

発表概要

Packrat Parsingの表引きによる高速化

矢口 拓実^{1,a)} 池袋 教誉¹ 前田 敦司¹

2015年8月6日発表

Packrat Parsing は、広範囲な構文規則を解析できる構文解析手法である。この手法では、解析が成功するまで構文を探索するバックトラックを用い、一度解析した解析対象文字列の位置と結果を保存するメモ化と呼ばれる手法によって解析時間を線形時間に保っている。既存の多くの Packrat Parsing 実装は、PEG (Parsing Expression Grammar) で記述された文法規則を、再帰呼び出しによるバックトラックとメモ化表を用いるプログラムに変換するが、このような処理系の実行速度は、正規表現に基づいて表引きを用いた字句解析器と、同じく表引きとスタックを用いて処理を進める LALR(1) などのパーサルゴリズムの組み合わせと比較すると、一般的に劣っている。本発表では、Packrat Parsing の高速化のため、解析の意味が変化しない範囲で表引きによって構文解析を進める手法を検討する。本発表では、Packrat Parsing の処理の中に可能な限り表引きをとり入れることで、実行速度を向上させる手法を提案する。また、基本的な実行方式としては Medeiros らの提案した仮想マシン方式を改良したものを採用している。既存の手法と比較した性能評価の結果を示す。

Speed Improvement of Packrat Parsing Using Table Look-Up

TAKUMI YAGUCHI^{1,a)} KANATA IKEBUKURO¹ ATUSI MAEDA¹

Presented: August 6, 2015

Packrat Parsing is a parsing technique which is applicable to wide range of grammars. In this method, the parser tries to find a match in grammar rules and backtracks when necessary. The algorithm is guaranteed to run in linear time with a technique called memoization, which records every position in input string it tried to parse and result of parsing procedure for that position. Many of existing packrat parser implementation translates grammar rules described in PEG (Parsing Expression Grammar) to recursive programs which performs backtracking and memoization. Throughput of resulting parser generally tends to be slower than conventional parser implementation which combines a lexical analyzer, which uses table-based automata generated from regular expressions, and a parser, which also uses tables and stacks based on non-backtracking algorithms e.g. LALR(1). In this paper, we propose an implementation method for Packrat Parsing which relies on extensive use of table look-up. In our implementation, we have tried to use table look-up operation as far as possible to enhance throughput. It is based on a variant of PEG abstract machine proposed by Medeiros and Ierusalimsky. We also present performance measurement result in comparison with other existing implementations.

¹ 筑波大学
University of Tsukuba

^{a)} yaguchi@ialab.cs.tsukuba.ac.jp