Conjunctive Categorial Grammars

Stepan Kuznetsov
Steklov Mathematical Institute
Moscow, Russia
sk@mi.ras.ru

Alexander Okhotin
St. Petersburg State University
St. Petersburg, Russia
alexander.okhotin@spbu.ru

Abstract

Basic categorial grammars are enriched with a conjunction operation, and it is proved that the formalism obtained in this way has the same expressive power as conjunctive grammars, that is, context-free grammars enhanced with conjunction. It is also shown that categorial grammars with conjunction can be naturally embedded into the Lambek calculus with conjunction and disjunction operations. This further implies that a certain NP-complete set can be defined in the Lambek calculus with conjunction.

1 Introduction

This paper establishes a connection between two formal grammar models that emerged within two different theories of syntax.

One theory is the *immediate constituent analysis*, which has its roots in the traditional grammar, and was investigated in the early 20th century by the structural linguists. It reached universal recognition under the name "context-free grammars," introduced in Chomsky's early work. In this paradigm, a grammar assigns certain properties to groups of words, such as "noun phrase" (NP), "verb phrase" (VP) or "sentence" (S). These properties are known as *syntactic categories*, or, in Chomsky's terminology, *nonterminal symbols*. Rules of a grammar, such as S $\to$ NP VP, show how shorter substrings with known properties can be concatenated to form longer strings belonging to a certain category.

In the other theory, which was discovered by Ajdukiewicz (1935), further developed by Bar-Hillel et al. (1960), and is nowadays known as *categorial grammars*, syntactic categories are defined in different way. "Noun phrases" are treated as "the category of phrases *equivalent* to nouns," whereas verb phrases are defined as "the category of phrases, which would form a complete sentence, if a noun, or anything equivalent to a noun, is concatenated on the left," denoted by NOUN$\backslash$SENTENCE. A categorial grammar explicitly assigns categories to individual words; and then, by definition, a concatenation of any string of type NOUN with any string of type NOUN$\backslash$SENTENCE forms a complete sentence. The crucial feature of this approach is that the laws that govern reduction of categories, namely, $A\,(A \setminus B)$ to B, and $(B \,/\, A)\,A$ to B, are universal. In contrast, Chomsky's context-free formalism uses different rules for different categories.

A formal connection between these two models was established by Bar-Hillel et al. (1960), who proved them to be equivalent in power: a language is defined by a context-free grammar if and only if it is defined by a basic categorial grammar (assuming languages do not contain the empty string).

More than half a century of research gave birth to many extensions of both basic models. Categorial grammars were the first to get an interesting extension: the *Lambek calculus*, introduced by Lambek (1958), augments the model with additional derivation rules. Later, Pentus (1993) established that this extended model is still equivalent in power to context-free grammars. Pentus' translation yields a context-free grammar of exponential size with respect to the original Lambek grammar. For the special case of *unidirectional* Lambek grammars, which use only one kind of division operators ($\backslash$, $/$), but not both, Kuznetsov (2016), using the ideas of Savateev (2009), presents a polynomial translation into context-free grammars. Other generalizations of categorial grammars include combinatory categorial grammars by Steedman (1996), categorial dependency grammars by Dekhtyar and Dikovsky

(2008), and others.

Lambek grammars, in their turn, can be generalized further. Modern extensions of the Lambek calculus with new operations, such as those considered by Carpenter (1997), Morrill (2011), and Moot and Retoré (2012), are capable of describing quite sophisticated syntactic phenomena.

From the point of view of modern logic, the Lambek calculus is a substructural logical system, namely, a non-commutative variation of *linear logic*, introduced by Girard (1987), see Abrusci (1990), Yetter (1990). Linear logic offers many logical operations, and some of them can be used in the non-commutative case for extending Lambek grammars.

Morrill (2011) and his collaborators, following and extending Moortgat (1996), consider a system, based on the Lambek calculus, with discontinuous connectives, subexponentials for controlled non-linearity, brackets for controlled non-associativity, and many other operations. The use of negation in categorial grammars was considered by Buszkowski (1996). Kanazawa (1992) investigated the power of Lambek grammars with conjunction and disjunction operations that are "additive operations" in terms of linear logic.

Numerous generalized models have also been introduced in the paradigm of immediate constituent analysis, as extensions of the context-free formalism. One direction is to extend the form of constituents, that is, sentence fragments to which syntactic categories are being assigned in a grammar. The most well-known of these models are the multi-component grammars, introduced by Vijay-Shanker et al. (1987) and by Seki et al. (1991), inspired by an earlier model by Fischer (1968): these grammars define the properties of discontinuous constituents, that is, substrings with a bounded number of gaps. Extensions of another kind augment the model by introducing new logical operators to be used in grammar rules: for instance, *conjunctive grammars*, featuring a conjunction operation, and *Boolean grammars*, further equipped with negation, were introduced by Okhotin (2001, 2004). Earlier, Latta and Wall (1993) argued for the relevance of such operations in linguistic descriptions. The main results on conjunctive grammars indicate that they preserve the practically useful properties of context-free grammars, such as efficient parsing algorithms, while substantially extending their expressive power. The known results on conjunctive grammars are presented in a fairly recent survey by Okhotin (2013).

A few years ago, Kuznetsov (2013) compared the expressive power of Lambek grammars with conjunction, as considered by Kanazawa (1992), with that of conjunctive grammars. It was proved that a large subclass of conjunctive grammars (namely, conjunctive grammars in Greibach normal form) can be simulated in the Lambek calculus with conjunction, but the exact power of the latter remains undetermined.

This paper makes a fresh attempt at introducing conjunction in categorial grammars. The new model extends basic categorial grammars, rather than Lambek grammars, and for that reason it uses categories and rules of a simpler form than in the earlier model by Kanazawa (1992) and Kuznetsov (2013). Yet, it is shown that this model can simulate every conjunctive grammar. A converse simulation is presented as well, which implies the equivalence of the two models.

As compared to the classical equivalence result for context-free grammars and basic categorial grammars, the new result requires a more elaborate construction. One particular difficulty is that the normal form theorems for conjunctive grammars are weaker than those for the context-free grammars: in particular, no analogue of the Greibach normal form is known for conjunctive grammars. For this reason, the simulation of conjunctive grammars by the proposed conjunctive categorial grammars relies on a different normal form by Okhotin and Reitwießner (2010). This leads to a representation of the whole class of conjunctive grammars, in contrast to the result by Kuznetsov (2013), which is valid only for grammars in Greibach normal form.

The second result is that conjunctive categorial grammars, as defined in this paper, can be represented in the Lambek calculus with the conjunction operation, as considered by Kanazawa (1992), and therefore this extension of the Lambek calculus is at least as powerful as are the conjunctive grammars. Furthermore, it is proved that Kanazawa's model (Kanazawa, 1992) can describe an NP-complete language, which conjunctive grammars cannot describe unless $\mathrm{P} = \mathrm{NP}$.

2 Basic Categorial Grammars and Context-Free Grammars

Let Σ be a finite *alphabet* of the language being defined. Its elements are called *symbols*. In linguistic descriptions, symbols typically represent words of the language. The set of non-empty strings over Σ is denoted by Σ^+. Throughout this paper, any subset of Σ^+ is a *language*, that is, all languages are assumed to be without the empty string.

The models considered in this paper are derived from two classical formal grammar frameworks: basic categorial grammars and context-free grammars.

Basic categorial grammars (BCG) have their roots in the works of Ajdukiewicz (1935). Let Σ be an alphabet. Let $\mathbf{Pr} = \{p, q, r, \dots\}$ be a finite set of *primitive categories*, and let $s \in \mathbf{Pr}$ be a designated *target category* of all syntactically correct sentences.

The set $\mathbf{BCat}$ of *basic categories* is defined as follows.

- Every primitive category is a basic category.
- If $A \in \mathbf{BCat}$ and $p \in \mathbf{Pr}$, then $(p \,\backslash\, A) \in \mathbf{BCat}$ and $(A \,/\, p) \in \mathbf{BCat}$.

The definition of a basic categorial grammar is given in terms of logical *propositions*, which are expressions of the form $B(v)$, where $B \in \mathbf{BCat}$ and $v \in \Sigma^+$. This proposition states that v is a string of syntactic category B.

The language of propositions is indeed very simple: all propositions are atomic, there are no variables and quantifiers (if the syntactic category B is considered, in the spirit of first-order logic, as a predicate, then its argument, v, is a constant term).

A categorial grammar is regarded as a *logical calculus* for deriving categorial propositions. It includes a finite set of axioms (axiomatic propositions) of the form $A(a)$, where $A \in \mathbf{BCat}$ and $a \in \Sigma$, and the following inference rules.

$$\frac{p(u) \quad (p \,\backslash\, A)(v)}{A(uv)} \qquad \frac{(A \,/\, p)(u) \quad p(v)}{A(uv)}$$

The string w belongs to the language generated by the BCG if and only if the proposition $s(w)$ is derivable by means of this calculus.

Example 1. *The basic categorial grammar with the following axiomatic propositions and with s as the target category ($\mathbf{Pr} = \{s, p, q\}$) describes the language $\{ba^n ca^n \mid n \geqslant 0\}$.*

$$(s \,/\, p)(b), \quad p(c), \quad (p \,/\, q)(a), \quad (p \,\backslash\, q)(a)$$

Another, more well-known formal grammar framework is the phrase-structure formalism, defined by Chomsky (1956) and later renamed into *context-free grammars* (CFG). In a CFG, there is a fixed finite set of categories N (usually called "non-terminal symbols"), and one of them is designated as the initial symbol $S \in N$. The grammar is defined by a finite set of rules (or "productions") of the form $A \to \beta$, where $A \in N$ and $\beta \in (\Sigma \cup N)^+$.

Even though Chomsky's original definition of context-free grammars was given in terms of string rewriting, it is more convenient—at least in this paper—to present it as a logical derivation similar to the one in categorial grammars. Propositions in the context-free framework are of the form $\beta(u)$, where $\beta \in (\Sigma \cup N)^+$ and $u \in \Sigma^+$. Intuitively, such a proposition means that u can be derived from β using the rules of the CFG. Axioms of the calculus of propositions are of the form $a(a)$, $a \in \Sigma$, and the rules of inference are as follows.

$$\frac{\beta_1(u_1) \quad \beta_2(u_2)}{(\beta_1\beta_2)(u_1u_2)}$$

$$\frac{\beta(v)}{A(v)} \qquad \text{for each rule } A \to \beta$$

Again, the string w belongs to the language generated by this grammar if and only if the proposition $S(w)$ is derivable.

Example 2. *The language from Example 1 is described by the following CFG.*

$$S \to bA$$
$$A \to aAa \mid c$$

As usual, "$A \to aAa \mid c$" is a short-hand notation for two rules, $A \to aAa$ and $A \to c$.

There is an important difference between BCGs and CFGs: in BCGs, the linguistic information is stored in the axioms (in other words, it is *lexicalized*), while the inference rules are the same for all BCGs. For CFGs, the situation is opposite: axioms are trivial, and all information is kept in the rules. However, these two formalisms are equivalent in power.

Theorem A. *A language is generated by a BCG if and only if it is generated by a CFG.* (Bar-Hillel et al., 1960)

$$\dfrac{\dfrac{b(b) \quad \dfrac{a(a) \quad \dfrac{c(c)}{B(c)} \quad a(a)}{B(aca)} \quad c(a) \quad \dfrac{a(a)}{A(a)}}{bBcA(bacaca)} \qquad \dfrac{b(b) \quad \dfrac{a(a)}{A(a)} \quad c(a) \quad \dfrac{a(a) \quad \dfrac{c(c)}{B(c)} \quad a(a)}{B(aca)}}{bAcB(bacaca)}}{S(bacaca)}$$

Figure 1

3 Conjunction in Grammars

In this section, both grammar formalisms are enriched with a *conjunction* operation. Using conjunction, one can impose multiple syntactic constraints on the same phrase at the same time. The extension of context-free grammars with conjunction, called *conjunctive grammars,* was introduced by Okhotin (2001).

Let Σ be the alphabet, and let N be the set of categories ("non-terminal symbols"), with $S \in N$ representing all well-formed sentences. A conjunctive grammar is defined by a finite set of rules of the form $A \to \beta_1 \,\&\, \dots \,\&\, \beta_k$, with $\beta_i \in (\Sigma \cup N)^+$. If k is 1, then this is an ordinary rule $A \to \beta$, as in an ordinary context-free grammar.

Propositions in a conjunctive grammar are of the form $\beta(u)$, where $u \in \Sigma^+$ and $\beta \in (\Sigma \cup N)^+$. Axioms of the calculus of propositions are of the form $a(a)$, where $a \in \Sigma$. The first inference rule is as follows.

$$\frac{\beta_1(u_1) \quad \beta_2(u_2)}{(\beta_1\beta_2)(u_1u_2)}$$

The other inference rule is valid for each grammar rule $A \to \beta_1 \,\&\, \dots \,\&\, \beta_k$ and for each string v.

$$\frac{\beta_1(v) \quad \dots \quad \beta_k(v)}{A(v)}$$

The string w belongs to the language generated by the grammar if and only if the proposition $S(w)$ is derivable from the axioms.

Example 3. *The following conjunctive grammar describes the language* $\{ba^nca^nca^n \mid n \geqslant 1\}$.

$$\begin{aligned} S &\to bBcA \,\&\, bAcB \\ A &\to aA \mid a \\ B &\to aBa \mid c \end{aligned}$$

The rules for A and B use no conjunction, and have the same effect as in ordinary context-free grammars. Thus, $bBcA(w)$ is true for all strings of the form $w = ba^nca^nca^i$, with $n \geqslant 0$, $i \geqslant 1$, whereas $bAcB(w)$ holds true for strings of the form $w = ba^ica^nca^n$. The conjunction of these two conditions is exactly the condition of membership in the desired language, and the rule for S ensures it by derivations of the following form.

$$\frac{bBcA(ba^nca^nca^n) \quad bAcB(ba^nca^nca^n)}{S(ba^nca^nca^n)}$$

A full derivation of the string $w = bacaca$ is given in Figure 1.

The notion of a *conjunctive categorial grammar* is defined by extending basic categorial grammars with the conjunction operation. Let $\mathbf{Pr} = \{p, q, r, \dots\}$ be the set of primitive categories, $s \in \mathbf{Pr}$ is the target category.
The set of *conjuncts,* $\mathbf{Conj}$, is defined as follows:

1. every primitive category is a conjunct;
2. if $p_1, \dots, p_k \in \mathbf{Pr}$, then $(p_1 \wedge \dots \wedge p_k) \in \mathbf{Conj}$.

The set of *basic categories with conjunction,* $\mathbf{BCat}_\wedge$, is defined as follows.

1. Every primitive category belongs to $\mathbf{BCat}_\wedge$.
2. If $\mathcal{C} \in \mathbf{Conj}$ and $A \in \mathbf{BCat}_\wedge$, then $(\mathcal{C} \setminus A) \in \mathbf{BCat}_\wedge$ and $(A \,/\, \mathcal{C}) \in \mathbf{BCat}_\wedge$.

Categorial propositions are expressions of the form $B(v)$, where $v \in \Sigma^+$ and $B \in \mathbf{BCat}_\wedge \cup \mathbf{Conj}$. A conjunctive categorial grammar is a logical theory deriving categorial propositions. It includes an arbitrary finite set of axioms of the form $A(a)$, with $A \in \mathbf{BCat}_\wedge$ and $a \in \Sigma$, and the following inference rules.

$$\frac{p_1(v) \quad \dots \quad p_k(v)}{(p_1 \wedge \dots \wedge p_k)(v)}$$

$$\frac{\mathcal{C}(u) \quad (\mathcal{C} \setminus A)(v)}{A(uv)} \qquad \frac{(A \,/\, \mathcal{C})(v) \quad \mathcal{C}(u)}{A(vu)}$$

$$
\dfrac{(s\,/(x\wedge y))(b) \qquad \dfrac{\dfrac{\dfrac{\dfrac{(p\,/\,q)(a) \quad \dfrac{p(c) \quad (p\,\backslash\,q)(a)}{q(ca)}}{p(aca)} \quad (p\,\backslash(x\,/\,r))(c)}{(x\,/\,r)(acac)} \quad r(a)}{x(acaca)} \qquad \dfrac{r(a) \quad \dfrac{((r\,\backslash\,y)\,/\,p)(c) \quad \dfrac{(p\,/\,q)(a) \quad \dfrac{p(c) \quad (p\,\backslash\,q)(a)}{q(ca)}}{p(aca)}}{(r\,\backslash\,y)(caca)}}{y(acaca)}}{(x\wedge y)(acaca)}}{s(bacaca)}
$$

Figure 2

The string w belongs to the language generated by this grammar if and only if the proposition $s(w)$ is derivable.

Example 4. *The categorial conjunctive grammar with the set of primitive categories* $\mathbf{Pr} = \{s, x, y, p, q, r\}$*, s as the target category, and with the following set of axioms describes the language* $\{ba^n ca^n ca^n \mid n \geqslant 1\}$*, the same as in Example 3.*

$$
\begin{array}{lll}
r(a), & (r\,/\,r)(a), & \\
p(c), & (p\,/\,q)(a), & (p\,\backslash\,q)(a), \\
(p\,\backslash(x\,/\,r))(c), & ((r\,\backslash\,y)\,/\,p)(c), & \\
(s\,/(x\wedge y))(b). & &
\end{array}
$$

The category p is defined in the same way as in Example 1. Then, using further categories without conjunction, the propositions $x(a^n ca^n ca^i)$ and $y(a^i ca^n ca^n)$, for all $n \geqslant 0$, $i \geqslant 1$, are derived as in an ordinary categorial grammar. The only strings that satisfy both conditions, x and y, are those of the form $a^n ca^n ca^n$, and these are therefore all strings in the category s, derived as follows.

$$
\dfrac{(s\,/(x\wedge y))(b) \qquad \dfrac{x(a^n ca^n ca^n) \quad y(a^n ca^n ca^n)}{(x\wedge y)(a^n ca^n ca^n)}}{s(ba^n ca^n ca^n)}
$$

A complete derivation of the proposition $s(bacaca)$ is presented in Figure 2.

The calculus used in the conjunctive categorial grammar formalism enjoys the following *inverted subformula property* (ISF): if a category of the form $(\mathcal{C}\setminus A)$ or $(A\,/\,\mathcal{C})$ appears somewhere in the derivation, then it is a subexpression of some category used in an axiom. (The notion of subexpression on categories is defined in a standard way: each conjunct (in particular, primitive category) is a subexpression of itself, and subexpressions of $(\mathcal{C}\setminus A)$ include $\mathcal{C}\setminus A$, $\mathcal{C}$, and all subexpressions of A; symmetrically for $(A\,/\,\mathcal{C})$. To prove the ISF, we trace the rightmost branch of the derivation upwards; finally we reach an axiom that includes the goal category as a subexpression.)

Another useful property is the fact that the rule for $\wedge$ is invertible: if $(p_1 \wedge \ldots \wedge p_k)(v)$ is derivable, then so are $p_1(v), \ldots, p_k(v)$. Indeed, the only way to derive $(p_1 \wedge \ldots \wedge p_k)(v)$ is by applying this rule.

The calculus used in conjunctive categorial grammars also enjoys the following *cut elimination* property.

Lemma 1. *Let $A(u)$ (for some $A \in \mathbf{BCat}_\wedge$ and $u \in \Sigma^+$) be derivable in the given conjunctive categorial grammar. Consider a new conjunctive categorial grammar over an extended alphabet $\Sigma \cup \{b\}$, where $b \notin \Sigma$. The new grammar has all the same axioms as the original grammar, and an additional axiom $A(b)$. Then, if the new grammar derives $B(v_1 b v_2)$, for some $B \in \mathbf{BCat}_\wedge$ and arbitrary, possibly empty, strings v_1, v_2 over Σ, then $B(v_1 u v_2)$ is derivable in the original grammar.*

Proof. Consider the derivation of $B(v_1 b v_2)$ in the extended grammar and substitute u for all occurrences of b. Applications of inference rules remain valid; the same for axioms of the old grammar (they don't include b). The new axiom $A(b)$ becomes $A(u)$, which is derivable in the old grammar by assumption. $\square$

4 Equivalence of Conjunctive Grammars and Conjunctive Categorial Grammars

The main result of this paper is an extension of Theorem A for grammars with conjunction, stated as follows.

Theorem 1. *A language is generated by a conjunctive grammar if and only if it is generated by a conjunctive categorial grammar.*

The proof uses the following two known properties of conjunctive grammars. The first result is their closure under quotient with a single symbol.

Lemma B. *If L is a language over Σ described by a conjunctive grammar, and $a \in \Sigma$ is any symbol, then there exists a conjunctive grammar that describes the language $a^{-1}L = \{w \mid aw \in L\}$.* (Okhotin and Reitwießner, 2010, Thm. 2)

The other result is a normal form theorem. A conjunctive grammar G with the initial symbol S is in the *odd normal form*, if all its rules are of the following form, with $A \in N$, $a \in \Sigma$, $B_i, C_i \in N$, and $a_i \in \Sigma$.

$$\begin{aligned} &A \to a \\ &A \to B_1 a_1 C_1 \,\&\, \ldots \,\&\, B_k a_k C_k \\ &S \to aA \end{aligned}$$

Rules of the latter kind are allowed only if S is never referenced in any rules.

Theorem C. *Every language described by a conjunctive grammar can be described by a conjunctive grammar in odd normal form.* (Okhotin and Reitwießner, 2010)

Proof of Theorem 1. **The "if" part** is easier. A conjunctive grammar equivalent to a given categorial grammar G has the set N comprised of all categories used in the axioms of G, and of all their subexpressions (categories and conjuncts). The conjunctive rules are now as follows.

$$\begin{aligned} (p_1 \wedge \ldots \wedge p_k) &\to p_1 \,\&\, \ldots \,\&\, p_k, \\ A &\to \mathcal{C}\,(\mathcal{C} \setminus A), \\ A &\to (A \,/\, \mathcal{C})\,\mathcal{C}, \end{aligned}$$

for all $(p_1 \wedge \ldots \wedge p_k)$, $(\mathcal{C} \setminus A)$, $(A \,/\, \mathcal{C}) \in N$, and

$$A \to a, \quad \text{if } A(a) \text{ is an axiom in } G.$$

For **the "only if" part** of the proof, the first step is to transform a given conjunctive grammar. Let $\Sigma = \{a_1, \ldots, a_n\}$. For each symbol a_i, by Lemma B, there is a conjunctive grammar G_i that describes the quotient $a_i^{-1}L$. By Theorem C, this grammar can be assumed to be in the odd normal form. It can also be assumed that, for $i \neq j$, the grammars G_i and G_j have disjoint sets of non-terminal symbols. Let S_i be the initial symbol of G_i. Then this grammar is further modified as follows. Every rule

$$A \to B_1 a_1 C_1 \,\&\, \ldots \,\&\, B_k a_k C_k$$

is replaced with $k + 1$ new rules:

$$A \to \widetilde{X}_1 \,\&\, \ldots \,\&\, \widetilde{X}_k \qquad \text{and} \qquad \widetilde{X}_i \to B_i a_i C_i,$$

where $\widetilde{X}_i$ are fresh non-terminals. For the sake of uniformity, rules of the form

$$S_i \to aA$$

are replaced with

$$S_i \to \widetilde{Y} \qquad \text{and} \qquad \widetilde{Y} \to aA,$$

and rules of the form

$$A \to a$$

are replaced with

$$A \to \widetilde{Z} \text{ and } \widetilde{Z} \to a.$$

Finally, a new conjunctive grammar for L is obtained by joining these grammars together, for all i, adding the following extra rules for the new initial symbol $\widetilde{S}$.

$$\widetilde{S} \to a_1 S_1, \qquad \ldots, \qquad \widetilde{S} \to a_n S_n$$

In the resulting grammar, all non-terminals are of two sorts (with and without a tilde), and the rules have the following form.

$$\begin{gathered} A \to \widetilde{X}_1 \,\&\, \ldots \,\&\, \widetilde{X}_k \qquad \text{(here } k \text{ could be 1)} \\ \widetilde{X} \to BaC, \qquad \widetilde{Y} \to aA, \qquad \text{and} \qquad \widetilde{Z} \to a \end{gathered}$$

It is then transformed to a conjunctive categorial grammar, with the set of primitive categories $\mathbf{Pr} = \{p_{\widetilde{X}} \mid \widetilde{X}$ is a non-terminal decorated with a tilde $\}$, and with the following axioms.

1. For each rule $\widetilde{Z} \to a$, there is an axiom $p_{\widetilde{Z}}(a)$.
2. For each pair of rules $\widetilde{Y} \to aA$ and $A \to \widetilde{X}_1 \,\&\, \ldots \,\&\, \widetilde{X}_k$, the axiom is $\Big(p_{\widetilde{Y}} \,/ (p_{\widetilde{X}_1} \wedge \ldots \wedge p_{\widetilde{X}_k})\Big)(a)$.
3. For each triple of rules $\widetilde{X} \to BaC$, $B \to \widetilde{Y}_1 \,\&\, \ldots \,\&\, \widetilde{Y}_k$, and $C \to \widetilde{Z}_1 \,\&\, \ldots \,\&\, \widetilde{Z}_m$, the axiom is

$$\Big(\big((p_{\widetilde{Y}_1} \wedge \ldots \wedge p_{\widetilde{Y}_k}) \setminus p_{\widetilde{X}}\big) \,/ (p_{\widetilde{Z}_1} \wedge \ldots \wedge p_{\widetilde{Z}_m})\Big)(a).$$

The target category is $p_{\widetilde{S}}$.

Claim. *For every non-terminal $\widetilde{X}$ decorated with a tilde, the proposition $p_{\widetilde{X}}(v)$ is derivable in the newly constructed conjunctive categorial grammar if and only if $\widetilde{X}(v)$ is derivable in the original conjunctive grammar.*

Proof. **The "if" part.** The proof proceeds by induction on the derivation size. There are three possible cases.

Case 1. The proposition $\widetilde{X}(v)$ is actually of the form $\widetilde{Z}(a)$ and is derived from $a(a)$ using the rule $\widetilde{Z} \to a$. Then $p_{\widetilde{Z}}(a)$ is an axiom in the conjunctive categorial grammar.

Case 2. The proposition $\widetilde{X}(v)$ is of the form $\widetilde{Y}(av_1)$ and is derived from $a(a)$ and $A(v_1)$ using a rule of the form $\widetilde{Y} \to aA$. Next, a rule of the form $A \to \widetilde{X}_1 \,\&\, \ldots \,\&\, \widetilde{X}_k$ should be applied for A. Therefore, the propositions $\widetilde{X}_i(v_1)$ are derivable (for all i) in the original conjunctive grammar. Then, by induction hypothesis, $p_{\widetilde{X}_i}(v_1)$ are derivable in the conjunctive categorial grammar, and there is a derivation for $p_{\widetilde{Y}}(av_1)$ shown in Figure 3.

Case 3. The proposition $\widetilde{X}(v)$ is of the form $\widetilde{X}(v_1av_2)$ and is derived from some propositions of the form $B(v_1)$, $a(a)$, and $C(v_2)$, following the rule $\widetilde{X} \to BaC$. Next, for B and C, some rules for the form $B \to \widetilde{Y}_1 \,\&\, \ldots \,\&\, \widetilde{Y}_k$ and $C \to \widetilde{Z}_1 \,\&\, \ldots \,\&\, \widetilde{Z}_m$ should be applied. Therefore, the propositions $\widetilde{Y}_i(v_1)$ and $\widetilde{Z}_j(v_2)$ are derivable (for all i, j) in the original conjunctive grammar. Then, by induction hypothesis, $p_{\widetilde{Y}_i}(v_1)$ and $p_{\widetilde{Z}_j}(v_2)$ are derivable in the conjunctive categorial grammar, and there is a derivation for $p_{\widetilde{X}}(v_1av_2)$, as shown in Figure 4.

The "only if" part. This time, it is assumed that $p_{\widetilde{X}}(v)$ is derivable in the newly constructed conjunctive categorial grammar. The proof is by induction on its derivation.

The *axiom case* is trivial: any axiom of the form $p_{\widetilde{Z}}(a)$ is associated with a rule $\widetilde{Z} \to a$ in the original conjunctive grammar, and then $\widetilde{Z}(a)$ is derivable from $a(a)$.

In the *left division* case, $v = v_1w$, and the last step of the derivation is as follows.

$$\frac{\mathcal{C}(v_1) \quad (\mathcal{C} \setminus p_{\widetilde{X}})(w)}{p_{\widetilde{X}}(v_1w)}$$

By the ISF (see above), $(\mathcal{C} \setminus p_{\widetilde{X}})$ is a subexpression of the category in one of the axioms. The only possibility is that $(\mathcal{C} \setminus p_{\widetilde{X}})$ is a subexpression $((p_{\widetilde{Y}_1} \wedge \ldots \wedge p_{\widetilde{Y}_k}) \setminus p_{\widetilde{X}})$ of an axiom $((p_{\widetilde{Y}_1} \wedge \ldots \wedge p_{\widetilde{Y}_k}) \setminus p_{\widetilde{X}}) / (p_{\widetilde{Z}_1} \wedge \ldots \wedge p_{\widetilde{Z}_m})$. Moreover, again by the ISF, the only way to derive $(\mathcal{C} \setminus p_{\widetilde{X}})(w)$ is to apply the right division rule to the category used in the axiom. This analysis shows that the derivation must end in the way depicted in the earlier Figure 4, where $w = av_2$.

Since the rules for $\wedge$ are invertible (see above), the propositions $p_{\widetilde{Y}_1}(v_1)$, ..., $p_{\widetilde{Y}_k}(v_1)$, $p_{\widetilde{Z}_1}(v_2)$, ..., $p_{\widetilde{Z}_m}(v_2)$ are derivable. By induction hypothesis, $\widetilde{Y}_i(v_1)$ and $\widetilde{Z}_j(v_2)$, for all i, j, are derivable in the original conjunctive grammar. Then, the derivation uses the rules $\widetilde{X} \to BaC$, $B \to \widetilde{Y}_1 \,\&\, \ldots \,\&\, \widetilde{Y}_k$, and $C \to \widetilde{Z}_1 \,\&\, \ldots \,\&\, \widetilde{Z}_m$, and is of the following form.

$$\frac{\dfrac{\widetilde{Y}_1(v_1) \;\ldots\; \widetilde{Y}_k(v_1)}{B(v_1)} \qquad a(a) \qquad \dfrac{\widetilde{Z}_1(v_2) \;\ldots\; \widetilde{Z}_m(v_2)}{C(v_2)}}{\widetilde{X}(v_1av_2)}$$

The *right division* case is even easier. Here a derivation ends as follows.

$$\frac{(p_{\widetilde{Y}} / \mathcal{C})(w) \quad \mathcal{C}(v_1)}{p_{\widetilde{Y}}(wv_1)}$$

By the ISF, the left premise could be nothing but an axiom of the form $\big(p_{\widetilde{Y}} / (p_{\widetilde{X}_1} \wedge \ldots \wedge p_{\widetilde{X}_k})\big)(a)$ (and $w = a$). Then, $\mathcal{C}(v_1)$ is $(p_{\widetilde{X}_1} \wedge \ldots \wedge p_{\widetilde{X}_k})(v_1)$, and by the invertibility of the $\wedge$ rule, all $p_{\widetilde{X}_i}(v_1)$ are derivable. By the induction hypothesis, $\widetilde{X}_i(v_1)$, for all i, are derivable in the original conjunctive grammar, and there is the following derivation for $\widetilde{Y}(av_1)$, using the rules $\widetilde{Y} \to aA$ and $A \to \widetilde{X}_1 \,\&\, \ldots \,\&\, \widetilde{X}_k$.

$$\frac{a(a) \qquad \dfrac{\widetilde{X}_1(v_1) \quad \ldots \quad \widetilde{X}_k(v_k)}{A(v_1)}}{\widetilde{Y}(av_1)}$$

□

This claim immediately yields the main result, since $\widetilde{S}(w)$ is derivable in the original conjunctive grammar if and only if $p_{\widetilde{S}}(w)$ is derivable in the constructed conjunctive categorial grammar. □

5 Conjunctive Categorial Grammars and Lambek Grammars with Additives

Lambek (1958) suggested a richer logic as a background for categorial grammars, called *the Lambek calculus*. In the Lambek calculus, or **L** for short, syntactic categories built from a set of $\mathbf{Pr} =$

$$\dfrac{(p_{\widetilde{Y}} \,/\, (p_{\widetilde{X}_1} \wedge \ldots \wedge p_{\widetilde{X}_k}))(a) \qquad \dfrac{p_{\widetilde{X}_1}(v_1) \quad \ldots \quad p_{\widetilde{X}_k}(v_1)}{(p_{\widetilde{X}_1} \wedge \ldots \wedge p_{\widetilde{X}_k})(v_1)}}{p_{\widetilde{Y}}(av_1)}$$

Figure 3

$$\dfrac{(p_{\widetilde{Y}_1} \wedge \ldots \wedge p_{\widetilde{Y}_k})(v_1) \qquad \dfrac{(((p_{\widetilde{Y}_1} \wedge \ldots \wedge p_{\widetilde{Y}_k}) \setminus p_{\widetilde{X}}) \,/\, (p_{\widetilde{Z}_1} \wedge \ldots \wedge p_{\widetilde{Z}_m}))(a) \quad (p_{\widetilde{Z}_1} \wedge \ldots \wedge p_{\widetilde{Z}_m})(v_2)}{((p_{\widetilde{Y}_1} \wedge \ldots \wedge p_{\widetilde{Y}_k}) \setminus p_{\widetilde{X}})(av_2)}}{p_{\widetilde{X}}(v_1 a v_2)}$$

Figure 4

$\{p_1, p_2, p_3, \ldots\}$ of primitive categories using three binary operations: product ($\cdot$), which means concatenation, left division ($\backslash$), and right division ($/$). The formal recursive definition is as follows.

1. Every primitive category is a category.
2. If A and B are categories, then $(A \cdot B)$, $(A \setminus B)$, and $(B \,/\, A)$ are also categories.

The set of all Lambek categories is denoted by $\mathbf{Cat}$. As opposed to basic categories, deep nesting of division operations is allowed here, that is denominators are allowed to be non-primitive.

A Lambek categorial grammar consists of a target category $s \in \mathbf{Cat}$ (usually s is required to be a primitive category) and a finite number of axiomatic propositions of the form $A(a)$, where A is a category and a is a letter of the alphabet.

A string $w = a_1 \ldots a_n$ is considered accepted by the grammar, if, for some categories A_1, $\ldots$, A_n, the propositions $A_i(a_i)$ are included in the grammar as axiomatic ones, and the *sequent* $A_1, \ldots, A_n \to s$ is derivable in the Lambek calculus, which consists of the axioms and inference rules listed in Figure 5. In all rules, left-hand sides of the sequents are required to be non-empty.

Note that in Lambek grammars, arrows tranditionally point in an opposite direction than in context-free grammars ($\ldots \to s$ vs. $S \to \ldots$).

The following *cut rule* is not officially included in the system, but is admissible (Lambek, 1958).

$$\dfrac{\Pi \to A \quad \Gamma, A, \Delta \to D}{\Gamma, \Pi, \Delta \to D} \ (\mathrm{cut})$$

As one can easily see, all basic categories, as defined in Section 2, are also Lambek categories: $\mathbf{BCat} \subset \mathbf{Cat}$. Moreover, as noticed by Buszkowski (1985), if a basic categorial grammar is regarded as a Lambek categorial grammar with the same set of axiomatic propositions, then it describes the same language.

Next, the Lambek calculus is extended with the so-called "additive" conjunction and disjunction, as defined by Kanazawa (1992). These new operations correspond to the additive operations in linear logic by Girard (1987). Inference rules for these operations are depicted in Figure 6.

This calculus, denoted by $\mathbf{MALC}$ ("multiplicative-additive Lambek calculus"), also enjoys cut elimination and the subformula property.

Lambek categories with $\wedge$ and $\vee$ generalize conjunctive categories (and conjuncts):

$$\mathbf{BCat}_\wedge \cup \mathbf{Conj} \subset \mathbf{Cat}_{\wedge,\vee},$$

and every conjunctive categorial grammar can be translated into a Lambek grammar with $\wedge$ and $\vee$. However, one cannot simply take the axiomatic propositions of a conjunctive categorial grammar and use them as axiomatic propositions in the sense of Lambek grammars: this would yield a grammar that is not equivalent to the original one (for instance, the Lambek grammar with the axiomatic propositions from Example 4 does not accept any strings at all). The construction has to be more subtle.

Theorem 2. *Let $\Sigma = \{a_1, \ldots, a_n\}$ and consider a conjunctive categorial grammar with the following axiomatic propositions.*

$$\begin{array}{llll} A_{1,1}(a_1), & A_{1,2}(a_1), & \ldots & A_{1,k_1}(a_1), \\ A_{2,1}(a_2), & A_{2,2}(a_2), & \ldots & A_{2,k_2}(a_2), \\ \vdots & & & \\ A_{n,1}(a_n), & A_{n,2}(a_n), & \ldots & A_{n,k_n}(a_n). \end{array}$$

$$A \to A$$

$$\frac{\Pi \to A \quad \Gamma, B, \Delta \to D}{\Gamma, \Pi, A \setminus B, \Delta \to D}\ (\setminus \to) \qquad \frac{A, \Pi \to B}{\Pi \to A \setminus B}\ (\to \setminus) \qquad \frac{\Gamma, A, B, \Delta \to D}{\Gamma, A \cdot B, \Delta \to D}\ (\cdot \to)$$

$$\frac{\Pi \to A \quad \Gamma, B, \Delta \to D}{\Gamma, B \,/\, A, \Pi, \Delta \to D}\ (/ \to) \qquad \frac{\Pi, A \to B}{\Pi \to B \,/\, A}\ (\to /) \qquad \frac{\Gamma \to A \quad \Delta \to B}{\Gamma, \Delta \to A \cdot B}\ (\to \cdot)$$

Figure 5: The Lambek Calculus

$$\frac{\Gamma, A_1, \Delta \to D}{\Gamma, A_1 \wedge A_2, \Delta \to D}\ (\wedge \to)_1 \qquad \frac{\Gamma, A_2, \Delta \to D}{\Gamma, A_1 \wedge A_2, \Delta \to D}\ (\wedge \to)_2 \qquad \frac{\Pi \to A_1 \quad \Pi \to A_2}{\Pi \to A_1 \wedge A_2}\ (\to \wedge)$$

$$\frac{\Gamma, A_1, \Delta \to D \quad \Gamma, A_2, \Delta \to D}{\Gamma, A_1 \vee A_2, \Delta \to D}\ (\vee \to) \qquad \frac{\Pi \to A_1}{\Pi \to A_1 \vee A_2}\ (\to \vee)_1 \qquad \frac{\Pi \to A_2}{\Pi \to A_1 \vee A_2}\ (\to \vee)_2$$

Figure 6: Rules for Conjunction and Disjunction

Then the Lambek grammar with atomic propositions $(A_{i,1} \wedge A_{i,2} \wedge \ldots \wedge A_{i,k_i})(a_i)$ (for $i = 1, \ldots, n$) describes the same language as the original conjunctive categorial grammar. (If $k_i = 1$, we take just $A_{i,1}(a_i)$.)

Proof. Let $B_i = A_{i,1} \wedge A_{i,2} \wedge \ldots \wedge A_{i,k_i}$. The new Lambek grammar uses axiomatic propositions of the form $B_i(a_i)$, one for each symbol in Σ. It is sufficient to prove the following: for the target category $s \in \mathbf{Pr}$ and for a string $a_{i_1} \ldots a_{i_m}$, the proposition $s(a_{i_1} \ldots a_{i_m})$ is derivable in the conjunctive categorial grammar if and only if the sequent $B_{i_1}, \ldots, B_{i_m} \to s$ is derivable in $\mathbf{MALC}$.

The "only if" part. In order to use induction on the length of derivation in the conjunctive categorial grammar, the statement is proved not only for s, but for an arbitrary category $D \in \mathbf{BCat}_\wedge \cup \mathbf{Conj}$.

The proof in the base case is immediate: if $D(a_i)$ is an axiom, then D is one of the $A_{i,j}$ in the conjunction B_i, and the sequent $B_i \to A_{i,j}$ is derivable by several applications of the $(\wedge \to)$ rules.

For the induction step, there are three cases.

Case 1: $D = (p_1 \wedge \ldots \wedge p_k)$. Then, by the induction hypothesis, $B_{i_1}, \ldots, B_{i_m} \to p_j$ is derivable in $\mathbf{MALC}$ for every j, and $B_{i_1}, \ldots, B_{i_m} \to p_1 \wedge \ldots \wedge p_k$ is derived by the $(\to \wedge)$ rule.

Case 2: $D(a_{i_1} \ldots a_{i_m})$ is derived from $\mathcal{C}(a_{i_1} \ldots a_{i_\ell})$ and $(\mathcal{C} \setminus D)(a_{i_{\ell+1}} \ldots a_{i_m})$ for some $\mathcal{C} \in \mathbf{Conj}$. Then, by the induction hypothesis, the sequents $B_{i_1}, \ldots, B_{i_\ell} \to \mathcal{C}$ and $B_{i_{\ell+1}}, \ldots, B_{i_m} \to \mathcal{C} \setminus D$ are derivable, and then $B_{i_1}, \ldots, B_{i_m} \to D$ can be derived in the following way. First, $B_{i_1}, \ldots, B_{i_\ell}, \mathcal{C} \setminus D \to D$ is derived from $B_{i_1}, \ldots, B_{i_\ell} \to \mathcal{C}$ and $D \to D$, and then it is combined with $B_{i_{\ell+1}}, \ldots, B_{i_m} \to \mathcal{C} \setminus D$ using the cut rule, to get $B_{i_1}, \ldots, B_{i_\ell}, B_{i_{\ell+1}}, \ldots, B_{i_m} \to D$.

Case 3: $D(a_{i_1} \ldots a_{i_m})$ is derived from $(D \,/\, \mathcal{C})(a_{i_1} \ldots a_{i_\ell})$ and $\mathcal{C}(a_{i_{\ell+1}} \ldots a_{i_m})$. The proof is symmetric.

The "if" part. The following more general statement is claimed. *For every $j = 1, \ldots, m$, let B'_{i_j} be a conjunction of an arbitrary subset of formulae $A_{i_j,k}$ used in the conjunction B_{i_j}; in other words, B'_{i_j} may coincide with B_{i_j} or lack some of the conjuncts. Then, for any $\mathcal{C} \in \mathbf{Conj}$ (in particular, for $\mathcal{C} = s \in \mathbf{Pr} \subset \mathbf{Conj}$), if $B'_{i_1}, \ldots, B'_{i_m} \to \mathcal{C}$ is derivable in $\mathbf{MALC}$, then the proposition $\mathcal{C}(a_{i_1} \ldots a_{i_m})$ is derivable in the original conjunctive categorial grammar.*

The claim is proved by induction on the cut-free derivation of the sequent $B'_{i_1}, \ldots, B'_{i_m} \to \mathcal{C}$ in $\mathbf{MALC}$.

Case 1. $\mathcal{C} = p_1 \wedge \ldots \wedge p_k$, $k \geqslant 2$. Since the $(\to \wedge)$ rule in $\mathbf{MALC}$ is invertible (this follows from the cut elimination), it can be assumed that all $k-1$

applications of this rule were applied immediately.

$$\frac{B'_{i_1}, \dots, B'_{i_m} \to p_1 \quad \dots \quad B'_{i_1}, \dots, B'_{i_m} \to p_k}{B'_{i_1}, \dots, B'_{i_m} \to p_1 \wedge \dots \wedge p_k}$$

Then, by the induction hypothesis, all propositions $p_j(a_1 \dots a_m)$ are derivable in the conjunctive categorial grammar, and from them one can derive $(p_1 \wedge \dots \wedge p_k)(a_1 \dots a_m)$.

In all other cases, $\mathcal{C} \in \mathbf{Pr}$.

Case 2: an axiom. Then, $m = 1$, $B'_{i_1} = \mathcal{C}$, and, since all elements of B'_{i_1} should be of the form $A_{i_1,k}$, the proposition $\mathcal{C}(a_1)$ is an axiom of the conjunctive categorial grammar, and therefore derivable.

Case 3: the last rule of the derivation is $(\wedge \to)$. Then, $B'_{i_\ell} = B''_{i_\ell} \wedge A_{i_\ell,k}$:

$$\frac{B'_{i_1}, \dots, B''_{i_\ell}, \dots, B'_{i_m} \to \mathcal{C}}{B'_{i_1}, \dots, B''_{i_\ell} \wedge A_{i_\ell,k}, \dots, B'_{i_m} \to \mathcal{C}}$$

or

$$\frac{B'_{i_1}, \dots, A_{i_\ell,k}, \dots, B'_{i_m} \to \mathcal{C}}{B'_{i_1}, \dots, B''_{i_\ell} \wedge A_{i_\ell,k}, \dots, B'_{i_m} \to \mathcal{C}}$$

In both cases the induction hypothesis is applied: since B''_{i_ℓ} or $A_{i_\ell,k}$ can act as B'_{i_ℓ}, the proposition $\mathcal{C}(a_{i_1} \dots a_{i_\ell} \dots a_{i_m})$ is derivable in the conjunctive categorial grammar.

Case 4: the last rule is $(\backslash \to)$. In this case, $B'_{i_h} = A_{i_h,k} = \mathcal{C}' \backslash A'$, for some h and for $\mathcal{C}' \in \mathbf{Conj}$ and $A' \in \mathbf{BCat}_\wedge$, and the sequent $B'_{i_1}, \dots, B'_{i_{\ell-1}}, B'_{i_\ell}, \dots, B'_{i_{h-1}}, \mathcal{C}' \backslash A', B'_{i_{h+1}}, \dots, B'_{a_m} \to \mathcal{C}$ is derived from $B'_{i_\ell}, \dots, B'_{i_{h-1}} \to \mathcal{C}'$ and $B'_{i_1}, \dots, B'_{i_{\ell-1}}, A', B'_{i_{h+1}}, \dots, B'_{i_m} \to \mathcal{C}$. By the induction hypothesis, the proposition $\mathcal{C}'(a_{i_\ell} \dots a_{i_{h-1}})$ can be derived in the conjunctive categorial grammar, and, since $(\mathcal{C}' \backslash A')(a_{i_h})$ is an axiom, the proposition $A'(a_{i_\ell} \dots a_{i_{h-1}} a_{i_h})$ is also derivable.

Now, the conjunctive categorial grammar is extended by adding a new symbol a_{n+1} to the original alphabet $\Sigma = \{a_1, \dots, a_n\}$, with a new axiom, $A'(a_{n+1})$. For the new grammar, we have the same B_j for $j = 1, \dots, n$, and $B_{n+1} = A'$. Since $B'_{i_1}, \dots, B'_{i_{\ell-1}}, A', B'_{i_{h+1}}, \dots, B'_{i_m} \to \mathcal{C}$ is derivable in $\mathbf{MALC}$, by the induction hypothesis, the proposition $\mathcal{C}(a_{i_1} \dots a_{i_{\ell-1}} a_{n+1} a_{i_{h+1}} \dots a_{i_m})$ is derivable in the extended conjunctive categorial grammar.

By Lemma 1, the desired proposition $\mathcal{C}(a_{i_1} \dots a_{i_{\ell-1}} a_{i_\ell} \dots a_{i_{h-1}} a_{i_h} a_{i_{h+1}} \dots a_{i_m})$, where the string $u = a_{i_\ell} \dots a_{i_{h-1}} a_{i_h}$ has been substituted for a fresh symbol $b = a_{n+1}$, can be derived in the original conjunctive categorial grammar.

Case 5: the last rule is $(/ \to)$. Symmetric. □

This embedding immediately implies that every language generated by a conjunctive grammar can be generated by an $\mathbf{MALC}$-grammar. This supersedes the result by Kuznetsov (2013).

In the classical case without the conjunction, a converse result was shown by Pentus (1993): every language generated by a Lambek grammar is context-free. Whether an analogous property holds for $\mathbf{MALC}$ (that is, whether every $\mathbf{MALC}$-language is generated by a conjunctive grammar) remains an open problem. Establishing any such upper bound on the power of the new model would require proving a non-trivial variant of the famous theorem by Pentus (1993), which would likely be difficult.

However, there is some evidence that $\mathbf{MALC}$ should be strictly more powerful than conjunctive grammars. First, there is a result by Okhotin (2011) that conjunctive grammars can describe a certain P-complete language representing the Circuit Value Problem (CVP) under a suitable encoding. On the other hand, the class of languages generated by $\mathbf{MALC}$-grammars is, by definition, closed under symbol-to-symbol homomorphisms. These two facts are sufficient to develop a $\mathbf{MALC}$ represenation for an NP-complete language, which is the last result of this paper.

Theorem 3. *The family of languages generated by* $\mathbf{MALC}$*-grammars contains an* NP*-complete language.*

Sketch of proof. It is not difficult to transform the grammar for the CVP given by Okhotin (2011), so that each CVP instance is represented in the form $u_{k,C}v$, where $u_{k,C} \in \Sigma^*$ is a description of a circuit C with k inputs, while $v \in \{0,1\}^k$ contains the input values, and $0, 1 \notin \Sigma$. The grammar then describes the set of all such strings, on which the circuit evaluates to 1 on the given input values.

$$\text{CVP} = \{u_{k,C}v \mid C(v) = 1\}$$

Let $h \colon \Sigma \cup \{0,1\} \to \Sigma \cup \{?\}$ be a homomorphism that maps both digits to the question mark symbol, leaving all other symbols intact: $h(0) = h(1) = ?$, $h(a) = a$ for all $a \in \Sigma$. This transforms the

Circuit Value Problem to the Circuit Satisfiability Problem, which is NP-complete.

$$h(\mathrm{CVP}) = \{u_{k,C}?^k \mid \exists v \in \{0,1\}^k : C(v) = 1\}$$

Since the language CVP is described by a conjunctive grammar, by Theorem 1, it is also described by a conjunctive categorial grammar, and then, by Theorem 2, also by an $\mathbf{MALC}$-grammar. Next, as observed by Kanazawa (1992), its symbol-to-symbol homomorphic image $h(\mathrm{CVP})$ must have an $\mathbf{MALC}$-grammar as well. □

On the other hand, every language described by a conjunctive grammar can be parsed in polynomial time—to be exact, in time $O(n^\omega)$, where $\omega < 3$ is the exponent in the complexity of matrix multiplication (Okhotin, 2014). This leads to the following corollary.

Corollary 1. *Under the assumption that* $\mathrm{P} \neq \mathrm{NP}$*, conjunctive categorial grammars are strictly weaker in power than* $\mathbf{MALC}$.

It would be interesting to establish an unconditional separation of these two classes.

Acknowledgments

Stepan Kuznetsov's research was supported by the Russian Foundation for Basic Research (grants 15-01-09218-a and 14-01-00127-a) and the Presidential Council for Support of Leading Scientific Schools of Russia (grant NŠ-9091.2016.1).

References

V. M. Abrusci. 1990. A comparison between Lambek syntactic calculus and intuitionistic linear propositional logic. *Zeitschrift für mathematische Logik und Grundlagen der Mathematik* 36(1):11–15.

K. Ajdukiewicz. 1935. Die syntaktische Konnexität. *Studia Philosophica* 1:1–27.

Y. Bar-Hillel, C. Gaifman, and E. Shamir. 1960. On categorial and phrase-structure grammars. *Bulletin of the Research Council of Israel, Section F* 9F:1–16.

W. Buszkowski. 1985. The equivalence of unidirectional Lambek categorial grammars and context-free grammars. *Zeitschrift für mathematische Logik und Grundlagen der Mathematik* 31:369–384.

W. Buszkowski. 1996. Categorial grammars with negative information. In H. Wansing, editor, *Negation. A Notion in Focus*, De Gruyter, pages 107–126.

B. Carpenter. 1997. *Type-Logical Semantics*. MIT Press, Cambridge, MA.

N. Chomsky. 1956. Three models for the description of language. *IRE Transactions on Information Theory* 2:113–124.

M. Dekhtyar and A. Dikovsky. 2008. Generalized categorial dependency grammars. In *Pillars of Computer Science (Trakhtenbrot Festschrift)*, Springer, volume 4800 of *LNCS*, pages 230–255.

M. J. Fischer. 1968. Grammars with macro-like productions. In *Proc. 9th Annual Symposium on Switching and Automata Theory (SWAT 1968)*. New York, pages 131–142.

J.-Y. Girard. 1987. Linear logic. *Theoretical Computer Science* 50(1):1–102.

M. Kanazawa. 1992. The lambek calculus enriched with additional connectives. *Journal of Logic, Language, and Information* 1:141–171.

S. Kuznetsov. 2013. Conjunctive grammars in greibach normal form and the lambek calculus with additive connectives. In *Proc. Formal Grammar 2012 and Formal Grammar 2013*. Springer, volume 8036 of *LNCS*, pages 242–249.

S. L. Kuznetsov. 2016. On translating Lambek grammars with one division into context-free grammars. *Proceedings of Steklov Institute of Mathematics* 294:129–138.

J. Lambek. 1958. The mathematics of sentence structure. *American Mathematical Monthly* 65:154–170.

M. Latta and R. Wall. 1993. Intersective context-free languages. In *Lenguajes Naturales y Lenguajes Formales IX*. Barcelona, pages 15–43.

M. Moortgat. 1996. Multimodal linguistic inference. *Journal of Logic, Language, and Information* 5(3, 4):349–385.

R. Moot and C. Retoré. 2012. *The Logic of Categorial Grammars: A Deductive Account of Natural Language Syntax and Semantics*. Springer, Heidelberg.

G. V. Morrill. 2011. *Categorial Grammar: Logical Syntax, Semantics, and Processing*. Oxford University Press.

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

A. Okhotin. 2004. Boolean grammars. *Information and Computation* 194(1):19–48.

A. Okhotin. 2011. A simple p-complete problem and its language-theoretic representations. *Theoretical Computer Science* 412(1–2):68–82.

A. Okhotin. 2013. Conjunctive and boolean grammars: the true general case of the context-free grammars. *Computer Science Review* 9:27–59.

A. Okhotin. 2014. Parsing by matrix multiplication generalized to boolean grammars. *Theoretical Computer Science* 516:101–120.

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

M. Pentus. 1993. Lambek grammars are context-free. In *Proc. 8th Annual IEEE Symposium on Logic in Computer Science (LICS 1993)*. Montreal, pages 430–433.

Yu. V. Savateev. 2009. Recognition of derivability for the lambek calculus with one division. *Moscow University Mathematics Bulletin* 64(2):73–75.

H. Seki, T. Matsumura, M. Fujii, and T. Kasami. 1991. On multiple context-free grammars. *Theoretical Computer Science* 88(2):191–229.

M. Steedman. 1996. *Surface structure and interpretation*. MIT Press.

K. Vijay-Shanker, D. J. Weir, and A. K. Joshi. 1987. Characterizing structural descriptions produced by various grammatical formalisms. In *Proc. 25th Annual Meeting of the Association for Computational Linguistics (ACL 1987)*. pages 104–111.

D. N. Yetter. 1990. Quantales and (noncommutative) linear logic. *Journal of Symbolic Logic* 55(1):41–64.