Decomposition of CFT(S) Transformations with Look-ahead

Tamás Hornung and Sándor Vágvölgyi

We generalize Engelfriet's decomposition result: $T^R = T \circ LH$, where T^R, T, and LH denote the class of tree transformations induced by top-down tree transducers with regular look-ahead, top-down tree transducers, and linear homomorphisms, respectively (see [1]).

A top-down tree transducer can be considered as a nondeterministic recursive 'program' that acts on trees and generates trees. We can describe the 'program' as a grammar which operates on a special storage type. We consider Engelfriet's theorem replacing top-down tree transducers by CFT(S) transducers, that is by context-free tree transducers which operate on a storage type S.

In a regular tree (RT) grammar (see [5]), the nonterminals have rank 0. The context-free tree (CFT) grammars are obtained from RT grammars by allowing nonterminals of rank greater than 0. The CFT grammars can be considered with two modes of derivation: in side-out (call by name) and outside-in (call by value). A CFT grammar with these derivations is said to be inside-out tree (IO) grammar and outside-in tree (OI) grammar (cf. [3]).

The concept of storage type is introduced in [2] and [4]. Roughly speaking, a storage type S consists of a set of input elements and a set of configurations. The input elements can be encoded as configurations. The configurations can be tested by predicates and transformed by instructions.

Let MOD be the set RT, IO, OI, CF of modifiers where CF abbreviates the type of context-free grammars. Let X range over MOD. An X(S) transducer is an X grammar in which every rule is provided with a predicate, and every nonterminal of the right-hand side of the rule has an instruction. Considering a derivation of the X(S) transducer, each occurrence of a nonterminal A is associated with a configuration c (different occurrences may be associated with different configurations). A rule of the X(S) transducer can be applied to the tuple A(c) as a rule of an X grammar can be applied to the nonterminal A, provided the test specified by the rule holds for c. The new configurations for the nonterminals of the right-hand side of the rule are obtained by transforming c according to instructions also specified in the rule. The initial nonterminal of the grammar is associated with a configuration corresponding to an input element. Thus, the X(S) transducer induces a transformation from the input set to the set of terminal trees or strings. The class of transformations induced by X(S) transducers is also denoted by X(S) (cf. [4]).

Top-down tree transducers are RT(TR) transducers, where TR is a particular storage type, called the tree storage type. For this storage type, the root of trees can be tested and the trees can be transformed into their immediate subtrees. Top-down tree transducers are the same as RT(TR) transducers.

The concept of storage type S with look-ahead is introduced in [2] and [4] as a generalization of regular look-ahead, and denoted by S_{LA}. The storage type S_{LA} is obtained from S by adding special tests to the set of predicates of S, so-called look ahead tests. They have the form <B>, where B is a CF(S) transducer. The look-ahead test <B> is true on a configuration c if and only if the transducer B can derive a terminal string from $B_{in}(c)$ where B_{in} is the initial nonterminal of B. We allow Y(S) transducer s in the look-ahead tests (where Y is in MOD) and we write SY instead of SLA in this case (SLA = SCF). Since the class of domains of top-down tree transducers is equal to the class of regular tree languages (see [5]), Engelfriet's decomposition result takes the form RT(TRRT) = RT(TR) o LH. (1) In this paper, we study whether the equation X(SY) = X(S) o LH (2) is true for a storage type S and for the modifiers X, Y in MOD when X differs from CF. We show that X(S) o LH is a subclass of X(SX), but there exists a storage type S such that X(S) o LH is not equal to X(SX). An X(SY) transducer, in which look-ahead tests do not occur in the negated part of tests belonging to the rules, is denoted by X+(SY). The class of transformations induced by X+(SY) is denoted also by X+(SY). It holds for all storage types S and modifier X in MOD CF that X+(SX) = X(S) o LH. (3) Since RT+(TRRT) = RT(TRRT), (3) implies (1), and if X+(SX) = X(SY) then (2) holds true. We show that X(SIO) = X(SRT) = X(SLA) for all storage types S. We give examples for storage types such that X+(SX) = X(SX).

References

[1] J. Engelfriet, Top-down tree transducers with regular look ahead, Math. Systems Theory 10 (1977) 289-303.

[2] J. Engelfriet, Context-free grammars with storage, Technical Report Nr 86-11 (1986), University of Leiden, July 1986.

[3] J. Engelfriet and E.M. Schmidt, IO and OI, I., II., J. Comput. System Sci. 15(1977) 328-353, 16(1978) 67-99.

[4] J. Engelfriet and H. Vogler, Pushdown machines for the macro tree transducers, Theoret. Comput. Sci. 42(1986) 251-368.

[5] F. Gécseg and M. Steinby, Tree automata (Akadémiai Kiadó, Budapest, 1984).