



HOKKAIDO UNIVERSITY

Title	A Chomsky-Schützenberger-Weir Representation : Theorem for Simple Context-Free Tree Grammars
Author(s)	金沢, 誠
Description	ERATO 세미나2010 : No.31. 2011年1月11日
Relation	2010年度科学技術振興機構ERATO湊離散構造処理系プロジェクト講究録. p.191-192.
Issue Date	2011-06
Doc URL	https://hdl.handle.net/2115/48451
Type	conference presentation
File Information	31_all.pdf



ERATO セミナ 2010 - No. 31

A Chomsky-Schützenberger-Weir Representation Theorem for Simple Context-Free Tree Grammars

金沢 誠
国立情報学研究所

2011/1/11

概要

Analogues of the Chomsky-Schützenberger Theorem for context-free languages have been obtained for two distinct classes of languages, both of which have sometimes been identified with the informal notion of “mildly context-sensitive” languages.

In 1988, Weir showed that each tree-adjoining language is a homomorphic image of the intersection of two Dyck languages over $2n$ pairs of brackets and a regular set, where the two Dyck languages are related by a permutation of the alphabet that maps $[_{2i+1} \]_{2i+1} \ [_{2i+2} \]_{2i+2}$ to $[_{2i+1} \]_{2i+2} \ [_{2i+2} \]_{2i+1}$ for each $i = 0, \dots, n-1$.

Recently, Yoshinaka et al. (2010) obtained a representation theorem for each level in the two-dimensional hierarchy of multiple context-free languages, which subsumes tree adjoining languages and much more. They introduced the notion of “multiple Dyck languages” (for each dimension and rank), and showed that every multiple context-free language of a given dimension and rank is a homomorphic image of the intersection of a multiple Dyck language and a regular set.

In this talk, I generalize Weir’s representation theorem for the tree-adjoining grammars to the simple (i.e., non-deleting and non-duplicating) context-free tree grammars, which correspond at the level of string languages to the “well-nested” multiple context-free grammars. The well-nested multiple context-free languages form a large subclass of the multiple context-free languages with some interesting distinguishing properties. In this representation, each string language generated by a context-free tree grammar of rank m is obtained as a homomorphic image of the intersection of two Dyck languages over $(m+1)n$ pairs of brackets and a regular set, where the two Dyck languages are related by a permutation of the alphabet that maps $[_{m_i+1} \]_{m_i+1} \ [_{m_i+2} \]_{m_i+2} \ \dots \ [_{m_i+m} \]_{m_i+m}$ to $[_{m_i+1} \]_{m_i+2} \ [_{m_i+2} \]_{m_i+3} \ \dots \ [_{m_i+m} \]_{m_i+1}$

for each $i = 0, \dots, n - 1$. This is obtained as a consequence of a representation theorem at the level of tree languages of simple context-free tree grammars, where a tree analogue of Dyck languages plays the role of the usual Dyck languages.