Ph.D. Thesis Proposals (4) on formal Language Processing (MAPi)

Pedro Rangel Henriques
LP@CCTC – DI/Universidade do Minho

scholar year 2010/11

Language Quality assessment from an Attribute Grammar

Supervisor: Pedro Rangel Henriques
Keywords: análise de código, avaliação de software

Abstract:
Specification, Modeling or Design languages, as well as general purpose or domain specific programming languages, are created by Language Engineers to be for software analysts, designers or implementors

Being engineering products, languages must be *good* to satisfy the customer. However, to be fairly assessed it is mandatory to identify important parameters that can be objectively measured. Concerning languages, it is possible to identify the criteria that define the language quality—*legibility* (in terms of learning, writing and understanding) and efficiency—and also the factors that influence these criteria. But those impact factors are not measurable.

On the other hand, it is well known that formal languages are rigorously defined by Grammars (Context Free Grammars for syntax, and Attribute Grammars for semantics). However Grammars are mathematical objects precisely defined, what allows to identify a set of metrics that allow to assess the quality of grammars.

Previous studies have shown that there is no direct relation between context free grammars and the generated language.

The objective of this project proposal is to study metrics for Attribute Grammars (AG) and explore the possibility to infer the language quality from the AG measures. After creating a theoretical framework to reason about grammar quality and its influence over the language, it is important to implement different strategies to refactor both (grammar and language).

The Ph.D. is planned to be divided into the following tasks: analyze and write the state of the art on grammar metrics and language quality; propose a convenient set of AG metrics, and identify as much as possible factors the impact on the language quality; propose a way to relate grammars metrics and language quality; design the architecture of a system to automatize AG evaluation and language assessment; extend the previous system with refactoring facilities; choose case studies and make the system experimental validation.

Conferences (in ranks A to C, according to ERA2010) where the intermediate or final results should be published: PLDI, POPL, ICALP, SLE, LDTA, wLLIC