S4 Enriched Multimodal Categorial Grammars are Context-free: Corrigendum

Andrew R. Plummer

Department of Linguistics
The Ohio State University
222 Oxley Hall, 1712 Neil Avenue
Columbus, OH 43210, USA

Abstract

Plummer [3] showed that categorial grammars based on $\mathbf{NL}_{\mathrm{S4}}$, the non-associative multimodal Lambek Calculus enriched with S4 axioms, weakly recognize context-free languages. However, the proof contains a gap. We correct the earlier proof, utilizing a technique given in Buszkowski [1]. This technique immediately proves that $\mathbf{NL}_{\mathrm{S4}}$ is decidable in polynomial time.

1 $\mathbf{NL}_{\mathrm{S4}}$-grammars

Plummer [3] showed that categorial grammars based on $\mathbf{NL}_{\mathrm{S4}}$, the non-associative multimodal Lambek Calculus enriched with S4 axioms, weakly recognize context-free languages. However, the proof contains a gap. In the proof of Lemma 2.4 on p. 178, for rule 4, it is not true that an interpolant for Δ in the premise serves as an interpolant for Δ in $\Gamma[\Delta] \Rightarrow A$. The case $\Delta = \langle \Delta' \rangle$ is not treated properly. We cannot infer $\Gamma[\langle B \rangle] \Rightarrow A$ from $\Gamma[B] \Rightarrow A$. We correct the earlier proof, utilizing a technique given in Buszkowski [1].

We write $\mathbf{NL}_{\mathrm{S4}} \vdash \Gamma \Rightarrow A$ if the $\mathbf{NL}_{\mathrm{S4}}$-sequent $\Gamma \Rightarrow A$ is provable in $\mathbf{NL}_{\mathrm{S4}}$. Moortgat [2] proved Cut-elimination, the subformula property, and decidability for $\mathbf{NL}_{\mathrm{S4}}$. Let $\mathcal{T}$ be a finite set of formulas closed under subformulas. Let $\mathcal{T}' = \{\Diamond M \mid M \in \mathcal{T}\} \cup \mathcal{T}$. By a $\mathcal{T}'$-sequent we mean a sequent $\Gamma \Rightarrow A$ such that A and all formulas appearing in Γ belong to $\mathcal{T}'$. We write $\Gamma \Rightarrow_{\mathcal{T}'} A$ if $\Gamma \Rightarrow A$ has a proof in $\mathbf{NL}_{\mathrm{S4}}$ consisting of $\mathcal{T}'$-sequents only.

Since $\mathbf{NL}_{\mathrm{S4}}$ has the subformula property, every $\mathcal{T}'$-sequent provable in $\mathbf{NL}_{\mathrm{S4}}$ has a proof in $\mathbf{NL}_{\mathrm{S4}}$ such that all sequents appearing in this proof are $\mathcal{T}'$-sequents. We shall describe an effective procedure which produces all $\mathcal{T}'$-sequents $(A, B) \Rightarrow C$, $\langle A \rangle \Rightarrow B$, and $A \Rightarrow B$ which are provable in $\mathbf{NL}_{\mathrm{S4}}$. Furthermore, we show that every $\mathcal{T}'$-sequent provable in $\mathbf{NL}_{\mathrm{S4}}$ can be derived from these sequents by *Cut* only. We first prove an interpolation lemma for $\mathbf{NL}_{\mathrm{S4}}$-sequents.

Lemma 1.1. *Let S be a sequent $\Gamma[\Delta] \Rightarrow C$ provable in $\mathbf{NL}_{\mathrm{S4}}$. Let $\mathcal{T}_S$ be the set of formulas containing C and all formulas in Γ such that $\mathcal{T}_S$ is closed under subformulas, and let $\mathcal{T}'_S =$*

$\{\Diamond M \mid M \in \mathcal{T}_S\} \cup \mathcal{T}_S$. *Then there is a type* $D \in \mathcal{T}'_S$ *such that* $\mathbf{NL}_{\mathrm{S4}} \vdash \Delta \Rightarrow D$ *and* $\mathbf{NL}_{\mathrm{S4}} \vdash \Gamma[D] \Rightarrow C$.

Proof. The proof is by induction over cut-free sequent derivations. We provide details for the only case requiring attention. Assume the rule is 4. Suppose $\Delta = \langle\Delta'\rangle$. Let S' be the sequent $\Gamma[\langle\Delta'\rangle] \Rightarrow C$. By the induction hypothesis, there is a type $D \in \mathcal{T}'_{S'}$ such that $\Delta' \Rightarrow D$ and $\Gamma[\langle D\rangle] \Rightarrow C$.

Case 1. Suppose $D \in \mathcal{T}_{S'}$. Then $\Diamond D \in \mathcal{T}'_{S'}$. By applying 4 to $\Gamma[\langle D\rangle] \Rightarrow C$, we have $\Gamma[\langle\langle D\rangle\rangle] \Rightarrow C$. By applying $\Diamond L$, we have $\Gamma[\langle\Diamond D\rangle] \Rightarrow C$. By applying $\Diamond R$ to $\Delta' \Rightarrow D$, we have $\langle\Delta'\rangle \Rightarrow \Diamond D$. Since $\Diamond D \in \mathcal{T}'_{S'}$, then $\Diamond D \in \mathcal{T}'_S$. Hence $\Diamond D$ is an interpolant for Δ.

Case 2. Suppose $D = \Diamond E$, where $E \in \mathcal{T}_{S'}$. Hence, $\Delta' \Rightarrow \Diamond E$. By applying $\Diamond R$ to $\Delta' \Rightarrow \Diamond E$, we have $\langle\Delta'\rangle \Rightarrow \Diamond\Diamond E$. Since $\mathbf{NL}_{\mathrm{S4}} \vdash \Diamond\Diamond E \Rightarrow \Diamond E$, by *Cut* we have $\langle\Delta'\rangle \Rightarrow \Diamond E$. Hence, $\langle\Delta'\rangle \Rightarrow D$. Since $D \in \mathcal{T}'_{S'}$, then $D \in \mathcal{T}'_S$. Hence D is an interpolant for Δ. □

Let $S^{\mathcal{T}'}$ be the union of the sets

$$\{A \Rightarrow B \mid \mathbf{NL}_{\mathrm{S4}} \vdash A \Rightarrow B \text{ and } A, B \in \mathcal{T}'\},$$
$$\{\langle A\rangle \Rightarrow B \mid \mathbf{NL}_{\mathrm{S4}} \vdash \langle A\rangle \Rightarrow B \text{ and } A, B \in \mathcal{T}'\},$$
$$\{(A, B) \Rightarrow C \mid \mathbf{NL}_{\mathrm{S4}} \vdash (A, B) \Rightarrow C \text{ and } A, B, C \in \mathcal{T}'\}.$$

Clearly, $S^{\mathcal{T}'}$ is finite. Let $S(\mathcal{T}')$ be the closure of $S^{\mathcal{T}'}$ under *Cut*. We write $\Gamma \Rightarrow_{S(\mathcal{T}')} A$ if $\Gamma \Rightarrow A$ is provable in $S(\mathcal{T}')$.

Lemma 1.2. *For any* $\mathcal{T}'$*-sequent* $\Gamma \Rightarrow C$, $\Gamma \Rightarrow_{\mathcal{T}'} C$ *if and only if* $\Gamma \Rightarrow_{S(\mathcal{T}')} C$.

Proof. The nontrivial part of the proof is by induction on the number of structural operators, $(\cdot,\cdot)$ and $\langle\cdot\rangle$, in Γ. We provide details for the case $\Gamma = \Delta[\langle B\rangle]$, where B is a type. Let S be the sequent $\Gamma \Rightarrow C$. By Lemma 1.1, there is an interpolant $D \in \mathcal{T}'_S$ for $\langle B\rangle$ in $\Delta[\langle B\rangle] \Rightarrow C$. Moreover, $D \in \mathcal{T}'$ or $D = \Diamond\Diamond E$ where $E \in \mathcal{T}$.

Case 1. Suppose $D \in \mathcal{T}'$. Then $\langle B\rangle \Rightarrow_{S(\mathcal{T}')} D$. Since $\mathbf{NL}_{\mathrm{S4}} \vdash \Delta[D] \Rightarrow C$ where C and every formula in $\Delta[D]$ is in $\mathcal{T}'$, it follows that $\Delta[D] \Rightarrow_{\mathcal{T}'} C$. By the induction hypothesis, $\Delta[D] \Rightarrow_{S(\mathcal{T}')} C$. Applying *Cut* to the premises $\langle B\rangle \Rightarrow_{S(\mathcal{T}')} D$ and $\Delta[D] \Rightarrow_{S(\mathcal{T}')} C$, we have that $\Delta[\langle B\rangle] \Rightarrow_{S(\mathcal{T}')} C$.

Case 2. Suppose $D = \Diamond\Diamond E$ where $E \in \mathcal{T}$. Then $\mathbf{NL}_{\mathrm{S4}} \vdash \langle B\rangle \Rightarrow \Diamond\Diamond E$ and $\mathbf{NL}_{\mathrm{S4}} \vdash \Delta[\Diamond\Diamond E] \Rightarrow C$. Since $\mathbf{NL}_{\mathrm{S4}} \vdash \Diamond\Diamond E \Rightarrow \Diamond E$, by *Cut* we have $\mathbf{NL}_{\mathrm{S4}} \vdash \langle B\rangle \Rightarrow \Diamond E$. Since B and $\Diamond E$ are in $\mathcal{T}'$, it follows that $\langle B\rangle \Rightarrow_{S(\mathcal{T}')} \Diamond E$. Since $\mathbf{NL}_{\mathrm{S4}} \vdash \Diamond E \Rightarrow \Diamond\Diamond E$, by *Cut* we have $\mathbf{NL}_{\mathrm{S4}} \vdash \Delta[\Diamond E] \Rightarrow C$. Since C and every formula in $\Delta[\Diamond E]$ is in $\mathcal{T}'$, it follows that $\Delta[\Diamond E] \Rightarrow_{\mathcal{T}'} C$. By the induction hypothesis, $\Delta[\Diamond E] \Rightarrow_{S(\mathcal{T}')} C$. Applying *Cut* to the premises $\langle B\rangle \Rightarrow_{S(\mathcal{T}')} \Diamond E$ and $\Delta[\Diamond E] \Rightarrow_{S(\mathcal{T}')} C$, we have that $\Delta[\langle B\rangle] \Rightarrow_{S(\mathcal{T}')} C$. □

A *categorial grammar* based on a system S can be defined as a finite set of assignments $a \to A$ such that $a \in \Sigma$, Σ is an alphabet, and A is a formula. For a tree of formulas Γ, we denote by $s(\Gamma)$ the string of formulas which arises from Γ by dropping all occurrences of the structural operators and commas. For a categorial grammar G and a formula A, the language $L(G, A)$ consists of all strings $a_1 \ldots a_n$, for $n \geq 1$, satisfying the following conditions: there exist formulas A_i, $i = 1, \ldots, n$, and a tree of formulas Γ such that $s(\Gamma) = A_1 \ldots A_n$, all $a_i \to A_i$ belong to G, and $\Gamma \Rightarrow A$ is provable in S.

Theorem 1.3. *If G is a categorial grammar based on $\mathbf{NL}_{\mathrm{S4}}$, then for any formula A, $L(G, A)$ is a context-free language.*

Proof. Let $\mathcal{T}$ be the set of all subformulas of A and all subformulas of formulas appearing in G, and let $\mathcal{T}' = \{\Diamond M \mid M \in \mathcal{T}\} \cup \mathcal{T}$. For any $\mathcal{T}'$-sequent $\Gamma \Rightarrow A$, by the subformula property and Lemma 1.2, $\mathbf{NL}_{\mathrm{S4}} \vdash \Gamma \Rightarrow A$ if and only if $\Gamma \Rightarrow_{S(\mathcal{T}')} A$. By removing all structural operators, proofs in $S(\mathcal{T}')$ are derivations of a context-free grammar whose production rules are reversed sequents from $S^{\mathcal{T}'}$. We add lexical production rules $A \rightarrow a$ for $a \rightarrow A$ belonging to G. □

Corollary 1.4. $\mathbf{NL}_{\mathrm{S4}}$ *is decidable in polynomial time.*

Proof. Now, $\mathcal{T}$ is the set of all subformulas of formulas appearing in $\Gamma \Rightarrow A$. For a full proof see Buszkowski [1]. □

2 Acknowledgements

The author wishes to thank Wojciech Buszkowski for revealing the gap in the previous proof, and for all his help in correcting it.

References

[1] W. Buszkowski, Lambek calculus with nonlogical axioms, from *Language and Grammar: Studies in Mathematical Linguistics and Natural Language*, C. Casadio, P.J. Scott, R.A.G. Seely (eds.), pages 77–93, "Center for the Study of Language and Information", Lecture Notes 168, Stanford, 2005.

[2] M. Moortgat, Multimodal Linguistic Inference, *Journal of Logic, Language and Information* **5**, (1996) 349–385.

[3] A. R. Plummer, S4 enriched multimodal categorial grammars are context-free, *Theoretical Computer Science* **388**, (2007) 173–180.