Mikhail Barash | Alexander Okhotin

Grammars with two-sided contexts

Turku Centre *for* Computer Science

TUCS Technical Report
No 1090, October 2013

Grammars with two-sided contexts

Mikhail Barash
mikbar@utu.fi
Department of Mathematics and Statistics, University of Turku, *and*
Turku Centre for Computer Science
Turku FI–20014, Finland
Alexander Okhotin
alexander.okhotin@utu.fi
Department of Mathematics and Statistics, University of Turku, *and*
Turku Centre for Computer Science
Turku FI–20014, Finland

TUCS Technical Report
No 1090, October 2013

Abstract

In a recent paper (M. Barash, A. Okhotin, "Defining contexts in context-free grammars", LATA 2012), the authors introduced an extension of the context-free grammars equipped with an operator for referring to the left context of the substring being defined. This paper proposes a more general model, in which context specifications may be two-sided, that is, both the left and the right contexts can be specified by the corresponding operators. The paper gives the definitions and establishes the basic theory of such grammars, leading to a parsing algorithm working in time $\mathcal{O}(n^{\omega+1})$, where n is the length of the input string and ω is the exponent of matrix multiplication complexity.

Keywords: Context-free grammars, conjunctive grammars, contexts, context-sensitive grammars, parsing.

TUCS Laboratory
FUNDIM, Fundamentals of Computing and Discrete Mathematics

1 Introduction

The context-free grammars are a logic for representing the syntax of languages, in which the properties of longer strings are defined by concatenating shorter strings with known properties. Disjunction of syntactic conditions can be represented in this logic as multiple alternative concatenations defining a single symbol. One can further augment this logic with conjunction and negation operations, leading to *conjunctive grammars* [8] and *Boolean grammars* [10]. These grammars are context-free in the general sense of the word, as they define the properties of each substring independently of the context, in which it occurs. Furthermore, most of the practically important features of ordinary context-free grammars, such as efficient parsing algorithms, are preserved in their conjunctive and Boolean variants [10, 11, 13]. These grammars models have been a subject of recent theoretical studies [1, 4, 6, 7, 16, 21].

Recently, the authors [2] proposed an extension of the context-free grammars with special operators for expressing the form of the *left context*, in which the substring occurs. For example, a rule $A \to BC \,\&\, {\lhd}D$ asserts that every string representable in the form BC in a left context of the form described by D therefore has the property A. These grammars were motivated by the well-known Chomsky's [3] idea of a phrase-structure rule applicable only in some particular contexts [3, p. 142]. Chomsky's own attempt to implement this idea by string rewriting resulted in a model equivalent to linear-space Turing machines, which had nothing to do with the syntax of languages. In contrast, the model proposed by the authors [2] is defined using deduction systems, and properly maintains the underlying logic of the context-free grammars. It was found to have a cubic-time parsing algorithm [2]. However, the model allowed specifying contexts only on one side, and thus it implemented, so to say, one half of Chomsky's idea.

This paper continues the development of formal grammars with context specifications by allowing contexts in both directions. The proposed *grammars with two-sided contexts* may contain such rules as $A \to BC \,\&\, {\lhd}D \,\&\, {\rhd}H$, which define any substring of the form BC preceded by a substring of the form D and followed by a substring of the form H. If the grammar contains additional rules $B \to b$, $C \to c$, $D \to d$ and $H \to h$, then the above rule for A asserts that a substring bc of a string $w = dbch$ has the property A. However, this rule will not produce the same substring bc occurring in another string $w' = dbcd$, because its right context does not satisfy the conjunct ${\rhd}H$. Furthermore, the grammars allow expressing the so-called *extended right context* (${\trianglerighteq}\alpha$), which defines the form of the current substring concatenated with its right context, as well as the symmetrically defined *extended left context* (${\trianglelefteq}\alpha$).

The intuitive definition is formalized by deduction of propositions of the form $A\big(u\langle w\rangle v\big)$, which states that the substring w occurring in the context

$u\langle w\rangle v$ has the property A, where A is a syntactic category defined by the grammar ("nonterminal symbol" in Chomsky's terminology). Then, each rule of the grammar becomes a schema for deduction rules, and a string w is generated by the grammar, if there is a proof of the proposition $S\big(\varepsilon\langle w\rangle\varepsilon\big)$.

This paper gives the definition and basic examples of grammars with two-sided contexts, and then proceeds with developing a generalization of the Chomsky normal form for these grammars. Once the normal form is established, it is easy to obtain a parsing algorithm with the running time $\mathcal{O}(n^4)$, which can be improved by employing fast matrix multiplication using the method of Valiant [20].

2 Definition

Ordinary context-free grammars allow using the concatenation operation to express the form of a string, and disjunction to define alternative forms. In conjunctive grammars, the conjunction operation may be used to assert that a substring being defined must conform to several conditions at the same time. The grammars studied in this paper further allow operators for expressing the form of the left context ($\lhd$, $\lhdeq$) and the right context ($\rhd$, $\rhdeq$) of a substring being defined.

Definition 1. *A grammar with two-sided contexts is a quadruple* $G = (\Sigma, N, R, S)$*, where*

- Σ *is the alphabet of the language being defined;*
- N *is a finite set of auxiliary symbols ("nonterminal symbols" in Chomsky's terminology), which denote the properties of strings defined in the grammar;*
- R *is a finite set of grammar rules, each of the form*

$$\begin{aligned} A \to \alpha_1 \,\&\, \ldots \,\&\, \alpha_k \,\&\, \lhd\beta_1 \,\&\, \ldots \,\&\, \lhd\beta_m \,\&\, \lhdeq\gamma_1 \,\&\, \ldots \,\&\, \lhdeq\gamma_n \,\&\\ \&\, \rhdeq\kappa_1 \,\&\, \ldots \,\&\, \rhdeq\kappa_{m'} \,\&\, \rhd\delta_1 \,\&\, \ldots \,\&\, \rhd\delta_{n'}, \end{aligned} \tag{1}$$

  with $A \in N$*,* $k \geqslant 1$*,* $m, n, m', n' \geqslant 0$ *and* $\alpha_i, \beta_i, \gamma_i, \kappa_i, \delta_i \in (\Sigma \cup N)^*$*;*
- $S \in N$ *is a symbol representing well-formed sentences of the language.*

If all rules in a grammar have only left contexts (that is, if $m' = n' = 0$), then this is a grammar with one-sided contexts [2]. If no context operators are ever used ($m = n = m' = n' = 0$), this is a conjunctive grammar, and if the conjunction is also never used ($k = 1$), this is an ordinary context-free grammar.

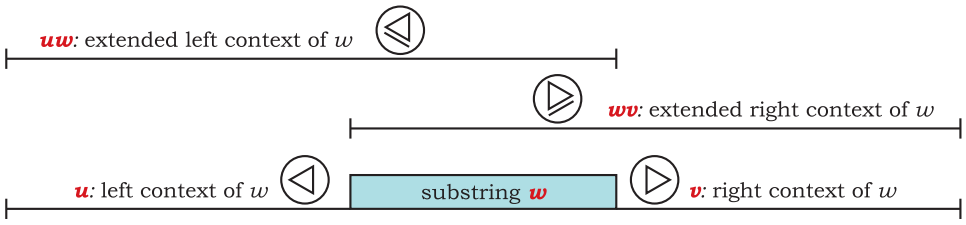


Figure 1: A substring w of a string uwv: four types of contexts.

For each rule (1), each term α_i, $\lhd\beta_i$, $\trianglelefteq\gamma_i$, $\trianglerighteq\kappa_i$ and $\rhd\delta_i$ is called a *conjunct*. Denote by $u\langle w\rangle v$ a substring $w \in \Sigma^*$, which is preceded by $u \in \Sigma^*$ and followed by $v \in \Sigma^*$, as illustrated in Figure 1. Intuitively, such a substring is generated by a rule (1), if

- each *direct conjunct* $\alpha_i = X_1 \dots X_\ell$ gives a representation of w as a concatenation of shorter substrings described by $X_1, \dots, X_\ell$, as in context-free grammars;
- each conjunct $\lhd\beta_i$ similarly describes the form of the *left context* u;
- each conjunct $\trianglelefteq\gamma_i$ describes the form of the *extended left context* uw;
- each conjunct $\trianglerighteq\kappa_i$ describes the *extended right context* wv;
- each conjunct $\rhd\delta_i$ describes the *right context* v.

The semantics of grammars with two-sided contexts are defined by a deduction system of elementary propositions (items) of the form "a string $w \in \Sigma^*$ written in a left context $u \in \Sigma^*$ and in a right context $v \in \Sigma^*$ has the property $X \in \Sigma \cup N$", denoted by $X\big(u\langle w\rangle v\big)$. The deduction begins with axioms: any symbol $a \in \Sigma$ written in any context has the property a, denoted by $a\big(u\langle a\rangle v\big)$ for all $u, v \in \Sigma^*$. Each rule in R is then regarded as a schema for deduction rules. For example, a rule $A \to BC$ allows making deductions of the form

$$B\big(u\langle w\rangle w'v\big), C\big(uw\langle w'\rangle v\big) \vdash_G A\big(u\langle ww'\rangle v\big) \qquad (\text{for all } u, w, w', v \in \Sigma^*),$$

which is essentially a concatenation of w and w' that respects the contexts. If the rule is of the form $A \to BC \,\&\, \lhd D$, this deduction requires an extra premise:

$$B\big(u\langle w\rangle w'v\big), C\big(uw\langle w'\rangle v\big), D\big(\varepsilon\langle u\rangle ww'v\big) \vdash_G A\big(u\langle ww'\rangle v\big).$$

And if the rule is $A \to BC \,\&\, \trianglerighteq F$, the deduction proceeds as follows:

$$B\big(u\langle w\rangle w'v\big), C\big(uw\langle w'\rangle v\big), F\big(u\langle ww'v\rangle\varepsilon\big) \vdash_G A\big(u\langle ww'\rangle v\big).$$

The general form of deduction schemata induced by a rule in R is defined below.

Definition 2. *Let $G = (\Sigma, N, R, S)$ be a grammar with two-sided contexts, and define the following deduction system of items of the form $X\big(u\langle w\rangle v\big)$, with $X \in \Sigma \cup N$ and $u, w, v \in \Sigma^*$. There is a single axiom scheme:*

$$\vdash_G a\big(x\langle a\rangle y\big) \quad (\textit{for all } a \in \Sigma \textit{ and } x, y \in \Sigma^*).$$

Each rule (1) in R defines the following scheme for deduction rules:

$$I \vdash_G A\big(u\langle w\rangle v\big),$$

for all $u, w, v \in \Sigma^*$ and for every set of items I satisfying the below properties:

- *For every direct conjunct $\alpha_i = X_1 \dots X_\ell$, with $\ell \geqslant 0$ and $X_j \in \Sigma \cup N$, there should exist a partition $w = w_1 \dots w_\ell$ with $X_j\big(uw_1 \dots w_{j-1}\langle w_j\rangle w_{j+1} \dots w_\ell v\big) \in I$ for all $j \in \{1, \dots, \ell\}$.*
- *For every conjunct $\lhd\beta_i = \lhd X_1 \dots X_\ell$ there should be such a partition $u = u_1 \dots u_\ell$, that $X_j\big(u_1 \dots u_{j-1}\langle u_j\rangle u_{j+1} \dots u_\ell wv\big) \in I$ for all $j \in \{1, \dots, \ell\}$.*
- *Every conjunct $\lhdeq\gamma_i = \lhdeq X_1 \dots X_\ell$ should have a corresponding partition $uw = x_1 \dots x_\ell$ with $X_j\big(x_1 \dots x_{j-1}\langle x_j\rangle x_{j+1} \dots x_\ell v\big) \in I$ for all j.*
- *For every conjunct $\rhd\delta_i$ and $\rhdeq\kappa_i$, the condition is defined symmetrically.*

Then the language generated by a symbol $A \in N$ is defined as

$$L_G(A) = \{\, u\langle w\rangle v \mid u, w, v \in \Sigma^*, \vdash_G A\big(u\langle w\rangle v\big) \,\}.$$

The language generated by the grammar G is the set of all strings with left and right contexts ε generated by S: $L(G) = \{\, w \mid w \in \Sigma^*, \vdash_G S\big(\varepsilon\langle w\rangle\varepsilon\big) \,\}$.

Consider the following grammar, which defines the language of all strings of the form $a^n b^n c^n d^n$ ($n \geqslant 0$), possibly with a symbol e inserted anywhere in d^n.

Example 1 (cf. example with one-sided contexts [2, Ex. 1])**.** The following grammar generates the language $\{\, a^n b^n c^n d^n \mid n \geqslant 0 \,\} \cup \{\, a^n b^n c^n d^\ell e d^{n-\ell} \mid n \geqslant \ell \geqslant 0 \,\}$:

$$\begin{aligned} S &\to aSd \mid bSc \mid \varepsilon \,\&\, \lhd A \mid Se \,\&\, \rhd D \\ A &\to aAb \mid \varepsilon \\ D &\to Dd \mid \varepsilon \end{aligned}$$

The rules $S \to aSd$ and $S \to bSc$ match each symbol a or b in the first part of the string to the corresponding d or c in its second part. In the middle of the string, the rule $S \to \varepsilon \,\&\, \lhd A$ ensures that the first half of the string is $a^n b^n$. These symbols must have matching symbols $c^n d^n$ in the second half. Furthermore, the rule $S \to Se \,\&\, \rhd D$ allows inserting the symbol e only in a right context of the form d^*.

The following deduction proves that the string *abced* has the property S.

$$\begin{array}{rll}
 & \vdash a\big(\varepsilon\langle a\rangle bced\big) & (axiom)\\
 & \vdash b\big(a\langle b\rangle ced\big) & (axiom)\\
 & \vdash c\big(ab\langle c\rangle ed\big) & (axiom)\\
 & \vdash e\big(abc\langle e\rangle d\big) & (axiom)\\
 & \vdash d\big(abce\langle d\rangle \varepsilon\big) & (axiom)\\
 & \vdash A\big(a\langle\varepsilon\rangle bced\big) & (A \to \varepsilon)\\
a\big(\varepsilon\langle a\rangle bced\big), A\big(a\langle\varepsilon\rangle bced\big), b\big(a\langle b\rangle ced\big) & \vdash A\big(\varepsilon\langle ab\rangle ced\big) & (A \to aAb)\\
A\big(\varepsilon\langle ab\rangle ced\big) & \vdash S\big(ab\langle\varepsilon\rangle ced\big) & (S \to \varepsilon \,\&\, \lhd A)\\
b\big(a\langle b\rangle ced\big), S\big(ab\langle\varepsilon\rangle ced\big), c\big(ab\langle c\rangle ed\big) & \vdash S\big(a\langle bc\rangle ed\big) & (S \to bSc)\\
 & \vdash D\big(abce\langle\varepsilon\rangle d\big) & (D \to \varepsilon)\\
D\big(abce\langle\varepsilon\rangle d\big), d\big(abce\langle d\rangle\varepsilon\big) & \vdash D\big(abce\langle d\rangle\varepsilon\big) & (D \to Dd)\\
S\big(a\langle bc\rangle ed\big), e\big(abc\langle e\rangle d\big), D\big(abce\langle d\rangle\varepsilon\big) & \vdash S\big(a\langle bce\rangle d\big) & (S \to Se \,\&\, \rhd D)\\
a\big(\varepsilon\langle a\rangle bced\big), S\big(a\langle bce\rangle d\big), d\big(abce\langle d\rangle\varepsilon\big) & \vdash S\big(\varepsilon\langle abced\rangle\varepsilon\big) & (S \to aSd)
\end{array}$$

This deduction can be represented as the tree in Figure 2.

Consider the problem of checking declaration of identifiers before their use, which can be expressed by a conjunctive grammar. However, if the identifiers may be declared *before or after* their use, then no Boolean grammar for such a language is known. A grammar with one sided contexts for declarations before or after use has recently been constructed by the authors. Using two-sided contexts, the same language can be defined in a much more natural way.

Example 2 (cf. grammar with one-sided contexts [2])**.** Consider the language

$$\{\, u_1 \dots u_n \mid \text{for every } i,\ u_i \in a^*c, \textbf{ or } \text{there exist } j,k \text{ with } u_i = b^k c \text{ and } u_j = a^k c \,\} \tag{2}$$

Substrings of the form $a^k c$ represent declarations, while every substring of the form $b^k c$ is a reference to a declaration of the form $a^k c$.

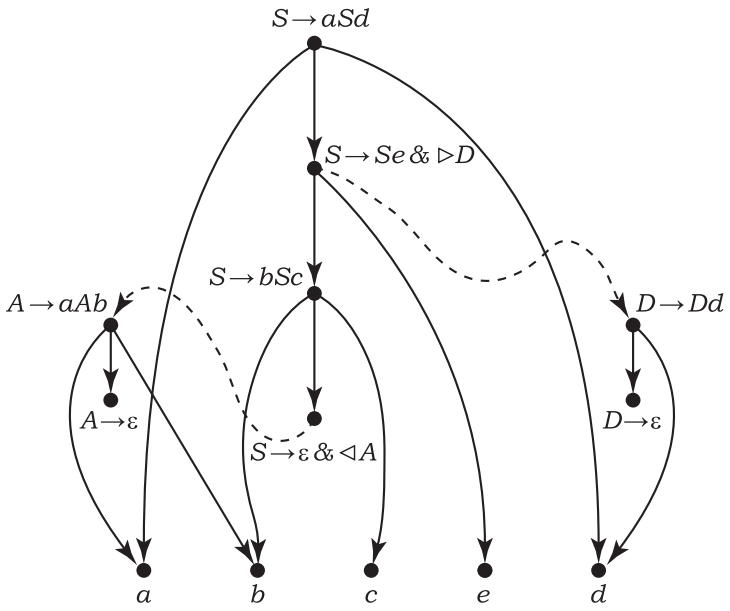


Figure 2: A parse tree of the string *abced* according to the grammar in Example 1.

This language is generated by the following grammar:

$$\begin{aligned}
S &\to AS \mid CS \mid DS \mid \varepsilon \\
A &\to aA \mid c \\
B &\to bB \mid c \\
C &\to B \,\&\, \trianglelefteq EFc \\
D &\to B \,\&\, \trianglerighteq HcE \\
E &\to AE \mid BE \mid \varepsilon \\
F &\to aFb \mid cE \\
H &\to bHa \mid cE
\end{aligned}$$

The idea of the grammar is that S should generate a string $u_1 \dots u_\ell \langle u_{\ell+1} \dots u_{\ell'} \rangle u_{\ell'+1} \dots u_n$, with $u_i \in a^*c \cup b^*c$, if every reference in $u_{\ell+1} \dots u_{\ell'}$ has a corresponding declaration somewhere in the whole string $u_1 \dots u_n$. This condition is defined inductively, and the rule $S \to \varepsilon$ defines the base case: the string $u_1 \dots u_n \langle \varepsilon \rangle \varepsilon$ has the desired property. The rule $S \to CS$ appends a reference of the form $(b^*c \,\&\, \trianglelefteq EFc)$, where the context specification ensures that this reference has a matching *earlier* declaration; here E represents the prefix of the string up to that earlier declaration, while F matches the symbols a in the declaration to the symbols b in the reference. The possibility of a *later* declaration is checked by another rule $S \to DS$, which adds a reference of the form $(b^*c \,\&\, \trianglerighteq HcE)$, where H is used to match

the bs forming this reference to the as in the later declaration.

The next example gives a grammar with contexts that defines reachability on graphs. Whereas Sudborough [19] defined a linear context-free grammar for a special encoding of the graph reachability problem on acyclic graphs with numbered nodes, the grammar presented below allows any graphs and uses a direct encoding. This example illustrates the ability of grammars with contexts to define various kinds of cross-references.

Example 3. Consider encodings of directed graphs as strings of the form

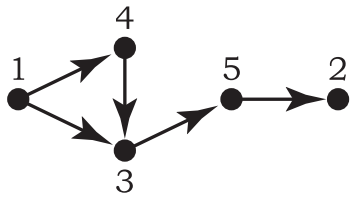
$$b^s\ a^{i_1}b^{j_1}\ a^{i_2}b^{j_2}\ \ldots\ a^{i_n}b^{j_n}\ a^t, \qquad \text{with } s,t \geqslant 1,\ n \geqslant 0,\ i_k, j_k \geqslant 1,$$

where each block a^ib^j denotes an arc from vertex number i to vertex number j, while the prefix b^s and the suffix a^t mark s as the source vertex and t as the target. Then the following grammar defines all graphs with a path from s to t.

$$\begin{aligned}
&S \to FDCA \mid F\\
&A \to aA \mid a\\
&B \to bB \mid b\\
&C \to ABC \mid \varepsilon\\
&D \to B \,\&\, \trianglerighteqslant S \mid B \,\&\, \trianglelefteqslant BCE\\
&E \to aEb \mid DCA\\
&F \to bFa \mid bCa
\end{aligned}$$

The rule $S \to F$ handles the case of s and t being the same node. The other alternative $S \to FDCA$ uses F to match a^s to the tail of an arc $a^sb^{s'}$, and then uses D to generate its head $b^{s'}$. The contexts in the rules for D ensure that there is a path from s' to t as follows: $\trianglerighteqslant S$ handles the case of the next arc located to the right of this point in the string, while $\trianglelefteqslant BCE$ uses E to continue the path by an arc earlier in the list.

For example, consider a directed graph with five vertices 1, ..., 5, and the arcs 3—5, 1—4, 4—3, 1—3, and 5—2, where the path from vertex 1 to vertex 2 is sought, and an encoding for this instance of the problem as a string $b^1\ a^3b^5\ a^1b^4\ a^4b^3\ a^1b^3\ a^5b^2\ a^2$. One of the paths from 1 to 2 is 1, 4, 3, 5, 2, and Figure 3 informally illustrates the parse of this string according to the grammar, which follows that path.



3 Definition by language equations

Ordinary context-free grammars have an equivalent definition by language equations, due to Ginsburg and Rice [5]. This definition is inherited by

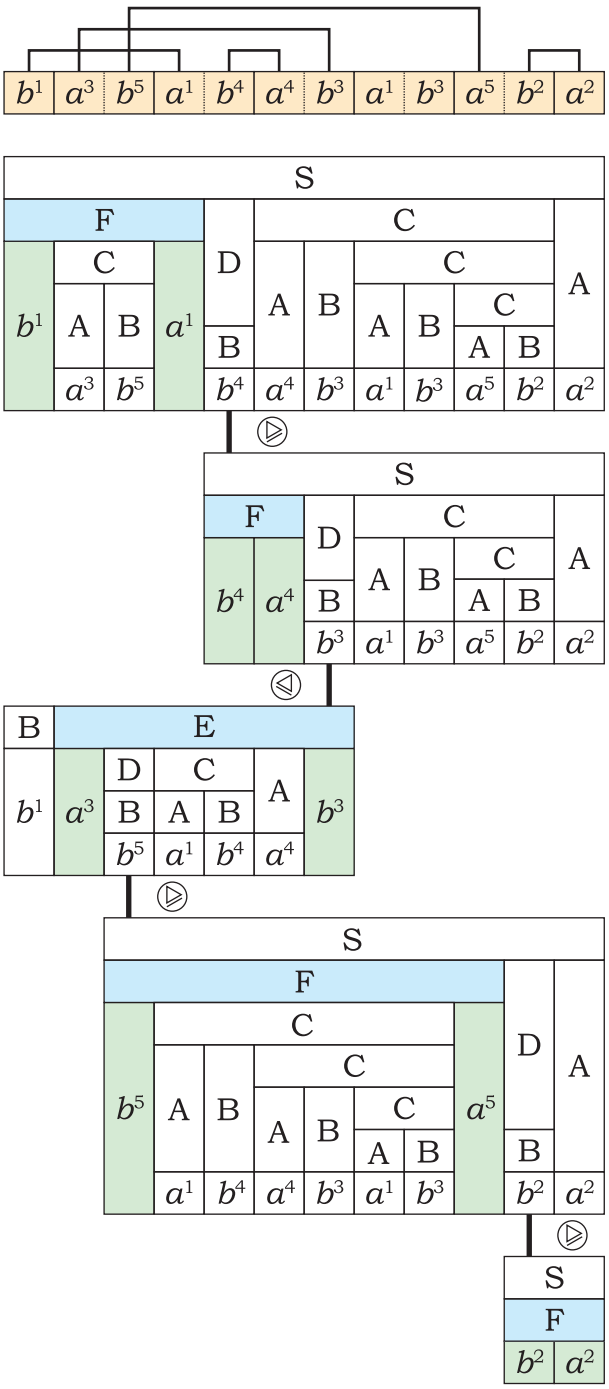


Figure 3: An informal diagram representing a parse of the string $b^1\ a^3b^5\ a^1b^4\ a^4b^3\ a^1b^3\ a^5b^2\ a^2$ according to the grammar in Example 3. Highlighted blocks correspond to the substrings constituting the path 1–4–3–5–2.

conjunctive grammars [9], as well as by grammars with one-sided contexts [2].

This representation is extended to grammars with two-sided contexts by replacing ordinary formal languages with sets of triples of the form $u\langle w\rangle v$, that is, to *languages of triples* $L \subseteq \Sigma^* \times \Sigma^* \times \Sigma^*$. All usual operations on languages used in equations are extended to languages of triples as follows: for all $K, L \subseteq \Sigma^* \times \Sigma^* \times \Sigma^*$, consider their

- *union* $K \cup L = \{\, u\langle w\rangle v \mid u\langle w\rangle v \in K \text{ or } u\langle w\rangle v \in L \,\}$;
- *intersection* $K \cap L = \{\, u\langle w\rangle v \mid u\langle w\rangle v \in K,\ u\langle w\rangle v \in L \,\}$;
- *concatenation* $K \cdot L = \{\, u\langle ww'\rangle v \mid u\langle w\rangle w'v \in K,\ uw\langle w'\rangle v \in L \,\}$;
- *left context* $\lhd L = \{\, u\langle w\rangle v \mid \varepsilon\langle u\rangle wv \in L \,\}$;
- *extended left context*, $\trianglelefteq L = \{\, u\langle w\rangle v \mid \varepsilon\langle uw\rangle v \in L \,\}$.
- *right context* $\rhd L = \{\, u\langle w\rangle v \mid uw\langle v\rangle\varepsilon \in L \,\}$;
- *extended right context*, $\trianglerighteq L = \{\, u\langle w\rangle v \mid u\langle wv\rangle\varepsilon \in L \,\}$.

Definition 3. *For every grammar with contexts $G = (\Sigma, N, R, S)$, the associated system of language equations is a system of equations in variables N, in which each variable assumes a value of a language of triples $L \subseteq \Sigma^* \times \Sigma^* \times \Sigma^*$, and which contains the following equations for every variable A:*

$$A = \bigcup_{\substack{A \to \alpha_1 \,\&\, \ldots \,\&\, \alpha_k \,\&\\ \&\, \lhd\beta_1 \,\&\, \ldots \,\&\, \lhd\beta_m \,\&\\ \&\, \trianglelefteq\gamma_1 \,\&\, \ldots \,\&\, \trianglelefteq\gamma_n \,\&\\ \&\, \trianglerighteq\kappa_1 \,\&\, \ldots \,\&\, \trianglerighteq\kappa_{m'} \,\&\\ \&\, \rhd\delta_1 \,\&\, \ldots \,\&\, \rhd\delta_{n'} \in R}} \left(\bigcap_{i=1}^{k} \alpha_i \ \cap \bigcap_{i=1}^{m} \lhd\beta_i \ \cap \bigcap_{i=1}^{n} \trianglelefteq\gamma_i \ \cap \bigcap_{i=1}^{m'} \trianglerighteq\kappa_i \ \cap \bigcap_{i=1}^{n'} \rhd\delta_i \right).$$

Each instance of a symbol $a \in \Sigma$ in such a system denotes the language $\{\, x\langle a\rangle y \mid x, y \in \Sigma^* \,\}$, and each empty string denotes the language $\{\, x\langle\varepsilon\rangle y \mid x, y \in \Sigma^* \,\}$. A solution of such a system is a vector of languages $(\ldots, L_A, \ldots)_{A\in N}$, such that the substitution of L_A for A, for all $A \in N$, turns each equation into an equality.

This system always has solutions, and among them the *least solution* with respect to the partial order $\sqsubseteq$ of componentwise inclusion on the set $(2^{\Sigma^*\times\Sigma^*\times\Sigma^*})^n$.

For any two n-tuples of languages of triples, let $(K_1, \ldots, K_n) \sqsubseteq (L_1, \ldots, L_n)$ if and only if $K_i \subseteq L_i$. Its least element is $\bot = (\varnothing, \ldots, \varnothing)$.

Consider a system of language equations of the form

$$X_i = \varphi_i(X_1, \ldots, X_n) \qquad (1 \leqslant i \leqslant n),$$

where $\varphi_i : (2^{\Sigma^* \times \Sigma^* \times \Sigma^*})^n \to 2^{\Sigma^* \times \Sigma^* \times \Sigma^*}$ are functions of $X_1, \ldots, X_n$, defined using the operations of concatenation, union, intersection, and any of the four context operators.

The right-hand sides of such a system can be represented as a vector function $\varphi = (\varphi_1, \ldots, \varphi_n)$, which has the following properties:

Lemma 1. *For every grammar with contexts, the vector function $\varphi = (\varphi_1, \ldots, \varphi_n)$ in the associated system of language equations is monotone, in the sense that for any two vectors K and L, the inequality $K \sqsubseteq L$ implies $\varphi(K) \sqsubseteq \varphi(L)$.*

Lemma 2. *For every grammar with contexts, the vector function $\varphi = (\varphi_1, \ldots, \varphi_n)$ in the associated system of language equations is continuous, in the sense that for every sequence of vectors of languages of triples $\{L^{(i)}\}_{i=1}^{\infty}$, it holds that*

$$\bigsqcup_{i=1}^{\infty} \varphi(L^{(i)}) = \varphi\Big(\bigsqcup_{i=1}^{\infty} L^{(i)}\Big).$$

The next result follows by the standard method of least fixed points.

Lemma 3. *If φ is monotone and continuous, then the system $X = \varphi(X)$ has a least solution, which is given by*

$$\bigsqcup_{k=0}^{\infty} \varphi^k(\bot).$$

Therefore, every system of equations corresponding to a grammar with contexts has a least solution, which shall be used to give an equivalent definition of the language generated by a grammar.

Definition 4. *Let $G = (\Sigma, N, R, S)$ be a grammar with two-sided contexts, let $X = \varphi(X)$ be the associated system of language equations, and let $(L_{A_1}, \ldots, L_{A_n})$ with $L_{A_i} \subseteq \Sigma^* \times \Sigma^* \times \Sigma^*$, where $N = \{A_1, \ldots, A_n\}$, be its least solution. Define the language generated by each nonterminal symbol $A \in N$ as the corresponding component of this solution: $L_G(A) = L_A$. Let $L(G) = \{\, w \mid \varepsilon\langle w\rangle\varepsilon \in L_S \,\}$.*

Definitions 2 and 4 are proved equivalent as follows. For every vector function $\varphi = (\varphi_1, \ldots, \varphi_n)$, denote the i-th component of the vector φ by $[\varphi]_i$.

Theorem 1. *Let $G = (\Sigma, N, R, S)$ be a grammar with two-sided contexts, and let $X = \varphi(X)$ be the associated system of language equations. Let $u\langle w\rangle v \in \Sigma^* \times \Sigma^* \times \Sigma^*$ be a string with contexts. Then, for every $A \in N$,*

$$u\langle w\rangle v \in \Big[\bigsqcup_{t \geqslant 0} \varphi^t(\varnothing, \ldots, \varnothing)\Big]_A \quad \text{if and only if} \quad \vdash_G A\big(u\langle w\rangle v\big).$$

4 Normal form

An ordinary context-free grammar can be transformed to the Chomsky normal form, with the rules restricted to $A \to BC$ and $A \to a$, with $B, C \in N$ and $a \in \Sigma$. This form has the following generalization to grammars with contexts.

Definition 5. *A grammar with two-sided contexts $G = (\Sigma, N, R, S)$ is said to be in the binary normal form, if each rule in R is of one of the forms*

$$A \to B_1C_1 \,\&\, \ldots \,\&\, B_kC_k \,\&\, \lhd D_1 \,\&\, \ldots \,\&\, \lhd D_m \,\&\, \trianglelefteq E_1 \,\&\, \ldots \,\&\, \trianglelefteq E_n \,\&\, \& \trianglerighteq F_1 \,\&\, \ldots \,\&\, \trianglerighteq F_{n'} \,\&\, \rhd H_1 \,\&\, \ldots \,\&\, \rhd H_{m'}, \tag{3a}$$

$$A \to a \,\&\, \lhd D_1 \,\&\, \ldots \,\&\, \lhd D_m \,\&\, \trianglelefteq E_1 \,\&\, \ldots \,\&\, \trianglelefteq E_n \,\&\, \& \trianglerighteq F_1 \,\&\, \ldots \,\&\, \trianglerighteq F_{n'} \,\&\, \rhd H_1 \,\&\, \ldots \,\&\, \rhd H_{m'}, \tag{3b}$$

where $k \geqslant 1$, $m, n, n', m' \geqslant 0$, $B_i, C_i, D_i, E_i, F_i, H_i \in N$, $a \in \Sigma$.

The transformation to the normal form consists of three stages: first, removing all direct *empty conjuncts* ε; secondly, eliminating *empty contexts* ($\lhd\varepsilon$, $\rhd\varepsilon$); finally, getting rid of *unit conjuncts* of the form B, with $B \in N$.

4.1 Null conjuncts

The first step is the removal of all rules $A \to \varepsilon \,\&\, \ldots$, so that no symbols generate ε, while all non-empty strings are generated as before. As generation of longer strings may depend on the generation of ε, already for ordinary context-free grammars, such a transformation requires adding extra rules that simulate the same dependence without actually generating any empty strings.

Example 4. Consider the following context-free grammar, which defines the language $\{abc, ab, ac, a, bcd, bd, cd, d\}$.

$$\begin{aligned} S &\to aA \mid Ad \\ A &\to BC \\ B &\to \varepsilon \mid b \\ C &\to \varepsilon \mid c \end{aligned}$$

Since B generates the empty string, the rule $A \to BC$ can be used to generate just C; therefore, once the rule $B \to \varepsilon$ is removed, one should add a new rule $A \to C$, in which B is omitted. Similarly one can remove the rule $C \to \varepsilon$ and add a "compensatory" rule $A \to B$. Since both B and C generate ε, so does A, by the rule $A \to BC$. Hence, extra rules $S \to a$ and $S \to d$, where A is omitted, have to be added.

An algorithm for carrying out such a transformation first calculates the set of nonterminals that generate ε, known as $\textsc{Nullable}(G) \subseteq N$, and then uses it to reconstruct the rules of the grammar. For the grammar in Example 4, $\textsc{Nullable}(G) = \{A, B, C\}$.

The same idea works for conjunctive grammars as well [8].

For grammars with contexts, this kind of transformation gets more complicated, because the generation of ε in the original grammar may depend on the contexts, and the same logical dependence must be ensured in the new grammar [2].

Example 5. Consider the following grammar with left contexts, which defines the language $L = \{abc, ab, ac, a, bcd, bd\}$:

$$\begin{array}{lcl} S & \to & aA \mid Ad \\ A & \to & BC \\ B & \to & \varepsilon \,\&\, {\lhd}D \mid b \\ C & \to & \varepsilon \mid c \\ D & \to & a \end{array}$$

The context specification ${\lhd}D$ in the rule $B \to \varepsilon \,\&\, {\lhd}D$ limits the generation of the empty string to left contexts of the form D, that is, to the left context a only. Then, in order to omit B in the rule $A \to BC$, one should add an extra rule $A \to C \,\&\, {\lhdeq}D$, in which the context operator ensures that B generates ε in this context. Since C generates ε in all contexts, a rule $A \to B$ is added as before. Furthermore, A generates the empty string only in the left context D, and hence A has to be omitted in the rules for S, as long as the context is of that form. This is done by two extra rules $A \to a \,\&\, {\lhdeq}D$ and $S \to d \,\&\, {\lhd}D$.

In this paper, the known method illustrated in Example 5 is further extended to grammars with two-sided contexts, where one has to consider both left and right contexts, in which a given nonterminal generates ε.

Example 6. Consider the following grammar with two-sided contexts, defining the language $L = \{abc, ac, bcd, bd, d\}$:

$$\begin{array}{lcl} S & \to & aA \mid Ad \\ A & \to & BC \\ B & \to & \varepsilon \,\&\, {\lhd}D \mid b \\ C & \to & \varepsilon \,\&\, {\rhd}E \mid c \\ D & \to & a \mid \varepsilon \,\&\, {\rhd}F \\ E & \to & d \\ F & \to & d \end{array}$$

In this grammar, the nonterminal B generates the empty string only in a left context of the form defined by D, while C defines ε only in a right context of the form E. Because of this, both B and C can be omitted in the rule $A \to BC$, giving two rules $A \to C \,\&\, {\lhd}D$ and $A \to B \,\&\, {\rhd}E$. Each of these rules ensures that B (or C) defines ε in this context. In those contexts where *both* B and C generate ε, so can A, by the rule $A \to BC$. Hence, in the rules for S, nonterminal A can be accordingly omitted by having rules $S \to a \,\&\, {\lhdeq}D \,\&\, {\rhd}E$ and $S \to d \,\&\, {\lhd}D \,\&\, {\rhdeq}E$.

By the combination of the rules $B \to \varepsilon \,\&\, {\lhd}D$ and $D \to \varepsilon \,\&\, {\rhd}F$, the grammar allows B to generate ε, but only in the left context ε and in any right context defined by F. When B is omitted in the rule $A \to BC$, this condition is simulated by a rule $A \to C \,\&\, {\lhd}\varepsilon \,\&\, {\rhdeq}F$. In a similar way, when omitting A in the rule $S \to Ad$, one has to have an extra rule $S \to d \,\&\, {\lhd}\varepsilon \,\&\, {\rhdeq}E \,\&\, {\rhdeq}F$, which contains the information about the right contexts, in which A defines the empty string in the empty left context.

After all null conjuncts have been eliminated from the grammar, its rules are as follows:

$$\begin{aligned}
S &\to aA \mid a \,\&\, {\lhdeq}D \,\&\, {\rhd}E \mid Ad \mid d \,\&\, {\lhd}D \,\&\, {\rhdeq}E \mid d \,\&\, {\lhd}\varepsilon \,\&\, {\rhdeq}E \,\&\, {\rhdeq}F \\
A &\to BC \mid B \,\&\, {\rhd}E \mid C \,\&\, {\lhd}D \mid C \,\&\, {\lhd}\varepsilon \,\&\, {\rhdeq}F \\
B &\to b \\
C &\to c \\
D &\to a \\
E &\to d \\
F &\to d
\end{aligned}$$

It order to define such a transformation for any given grammar with two-sided contexts, it is convenient to assume that in each rule of the grammar, the context operators are applied only to single nonterminal symbols rather than concatenations, that is, every rule is of the form

$$\begin{aligned}
A \to \alpha_1 \,\&\, \ldots \,\&\, \alpha_k \,\&\, {\lhd}D_1 \,\&\, \ldots \,\&\, {\lhd}D_m \,\&\, {\lhdeq}E_1 \,\&\, \ldots \,\&\, {\lhdeq}E_n \,\&\, \\
\&\, {\rhdeq}F_1 \,\&\, \ldots \,\&\, {\rhdeq}F_{m'} \,\&\, {\rhd}H_1 \,\&\, \ldots \,\&\, {\rhd}H_{n'},
\end{aligned} \tag{4}$$

with $A \in N$, $k \geqslant 1$, $m, n, m', n' \geqslant 0$, $\alpha_i \in (\Sigma \cup N)^*$ and $D_i, E_i, F_i, H_i \in N$.

Definition 6. *Let $G = (\Sigma, N, R, S)$ be a grammar with two-sided contexts with all rules of the form (4). For any set $\mathcal{S} \subseteq 2^N \times N \times 2^N$, denote by $\mathcal{S}^*$ the set of all pairs $(U_1 \cup \ldots \cup U_\ell,\ A_1 \ldots A_\ell,\ V_1 \cup \ldots \cup V_\ell)$ with $\ell \geqslant 0$ and $(U_i, A_i, V_i) \in \mathcal{S}$.*

Construct the sequence of sets $\textsc{Nullable}_i(G) \subseteq 2^N \times N \times 2^N$, for $i \geqslant 0$, as follows. Let $\textsc{Nullable}_0(G) = \varnothing$. Every next set $\textsc{Nullable}_{i+1}(G)$ contains all triples

$$\Big(\{D_1, \ldots, D_m\} \cup \{E_1, \ldots, E_n\} \cup \bigcup_i U_i,\ A,\ \{F_1, \ldots, F_{m'}\} \cup \{H_1, \ldots, H_{n'}\} \cup \bigcup_i V_i\Big)$$

for which there is exist a rule (4) and triples $(U_1, \alpha_1, V_1), \ldots, (U_k, \alpha_k, V_k) \in \textsc{Nullable}_i^*(G)$.

Finally, let $\textsc{Nullable}(G) = \bigcup_{i \geqslant 0} \textsc{Nullable}_i(G)$.

In the definition of $\mathcal{S}^*$, note that $\varnothing^* = \{(\varnothing, \varepsilon, \varnothing)\}$. This value of $\textsc{Nullable}_0^*(G)$ is used in the construction of $\textsc{Nullable}_1(G)$.

An element (U, A, V) of the set Nullable represents the intuitive idea that the nonterminal A generates ε in a left context of the form described by each nonterminal in the set U, and in a right context of the form given by nonterminals in V.

Example 7. For the grammar in Example 6, the set $\textsc{Nullable}(G)$ can be constructed according to this definition as follows:

$$\begin{aligned}
\textsc{Nullable}_0(G) &= \varnothing, \\
\textsc{Nullable}_1(G) &= \big\{(\{D\}, B, \varnothing), (\varnothing, C, \{E\}), (\varnothing, D, \{F\})\big\}, \\
\textsc{Nullable}_2(G) &= \big\{(\{D\}, B, \varnothing), (\varnothing, C, \{E\}), (\varnothing, D, \{F\}), (\{D\}, A, \{E\})\big\}
\end{aligned}$$

and $\textsc{Nullable}(G) = \textsc{Nullable}_2(G)$.

The elements $(\{D\}, B, \varnothing)$ and $(\varnothing, C, \{E\})$ are obtained directly from the rules of the grammar, and the element $(\{D\}, A, \{E\})$ represents the "concatenation" BC in the rule for A.

The set $\textsc{Nullable}(G)$ represents the generation of ε by different nonterminals in different contexts as follows.

Lemma 4. *Let* $G = (\Sigma, N, P, S)$ *be a grammar with contexts, let* $A \in N$ *and* $u, v \in \Sigma^*$. *Then,* $u\langle\varepsilon\rangle v \in L_G(A)$ *if and only if there exists such a triple* $(\{J_1, \ldots, J_\ell\}, A, \{K_1, \ldots, K_t\})$ *in* $\textsc{Nullable}(G)$, *with* $J_1, \ldots, J_\ell, K_1, \ldots, K_t \in N$, *for which* $\varepsilon\langle u\rangle v \in L_G(J_i)$ *for all* i *and* $u\langle v\rangle\varepsilon \in L_G(K_j)$ *for all* j.

Proof. $\Leftarrow$ The proof is an induction on p, the number of steps in the deduction of an item $A\big(u\langle\varepsilon\rangle v\big)$.

Basis. Let $p = 1$. Then, the item $A\big(u\langle\varepsilon\rangle v\big)$ is deduced by the rule $A \to \varepsilon$, and since such a rule exists, the triple $(\varnothing, A, \varnothing)$ is in $\textsc{Nullable}_1(G)$.

Induction step. Assume that $A\big(u\langle\varepsilon\rangle v\big)$ is deduced in $p \geqslant 2$ steps, and consider the rule (4) used at the last step of the deduction. For each direct conjunct α_i in this rule, let $\alpha_i = X_{i,1} \ldots X_{i,\ell}$ with $X_{i,j} \in \Sigma \cup N$. Then the last step of the deduction is

$$X_{i,j}\big(u\langle\varepsilon\rangle v\big), D_i\big(\varepsilon\langle u\rangle v\big), E_i\big(\varepsilon\langle u\rangle v\big), F_i\big(u\langle v\rangle\varepsilon\big), H_i\big(u\langle v\rangle\varepsilon\big) \vdash_G A\big(u\langle\varepsilon\rangle v\big).$$

By the induction hypothesis, for each symbol $X_{i,j}$, there exists a triple $(U_{i,j}, X_{i,j}, V_{i,j}) \in \textsc{Nullable}(G)$, in which $U_{i,j}, V_{i,j} \subseteq N$, and $\varepsilon\langle u\rangle v \in L_G(J)$

for all $J \in U_{i,j}$, and $u\langle v\rangle\varepsilon \in L_G(K)$ for all $K \in V_{i,j}$. Then, for each i, the pair (U_i, α_i, V_i) is in the set $\textsc{Nullable}^*(G)$, where $U_i = U_{i,1} \cup \ldots \cup U_{i,\ell}$, $V_i = V_{i,1} \cup \ldots \cup V_{i,\ell}$. Denote $U = \bigcup_i U_i \cup \{D_1, \ldots, D_m\} \cup \{F_1, \ldots, F_n\}$ and $V = \bigcup_i V_i \cup \{F_1, \ldots, F_{m'}\} \cup \{H_1, \ldots, H_{n'}\}$. Then the triple (U, A, V) is in $\textsc{Nullable}(G)$.

$\Leftarrow$ Consider nonterminals $J_1, \ldots, J_\ell$ and $K_1, \ldots, K_t \in N$, such that $(\{J_1, \ldots, J_\ell\}, A, \{K_1, \ldots, K_t\}) \in \textsc{Nullable}(G)$. Then, by Definition 6, $(\{J_1, \ldots, J_\ell\}, A, \{K_1, \ldots, K_t\}) \in \textsc{Nullable}_h(G)$ for some $h \geqslant 0$. The proof goes by induction on h.

Basis. Let $h = 0$. In this case, $\textsc{Nullable}(G)$ is an empty set and no nonterminals $A \in N$ satisfy the assumptions of the lemma.

Induction step. Let now $(\{J_1, \ldots, J_\ell\}, A, \{K_1, \ldots, K_t\}) \in \textsc{Nullable}_h(G)$ and $\varepsilon\langle u\rangle v \in L_G(J_i)$, $u\langle v\rangle\varepsilon \in L_G(K_j)$, for all $i \in \{1, \ldots, \ell\}$, $j \in \{1, \ldots, t\}$. Then, by Definition 6, the grammar has a rule of the form (4), such that for every direct conjunct α_i, the element $(U_i, \alpha_i, V_i) \in \textsc{Nullable}^*_{h-1}(G)$, and

$$\begin{array}{rcl} U_1 \cup \ldots \cup U_k \cup \{D_1, \ldots, D_m\} \cup \{E_1, \ldots, E_n\} & = & \{J_1, \ldots, J_\ell\}, \\ V_1 \cup \ldots \cup V_k \cup \{F_1, \ldots, F_{m'}\} \cup \{H_1, \ldots, H_{n'}\} & = & \{K_1, \ldots, K_t\}. \end{array}$$

For each conjunct α_i, let $\alpha_i = X_{i,1} \ldots X_{i,p}$ with $p \geqslant 0$ and $X_{i,1}, \ldots, X_{i,p} \in \Sigma \cup N$. Then, by the definition of a "star" of $\textsc{Nullable}_{h-1}(G)$, there exist sets $U_{i,1}, \ldots, U_{i,p} \subseteq N$, $V_{i,1}, \ldots, V_{i,p} \subseteq N$, such that $U_{i,1} \cup \ldots \cup U_{i,p} = U_i$, $V_{i,1} \cup \ldots \cup V_{i,p} = V_i$ and $(U_{i,j}, X_{i,j}, V_{i,j}) \in \textsc{Nullable}_{h-1}(G)$, for all j. By the induction hypothesis, applied to every symbol $X_{i,j}$, one gets that $u\langle\varepsilon\rangle v \in L_G(X_{i,j})$, for each j.

Repeating the same procedure for every element (U_i, α_i, V_i) of the set $\textsc{Nullable}^*_{h-1}(G)$, gives that $\vdash_G X_{i,1}\big(u\langle\varepsilon\rangle v\big), \ldots, X_{i,p}\big(u\langle\varepsilon\rangle v\big)$.

Now the item $A\big(u\langle\varepsilon\rangle v\big)$ can be deduced in the grammar G using the rule (4) as follows:

$$\begin{array}{l} X_{1,1}\big(u\langle\varepsilon\rangle v\big), \ldots, X_{k,p}\big(u\langle\varepsilon\rangle v\big), D_1\big(\varepsilon\langle u\rangle v\big), \ldots, D_m\big(\varepsilon\langle u\rangle v\big), \\ \qquad E_1\big(\varepsilon\langle u\rangle v\big), \ldots, E_n\big(\varepsilon\langle u\rangle v\big), F_1\big(u\langle v\rangle\varepsilon\big), \ldots, F_{m'}\big(u\langle v\rangle\varepsilon\big), \\ \qquad\qquad H_1\big(u\langle v\rangle\varepsilon\big), \ldots, H_{n'}\big(u\langle v\rangle\varepsilon\big) \vdash_G A\big(u\langle\varepsilon\rangle v\big). \end{array}$$

□

Consider the second case described in Example 6, where a nonterminal (B) defines the empty string in the empty left context ($\lhd D$), and there are restrictions on the right contexts ($\rhd F$), in which it does so. This case has to be treated in a special way, using a variant of the set $\textsc{Nullable}(G)$ that assumes empty left contexts.

Definition 7. *Let $G = (\Sigma, N, R, S)$ be a grammar with two-sided contexts. Construct a sequence of sets $\lhd\varepsilon\text{-}\textsc{Nullable}_i(G) \subseteq N \times 2^N$, with $i \geqslant 0$, by setting $\lhd\varepsilon\text{-}\textsc{Nullable}_0(G) = \{\, (A, V) \mid (\varnothing, A, V) \in \textsc{Nullable}(G) \,\}$, and $\lhd\varepsilon\text{-}\textsc{Nullable}_{i+1}(G) = \big\{(A, V \cup V_1 \cup \ldots \cup V_\ell) \;\big|\; (\{J_1, \ldots, J_\ell\}, A, V) \in \textsc{Nullable}(G),\ \exists\, V_1, \ldots, V_\ell \subseteq N :\ (J_i, V_i) \in \lhd\varepsilon\text{-}\textsc{Nullable}_i(G)\big\}$. Define $\lhd\varepsilon\text{-}\textsc{Nullable}(G) = \bigcup_{i \geqslant 0} \lhd\varepsilon\text{-}\textsc{Nullable}_i(G)$.*

Example 8. For the grammar in Example 6,

$$\begin{aligned} \lhd\varepsilon\text{-}\textsc{Nullable}_0(G) &= \big\{(C, \{E\}), (D, \{F\})\big\}, \\ \lhd\varepsilon\text{-}\textsc{Nullable}_1(G) &= \big\{(C, \{E\}), (D, \{F\}), (B, \{F\}), (A, \{E, F\})\big\}, \end{aligned}$$

and $\lhd\varepsilon\text{-}\textsc{Nullable}(G) = \lhd\varepsilon\text{-}\textsc{Nullable}_1(G)$.

The element $(B, \{F\})$ means that every time B defines ε in the left context ε, its right context is of the form F. For $(A, \{E, F\})$, the right contexts come from the elements $(B, \{F\})$ and $(C, \{E\})$, corresponding to a concatenation BC in the rule for A.

Lemma 5. *Let $G = (\Sigma, N, R, S)$ be a grammar with contexts, let $A \in N$ and $v \in \Sigma^*$. Then $\varepsilon\langle\varepsilon\rangle v \in L_G(A)$ if and only if there exist $K_1, \ldots, K_t$ such that $(A, \{K_1, \ldots, K_t\}) \in \lhd\varepsilon\text{-}\textsc{Nullable}(G)$, and for all $i \in \{1, \ldots, t\}$ the string $\varepsilon\langle v\rangle\varepsilon \in L_G(K_i)$.*

Proof. $\Leftarrow$ Let $K_1, \ldots, K_t \in N$ and $(A, \{K_1, \ldots, K_t\}) \in \lhd\varepsilon\text{-}\textsc{Nullable}(G)$. Then, by Definition 6, $(A, \{K_1, \ldots, K_t\}) \in \lhd\varepsilon\text{-}\textsc{Nullable}_h(G)$ for some $h \geqslant 0$. The proof is an induction on h.

Basis. Let $h = 0$. Then, by Definition 6, $\lhd\varepsilon\text{-}\textsc{Nullable}(G) = \{\, (A, \{K_1, \ldots, K_t\}) \mid (\varnothing, A, \{K_1, \ldots, K_t\}) \in \textsc{Nullable}(G) \,\}$. Since $\varepsilon\langle v\rangle\varepsilon \in L_G(K_i)$ by the assumption and $(\varnothing, A, \{K_1, \ldots, K_t\})$ is in $\textsc{Nullable}(G)$, applying Lemma 4 gives that $u\langle\varepsilon\rangle v \in L_G(A)$ where $u \in \Sigma^*$. Thus, $\varepsilon\langle\varepsilon\rangle v \in L_G(A)$, as desired.

Induction step. Let $h > 0$. By Definition 7, the element $(A, V \cup V_1 \cup \ldots \cup V_\ell)$ is in $\lhd\varepsilon\text{-}\textsc{Nullable}_h(G)$ if there exist such nonterminals $J_1, \ldots, J_\ell \in N$ that $(\{J_1, \ldots, J_\ell\}, A, V \cup V_1 \cup \ldots \cup V_\ell)$ is in $\textsc{Nullable}(G)$ and every pair (J_i, V_i) is in $\lhd\varepsilon\text{-}\textsc{Nullable}_{h-1}(G)$.

Consider a pair $(J_i, V_i) \in \lhd\varepsilon\text{-}\textsc{Nullable}_{h-1}(G)$, for all $i \in \{1, \ldots, \ell\}$. By the induction hypothesis, $\varepsilon\langle\varepsilon\rangle v \in L_G(J_i)$. Then $\varepsilon\langle\varepsilon\rangle v \in L_G(A)$ by Lemma 4, as desired.

$\Rightarrow$ The proof goes by induction on p, the number of steps used in deduction of the item $A\big(\varepsilon\langle\varepsilon\rangle v\big)$.

Basis. Let $p = 2$ and consider the two rules $A \to \varepsilon \,\&\, \lhd D$, $D \to \varepsilon$. The item $A\big(\varepsilon\langle\varepsilon\rangle v\big)$ can be deduced as $D\big(\varepsilon\langle\varepsilon\rangle v\big) \vdash_G A\big(\varepsilon\langle\varepsilon\rangle v\big)$. By Definition 6, $(\{D\}, A, \varnothing), (\varnothing, D, \varnothing) \in \textsc{Nullable}(G)$. Hence, $(A, \varnothing) \in \lhd\varepsilon\text{-}\textsc{Nullable}(G)$ by Definition 7 and the statement of the lemma is thus satisfied.

Induction step. One has to prove that there exists a set $V \subseteq N$, such that $(A, V) \in \lhd\varepsilon\text{-}\textsc{Nullable}(G)$ and $\varepsilon\langle v\rangle\varepsilon$ belongs to the language of each element of V. Let the item $A\big(\varepsilon\langle\varepsilon\rangle v\big)$ be deduced in p steps and let the last step of its deduction use a rule of the form (4), with $\{F_1, \ldots, F_{m'}, H_1, \ldots, H_{n'}\} \subseteq V$.

By Definition 7, $(A, V) \in \lhd\varepsilon\text{-}\textsc{Nullable}(G)$ means that there exist nonterminals $J_1, \ldots, J_\ell \in N$ and sets $V_1, \ldots, V_\ell \subseteq N$, such that $(\{J_1, \ldots, J_\ell\}, A, V \cup V_1 \cup \ldots \cup V_\ell) \in \textsc{Nullable}(G)$ and $(J_i, V_i) \in \lhd\varepsilon\text{-}\textsc{Nullable}(G)$.

By the induction hypothesis, $\varepsilon\langle v\rangle\varepsilon$ is in $L_G(K)$, for all $K \in V_i$ and $i \in \{1, \ldots, \ell\}$. Thus, $\varepsilon\langle v\rangle\varepsilon \in L_G(K)$ (for all $K \in V \cup V_1 \cup \ldots \cup V_\ell$). $\square$

Similarly to the set $\lhd\varepsilon\text{-}\textsc{Nullable}(G)$, define the set $\rhd\varepsilon\text{-}\textsc{Nullable}(G)$ as follows.

Definition 8. *Let $G = (\Sigma, N, R, S)$ be a grammar with two-sided contexts. Define symmetrically to Definition 7 the set $\rhd\varepsilon\text{-}\textsc{Nullable}(G) \subseteq 2^N \times N$, by setting $\rhd\varepsilon\text{-}\textsc{Nullable}(G) = \bigcup_{i\geqslant 0} \rhd\varepsilon\text{-}\textsc{Nullable}_i(G)$, with $\rhd\varepsilon\text{-}\textsc{Nullable}_0(G) = \{\, (U, A) \mid (U, A, \varnothing) \in \textsc{Nullable}(G) \,\}$, and $\rhd\varepsilon\text{-}\textsc{Nullable}_{i+1}(G) = \big\{(U \cup U_1 \cup \ldots \cup U_\ell, A) \;\big|\; (U, A, \{K_1, \ldots, K_t\}) \in \textsc{Nullable}(G),\ \exists\, U_1, \ldots, U_t \subseteq N :\ (U_i, K_i) \in \rhd\varepsilon\text{-}\textsc{Nullable}_i(G)\big\}$.*

Again, the following characterization for this set is established.

Lemma 6. *Let $G = (\Sigma, N, R, S)$ be a grammar with contexts, let $A \in N$ and $u \in \Sigma^*$. Then $u\langle\varepsilon\rangle\varepsilon \in L_G(A)$ if and only if there exist $J_1, \ldots, J_\ell$ such that $(\{J_1, \ldots, J_\ell\}, A) \in \rhd\varepsilon\text{-}\textsc{Nullable}(G)$, and for all $i \in \{1, \ldots, t\}$ the string $\varepsilon\langle u\rangle\varepsilon \in L_G(J_i)$.*

Now a general transformation to the normal form can be given.

Construction 1. Let $G = (\Sigma, N, R, S)$ be a grammar with two-sided contexts, with all rules of the form

$$A \to a \tag{5a}$$

$$A \to B_1 \,\&\, \ldots \,\&\, B_k \,\&\, \lhd D_1 \,\&\, \ldots \,\&\, \lhd D_m \,\&\, \trianglelefteqslant E_1 \,\&\, \ldots \,\&\, \trianglelefteqslant E_n \,\&\, \& \trianglerighteqslant F_1 \,\&\, \ldots \,\&\, \trianglerighteqslant F_{m'} \,\&\, \rhd H_1 \,\&\, \ldots \,\&\, \rhd H_{n'} \tag{5b}$$

$$A \to BC \tag{5c}$$

$$A \to \varepsilon, \tag{5d}$$

where $a \in \Sigma$ and $A, B, C, D_i, E_i, F_i, H_i \in N$. Construct another grammar with two-sided contexts $G' = (\Sigma, N, R', S)$, with the following rules.

1. All rules of the form (5a) in R are added to R':

$$A \to a \in R. \tag{6}$$

2. For every "long" rule of the form (5b) in R, the set R' shall contain this rule:

$$\begin{aligned} A \to B_1 \,\&\, \ldots \,\&\, B_k \,\&\, \lhd D_1 \,\&\, \ldots \,\&\, \lhd D_m \,\&\, \trianglelefteq E_1 \,\&\, \ldots \,\&\, \trianglelefteq E_n \,\&\, \\ \&\, \trianglerighteq F_1 \,\&\, \ldots \,\&\, \trianglerighteq F_{m'} \,\&\, \rhd H_1 \,\&\, \ldots \,\&\, \rhd H_{n'}, \end{aligned} \tag{7}$$

as well as the following additional ones.

If $m \geqslant 1$, then, for any collection of m pairs (D_1, V_1), …, $(D_m, V_m) \in \lhd\varepsilon$-Nullable$(G)$, let $\bigcup_{i=1}^{m} V_i = \{K_1, \ldots, K_t\}$ and add the rule

$$\begin{aligned} A \to B_1 \,\&\, \ldots \,\&\, B_k \,\&\, E_1 \,\&\, \ldots \,\&\, E_n \,\&\, \lhd\varepsilon \,\&\, \trianglerighteq K_1 \,\&\, \ldots \,\&\, \trianglerighteq K_t \,\&\, \\ \&\, \trianglerighteq F_1 \,\&\, \ldots \,\&\, \trianglerighteq F_{m'} \,\&\, \rhd H_1 \,\&\, \ldots \,\&\, \rhd H_{n'}. \end{aligned} \tag{8a}$$

Symmetrically, if (U_1, H_1), …, $(U_{n'}, H_{n'}) \in \rhd\varepsilon$-Nullable$(G)$ (with $n' \geqslant 1$) and $\bigcup_{i=1}^{n'} U_i = \{K_1, \ldots, K_t\}$, then there is a rule

$$\begin{aligned} A \to B_1 \,\&\, \ldots \,\&\, B_k \,\&\, F_1 \,\&\, \ldots \,\&\, F_{m'} \,\&\, \lhd D_1 \,\&\, \ldots \,\&\, \lhd D_m \,\&\, \\ \&\, \trianglelefteq E_1 \,\&\, \ldots \,\&\, \trianglelefteq E_n \,\&\, \trianglelefteq K_1 \,\&\, \ldots \,\&\, \trianglelefteq K_t \,\&\, \rhd\varepsilon. \end{aligned} \tag{8b}$$

Finally, if (D_1, V_1), …, $(D_m, V_m) \in \lhd\varepsilon$-Nullable$(G)$, (U_1, H_1), …, $(U_{n'}, H_{n'}) \in \rhd\varepsilon$-Nullable$(G)$ (with m, $n' \geqslant 1$) and $\bigcup_{i=1}^{m} V_i \cup \bigcup_{j=1}^{n'} U_j = \{K_1, \ldots, K_t\}$, then the set R' contains a rule

$$\begin{aligned} A \to B_1 \,\&\, \ldots \,\&\, B_k \,\&\, E_1 \,\&\, \ldots \,\&\, E_n \,\&\, F_1 \,\&\, \ldots \,\&\, F_{m'} \,\&\, \\ \&\, K_1 \,\&\, \ldots \,\&\, K_t \,\&\, \lhd\varepsilon \,\&\, \rhd\varepsilon. \end{aligned} \tag{8c}$$

3. Every rule of the form (5c) in R is added to R':

$$A \to BC, \tag{9}$$

along with the following extra rules:

$$A \to B \,\&\, \trianglelefteq J_1 \,\&\, \ldots \,\&\, \trianglelefteq J_\ell \,\&\, \rhd K_1 \,\&\, \ldots \,\&\, \rhd K_t, \tag{10a}$$

for all $(\{J_1, \ldots, J_\ell\}, C, \{K_1, \ldots, K_t\}) \in$ Nullable(G);

$$A \to B \,\&\, \trianglelefteq J_1 \,\&\, \ldots \,\&\, \trianglelefteq J_\ell \,\&\, \rhd\varepsilon, \tag{10b}$$

for all $(\{J_1, \ldots, J_\ell\}, C) \in \rhd\varepsilon$-Nullable$(G)$ with $\ell \geqslant 1$;

$$A \to C \,\&\, \lhd J_1 \,\&\, \ldots \,\&\, \lhd J_\ell \,\&\, \trianglerighteq K_1 \,\&\, \ldots \,\&\, \trianglerighteq K_t, \tag{10c}$$

for all $(\{J_1, \ldots, J_\ell\}, B, \{K_1, \ldots, K_t\}) \in$ Nullable(G);

$$A \to C \,\&\, \lhd\varepsilon \,\&\, \trianglerighteq K_1 \,\&\, \ldots \,\&\, \trianglerighteq K_t, \tag{10d}$$

for all $(B, \{K_1, \ldots, K_t\}) \in \lhd\varepsilon$-Nullable$(G)$ with $t \geqslant 1$.

Lemma 7. *Let $G = (\Sigma, N, R, S)$ be a grammar with two-sided contexts. Then the grammar $G' = (\Sigma, N', R', S)$ obtained by Construction 1 generates the language $L(G') = L(G) \setminus \{\varepsilon\}$.*

Proof. $\subseteq$ It is claimed that whenever an item $A\big(u\langle w\rangle v\big)$ (with $A \in N$, $u, v \in \Sigma^*$ and $w \in \Sigma^+$) can be deduced in the grammar G, it can also be deduced in the new grammar G'. The proof is by induction on p, the number of steps in the deduction of $A\big(u\langle w\rangle v\big)$ in G.

Basis. Let $p = 1$. Consider an item $A\big(u\langle w\rangle v\big)$ deduced in G by a rule of the form $A \to a$. Then $w = a \in \Sigma$ and the last step of the deduction takes form $a\big(u\langle a\rangle v\big) \vdash_G A\big(u\langle a\rangle v\big)$. The same rule $A \to a$ is also contained in the new grammar G', and one can deduce the item $A\big(u\langle a\rangle v\big)$ in G' by this rule: $a\big(u\langle a\rangle v\big) \vdash_{G'} A\big(u\langle a\rangle v\big)$.

Induction step. Let the item $A\big(u\langle w\rangle v\big)$ be deduced in G by some rule of the form (5b) or (5c).

If this is a rule $A \to BC$, then the last step of the deduction is $B\big(u\langle w_1\rangle w_2 v\big)$, $C\big(uw_1\langle w_2\rangle v\big) \vdash_G A\big(u\langle w_1 w_2\rangle v\big)$, for some partition $w = w_1 w_2$. Each of the premises is deduced in fewer than p steps. Though the string w is non-empty, one of w_1, w_2 may be empty, and the proof splits into three cases, depending on whether any of these strings is empty, and if so, then which of them.

- If both w_1 and w_2 are non-empty, then both items $B\big(u\langle w_1\rangle w_2 v\big)$ and $C\big(uw_1\langle w_2\rangle v\big)$ can be deduced in the grammar G' by the induction hypothesis. Since G' has the same rule $A \to BC$, it can be used to deduce the item $A\big(u\langle w\rangle v\big)$ in G' in the same way: $B\big(u\langle w_1\rangle w_2 v\big)$, $C\big(uw_1\langle w_2\rangle v\big)$ $\vdash_{G'} A\big(u\langle w_1 w_2\rangle v\big)$.

- Let $w_1 = w$ and $w_2 = \varepsilon$. Then the last step of the deduction is $B\big(u\langle w\rangle v\big)$, $C\big(uw\langle\varepsilon\rangle v\big) \vdash_G A\big(u\langle w\rangle v\big)$. By the induction hypothesis, the item $B\big(u\langle w\rangle v\big)$ can be deduced in G'. Though the other item $C\big(uw\langle\varepsilon\rangle v\big)$ can be deduced only in G, it is reflected by other items in G'. The proof splits into two cases, depending on whether v is empty or not.

  First consider the case of $v \neq \varepsilon$. Then, since the item $C\big(uw\langle\varepsilon\rangle v\big)$ can be deduced in G, by Lemma 4, there exist such nonterminals $J_1, \ldots, J_\ell$, $K_1, \ldots, K_t \in N$, that $(\{J_1, \ldots, J_\ell\}, C, \{K_1, \ldots, K_t\}) \in \textsc{Nullable}(G)$ and all items $J_i\big(\varepsilon\langle uw\rangle v\big)$ and $K_j\big(uw\langle v\rangle\varepsilon\big)$, for all applicable i and j, can be deduced in the grammar G. By the induction hypothesis, each of these items can be deduced in the new grammar G'. From these premises, using the rule (10a), the desired item $A\big(u\langle w\rangle v\big)$ can be deduced as follows: $B\big(u\langle w\rangle v\big)$, $J_1\big(\varepsilon\langle uw\rangle v\big)$, …, $J_\ell\big(\varepsilon\langle uw\rangle v\big)$, $K_1\big(uw\langle v\rangle\varepsilon\big)$, …, $K_t\big(uw\langle v\rangle\varepsilon\big) \vdash_{G'} A\big(u\langle w\rangle v\big)$.

Let now $v = \varepsilon$. In this case, the last step of the deduction of $A\big(u\langle w\rangle\varepsilon\big)$ is as follows: $B\big(u\langle w\rangle\varepsilon\big)$, $C\big(uw\langle\varepsilon\rangle\varepsilon\big) \vdash_G A\big(u\langle w\rangle\varepsilon\big)$. By the induction hypothesis, the item $B\big(u\langle w\rangle\varepsilon\big)$ can be deduced in the new grammar G'. Since $uw\langle\varepsilon\rangle\varepsilon \in L_G(C)$, by Lemma 6, there exist $J_1, \ldots, J_\ell \in N$, such that $(\{J_1, \ldots, J_\ell\}, C) \in \rhd\varepsilon\text{-}\textsc{Nullable}(G)$, and the items $J_1\big(\varepsilon\langle uw\rangle\varepsilon\big)$, $\ldots$, $J_\ell\big(\varepsilon\langle uw\rangle\varepsilon\big)$ can be deduced in the grammar G. By the induction hypothesis, all these items can also be deduced in the grammar G. From these premises, using the rule (10b), one can deduce the desired item in G' as follows: $B\big(u\langle w\rangle\varepsilon\big)$, $J_1\big(\varepsilon\langle uw\rangle\varepsilon\big)$, $\ldots$, $J_\ell\big(\varepsilon\langle uw\rangle\varepsilon\big) \vdash_{G'} A\big(u\langle w\rangle\varepsilon\big)$.

- The case of $w_1 = \varepsilon$ and $w_2 = w$, where the last step of deduction of $A\big(u\langle w\rangle v\big)$ in G is $B\big(u\langle\varepsilon\rangle wv\big)$, $C\big(u\langle w\rangle v\big) \vdash_G A\big(u\langle w\rangle v\big)$, is handled symmetrically to the above case.

  If $u \neq \varepsilon$, then Lemma 4 gives a triple $(\{J_1, \ldots, J_\ell\}, B, \{K_1, \ldots, K_t\}) \in \textsc{Nullable}(G)$, for which $\vdash_{G'} C\big(u\langle w\rangle v\big)$ and $\vdash_{G'} J_1\big(\varepsilon\langle u\rangle wv\big)$, $\ldots$, $J_\ell\big(\varepsilon\langle u\rangle wv\big)$, $K_1\big(u\langle wv\rangle\varepsilon\big)$, $\ldots$, $K_t\big(u\langle wv\rangle\varepsilon\big)$. From these premises, one can then deduce $A\big(u\langle w\rangle v\big)$ in G' by the rule (10c).

  If $u = \varepsilon$, then the last step of the deduction of $A\big(u\langle w\rangle\varepsilon\big)$ took the form $B\big(\varepsilon\langle\varepsilon\rangle wv\big)$, $C\big(\varepsilon\langle w\rangle v\big) \vdash_G A\big(\varepsilon\langle w\rangle v\big)$. Then, by Lemma 5, there is a pair $(B, \{K_1, \ldots, K_t\}) \in \lhd\varepsilon\text{-}\textsc{Nullable}(G)$, for which all items $K_1\big(\varepsilon\langle wv\rangle\varepsilon\big)$, $\ldots$, $K_t\big(\varepsilon\langle wv\rangle\varepsilon\big)$ are deducible in G'. These items, together with the item $C\big(\varepsilon\langle w\rangle v\big)$ are the premises for the deduction of $A\big(\varepsilon\langle w\rangle v\big)$ by the rule (10d).

Consider now the other case, when the last step of the deduction of $A\big(u\langle w\rangle v\big)$ is by some rule (5b), and is accordingly of the form

$$B_1\big(u\langle w\rangle v\big), \ldots, B_k\big(u\langle w\rangle v\big), \tag{11a}$$
$$D_1\big(\varepsilon\langle u\rangle wv\big), \ldots, D_m\big(\varepsilon\langle u\rangle wv\big), \tag{11b}$$
$$E_1\big(\varepsilon\langle uw\rangle v\big), \ldots, E_n\big(\varepsilon\langle uw\rangle v\big), \tag{11c}$$
$$F_1\big(u\langle wv\rangle\varepsilon\big), \ldots, F_{m'}\big(u\langle wv\rangle\varepsilon\big), \tag{11d}$$
$$H_1\big(uw\langle v\rangle\varepsilon\big), \ldots, H_{n'}\big(uw\langle v\rangle\varepsilon\big) \tag{11e}$$
$$\vdash_G A\big(u\langle w\rangle v\big).$$

- If $u \neq \varepsilon$ and $v \neq \varepsilon$, then, by the induction hypothesis, each of the premises can be deduced in the grammar G', and thus $A\big(u\langle w\rangle v\big)$ can be deduced in G' by the same rule (5b), which is in R' by the construction.

- Let in the rule (5b) $m \geqslant 1$ and let $u = \varepsilon$ and $v \neq \varepsilon$. The items (11a), (11c)–(11e) can still be deduced in G' by the induction hypothesis.

  Since $\varepsilon\langle u\rangle wv \in L_G(D_i)$ for all $i \in \{1, \ldots, m\}$, then, by Lemma 5, there exist $V_i \subseteq N$ such that $(D_i, V_i) \in \lhd\varepsilon\text{-}\textsc{Nullable}(G)$, and

$$\vdash_G J\big(\varepsilon\langle v\rangle\varepsilon\big) \qquad \text{(for every } J \in V_i\text{)}. \tag{12}$$

By the induction hypothesis, such an item can be deduced in the grammar G', as well. Now the item $A\big(u\langle w\rangle v\big)$ can be deduced in G' out of the premises (11a), (12), and (11c)–(11e) by the rule (8a), which is added to R' by Construction 1.

- Let in the rule (5b) $n' \geqslant 1$ and let $u \neq \varepsilon$ and $v = \varepsilon$. The items (11a), (11b)–(11d) can be deduced in G' by the induction hypothesis.

  Since the items $H_i\big(uw\langle v\rangle\varepsilon\big)$ are deducible in G, by Construction 1, the grammar G' has a rule of the form (8b). Similarly to the previous case, by Lemma 6, the item

  $$K_i\big(\varepsilon\langle u\rangle\varepsilon\big), \tag{13}$$

  for all $i \in \{1, \ldots, t\}$, is deducible in G. By the induction hypothesis, one can deduce each of them in the grammar G', as well. The deduction of the desired item $A\big(u\langle w\rangle v\big)$ in G' out of the premises (11a), (11b)–(11d) and (13) can be done by the rule (8b), which is in R' by the construction.

- Let now $u = v = \varepsilon$. By the induction hypothesis, the items (11a), (11c) and (11d) are deducible in G'.

  Similarly to the previous cases, one can show that

  $$\vdash_G J_i\big(\varepsilon\langle u\rangle\varepsilon\big), \vdash_G K_i\big(\varepsilon\langle v\rangle\varepsilon\big). \tag{14}$$

  The deduction of $A\big(u\langle w\rangle v\big)$ can now be carried out by a rule (8c), which is in G' by the construction, using the premises (11a), (11c)–(11d), and (14).

☺ Conversely, it has to be proved that $\vdash_{G'} A\big(u\langle w\rangle v\big)$ implies that $\vdash_G A\big(u\langle w\rangle v\big)$ and $w \neq \varepsilon$. The proof is by induction on p, the number of steps in deduction of the item $A\big(u\langle w\rangle v\big)$ in G'.

Basis. Let $p = 1$, and let an item $A\big(u\langle w\rangle v\big)$ be deduced in G' by a rule $A \to a$. Then $w = a \in \Sigma$ and the deduction takes the form $a\big(u\langle a\rangle v\big) \vdash_G A\big(u\langle a\rangle w\big)$. The deduction of $A\big(u\langle w\rangle v\big)$ in the old grammar G is exactly the same and uses a rule $A \to a$, which is in R by virtue of Construction 1.

Induction step. Let an item $A\big(u\langle w\rangle v\big)$ be deduced in G', and the last step of this deduction use a rule of the form r'. Then the following cases are possible.

- Let the rule r' be of the form (7). That is, the last step of deduction of $A\big(u\langle w\rangle v\big)$ in G' uses the premises (11a)–(11e). By the induction hypothesis, each of these premises can also be deduced in the grammar G. The item $A\big(u\langle w\rangle v\big)$ can be deduced in G out of these premises using the same rule r', which is in R by the construction.

- Let the rule r' be of the form (8a). That is, $u = \varepsilon$ and the last step of deduction of $A\big(\varepsilon\langle w\rangle v\big)$ takes form $B_1\big(\varepsilon\langle w\rangle v\big), \ldots, B_k\big(\varepsilon\langle w\rangle v\big), E_1\big(\varepsilon\langle w\rangle v\big), \ldots, E_n\big(\varepsilon\langle w\rangle v\big), K_1\big(\varepsilon\langle wv\rangle\varepsilon\big), \ldots, K_t\big(\varepsilon\langle wv\rangle\varepsilon\big), F_1\big(\varepsilon\langle wv\rangle\varepsilon\big), \ldots, F_{m'}\big(\varepsilon\langle wv\rangle\varepsilon\big), H_1\big(w\langle v\rangle\varepsilon\big), \ldots, H_{n'}\big(w\langle v\rangle\varepsilon\big) \vdash_{G'} A\big(\varepsilon\langle w\rangle v\big)$.

  By the induction hypothesis, all of the premises can be deduced in the grammar G. Construction 1 only adds the rule (8a) to R', when R contains a rule (7) (with $m \geqslant 1$) and there exist $(D_1, V_1), \ldots, (D_m, V_m) \in \lhd\varepsilon\text{-}\textsc{Nullable}(G)$, such that $\bigcup_{i=1}^{m} V_i = \{K_1, \ldots, K_t\}$. Applying Lemma 5 to every pair (D_i, V_i), one can obtain that $\varepsilon\langle\varepsilon\rangle v \in L_G(D_i)$. Thus, the item $A\big(\varepsilon\langle w\rangle v\big)$ can be deduced in G out of the premises $B_i\big(\varepsilon\langle w\rangle v\big)$, $D_i\big(\varepsilon\langle\varepsilon\rangle v\big)$, $E_i\big(\varepsilon\langle w\rangle v\big)$, $F_i\big(\varepsilon\langle wv\rangle\varepsilon\big)$, and $H_i\big(w\langle v\rangle\varepsilon\big)$ by the rule (7).

- Let r' be of the form (8b). Symmetrically to the previous case, this rule requires $v = \varepsilon$, and the last step of deduction of $A\big(u\langle w\rangle\varepsilon\big)$ is $B_i\big(u\langle w\rangle\varepsilon\big)$, $F_i\big(u\langle w\rangle\varepsilon\big)$, $D_i\big(\varepsilon\langle u\rangle w\big)$, $E_i\big(\varepsilon\langle uw\rangle\varepsilon\big)$, $K_i\big(\varepsilon\langle uw\rangle\varepsilon\big) \vdash_{G'} A\big(u\langle w\rangle\varepsilon\big)$. All of the premises can be deduced in G by the induction hypothesis.

  The rule (8b) is only added to G', when G has a rule (7) (with $n' \geqslant 1$) and there exist $(U_1, H_1), \ldots, (U_{n'}, H_{n'}) \in \rhd\varepsilon\text{-}\textsc{Nullable}(G)$, such that $\bigcup_{i=1}^{m} U_i = \{K_1, \ldots, K_t\}$. Similarly to the previous case, one can obtain by Lemma 6, that $u\langle\varepsilon\rangle\varepsilon \in L_G(H_i)$ and deduce the item $A\big(u\langle w\rangle\varepsilon\big)$ out of the premises $B_i\big(u\langle w\rangle\varepsilon\big)$, $D_i\big(\varepsilon\langle u\rangle w\big)$, $H_i\big(u\langle\varepsilon\rangle\varepsilon\big)$, $F_i\big(u\langle w\rangle\varepsilon\big)$, $E_i\big(\varepsilon\langle uw\rangle\varepsilon\big)$ by the rule (7).

- Let r' be of the form (8c). In this case $u = v = \varepsilon$ and the last step of deduction of $A\big(\varepsilon\langle w\rangle\varepsilon\big)$ in G is $B_i\big(\varepsilon\langle w\rangle\varepsilon\big)$, $E_i\big(\varepsilon\langle w\rangle\varepsilon\big)$, $F_i\big(\varepsilon\langle w\rangle\varepsilon\big)$, $K_i\big(\varepsilon\langle w\rangle\varepsilon\big) \vdash_{G'} A\big(\varepsilon\langle w\rangle\varepsilon\big)$.

  Similarly to the two previous cases, one can conclude that the items $D_i\big(\varepsilon\langle\varepsilon\rangle v\big)$ and $H_i\big(u\langle\varepsilon\rangle\varepsilon\big)$ can be deduced in G.

  Finally, the deduction of the item $A\big(\varepsilon\langle w\rangle\varepsilon\big)$ in G can be carried out using the rule (7) (which is in R by the construction) as follows: $B_i\big(\varepsilon\langle w\rangle\varepsilon\big)$, $D_i\big(\varepsilon\langle\varepsilon\rangle v\big)$ $E_i\big(\varepsilon\langle w\rangle\varepsilon\big)$, $F_i\big(\varepsilon\langle w\rangle\varepsilon\big)$, $H_i\big(u\langle\varepsilon\rangle\varepsilon\big) \vdash_G A\big(\varepsilon\langle w\rangle\varepsilon\big)$.

- Let r' be of the form (9). In the grammar G', the last step of deduction of $A\big(u\langle w\rangle v\big)$ takes the form $B\big(u\langle w_1\rangle w_2 v\big), C\big(uw_1\langle w_2\rangle v\big) \vdash_{G'} A\big(u\langle w\rangle v\big)$, for some partition $w_1 w_2 = w$. By the induction hypothesis, both of the premises can be deduced in the grammar G. Then, the item $A\big(u\langle w\rangle v\big)$ can be deduced out these premises in the grammar G using a rule $A \to BC$, which is in G by the construction: $B\big(u\langle w_1\rangle w_2 v\big)$, $C\big(uw_1\langle w_2\rangle v\big)$ $\vdash_G A\big(u\langle w\rangle v\big)$.

- Let r' be of the form (10a). Then $(\{J_1, \ldots, J_\ell\}, C, \{K_1, \ldots, K_t\}) \in \textsc{Nullable}(G)$, and the item $A\big(u\langle w\rangle v\big)$ is deduced in G' out of the

premises $B\big(u\langle w\rangle v\big)$, $J_1\big(\varepsilon\langle uw\rangle v\big)$, …, $J_\ell\big(\varepsilon\langle uw\rangle v\big)$, $K_1\big(u\langle wv\rangle\varepsilon\big)$, …, $K_t\big(u\langle wv\rangle\varepsilon\big)$. By the induction hypothesis, the item $B\big(u\langle w\rangle v\big)$ can be deduced in G. By Lemma 4, the item $C\big(uw\langle\varepsilon\rangle v\big)$ can be deduced in G, as well. Then, using these two premises, one can carry out the deduction of $A\big(u\langle w\rangle v\big)$ in the grammar G by the rule $A \to BC$: $B\big(u\langle w\rangle v\big)$, $C\big(uw\langle\varepsilon\rangle v\big) \vdash_G A\big(u\langle w\rangle v\big)$.

- Let r' be of the form (10b). In order to deduce $A\big(u\langle w\rangle v\big)$ by this rule, v must be empty, and the last step of deduction of $A\big(u\langle w\rangle\varepsilon\big)$ is thus $B\big(u\langle w\rangle\varepsilon\big)$, $J_1\big(\varepsilon\langle uw\rangle\varepsilon\big)$, …, $J_\ell\big(\varepsilon\langle uw\rangle\varepsilon\big) \vdash_G A\big(u\langle w\rangle\varepsilon\big)$.

  Construction 1 only adds the rule (10b) to R', when $(\{J_1, \ldots, J_\ell\}, C) \in \rhd\varepsilon\text{-}\textsc{Nullable}(G)$. Then, by Lemma 6, the item $C\big(u\langle\varepsilon\rangle\varepsilon\big)$ can be deduced in the grammar G. The item $B\big(u\langle w\rangle\varepsilon\big)$ can be deduced in G by the induction hypothesis. Now the item $A\big(u\langle w\rangle\varepsilon\big)$ can be deduced in G out of these premises: $B\big(u\langle w\rangle\varepsilon\big)$, $C\big(u\langle\varepsilon\rangle\varepsilon\big) \vdash_G A\big(u\langle w\rangle v\big)$.

- Let r' be of the form (10c). The last step of deduction of $A\big(u\langle w\rangle v\big)$ in G' is $C\big(u\langle w\rangle v\big)$, $J_1\big(\varepsilon\langle u\rangle wv\big)$, …, $J_\ell\big(\varepsilon\langle u\rangle wv\big)$, $K_1\big(u\langle wv\rangle\varepsilon\big)$, …, $K_t\big(u\langle wv\rangle\varepsilon\big)$. Since the rule (10c) is in R', the set $\textsc{Nullable}(G)$ contains an element $(\{J_1, \ldots, J_\ell\}, B, \{K_1, \ldots, K_t\})$, and therefore, by Lemma 4, one can deduce an item $B\big(u\langle\varepsilon\rangle wv\big)$ in G. Finally, the item $A\big(u\langle w\rangle v\big)$ can be deduced in G by the rule $A \to BC$ (which is in R by the construction) as follows: $B\big(u\langle\varepsilon\rangle wv\big)$, $C\big(u\langle w\rangle v\big) \vdash_G A\big(u\langle w\rangle v\big)$.

- Let r' be of the form (10d). This rule requires $u = \varepsilon$, and thus the last step of deduction of $A\big(u\langle w\rangle\varepsilon\big)$ takes the form $C\big(\varepsilon\langle w\rangle v\big)$, $K_1\big(\varepsilon\langle wv\rangle\varepsilon\big)$, …, $K_t\big(\varepsilon\langle wv\rangle\varepsilon\big) \vdash_G A\big(\varepsilon\langle w\rangle v\big)$. The item $C\big(\varepsilon\langle w\rangle v\big)$ can be deduced in G by the induction hypothesis.

  The rule (10b) is only added to R', when $(B, \{K_1, \ldots, K_t\}) \in \lhd\varepsilon\text{-}\textsc{Nullable}(G)$. Then, by Lemma 5, the item $B\big(\varepsilon\langle\varepsilon\rangle wv\big)$ can be deduced in the grammar G. Finally, the item $A\big(\varepsilon\langle w\rangle v\big)$ can be deduced in G out of the premises $B\big(\varepsilon\langle\varepsilon\rangle wv\big)$, $C\big(\varepsilon\langle w\rangle v\big)$ by the rule $A \to BC$. □

The construction eliminates the empty string in all direct conjuncts, but the resulting grammar may still contain epsilon contexts ($\lhd\varepsilon$ and $\rhd\varepsilon$), and the next step is to eliminate these contexts.

4.2 Null contexts

Example 9. Consider the following grammar with two-sided contexts:

$$\begin{aligned} S &\to Ab \mid bA \\ A &\to a \,\&\, \lhd\varepsilon \mid c \,\&\, \rhd\varepsilon \mid d, \end{aligned}$$

which defines the language $\{ab, db, bc, bd\}$. The conjunct $\lhd\varepsilon$ in the first rule for nonterminal A ensures that A can only generate a in the sentence of the form Ab (where A has left context ε), and not in bA. Similarly, $\rhd\varepsilon$ in the second rule for A restricts the generation of c only in the very end of the string.

In order to eliminate the contexts $\lhd\varepsilon$ and $\rhd\varepsilon$ from the grammar, one has to add two variants of nonterminal A: one with the empty left context (${}^0A^1$) and another with the empty right context (${}^1A^0$), having the rules ${}^0A^1 \to a$ and ${}^1A^0 \to c$, respectively. Since d can be generated by A in any contexts, including the empty ones, the nonterminals ${}^0A^1$ and ${}^1A^0$ also have rules ${}^0A^1 \to d$ and ${}^1A^0 \to d$.

The rules for the initial symbol are changed to $S \to {}^0A^1b \mid b\,{}^1A^0$, so that the knowledge on the context emptiness is passed from S down to A.

Finally, the rules of the new grammar are as follows:

$$\begin{aligned} S &\to {}^0A^1b \mid b{}^1A^0 \\ {}^0A^1 &\to a \mid d \\ {}^1A^0 &\to c \mid d \end{aligned}$$

The general construction makes four versions of each nonterminal: ${}^\ell A^r$, where $\ell, r \in \{0, 1\}$ determine the emptiness of the left and the right context.

Construction 2. Let $G = (\Sigma, N, R, S)$ be a grammar with two-sided contexts with all rules of the form

$$A \to a \tag{15a}$$
$$A \to BC \tag{15b}$$
$$A \to B_1 \,\&\, \ldots \,\&\, B_k \,\&\, \lhd D_1 \,\&\, \ldots \,\&\, \lhd D_m \,\&\, \trianglelefteq E_1 \,\&\, \ldots \,\&\, \trianglelefteq E_n \,\&\, \trianglerighteq F_1 \,\&\, \ldots \,\&\, \trianglerighteq F_{m'} \,\&\, \rhd H_1 \,\&\, \ldots \,\&\, \rhd H_{n'} \,\&\, {}^x\Delta^y, \tag{15c}$$

where $a \in \Sigma$, $A, B, C, B_i, D_i, E_i, F_i, H_i \in N$, and

$${}^x\Delta^y = \begin{cases} \Sigma^*, & \text{if } x = 1 \text{ and } y = 1 \\ \rhd\varepsilon, & \text{if } x = 1 \text{ and } y = 0 \\ \lhd\varepsilon, & \text{if } x = 0 \text{ and } y = 1 \\ \lhd\varepsilon \,\&\, \rhd\varepsilon, & \text{if } x = 0 \text{ and } y = 0. \end{cases}$$

Let $N' = \{\, {}^xA^y \mid A \in N,\ x, y \in \{0, 1\} \,\}$. Construct a grammar with two-sided contexts $G' = (\Sigma, N', R', {}^0S^0)$ with the following set of rules.

1. For every rule of the form (15a) in R, add to R' the four rules

$${}^\ell A^r \to a, \qquad \text{with } \ell, r \in \{0, 1\}. \tag{16a}$$

2. For every rule of the form (15b) in R, add to R' the four rules

$$ {}^{\ell}A^{r} \to {}^{\ell}B^{1}\ {}^{1}C^{r}, \qquad \text{with } \ell, r \in \{0,1\}. \tag{16b} $$

3. For every rule of the form (15c) in R, add to R' all possible rules of the form:

$$ {}^{\ell}A^{r} \to {}^{\ell}B_1{}^{r} \,\&\, \ldots \,\&\, {}^{\ell}B_k{}^{r} \,\&\, \lhd^{0}D_1{}^{1} \,\&\, \ldots \,\&\, \lhd^{0}D_m{}^{1} \,\&\, \trianglelefteqslant^{0}E_1{}^{r} \,\&\, \ldots \,\&\, \trianglelefteqslant^{0}E_n{}^{r} \,\&\, \\ \&\, \trianglerighteq^{\ell}F_1{}^{0} \,\&\, \ldots \,\&\, \trianglerighteq^{\ell}F_{m'}{}^{0} \,\&\, \rhd^{1}H_1{}^{0} \,\&\, \ldots \,\&\, \rhd^{1}H_{n'}{}^{0}, \tag{16c} $$

with $\ell, r \in \{0,1\}$, $\ell \leqslant x$ and $r \leqslant y$.

Lemma 8. *Let $G = (\Sigma, N, R, S)$ be a grammar with two-sided contexts, and let G' be the grammar obtained by Construction 2. Then for all $A \in N$ and for all $x, y \in \{0,1\}$, $L_{G'}({}^{x}A^{y}) = \{\, u\langle w\rangle v \mid u\langle w\rangle v \in L_G(A),\ \operatorname{sgn}|u| = x,\ \operatorname{sgn}|v| = y \,\}$.*

Claim 1.1. *Let $G = (\Sigma, N, R, S)$ be a grammar with two-sided contexts, and let G' be the grammar obtained by Construction 2. Then for all $A \in N$, $\vdash_G A\big(u\langle w\rangle v\big)$ implies that $\vdash_{G'} {}^{\operatorname{sgn}|u|}A^{\operatorname{sgn}|v|}\big(u\langle w\rangle v\big)$.*

Proof. The proof is by induction on d, the number of steps used in deduction of the item $A\big(u\langle w\rangle v\big)$ in G.

Basis. Let $d = 1$. Then $w = a \in \Sigma$ and the deduction of the item $A\big(u\langle w\rangle v\big)$ in the grammar G uses a rule $A \to a \in R$. According to Construction 2, the grammar G' has a rule of the form ${}^{\ell}A^{r} \to a$, with $\ell, r \in \{0,1\}$. Then $\vdash_{G'} {}^{\ell}A^{r}\big(u\langle a\rangle v\big)$ by this rule.

Induction step. Let $d > 1$ and let the item $\vdash_G A\big(u\langle w\rangle v\big)$ be deduced in the grammar G. Then the last step of its deduction can use a rule $p \in R$ which is either of the form (15b) or (15c).

1. Let p be of the form (15b). Then the item $A\big(u\langle w\rangle v\big)$ is deduced in the grammar G as $B\big(u\langle w_1\rangle w_2 v\big)$, $C\big(uw_1\langle w_2\rangle v\big) \vdash_G A\big(u\langle w\rangle v\big)$ for some partition $w_1 w_2 = w$. By induction hypothesis, $\vdash_{G'} {}^{\ell}B^{1}\big(u\langle w_1\rangle w_2 v\big)$ and $\vdash_{G'} {}^{1}C^{r}\big(uw_1\langle w_2\rangle v\big)$. Then ${}^{\ell}B^{1}\big(u\langle w_1\rangle w_2 v\big)\ {}^{1}C^{r}\big(uw_1\langle w_2\rangle v\big) \vdash_{G'} {}^{l}A^{r}\big(u\langle w\rangle v\big)$ (with $\ell = \operatorname{sgn}|u|$ and $r = \operatorname{sgn}|v|$), by the rule of the form (16b), which is added to R according to Construction 2.

2. Let p be of the form (15c). Then the item $A\big(u\langle w\rangle v\big)$ is deduced in the grammar G as follows: $B_i\big(u\langle w\rangle v\big)$, $D_i\big(\varepsilon\langle u\rangle wv\big)$, $E_i\big(\varepsilon\langle uw\rangle v\big)$, $F_i\big(u\langle wv\rangle\varepsilon\big)$ $H_i\big(uw\langle v\rangle\varepsilon\big) \vdash_G A\big(u\langle w\rangle v\big)$. By induction hypothesis, $\vdash_{G'} {}^{\ell}B_i{}^{r}\big(u\langle w\rangle v\big)$, $\vdash_{G'} {}^{0}D_i{}^{1}\big(\varepsilon\langle u\rangle wv\big)$, $\vdash_{G'} {}^{0}E_i{}^{r}\big(\varepsilon\langle uw\rangle v\big)$, $\vdash_{G'} {}^{\ell}F_i{}^{0}\big(u\langle wv\rangle\varepsilon\big)$ $\vdash_{G'} {}^{1}H_i{}^{0}\big(uw\langle v\rangle\varepsilon\big)$, where $l = \operatorname{sgn}|u|$ and $r = \operatorname{sgn}|v|$.

   According to Construction 2, the grammar G' has a rule of the form (16c), by which the item ${}^{\ell}A^{r}\big(u\langle w\rangle v\big)$ can be deduced as follows: ${}^{\ell}B_i{}^{r}\big(u\langle w\rangle v\big)$, ${}^{0}D_i{}^{1}\big(\varepsilon\langle u\rangle wv\big)$, ${}^{0}E_i{}^{r}\big(\varepsilon\langle uw\rangle v\big)$, ${}^{\ell}F_i{}^{0}\big(u\langle wv\rangle\varepsilon\big)$ ${}^{1}H_i{}^{0}\big(uw\langle v\rangle\varepsilon\big)$ $\vdash_{G'} {}^{\ell}A^{r}\big(u\langle w\rangle v\big)$. □

Claim 1.2. *Let $G = (\Sigma, N, R, S)$ be a grammar with two-sided contexts, and $G' = (\Sigma, N', R', {}^0S^0)$ be the grammar obtained by Construction 2. Moreover, let $\vdash_{G'} {}^0S^0\big(\varepsilon\langle w_0\rangle\varepsilon\big)$ (with $w_0 \in \Sigma^+$), and let T be a deduction tree of this item. Let I be a downward closed set of nodes in T. Then:*

1. *for each node $\tau = {}^\ell A^r\big(u\langle w\rangle v\big)$ in T such that $\tau \notin I$, it holds that $\ell = \operatorname{sgn}|u|$ and $r = \operatorname{sgn}|v|$;*

2. *$\hat{I} \vdash_G S\big(\varepsilon\langle w_0\rangle\varepsilon\big)$, where $\hat{I} = \{\, A\big(u\langle w\rangle v\big) \mid {}^\ell A^r\big(u\langle w\rangle v\big),\ \ell, r \in \{0,1\}\,\}$.*

Proof. The proof is by induction on d, the number of nodes in T which do not belong to the set I.

Basis. Let $d = 0$.

That is, there are no nodes τ in T such that $\tau \notin I$, and the first part of the claim is thus proved.

Since the set I contains every node of the tree T, it contains the element ${}^0S^0\big(\varepsilon\langle w_0\rangle\varepsilon\big)$. Then the set $\hat{I}$ contains the item $S\big(\varepsilon\langle w_0\rangle\varepsilon\big)$. Therefore, $\hat{I} \vdash_G S\big(\varepsilon\langle w_0\rangle\varepsilon\big)$ in zero steps, which proves the second part of the claim.

Induction step. Let $d \geqslant 1$ and let I be some downward closed set of nodes in T.

Consider any node τ in T such that $\tau \notin I$ and $I \vdash_{G'} {}^\ell A^r\big(u\langle w\rangle v\big)$ in one step. Such an item exists, since the set of nodes of T not in I forms a non-empty tree, which has leaves.

Define the set $I' = I \cup \tau$. This set is downward closed, since the deduction of the item τ out of the premises I is made in one step.

Let us prove the first part of the claim.

Consider any node in T which is not in I. If that node is not τ, then the first part of claim is given in the induction hypothesis for the set I'. Otherwise, that node is τ and it can either be a root of the tree or its internal node.

If τ is the root of the tree T, then $\tau = {}^0S^0\big(\varepsilon\langle w_0\rangle\varepsilon\big)$. Then $0 = \operatorname{sgn}|u|$, as claimed.

Assume that τ is in I' and is not the root of the tree T. Then τ has some parent node ${}^\ell A^r\big(u\langle w\rangle v\big)$ in T, such that it is deducible in G' by some rule $p \in R'$, and the item τ is among the premises for such deduction.

Let p be of the form (16b). Then ${}^\ell B^1\big(u\langle w_1\rangle w_2 v\big)$, ${}^1C^r\big(uw_1\langle w_2\rangle v\big) \vdash_{G'} {}^\ell A^r\big(u\langle w\rangle v\big)$ for some partition $w = w_1 w_2$.

1. Assume $\tau = {}^\ell B^1\big(u\langle w_1\rangle w_2 v\big)$. Since w_2 is not empty, this item satisfies $\operatorname{sgn}|w_2 v| = 1$, as desired. Since the first part of the claim holds for the item ${}^\ell A^r\big(u\langle w\rangle v\big)$, it also holds for the item ${}^\ell B^1\big(u\langle w_1\rangle w_2 v\big)$, that is, $\ell = \operatorname{sgn}|u|$.

2. The case when $\tau = {}^1C^r\big(uw_1\langle w_2\rangle v\big)$ can be proved in the same way.

Let p be of the form (16c). Then ${}^{\ell}B_i{}^{r}\big(u\langle w\rangle v\big)$, ${}^{0}D_i{}^{1}\big(\varepsilon\langle u\rangle wv\big)$, ${}^{0}E_i{}^{r}\big(\varepsilon\langle uw\rangle v\big)$, ${}^{\ell}F_i{}^{0}\big(u\langle wv\rangle\varepsilon\big)$ ${}^{1}H_i{}^{0}\big(uw\langle v\rangle\varepsilon\big) \vdash_{G'} {}^{\ell}A^{r}\big(u\langle w\rangle v\big)$.

1. Let $\tau = {}^{\ell}B_i{}^{r}\big(u\langle w\rangle v\big)$. Since the first part of the claim holds for the item ${}^{\ell}A^{r}\big(u\langle w\rangle v\big)$, it should also hold for ${}^{\ell}B_i{}^{r}\big(u\langle w\rangle v\big)$, that is, $\ell = \operatorname{sgn}|u|$ and $r = \operatorname{sgn}|v|$.

2. Let $\tau = {}^{0}D_i{}^{1}\big(\varepsilon\langle u\rangle wv\big)$. The left context indicator of the nonterminal ${}^{0}D_i{}^{1}$ is $\operatorname{sgn}|\varepsilon| = 0$, and its right context indicator is $\operatorname{sgn}|wv| = 1$ (since w cannot be empty), as desired.

3. The case when $\tau = {}^{1}H_i{}^{0}\big(uw\langle v\rangle\varepsilon\big)$ can be proved similarly.

4. Let $\tau = {}^{0}E_i{}^{r}\big(\varepsilon\langle uw\rangle v\big)$. The left context indicator for the nonterminal ${}^{0}E_i{}^{r}$ is $\operatorname{sgn}|\varepsilon| = 0$, as desired. Since the first part of the claim holds for the item ${}^{\ell}A^{r}\big(u\langle w\rangle v\big)$, it also holds for the item ${}^{0}E_i{}^{r}\big(\varepsilon\langle uw\rangle v\big)$. That is, $r = \operatorname{sgn}|v|$.

5. The case when $\tau = {}^{\ell}F_i{}^{0}\big(u\langle wv\rangle\varepsilon\big)$ is considered in an analogous way.

Let us now prove the second part of the claim. By induction hypothesis, applied to the set I',

$$\hat{I}' \vdash_G S\big(\varepsilon\langle w_0\rangle\varepsilon\big). \tag{17}$$

Let $\tau = {}^{\ell}A^{r}\big(u\langle w\rangle v\big)$ (with $\ell, r \in \{0,1\}$).
Define

$$\hat{I} = \hat{I}' \setminus \{A\big(u\langle w\rangle v\big)\}. \tag{18}$$

and show that

$$\hat{I} \vdash_G A\big(u\langle w\rangle v\big). \tag{19}$$

Consider the tree T. The node ${}^{\ell}A^{r}\big(u\langle w\rangle v\big)$ should have child nodes which represent a deduction of this item according to some rule $p \in R'$.

Depending on the form of this rule, the following cases are possible.

1. Let p be of the form (16a). Then $w = a \in \Sigma$ and $a\big(u\langle a\rangle v\big) \vdash_{G'} {}^{\ell}A^{r}\big(u\langle a\rangle v\big)$.

   According to Construction 2, the grammar G has the rule $A \to a \in R$, by which the item $A\big(u\langle w\rangle v\big)$ can be deduced: $a\big(u\langle a\rangle v\big) \vdash_G A\big(u\langle a\rangle v\big)$.

2. Let p be of the form (16b). Then ${}^{\ell}B^{1}\big(u\langle w_1\rangle w_2 v\big)$, ${}^{1}C^{r}\big(uw_1\langle w_2\rangle v\big) \vdash_{G'} {}^{\ell}A^{r}\big(u\langle w\rangle v\big)$ (for some partition $w = w_1 w_2$).

   The set $\hat{I}$ contains the items $B\big(u\langle w_1\rangle w_2 v\big)$ and $C\big(uw_1\langle w_2\rangle v\big)$.

   According to Construction 2, the grammar G has a rule $A \to BC$, by which the item $A\big(u\langle w\rangle v\big)$ can be deduced in the grammar G: $B\big(u\langle w_1\rangle w_2 v\big)$, $C\big(uw_1\langle w_2\rangle v\big) \vdash_G A\big(u\langle w\rangle v\big)$.

3. Let p be of the form (16c).

   In this case, ${}^{\ell}B_i{}^{r}\big(u\langle w\rangle v\big)$, ${}^{0}D_i{}^{1}\big(\varepsilon\langle u\rangle wv\big)$, ${}^{0}E_i{}^{r}\big(\varepsilon\langle uw\rangle v\big)$, ${}^{\ell}F_i{}^{0}\big(u\langle wv\rangle\varepsilon\big)$ ${}^{1}H_i{}^{0}\big(uw\langle v\rangle\varepsilon\big) \vdash_{G'} {}^{\ell}A^{r}\big(u\langle w\rangle v\big)$ and the set $\hat{I}$ contains the items $B_i\big(u\langle w\rangle v\big)$, $D_i\big(\varepsilon\langle u\rangle wv\big)$, $E_i\big(\varepsilon\langle uw\rangle v\big)$, $F_i\big(u\langle wv\rangle\varepsilon\big)$ $H_i\big(uw\langle v\rangle\varepsilon\big)$.

   According to Construction 2, the grammar G has a rule of the form

$$A \to B_1 \,\&\, \ldots \,\&\, B_k \,\&\, \lhd D_1 \,\&\, \ldots \,\&\, \lhd D_m \,\&\, \lhdeq E_1 \,\&\, \ldots \,\&\, \lhdeq E_n \,\& \\ \& \,\rhdeq F_1 \,\&\, \ldots \,\&\, \rhdeq F_{m'} \,\&\, \rhd H_1 \,\&\, \ldots \,\&\, \rhd H_{n'} \,\&\, {}^{x}\Delta^{y}, \tag{20}$$

   with $x, y \in \{0, 1\}$.

   - Let $x = 1$ and $y = 1$. That is, the rule (20) does not have $\lhd\varepsilon$- or $\rhd\varepsilon$-conjuncts. Then the rule p may have any $\ell, r \in \{0, 1\}$.

     The item $A\big(u\langle w\rangle v\big)$ can be deduced in the grammar G as follows: $B_i\big(u\langle w\rangle v\big)$, $D_i\big(\varepsilon\langle u\rangle wv\big)$, $E_i\big(\varepsilon\langle uw\rangle v\big)$, $F_i\big(u\langle wv\rangle\varepsilon\big)$, $H_i\big(uw\langle v\rangle\varepsilon\big) \vdash_G A\big(u\langle w\rangle v\big)$.
   - Let $x = 1$ and $y = 0$. Then the rule (20) has a $\rhd\varepsilon$-conjunct and the condition for the rule (16c) requires that $r = 0$. By the first part of this claim for ${}^{l}A^{r}\big(u\langle w\rangle v\big)$, it holds that $v = \varepsilon$ and $u \in \Sigma^*$. Thus, the item $A\big(u\langle w\rangle v\big)$ is deduced as $B_i\big(u\langle w\rangle\varepsilon\big)$, $D_i\big(\varepsilon\langle u\rangle w\big)$, $E_i\big(\varepsilon\langle uw\rangle\varepsilon\big)$, $F_i\big(u\langle w\rangle\varepsilon\big) \vdash_G A\big(u\langle w\rangle\varepsilon\big)$.
   - Let $x = 0$ and $y = 1$, that is, the rule (20) has a $\lhd\varepsilon$-conjunct. This case is proved similarly to the previous one, and the item $A\big(u\langle w\rangle v\big)$ can be deduced as: $B_i\big(\varepsilon\langle w\rangle v\big)$, $E_i\big(\varepsilon\langle w\rangle v\big)$, $F_i\big(\varepsilon\langle wv\rangle\varepsilon\big)$, $H_i\big(w\langle v\rangle\varepsilon\big) \vdash_G A\big(\varepsilon\langle w\rangle v\big)$.
   - Let $x = y = 0$, that is the rule (20) has both $\lhd\varepsilon$- and $\rhd\varepsilon$-conjuncts. In this case $\ell = r = 0$. Hence, $u = v = \varepsilon$ and the item $A\big(\varepsilon\langle w\rangle\varepsilon\big)$ can be deduced in the grammar G out of the premises $B_i\big(\varepsilon\langle w\rangle\varepsilon\big)$, $E_i\big(\varepsilon\langle w\rangle\varepsilon\big)$, $F_i\big(\varepsilon\langle w\rangle\varepsilon\big)$.

   Finally, it follows from (17)–(19) that $\hat{I} \vdash_G S\big(\varepsilon\langle w_0\rangle\varepsilon\big)$. □

4.3 Unit conjuncts

The third stage of the transformation to the normal form is removing the *unit conjuncts* in rules of the form $A \to B \,\&\, \ldots$. Already for conjunctive grammars [8], the only known transformation involves substituting all rules for B into all rules for A; in the worst case, this results in an exponential blowup. The same construction applies verbatim to grammars with contexts.

Theorem 2. *For each grammar with left contexts $G = (\Sigma, N, R, S)$ there exists and can be effectively constructed a grammar with left contexts $G' = (\Sigma, N', R', S)$ in the binary normal form, such that $L(G) = L(G') \setminus \{\varepsilon\}$.*

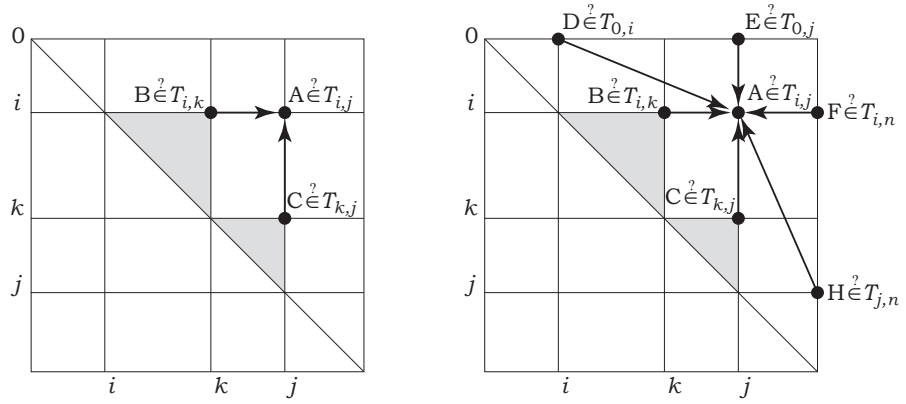


Figure 4: How the membership of A in $T_{i,j}$ depends on other data, for rules (a) $A \to BC$ and (b) $A \to BC \mathbin{\&} {\lhd}D \mathbin{\&} {\trianglelefteq}E \mathbin{\&} {\trianglerighteq}F \mathbin{\&} {\rhd}H$.

5 Parsing algorithm

Let $G = (\Sigma, N, R, S)$ be a grammar with two-sided contexts in the binary normal form, and let $w = a_1 \dots a_n \in \Sigma^+$ with $n \geqslant 1$ and $a_i \in \Sigma$ be an input string to be parsed. For every two positions i, j with $0 \leqslant i < j \leqslant n$, let

$$T_{i,j} = \left\{ A \;\middle|\; A \in N,\ \vdash_G A\big(a_1 \dots a_i \langle a_{i+1} \dots a_j \rangle a_{j+1} \dots a_n\big) \right\}$$

be the set of nonterminals generating the corresponding substring. In particular, the string w is in $L(G)$ if and only if $S \in T_{0,n}$.

In ordinary context-free grammars, as well as in their conjunctive variant, each set $T_{i,j}$ depends only on the sets $T_{i',j'}$ with $j' - i' < j - i$, and hence all these sets may be constructed inductively, beginning with shorter substrings and eventually reaching the set $T_{0,n}$ [8]. In grammars with only left contexts, each set $T_{i,j}$ additionally depends on the sets $T_{0,i}$ and $T_{0,j}$ via the conjuncts of the form ${\lhd}D$ and ${\trianglelefteq}E$, respectively, which allows constructing these sets progressively for $j = 1, \dots, n$ [2]. The structure of logical dependencies in grammars with two-sided contexts is more complicated, as shown in the following example.

Example 10. Consider the grammar with the rules $S \to AB$, $A \to a \mathbin{\&} {\rhd}B$, $B \to b \mathbin{\&} {\lhd}C$ and $C \to a$, and the input string $w = ab$. It is immediately seen that $C \in T_{0,1}$. From this one can infer that $B \in T_{1,2}$, and that knowledge can in turn be used to show that $A \in T_{0,1}$. These data imply that $S \in T_{0,2}$. Thus, none of the sets $T_{0,1}$ and $T_{1,2}$ can be constructed before approaching the other.

The proposed algorithm for constructing the sets $T_{i,j}$ works as follows. At the first pass, it makes all deductions $\vdash_G A\big(a_1 \dots a_i \langle a_{i+1} \dots a_j \rangle a_{j+1} \dots a_n\big)$

that do not involve any contexts, and accordingly puts A to the corresponding $T_{i,j}$. This pass is done progressively for longer and longer substrings, as in the case of context-free grammars. During this first pass, some symbols may be added to any $T_{0,j}$ and $T_{i,n}$, so that some contexts are known to be true. Then the algorithm makes another pass over all entries $T_{i,j}$, from shorter substrings to longer ones, this time using the known true contexts in the deductions. This may result in adding more elements to $T_{0,j}$ and $T_{i,n}$, which will require another pass. Since a new pass is needed only if any new element is added to any of $2n-1$ subsets of N, the total number of passes is at most $(2n-1) \cdot |N| + 1$.

For succinctness, the algorithm uses the following notation for multiple context operators. For a set $\mathcal{X} = \{X_1, \ldots, X_\ell\}$, with $X_i \in N$, and for an operator $Q \in \{\lhd, \trianglelefteq, \trianglerighteq, \rhd\}$, denote $Q\mathcal{X} := \{QX_1, \ldots, QX_\ell\}$.

Algorithm 1. Let $G = (\Sigma, N, R, S)$ be a grammar with contexts in the binary normal form. Let $w = a_1 \ldots a_n \in \Sigma^+$ (with $n \geqslant 1$ and $a_i \in \Sigma$) be the input string. Let $T_{i,j}$ with $0 \leqslant i < j \leqslant n$ be variables, each representing a subset of N, and let $T_{i,j} = \varnothing$ be their initial values.

```
1: while any of T_{0,j} (1 ⩽ j ⩽ n) or T_{i,n} (1 ⩽ i < n) change do
2:     for j = 1, ..., n do
3:         for all A → a & ⊲D & ⋖E & ⋗F & ⊳H ∈ R do
4:             if a_j = a ∧ D ⊆ T_{0,j−1} ∧ E ⊆ T_{0,j} ∧ F ⊆ T_{j,n} ∧ H ⊆ T_{j+1,n}
               then
5:                 T_{j−1,j} = T_{j−1,j} ∪ {A}
6:         for i = j − 2 to 0 do
7:             let U = ∅ (U ⊆ N × N)
8:             for k = i + 1 to j − 1 do
9:                 U = U ∪ (T_{i,k} × T_{k,j})
10:            for all A → B_1C_1 & ... & B_mC_m & ⊲D & ⋖E & ⋗F & ⊳H ∈
               R do
11:                if (B_1, C_1), ..., (B_m, C_m) ∈ U ∧ D ⊆ T_{0,i} ∧ E ⊆ T_{0,j} ∧
                   F ⊆ T_{i,n} ∧ H ⊆ T_{j,n} then
12:                    T_{i,j} = T_{i,j} ∪ {A}
13: accept if and only if S ∈ T_{0,n}
```

Theorem 3. *For every grammar with two-sided contexts G in the binary normal form, Algorithm 1, given an input string $w = a_1 \ldots a_n$, constructs the sets $T_{i,j}$ and determines the membership of w in $L(G)$, and does so in time $\mathcal{O}(|G|^2 \cdot n^4)$, using space $\mathcal{O}(|G| \cdot n^2)$.*

Each pass of Algorithm 1 is the same as the entire parsing algorithm for grammars without contexts [8], and that algorithm can be accelerated

by changing the order of computing the entries $T_{i,j}$, so that most of the calculations can be offloaded to a procedure for multiplying Boolean matrices [20, 13]. If $\mathrm{BMM}(n)$ is the complexity of multiplying two $n \times n$ Boolean matrices, the resulting algorithm works in time $\mathrm{BMM}(n)$. By the same method, Algorithm 1 can be restated to make $\mathcal{O}(n)$ such passes, with the following improvement in running time.

Theorem 4. *For every grammar with two-sided contexts G in the binary normal form, there is an algorithm to determine whether a given string $w = a_1 \dots a_n$ is in $L(G)$, which works in time $\mathcal{O}(|G|^2 \cdot n \cdot \mathrm{BMM}(n))$, using space $\mathcal{O}(|G| \cdot n^2)$.*

Let ω be the infimum of all real numbers, for which $\mathrm{BMM}(n) = n^{\omega+o(1)}$; it is known that $2 \leqslant \omega < 2.373$. Then the complexity of known parsing algorithms for different families of formal grammars is as follows (with $n^{o(1)}$ factor omitted):

- n^2 for unambiguous conjunctive and Boolean grammars, as well as for all their subclasses down to linear grammars with disjunction only [12];
- n^ω for the general case of context-free grammars, including their conjunctive and Boolean variants [20, 13];
- n^3 for grammars with one-sided contexts [2];
- $n^{\omega+1}$ *for grammars with two-sided contexts (Theorem 4);*
- $n^{2\omega}$ for tree-adjoining grammars [17].

All these grammar formalisms can be described in the much more general logic ILFP, introduced by Rounds [18], which allows recursive definitions of n-ary predicates on positions in a string, existential and universal quantification over such positions, as well as conjunction and disjunction. That logic is the natural general concept for all known meaningful families of formal grammars.

The complexity of parsing for different families of formal grammars is illustrated in Figure 5.

To conclude, this paper has developed a formal representation for the idea of phrase-structure rules applicable in a context, featuring in the early work of Chomsky [3]. This idea did not receive adequate treatment before, due to the inappropriate string-rewriting approach. Perhaps there were other good ideas in the theory of formal grammars, which were incorrectly formalized before, and could be re-investigated using the logical approach?

Another possibility for further studies is investigating Boolean and stochastic variants of grammars with contexts, following the recent related work [4, 7, 21].

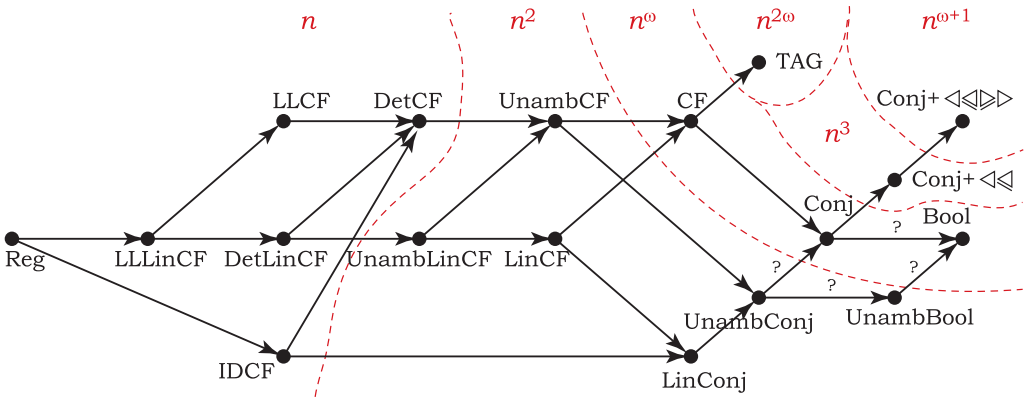


Figure 5: Hierarchy of grammar families. Complexity of known parsing algorithms.

Acknowledgements

Supported by the Academy of Finland under grant 257857.

References

[1] T. Aizikowitz, M. Kaminski, "LR(0) conjunctive grammars and deterministic synchronized alternating pushdown automata", *Computer Science in Russia* (CSR 2011, St. Petersburg, Russia, 14–18 June 2011), LNCS 6651, 345–358.

[2] M. Barash, A. Okhotin, "Defining contexts in context-free grammars", *Language and Automata Theory and Applications* (LATA 2012, A Coruña, Spain, 5–9 March 2012), LNCS 7183, 106–118.

[3] N. Chomsky, "On certain formal properties of grammars", *Information and Control*, 2:2 (1959), 137–167.

[4] Z. Ésik, W. Kuich, "Boolean fuzzy sets", *International Journal of Foundations of Computer Science*, 18:6 (2007), 1197–1207.

[5] S. Ginsburg, H. G. Rice, "Two families of languages related to ALGOL", *Journal of the ACM*, 9 (1962), 350–371.

[6] A. Jeż, "Conjunctive grammars can generate non-regular unary languages", *International Journal of Foundations of Computer Science*, 19:3 (2008), 597–615.

[7] V. Kountouriotis, Ch. Nomikos, P. Rondogiannis, "Well-founded semantics for Boolean grammars", *Information and Computation*, 207:9 (2009), 945–967.

[8] A. Okhotin, "Conjunctive grammars", *Journal of Automata, Languages and Combinatorics*, 6:4 (2001), 519–535.

[9] A. Okhotin, "Conjunctive grammars and systems of language equations", *Programming and Computer Software*, 28:5 (2002), 243–249.

[10] A. Okhotin, "Boolean grammars", *Information and Computation*, 194:1 (2004), 19–48.

[11] A. Okhotin, "Generalized LR parsing algorithm for Boolean grammars", *International Journal of Foundations of Computer Science*, 17:3 (2006), 629–664.

[12] A. Okhotin, "Unambiguous Boolean grammars", *Information and Computation*, 206 (2008), 1234–1247.

[13] A. Okhotin, "Fast parsing for Boolean grammars: a generalization of Valiant's algorithm", *Developments in Language Theory* (DLT 2010, London, Ontario, Canada, August 17–20, 2010), LNCS 6224, 340–351.

[14] A. Okhotin, "Conjunctive and Boolean grammars: the true general case of the context-free grammars", *Computer Science Review*, 9 (2013), 27–59.

[15] A. Okhotin, "Improved normal form for grammars with one-sided contexts", *Descriptional Complexity of Formal Systems* (DCFS 2013, London, Ontario, Canada, 22-25 July 2013), LNCS 8031, 205–216.

[16] A. Okhotin, C. Reitwießner, "Conjunctive grammars with restricted disjunction", *Theoretical Computer Science*, 411:26–28 (2010), 2559–2571.

[17] S. Rajasekaran, S. Yooseph, "TAL recognition in $\mathcal{O}(M(n^2))$ time", *Journal of Computer and System Sciences*, 56:1 (1998), 83–89.

[18] W. C. Rounds, "LFP: A logic for linguistic descriptions and an analysis of its complexity", *Computational Linguistics*, 14:4 (1988), 1–9.

[19] I. H. Sudborough, "A note on tape-bounded complexity classes and linear context-free languages", *Journal of the ACM*, 22:4 (1975), 499–500.

[20] L. G. Valiant, "General context-free recognition in less than cubic time", *Journal of Computer and System Sciences*, 10:2 (1975), 308–314.

[21] R. Zier-Vogel, M. Domaratzki, "RNA pseudoknot prediction through stochastic conjunctive grammars", *Computability in Europe 2013. Informal Proceedings*, 80–89.

Turku
Centre *for*
Computer
Science

Joukahaisenkatu 3-5 A, 20520 TURKU, Finland | www.tucs.fi

University of Turku
Faculty of Mathematics and Natural Sciences
- Department of Information Technology
- Department of Mathematics

Turku School of Economics
- Institute of Information Systems Sciences

Åbo Akademi University
- Department of Computer Science
- Institute for Advanced Management Systems Research

ISBN 978-952-12-2963-3
ISSN 1239-1891