Lambek Calculus with Nonlogical Axioms

Wojciech Buszkowski
Faculty of Mathematics and Computer Science
Adam Mickiewicz University
Poznań Poland

in: Language and Grammar, (C. Casadio et al., eds.), CSLI
Lecture Notes 168:77–93, 2005

Abstract

We study Nonassociative Lambek Calculus and Associative Lambek Calculus enriched with finitely many nonlogical axioms. We prove that the nonassociative systems are decidable in polynomial time and generate context-free languages. In [Buszkowski 1982] it has been shown that finite axiomatic extensions of Associative Lambek Calculus generate all recursively enumerable languages; here we give a new proof of this fact. We also obtain similar results for systems with permutation and $n-$ary operations.

1 Introduction and preliminaries

In [Buszkowski 1982], systems of Associative Lambek Calculus with finitely many nonlogical axioms, not containing product, are shown to be undecidable (in general) and to generate all recursively enumerable languages. In section 3 of the present paper we give a new version of this result. In section 2 we show that for Nonassociative Lambek Calculus the situation is different: all systems of Nonassociative Lambek Calculus with finitely many nonlogical axioms are decidable in polynomial time and generate context-free languages. The same holds for systems with unary modalities, studied in [Moortgat 1995, Moortgat 1997], $n-$ary operations (i.e. for the Generalized Lambek Calculus of [Buszkowski 1989], studied in [Kołowska 1997, Kandulski 1997, Jäger 2002]), and the rule of permutation [Kandulski 1995, Jäger 2002]. These results are new; they do not rely on cut elimination which is not available for systems with nonlogical axioms. Further, our results from section 2 provide a new proof of the context freeness of categorial grammars based on Nonassociative Lambek Calculus, first proven in [Buszkowski 1986a] for the product-free system, then in [Kandulski 1988] for the full system, and recently in [Jäger 2002] by a modification of results of [Roorda 1991] and [Pentus 1993], the same system with permutation [Kandulski 1995, Jäger 2002] and Generalized Lambek Calculus

[Kandulski 1997, Jäger 2002]. The polynomial time decidability of Nonassociative Lambek Calculus with modalities and/or permutation and Generalized Lambek Calculus seems to be new; [de Groote and Lamarche 2002] prove it for (Classical) Nonassociative Lambek Calculus. Recently, [Pentus 2003] proves that Associative Lambek Calculus is NP-complete.

Our interest in adding nonlogical axioms to the Lambek calculus can be motivated in various ways. First, it is an obvious logical thread of theories based on the given logic; in other words, one studies the consequence relation associated with this logic. Second, for the associative case, nonlogical axioms enable one to surpass the limitations of the context-free languages. Third, there are many types of evidence for the usefullness of nonlogical axioms in natural language description. For example, Lambek [Lambek 1999, Lambek 2001, Casadio and Lambek 2002] uses axioms of the form $\pi_i \to \pi$ to express the inclusion of the class of personal pronouns in $i-$th Person in the class of personal pronouns; different kinds of subcategorization can be found in Keenan and Faltz [Keenan and Faltz 1985]. More in the style of Moortgat [Moortgat 1997], one might take Nonassociative Lambek Calculus as the basic logic and, besides modalities, use axioms of the form $(A \cdot B) \cdot C \leftrightarrow A \cdot (B \cdot C)$ or $A \cdot B \leftrightarrow B \cdot A$, for some concrete types A, B, C, to admit associativity or permutation in some special cases. A limited usage of contraction can also be helpful: Lambek Calculus does not lift S\(S/S) (the type of sentence conjunction) to VP\(VP/VP), VP=PN\S (the type of verb phrase conjunction), but one may stipulate S\(S/S)→VP\(VP/VP). Fourth, Lambek logics can also be treated as a machinery of grammar transformation: for instance, a context-free grammar can be transformed into an equivalent basic categorial grammar (see [Buszkowski 1996]).

We describe the formalism of Nonassociative Lambek Calculus (NL). Formulas (also called types) are formed out of (denumerably many) atoms $p.q, r, \ldots$ by means of binary operation symbols $\bullet$ (product), \ (left residuation), / (right residuation); these symbols are called multiplicative conjunction and implications in substructural logics [Restall 2000]. Formula structures are recursively defined as follows: (i) all formulas are (atomic) formula structures, (ii) if X, Y are formula structures, then $(X \circ Y)$ is a formula structure. $X[Y]$ denotes a formula structure X with a distinguished substructure Y, and $X[Z]$ stands for the substitution of Z for Y in X. Sequents are formal expressions $X \to A$ such that X is a formula structure and A is a formula. Axioms and inference rules of NL (in a sequential form) are the following:

$$\text{(Id) } A \to A$$

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

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

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

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

NL is due to Lambek [Lambek 1961] who has proven cut elimination theorem and decidability of this system. It is easy to show that NL is complete with respect to residuated groupoids. It is also complete with respect to powerset structures over groupoids [Kołowska 1997], though not free groupoids [Došen 1992].

We shall consider extensions of NL by structural rules:

$$\text{(ASS)}\ \frac{X[(Y \circ Y') \circ Y''] \to A}{X[Y' \circ (Y' \circ Y'')] \to A}$$

$$\text{(PER)}\ \frac{X[Y \circ Z] \to A}{X[Z \circ Y] \to A}$$

i.e. the rule of associativity and the rule of permutation. (ASS) acts up-down and down-up. Associative Lambek Calculus (L) is NL plus (ASS); it is due to Lambek [Lambek 1958]. NL plus (PER) will be denoted by NLP. L is complete with respect to residuated semigroups, powerset structures over semigroups [Buszkowski 1986] and over free semigroups [Pentus 1995]. Cut elimination and decidability hold for both systems.

They, however, need not hold, if we affix new nonlogical axioms of the form $A \to B$. For a set Φ, of formulas $A \to B$, NL(Φ) denotes the system NL with all formulas from Φ as new axioms, and similarly for L(Φ), NLP(Φ). Caution: we always assume (CUT) to be a rule in these extended systems. This rule is necessary to prove strong completeness: the sequents provable in the extended system are precisely those which are valid in the appropriate frames under all assignments (of elements of the frame for variables) which satisfy the new axioms. Since new axioms are not supposed to be closed under substitution, atoms appearing in them have to be treated as constants rather than variables, but this will not be regarded in notation.

It has been shown in [Buszkowski 1982] that there exist finite Φ such that L(Φ) is undecidable, and the formulas in Φ are of the form p or p/q. In section 2 of the present paper we prove that all systems NL(Φ), with Φ finite, are decidable in polynomial time and generate context-free languages. The same holds for NLP, Generalized Lambek Calculus and NL, NLP with unary modalities. In section 3 we give a new discussion of the situation for L(Φ).

Associative Lambek Calculus with (PER) was studied in [van Benthem 1986, van Benthem 1991] as a logic of semantic types (the system is known as the Lambek-van Benthem calculus). It is decidable, but its complexity and the decidability of its extensions by means of nonlogical axioms remain unknown.

2 NL with nonlogical axioms

Since (CUT) is a rule in NL(Φ), it is not obvious that the system has the subformula property. A slightly generalized form of this property can be shown by a refinement of standard proofs of the completeness theorem. Actually, in

[Buszkowski 1986] the subformula property for L(Φ) has been proven in this way. Here we give a simpler proof, based on more general algebraic models of NL.

A residuated groupoid is a structure $\mathcal{M} = (M, \leq, \cdot, \backslash, /)$ such that $(M, \leq)$ is a poset, $(M, \cdot)$ is a groupoid, and $\backslash, /$ are binary operations on M, satisfying the equivalences:

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

for all $a, b, c \in M$. It is easy to show that every residuated groupoid fulfills:

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

for all elements a, b, c. Also, the following monotonicity laws for residuals hold true:

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

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

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

For any formula structure X, define the formula $F(X)$, by setting: (i) $F(A) = A$, (ii) $F(X \circ Y) = F(X) \bullet F(Y)$. A sequent $X \to A$ is said to be *true* in model $(\mathcal{M}, \mu)$ if $\mu(F(X)) \leq \mu(A)$.

All sequents provable in NL(Φ) are true in all models $(\mathcal{M}, \mu)$ such that all sequents in Φ are true. This can easily be proven by induction on proofs in NL(Φ). The converse is also true: if $X \to A$ is not provable in NL(Φ), then there exists a model $(\mathcal{M}, \mu)$ in which $X \to A$ is not true. This can be proven by a Lindenbaum model.

To get the subformula property, we need more special models. Let $(M, \leq, \cdot)$ be a preordered groupoid, that means, it is a groupoid with a preordering (i.e. a reflexive and transitive relation), satisfying (MON). A set $P \subseteq M$ is called *a cone* on M if $a \leq b$ and $b \in P$ entails $a \in P$. Let $C(M)$ denote the set of cones on M. The operations $\cdot, \backslash, /$ on $C(M)$ are defined as follows:

$$\text{(M1) } P_1P_2 = \{c \in M : (\exists a \in P_1, b \in P_2)\, c \leq ab\}$$

$$\text{(M2) } P_1\backslash P_2 = \{c \in M : (\forall a \in P_1)\, ac \in P_2\}$$

$$\text{(M3) } P_1/P_2 = \{c \in M : (\forall b \in P_2)\, cb \in P_1\}$$

It is easy to see that $(C(M), \subseteq, \cdot, \backslash, /)$ is a residuated groupoid; we call it the residuated groupoid of cones over the given preordered groupoid.

Let T be a set of formulas, closed under subformulas and such that all formulas appearing in Φ 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. We prove the subformula property for NL(Φ).

Lemma 1 *Every $T-$sequent provable in NL(Φ) has a proof in NL(Φ) such that all sequents appearing in this proof are $T-$sequents.*

PROOF. We write $X \rightarrow_T A$ if $X \rightarrow A$ has a proof in NL(Φ), consisting of $T-$sequents only. Let M be the set of all formula structures all of whose atomic substructures belong to T. We define a preordering $\leq$ on M. First, we say that X *directly reduces* to Y if either $Y = Y[A]$ and $X = Y[Z]$, for some Z, A such that $Z \rightarrow_T A$, or $Y = Y[A \circ B]$ and $X = Y[A \bullet B]$, for some A, B such that $(A \bullet B) \in T$. Now, $X \leq Y$ holds iff there exist $Y_0, \ldots, Y_n$, $n \geq 0$, such that $Y = Y_0$, $X = Y_n$ and Y_i directly reduces to Y_{i-1}, for each $i = 1, \ldots, n$. Clearly, $X \leq Y$ entails $Z \circ X \leq Z \circ Y$ and $X \circ Z \leq Y \circ Z$, hence $(M, \leq, \circ)$ is a preordered groupoid. Further, if $Y \rightarrow_T A$ and $X \leq Y$ then $X \rightarrow_T A$.

We consider the residuated groupoid of cones $C(M)$. An assignment μ is defined by setting:

$$\mu(p) = \{X \in M : X \rightarrow_T p\}, \text{ for all atoms } p.$$

Clearly, $\mu(p)$ is a cone. We prove:

$$\mu(A) = \{X \in M : X \rightarrow_T A\}, \text{ for all } A \in T.$$

We proceed by induction on A. For atomic A, this follows from the definition of μ. Let $A = B \bullet C$. Let $X \in \mu(A)$. Then, there exist $Y \in \mu(B)$, $Z \in \mu(C)$ such that $X \leq Y \circ Z$. By the induction hypothesis, $Y \rightarrow_T B$, $Z \rightarrow_T C$, hence $Y \circ Z \rightarrow_T B \bullet C$, by ($\bullet$R). We get $X \rightarrow_T A$. For the other direction, assume $X \rightarrow_T A$. We have $B \in \mu(B)$ and $C \in \mu(C)$, by the induction hypothesis and (Id), which yields $B \circ C \in \mu(A)$. 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 $A = B\backslash C$. Let $X \in \mu(A)$. By the induction hypothesis and (Id), $B \in \mu(B)$, which yields $B \circ X \in \mu(C)$. By the induction hypothesis again, $B \circ X \rightarrow_T C$, and consequently, $X \rightarrow_T A$, by ($\backslash$R). Now, assume $X \rightarrow_T A$. Let $Y \in \mu(B)$. By the induction hypothesis, $Y \rightarrow_T B$, which yields $Y \circ X \rightarrow_T C$, by (CUT) and $B \circ (B\backslash C) \rightarrow_T C$. Then, $Y \circ X \in \mu(C)$, by the induction hypothesis. We have shown $X \in \mu(A)$. For $A = B/C$, the argument is dual. The equality has been proven.

If $A \rightarrow B$ is in Φ, then $\mu(A) \subseteq \mu(B)$: if $X \in \mu(A)$ then $X \rightarrow_T A$, hence $X \rightarrow_T B$, which yields $X \in \mu(B)$. So, $(C(M), \mu)$ satisfies all axioms in Φ. Consequently, all sequents provable in NL(Φ) must be true in this model. Let $X \rightarrow A$ be a $T-$sequent provable in NL(Φ). Since $B \in \mu(B)$, for all $B \in T$, we get $X \in \mu(F(X))$, hence $X \in \mu(A)$. Then, $X \rightarrow_T A$, which finishes the proof. Q.E.D.

Notice that the above proof yields the completeness of NL(Φ) with respect to residuated groupoids of cones over preordered groupoids, hence also with respect to residuated groupoids. If $A \rightarrow B$ is not provable, then, for an appropriate T and the corresponding M, μ, we have $A \in \mu(A)$ and $A \notin \mu(B)$, hence $A \rightarrow B$ is not true in the model.

Let Φ be finite, and let T be a finite set of formulas, closed under subformulas and such that T contains all formulas appearing in Φ. We shall describe an effective procedure which produces all $T-$sequents $A \circ B \to C$ and $A \to B$ which are provable in NL(Φ). Furthermore, we show that every $T-$sequent provable in NL(Φ) can be derived from those $T-$sequents by (CUT) only.

Let S_0 consist of all $T-$sequents of the form (Id), all sequents from Φ and all $T-$sequents of the form:

$$\text{(S0) } A \circ (A\backslash B) \to B,\ (A/B) \circ B \to A,\ A \circ B \to A \bullet B.$$

$T-$sequents of the form $A \circ B \to C$ or $A \to B$ are called *basic sequents*. If S is a set of basic sequents, then S' denotes the closure of S under (CUT), that means, S' is the smallest set of basic sequents containing S and such that, if the premises of (CUT) are in S' and the conclusion is a basic sequent, then the conclusion is in S'. Assume S_n has already been defined. Let S_{n+1} be the smallest set S, of basic sequents, containing S'_n, which is closed under the following rules:

(R1) if $(A \circ B \to C) \in S$ and $(A \bullet B) \in T$ then $(A \bullet B \to C) \in S$,

(R2) if $(A \circ B \to C) \in S$ and $(A\backslash C) \in T$ then $(B \to A\backslash C) \in S$,

(R3) if $(A \circ B \to C) \in S$ and $(C/B) \in T$ then $(A \to C/B) \in S$.

Clearly, $S_n \subseteq S_{n+1}$, for all $n \geq 0$. We define S^T as the union of all S_n, for $n \geq 0$. Evidently, S^T is closed under rules (R1), (R2), (R3). S^T is a set of basic sequents, hence it must be finite. Then, there exists $n \geq 0$ such that $S_n = S_{n+1}$, and $S^T = S_n$, for the least n satisfying $S_n = S_{n+1}$.

Let m denote the cardinality of T. There are $f(m) = m^3 + m^2$ $T-$sequents whose antecedents are of length at most 2, hence the least n such that $S^T = S_n$ must not exceed $f(m)$. The construction of S'_i from S_i can be done in at most $f(m)^2$ steps, and similarly the construction of S_{i+1} from S'_i. Accordingly, we can construct S^T from T in time $O(m^9)$.

By $S(T)$ we denote the system whose axioms are all sequents from S^T and whose only inference rule is (CUT). Then, every proof in $S(T)$ consists of $T-$sequents only. We write $X \to_{S(T)} A$ if $X \to A$ is provable in $S(T)$. Observe that every basic sequent provable in $S(T)$ belongs to S^T (use induction on proofs in $S(T)$). We state an interpolation lemma for $S(T)$.

Lemma 2 *If $X[Y] \to_{S(T)} A$, then there exists $D \in T$ such that $Y \to_{S(T)} D$ and $X[D] \to_{S(T)} A$.*

Proof. We proceed by induction on proofs in $S(T)$. Assume $X[Y] \to A$ belongs to S^T. If $Y = X$, then $D = A$. Otherwise, $X[Y] = B \circ C$, and $Y = B$ or $Y = C$, hence $D = B$ or $D = C$, respectively. Assume $X[Y] \to A$ is the conclusion of (CUT). Then, $X[Y] = Z[Y']$and, for some $B \in T$, $Y' \to_{S(T)} B$ and $Z[B] \to_{S(T)} A$. We consider three cases.

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

(2) Y' is contained in Y. Then, $X[Y] = X[Y[Y']]$ and $Z[B] = X[Y[B]]$. By the induction hypothesis, there exists $D \in T$ such that $Y[B] \rightarrow_{S(T)} D$ and $X[D] \rightarrow_{S(T)} A$. Clearly, $Y \rightarrow_{S(T)} D$, by (CUT).

(3) Y and Y' do not overlap. Then, Y is contained in Z and does not overlap B in Z. We write $Z[B] = Z[B, Y]$. By the induction hypothesis, there exists $D \in T$ such that $Y \rightarrow_{S(T)} D$ and $Z[B, D] \rightarrow_{S(T)} A$. By (CUT), $Z[Y', D] \rightarrow_{S(T)} A$, which means $X[D] \rightarrow_{S(T)} A$. Q.E.D.

Lemma 3 *For any $T-$sequent $X \rightarrow A$, $X \rightarrow_T A$ iff $X \rightarrow_{S(T)} A$.*

PROOF. The 'if' direction is easy. One shows $X \rightarrow_T A$, for all sequents $X \rightarrow A$ in S^T, by a direct inspection of the construction. $S(T)$ uses (CUT) restricted to $T-$sequents.

The $T-$sequents which are axioms of NL(Φ) belong to S_0. Thus, to prove the 'only if' direction it suffices to show that all inference rules of NL(Φ), restricted to $T-$sequents, are admissible in $S(T)$. This is obvious for (CUT). Let us consider (\L). Assume $X[B] \rightarrow_{S(T)} C$, $Y \rightarrow_{S(T)} A$ and $(A\backslash B) \in T$. Since $A \circ (A\backslash B) \rightarrow B$ is in S_0, then $X[Y \circ (A\backslash B)] \rightarrow_{S(T)} C$, by two applications of (CUT). Let us consider (\R). Assume $A \circ X \rightarrow_{S(T)} B$ and $(A\backslash B) \in T$. By lemma 2, there exists $D \in T$ such that $X \rightarrow_{S(T)} D$ and $A \circ D \rightarrow_{S(T)} B$. Since $A \circ D \rightarrow B$ is basic, then it belongs to S^T, and consequently, $D \rightarrow A\backslash B$ also belongs to S^T. So, $X \rightarrow_{S(T)} A\backslash B$, by (CUT). For rules (/L) and (/R), the argument is dual. Let us consider ($\bullet$L). Assume $X[A \circ B] \rightarrow_{S(T)} C$ and $(A \bullet B) \in T$. By lemma 2, there exists $D \in T$ such that $A \circ B \rightarrow_{S(T)} D$ and $X[D] \rightarrow_{S(T)} C$. Again $A \circ B \rightarrow D$ belongs to S^T, hence $A \bullet B \rightarrow D$ belongs to S^T. So, $X[A \bullet B] \rightarrow_{S(T)} C$, by (CUT). Let us consider ($\bullet$R). Assume $X \rightarrow_{S(T)} A$, $Y \rightarrow_{S(T)} B$ and $(A \bullet B) \in T$. Then, $A \circ B \rightarrow A \bullet B$ belongs to S_0, hence $X \circ Y \rightarrow_{S(T)} A \bullet B$, by two applications of (CUT). Q.E.D.

We are ready to prove main results of this section.

Theorem 1 *If Φ is finite, then NL(Φ) is decidable in polynomial time.*

PROOF. Fix a finite set Φ. Let $X \rightarrow A$ be a sequent. Let n be the number of logical constants and atoms in $X \rightarrow A$. It is easy to show that the number of subformulas of a formula B equals the number of logical constants and atoms in B. Let T be the set of all subformulas of formulas appearing in $X \rightarrow A$ and formulas appearing in Φ. We can construct T in time $O(n^2)$, and T has $O(n^2)$ elements. By lemma 1, $X \rightarrow A$ is provable in NL(Φ) iff $X \rightarrow_T A$. By lemma 3, $X \rightarrow_T A$ iff $X \rightarrow_{S(T)} A$. Since proofs in $S(T)$ are actually derivation trees of a context-free grammar whose production rules are the reversed sequents from S^T, then $X \rightarrow_{S(T)} A$ can be checked in time pn^3, where p is the size of S^T. S^T can be constructed in $O(n^{18})$ steps, and the size of S^T is at most $O(n^6)$. Accordingly, the total time equals $O(n^{18})$. The same computation can be done for variable Φ with n being the number of logical constants and atoms in $X \rightarrow A$ and Φ. Q.E.D.

A categorial grammar, based on a system S, can be defined as a finite set of assignments $a \rightarrow A$ such that $a \in \Sigma$, Σ is an alphabet, and A is a formula. 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. For a categorial grammar G and a formula A, *the language* $L(G, A)$ consists of all strings $a_1 \ldots a_n$, $n \geq 1$, satisfying the following condition: there exist formulas A_i, $i = 1, \ldots, n$, and a formula structure X such that $s(X) = A_1 \ldots A_n$, all $a_i \to A_i$ belong to G, and $X \to A$ is provable in S.

Theorem 2 *If G is a categorial grammar, based on NL(Φ), for a finite Φ, then, for any formula A, $L(G, A)$ is a context-free language.*

PROOF. We construct a context-free grammar for $L(G, A)$ as in the proof of theorem 1. Now, T is the set of all subformulas of A and all subformulas of formulas appearing in G. We add lexical production rules $A \to a$ for $a \to A$ belonging to G. Q.E.D.

Theorem 1 generalizes the result of [de Groote and Lamarche 2002] who prove polynomial time complexity of the decision problem for NL. Theorem 2 generalizes the results of [Buszkowski 1986a, Kandulski 1988] that NL generates context-free languages. It, however, should be noticed that the latter results were stronger: they established the equivalence of categorial grammars based on NL and basic categorial grammars in the scope of phrase-structure languages, which could not directly be obtained by the above method. On the other hand, the methods of [Buszkowski 1986a, Kandulski 1988], using a normalization procedure for derivations of unary sequents in NL, are not extendible to axiomatic extensions of NL.

Since $S(T)$ is equivalent to NL(Φ) for $T-$sequents, then lemma 2 yields an interpolation lemma for NL(Φ): if $X[Y] \to A$ is derivable in NL(Φ), then there exists $D \in T$ such that both $Y \to D$ and $X[D] \to A$ are derivable in NL(Φ), where T is the smallest set of formulas, containing all formulas appearing in Φ and $X \to A$ and being closed under subformulas. For $\Phi = \emptyset$, this is essentially the interpolation lemma proven in Jäger [Jäger 2002] who uses a standard induction on cut-free derivations in NL. This author observes that theorem 2 for NL follows from this lemma: if $X \to A$ is derivable in NL, then it can be derived from the basic sequents (in the sense defined above), provable in NL, by means of (CUT), which can easily be simulated by a context-free grammar. This grammar is effectively constructed, since the decision procedure for NL, based on cut elimination, yields all basic sequents, provable in NL (of course, cut elimination also guarantees the subformula property).

For the case of NL(Φ), cut elimination is not possible, hence we have proven the subformula property in a different way (lemma 1). Now, with the subformula property already proven, we could prove the interpolation lemma for NL(Φ) by induction on proofs in NL(Φ) with (CUT). Then, as in [Jäger 2002], theorem 2 follows from the interpolation lemma, but the resulting context-free grammar is not effectively constructed, since we have not provided any decision procedure for NL(Φ). Therefore, we have gone a completely different way: we explicitly construct all basic sequents (for a fixed T), provable in NL(Φ), prove the interpolation lemma for the auxiliary system $S(T)$, and prove the equivalence of

$S(T)$ and NL(Φ) for $T-$systems. This yields both goals: a polynomial time decision procedure for NL(Φ) and the equivalence with context-free grammars.

Theorems 1 and 2 can also be proven for systems NLP(Φ), with Φ finite. We outline the proofs.

A residuated groupoid $(M, \leq, \cdot, \backslash, /)$ is said to be *commutative* , if $ab = ba$, for all $a, b \in M$ (then, $a\backslash b = b/a$, for all $a, b \in M$, so one may abandon one of $\backslash, /$). Clearly, all sequents provable in NLP(Φ) are true in all models $(\mathcal{M}, \mu)$ such that $\mathcal{M}$ is a commutative residuated groupoid and all sequents from Φ are true in the model.

We modify the proof of lemma 1. We add a new disjunct to the definition of direct reducibility: $Y = Z[Y', Y'']$ and $X = Z[Y'', Y']$, for some Z, Y', Y''. Then, $C(M)$ is a commutative residuated groupoid. Let $P_1, P_2 \in C(M)$. Let $X \in P_1P_2$. By (M1), $X \leq Y \circ Z$, for some $Y \in P_1$, $Z \in P_2$. But $Y \circ Z \leq Z \circ Y$, hence $X \in P_2P_1$, which yields $P_1P_2 \subseteq P_2P_1$. The converse inclusion can be shown in a similar way. The remainder of the proof goes without change.

The construction of sets S_n needs one change: S' is the smallest set of basic sequents which contains S and is closed under (CUT) and (PER). The proof of lemma 2 needs no change. In the proof of lemma 3, we must show that $S(T)$ is closed under (PER). Assume $X[Y \circ Z] \rightarrow_{S(T)} A$. By lemma 2, there exist $D', D'' \in T$ such that $Y \rightarrow_{S(T)} D'$, $Z \rightarrow_{S(T)} D''$ and $X[D' \circ D''] \rightarrow_{S(T)} A$. By lemma 2 again, there exists $D \in T$ such that $D' \circ D'' \rightarrow_{S(T)} D$ and $X[D] \rightarrow_{S(T)} A$. Since $D' \circ D'' \rightarrow D$ belongs to S^T, then also $D'' \circ D' \rightarrow D$ belongs to S^T, and consequently, $X[Z \circ Y] \rightarrow_{S(T)} A$, by three applications of (CUT).

Generalized Lambek Calculus (GLC) admits a finite number of product symbols f, of arity $n(f) \geq 1$, each of them giving rise to residuation symbols f/i, for $i = 1, \ldots, n(f)$. Formulas are atoms and complex formulas $f(A_1, \ldots, A_{n(f)})$, $(f/i)(A_1, \ldots, A_{n(f)})$. Formula structures are formed out of formulas (atomic structures) by means of structure operations $\circ_f$, of arity $n(f)$, for every product symbol f. The axioms of GLC are (Id), for every formula A, and the inference rules are the following:

$$(f\mathrm{L})\ \frac{X[\circ_f(A_1, \ldots, A_{n(f)})] \rightarrow B}{X[f(A_1, \ldots, A_{n(f)})] \rightarrow B}$$

$$(f\mathrm{R})\ \frac{X_1 \rightarrow A_1; \ldots; X_{n(f)} \rightarrow A_{n(f)}}{\circ_f(X_1, \ldots, X_{n(f)}) \rightarrow f(A_1, \ldots, A_{n(f)})}$$

$$(f/i\mathrm{L})\ \frac{X[A_i] \rightarrow B;\ Y_1 \rightarrow A_1, \ldots; Y_{n(f)} \rightarrow A_{(n(f)}}{X[\circ_f(Y_1, \ldots, (f/i)(A_1, \ldots, A_{n(f)}), \ldots, Y_{n(f)})] \rightarrow B}$$

$$(f/i\mathrm{R})\ \frac{\circ_f(A_1, \ldots, X, \ldots, A_{n(f)}) \rightarrow A_i}{X \rightarrow (f/i)(A_1, \ldots, A_{n(f)})}$$

and (CUT). In rule $(f/i\mathrm{L})$, premise $Y_i \rightarrow A_i$ is dropped, and in the conclusion the formula $(f/i)(A_1, \ldots, A_{n(f)})$ takes the place of the $i-$th argument of $\circ_f$. In rule $(f/i\mathrm{R})$, X takes the place of the $i-$th argument of $\circ_f$ in the premise.

Actually, the cited papers regard the case $n(f) \geq 2$ only, but all results can easily be generalized to the case $n(f) \geq 1$. For $n(f) = 1$, f and $f/1$ are precisely the minimal modalities diamond and box, studied in [Moortgat 1997, Restall 2000]. Clearly, NL is GLC with exactly one product symbol f, $n(f) = 2$.

Frames for GLC are residuated algebras [Buszkowski 1989]. *A residuated algebra* is a structure $\mathcal{M} = (M, \leq, F)$ such that $(M, \leq)$ is a poset, F is a set of operations on M, and the following condition holds: for every $f \in F$, of arity $n(f)$, there exist $n(f)-$ary operations f/i on M, for $i = 1, \ldots, n(f)$, satisfying the equivalence:

$$\text{(GRES) } f(a_1, \ldots, a_{n(f)}) \leq b \text{ iff } a_i \leq (f/i)(a_1, \ldots, b, \ldots, a_{n(f)}),$$

for all $i = 1, \ldots, n(f)$, $a_j, b \in M$. In (GRES), b takes the place of the $i-$th argument of f/i. One easily shows that every residuated algebra satisfies:

$$\text{(GMON) if } a_i \leq b_i, \text{ for } i = 1, \ldots, n, \text{ then } f(a_1, \ldots, a_n) \leq f(b_1, \ldots, b_n),$$

for all $f \in F$ with $n(f) = n$ and $a_j, b_j \in M$. Models are defined as for the case of NL with residuated algebras instead of residuated groupoids. One can show that the sequents provable in GLC are precisely those which are true in all residuated algebras.

GLC(Φ) can be defined in a similar way as NL(Φ). Theorems 1 and 2 can easily be proven for GLC(Φ), Φ finite.

In the proof of lemma 1, we must consider preordered algebras, i.e. structures $(M, \leq, F)$ such that (M, F) is an algebra and $\leq$ is a preordering on M, satisfying (GMON), for all $f \in F$. The set $C(M)$, of cones on $(M, \leq)$, can be supplied with operations f and f/i, defined as follows:

$$\text{(GM1) } f(P_1, \ldots, P_n) = \{a \in M : (\exists a_1 \in P_1, \ldots a_n \in P_n)\, b \leq f(a_1, \ldots, a_n)\},$$

$$\text{(GM2) } (f/i)(P_1, \ldots, P_n) = \{a \in M : \forall\, [a_j \in P_j]\, f(a_1, \ldots, a, \ldots, a_n) \in P_i\}.$$

In (GM2), $[a_j \in P_j]$ is the list $a_1 \in P_1, \ldots, a_n \in P_n$, with $a_i \in P_i$ omitted, and a takes the place of the $i-$th argument of f. One easily shows that $C(M)$ with $\subseteq$ and the operations defined by (GM1) and (GM2) is a residuated algebra.

For a fixed set T, we construct a preordered algebra $(M. \leq, F)$ in the following way. M is the set of all formula structures whose atomic substructures are in T. F consists of all operations $\circ_f$ admissible in GLC. X directly reduces to Y if either $X \to B$ arises from $Y \to B$ (B is arbitrary) by some rule (f/iL) with $Y_j \to_T A_j$, for all $j \neq i$, or $X \to B$ arises from $Y \to B$ (B is arbitrary) by some rule (fL). The remainder of the proof is essentially the same as in the proof of lemma 1 except for obvious replacements of binary product symbols and operations with $n-$ary product symbols and operations. Let us only consider the proof of:

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

for $A = f(B)$, $A \in T$ (so, $n_f = 1$). Assume $X \in \mu(A)$. By (GM1), there exists $Y \in \mu(B)$ such that $X \leq \circ_f(Y)$. By the induction hypothesis, $Y \to_T B$, hence $\circ_f(Y) \to_T f(B)$, by rule (fR). We obtain $X \leq A$, and consequently, $X \to_T A$. Assume $X \to_T A$. We have $B \in \mu(B)$, by the induction hypothesis, hence

$\circ_f(B) \in \mu(A)$, by (GM1). Since $A \leq \circ_f(B)$, by the definition of $\leq$, and $X \leq A$, then $X \in \mu(A)$.

In the construction of S^T, conditions (R1), (R2) and (R3) must be replaced with:

(GR1) if $\circ_f(A_1, \ldots, A_n) \to A$ belongs to S and $f(A_1, \ldots, A_n) \in T$, then $f(A_1, \ldots, A_n) \to B$ belongs to S,

(GR2) if $\circ_f(A_1, \ldots, A_n) \to B$ belongs to S and $(f/i)(A_1, \ldots, B, \ldots, A_n) \in T$, then $A_i \to (f/i)(A_1, \ldots, B, \ldots, A_n)$ belongs to S;

in (GR2), B takes the place of the $i-$th argument of f/i. Further, basic sequents are $T-$sequents of the form $A \to B$ or $\circ_f(A_1, \ldots, A_{n(f)}) \to B$. S_0 consists of all basic sequents of the form (Id), all basic sequents from Φ, and all basic sequents:

$$\circ_f(A_1, \ldots, A_{n(f)}) \to f(A_1, \ldots, A_{n(f)}),$$

$$\circ_f(A_1, \ldots, (f/i)(A_1, \ldots, A_{n(f)}), \ldots, A_{n(f)}) \to A_i.$$

In the latter, $(f/i)(A_1, \ldots, A_{n(f)})$ takes the place of the $i-$th argument of $\circ_f$.

S^T can be constructed in time polynomial with respect to the cardinality of T. Now, lemmas 2 and 3 and theorems 1 and 2 can be proven by a straightforward modification of the proofs given above. Clearly, this also yields theorems 1 and 2 for Moortgat's calculus NL with minimal modalities and its finite axiomatic extensions (for the pure NL with minimal modalities, theorem 2 has been proven in [Jäger 2002]). The same holds for GLC(Φ) enriched with the permutation rule, for some product symbols f:

$$\text{(GPER)}\ \frac{X[\circ_f(\ldots, Y, \ldots, Z, \ldots)] \to A}{X[\circ_f(\ldots, Z, \ldots, Y, \ldots)] \to A}$$

3 L with nonlogical axioms

Recall that L equals NL plus (ASS). It has been shown in [Buszkowski 1982] that finite axiomatic extensions of L can generate arbitrary recursively enumerable languages, if even nonlogical axioms are of the form $p \circ q \to r$ and $p/q \to r$ (then, the $/-$fragment of L is sufficient). For the full L, an analogous result can be obtained in a more easy way. We give the proof.

A generative grammar is a quadruple $G = (\Sigma, N, s, P, V)$ such that Σ and N are disjoint finite alphabets, $s \in N$, P is a finite set of production rules $x \to y$ such that $x, y \in N^\star$, $x \neq y$, and V is a finite set of lexical rules $a \mapsto p$ such that $a \in \Sigma$ and $p \in N$. The relation $x \to^\star y$ is defined in the standard way (only production rules are taken into account). *The language* of G, denoted by $L(G)$, is the set of all $a_1 \ldots a_n$, $n \geq 0$, such that, for some lexical rules $a_1 \mapsto p_1, \ldots, a_n \mapsto p_n$, we have $s \to^\star p_1 \ldots p_n$. The languages of generative grammars are precisely the recursively enumerable languages (r.e. languages). Recall that elements of Σ and N are called *terminals* and *nonterminals*, respectively.

The $\epsilon-$free r.e. languages can be generated by generative grammars whose production rules are of the form $p \to q$, $p \to qr$ and $pq \to r$, for $p, q, r \in N$; we call them *binary grammars*. This can be shown by a routine construction. First, rules of the form $x \to \epsilon$ and $\epsilon \to x$ are replaced with $x \to E$ and $E \to x$, respectively, where E is a new nonterminal; one also adds rules $Ep \to p$, $pE \to p$, $p \to Ep$, $p \to pE$, for any $p \in N$. Clearly, the new grammar is equivalent to the initial one. Second, by introducing new nonterminals, every rule of the form $p \to p_1 \dots p_n$, $n > 2$, is replaced with a set of rules of the form $q \to q'q''$, and every rule of the form $p_1 \dots p_n \to p$, $n > 2$, is replaced with a set of rules of the form $q'q'' \to q$. To generate r.e. languages with ϵ it is sufficient to admit lexical rules of the form $\epsilon \to s$.

L(Φ) is L with (CUT) and all sequents from Φ as nonlogical axioms. The subformula property for L(Φ) can be proven by a modification of the proof of lemma 1. *A residuated semigroup* is a residuated groupoid, satisfying: $(ab)c = a(bc)$, for all elements a, b, c. Sequents provable in L(Φ) are precisely those which are true in all models $(\mathcal{M}, \mu)$ such that $\mathcal{M}$ is a residuated semigroup, and all sequents from Φ are true in the model. Let T be a set of formulas which contains all formulas from Φ and is closed under subformulas. M is defined as in the proof of lemma 1. We define: X directly reduces to Y iff either X directly reduces to Y in the sense of lemma 1, or $X \to B$ arises from $Y \to B$ (B is arbitrary), by (ASS). Then, $(M, \leq, \circ)$ is a preordered groupoid, but $C(M)$ is a residuated semigroup. The remainder of the proof of lemma 1 goes without change.

Let $G = (\Sigma, N, s, P, V)$ be a binary grammar. We construct a finite set $\Phi(G)$ of nonlogical axioms, to be added to L. Symbols from N are identified with some atomic formulas. $\Phi(G)$ contains all sequents $q \to p$, for $p \to q$ belonging to P, $q \bullet r \to p$, for $p \to qr$ belonging to P, and all $r \to p \bullet q$, for $pq \to r$ belonging to P.

Lemma 4 *For all $p_1, \dots, p_n, p' \in N$, $p' \to^\star p_1 \dots p_n$ in the sense of G iff $p_1 \circ \dots \circ p_n \to p'$ is provable in $L(\Phi(G))$.*

Proof. Due to (ASS), we may omit parentheses in $p_1 \circ \dots \circ p_n$. The 'only if' part is proven by induction on derivations in G. If $n = 1$ and $p' = p_1$, then $p' \to p'$ is (Id). Assume $p_1 \dots p_n$ arises by production rule $p \to q$. Then, $p_i = q$ and $p_1 \dots p \dots p_n$ is derivable from p'. We use the induction hypothesis and (CUT). Assume $p_1 \dots p_n$ arises by production rule $p \to qr$. Now, $p_i = q$, $p_{i+1} = r$, and $p_1 \dots p_{i-1} p p_{i+2} \dots p_n$ is derivable from p'. Since $q \bullet r \to p$ is a nonlogical axiom, and $q \circ r \to q \bullet r$ is provable in L, then our thesis holds, by the induction hypothesis and two applications of (CUT). Assume $p_1 \dots p_n$ arises by production rule $pq \to r$. Now, $p_i = r$ and $p_1 \dots p_{i-1} pq p_{i+1} \dots p_n$ is derivable from p'. Since $r \to p \bullet q$ is a nonlogical axiom, then our thesis holds, by the induction hypothesis, ($\bullet$L) and (CUT).

Let T be the smallest set of formulas which contains all atoms from N and all formulas from $\Phi(G)$ and is closed under subformulas. Clearly, T is finite, and each formula from T has the form either p, or $p \bullet q$, for $p, q \in N$. We define a string $a(X) \in N^\star$, for every formula structure X which consists of formulas from T: $a(p) = p$, $a(p \bullet q) = a(p)a(q)$, $a(X \circ Y) = a(X)a(Y)$.

The 'if' part is a consequence of the following claim: if $X \to A$ is a $T-$sequent provable in $\mathrm{L}(\Phi(G))$, then $a(A) \to^\star a(X)$ in G. We use induction on proofs in $\mathrm{L}(\Phi(G))$, consisting of $T-$sequents. If $X \to A$ is an axiom of $\mathrm{L}(\Phi(G))$, then the thesis is obvious. The only inference rules to be considered are ($\bullet$L), ($\bullet$R) and (CUT). Assume $X \to A$ arises by ($\bullet$L). Then, $X = X[p \bullet q]$, and the premise is $X[p \circ q] \to A$. By the induction hypothesis, $a(A) \to^\star a(X[p \circ q])$, hence $a(A) \to^\star a(X[p \bullet q])$, since $a(X[p \circ q]) = a(X[p \bullet q])$. For ($\bullet$R), we use the fact: $x \to^\star y$ and $x' \to^\star y'$ entail $xx' \to^\star yy'$. For (CUT), we use the fact: $x \to^\star yzy'$ and $z \to^\star z'$ entail $x \to^\star yz'y'$. Q.E.D.

Categorial grammars based on $\mathrm{L}(\Phi)$ are defined as in section 2 with $\mathrm{L}(\Phi)$ replacing $\mathrm{NL}(\Phi)$. We state the main result of this section.

Theorem 3 *For any binary grammar G, there exists a categorial G', based on $L(\Phi(G))$, such that $L(G) = L(G', s)$.*

Proof. Consider the grammar G', based on $\mathrm{L}(\Phi(G))$, whose lexical rules are the same as the lexical rules of G. By lemma 4, $L(G', s) = L(G)$. Q.E.D.

As a consequence, we obtain the undecidability of $\mathrm{L}(\Phi)$, for some finite Φ. Here, Φ may consist of sequents of the form $p \to q \bullet r$ and $p \bullet q \to r$. The latter sequents can be replaced with $p \circ q \to r$ or $p \to r/q$, since both are deductively equivalent to $p \bullet q \to r$ in L. The former sequents, however, essentially involve product. In [Buszkowski 1982], it has been shown that sequents $p \to q \bullet r$ can be replaced with a finite number of product-free sequents of the form $p' \to r'/q'$ and $p'/q' \to r'$; the new set of nonlogical axioms is not deductively equivalent to the initial one, but it yields the same derivable sequents of the form $p_1 \circ \cdots p_n \to p$. We refer the reader to [Buszkowski 1982] for details.

References

[Buszkowski 1982] W. Buszkowski, Some decision problems in the theory of syntactic categories, *Zeitschrift für mathematische Logik und Grundlagen der Mathematik* 28 (1982), 539-548.

[Buszkowski 1986] W. Buszkowski, Completeness results for Lambek Syntactic Calculus, *Zeitschrift für mathematische Logik und Grundlagen der Mathematik* 32 (1986), 13-28.

[Buszkowski 1986a] W. Buszkowski, Generative capacity of Nonassociative Lambek Calculus, *Bulletin of Polish Academy of Sciences. Mathematics* 34 (1986), 507-516.

[Buszkowski 1989] W. Buszkowski, *Logical Foundations of Ajdukiewicz and Lambek Categorial Grammars* (in Polish), Polish Scientific Publishers, Warsaw, 1989.

[Buszkowski 1996] W. Buszkowski, Extending Lambek Grammars to Basic Categorial Grammars, *Journal of Logic, Language and Information* 5 (1996), 279-295.

[Casadio and Lambek 2002] C. Casadio and J. Lambek, A Tale of Four Grammars, *Studia Logica* 71 (2002), 315-329, (special issue: The Lambek calculus in logic and linguistics).

[Došen 1992] K. Došen, A brief survey of frames for the Lambek calculus, *Zeitschrift für mathematische Logik und Grundlagen der Mathematik* 38 (1992), 179-187.

[de Groote and Lamarche 2002] P. de Groote and F. Lamarche, Classical Non-Associative Lambek Calculus, *Studia Logica* 71 (2002), 355-388, (special issue: The Lambek calculus in logic and linguistics).

[Jäger 2002] G. Jäger, Residuation, Structural Rules and Context Freeness, manuscript, 2002, to appear in: *Journal of Logic, Language and Information.*

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

[Kandulski 1995] M. Kandulski, On commutative and nonassociative syntactic calculi and categorial grammars, *Mathematical Logic Quarterly* 41 (1995), 217-235.

[Kandulski 1997] M. Kandulski, On generalized Ajdukiewicz and Lambek calculi and grammars, *Fundamenta Informaticae* 30 (1997), 169-181.

[Keenan and Faltz 1985] E.L. Keenan and L.M. Faltz, *Boolean Semantics for Natural Language*, D. Reidel, Dordrecht, 1985.

[Kołowska 1997] M. Kołowska-Gawiejnowicz, Powerset residuated algebras and Generalized Lambek Calculus, *Mathematical Logic Quarterly* 43 (1997), 60-72.

[Lambek 1958] J. Lambek, The mathematics of sentence structure, *American Mathematical Monthly* 65 (1958), 154-170.

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

[Lambek 1999] J. Lambek, Type grammars revisited, in: A. Lecomte, F. Lamarche and G. Perrier, *Logical Aspects of Computational Linguistics*, LNAI 1582, Springer, Berlin, 1999, 1-27.

[Lambek 2001] J. Lambek, Type Grammars as Pregroups, *Grammars* 4 (2001), 21-39.

[Moortgat 1995] M. Moortgat, Multimodal linguistic inference, *Bulletin of IGPL* 3 (1995), 371-401.

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

[Pentus 1993] M. Pentus, Lambek grammars are context-free, *Proc. 8 IEEE Symposium on Logic in Computer Science*, 1993, 429-433.

[Pentus 1995] M. Pentus, Models for the Lambek calculus, *Annals of Pure and Applied Logic* 75 (1995), 179-213.

[Pentus 2003] M. Pentus, Lambek calculus is NP-complete, manuscript, Moscow, 2003.

[Restall 2000] G. Restall, *An Introduction to Substructural Logics*, Routledge, London and New York, 2000.

[Roorda 1991] D. Roorda, *Resource Logics: Proof Theoretical Investigations*, Ph.D. Thesis, University of Amsterdam, 1991.

[van Benthem 1986] J. van Benthem, *Essays in Logical Semantics*, D. Reidel, Dordrecht, 1986.

[van Benthem 1991] J. van Benthem, *Language in Action. Categories, Lambdas and Dynamic Logic*, North-Holland, Amsterdam, 1991.