Journal of Natural Language Processing Vol. 2 No. 2

Design of Efficient Unification for Natural Language

Hideto Tomabechi†

Graph unification has become a central processing mechanism of many natural language systems due to the popularity of unification-based theories of computational linguistics. Despite the popularity of graph unification as the central processing mechanism, it remains the most expensive part of unification-based natural language processing. Graph unification alone often takes over 90% of total parsing time. As the criteria for efficient unification, we focus on two elements in the design of an efficient unification algorithm: 1) elimination of excessive copying and 2) quick detection of unification failures. We propose a scheme to attain these criteria without expensive overhead for reversing the changes made to the graph node structures based on the notion of quasi-destructive unification. Our experiments using an actual large scale grammar and also using a simulated grammar producing different unification success rates show that the quasi-destructive graph unification algorithm runs roughly at twice as fast as Wroblewski's non-destructive unification algorithm.

KeyWords: *Unification, Graph Unification, Non-destructive Unification, Unification-based Grammer, Feature Structure, Parsing*

† Justsystem Scientific Institute, tomabech@justsystem.co.jp