

# Metaprogramming with Macros

Eugene Burmako  
LAMP, I&C, EPFL

**Abstract**—Macros realize the notion of textual abstraction. Textual abstraction consists of recognizing pieces of text that match a specification and replacing them according to a procedure.

In the focus of the study are syntactic macros in lexically scoped programming languages. We identify the problems of *hygiene* and *referential transparency* and describe the solutions employed in Template Haskell [1], Nemerle [2] and Racket [3].

We discuss integration of hygienic macros into statically typed languages and propose to improve upon state of the art by providing a type system for syntax templates and uncovering synergies with high-level language features such as path-dependent types and implicits [4].

**Index Terms**—metaprogramming, macros, quasiquotes, hygiene, referential transparency

## I. INTRODUCTION

**P**ROCEDURAL abstraction is pervasive. Factoring out parameterized fragments of programs into procedures is a conventional best practice.

Modern programming languages integrate the notion of procedures into their semantics. Procedures are viewed as independent programs that can communicate with the main program. As of such they can be manipulated as units, and big procedures can be built from the smaller ones. This is a powerful way to manage complexity of software systems.

Proposal submitted to committee: September 3rd, 2012;  
Candidacy exam date: September 10th, 2012; Candidacy exam committee: Christoph Koch, Martin Odersky, Viktor Kuncak.

This research plan has been approved:

Date: \_\_\_\_\_

Doctoral candidate: \_\_\_\_\_  
(name and signature)

Thesis director: \_\_\_\_\_  
(name and signature)

Thesis co-director: \_\_\_\_\_  
(if applicable) (name and signature)

Doct. prog. director: \_\_\_\_\_  
(R. Urbanke) (signature)

However procedural abstraction is sometimes not enough, because its manifestations are bound by language syntax and it operates within the semantics of the language.

For example, in most programming languages it is impossible to define short-circuiting logical operators as procedures, because procedures are usually not in control of operational semantics. Another example in this vein is a C-like `for` loop, which supports optional prologue that introduces variables visible in its body. Procedures typically cannot abstract over variable bindings, so they cannot express this language construct.

*Textual abstraction* consists of recognizing pieces of text that match a specification and replacing them according to a procedure. Matched fragments are referred to as macro calls or macro applications, and procedures that transform them are dubbed macros or macro transformers. The process of applying macros is called macro expansion [5].

Having means of textual abstraction in their toolbox, programmers can use a multitude of techniques, some of which are:

- reification (providing programs with ways to treat code as data),
- language virtualization (overloading/overriding semantics of the original programming language to enable deep embedding of DSLs),
- programmable optimization (application of optimizations such as inlining or fusion based on knowledge about the program being optimized),
- static verification (using reified representation of the program and, possibly, its contracts defined alongside the program, to verify invariants at compile time),
- algorithmic program construction (generation of code that is tedious to write with the abstractions supported by a programming language).

One possibility to implement a macro system is having it as a standalone tool operating on character streams. This gives rise to lexical macros. Such a design warrants simplicity of the implementation, but undermines robustness, because macros operating on lexical level have no mechanism that prevents generation of syntactically invalid programs.

Another approach would be to integrate a macro expander into the compiler and have macros work with syntax trees, introducing *syntactic macros*. In this model macro application is a node in the program tree, and macro expansion produces a new node that replaces the macro application without distorting the structure of the program.

Problems inherent to syntactic macros can be divided into two categories: inadvertent variable capture and semantically invalid expansions. The rest of the papers dwells upon these challenges.

## REFERENCES

- [1] T. Sheard and S. Peyton Jones, Template meta-programming for Haskell. ACM SIGPLAN Notices, 2002.
- [2] K. Skalski, M. Moskal and P. Olszta, Meta-programming in Nemerle. Third International Conference on Generative Programming and Component Engineering, 2004.
- [3] E. Barzilay, R. Culpepper and M. Flatt, Keeping it Clean with Syntax Parameters. Scheme and Functional Programming Workshop, 2011.
- [4] M. Odersky, L. Spoon and B. Venners, Programming in Scala 2nd Edition. Artima, 2010.
- [5] E. Kohlbecker, Syntactic Extensions in the Programming Language Lisp. PhD thesis, Indiana University, 1986.