Maria Bulińska

On the Complexity of Nonassociative Lambek Calculus with Unit

Abstract. Nonassociative Lambek Calculus (NL) is a syntactic calculus of types introduced by Lambek [8]. The polynomial time decidability of NL was established by de Groote and Lamarche [4]. Buszkowski [3] showed that systems of NL with finitely many assumptions are decidable in polynomial time and generate context-free languages; actually the P-TIME complexity is established for the consequence relation of NL. Adapting the method of Buszkowski [3] we prove an analogous result for Nonassociative Lambek Calculus with unit (NL1). Moreover, we show that any Lambek grammar based on NL1 (with assumptions) can be transformed into an equivalent context-free grammar in polynomial time.

Keywords: Nonassociative Lambek calculus, P-TIME decidability, Context-free grammar

1. Introduction

The Lambek calculus (associative or nonassociative) is a calculus of types introduced by Lambek [7, 8] in order to consider formal grammars as deductive systems. Lambek also proved that the derivability problem for associative and nonassociative Lambek calculus is decidable. The polynomial time decidability of Nonassociative Lambek Calculus (NL) was established by de Groote and Lamarche [4]. Buszkowski [3] showed that the consequence relation of NL is decidable in polynomial time, and the corresponding categorial grammars generate context-free languages; the same holds for systems with unary modalities, studied in Moortgat [10] and for Generalized Lambek Calculus (with n-ary operations). The context-freeness of the languages generated by NL was earlier proved by Buszkowski [2], Kandulski [6], and a new proof was proposed by Jäger [5]. The method of interpolation used in [5] is also essential in our work. We take into consideration Nonassociative Lambek Calculus with unit (NL1). NL1 enriched with a finite set of assumptions Γ is denoted by NL1(Γ). To show that the provability in NL1(Γ) is decidable in polynomial time we adapt the method of Buszkowski

Presented by **Name of Editor**; *Received* June 1, 2009

Studia Logica (2009) 93: 1–14
DOI: 10.1007/s11225-009-9205-2

[3]. Our modification of this method is essential, since we have to eliminate "nullary" sequents $\Lambda \to A$ as premises of the cut rule. The context-freeness of the languages generated by systems NL1(Γ) is also established. Further, we show that any categorial grammar based on NL1(Γ) can be transformed into an equivalent context-free grammar in polynomial time.

For certain reasons one may be interested in adding some assumptions to the calculus. What can we gain in this way? First of all, we can study the consequence relation associated with the given logic. Further, we can use additional assumptions to describe subcategorization in a natural language. For example, Lambek [9] uses axioms of the form $\pi_i \to \pi$ to express the inclusion of the class of pronouns in i-th Person in the class of pronouns. Bulińska [1] obtained the weak equivalence of context-free grammars and grammars based on the Associative Lambek Calculus with any finite set of simple assumptions of the form $p \to q$, where p, q are primitive types. With arbitrary assumptions, systems of Associative Lambek Calculus generate all recursively enumerable languages [3]. We can use this fact to surpass the limitations of context-free languages. NL is naturally related to tree structures of linguistic expressions. By enriching this calculus with finite set of assumptions, we can improve its expressibility without losing the nice computational simplicity. For example, one can take NL as the basic logic and add axioms of the form $(A \bullet B) \bullet C \leftrightarrow A \bullet (B \bullet C)$ and $A \bullet B \leftrightarrow B \bullet A$, for some concrete types A, B, C to admit associativity and permutation in some special cases. (Moortgat [10] prefers to apply axiom-schemes for types preceded by special modalities.) Sometimes a limited usage of other structural rules may be helpful. For instance, any sentence conjunction of type $S\backslash(S/S)$ can also act as a verb phrase conjunction of type $VP\backslash(VP/VP)$, where $VP = PN\backslash S$. But in Lambek calculus we cannot transform the former type into the latter (in Associative Lambek Calculus permutation and contraction are needed). So, the sequent $S\backslash(S/S) \to VP\backslash(VP/VP)$ can be added as an assumption.

2. Preliminaries

First we describe the formalism of NL1. Let At be a denumerable set of atoms (primitive types). Formulas (also called types) are built from atoms $p, q, r, \ldots$ and the constant $\mathbf{1}$ by means of three binary connectives $\backslash$, $/$, $\bullet$, called *left residuation, right residuation*, and *product*, respectively. We denote the set of all formulas by Tp1. The set of formula structures STR1 is defined recursively as follows: (i) $\Lambda \in$ STR1, where Λ denotes the empty structure, (ii) all formulas are (atomic) formula structures, i.e. Tp1 $\subseteq$ STR1,

(iii) if $X, Y \in$ STR1, then $(X \circ Y) \in$ STR1. We set $(X \circ \Lambda) = (\Lambda \circ X) = X$. Substructures of a formula structure are defined in the following way: (i) Λ is the only substructure of Λ, (ii) if X is an atomic formula structure, then Λ and X are the only substructures of X, (iii) if $X = (X_1 \circ X_2)$, then X and all substructures of X_1 and X_2 are substructures of X. By $X[Y]$ we denote a formula structure X with a distinguished substructure Y, and by $X[Z]$ - the substitution of Z for Y in X. Sequents are formal expressions $X \to A$ such that $A \in$ Tp1, $X \in$ STR1. The Gentzen-style axiomatization of the calculus NL1 employs the axioms:

$$\text{(Id)} \quad A \to A \qquad\qquad (\mathbf{1}\text{R}) \quad \Lambda \to \mathbf{1}$$

and the following rules of inference:

$$(\mathbf{1}\text{L}) \quad \frac{X[\Lambda] \to A}{X[\mathbf{1}] \to A},$$

$$(\bullet\text{L}) \quad \frac{X[A \circ B] \to C}{X[A \bullet B] \to C}, \qquad\qquad (\bullet\text{R}) \quad \frac{X \to A; \quad Y \to B}{X \circ Y \to A \bullet B},$$

$$(\backslash\text{L}) \quad \frac{Y \to A; \quad X[B] \to C}{X[Y \circ (A\backslash B)] \to C}, \qquad\qquad (\backslash\text{R}) \quad \frac{A \circ X \to B}{X \to A\backslash B},$$

$$(/\text{L}) \quad \frac{X[A] \to C; \quad Y \to B}{X[(B/A) \circ Y] \to C}, \qquad\qquad (/\text{R}) \quad \frac{X \circ B \to A}{X \to A/B},$$

$$\text{(CUT)} \quad \frac{Y \to A; \quad X[A] \to B}{X[Y] \to B}.$$

For any system S we write S $\vdash X \to A$ if the sequent $X \to A$ is derivable in S.

The algebraic models of NL1 are unital residuated groupoids . A *unital residuated groupoid* is a structure $\mathcal{M} = (M, \leq, \cdot, \backslash, /, 1)$ such that $(M, \leq)$ is a poset , $(M, \cdot, 1)$ is a groupoid with unit 1, satisfying $a \cdot 1 = a$, $1 \cdot a = a$ for all $a \in M$, and $\backslash, /$ are binary operations on M satisfying the equivalences :

$$\text{(RES)} \quad ab \leq c \quad \text{iff} \quad b \leq a\backslash c \quad \text{iff} \quad a \leq c/b$$

for all $a, b, c \in M$. Every residuated groupoid fulfills the following monotonicity laws:

$$\text{(MON)} \quad \text{if} \quad a \leq b \quad \text{then} \quad ca \leq cb \quad \text{and} \quad ac \leq bc$$

$$\text{(MRE)} \quad \text{if} \quad a \leq b \quad \text{then} \quad c\backslash a \leq c\backslash b, \quad a/c \leq b/c, \quad b\backslash c \leq a\backslash c, \quad c/b \leq c/a$$

for all $a, b, c \in M$. A *model* is a pair $(\mathcal{M}, \mu)$ such that $\mathcal{M}$ is a unital residuated groupoid and μ is an assignment of elements of M to atoms. One extends μ for all formulas :

$$\mu(\mathbf{1}) = 1, \quad \mu(A \bullet B) = \mu(A) \cdot \mu(B),$$

$$\mu(A \backslash B) = \mu(A) \backslash \mu(B), \quad \mu(A/B) = \mu(A)/\mu(B).$$

and formula structures:

$$\mu(\Lambda) = 1, \quad \mu(X \circ Y) = \mu(X) \cdot \mu(Y).$$

A sequent $X \to A$ is said to be true in the model $(\mathcal{M}, \mu)$ if $\mu(X) \leq \mu(A)$. In particular a sequent $\Lambda \to A$ is true in $(\mathcal{M}, \mu)$ if $1 \leq \mu(A)$. One can prove the following property for formula structures:

$$(\mathrm{MON-STR}) \quad \text{if} \quad \mu(Y) \leq \mu(Z) \quad \text{then} \quad \mu(X[Y]) \leq \mu(X[Z]).$$

3. NL1 with assumptions

Let Γ be a set of sequents of the form $A \to B$, where $A, B \in \mathrm{Tp1}$. By ($\mathbf{1}$L), ($\bullet$L), ($\bullet$R) and (CUT), every sequent is deductively equivalent in NL1 to a sequent $A \to B$. By NL1(Γ) we denote the calculus NL1 with the additional set Γ of assumptions. NL1 is strongly complete with respect to unital residuated groupoids , i.e. the sequents provable in NL1(Γ) are precisely those which are true in all models $(\mathcal{M}, \mu)$ in which all sequents from Γ are true. Soundness is easily provable by induction on derivations in NL1(Γ). Completeness follows from the fact that the Lindenbaum algebra of NL1 is a unital residuated groupoid.

In general, the calculus NL1(Γ) does not possess the standard subformula property, since (CUT) is a legal rule in this system. Thus, we take into consideration the subformula property in some extended form. Hereafter, we always assume that T is a set of formulas closed under subformulas and such that $\mathbf{1} \in T$, and all formulas appearing in Γ belong to T. By a T-sequent we mean a sequent $X \to A$ such that A and all formulas appearing in X belong to T. Now, we can reformulate the subformula property as follows: every T-sequent provable in NL1(Γ) has a proof in this system such that all sequents appearing in this proof are T-sequents.

To prove the subformula property for NL1(Γ) we will use special models, namely unital residuated groupoids of cones over given preordered unital groupoids. Let $(M, \leq, \cdot, 1)$ be a preordered unital groupoid, that means, it is a unital groupoid with a preordering (i.e. a reflexive and transitive relation), satisfying (MON). A set $P \subseteq M$ is called a *cone* (or: a *downset*) on M if $a \leq b$ and $b \in P$ entails $a \in P$. Let $C(M)$ denotes the set of cones on M. The cone I and the operations $\cdot, \backslash, /$ on $C(M)$ are defined as follows:

$$\text{(M1)} \quad I = \{a \in M : a \leq 1\}$$

$$\text{(M2)} \quad P_1P_2 = \{c \in M : (\exists a \in P_1, b \in P_2)\ c \leq ab\}$$

$$\text{(M3)} \quad P_1 \backslash P_2 = \{c \in M : (\forall a \in P_1)\ ac \in P_2\}$$

$$\text{(M4)} \quad P_1/P_2 = \{c \in M : (\forall b \in P_2)\ cb \in P_1\}.$$

The structure $(C(M), \subseteq, \cdot, \backslash, /, I)$ is a unital residuated groupoid. It is called the unital residuated groupoid of cones over the given preordered unital groupoid.

Let M be the set of all formula structures whose all atomic substructures belong to T and $\Lambda \in M$. If a sequent $X \to A$ has a proof in NL1(Γ) consisting of T-sequents only, we write: $X \to_T A$. First, we define a relation $\leq_b$ on M. $X \leq_b Y$ is read: X *directly reduces* to Y. The definition of this relation is as follows:

$$Y[Z] \leq_b Y[\Lambda] \quad \text{if} \quad Z \to_T \mathbf{1},$$

$$Y[Z] \leq_b Y[A] \quad \text{if} \quad Z \to_T A,$$

$$Y[A \bullet B] \leq_b Y[A \circ B] \quad \text{if} \quad A \bullet B \in T.$$

A preordering $\leq$ on M is defined as the reflexive and transitive closure of the relation $\leq_b$. Then $X \leq Y$ iff there exist $Y_0, \ldots, Y_n, n \geq 0$ such that $X = Y_0, Y = Y_n$ and $Y_{i-1} \leq_b Y_i$, for each $i = 1, \ldots, n$. $X \leq Y$ is read: X *reduces* to Y.

Clearly, $(M, \leq, \circ, \Lambda)$ is a preordered unital groupoid (with unit Λ). Moreover, we have the following fact.

FACT 3.1. *If $Y \to_T A$ and $X \leq Y$, then $X \to_T A$.*

PROOF. Assume $Y \to_T A$ and $X \leq Y$. Then, there exist $Y_0, \ldots, Y_n, n \geq 0$ such that $X = Y_0, Y = Y_n$ and $Y_{i-1} \leq_b Y_i$, for each $i = 1, \ldots, n$. We proceed by induction on n. For $n = 0$, $X = Y_0 = Y$ and of course, $X \to_T A$. For $n = 1$, we have $X \leq_b Y$. We consider three cases according to the definition of $\leq_b$.

(1) $X = X[Z]$, $Y = X[\Lambda]$, $Z \to_T \mathbf{1}$. Applying ($\mathbf{1}$L) to $X[\Lambda] \to_T A$ we get $X[\mathbf{1}] \to_T A$. By (CUT) from the last sequent and $Z \to_T \mathbf{1}$ we have $X[Z] \to_T A$.

(2) $X = X[Z]$, $Y = X[B]$, $Z \to_T B$. Then, we get $X[Z] \to_T A$, as a conclusion of (CUT) from premisses $X[B] \to_T A$ and $Z \to_T B$.

(3) $X = X[B \bullet C]$, $Y = X[B \circ C]$. Then, we get $X[B \bullet C] \to_T A$ applying ($\bullet$L) to $X[B \circ C] \to_T A$.

For $n > 1$, we have $X = Y_0 \leq_b Y_1$ and $Y_1 \leq Y_n = Y$. By induction hypothesis, $Y_1 \to_T A$. Applying the same argument as for $n = 1$ we get $X \to_T A$. ∎

Corollary 3.2. *$X \leq A$ (resp. $X \leq \Lambda$) iff $X \to_T A$ (resp. $X \to_T \mathbf{1}$).*

Proof. The 'if' direction is an immediate consequence of the definition of $\leq$. To prove the 'only if' direction, assume $X \leq A$ (resp. $X \leq \Lambda$). Using (Id) (resp. ($\mathbf{1}$R)) and Fact 3.1, we get $X \to_T A$ (resp. $X \to_T \mathbf{1}$). ∎

Next, we take into consideration the residuated groupoid of cones $\mathcal{C}(M) = (C(M), \subseteq, \cdot, \backslash, /, I)$ over $(M, \leq, \circ, \Lambda)$. An assignment μ on $\mathcal{C}(M)$ is defined by setting:

$$\mu(p) = \{X \in M : X \to_T p\},$$

for all atoms p. By Fact 3.1, $\mu(p)$ is a cone.

Fact 3.3. *For all $A \in T$, $\mu(A) = \{X \in M : X \to_T A\}$.*

Proof. We proceed by induction on A. For $A = \mathbf{1}$, $\mu(\mathbf{1}) = \{X \in M : X \leq \mathbf{1}\} = \{X \in M : X \to_T \mathbf{1}\}$, by Corollary 3.2. The further proof is analogous to that of Lemma 1 in [3]; we recall it for the sake of completeness.

Let $A = B \bullet C$. Then $\mu(A) = \mu(B \bullet C) = \mu(B) \circ \mu(C)$. Assume $X \in \mu(A)$. There exist $Y \in \mu(B)$, $Z \in \mu(C)$ such that $X \leq Y \circ Z$. By the induction hypothesis, $Y \to_T B$, $Z \to_T C$, whence $Y \circ Z \to_T B \bullet C$, by ($\bullet$R). Using Fact 3.1, we have $X \to_T A$. Now, assume $X \to_T A$. By Corollary 3.2, $X \leq A$. By the induction hypothesis and (Id), $B \in \mu(B)$, $C \in \mu(C)$. Hence, $B \circ C \in \mu(A)$. Using the definition of $\leq$, we get $A \leq B \circ C$, hence $X \leq B \circ C$ which yields $X \in \mu(A)$.

Let $A = B/C$. Then $\mu(A) = \mu(B)/\mu(C)$. Assume $X \in \mu(A)$. By the induction hypothesis and (Id), we get $C \in \mu(C)$. Hence $X \circ C \in \mu(B)$. Using the induction hypothesis once more, we have $X \circ C \to_T B$. Thus $X \to_T A$, by (/ R). Now, assume $X \to_T A$. Let $Y \in \mu(C)$. By the induction hypothesis, $Y \to_T C$. Using (CUT) to this sequent and $(B/C) \circ C \to_T B$, we get $(B/C) \circ Y \to_T B$, which means $A \circ Y \to_T B$. By (CUT) again, we have $X \circ Y \to_T B$. By the induction hypothesis, $X \circ Y \in \mu(B)$, which yields $X \in \mu(A)$. The case $A = B \backslash C$ is treated in a similar way.

∎

Fact 3.4. *Every sequent provable in* NL1(Γ) *is true in* $(\mathcal{C}(M), \mu)$.

Proof. It suffices to show that each assumption from Γ is true in $(\mathcal{C}(M), \mu)$. Assume that $A \rightarrow B$ belongs to Γ. It yields $A \rightarrow_T B$. We need to show that $\mu(A) \subseteq \mu(B)$. Let $X \in \mu(A)$. Then, $X \rightarrow_T A$. By (CUT), we get $X \rightarrow_T B$, which yields $X \in \mu(B)$. ■

Lemma 3.5. *The system* NL1(Γ) *possesses the extended subformula property.*

Proof. Let $X \rightarrow A$ be a T-sequent provable in NL1(Γ). By Fact 3.4, it is true in the model $(\mathcal{C}(M), \mu)$, i.e. $\mu(X) \subseteq \mu(A)$. Since $X \in \mu(X)$, we have $X \in \mu(A)$. But it yields $X \rightarrow_T A$. ■

Hereafter we assume that Γ and T are finite. A sequent is said to be *basic* if it is a T-sequent of the form $\Lambda \rightarrow A$, $A \rightarrow B$, $A \circ B \rightarrow C$. We describe an effective procedure which produces all basic sequents derivable in NL1(Γ).

Let S_0 consist of $\Lambda \rightarrow \mathbf{1}$, all T-sequents of the form (Id), all sequents from Γ, and all T-sequents of the form:

$$\mathbf{1} \circ A \rightarrow A, \quad A \circ \mathbf{1} \rightarrow A, \quad A \circ B \rightarrow A \bullet B,$$
$$A \circ (A \backslash B) \rightarrow B, \quad (A/B) \circ B \rightarrow A.$$

Assume S_n has already been defined. S_{n+1} is S_n enriched with all sequents resulting from the following rules:

(S1) if $(A \circ B \rightarrow C) \in S_n$ and $(A \bullet B) \in T$, then $(A \bullet B \rightarrow C) \in S_{n+1}$,

(S2) if $(A \circ X \rightarrow C) \in S_n$ and $(A \backslash C) \in T$, then $(X \rightarrow A \backslash C) \in S_{n+1}$,

(S3) if $(X \circ B \rightarrow C) \in S_n$ and $(C/B) \in T$, then $(X \rightarrow C/B) \in S_{n+1}$,

(S4) if $(\Lambda \rightarrow A) \in S_n$ and $(A \circ X \rightarrow C) \in S_n$, then $(X \rightarrow C) \in S_{n+1}$,

(S5) if $(\Lambda \rightarrow A) \in S_n$ and $(X \circ A \rightarrow C) \in S_n$, then $(X \rightarrow C) \in S_{n+1}$,

(S6) if $(A \rightarrow B) \in S_n$ and $(B \circ X \rightarrow C) \in S_n$, then $(A \circ X \rightarrow C) \in S_{n+1}$,

(S7) if $(A \rightarrow B) \in S_n$ and $(X \circ B \rightarrow C) \in S_n$, then $(X \circ A \rightarrow C) \in S_{n+1}$,

(S8) if $(A \circ B \rightarrow C) \in S_n$ and $(C \rightarrow D) \in S_n$, then $(A \circ B \rightarrow D) \in S_{n+1}$.

Clearly, $S_n \subseteq S_{n+1}$ and all sequents in S_n are basic for all $n \geq 0$. We define S^T as the join of this chain. S^T is a set of basic sequents, hence it must be finite. It yields $S^T = S_k$, for the least k such that $S_k = S_{k+1}$, and this k is not greater than the number of basic sequents. Of course, S^T is closed under the rules (S1)-(S8). The rules (S1), (S2), (S3) are in fact ($\bullet$L), ($\backslash$R), (/R) restricted to the basic sequents, and (S4)-(S8) describe the use of (CUT) with the same restriction.

FACT 3.6. *The set S^T can be constructed in polynomial time.*

PROOF. Let n be the cardinality of T. There are n, n^2 and n^3 basic sequents of the form $\Lambda \to A$, $A \to B$ and $A \circ B \to C$, respectively. Hence, we have $m = n^3 + n^2 + n$ basic sequents. The set S_0 can be constructed in time $O(n^2)$. To get S_{i+1} from S_i we must close S_i under the rules (S1)-(S8). For the rules (S1)-(S3) it can be done in at most $m^2 \cdot n$ steps for each rule. For example, to close S_i under (S1) we must check if $(A \circ B \to C) \in S_i$ and $(A \bullet B) \in T$ which needs at most m and n steps, respectively. The sequent $A \bullet B \to C$ is added to S_{i+1} only if it does not belong to this set. To check this fact the next m steps are needed. To close S_i under the rules (S4)-(S8) at most m^3 steps are needed. Hence, we can get S_{i+1} from S_i in time $O(m^3)$ The least k such that $S^T = S_k$ is not greater than m. So, we can construct S^T from T in time $O(m^4) = O(n^{12})$. ■

We take into consideration the system whose axioms are all sequents from S^T and whose only inference rule is (CUT). This system we denote by $S(T)$. It is clear that every proof in $S(T)$ consists of T-sequents only.

If in a proof in $S(T)$ of some sequent $X \to A$ only sequents without empty antecedents are used as premises of (CUT), the length of all sequents in this proof is not greater than the length of $X \to A$. But it does not hold if we allow in (CUT) the premises of the form $\Lambda \to A$. Therefore we introduce another system $S(T)^-$ whose axioms are all sequents from S^T and whose only inference rule is (CUT) with premises without empty antecedents. This restricted (CUT) is denoted by (CUT$^+$).

LEMMA 3.7. *For any sequent $X \to A$, $S(T) \vdash X \to A$ iff $S(T)^- \vdash X \to A$.*

PROOF. The 'if' direction is evident. To prove the 'only if' direction we show that $S(T)^-$ is closed under (CUT), i.e.

(*) If $S(T)^- \vdash X \to B$ and $S(T)^- \vdash Y[B] \to A$, then $S(T)^- \vdash Y[X] \to A$.

Assume $S(T)^- \vdash X \to B$ and $S(T)^- \vdash Y[B] \to A$. If $X \neq \Lambda$, then $S(T)^- \vdash Y[X] \to A$ by the definition of $S(T)^-$. Assume $X = \Lambda$. Then the sequent $X \to B$ is of the form $\Lambda \to B$ and $S(T)^- \vdash \Lambda \to B$, which means that $\Lambda \to B$ is an axiom of $S(T)^-$. To prove (*) we proceed by induction on derivations of the second premise: $Y[B] \to A$. If $Y[B] \to A$ is an axiom of $S(T)^-$, then $(Y[B] \to A) \in S^T$. By (S4) or (S5), $(Y[\Lambda] \to A) \in S^T$ which yields $S(T)^- \vdash Y[\Lambda] \to A$. Assume that $Y[B] \to A$ is a conclusion of (CUT$^+$). Then, $Y[B] = Z[Y']$ and, for some $C \in T$, $S(T)^- \vdash Y' \to C$ and $S(T)^- \vdash Z[C] \to A$. We consider the following cases.

(1) B is contained in Y'. Then $Y' = Y'[B]$. Let $Y'[B] \neq B$. By the induction hypothesis, (*) holds for $\Lambda \to B$ and $Y'[B] \to C$, so $S(T)^- \vdash Y'[\Lambda] \to C$. Since $Y'[B] \neq B$, we have $Y'[\Lambda] \neq \Lambda$. Using (CUT$^+$), we get $S(T)^- \vdash Z[Y'[\Lambda]] \to A$, which means $S(T)^- \vdash Y[\Lambda] \to A$. Now, let $Y'[B] = B$. By the induction hypothesis, (*) holds for $\Lambda \to B$ and $B \to C$, so $S(T)^- \vdash \Lambda \to C$. Using the induction hypothesis to $\Lambda \to C$ and $Z[C] \to A$, we get $S(T)^- \vdash Z[\Lambda] \to A$, which means $S(T)^- \vdash Y[\Lambda] \to A$.

(2) B and Y' do not overlap. Then B is contained in Z and does not overlap C in Z. We write $Z[C] = Z[B, C]$. By the assumption, we have $Y' \neq \Lambda$. By the induction hypothesis, (*) holds for $\Lambda \to B$ and $Z[B, C] \to A$, so $S(T)^- \vdash Z[\Lambda, C] \to A$. By (CUT$^+$), $S(T)^- \vdash Z[\Lambda, Y'] \to A$, which means $S(T)^- \vdash Y[\Lambda] \to A$.

■

Corollary 3.8. *Every basic sequent provable in $S(T)$ belongs to S^T.*

Proof. We proceed by induction on proofs in $S(T)^-$. Assume that $X \to A$ is a basic sequent derivable in $S(T)^-$. If $X \to A$ is an axiom of $S(T)^-$, then $(X \to A) \in S^T$. If $X \to A$ is a conclusion of (CUT$^+$) then $X \neq \Lambda$. We consider two cases.

(1) $X = B$. There exists a proof such that $B \to A$ is a conclusion from premises $B \to C$ and $C \to A$. Since proofs in $S(T)^-$ consist of T-sequents only, $B \to C$ and $C \to A$ are both basic sequents. By the induction hypothesis, $(B \to C) \in S^T$ and $(C \to A) \in S^T$. By (S6) or (S7), $(B \to A) \in S^T$.

(2) $X = B \circ C$. There exists a proof such that $B \circ C \to A$ is a conclusion of (CUT$^+$) with premises of the form: $(B \circ C \to D,\ D \to A)$ or $(B \to D,\ D \circ C \to A)$ or $(C \to D,\ B \circ D \to A)$. By the induction hypothesis these sequents belong to S^T. By (S6), (S7), (S8) , in each case, we get $(B \circ C \to A) \in S^T$.

■

Now, we formulate an interpolation lemma for $S(T)$.

Lemma 3.9. *(i) If $S(T) \vdash X[Y] \to A$, $Y \neq \Lambda$, then there exists $D \in T$ such that $S(T) \vdash Y \to D$ and $S(T) \vdash X[D] \to A$. (ii) If $S(T) \vdash X[\Lambda] \to A$, then $S(T) \vdash X[\mathbf{1}] \to A$; it means that $D = \mathbf{1}$ for $Y = \Lambda$.*

Proof. The proof of (i) is analogous to that of Lemma 2 in [3]. Assume $S(T) \vdash X[Y] \to A$ and $Y \neq \Lambda$. We proceed by induction on proofs in $S(T)$.

Assume, that $X[Y] \to A$ is an axiom of $S(T)$. If $Y = X$, then $(X[Y] \to A) = (X \to A) = (Y \to A)$. We set $D = A$. Then $(Y \to D) = (Y \to A)$ and $(X[D] \to A) = (A \to A)$. By the assumption of the lemma, we have $S(T) \vdash Y \to A$ and $S(T) \vdash A \to A$, since $(A \to A) \in S^T$. If $Y \neq X$, then $X[Y] = B \circ C$ and $Y = B$ or $Y = C$. Hence $D = B$ or $D = C$, respectively.

Assume, that $X[Y] \to A$ is a conclusion of (CUT$^+$). Then $X[Y] = Z[Y']$, $Y' \neq \Lambda$, and for some $B \in T$, $S(T) \vdash Y' \to B$ and $S(T) \vdash Z[B] \to A$. The following cases are considered.

(1) Y is contained in Y'. Then $Y' = Y'[Y]$. By the induction hypothesis, there exists $D \in T$ such that $S(T) \vdash Y \to D$ and $S(T) \vdash Y'[D] \to B$. Using (CUT) with the premises $Z[B] \to A$ and $Y'[D] \to B$ we get $S(T) \vdash Z[Y'[D]] \to A$, which means $S(T) \vdash X[D] \to A$.

(2) Y' is contained in Y. Then $X[Y] = X[Y[Y']] = Z[Y']$ and $Z[B] = X[Y[B]]$. By the induction hypothesis, there exists $D \in T$ such that $S(T) \vdash Y[B] \to D$ and $S(T) \vdash X[D] \to A$. Using (CUT) with the premises $Y' \to B$ and $Y[B] \to D$ we get $S(T) \vdash Y[Y'] \to D$.

(3) Y and Y' do not overlap. Then Y is contained in $Z[B]$ and does not overlap B in $Z[B]$. We write $Z[B] = Z[B, Y]$. By the induction hypothesis, there exists $D \in T$ such that $S(T) \vdash Y \to D$ and $S(T) \vdash Z[B, D] \to A$. Using (CUT) with the premises $Y' \to B$ and $Z[B, D] \to B$ we get $S(T) \vdash Z[Y', D] \to A$, which means $S(T) \vdash X[D] \to A$.

Now, we prove (ii). Assume $S(T) \vdash X[\Lambda] \to A$. We consider three cases.

(1) $X[\Lambda] = \Lambda$. Then $\Lambda \to A$ is a basic sequent derivable in $S(T)$. By Corollary 3.8, $(\Lambda \to A) \in S^T$. By the construction of S^T, $(A \circ \mathbf{1} \to A) \in S^T$. Applying (S4) we have $(\mathbf{1} \to A) \in S^T$, which means $S(T) \vdash X[\mathbf{1}] \to A$.

(2) $X[\Lambda] = X[\Lambda \circ Z]$, $Z \neq \Lambda$. By (i), there exists $D \in T$ such that $S(T) \vdash Z \to D$ and $S(T) \vdash X[\Lambda \circ D] \to A$. Since $D \in T$, $(\mathbf{1} \circ D \to D) \in S^T$. By two applications of (CUT), using $\Lambda \circ D = D$, we get $S(T) \vdash X[\mathbf{1} \circ Z] \to A$.

(3) $X[\Lambda] = X[Z \circ \Lambda]$, $Z \neq \Lambda$. Similar to the previous case.

■

Lemma 3.10. *For any T-sequent $X \to A$, $X \to_T A$ iff $S(T) \vdash X \to A$.*

Proof. Recall that $X \to_T A$ means that $X \to A$ has a proof in NL1(Γ) consisting of T-sequents only. The 'if' direction is obvious, since for all sequents $(X \to A) \in S^T$, we have $X \to_T A$, by the construction of S^T.

The T-sequents which are axioms of NL1(Γ) belong to S_0. Thus, to prove the 'only if' direction it suffices to show that all inference rules of NL1(Γ), restricted to T-sequents, are admissible in $S(T)$. This is obvious for (CUT). Let us consider (**1**L). By Lemma 3.9, (**1**L) is admissible in $S(T)$.

Other rules are treated as in [3]. Let us consider (/L). Assume $S(T) \vdash X[A] \to C$, $S(T) \vdash Y \to B$ and $(A/B) \in T$. We have $((A/B) \circ B \to A) \in S^T$, so it is an axiom of $S(T)$. By two applications of (CUT) we get $S(T) \vdash X[(A/B) \circ Y] \to C$. Let us consider (/R). Assume $S(T) \vdash X \circ B \to A$ and $(A/B) \in T$. By Lemma 3.9, there exists $D \in T$ such that $S(T) \vdash X \to D$ and $S(T) \vdash D \circ B \to A$. Since $D \circ B \to A$ is basic, then it belongs to S^T (Corollary 3.8). By (S3) we have $(D \to A/B) \in S^T$. So, $S(T) \vdash X \to A/B$, by (CUT). For rules (\L) and (\R) the argument is dual. Let us consider ($\bullet$L). Assume $S(T) \vdash X[A \circ B] \to C$ and $(A \bullet B) \in T$. By Lemma 3.9, there exists $D \in T$ such that $S(T) \vdash A \circ B \to D$ and $S(T) \vdash X[D] \to C$. $(A \circ B \to D) \in S^T$ as a basic sequent provable in $S(T)$, and consequently $(A \bullet B \to D) \in S^T$, by (S1). Using (CUT) we get $S(T) \vdash X \to C$. Let us consider ($\bullet$R). Assume $S(T) \vdash X \to A$, $S(T) \vdash Y \to AB$ and $(A \bullet B) \in T$. We have $(A \circ B \to A \bullet B) \in S_0$. So, $S(T) \vdash X \circ Y \to A \bullet B$, by two applications of (CUT). ■

THEOREM 3.11. NL1(Γ) *is decidable in polynomial time.*

PROOF. Let Γ be a finite set of sequents of the form $C \to D$. It suffices to show the thesis for sequents $B \to A$.

Let n be the number of logical constants and atoms occurring in $B \to A$ and Γ. As T we take the set consisting of **1**, all subformulas of formulas appearing in $B \to A$ and Γ. Since the number of subformulas of any formula is equal to the number of logical constants and atoms in it, T has at most $n+1$ elements and we can construct it in time $O(n^2)$. By Lemma 3.5, NL1(Γ) $\vdash B \to A$ iff $B \to_T A$. By Lemma 3.10, $B \to_T A$ iff $S(T) \vdash B \to A$. Since $B \to A$ is basic, then $S(T) \vdash B \to A$ iff $(B \to A) \in S^T$. Consequently, NL1(Γ) $\vdash B \to A$ iff $(B \to A) \in S^T$. By the proof of Fact 3.6, the size of S^T is at most $O(n^3)$ and S^T can be constructed in time $O(n^{12})$. Hence, the total time of deciding if NL1(Γ) $\vdash B \to A$ is $O(n^{12})$. ■

For any type $A \in$ Tp1, a *Lambek categorial grammar* (based on the system NL1(Γ)) with distinguished type A is a quadruple $G = (V_G, \text{NL1}(\Gamma), A, f)$, where V_G is a nonempty finite lexicon (alphabet), f is a mapping such that for all $v \in V_G$, $f(v) \subset$ Tp1 and $f(v)$ is a finite set. For a formula structure X, by $s(X)$ we denote the string of formulas which arises from X

by dropping all occurrences of $\circ$ and the corresponding parentheses. The language $L(G, A)$, generated by G, is defined as the set of all nonempty strings $v = v_1 \dots v_n$ over the alphabet V_G such that there exists a sequent $X \to A$ derivable in NL1(Γ), such that $s(X) = A_1 \dots A_n$ and $A_i \in f(v_i)$ for all $1 \leq i \leq n$.

A *context-free grammar* is a quadruple $G = (N_G, V_G, S_G, R_G)$ such that N_G and V_G are disjoint, nonempty and finite sets, $S_G \in N_G$, and R_G is a finite set of expressions of the form $A \mapsto v$, where $A \in N_G$, $v \in (N_G \cup V_G)^*$. We refer to N_G, V_G, S_G and R_G as the set of nonterminal symbols, the set of terminal symbols, the initial symbol and the set of production rules of G respectively. We say that a string w is *directly derivable* from a string v in G (write $v \Rightarrow_G w$) if there exist strings r, s, t and rule $A \mapsto s$ in R_G such that $v = rAt$, $w = rst$. We say that a string w is *derivable* from a string v in G (write $v \Rightarrow_G^* w$) if there exists a sequence $(v_0, \dots, v_n)$ such that $n \geq 0$, $v_0 = v$, $v_n = w$ and $v_{i-1} \Rightarrow_G v_i$, for all $i = 1, \dots, n$. The language $L(G)$ generated by G is defined as the set of all strings $v \in V_G^*$ such that $S_G \Rightarrow_G^* v$.

THEOREM 3.12. *If G_1 is a Lambek categorial grammar based on the system* NL1(Γ)*, then for any $A \in$* Tp1 *there is a context-free grammar G_2, such that $L(G_1, A) = L(G_2)$ and the transformation of G_1 into G_2 can be done in polynomial time with respect to the size of G_1.*

PROOF. Let Γ be a finite set of sequents of the form $B \to C$ and $A \in$ Tp1. Fix a Lambek grammar G_1 based on the system NL1(Γ). Let Tp1(G_1) be the set of types B, for which there is $v \in V_{G_1}$, such that $B \in f(v)$. The set Tp1(G_1) is finite. Let T be the set of all subformulas appearing in A, in formulas from the set Tp1(G_1) and in formulas appearing in Γ. We also assume $\mathbf{1} \in T$. The context-free grammar $G_2 = (N_{G_2}, V_{G_2}, S_{G_2}, R_{G_2})$ is defined as follows:

$$
\begin{aligned}
V_{G_2} &= V_{G_1} = V, \quad N_{G_2} = T, \quad S_{G_2} = A, \\
R_{G_2} &= \{B \mapsto v : \ v \in V_{G_2},\ B \in f(v)\} \cup \\
&\cup \ \{B \mapsto C : \ B, C \in N_{G_2},\ \mathrm{NL1}(\Gamma) \vdash C \to B\} \cup \\
&\cup \ \{B \mapsto CD : \ B, C, D \in N_{G_2},\ \mathrm{NL1}(\Gamma) \vdash C \circ D \to B\}.
\end{aligned}
$$

To establish that the languages generated by the grammars G_1 and G_2 coincide, first we prove that $L(G_1, A) \subset L(G_2)$. Suppose $v_1 \dots v_n \in L(G_1, A)$. According to the definition of $L(G_1, A)$ there is a formula structure X such that NL1(Γ) $\vdash X \to A$, $s(X) = A_1 \dots A_n$ and $A_i \in f(v_i)$, for all $1 \leqslant i \leqslant n$. By the construction of G_2, $(A_i \mapsto v_i) \in R_{G_2}$, for all $1 \leqslant i \leqslant n$, i.e.

$A_1 \ldots A_n \Rightarrow^*_{G_2} v_1 \ldots v_n$. Thus it suffices to prove that $A \Rightarrow^*_{G_2} A_1 \ldots A_n$. By the definition of T, $X \to A$ is a T-sequent. Using the extended subformula property we have $X \to_T A$. In view of Lemma 3.7 and Lemma 3.10, $X \to A$ is $S(T)^-$-derivable. An induction on the length of a $S(T)^-$-derivations shows that if $S(T^-) \vdash X \to A$ then $A \Rightarrow^*_{G_2} A_1 \ldots A_n$. Assume that $X \to A$ is an axiom of $S(T)^-$. If $X = A_1$, $A_1 \in T$, then $(A \mapsto A_1) \in R_{G_2}$, which means $A \Rightarrow^*_{G_2} A_1$. Using a similar argument for $X = A_1 \circ A_2$, $A_1, A_2 \in T$ we get $A \Rightarrow^*_{G_2} A_1 A_2$. Now, assume that $X \to A$ is a conclusion of (CUT$^+$). Then $X = X[Y]$, where $s(Y) = A_i A_{i+1} \ldots A_k$, $1 \leq i \leq k \leq n$, and (CUT$^+$) is applied to $Y \to B$ and $X[B] \to A$, where $B \in T$ and $s(X[B]) = A_1 \ldots A_{i-1} B A_{k+1} \ldots A_n$. By the induction hypothesis, $A \Rightarrow^*_{G_2} A_1 \ldots A_{i-1} B A_{k+1} \ldots A_n$ and $B \Rightarrow^*_{G_2} A_i \ldots A_k$. Hence $A \Rightarrow^*_{G_2} A_1 \ldots A_n$.

The inclusion $L(G_2) \subset L(G_1, A)$ is easy. Every derivation of $A \Rightarrow^*_{G_2} A_1 \ldots A_n$ can be treated as a derivation of $X \to A$ in $S(T)^-$, for some X such that $s(X) = A_1 \ldots A_n$.

It remains to prove that the transformation of G_1 into G_2 can be done in polynomial time with respect to the size of G_1. Let n be the number of logical constants and atoms in A, Γ and $\mathrm{Tp1}(G_1)$. Then the set $N_{G_2} = T$, defined above, has at most $n+1$ elements and can be constructed in time $O(n^2)$. The number of production rules of the form $B \to v$ equals the cardinal number of the set $\mathrm{Tp1}(G_1)$ and do not exceed n. The remaining production rules are the reversed sequents from the set S^T. By the proof of Theorem 3.11, the construction of this set can be done in time $O(n^{12})$. Hence the total time of construction of the context-free grammar G_2 is $O(n^{12})$.

■

References

[1] Bulińska, M., 'The Pentus Theorem for Lambek Calculus with Simple Nonlogical Axioms', *Studia Logica*, 81:43–59, 2005.

[2] Buszkowski, W., 'Generative Capacity of Nonassociative Lambek Calculus', *Bulletin of Polish Academy of Sciences. Mathematics*, 34:507–516, 1986.

[3] Buszkowski, W., 'Lambek Calculus with Nonlogical Axioms', in: *Language and Grammar. Studies in Mathematical Linguistics and Natural Language*, (C. Casadio, P. J. Scott, R. A. G. Seely, Eds.),CSLI Publications, 77–93, 2005.

[4] de Groote, P., Lamarche F.: Clasical Non-Associative Lambek Calculus , *Studia Logica*, 71:355–388, 2002.

[5] Jäger, G., 'Residuation, Structural Rules and Context Freeness', *Journal of Logic, Language and Information*, 13:47–59, 2002.

[6] Kandulski, M., 'The equivalence of Nonassociative Lambek Categorial Grammars and Context-Free Grammars', *Zeitschrift für mathematische Logik und Grundlagen der Mathematik*, 34:41–52, 1988.

[7] Lambek, J., 'The mathematics of sentence structure', *The American Mathematical Monthly*, 65:154–170, 1958.

[8] Lambek, J., 'On the calculus of syntactic types', in: *Structure of Language and Its Mathematical Aspects*, (R. Jacobson, Ed.), *Proc. Symp. Appl. Math.*, AMS, Providence, 166–178, 1961.

[9] Lambek, J., 'Type grammars revisited', in: *Logical Aspects of Computational Linguistics*, (A. Lecomte, F. Lamarche, G. Perrier, Eds.), LNAI 1582, Springer, Berlin , 1–27, 1999.

[10] Moortgat, M., 'Categorial Type Logics', in: (J. van Benthem, A. ter Meulen, Eds.), *Handbook of Logic and Language*, Elsevier, Amsterdam, The MIT Press, Cambrigde Mass., 93–177, 1997.

Maria Bulińska
Faculty of Mathematics and Computer Science
University of Warmia and Mazury
Żołnierska 14
Olsztyn, Poland
bulma@uwm.edu.pl