Droste, Manfred; **Vogler, Heiko**
The Chomsky-Schützenberger theorem for quantitative context-free languages. (English)
Zbl 1315.68169
Int. J. Found. Comput. Sci. 25, No. 8, 955-969 (2014).

This contribution presents the Chomsky-Schützenberger theorem for a rather wide class of weighted context-free languages together with some flanking results. *N. Chomsky* and *M. P. Schützenberger* [in: Computer programming and formal systems, Stud. Logic Found. Math., 118–161 (1963; Zbl 0148.00804)] showed that each context-free language can be presented as a homomorphic image of the intersection of a Dyck language (language of well-bracketed expressions over k parentheses) with a regular language. This result was later extended to weighted context-free languages in [*A. Salomaa* and *M. Soittola*, Automata-theoretic aspects of formal power series. New York, NY: Springer-Verlag (1978; Zbl 0377.68039)], where the weights were taken from a commutative semiring. A commutative semiring is a weight structure consisting of two binary, commutative and associative operations, called sum and product, and their units such that the product distributes over arbitrary finite sums. The present contribution extends the Chomsky-Schützenberger result to weighted context-free languages weighted over unital valuation monoids, which are commutative monoids $(A, +, 0)$ togehter with a valuation $v : A^* \to A$ such that (i) $v(\ldots, 0, \ldots) = 0$ and (ii) there exists $1 \in A$ with $v(a_1, \ldots, a_{i-1}, 1, a_{i+1}, \ldots, a_n) = v(a_1, \ldots, a_{i-1}, a_{i+1}, \ldots, a_n)$ for all integers $i \leq n$ and $a_1, \ldots, a_n \in A$. In other words, 1 acts like a unit for the valuation.

In detail, the paper first introduces weighted context-free grammars and weighted pushdown automata both using unital valuation monoids as weight structure. As usual, each production (transition) is assigned a weight from the weight structure. Along a derivation (computation), the weights are combined using the valuation function and the weights of several derivations for the same string are summed up. In the special case where an input string has infinitely many derivations, additional completeness requirements are imposed on the weight structure. It is demonstrated that both models are equally expressive using essentially the standard approach. Next, it is shown that the weights can be separated and applied only by a weighted alphabetic homomorphism. In this way, an essentially unweighted context-free language is obtained, to which the classical result is now applicable. In the end, the homomorphisms (i) obtained as part of the unweighted characterization and (ii) the homomorphism assigning the weights are composed to a new homomorphism. In this manner, the representation of each weighted context-free language as (weighted) homomorphic image of the intersection of an (unweighted) Dyck language and an (unweighted) regular language is obtained. Finally, the paper also studies context-free step functions. A constant context-free function is a weighted context-free language whose support (i.e., the set of words mapped to non-zero elements) forms a context-free language and all weights of the support elements are equal. A context-free step function is a finite sum of constant context-free functions. Step functions proved to be important for logical characterizations of weighted regular languages, and it is demonstrated here that context-free step functions behave differently than regular step functions.

Overall, the paper is very readable, self-contained (together with the companion proofs in the arXiv-version) and should be understandable to anyone with some background on context-free languages and weighted languages. A detailed background on the used weight structures is not necessary to appreciate the results.

Reviewer: Andreas Maletti (Stuttgart)

MSC:

68Q45	Formal languages and automata
68Q42	Grammars and rewriting systems
68Q70	Algebraic theory of languages and automata

Cited in **5** Documents

Keywords:

weighted context-free grammars; weighted pushdown automata; valuation monoids; context-free step functions; homomorphisms; Dyck languages

Full Text: DOI

References:

[1]   Autebert J., Grammar 1 pp 111– (1997)

[2]   Bar-Hillel Y., Z. Phonetik. Sprach. Komm. 14 pp 143– (1961)

[3]   DOI: 10.1145/1805950.1805953 · Zbl 1351.68155 · doi:10.1145/1805950.1805953

[4]   DOI: 10.1016/j.ins.2010.05.020 · Zbl 1205.68198 · doi:10.1016/j.ins.2010.05.020

[5]   DOI: 10.1016/j.tcs.2007.02.055 · Zbl 1118.68076 · doi:10.1016/j.tcs.2007.02.055

[6]   DOI: 10.1142/S0129054111009069 · Zbl 1251.68130 · doi:10.1142/S0129054111009069

[7]   DOI: 10.1016/j.ins.2009.09.003 · Zbl 1183.68337 · doi:10.1016/j.ins.2009.09.003

[8]   DOI: 10.1016/j.tcs.2011.11.008 · Zbl 1245.03060 · doi:10.1016/j.tcs.2011.11.008

[9]   Petre I., Ch. 7 pp 257– (2009)

[10]   DOI: 10.1016/j.scico.2005.02.009 · Zbl 1088.68040 · doi:10.1016/j.scico.2005.02.009

[11]   DOI: 10.1016/S0019-9958(61)80020-X · Zbl 0104.00702 · doi:10.1016/S0019-9958(61)80020-X