The Turing Completeness of Multimodal Categorial Grammars

Bob Carpenter

May 25, 1999

Abstract

In this paper, we demonstrate that the multimodal categorial grammars are in fact Turing-complete in their weak generative capacity. The result follows from a straightforward reduction of generalized rewriting systems to a mixed associative and modal categorial calculus. We conclude with a discussion of a restriction to the so-caled weak Sahlqvist lexical rules, for which we can ensure decidability.

Contents

1 Introduction

In this paper, we demonstrate that the multimodal categorial grammars are in fact Turing-complete in their weak generative capacity. The result follows from a straightforward reduction of generalized rewriting systems to a mixed associative and modal categorial calculus.

It turns out that we should not be surprised that seemingly arbitrary kinds of operations can be coded in multimodal categorial grammars. In this paper, we show that any computable grammar can be coded as a multimodal categorial grammar. From the standpoint of formal linguistics, this opening of the computational floodgates might even appear to be inevitable. Simply compare the introduction of general transformations in transformational grammars [Peters and Ritchie 1973], metarules in phrase structure grammars [Uszkoreit and Peters 1986], and lexical rules in categorial and phrase structure systems [Carpenter 1991], all of which have been shown to be Turing-complete. Although steps may be taken to restrict the power of these systems to ensure decidability, such moves appear rather ad hoc because of their lack of linguistic motivation. For instance, consider the restrictions against metarule self application [Gazdar et al. 1985], or the finite bound placed on unary phrase structure rules (and by association, empty categories) by [Kaplan and Bresnan 1982].

Natural language syntax is a difficult matter, and no formalism has even come close to providing a universal system in which all and only natural language grammars can be expressed. Perhaps even more discouraging is the fact that no grammars for particular languages have ever been developed that even come close to covering a naturally occurring range of data in a theoretically clean fashion. On the other hand, the grammar fragments that are typically proposed in formalisms such as GB, HPSG, and LFG are much better behaved computationally than the worst-case analysis would suggest. But when observed phenomena begin to exceed the natural coverage of a formalism, more powerful mechanisms are typically introduced.

2 Multimodal Categorial Grammar

The general paradigm of multimodal categorial grammars was introduced by Moortgat [1994], following earlier multimodal developments by Hepple [1990], Morrill [1991, 1994], and Moortgat and Oehrle [1994]. The fundamental idea underlying these systems is that languages allow many different *modes* of combination for linguistic expressions (which have often been called *resources* in the literature, following common usage in linear logic). In addition to Lambek's [1958, 1961] original modes of associative and non-associative concatenation, several additional mechanisms have been proposed. For instance, commutative operations have been applied to free word-order languages [Hepple 1990; Moortgat and Oehrle 1994], wrapping and infixing have been used to deal with scoping and unbounded dependencies [Moortgat 1991; Morrill 1994, 1995] and to deal with gapping and ellipsis [Solias 1992]. In some systems, unary modes of combination have been used to deal with permuting for unbounded dependencies [Morrill 1994], islands and locality constraints [Hepple 1990; Morrill 1990, 1994], for encoding syntactic features [Kraak 1995], and for relaxing control in order to copy resources for parasitic gaps [Morrill 1994]. More recently, analyses of clitics [Kraak 1995] and general word-order domains [Versmissen 1996] have been proposed in even more elaborate multi-modal systems.

We will provide a sequent-based proof theoretic presentation of multimodal categorial grammar. There are several other presentations possible, including semantic ones, of the same logic [Morrill 1994; Moortgat 1994]. As usual, we begin with a finite set AtCat of *atomic category symbols*. We also assume a finite set UnMod of *unary modes* of combination and a finite set BinMod of *binary modes* of combination. We then define the set of Cat of *categories* to be the least such that:

- $\mathsf{AtCat} \subseteq \mathsf{Cat}$
- $\Box_u A, \Diamond_u A \in \mathsf{Cat}$ if $A \in \mathsf{Cat}$ and $u \in \mathsf{UnMod}$
- $A/_b B, B\backslash_b A, A{\cdot_b}B \in \mathsf{Cat}$ if $A, B \in \mathsf{Cat}$ and $b \in \mathsf{BinMod}$

The proof theory will be presented in Belnap's [1981] *display logic*, a generalization of Gentzen-style sequent proofs. The antecedents of sequents are structured according to their mode of combination. The collection Ant of *sequent antecedents* is the least such that:

- $\mathsf{Cat} \subseteq \mathsf{Ant}$
- $(\Gamma)^u \in \mathsf{Ant}$ if $\Gamma \in \mathsf{Ant}$ and $u \in \mathsf{UnMod}$
- $(\Gamma, \Delta)^b \in \mathsf{Ant}$ if $\Gamma, \Delta \in \mathsf{Ant}$ and $b \in \mathsf{BinMod}$

The set Seq of *sequents* is the least such that:

- $\Gamma \Rightarrow A \in \mathsf{Seq}$ if $\Gamma \in \mathsf{Ant}$ and $A \in \mathsf{Cat}$

The inference schemes then follow the general scheme of *residuation* for unary and binary operations [Moortgat 1994]. In terms of sequent rule schemes, this amounts to left and right rules for each of our binary and unary connectives. These are as follows.

- $$\frac{}{A \Rightarrow A}ID$$
- $$\frac{\Gamma[(A)^u] \Rightarrow B}{\Gamma[\Diamond_u A] \Rightarrow B}\Diamond_u L \qquad \frac{\Gamma \Rightarrow A}{(\Gamma)^a \Rightarrow \Diamond_u A}\Diamond_u R$$
- $$\frac{\Gamma[A] \Rightarrow B}{\Gamma[(\Box_u A)^u] \Rightarrow B}\Box_u L \qquad \frac{(\Gamma)^u \Rightarrow A}{\Gamma \Rightarrow \Box_u A}\Box_u R$$
- $$\frac{\Delta \Rightarrow B \qquad \Gamma[A] \Rightarrow C}{\Gamma[(A/_b B, \Delta)^b] \Rightarrow C}/_b L \qquad \frac{(\Gamma, B)^b \Rightarrow A}{\Gamma \Rightarrow A/_b B}/_b R$$
- $$\frac{\Delta \Rightarrow B \qquad \Gamma[A] \Rightarrow C}{\Gamma[(\Delta, B\backslash_b A)^b] \Rightarrow C}\backslash_b L \qquad \frac{(B, \Gamma)^b \Rightarrow A}{\Gamma \Rightarrow B\backslash_b A}\backslash_b R$$

- $$\frac{\Gamma(A,B)^b \Rightarrow C}{\Gamma[A \cdot_b B] \Rightarrow C}\cdot_b L \qquad \frac{\Gamma \Rightarrow A \qquad \Delta \Rightarrow B}{(\Gamma,\Delta)^b \Rightarrow A \cdot_b B}\cdot_b R$$

A *proof* of a sequent $\Gamma \Rightarrow A$ consists of a tree rooted at $\Gamma \Rightarrow A$, every local tree of which matches some inference rule, and every branch of which is terminated with an application of the identity scheme.

Lambek's [1961] non-associative calculus, often called **NL**, is given by simply taking $\mathsf{UnMod} = \{\}$ and $\mathsf{BinMod} = \{n\}$. The rules for the product and the left and right slash then correspond to Lambek's own presentation. **NL** is the weakest possible logic based on residuation, thus making it the one that is most sensitive to the structure of expressions. Lambek's [1958] associative calculus, often called **L**, on the other hand, can be defined by taking $\mathsf{UnMod} = \{\}$ and $\mathsf{BinMod} = \{a\}$, along with the following pair of *structural postulates* for *associativity*.

- $$\frac{\Gamma[((A,B)^a,C)^a] \Rightarrow D}{\Gamma[(A,(B,C)^a)^a] \Rightarrow D}A_1(a) \qquad \frac{\Gamma[(A,(B,C)^a)^a] \Rightarrow D}{\Gamma[((A,B)^a,C)^a] \Rightarrow D}A_2(a)$$

These postulates encode the associativity of the mode of combination a. Similarly, the Lambek-van Benthem calculus [van Benthem 1983], often called **LP**, is derived with a single binary mode of combination p, along with the postulates of associativity and permutation.

- $$\frac{\Gamma[((A,B)^p,C)^p] \Rightarrow D}{\Gamma[(A,(B,C)^p)^p] \Rightarrow D}A_1(p) \qquad \frac{\Gamma[(A,(B,C)^p)^p] \Rightarrow D}{\Gamma[((A,B)^p,C)^p] \Rightarrow D}A_2(p)$$
- $$\frac{\Gamma[(B,A)^p] \Rightarrow C}{\Gamma[(A,B)^p] \Rightarrow C}P(p)$$

Morrill defined a wrapping mode w, and allowed it to interact with the non-associative mode n and the associative mode a. The logic is determined by the binary modes $\mathsf{BinMod} = \{a, n, w\}$ and no unary modes. The non-associative mode is not subject to structural rules (on its own) the associative mode a is subject to the associative rules above, and the wrapping mode interacts with the other two modes according to the following schemes.

- $$\frac{\Gamma[((\Delta_1,\Delta_2)^n,\Delta_3)^w] \Rightarrow A}{\Gamma[((\Delta_1,\Delta_3)^a,\Delta_2)^a] \Rightarrow A}$$
- $$\frac{\Gamma[((\Delta_1,\Delta_3)^a,\Delta_2)^a] \Rightarrow A}{\Gamma[((\Delta_1,\Delta_2)^n,\Delta_3)^w] \Rightarrow A}$$

Next, consider the following examples of structural postulates for unary modes of combination given by Moortgat [1994].

- $$\frac{\Gamma[(\Delta)^u] \Rightarrow A}{\Gamma[((\Delta)^u)^u] \Rightarrow A}4(u)$$
- $$\frac{\Gamma[(\Delta)^u] \Rightarrow A}{\Gamma[\Delta] \Rightarrow A}T(u)$$

In addition to postulates governing the behavior of a single mode of combination, Moortgat allows for mixed postulates, which allow the binary associative mode a to interact or *communicate* with a unary mode u (in order to encode the Hepple/Morrill approach to bounding domains).

- $\dfrac{\Gamma[((\Delta_1)^u, \Delta_2)^a] \Rightarrow A}{\Gamma[((\Delta_1, \Delta_2)^a)^u] \Rightarrow A} K_1(u, a)$
- $\dfrac{\Gamma[(\Delta_1, (\Delta_2)^u)^a] \Rightarrow A}{\Gamma[((\Delta_1, \Delta_2)^a)^u] \Rightarrow A} K_2(u, a)$
- $\dfrac{\Gamma[((\Delta_1)^u, (\Delta_2)^u)^a] \Rightarrow A}{\Gamma[((\Delta_1, \Delta_2)^a)^u] \Rightarrow A} K(u, a)$

It is also possible to use a unary operator p in conjunction with a binary operator a in order to condition permutation [Morrill 1994].

- $\dfrac{\Gamma[((\Gamma)^p, \Delta)^a] \Rightarrow A}{\Gamma[(\Delta, (\Gamma)^p)^a] \Rightarrow A} P(a, p)$

In linear logic, unary modes have been used to import more liberal control over resources into a more restrictive logic. For instance, there is a mode, typically written !, which allows both weakening and contraction, as below. It is typically subject to the K structural rule and mixed with the commutative and associative binary mode p.

- $\dfrac{\Gamma[(\Delta, (\Delta)^!)^p] \Rightarrow A}{\Gamma[(\Delta)^!] \Rightarrow A}$
- $\dfrac{\Gamma[\Phi] \Rightarrow A}{\Gamma[((\Delta)^!, \Phi)^p] \Rightarrow A}$

The first rule allows the Δ resource to be duplicated, and the second allows it to be eliminated. Often, the second rule above is used in the context of a nullary mode corresponding to the empty set of assumptions, but we will not need recourse to nullary modes in our development.

Assuming an alphabet AtExp of *atomic expressions*, a *lexicon* is a finite set Lex whose elements are of the form $e \Rightarrow A$, where $A \in \mathsf{Cat}$ and $e \in \mathsf{Exp}$. In general, the set Exp of *expressions* is the least such that:

- $\mathsf{AtExp} \subseteq \mathsf{Exp}$
- $(e)^u \in \mathsf{Exp}$ if $e \in \mathsf{Exp}$
- $(e_1, e_2)^b \in \mathsf{Exp}$ if $e_1, e_2 \in \mathsf{Exp}$

Assuming a non-associative binary mode of combination, expressions behave rather like bracketed strings, and with associative combination, behave like strings. We can then extend our sequents to include arbitrary expressions as follows.

- $\Gamma[A_1 \mapsto e_1, \ldots, A_n \mapsto e_n] \Rightarrow C$ iff $\Gamma \Rightarrow C$ and $e_i \Rightarrow A_i \in \mathsf{Lex}$ for $1 \leq i \leq n$

Here we take $\Gamma[A_1 \mapsto e_1, \ldots, A_n \mapsto e_n]$ to be the result of replacing all occurrences of A_i with e_i.

3 Turing Completeness

For the purposes of this paper, we note a few relevant points about the form of these rules. First, they may either add or delete structure, as evidenced by the schemes 4 and T above. Second, they may work in only one direction, as evidenced by the 4, K, and T schemes. No one has suggested any metatheoretical constraints on structural postulates other than that they are not conditioned by particular categories (though they can obviously be conditioned on the modes themselves). If arbitrary structural schemes are allowed, we have the following immediate result.

Theorem 1 (Turing Completeness) *A set $S \subseteq \mathsf{AtExp}^*$ of expressions is enumerable by a Turing machine if and only if there is a categorial grammar, containing a category A, such that $e \in S$ if and only if $e \Rightarrow A$ is provable.*

Proof:
($\Leftarrow$) This is obvious because a Turing machine can effectively enumerate all possible proofs by breadth-first search in the usual way.
($\Rightarrow$) Suppose S can be enumerated by a Turing machine. Then there is a general rewriting (Type 0) grammar G with a distinguished non-terminal C_0 (usually called the *start symbol*) such that $e \in S$ if and only if $C \stackrel{*}{\rightarrow} e$ (where $\stackrel{*}{\rightarrow}$ is the usual transitive reflexive closure of the one-step substring rewriting relation). Suppose the finite set of non-terminals of G is $V_1, \ldots, V_n$. Then we define a categorial grammar as follows. We take the set of unary modes to be $u_1, \ldots, u_n$ (one for each non-terminal), and combine this with a single binary mode a. We take exactly one category symbol, C. We assume the structural postulates of associativity for a, and will thus omit the binary bracketings; that is, we use the notation $A_1 \cdots A_n$ for $((A_1, A_2)^a, \cdots A_n)^a$. Now for every production $V_{i_1} \cdots V_{i_k} \Rightarrow V_{j_1} \cdots V_{j_m}$ in the rewriting system, we assume the following structural postulate:

$$\frac{\Gamma[(\Delta)^{j_1} \cdots (\Delta)^{j_m}] \Rightarrow A}{\Gamma[(\Delta)^{i_1} \cdots (\Delta)^{i_k}] \Rightarrow A}$$

Note that this scheme will entail, via the left rules for the various $\Diamond$, the following:

$$\frac{\Gamma[\Diamond_{j_1} C \cdots \Diamond_{j_m} C] \Rightarrow A}{\Gamma[\Diamond_{i_1} C \cdots \Diamond_{i_k} C] \Rightarrow A}$$

For every lexical relation $V_i \rightarrow e$ in the rewriting system, we assume a lexical entry $e \Rightarrow \Diamond_i C$ in the categorial grammar. Assuming $C_0 = V_k$ is the distinguished (start) symbol of the rewriting system, we must show that $e_1 \cdots e_p \Rightarrow \Diamond_k C$ if and only if $V_k \stackrel{*}{\rightarrow} e_1 \cdots e_p$. But this result is trivial, because there is a one-to-one correspondence between allowable one-step rewritings in the rewriting system and applications of structural rules in the categorial grammar. Simply recall that we have a one-step rewriting $\sigma\tau\rho \rightarrow \sigma\tau'\rho'$ in a rewriting grammar if $\tau \rightarrow \tau'$ is a rewriting rule in the grammar, and σ, τ, τ' and ρ are sequences of terminals and non-terminals. Finally, we note that it will not be possible to apply a right rule of any sort, because no connectives will ever be generated on the right-hand side of a sequent. Further, no left rule can be used because there will never be a pattern that maches the left unary or binary modal rules. $\square$

Usually a proof of Turing-completeness for a grammar formalism leads its proponents to think more carefully about the metatheory in terms of exactly what is to count as a grammar. One rather unnatural aspect of the reduction in the proof is that it allows arbitrary amounts of structure to be inserted and retracted. In the original Lambek system, decidability was ensured by requiring subproofs to only involve subcategories of the categories in the sequent being proved. Unfortunately, such restrictions have been violated in more recent proposals.

4 Sahlqvist Axioms and Decidability

An interesting restriction on structural rules was suggested by Kurtonina [1996] in her thesis as a means of guaranteeing completeness in a particular logical formulation of multi-modal categorial grammars. Kurtonina provides the following definition:

> A structural rule of the form $\alpha \vdash \beta$ is said to be a *weak Sahlqvist axiom* if and only if α is a pure product formula, associated in any order, without repetitions of proposition letters, and β is also a pure product formula containing at least one product, all of whose atoms occur in α.

Kurtonina's motivation was to enable the proof of a powerful completeness theorem concerning the non-associative Lambek Calculus (**NL**) combined with a weak Sahlqvist axiom, the combination of which leads to a complete proof theory for the models satisfying the corresponding frame conditions derived from the axiom.

¿From the point of view of representational power, it is easy to see that the restriction to Sahlqvist axioms defeats our coding of Turing machines using modal operators. Because all applications of structural rules will never increase the size of the proof space beyond a finite limit, decidability for the non-associative calculus combined with any weak Sahlqvist axiom is guaranteed.

For now, we have to leave aside the important issue of whether the restriction to weak Sahlqvist axioms will be too restrictive or whether such axioms will be sufficiently rich to allow us to adequately model natural language syntax and semantics.

5 Conclusion

We have shown that mutlimodal categorial grammars can simulate arbitrary generalized rewriting systems by modeling their productions using structural interaction postulates. Without some restriction, such as to the weak Sahlqvist axioms, the universal decision problem for the grammaticality of a string with respect to a grammar is undecidable. This places multimodal categorial grammar squarely in the camp of unconstrained grammatical theories such as transformational grammars, phrase-structure grammars with metarules, and simple applicative categorial grammars with lexical rules.

References

Belnap, N. 1981. Display logic. *Journal of Philosophical Logic* 11: 375–417.

Carpenter, B., 1991. The Generative Power of Categorial Grammars and Head-driven Phrase Structure Grammars with Lexical Rules. *Computational Linguistics* 17:3, 301–314.

Gazdar, G., E. Klein, G. Pullum, and I. Sag, 1985. *Generalized Phrase Structure Grammar*. Basil Blackwell, Oxford.

Hepple, M., 1990. *The Grammar and Processing of Order and Dependency: A Categorial Approach*. Ph.D. Dissertation, Centre for Cognitive Science, University of Edinburgh: Edinburgh.

Kaplan, R. and Bresnan, J. 1982. Lexical-functional grammar: A formal system for grammatical representation. In Bresnan, J., editor, *The Mental Representation of Grammatical Relations*, 173–281. MIT Press, Cambridge, Massachusetts.

Kraak, E. 1995. French object clitics: a mutlimodal analysis. In *Proceedings of the First International Conference on Formal Grammars*. Barcelona.

Kurtonina, N. 1996. *Frames and Labels: A Modal Analysis of Categorial Inference*. Doctoral Dissertation, Research Institute for Language and Speech, University of Utrecht.

Lambek, J., 1958. The mathematics of sentence structure. *American Mathematical Monthly* 65:154–169.

Lambek, J., 1961. On the calculus of syntactic types. In Jakobson, R., editor, *Structure of Language and its Mathematical Aspects, Proceedings of Symposia in Applied Mathematics*, 166–178, American Mathematical Society, Providence, Rhode Island.

Moortgat, M., 1991. Generalized Quantification and Discontinuous Type Constructors. To appear in Sijtsma and van Horck, eds., *Proceedings of the Tilburg Symposium on Discontinuous Constituency*. De Gruyter: Berlin.

Moortgat, M. 1994. Residuation in mixed Lambek systems. Paper presented to the Deduction and Language Workshop, London. Manuscript, Institute for Speech and Language (OTS), Utrecht.

Moortgat, M. and R. Oehrle. 1994. Adjacency, dependency and order. In P. Dekker and M. Stokhof, eds., *Proceedings of the 9th Amsterdam Colloquium*. Dordrecht: Foris.

Morrill, G., 1990. Intensionality and Boundedness. *Linguistics and Philosophy* 13, 699–726.

Morrill, G. 1994. *Type-Logical Grammar*. Kluwer Academic Publishers, Dordrecht.

Morrill, G. 1995. Discontinuity in categorial grammar. *Linguistics and Philosophy* 18:175–219.

Peters, P. S. and R. W. Ritchie. 1973. Context-sensitive immediate constituent analysis: context-free languages revisited. *Mathematical Systems Theory* 6:324–333.

Solias, T., 1992. *Gramáticas Categoriales, Coordinación Generalizada y Elisión.* Doctorate Thesis. Departamento de Lingüísticia, Lógica, Lenguas Modernas y Filosofía de la Ciencia, Universidad Autónoma de Madrid.

Uszkoreit, H. and P. S. Peters, 1986 On some formal properties of metarules. *Linguistics and Philosophy* **9**(4):477–494.

Versmissen, K. 1996. *Grammatical Composition: Modes, Models, Modalities.* Doctoral Dissertation, Research Institute for Language and Speech, University of Utrecht.