specific Languages in Practice |
| Tools in Study | Jetbrains MPS |
| Sentiment | <pre>generallearnings -- negative = 23, neutral = 28, positive = 13 evaluation -- negative = 45, neutral = 69, positive = 23 conclusion -- negative = 3, neutral = 10, positive = 4</pre> |
| Study ID | 2 |
| Title of Study | A framework for projectional multi-variant model editors |
| Year of Publication | 2021 |
| Author(s) Names | J. Schröpfer, T. Buchmann, and B. Westfecht |
| Source of Study | Google Scholar, SCOPUS |
| Type of Study | Action Design Research |
| Name of Venue | MODELSWARD |
| Tools in Study | EMF & Ecore |
| Sentiment | <pre>intro -- negative = 2, neutral = 23, positive = 2 conclusion -- negative = 2, neutral = 9, positive = 3</pre> |
| Study ID | 3 |
| Title of Study | A generic projectional editor for EMF models |
| Year of Publication | 2020 |
| Author(s) Names | J. Schröpfer, T. Buchmann, and B. Westfecht |
| Source of Study | Google Scholar, Microsoft Academic, SCOPUS, Semantic Scholar |
| Type of Study | Action Design Research |
| Name of Venue | MODELSWARD |
| Tools in Study | EMF & Ecore |
| Sentiment | <pre>intro -- negative = 12, neutral = 43, positive = 8 conclusion -- negative = 2, neutral = 6, positive = 1</pre> |
| Study ID | 4 |
| Title of Study | A model-driven approach towards automatic migration to microservices |
| Year of Publication | 2020 |
| Author(s) Names | A. Buccharone, K. Soysal, and C. Guidi |
| Source of Study | SpringerLink |
| Type of Study | Action Design Research |
| Name of Venue | International Workshop on Software Engineering Aspects of Continuous Development and New Paradigms of Software Production and Deployment |
| Tools in Study | Jetbrains MPS |
| Sentiment | <pre>intro -- negative = 1, neutral = 15, positive = 3 conclusion -- neutral = 4</pre> |

FIGURE E.1: Data extraction results 1 - 4

| | |
|---------------------|---|
| Study ID | 5 |
| Title of Study | AdaptiveVLE: An integrated framework for personalized online education using MPS JetBrains domain-specific modeling environment |
| Year of Publication | 2020 |
| Author(s) Names | S. Meacham, V. Pech, and D. Nauck |
| Source of Study | Google Scholar, IEEEExplores |
| Type of Study | Action Design Research |
| Name of Venue | IEEE Access |
| Tools in Study | Jetbrains MPS |
| Sentiment | intro -- negative = 3, neutral = 17, positive = 2 conclusion -- negative = 2, neutral = 3, positive = 3 |
| Study ID | 6 |
| Title of Study | Adding interactive visual syntax to textual code |
| Year of Publication | 2020 |
| Author(s) Names | L. Andersen, M. Ballantyne, and M. Felleisen |
| Source of Study | ACM |
| Type of Study | Action Design Research |
| Name of Venue | Proceedings of the ACM on Programming Languages OOPSLA |
| Tools in Study | Racket and DrRacket |
| Sentiment | conclusion -- negative = 1, neutral = 14, positive = 4 |
| Study ID | 7 |
| Title of Study | Blended graphical and textual modelling for UML profiles: A proof-of-concept implementation and experiment |
| Year of Publication | 2021 |
| Author(s) Names | L. Addazi and F. Ciccozzi |
| Source of Study | Google Scholar |
| Type of Study | Qualitative Study |
| Name of Venue | Journal of Systems and Software |
| Tools in Study | Xtext & Ecore |
| Sentiment | intro -- negative = 6, neutral = 21, positive = 5 projectionalediting -- neutral = 6 discussion -- negative = 18, neutral = 26, positive = 6 conclusion -- negative = 1, neutral = 5, positive = 4 |
| Study ID | 8 |
| Title of Study | Classification algorithms framework (CAF) to enable intelligent systems using JetBrains MPS domain-specific languages environment |
| Year of Publication | 2020 |
| Author(s) Names | S. Meacham, V. Pech, and D. Nauck |
| Source of Study | IEEEExplores, CORE |
| Type of Study | Action Design Research |
| Name of Venue | IEEE Access |
| Tools in Study | Jetbrains MPS |
| Sentiment | intro -- negative = 2, neutral = 19 evaluation -- negative = 3, neutral = 10, positive = 10 conclusion -- negative = 1, neutral = 6 |
| Study ID | 9 |
| Title of Study | DSL based approach for building model-driven questionnaires |
| Year of Publication | 2020 |
| Author(s) Names | A. L. Furtado |
| Source of Study | Google Scholar, SpringerLink |
| Type of Study | Action Design Research + Qualitative Study |
| Name of Venue | Enterprise Information Systems: 22nd International Conference ICEIS 2020 |
| Tools in Study | Jetbrains MPS |
| Sentiment | intro -- negative = 4, neutral = 11, positive = 3 Implementations -- negative = 4, neutral = 16 conclusion -- negative = 3, neutral = 5, positive = 2 |

FIGURE E.2: Data extraction results 5 - 9

| | |
|---------------------|---|
| Study ID | 10 |
| Title of Study | Efficient editing in a tree-oriented projectional editor |
| Year of Publication | 2020 |
| Author(s) Names | T. Beckmann |
| Source of Study | Google Scholar, ACM, Microsoft Academic, SCOPUS, Semantic Scholar |
| Type of Study | Action Design Research |
| Name of Venue | Conference Companion of the 4th International Conference on Art, Science, and Engineering of Programming |
| Tools in Study | sandblocks (Research Project) |
| Sentiment | intro -- negative = 1, neutral = 7, positive = 2 design -- neutral = 9 conclusion -- negative = 1, neutral = 1, positive = 1 |
| Study ID | 11 |
| Title of Study | Efficient generation of graphical modelviews via lazy model-to-text transformation |
| Year of Publication | 2020 |
| Author(s) Names | D. Kolovos, A. De La Vega, and J. Cooper |
| Source of Study | Google Scholar, BASE |
| Type of Study | Action Design Research |
| Name of Venue | Proceedings of the 23rd ACM/IEEE International Conference on Model Driven Engineering Languages and Systems |
| Tools in Study | Picto (Research Project using EGL, Graphviz, PlantUML) |
| Sentiment | intro -- negative = 3, neutral = 5, positive = 1 evaluation -- negative = 5, neutral = 43, positive = 1 results -- negative = 5, neutral = 34, positive = 5 conclusion -- negative = 1, neutral = 3, positive = 3 |
| Study ID | 13 |
| Title of Study | Engineering gameful applications with MPS |
| Year of Publication | 2021 |
| Author(s) Names | A. Bucciarone, A. Cicchetti, and A. Marconi |
| Source of Study | Google Scholar, SpringerLink |
| Type of Study | Action Design Research |
| Name of Venue | Domain-Specific Languages in Practice |
| Tools in Study | Jetbrains MPS |
| Sentiment | intro -- negative = 6, neutral = 20, positive = 5 engineering -- negative = 2, neutral = 34, positive = 7 mpsprojectional -- negative = 2, neutral = 18 lessonslearned -- negative = 11, neutral = 16, positive = 5 conclusion -- negative = 1, neutral = 5, positive = 2 |
| Study ID | 15 |
| Title of Study | Fasten: An extensible platform to experiment with rigorous modeling of safety-critical systems |
| Year of Publication | 2021 |
| Author(s) Names | D. Ratiu, A. Nordmann, P. Munk, C. Carlan, and M. Voelter |
| Source of Study | Google Scholar, SpringerLink |
| Type of Study | Action Design Research |
| Name of Venue | Domain-Specific Languages in Practice |
| Tools in Study | Jetbrains MPS |
| Sentiment | intro -- negative = 18, neutral = 40, positive = 11 platform -- neutral = 13, positive = 2 discussion -- negative = 3, neutral = 21, positive = 2 discussionMPS -- negative = 14, neutral = 25, positive = 8 conclusion -- negative = 3, neutral = 11, positive = 4 |
| Study ID | 16 |
| Title of Study | Gentleman: A light-weight web-based projectional editor generator |
| Year of Publication | 2020 |
| Author(s) Names | L.E. Lafontant and E. Syriani |
| Source of Study | Google Scholar, ACM, BASE, Microsoft Academic, SCOPUS, Semantic Scholar |
| Type of Study | Action Design Research |
| Name of Venue | Proceedings of the 23rd ACM/IEEE International Conference on Model Driven Engineering Languages and Systems: Companion Proceedings |
| Tools in Study | Gentleman (Research Project) |
| Sentiment | intro -- negative = 8, neutral = 12, positive = 2 editor -- negative = 5, neutral = 43, positive = 5 implementation -- neutral = 10, positive = 1 projections -- neutral = 40, positive = 6 conclusion -- neutral = 4, positive = 2 |

FIGURE E.3: Data extraction results 10 - 16

| | |
|---------------------|--|
| Study ID | 17 |
| Title of Study | Integrating UML and ALF: An approach to overcome the code generation dilemma in model-driven software engineering |
| Year of Publication | 2020 |
| Author(s) Names | J. Schröpfer and T. Buchmann |
| Source of Study | SpringerLink |
| Type of Study | Action Design Research |
| Name of Venue | International Conference on Model-Driven Engineering and Software Development |
| Tools in Study | ALF (Xtext) + Valkyrie (GMF) |
| Sentiment | intro -- negative = 4, neutral = 24, positive = 1 discussion -- negative = 4, neutral = 12, positive = 5 UI -- neutral = 9, positive = 2 conclusion -- neutral = 5 |
| Study ID | 18 |
| Title of Study | Javardise: A structured code editor for programming pedagogy in Java |
| Year of Publication | 2020 |
| Author(s) Names | A. L. Santos |
| Source of Study | Google Scholar |
| Type of Study | Action Design Research |
| Name of Venue | Conference Companion of the 4th International Conference on Art, Science, and Engineering of Programming |
| Tools in Study | Javardise (Research Project) |
| Sentiment | discussion -- negative = 16, neutral = 13, positive = 4 intro -- negative = 11, neutral = 11, positive = 1 |
| Study ID | 19 |
| Title of Study | Jetbrains MPS as core DSL technology for developing professional digital printers |
| Year of Publication | 2021 |
| Author(s) Names | E. Schindler, H. Moneva, J. van Pinxten, L. van Gool, B. van der Meulen, N. Stotz, and B. Theelen |
| Source of Study | Google Scholar, SpringerLink |
| Type of Study | Case Study |
| Name of Venue | Domain-Specific Languages in Practice |
| Tools in Study | Jetbrains MPS |
| Sentiment | CollaborativeDSM -- negative = 5, neutral = 29, positive = 13 conclusion -- negative = 2, neutral = 22, positive = 2 intro -- negative = 1, neutral = 21, positive = 3 MPS -- negative = 9, neutral = 9, positive = 3 |
| Study ID | 20 |
| Title of Study | Learning data analysis with metaR |
| Year of Publication | 2021 |
| Author(s) Names | M. Simi |
| Source of Study | Google Scholar, SpringerLink |
| Type of Study | Action Design Research |
| Name of Venue | Domain-Specific Languages in Practice |
| Tools in Study | Jetbrains MPS |
| Sentiment | intro -- negative = 1, neutral = 4, positive = 1 languagecomposition -- neutral = 22, positive = 3 MPS -- negative = 9, neutral = 56, positive = 6 conclusion -- negative = 2, neutral = 4, positive = 5 |
| Study ID | 21 |
| Title of Study | Migrating insurance calculation rule descriptions from Word to MPS |
| Year of Publication | 2021 |
| Author(s) Names | N. Stotz and K. Birken |
| Source of Study | Google Scholar, SpringerLink |
| Type of Study | Case Study |
| Name of Venue | Domain-Specific Languages in Practice |
| Tools in Study | Jetbrains MPS |
| Sentiment | intro -- neutral = 5, positive = 1 solutiontechnology -- negative = 20, neutral = 26, positive = 11 evaluation -- negative = 37, neutral = 49, positive = 10 conclusion -- negative = 9, neutral = 8, positive = 7 |

FIGURE E.4: Data extraction results 17 - 21

| | |
|---------------------|--|
| Study ID | 22 |
| Title of Study | Model-based safety assessment with sysml and component fault trees: Application and lessons learned |
| Year of Publication | 2020 |
| Author(s) Names | P. Munk and A. Nordmann |
| Source of Study | Google Scholar, SpringerLink |
| Type of Study | Action Design Research + Case Study |
| Name of Venue | Software and Systems Modeling |
| Tools in Study | Jetbrains MPS |
| Sentiment | intro -- negative = 8, neutral = 11, positive = 6 realisation -- negative = 29, neutral = 46, positive = 11 discussion -- negative = 6, neutral = 4, positive = 2 conclusion -- negative = 4, neutral = 3, positive = 1 |
| Study ID | 23 |
| Title of Study | Papyrus for gamers, let's play modeling |
| Year of Publication | 2020 |
| Author(s) Names | A. Buchiarone, M. Savary-Leblanc, X. L. Pallec, J.M. Bruel, A. Cicchetti, J. Cabot,S. Gerard, H. Aslam, A. Marconi, and M. Perillo |
| Source of Study | snowball |
| Type of Study | Action Design Research |
| Name of Venue | Proceedings of the 23rd ACM/IEEE International Conference on Model-Driven Engineering Languages and Systems: Companion Proceedings |
| Tools in Study | Jetbrains MPS |
| Sentiment | intro -- negative = 5, neutral = 8, positive = 3 |
| Study ID | 24 |
| Title of Study | Projecting textual languages |
| Year of Publication | 2021 |
| Author(s) Names | M. V. Merino, J. Bartels, M. van den Brand, T. van der Storm, and E. Schindler |
| Source of Study | Google Scholar, SpringerLink |
| Type of Study | Action Design Research + Case Study |
| Name of Venue | Domain-Specific Languages in Practice |
| Tools in Study | Jetbrains MPS + Rascal |
| Sentiment | intro -- negative = 9, neutral = 17, positive = 3 projectionalApproach -- negative = 5, neutral = 128, positive = 4 projectionalSyntax -- negative = 4, neutral = 42 limitation -- negative = 10, neutral = 25, positive = 6 discussion -- negative = 6, neutral = 16, positive = 9 conclusion -- negative = 4, neutral = 14, positive = 5 |
| Study ID | 25 |
| Title of Study | SpecEdit: Projectional editing for TLA+ specifications |
| Year of Publication | 2020 |
| Author(s) Names | R. Cuinat, C. Teodorov, and J. Champeau |
| Source of Study | Google Scholar, BASE, Microsoft Academic |
| Type of Study | Action Design Research + Case Study |
| Name of Venue | 2020 IEEE Workshop on Formal Requirements (FORMREQ) |
| Tools in Study | Jetbrains MPS |
| Sentiment | intro -- negative = 4, neutral = 16, positive = 6 projectionalEditor -- negative = 7, neutral = 78, positive = 17 lessonslearned -- negative = 4, neutral = 6, positive = 4 discussion -- neutral = 5, positive = 1 |
| Study ID | 26 |
| Title of Study | Teaching language engineering using MPS |
| Year of Publication | 2021 |
| Author(s) Names | A. Prinz |
| Source of Study | SpringerLink |
| Type of Study | Case Study |
| Name of Venue | Domain-Specific Languages in Practice |
| Tools in Study | Jetbrains MPS |
| Sentiment | intro -- negative = 6, neutral = 25, positive = 8 lessonslearned -- negative = 2, neutral = 12, positive = 2 metalanguages -- negative = 11, neutral = 25, positive = 11 MPSinTeaching -- negative = 14, neutral = 38, positive = 9 selectingTools -- negative = 6, neutral = 11, positive = 6 evaluation -- negative = 9, neutral = 30, positive = 20 conclusion -- negative = 2, neutral = 3, positive = 3 |

FIGURE E.5: Data extraction results 22 - 26

| | |
|---------------------|--|
| Study ID | 27 |
| Title of Study | Teaching MPS: Experiences from industry and academia |
| Year of Publication | 2021 |
| Author(s) Names | M. Barash and V. Pech |
| Source of Study | Google Scholar, SpringerLink |
| Type of Study | Case Study |
| Name of Venue | Domain-Specific Languages in Practice |
| Tools in Study | Jetbrains MPS |
| Sentiment | explainingprojectional -- negative = 1, neutral = 6, positive = 1 conclusion -- negative = 1, neutral = 5, positive = 1 |
| Study ID | 28 |
| Title of Study | Tiny structure editors for low, low prices! (generating guis from toString functions) |
| Year of Publication | 2020 |
| Author(s) Names | B. Hempel and R. Chugh |
| Source of Study | snowball |
| Type of Study | Action Design Research + Case Study |
| Name of Venue | 2020 IEEE Symposium on Visual Languages and Human-Centric Computing (VL/HCC) |
| Tools in Study | Tiny Structure Editor (Research Project) |
| Sentiment | intro -- negative = 7, neutral = 30, positive = 2 discussion -- negative = 5, neutral = 11, positive = 1 |
| Study ID | 29 |
| Title of Study | Towards ontology-based domain specific language for internet of things |
| Year of Publication | 2020 |
| Author(s) Names | E. Negm, S. Makady, and A. Salah |
| Source of Study | Google Scholar, ACM |
| Type of Study | Action Design Research |
| Name of Venue | Proceedings of the 2020 9th International Conference on Software and Information Engineering (ICSIE) |
| Tools in Study | Jetbrains MPS |
| Sentiment | intro -- negative = 8, neutral = 22, positive = 5 implementation -- neutral = 20, positive = 2 conclusion -- neutral = 7 |
| Study ID | 30 |
| Title of Study | Type-directed program transformations for the working functional programmer |
| Year of Publication | 2020 |
| Author(s) Names | J. Lubin and R. Chugh |
| Source of Study | Google Scholar |
| Type of Study | Action Design Research |
| Name of Venue | 10th Workshop on Evaluation and Usability of Programming Languages and Tools (PLATEAU 2019) |
| Tools in Study | DEUCE |
| Sentiment | intro -- negative = 7, neutral = 9, positive = 2 implementation -- negative = 13, neutral = 16, positive = 6 usabilityChallenges -- negative = 3, neutral = 5, positive = 2 |
| Study ID | 31 |
| Title of Study | What do practitioners expect from the meta-modeling tools? a survey |
| Year of Publication | 2021 |
| Author(s) Names | M. Ozkaya and D. Akdur |
| Source of Study | Google Scholar, Microsoft Academic, SCOPUS, Semantic Scholar |
| Type of Study | Survey |
| Name of Venue | Journal of Computer Languages |
| Tools in Study | MPS, MetaEdit+, WebGME, GEMS, sirius,Xtext, MS DSL ToolsMelange, GME* |
| Sentiment | intro -- negative = 11, neutral = 62, positive = 8 editorservice -- negative = 8, neutral = 20, positive = 6 lessonslearned -- negative = 10, neutral = 24, positive = 8 challenges -- negative = 5 toolusage -- negative = 1, neutral = 14, positive = 2 conclusion -- negative = 10, neutral = 20, positive = 3 |

FIGURE E.6: Data extraction results 27 - 31

Appendix F

Drools Concept hierarchy

The concept hierarchy presented on the following pages was extracted and interpreted from Drools railroad diagrams.

The diagram in figure F1 represents the file level and can be considered the root of concept hierarchy. This represents the concepts that are available to the rule file. As the only concept we will examine in depth is the rule, we show some concepts that are shared or are children of, for example, function, query and type declaration.

In our final implementation only the import, global and Rule concepts that were children of the rule file were implemented.

The diagram in figure F2 shows the children of a rule. Each attribute has a different behavior and structure and are thus all represented separately.

In These diagrams we do not show a concept diagram for the RHS. This is because it would be more or less the concept diagram for Java Statements,with the addition of Rule Variables and some special Drools functions. the concept diagram for a General Purpose Language, like Java, would be orders of magnitude bigger and more complex. Luckily, as MPS allows for almost seamless extension and integration of different languages, we are able to just import JetBrains implementation of Java for the RHS.

The hierarch for the LHS is shown in the diagram in figure F3. Because of the number of concepts being represented, it may be a little hard to read.

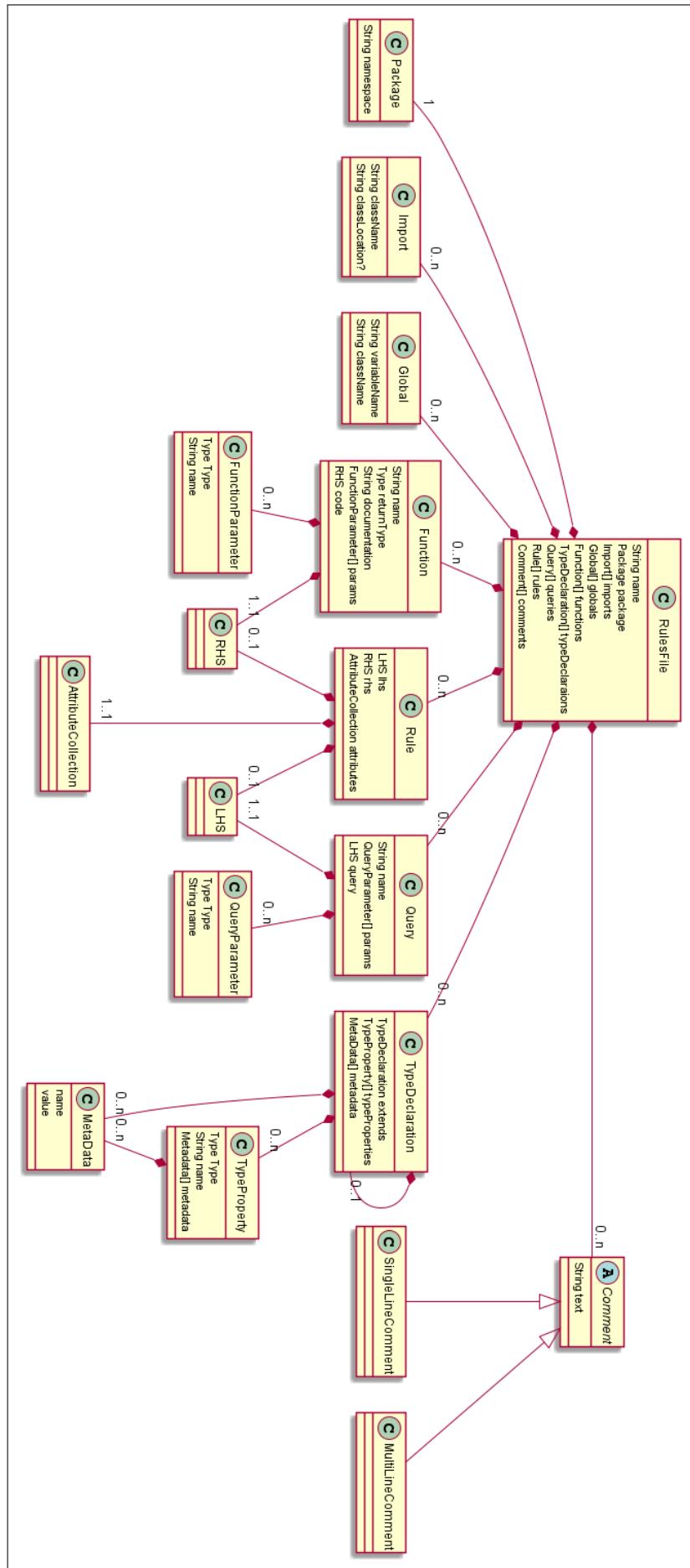


FIGURE F.1: Rule file concept hierarchy diagram

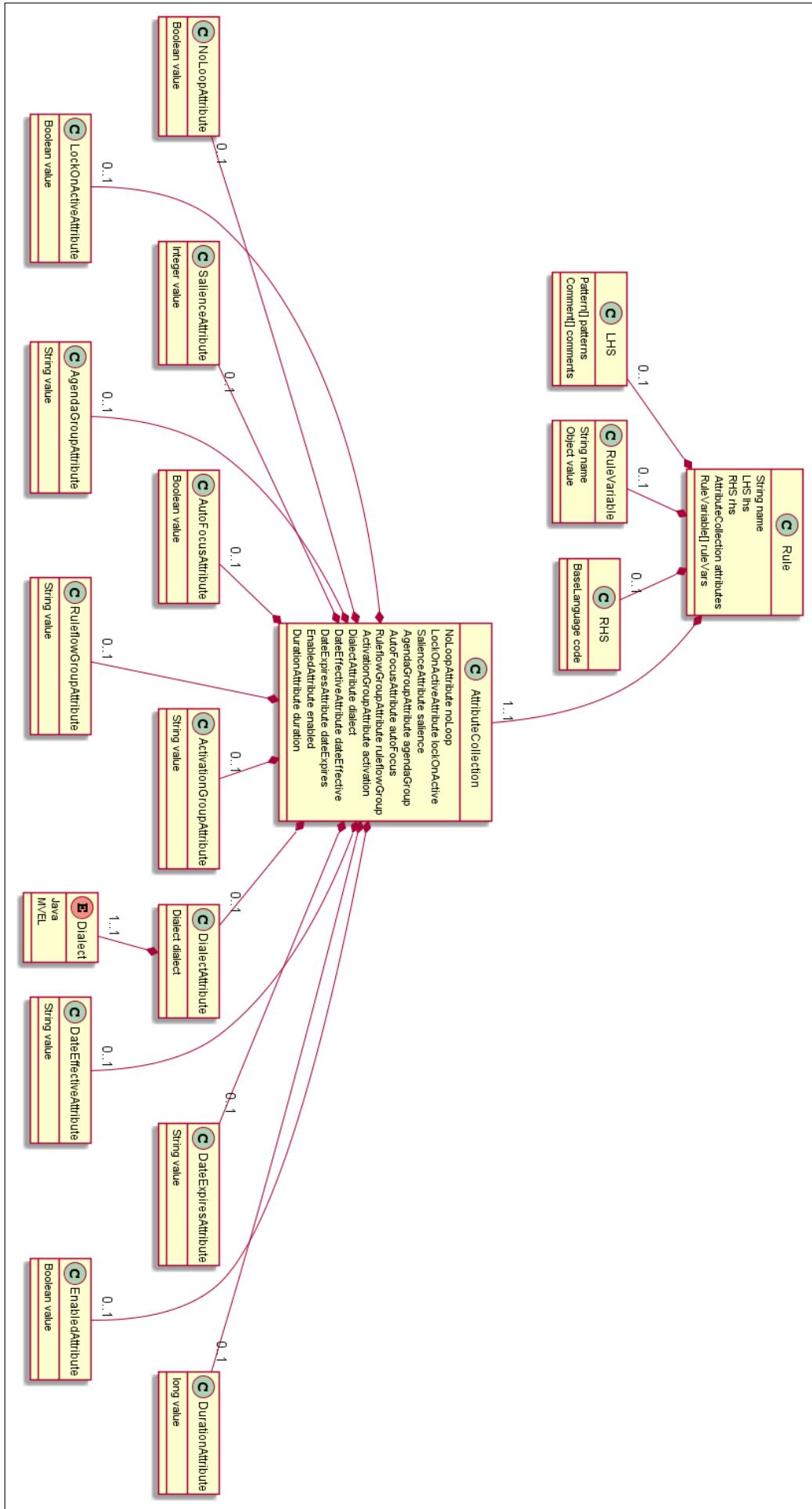


FIGURE F.2: Rules concept hierarchy diagram

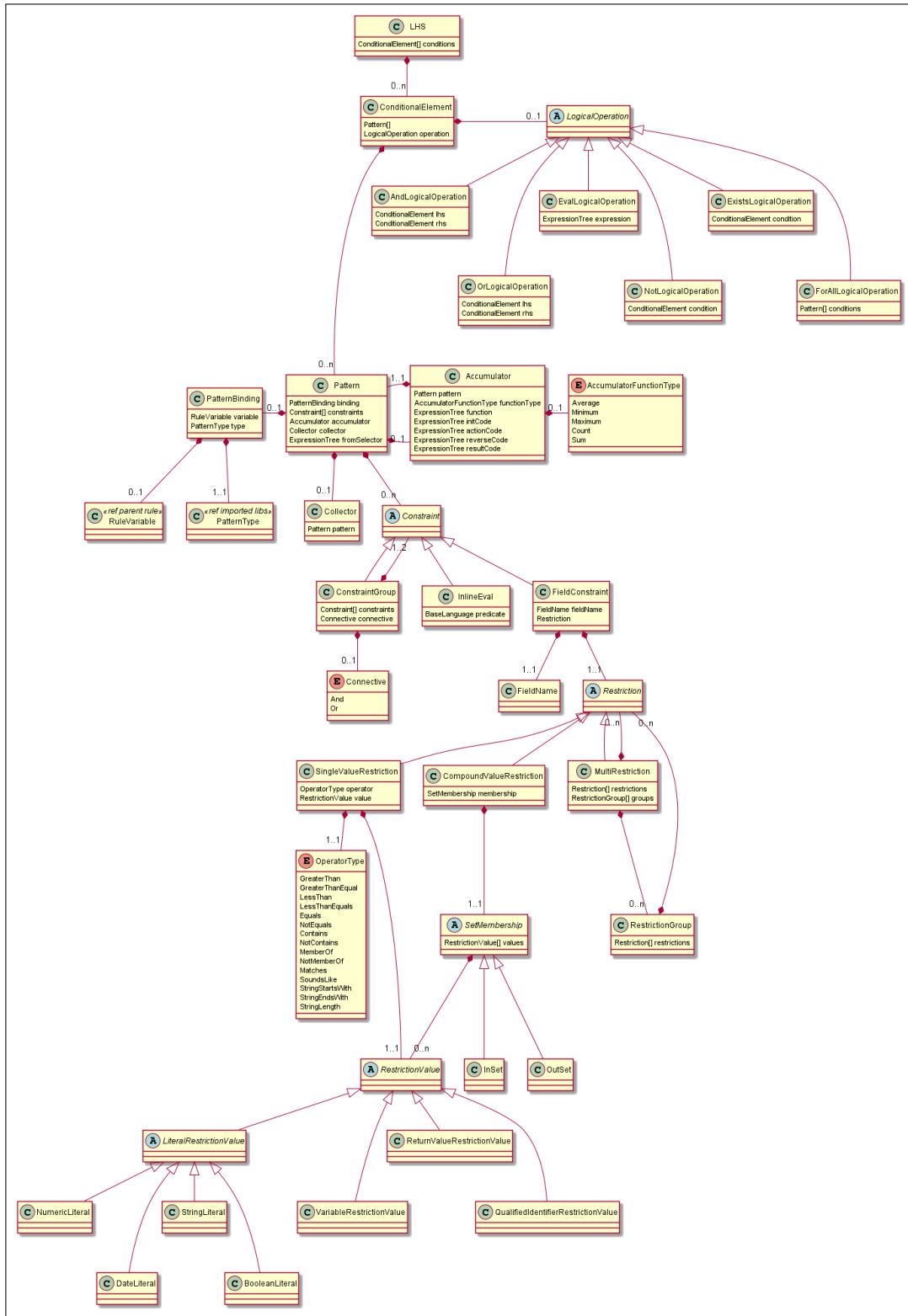


FIGURE F.3: Rule LHS concept hierarchy diagram

Appendix G

Questionnaire Text

Note: This questionnaire was presented on Survey Monkey and thus the text here is a best approximation of their paging system.

Page 1 - Introduction

Thank you for taking part in this research.

According to Survey Monkey, this survey should take 6 minutes to complete, when we tested it, the average was closer to 10 minutes.

This survey is for the validation section of a research master's project by Paul Spencer at the University of Amsterdam.

The purpose is to determine whether projectional editing can be used to aid the comprehensibility of business rules.

We are using Drools as our example business rules language.

You were selected as you asked or answered a Drools question on StackOverflow, listed Drools as a skill on your LinkedIn profile, or were referred to this survey by someone who previously answered this survey. (please feel free to forward this survey to anyone you know with Drools experience).

It is therefore assumed you are aware of what Drools is.

Projectional editing is a form of writing computer programs directly rather than writing text and having that parsed to create the program. This allows the developer multiple views and editors for the same code.

In this survey, we will present you with a few of these views.

On the following page, there is an animated GIF that will give a small demonstration of what this means.

Figure G.1 shows how this is presented to the subject.

Page 2 - Example of Projectional editing in Drools

Below is an animated GIF showing an example of a projectional implementation of Drools.

The top section is a tabular projection of the program.

The bottom part is a textual projection of the same program shown at the same time.

In this recording, we are editing in the tabular projection, which automatically updates the textual projection.

Here is placed an animated GIF of a demonstration of our prototype

Question: What is your first reaction to this mode of code editing?

Options: Very positive, Somewhat positive, Neutral, Somewhat negative, Very negative

the order of the options will be randomly presented as either “Very positive” to “Very negative” or “Very nega-

Figure G.2 shows how this is presented to the subject.

Page 3 - Positive about projectional editing

This page is only selected if the user chose very positive or somewhat positive

This question is optional.

you may use the Green “PREV” button to review the previous page.

Question: how would this coding style be useful to your interactions with Drools?

This is an open question with a text box.

Figure G.3 shows how this is presented to the subject.

Page 4 - negative about projections

This page is only selected if the user chose very positive or somewhat positive

This question is optional.

you may use the Green “PREV” button to review the previous page.

Question: What do you find negative with this style of coding

This is an open question with a text box.

Page 5 - Testing a projection

In questionnaire version A & D page 5 will be Testing a projection

In questionnaire version B & C page 5 will be Testing textual projection

On this page, we present you with an example projection of a collection of Drools rules, in this case, as a sort of decision table.

We will ask you to describe what you think it does, if you can't that is also good data for us.

A brief description of how this projection works follows:

for the decision table the following text:

- 1) each row is a rule
- 2) each column is a fact, or, when indented, a selection criteria of that fact
- 3) smiley faces indicate that a fact has been selected for a rule
- 4) if a fact has been selected and a variable is bound to it then the variable name appears instead of the smiley face.
- 5) the "Then" part of the rule appears in the "Actions" column

for the other table the following text:

- 1) each row is a rule
- 2) each column is for a variable or a property of a fact
- 3) if a property is selected then the selection criteria is in the appropriate cell
- 4) unselected cells are indicated by a grey/beige color
- 5) the "Then" part of the rule appears in the "Actions" column

depending on the version of this questionnaire the respondent will see one of the following pictures

Version A - decision table showing rule set 1 (FNWI)

Version B - decision table showing rule set 2 (LAW)

Version C - new table showing rule set 1

Version D - new table showing rule set 2

Question: Please describe what you think this group of rules does

This is an open question with a text box.

Question: How easy or difficult was it to describe this rule set?

Options: Very easy, Somewhat easy, Neutral, Somewhat difficult, Very difficult

the order of the options will be randomly presented as either "Very easy" to "Very difficult" or "Very difficult" to "Very easy"

Figure G.4 shows how this is presented to the subject.

Page 6 - Testing textual projection

In questionnaire version A & D page 6 will be Testing textual projection

In questionnaire version B & C page 6 will be Testing a projection

Here we present you a textual projection of Drools rules.

[Note: These are not the same rules as on the previous page]

depending on the version of this questionnaire the respondent will see one of the following pictures

Version A & C - a text projection of rule set 2 (LAW)

Version B & D - a text projection of rule set 1 (FNWI)

Question: Please describe what you think this group of rules does

This is an open question with a text box.

Question: How easy or difficult was it to describe this rule set?

Options: Very easy, Somewhat easy, Neutral, Somewhat difficult, Very difficult

the order of the options will be randomly presented as either “Very easy” to “Very difficult” or “Very difficult” to “Very easy”

Figure G.5 shows how this is presented to the subject.

Page 7 - Comparing projections 1

In this question, we ask to compare a new projection to a previously shown projection, on the page named “Testing a projection”.

If you wish to reacquaint yourself with the previous projection, you can use the Green “PREV” button at the bottom of this page.

A brief description of how this new projection works follows:

for the decision table the following text:

- 1) each row is a rule
- 2) each column is a fact, or, when indented, a selection criteria of that fact
- 3) smiley faces indicate that a fact has been selected for a rule
- 4) if a fact has been selected and a variable is bound to it then the variable name appears instead of the smiley face.
- 5) the “Then” part of the rule appears in the “Actions” column

for the other table the following text:

- 1) each row is a rule

- 2) each column is for a variable or a property of a fact
- 3) if a property is selected then the selection criteria is in the appropriate cell
- 4) unselected cells are indicated by a grey/beige color
- 5) the “Then” part of the rule appears in the “Actions” column

depending on the version of this questionnaire the respondent will see one of the following pictures

Version A - new table showing rule set 1

Version B - new table showing rule set 2

Version C - decision table showing rule set 1

Version D - decision table showing rule set 2

Question: How does the above projection compare to the first projection you described?

Options: Much easier to understand, Somewhat easier to understand, Neutral, Somewhat harder to understand, Much harder to understand

the order of the options will be randomly presented as either “Much easier to understand” to “Much harder to understand”

Figure G.6 shows how this is presented to the subject.

Page 8 - Comparing projections 2

In this question, we again ask to compare the new projection, this time to the textual projection, on the page named “Testing textual projection”.

If you wish to reacquaint yourself with the textual projection, you can, of course, use the Green “PREV” button at the bottom of this page again.

depending on the version of this questionnaire the respondent will see one of the following pictures

Version A - new table showing rule set 2

Version B - new table showing rule set 1

Version C - decision table showing rule set 2

Version D - decision table showing rule set 1

Question: How does the above projection compare to the text Drools rules you described?

Options: Much easier to understand, Somewhat easier to understand, Neutral, Somewhat harder to understand, Much harder to understand

the order of the options will be randomly presented as either “Much easier to understand” to “Much harder to understand”

Figure G.7 shows how this is presented to the subject.

Page 9 - Single rule helper 1 - Truth table

In questionnaire version A & D page 9 will be the Truth Table

In questionnaire version B & C page 9 will be the Circuit Diagram

Below we present another projection. This is a truth table projection. It highlights the conditions that have to be true for a rule to be selected.

The GIF shows the rule selected and the developer pressing the up and down arrow keys to step through the different true (highlighted in green) and false (highlighted in red) fact selections that result in a true outcome.

An animated GIF of the truth table example

Question: Would this help you with understanding your Drools rules?

Options: It would really help understanding, it would somewhat help understanding, Neutral, It would add a little confusion, It would add a lot of confusion

the order of the options will be randomly presented as either "It would really help understanding" to "It wo

Figure G.8 shows how this is presented to the subject.

Page 10 - Single rule helper 2 - Circuit Diagram

In questionnaire version A & D page 10 will be the Circuit Diagram

In questionnaire version B & C page 10 will be the Truth Table

This is a circuit diagram of the selection conditions. choosing a different condition highlights how they are related to each other.

The GIF shows the rule selected and the developer pressing the up and down arrow keys to step through the different fact selections (highlighted in yellow) and shown in the circuit diagram, thus showing how the facts relate to each other.

An animated GIF of the Circuit Diagram example

Question: Would this help you with understanding your Drools rules?

Options: It would really help understanding, it would somewhat help understanding, Neutral, It would add a little confusion, It would add a lot of confusion

the order of the options will be randomly presented as either "It would really help understanding" to "It wo

Figure G.9 shows how this is presented to the subject.

Page 11 - The Statistics page

Here we ask for data that we can use to slice and dice results.

Question: How long was/is your career as a developer?

Options: 0-1 year, 1-3 years, 3-10 years, greater than 10 years, none of the above

Question: When was the last time you had a coding interaction with Drools?

Options: during this week, some time after July 1st 2021, some time after Jan 1st 2021, some time after 2016, some time before 2016

Question: how long did you work with Drools?

Options: for years and intensely, for years but occasionally, not for long but intensely, I barely touched it

Question: Which tools have you used to edit Drools rules?

Checkboxes: Drools workbench, eclipse (with Drools plug-in), IntelliJ IDEA (with Drools plug-in), IDE or text editor without Drools assistance, other (please specify) has textbox, none of the above

Figure G.10 shows how this is presented to the subject.

Page 12 - So long, and thanks for all the fish

Thank you for your time. We leave you with a box where you can put in any thoughts about this if you feel like it.

Question: Do you have any thoughts or opinions you would like to share about what you have seen in this questionnaire?

This is an open question with a text box.

Figure G.11 shows how this is presented to the subject.

Projectional Drools Survey: Version D

Introduction

Thank you for taking part in this research.

According to Survey Monkey, this survey should take 6 minutes to complete, when we tested it, the average was closer to 10 minutes.

This survey is for the validation section of a research master's project by Paul Spencer at the University of Amsterdam.

The purpose is to determine whether projectional editing can be used to aid the comprehensibility of business rules.

We are using Drools as our example business rules language.

You were selected as you asked or answered a Drools question on StackOverflow, listed Drools as a skill on your LinkedIn profile, or were referred to this survey by someone who previously answered this survey. (please feel free to forward this survey to anyone you know with Drools experience).

It is therefore assumed you are aware of what Drools is.

Projectional editing is a form of writing computer programs directly rather than writing text and having that parsed to create the program. This allows the developer multiple views and editors for the same code.

In this survey, we will present you with a few of these views.

On the following page, there is an animated GIF that will give a small demonstration of what this means.

OK

FIGURE G.1: Screen 1 - introduction text

Projectional Drools Survey: Version D

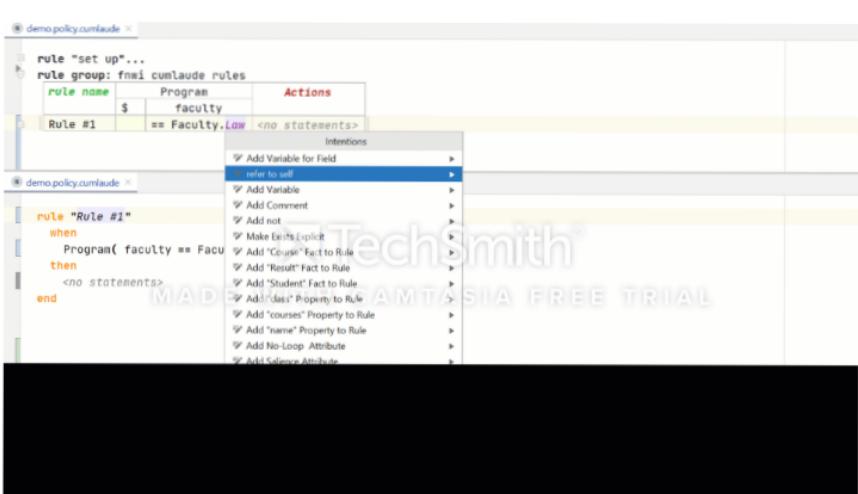
Example of Projectional editing in Drools

Below is an animated GIF showing an example of a projectional implementation of Drools.

The top section is a tabular projection of the program.

The bottom part is a textual projection of the same program shown at the same time.

In this recording, we are editing in the tabular projection, which automatically updates the textual projection.



* 1. What is your first reaction to this mode of code editing?

Very negative Somewhat negative Neutral Somewhat positive Very positive

[PREV](#) [NEXT](#)

FIGURE G.2: Screen 2 - first impression

Projectional Drools Survey: Version D

Positive about projectional editing

This question is optional.

you may use the Green "PREV" button to review the previous page.

2. How would this coding style be useful to your interactions with Drools?

[Large gray rectangular input area]

PREV **NEXT**

FIGURE G.3: Screen 3 - positive response

Projectional Drools Survey: Version D

Testing a projection

On this page, we present you with an example projection of a collection of Drools rules, in this case, as a sort of decision table.

We will ask you to describe what you think it does, if you can't that is also good data for us.

A Brief description of how this projection works follows:

- 1) each row is a rule
- 2) each column is for a variable or a property of a fact
- 3) if a property is selected then the selection criteria is in the appropriate cell
- 4) unselected cells are indicated by a grey/beige color
- 5) the "Then" part of the rule appears in the "Actions" column

| rule group: law cumlaude rules | | | | | | |
|--------------------------------|---------|-----------------------------|------------------------|---------------------|--|---|
| rule name | Program | | Student | | Result | Actions |
| | \$ | faculty | +\$ avg | closeCount | +\$ grade | |
| Rule #1 | | <code>== Faculty.Law</code> | <code>s >= 8</code> | | | <code>modify(s) { setCumlaude(true) };</code> |
| Rule #2 | | <code>== Faculty.Law</code> | <code>s</code> | | <code>>= 7 && < 8</code> | <code>int closeCnt = s.getCloseCount() + 1;</code> <code>modify(s) { setCloseCount(closeCnt) };</code> |
| Rule #3 | | <code>== Faculty.Law</code> | <code>s</code> | <code>> 1</code> | | <code>modify(s) { setCumlaude(false) };</code> <code>halt();</code> |

3. Please describe what you think this group of rules does

4. How easy or difficult was it to describe this rule set?

Very easy Somewhat easy Neutral Somewhat difficult Very difficult

PREV NEXT

FIGURE G.4: Screen 4 - describe projection

Projectional Drools Survey: Version D

Testing textual projection

Here we present you a textual projection of Drools rules.

[Note: These are not the same rules as on the previous page]

```
rule "Rule #1"
when
    Program( faculty == Faculty.FNWI )
    S : Student( )
    Result( grade < 8, exempted == false )
then
    modify( s ) { setcumlaude( false ) };
    halt();
end

rule "Rule #2"
when
    Program( faculty == Faculty.FNWI )
    S : Student( )
    c : Course( name == "Thesis" )
    Result( course == c, grade >= 8 )
then
    modify( s ) { setcumlaude( false ) };
    halt();
end

rule "Rule #3"
when
    Program( faculty == Faculty.FNWI )
    S : Student( avg >= 8 )
    Result( )
then
    modify( s ) { setcumlaude( true ) };
end
```

5. Please describe what you think this group of rules does

6. How easy or difficult was it to describe this rule set?

Very easy Somewhat easy Neutral Somewhat difficult Very difficult

[PREV](#) [NEXT](#)

FIGURE G.5: Screen 5 - describe text

Projectional Drools Survey: Version D

Comparing projections 1

In this question, we ask to compare a new projection to a previously shown projection, on the page named "Testing a projection".

If you wish to reacquaint yourself with the previous projection, you can use the Green "PREV" button at the bottom of this page.

A Brief description of how this new projection works follows:

- 1) each row is a rule
- 2) each column is a fact, or, when indented, a selection criteria of that fact
- 3) smiley faces indicate that a fact has been selected for a rule
- 4) if a fact has been selected and a variable is bound to it then the variable name appears instead of the smiley face.
- 5) the "Then" part of the rule appears in the "Actions" column

| rule group: law cumlaude rules | | | | | | | |
|--------------------------------|---------|------------------------|--------|---------|----------|----------------|---|
| rule name | Program | faculty == Faculty.Law | Result | Student | avg >= 8 | closeCount > 1 | Actions |
| Rule #1 | (| | | s | | | modify(s) { setCumlaude(true) }; |
| Rule #2 | (| | (| s | | | int closeCnt = s.getCloseCount() + 1; modify(s) { setCloseCount(closeCnt) }; |
| Rule #3 | (| | | | s | | modify(s) { setCumlaude(false) }; halt(); |

7. How does the above projection compare to the first projection you described?

- Much easier to understand
 Somewhat easier to understand
 Neutral
 Somewhat harder to understand
 Much harder to understand

PREV NEXT

FIGURE G.6: Screen 6 - compare projections

Projectional Drools Survey: Version D

Comparing projections 2

In this question, we again ask to compare the new projection, this time to the textual projection, on the page named "Testing textual projection".

If you wish to reacquaint yourself with the textual projection, you can, of course, use the Green "PREV" button at the bottom of this page again.

| rule group: fnwi cumlaude rules | | | | | | | | | |
|---------------------------------|--------|------------------|---------|-------------------------|--------|-----------------------------|--|---------|----------|
| rule name | Course | name == "Thesis" | Program | Faculty == Faculty.FNWI | Result | course == [Course Variable] | Actions | Student | avg >= 8 |
| Rule #1 | | | (C) | | | (C) exempted == false | modify(\$) { setCumlaude(false) }; halt(); | S | |
| Rule #2 | C | | (C) | | | (C) grade < 8 | modify(\$) { setCumlaude(false) }; halt(); | S | |
| rule #3 | | | (C) | (C) | | | modify(\$) { setCumlaude(true) }; | S | |

8. How does the above projection compare to the text Drools rules you described?

Much easier to understand
 Somewhat easier to understand
 Neutral
 Somewhat harder to understand
 Much harder to understand

[PREV](#)
 [NEXT](#)

FIGURE G.7: Screen 7 - compare projection to text

Projectional Drools Survey: Version D

Single rule helper 1 - Truth table

Below we present another projection.

This is a truth table projection.

It highlights the conditions that have to be true for a rule to be selected.

The GIF shows the rule selected and the developer pressing the up and down arrow keys to step through the different true (highlighted in green) and false (highlighted in red) fact selections that result in a true outcome.

```

rule "Weird blanket"
when
    Program( faculty == Faculty.FNWI || == Faculty.LAW )
    (Result( grade < 8 ) || not Result( exempted ) )
    and Student( yearsStudied < 5 )
    Course( name != "Thesis" ) || Result( exempted )
then
    halt();
end

```

| A | B | C | D | E | F | |
|---|---|---|---|---|---|---|
| F | T | F | F | T | T | T |
| F | T | F | T | T | T | T |
| F | T | T | T | F | T | T |
| T | I | T | T | T | T | T |
| T | F | F | F | T | T | T |
| T | E | T | F | T | T | T |
| T | F | T | T | F | F | T |
| T | F | T | T | T | T | T |
| T | T | F | F | T | T | T |
| T | T | T | F | T | T | T |
| T | T | T | T | F | T | T |
| T | T | T | T | T | T | T |

9. Would this help you with understanding your Drools rules?

- It would really help understanding
 It would somewhat help understanding
 Neutral
 It would add a little confusion
 It would add a lot of confusion

PREV NEXT

FIGURE G.8: Screen 8 - truth table

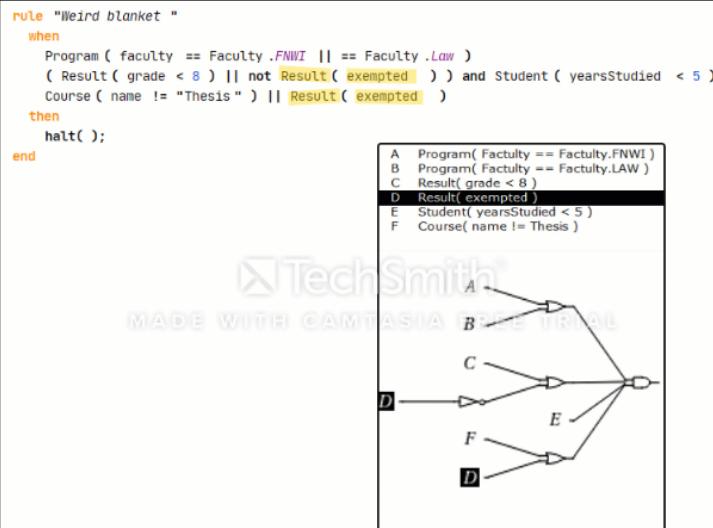
Projectional Drools Survey: Version D

Single rule helper 2 - Circuit Diagram

This is a circuit diagram of the selection conditions.

Choosing a different condition highlights how they are related to each other.

The GIF shows the rule selected and the developer pressing the up and down arrow keys to step through the different fact selections (highlighted in yellow) and shown in the circuit diagram, thus showing how the facts relate to each other.



10. Would this help you with understanding your Drools rules?

- It would really help understand
 it would somewhat help understand
 Neutral
 It would add a little confusion
 It would add a lot of confusion

[PREV](#) | [NEXT](#)

FIGURE G.9: Screen 9 - circuit diagram

Projectional Drools Survey: Version D

The Statistics page

Here we ask for data that we can use to slice and dice results.

11. How long was/is your career as a developer?

0-1 year greater than 10 years
 1-3 years None of the above
 3-10 years

12. When was the last time you had a coding interaction with Drools?

during this week some time after 2016
 some time after July 1st 2021 some time before 2016
 some time after Jan 1st 2021

13. how long did you work with Drools?

for years and intensely
 for years, but occasionally
 not for long, but intensely
 I barely touched it

14. Which tools have you used to edit Drools rules?

Drools workbench IDE or text editor without Drools assistance
 eclipse (with drools plugin) None of the above
 IntelliJ IDEA (with drools plugin)

Other (please specify)

[PREV](#) [NEXT](#)

FIGURE G.10: Screen 10 - personal details page

Projectional Drools Survey: Version D

So long, and thanks for all the fish

Thank you for your time. We leave you with a box where you can put in any thoughts about this if you feel like it.

15. Do you have any thoughts or opinions you would like to share about what you have seen in this questionnaire?

[Redacted text area]

[PREV](#) [DONE](#)

FIGURE G.11: Screen 11 - further comments