

Title	自己反映的構文解析系とコンパイルタイムリフレクションに関する研究
Author(s)	加藤, 大志朗
Citation	
Issue Date	2002-06
Type	Thesis or Dissertation
Text version	author
URL	http://hdl.handle.net/10119/930
Rights	
Description	Supervisor:片山 卓也, 情報科学研究科, 博士

A Research on Reflective Parsing System and Compile-time Reflection

Daijiro Kato

School of Information Science,
Japan Advanced Institute of Science and Technology

April 1, 2002

Abstract

The purpose of this paper is to construct a frame work of compiler-compiler for compile-time extensible language systems. This paper consists of three parts. Firstly, a self-extensible formal language system, named Reflective Context-Free Grammar (RCFG), is proposed. Secondly, an incremental construction method for LALR(1) parser is proposed. And additionally, we discuss on a frame work of descriptions of semantics for production rules which are newly introduced and embedded in program texts by users of compilers.

RCFG is one of self-extensible formal language systems, and an extension of Context-Free Grammar (CFG). The extensibility is obtained so as to embed new production rules which are desired to be used in the text following of the embedment. Typical point of RCFG is self-extensibility. To formalize the self-extensibility on the framework of CFG, we introduce a notion, named Augmented Forms (AF), and define derivation relation on AF sequences. In this paper, we establish some properties on RCFG, including the language class which is middle between Context-Free Language and Context-Sensitive Language, and the property that each word derived by given RCFG is also derived by a CFG which production rule set is identical to initial rule set of RCFG augmented with embedded production rules in the word. Especially, the latter property designates the characteristic of RCFG. Additionally, general parsing algorithm for RCFG, which is an extension of Earley's parsing algorithm for CFG, is given. Soundness and completeness of the algorithm are established, and also, the complexity of the algorithm and some restriction methods in order to accelerate parsing process are discussed.

In the second part of the paper, an incremental construction method for LALR(1) parser is discussed. In the method, both of LR(0) graphs and Look Ahead Symbol Sets are calculated in fully incremental manner, without use of any item set informations. Algorithms for the method are presented, and the worst case complexity is discussed. Adding to these, applications of the method for RCFG are discussed. To realize them, we introduce some notions. Firstly, we rearrange a method for incremental construction of LR(0) graphs so as not to use item set information. In conventional method for LR(0) parse table construction, some kind of subsets of items, called core or kernel, are used to identify states. Secondly, we introduce a notion for identification of states, named *MaxInc* which does not contain any information on item sets, instead of kernels. For incremental calculation of Look Ahead Symbol Set (LA), we introduce notions, named *Dependency Domain*, $E\Delta$ for ε -productivity judgment, and some other functions for calculation of (LA) with ε -productivity conditions. These notions essentially contribute to the incremental method proposed in this paper. We establish the equivalence between results obtained by conventional method and by our method.

In the additional part of the paper, we introduce a variation of Syntax Directed Translation (SDT) as a frame work for descriptions of semantics of new production rules which are introduced by users of compilers. As appendent arguments, we discuss on two kinds of implementation model of extended SDT from the point of view of development stages of programming language systems.

Key Words: formal language system, context-free grammar, self-extensible grammar, reflection, parsing algorithm, LALR(1), incremental construction method, compiler-compiler