**Frege IDE**

Master Thesis

*(Please note, this is a work in progress, the text and the graphical layout will be updated)*

**Outline**

* Chapter 1: Introduction
  + Motivation / Introduction *(what was already done, why this work even exists, how this relates to other IDEs, what is Frege)*
  + Goals *(What I strived to achieve in this work, what subset of features should make / made it here, why exactly this, how this relates to projectional IDEs – MPS, why choose a projectional IDE, what is MPS – very shortly)*
  + Outline *(what other chapters are here, what can be found there)*
* Chapter 2: Meta-programming system / JetBrains MPS
  + - *what it is, what it can do, links & references*
    - *difference between projectional and classic text-based IDEs, how this relates to MPS, what makes MPS different*
    - *examples (images), what it would be like to use MPS, so the reader has a clearer picture*
    - *working with MPS, structure, editor, constraints, behaviour, checking rules, typesystem, textgen – generic explanation + exemplar language (not necessarily tutorial, but the explanation should be clear and how it would be used)*
      * *Probably shouldn't go into too much detail for typesystem and editor? - will be explained again on Frege-IDE (how it was done + options and possibilities)*
* Chapter 3: Frege
  + *what is Frege*
  + *Frege vs. Haskell*
  + *Frege language, examples, and what we decided to support in 'Frege-IDE'*
* Chapter 4: The grammar
  + *Official Frgege grammar, or references (appendix?)*
    - *This is too abstract. Maybe pick only some interesting parts?e.g. “this is usually solved as...”, “on this level it means this and this is supported”...?*
  + *Subset related to what we decided to support in “Frege-IDE”*
    - *(Analysis, what parts of the grammar we decided to cut off, what couldn't be cut off, hot it relates to everything)*
* Chapter 5: Frege in MPS
  + *Grammar transformation for MPS structure aspect, design of the Frege structure, how and why (analysis + design decisions)*
  + *Editor*
    - *“Visuals” (FunctionName FunctionParameters = Expression)*
    - *Editor transformation and substitution actions*
    - *Lacking documentation on the MPS part.*
  + *References, scopes*
    - *(Constraints + behaviour)*
  + *Typesystem*
    - *Expression evaluation*
    - *TBA*
* *Chapter 6: Evaluation*
  + *What we strived to achieve, again? Recapitulation?*
  + *User-friendliness of the “Frege-IDE”, against classical text-based IDEs, advantages and disadvantages*
    - *(Should be probably mentioned user's time investment into learning to use the IDE)*
  + *Limitations*
    - *Something couldn't be done easily, e.g. Enter key-press does not always create a new line, ambiguity*
    - *Not all transformations are possible, e.g. rewriting f x y = x + y to f , x, y :: Int -> Int -> Int (by removing = and adding , , ::) is basically impossible to cover – it is not a text editor*
    - *Intentions – for the example above intention is an option to cover such transformation*
    - *(Mention built-in JAVA-like language, that also has these limitations)*
  + *Experience*
    - *What I found MPS lacking about, flexibility, not much detailed documentation*
* *Conclusion*
  + *Are projectional IDEs good for functional languages?*
  + *Is “Frege-IDE” usable?Future work, possible extensions.*
* *References*
* *Appendix*
  + *Frege formal grammar*
  + *Source code description*
  + *Examples*
  + *User manual for “Frege-IDE”*
* Chapter 1: Introduction
* Motivation / Introduction *(what was already done, why this work even exists, how this relates to other IDEs, what is Frege)*

(Classic intro,a lot of wiki references to say the obvious and well known.)

(what are IDES) (<https://en.wikipedia.org/wiki/Integrated_development_environment>) Integrated development environments (abbreviated IDEs) are a set of software applications that provide tools and facilities to computer programmers. They greatly ease the process of software development, providing features like intelligent code completion, syntax highlighting, build automation tools, debugger, etc. (…add anything??)

(what is Frege) (<https://en.wikipedia.org/wiki/Frege_(programming_language)> Frege, named after the German mathematician, ~~logician and philosopher~~ [Gottlob Frege](https://en.wikipedia.org/wiki/Gottlob_Frege), is a functional language heavily based on Haskell (add ref. link) for the Java virtual machine. It is considered a Haskell dialect, sometimes called “a Haskell for the JVM”. (…todo)

known IDEs

There are several IDEs for Haskell, not many specifically built to support Frege. Examples can be found at <https://wiki.haskell.org/IDEs>, most of them providing mainly syntax highlighting, macros and project management features, while some also provide more advanced features like code completion or type evaluation and inspection.

That being said, all of the notable IDEs are text-based, (add explanation?

See <https://en.wikipedia.org/wiki/Structure_editor>

…most source code editors are instead text editors with additional features such as syntax highlighting and code folding, rather than structure editors. The editors in some integrated development environments parse the source code and generate a parse tree, allowing the same analysis as by a structure editor, but the actual editing of the source code is generally done as raw text.

). In this work, however, we try to approach the topic from a different view and try to design a projectional editor (/IDE) specifically for the Frege programming language.

(what is projectional editor) quoting <https://en.wikipedia.org/wiki/Structure_editor>

A projectional editor (aka structure, structured editor), is a document editor that is cognizant of the document’s underlying structure. They are usually used to edit hierarchical or marked up text, computer programs, diagrams, and any other type of content with clear and well-defined structure. While for the most computer programs a conventional text-based IDE may be more suitable, for specific programming languages, especially DSL (<https://en.wikipedia.org/wiki/Domain-specific_language>) a projectional editor might prove to be a more effective tool.

<https://martinfowler.com/bliki/ProjectionalEditing.html> *- additional info sources*

In this work we want to design a projectional editor supporting a certain subset of the (features of) Frege programming language and examine what are the advantages and disadvantages of such approach over creating a conventional text-based IDE.

* Goals (What I strived to achieve in this work, what subset of features should make / made it here, why exactly this, how this relates to projectional IDEs – MPS, why choose a projectional IDE, what is MPS – very shortly)

As was already mentioned in this paper, in this work we design a projectional editor for a subset of (features of) Frege programming language.

We have chosen JetBrains® *(legality, trademarks, etc.? check)* MPS platform as an underlying tool for designing our IDE. (briefly - why:) MPS (standing for “Meta-programming system”) is an open source software solution allowing developers and language designers to create different kinds of projectional IDEs. It is a so called language workbench (A language workbench is a software development tool designed to define, reuse and compose domain-specific languages together with their integrated development environment. - <https://en.wikipedia.org/wiki/Language_workbench>)

It allows for creating both simple and complex languages, especially DSLs, while also allowing extending existing ones, when the languages available do not meet the needs of a developer. *(quoting MPSLW I.)* MPS has a large set of features allowing for designing editors which closely resemble those from conventional, text-based, IDEs.

Frege, based on Haskell language, has rather many syntactic (and semantic) constructs for this work to be able to include them all. We have therefore focused our attention only on the most important features worth examining, such as function declaration and definition, operators and (custom) datatypes. Our ideal IDE will have a user-friendly editor that should emulate normal text editing and writing code in the way that most Frege and Haskell developers are used to. This should be accompanied by a context help, sometimes referred to as ‘intellisense’, which would allow for referencing already defined functions, operators, variables, etc., in the corrects spots in the code. Last, but not least, we will strive for a type checker, which would be able to find small mistakes in the code, such as calling a function with illegal arguments, or evaluate type of an expression.

* *Outline (what other chapters are here, what can be found there)*

More about MPS tool is written in chapter 2, which examines what MPS is, what it can do and what are its limitations. The chapter examines a project structure in MPS, how to define an editor for a simple language and how to tackle certain problems.

Frege is described more in chapter 3, where we also take a look into what features we actually wanted to implement in our IDE.

Chapter 4 examines Frege grammar and how it relates to the features we want to support. We describe how the grammar should be transformed into concepts in MPS tool on the high-level. We provide a brief analysis and try to explain our decisions.

Chapter 5 is dedicated to the concrete work implementation. We show how exactly the grammar from chapter 4 was transformed into MPS concepts, how the editor was designed, how we tackled problems with references and context help. We conclude the chapter with type system, where we describe some of the more interesting algorithms used in the work, such as the one to evaluate type of expressions or inference types of function arguments, when provided with the function’s annotation.

We evaluate our decisions in chapter 6 where we also examine the advantages and disadvantages of the implemented projectional editor over standard text-based IDEs.

A brief summary of the whole work may be found in the conclusion, where we strive to answer the final question, whether projectional IDEs are good for functional languages.

After that, references of the used literature may be found and appendix, which provides some of the more interesting source materials.

* Chapter 2: Meta-programming system / JetBrains MPS
  + - *what it is, what it can do, links & references, difference between projectional and classic text-based IDEs, how this relates to MPS, what makes MPS different*

<https://www.jetbrains.com/mps/>, <https://www.youtube.com/watch?v=XGm_khXZl44&feature=youtu.be>

Meta-programming system, or MPS, is an open-source language workbench that focuses on DSL (domain-specific languages). [JB-YT] It is a software engineering tool which helps its users to create a new language and then to write other programs in that language. [FC] MPS has a wide range of users, with areas currently being applied in including electrical engineering, data mining, insurance industry and other. The tool can be used to create new languages as well as extending existing ones. (It should be noted that only the languages created in MPS tool may be further extended.) Programs written in the defined languages may then be conveniently transformed into pure text in a specific, usually generic-purpose language. [JB-YT]

MPS is a projectional editor, which means it does not treat the document as a text, but rather as structured concepts. This allows its users to create languages which involve non-parsable notations, such as decision tables, diagrams, and other controls. [JB-YT] Additionally, several editors may be specified for a single language, thus allowing users to switch between different visual representations of a document.



2.1 View of ‘math’ demo language in MPS which includes non-text elements to extend the existing Java language

Traditional process of compiling written code involves lexers and parsers to read programs, which are expressed as text files. The two combined then transform the code into abstract syntax tree (AST) data structures [FC], from which in the process of semantic analysis and code generation an executable program is created. A similar technique is used in the text-based IDEs. In contrast, in MPS, the user works with AST directly, therefore completely omitting the process of lexical analysis and subsequent parsing. [FC] This brings certain advantages:

* It may be easier to extend an existing language.

Extending lexers and parsers to accommodate for the changes in a language requires a certain set of skills and a deeper knowledge of the language’s grammar. The process is complicated, since it requires a programmer to keep track of the possible ambiguities that may arise when defining new grammar rules for a parser.[FC] (See, so called, ‘Dangling else’)

On the other hand, in MPS, the process usually only requires defining new concepts that can act as AST nodes and specifying places in the corresponding AST, where the new nodes can be created.

This also means that in MPS we can combine syntax of several different languages and introduce no syntax ambiguities whatsoever. (This however still may look ambiguous to the user, if there are several different concepts with the same textual representation.)

* We can check for type errors and other mistakes in the code at almost any time.

When writing a code in a traditional environment, one has to define a specific set of rules to deal with the incorrect syntax. Code being currently written means, it almost certainly cannot be correctly evaluated by the defined parser, so in an example case such as below, we will not be able to tell the user right away that the integral and string types are incomparable between themselves:

f = if 1 = "hello" then |

(syntax error on '=')

This cannot happen in MPS, since the code is already ‘parsed’. Even though ‘then’ and ‘else’ branches are not set yet, we can already check for errors in ‘condition’ node (1 = "hello"). This is also useful when designing a smart code completion feature, which requires a certain knowledge of the surrounding code context. This is impossible if we cannot parse the code beforehand, unless certain ‘tricks’ are used.

****

2.2 Exemplar AST for a simple arithmetic expression

Since a user may in MPS work only with AST directly, several features are supported to allow easy transformations of ASTs and allow creating IDEs that may even closely resemble the traditional, text-based, ones. What this means is that we can define actions how the AST should transform when, for instance, a certain node (or a whole subtree) is deleted, a certain specific text is written at the end or a beginning of a visual representation of a node, etc. The MPS actions are described further in this chapter.

* *working with MPS, structure, editor, constraints, behaviour, checking rules, typesystem, textgen – generic explanation + exemplar language (not necessarily tutorial, but the explanation should be clear and how it would be used)*

A project in MPS is divided into two main categories: solutions and languages.

A ‘language’ is the user defined language. It may represent a completely new language or an extension of an existing one. There may be several different languages defined in a single MPS project. They can act as an extension of each other or be completely independent languages.

A ‘solution’, on the other hand, is a part of the project that represents documents, or a code, written in (usually) one or more of the defined languages. Sometimes the solution only acts as a runtime support for one (or more) of the defined languages to be used, for example, in a code generation process.

******

2.3 Logical view of a project in MPS

Solution is a set of models. They act as packaging units that make it possible to reference the corresponding set of models from other solutions of languages. [FC]

Model is simply a set of ASTs. In the languages used by a solution, a certain AST nodes may be set to act as a root node. Model then basically consists of such AST root nodes. Each AST is basically analogous to a single source file in traditional programming paradigm. [FC]

Language describes what types of ASTs can be created with the language. [FC] Furthermore, it includes visual representation of each node, AST transformation actions, syntax and semantic rules and many other ‘settings’. It is separated into several categories, named aspects. We will describe only the most important ones that were also used in this work.

**Structure**

<https://confluence.jetbrains.com/display/MPSD20173/Basic+notions>

Structure aspect of a language allows to define structure of possible ASTs that can be expressed with the corresponding language. [FC] It defines what kind of nodes may be used in user models, what properties, children and references they may have. [JB-D]

Key notion in this aspect is a ‘concept’. Concept represents a sort of a class of AST nodes. It specifies what children and properties all nodes of the ‘class’ have, or can have. It closely resembles working with classes and instances in many popular object-oriented programming languages. In this analogy, a concept is basically a class, which extends another class. (All concepts have to extend at least BaseConcept, similar to Java language, where each class extends Object class.) It may be defined as abstract, in which case no AST nodes may be created directly for such concept. We may define interfaces, which concepts may implement, concept properties, which are analogous to fields and attributes of primitive types (this is not exactly true as property may be of type string or enumeration), children, which represent public fields of object types (i.e. referencing an AST node of a certain type with ownership, not necessarily exclusive) and references, which are similar to children but without the ‘ownership’ part.

Each concept may specify visual appearance of the node (more in ‘editor aspect’), implement methods (behavior aspect) and specify additional syntax constraints on where they can actually be placed in AST (constraints aspect).

******

2.4 Definition of a new concept in MPS

We start by naming a concept. This is similar to naming a class in languages like Java and must follow a similar set of naming rules.

The ‘extends’ clause provides a reference to the super-concept. By default, concepts are created with BaseConcept as their super concept, but this can be changed to a more specific super-concept. This basically encodes ‘is-a’ relationship in UML. [FC] If concept A extends concept B, it indicates that concept A has all of properties, children, references, methods, and definitions from all other language aspects, as B.

Concept may implement interface(s) by using ‘implements’ clause. More about interfaces is discussed later in ‘structure’ chapter.

A concept may be set for its instances to act as root AST nodes. As it was mentioned before, a solution is a set of models all of which consist of root AST nodes. Only the concepts denoted as such may be used that way.

Alias acts as string that triggers a built-in auto-completion menu. If the name is unambiguous (i.e. not a prefix of another item in the menu), an instance of the concept is immediately created. More about menu is discussed in editor aspect of language.



2.5 Auto-completion menu in MPS

Properties hold values that are owned by the concept. They are analogous to UML class properties. [FC] They are either:

* Primitive type – integer, boolean, or string
* Enumeration type – a custom ‘enum data type’ may be created in structure aspect
* Constrained data type - a custom ‘constrained data type’ may be created in structure aspect, which is basically a string compared to a defined matching regular expression

It may be interesting to note that primitive types can be derived from using the other two options.

Children resemble aggregation relationship known from UML. They are other concepts with defined cardinality ([1], [0..1], [1..\*] or [0..\*]) and type. A concrete instance (i.e. a node) must be a child of an exactly 1 other node, with an exception of root nodes. This way their life depends on their parent’s life meaning that removal of a certain node from AST removes all of its children as well as children of their children, recursively down to leaf nodes.

Expressing a relationship between nodes can be also done via references. It is only possible to create a reference to a node if this node already exists in the corresponding AST. [FC] Contrary to children, cardinality can take here only two forms: [0..1] meaning an optional reference and [1] meaning a required reference to a node of the given type.

Where would a reference be used? Consider a following piece of code in Frege:

f = 7

g = 1 + f

We could express the formed AST in many different ways, but let us imagine for the sake of simplicity a root node consisting of statement nodes. Both f = 7 and g = 1 + f are statements. It is easy to imagine 1 + 2 as a tree with node + having 2 children: “1” and “2”. But in the case of 1 + f it is less clear what f is. Using a reference here might be helpful. We already have a statement declaring, what f is, in the AST. In 1 + f we are only applying an existing function f. Such AST might then look like this:



2.6 Simplified AST for statements with a reference

Interface concepts are a mechanism to declare concept characteristics that can be used across several concept types. [FC] Unlike concepts, we cannot define alias for them nor can they extend concepts, only other interfaces. They are mostly used for grouping properties that are commonly used together and passing them onto necessary concepts. [FC]

**Editor**

[**https://confluence.jetbrains.com/display/MPSD33/Editor**](https://confluence.jetbrains.com/display/MPSD33/Editor)

Editor aspect is responsible for rendering and editing ASTs by the user of the language being created. This includes textual and graphical representation of each AST node and certain AST transformation actions. This aspect is what makes MPS a projectional editor, rather than using lexers and parsers to process the user-written code.

****

2.7 Concept editor in MPS

The easiest way to define an editor for a language is to define an editor for each concept (called ‘concept editor’). There may be several different editors defined for a single concept, which offers different views of the same concept for different needs. [FC] If no editor is defined for a given concept, a default one will be provided by MPS.

On the other hand, editor component is an editor responsible for rendering and editing only a part of a node. [FC] It does not focus on any single concept and as such may be reused across several concept editors to render certain parts similarly. [FC]

As most aspects in MPS, editor is too defined by writing a code in a special language provided by MPS – in this case jebtrains.mps.lang.editor. Furthermore, the code is written inside a projectional editor – to define a specific projectional editor for the new language being created.

The editor language is used to specify editor cells. An editor cell is typically used to render a certain portion of a node over a rectangular region in MPS editor window. The types of editor cells include:

* Constant cells – These are used to render keywords and other constant text in editor. [FC]
* Property cells – They render content of a specific property of a concept for which the editor is being defined. Editing such a cell in the editor window for a concrete AST is immediately reflected in the given property of the corresponding AST node. The cell provides automatic binding to the concept’s property.
* Child cells – Child cells delegate the rendering of a specific concept’s child (or a set of children) to their corresponding concept editors. The concrete behavior of such a cell depends on the child’s cardinality:
  + [1] – the editor cell is always present
  + [0..1], [0..\*], [1..n] – child nodes are bound to their corresponding editors and removing a child in MPS editor window results in removing it from the parent node of the corresponding AST as well
  + [0..\*], [1..\*] – the children’s corresponding concept editors are separated by a specified (textual) delimiter
* Referent cells – These cells are used to display an attribute of a referenced node from the given concept (see references in “structure” above). As in case of property cells, they are mapped to a certain property of the referenced node in the AST. However, they can only reflect the property of the original node, but not affect it.

******

2.8 Exemplar usage of the editor defined for Java-like language - changes in the name property of the original node (left) result in immediate re-rendering of the referent cell (right)

* Collection cells – Wrapper-like cells to contain other editor cells. They affect visual arrangement of the cells being rendered. There are 3 types:
  + Horizontal cells
  + Vertical cells
  + Indent cells
* Model access cell – This type of cell is used to modify a state of a node(s). We can define custom getter and setter for the cell, or just a custom getter, to change the behavior of our editor.

There are several other types of editor cells. Here we only described the most-used ones.

To change the way an editor cell is rendered, we can use editor styles. Applying editor style could be described as analogous to applying CSS *(Cascading Style Sheets)* styles to DOM nodes in HTML and XML documents. This allows us to change colors of a text, background color, spacing, padding, functional aspects such as editor cell being editable, read-only, etc.

******

2.9 Editor style for a selected editor cell

While all of this allows us to create a visually appealing editor, it still may not come as a user-friendly, especially to developers used to program in text-based IDEs. To simulate many types of behavior of classic text editors, MPS comes with a notion of actions. These were originally part of Action aspect of MPS, but since they are tightly related to the editor, they were moved in newer versions of MPS into this aspect.

To better explain the need of editor actions, consider an exemplar expression from figure (image?) 2.2: (7 + 1) \* 2 + 3. We already know what an AST would look like for such an expression, however, in MPS, without actions, a user would need to specify the AST him or herself.

First, we would need to add addition binary expression node, then for the right operand, literal ‘3’ while for the left operand another binary expression - multiplication, for which we would need to specify the operands again, and so on.

With a careful implementation of actions, however, we could let the user type the expression from ‘left to right’ while transforming the AST automatically for him. To simplify, let us forget about brackets for a while and try to type a slightly different arithmetic expression: 7 – 1 \* 2 + 3.

|  |  |
| --- | --- |
| First, the user types in ‘7’. That is a very simple unary expression and no further work is to be done here. |  |
| Then, user hits ‘-’. We immediately create a binary expression subtree, where root is the operator ‘-‘. ‘7’ is put as its left operand and we set focus on the right operand, so the user may edit that. |  |
| User types in ‘1’, which only concludes the editing of the right operand and we are done there. |  |
| Then, however, follows ‘\*’. We are now editing the right child of AST of the expression “7 – 1”. So we take a look at the parent’s operator’s precedence: - is definitely less precedent than \*. So we create a subtree for binary operator \*, put 1 as its left child and set focus on the right child. The subtree is placed in 1’s stead (i.e. right child of AST of the expression “7 – 1”). |  |
| Then comes 2, which only finishes editing of the right child of \* subtree. The twist comes, however, when user types in ‘+’. We take a look on the parent’s operator, which is \*. This has a higher precedence, so a new subtree has to be created elsewhere now. The parent of the \* node is, however, -. While – has the same precedence, as +, all of the operators are left associative, which means, we have to create our new subtree even on the higher level. We create the subtree, put the current AST for the expression “7 – 1 \* 2” as its left child and set focus on its right child. Typing ‘3’ only finishes editing of the right child and we are done. |  |

**Frege features – later!**

~~Frege, based on Haskell language, has rather many syntactic (and semantic) constructs for this work to be able to include them all.~~ We, therefore, focus only on the parts of the Frege that make it so popular functional language. ‘Syntactic sugars’ are for the most part omitted, as well as monads, which make Frege appear less functional and a more of an imperative programming language. To include more advanced features, like context help (references) and type system checking, we also had to keep the complexity of the work reasonably small (feasible) and thus concepts like classes and instances had to be omitted as well.

The following part includes the high level concepts of the supported Frege features.

(graphics / visuals.)

**module** fregeide.Skeleton

* A Frege program, or more specifically, a module, has to begin with a concept declaring its name. This allows for a module to be imported in other Frege modules while also making it clear to a developer (reading a program) what it does.

**module** fregeide.Skeleton(module fregeide.Submodule, ff, gg, hh)

* A module may specify which parts of the definition are visible to the outside (meaning: encapsulation). We may export functions, operators and declared datatypes. Additionally it is possible to re-export imported modules (see below).

**import** fregeide.Submodule

* We may import (one or more) modules inside the current one. This allows for usage of the imported functions, operators and datatypes.

**import** fregeide.Submodule as Sub (xx, yy)

* It is possible to specify which functions, operators and datatypes should be brought into the current namespace (we assume for simplicity there will be no naming conflicts), and which should be referenced by providing their full qualified names.

Infixr, function definition, annotation, type, data