

Expressive Power of Grammatical Formalisms

Alexis Manaster-Ramer & Wlodek Zadrozny
IBM Research
T. J. Watson Research Center
Yorktown Heights, NY 10598
AMR @ IBM.COM WLODZ @ IBM.COM

Abstract

We propose formalisms and concepts which allow to make precise the arguments in controversies over the adequacy of competing models of language, and over their formal equivalence.

1. Introduction

It is customary to judge the success of scientific models by their agreement or otherwise with the observed data. For example, linguists require of grammars that they generate the right sentences, but also that they correctly classify the sentences and phrases as to the categories and constructions they belong to. Our purpose is provide a formal account of the elusive concept of expressive power with respect to the kinds of categories and constructions that a grammar (of a given type) can reflect. The principal concept will be definability of relations in a logical formalism corresponding to a given grammar type in this language, specifically definability without the use of disjunction in the defining formula. Our results can be summarized as follows:

1. We cast CFGs in a logical formalism. We then progressively enrich the formalism to express the parametrization of categories and of constructions in various ways as well as by allowing metarules and transformations.
2. We then prove a number of theorems about what can and cannot be done in a given formalism, focusing on the definability of categories and constructions (both taken as relations in the logic).
 - a. Definability is characterized precisely for the first time, and we distinguish various kinds, of which nondisjunctive definability corresponds closely the notion of capturing a linguistic generalization in a grammar.

- b. Agreement is not definable in CF theories, but can be defined in theories/grammars with attributes.
- c. Constructions whose variants differ in word order and/or in the number of constituents cannot be captured even by CFGs with attributes, but can in slightly more powerful models.
- d. Constructions as above but where the order and/or the number of constituents correlate ('agree') with some other feature require inherently more powerful systems. We show how such patterns can be captured if we parametrize concatenation and the number of constituents (for the first time, something other than categories gets parametrized).
- e. The same generalizations can be captured via transformations or metarules. Grammars with transformations and metarules can be treated as particular cases of a certain formal proof system.
- f. Various extended notions of definability are considered; for example, the binary relation between pairs of trees related by a metarule or transformation, and the notion of definability across a class of grammars.

2. Definability

2.1 Expressive power

The expressive power of a logical theory depends on four factors:

1. The formal language L in which the theory is written;
Example: FOL (the First Order Logic) is more expressive than the Propositional Calculus

2. The class of WFF of well-formed formulas of L;
Example: FOL is more expressive than Horn Clauses, but the latter are easily computable (no function symbols in both cases).
3. Axioms of the theory;
Example. Two theories expressed in the same language, such as $L = (+, *, =, <, 1)$, can have completely different properties. For instance the axiomatization of real numbers in L gives a decidable theory, and hence any formula written in L is either provable or disprovable from the axioms. But the axiomatization of natural numbers in the same language does not decide all formulas, i.e. there are formulas written in L which cannot be proved or disproved from the axioms.
4. Rules of inference;
Example. With Modus Ponens $A, A \rightarrow B // B$ one can prove more than just with the rule $A // A \vee B$.

We shall deal mostly with 1,2 and 4, and allow 3 to correspond to a translation into a logical language of the context free part of a grammar. As we have already mentioned, the difference between TGs and MGs has something to do with 4.

2.2 Context Free Theories

Using grammars we can talk about which notions can be defined, or expressed, thereby only in an intuitive sense. The reason for that lies in the fact that "being defined/definable" is a property of a predicate, therefore a proper language for studying expressive power of different grammars, and grammatical formalisms is logic. Then we can talk about non-definability or definability of notions such as *Passive(x)* or *Passive_of(x,y)* formally, in a logical system. It turns out that definability of such a concept depends on a formal language in which a grammar is written. Thus it may happen that two grammars prove "sentencehood" of the exactly the same classes of strings, but it is possible to define such a predicate in one of those theories and not in the other.

We begin with CFGs. Since we are going to be concerned with definability, we first translate CFGs into CFTs (Context Free Theories). The translation works as follows. A CF production like $H \rightarrow B_1 B_2 \dots B_n$ goes into $H(x_1, x_2, \dots, x_n) \leftarrow B_1(x_1) \& B_2(x_2) \& \dots \& B_n(x_n)$ (note

that all the variables are different); $H \rightarrow a.G$ and $H \rightarrow a.b$ are replaced by $H(a.x) \leftarrow G(x)$ and $H(a.b)$, respectively.

For the sake of the uniformity of notation we will represent a rule of the form $H \leftarrow B_1 \& B_2$ as $[H, [B_1, B_2]]$. (The the rule of substitution will correspond to the resolution).

Proposition 1. Let G be a CFG and G' its translation. A string s belongs to L(G) iff G' proves S(s).

2.3 The Undefinability of Agreement

We turn to a simple example, namely, agreement between NPs and VPs. Consider a sample CFT:

$S(x,y) \leftarrow \text{NP_sing_fem}(x) \& \text{VP_sing_fem}(y).$
 $S(x,y) \leftarrow \text{NP_sing_mal}(x) \& \text{VP_sing_mal}(y).$
 $S(x,y) \leftarrow \text{NP_sing_neu}(x) \& \text{VP_sing_neu}(y).$
 $S(x,y) \leftarrow \text{NP_plur_fem}(x) \& \text{VP_plur_fem}(y).$
 $S(x,y) \leftarrow \text{NP_plur_mal}(x) \& \text{VP_plur_mal}(y).$
 $S(x,y) \leftarrow \text{NP_plur_neu}(x) \& \text{VP_plur_neu}(y).$
 ...

In order to prove that agreement cannot be captured here, we need to specify what that would mean. It is easy to show that well-defined relations such as *number_agree(number,x,y)* or *agree(number,x,y)* are only definable in CFT if the defining formula uses disjunction. But traditionally disjunction in grammatical description is a standard notation for two (or more) unrelated phenomena. Thus, disjunction is not forbidden, but when it occurs, it implies the factual claim about the referents of the disjuncts are distinct linguistic phenomena. In the case before us, that would be saying that singular and plural agreement are not the same phenomenon. It is not, of course, the business of logic to inquire into whether in fact number agreement in some language is a unitary phenomenon. It is rather the business of logic to provide the tools for the linguist scientist who, on whatever basis, makes such determinations, to capture formally the theories that he develops.

Accordingly, we first assume a special notion of definability, defined as follows:

Definition. A relation p is $\&$ -definable in a CFT theory T if there is a formula

(*) $P(x_1, \dots, x_N) \leftarrow B_1(x_1) \& \dots \& B_N(x_n) \& B(x_i, x_j, \dots) \& \dots$

s.t. the tuples (a_1, \dots, a_N) , which can be proven from $T + (*)$ to satisfy $P(\dots)$, are exactly those belonging to the relation p .

A category c is $\&$ -definable in a CFT theory T if the one argument relation corresponding to c is.

However, some categories thus defined are spurious in that they cannot be used in proving sentencehood. We want to rule these out.

Let $Lnf(G)$ be the language that contains only those categories which appear in the formulas $[S, [...]]$ which are derivable in (the CFT corresponding to) the grammar. From now on, by $\&$ -definability we shall understand the $\&$ -definability in $Lnf(G)$. Moreover, the notion of 'category' will be analogously restricted. This allows us to avoid spurious categories as in 2.4 below.

We will also refer to constructions which are two-place relations between a grammatical category (the category the construction yields) and a string of grammatical categories (which the construction is made up of).

2.4 Spurious categories

Consider a grammar like:

$S \rightarrow NP_{psg} VP_{psg}$
 $S \rightarrow NP_{pl} VP_{pl}$
 $NP_s \rightarrow Det N_s$
 $NP_p \rightarrow Det N_p$
 $NP \rightarrow Det N$
 $N_s \rightarrow dog, cat, \dots$
 $N_p \rightarrow dogs, cats, \dots$
 $N \rightarrow dog, dogs, cat, cats, \dots$

It is possible to introduce a symbol that corresponds to the category NP, but it cannot be used in deriving sentences from the start symbol. Our definition allows such spurious category symbols to appear in formulas of CFT, but at the same time it prevents them from having any influence on what categories are definable in the formalism. This simple example clearly shows that an appeal to intuitions would be insufficient to talk about expressive power of the two grammars.

2.5 Some Results

Theorem 2. If a CFT T contains

$$S(x,y) \leftarrow NP_1 - at_1(x) \ \& \ VP_1 - at_1(y)$$

$S(x,y) \leftarrow NP_2 - at_1(x) \ \& \ VP_2 - at_1(y)$
and both $NP_i - at_1(x)$ and $VP_i - at_1(x)$ are satisfiable for both i . Then the relation of agreement $agree(at_1, x, y)$ is not $\&$ -definable.

Proposition 3. The relation $agree(number, x, y)$ is not $\&$ -definable in the above CFT.

Theorem 3. The category NP is not $\&$ -definable or definable for this language in CFT.

Notice that linguists often do allow categories such as NP, but not constructions (such as the subject-predicate construction or passive) or features of constructions (such as agreement) to be described disjunctively. This is especially true of lexical categories. It is thus instructive that, as shown, the phenomenon of agreement means that certain categories are not definable at all, even with disjunction.

3. Attributes and Constructions

We will talk now about a logic corresponding to CFGs with attributes. We will show that such logics provide an inherently more expressive theory of categories, and in particular allow us to define the category NP and VP in a language with number or gender agreement.

However, the use of attributes does not lead to an all-powerful theory of constructions, and consequently certain linguistic generalizations are missed by grammars with attributes. This leads to the introduction of more powerful devices. All the formal languages we consider, if not first order, can be formalized as weak second order system in an obvious way.

3.1 Word Order and Selection Variation

Attribute theories clearly cannot treat as a single construction two forms with different word order. That is, they cannot $\&$ -define a relation $R(Cat, x)$, where Cat is some grammatical category and x ranges over a set of strings of categories identical except with respect to word order. This defect can be remedied by allowing an ID/LP format for rules, which we formalize in a very similar way. (Details omitted.) However, this formalism if the variants with different orders do not differ in any other way. What is more interesting is what happens with examples in which the choice of

attributes for some element is correlated with a choice of word order. For example, English has a slightly different class of verbs in "inverted" sentences than others, e.g. Aren't I smart? vs. *I aren't smart. There is no way to connect

$\left[\begin{array}{c} V \\ + \text{invert} \end{array} \right]$ with the $V < NP$ order, as opposed to

$\left[\begin{array}{c} V \\ - \text{invert} \end{array} \right]$ and the $NP < V$ order. It was precisely

to handle cases like this that metarules were introduced in GPSG model of grammar, and it is one of the reasons for transformations as well.

By parametrizing word order, we can capture the word order phenomena like those in English inverted sentences. Thus, we could have a grammar with a parametrized concatenation operator *conc* with values, such that $0 \text{ conc}(a,b,c) = abc$, whereas $1 \text{ conc}(a,b,c) = bac$, for example. We can now state a single rule of the form

$$S \rightarrow \alpha \text{ conc}(NP, \left[\begin{array}{c} AUX \\ \alpha \text{invert} \end{array} \right], VP)$$

to handle the subject-aux inversion facts.

Attribute theories also cannot handle variation in *selection*, i.e., the arity of a construction. This is easily remedied by formalizing the parenthesis notation of BNF which is often used to abbreviate CFG's--when we write, informally, a rule like $A \rightarrow B(C)$, for example. It is harder, to handle the correlation between some attribute of one element and the presence or absence of some other element. For instance, many analyses of English postulate separate constructions of the VP depending on the class of the verb, e.g., transitive (V NP), ditransitive (V NP NP), transitive-prepositional (V NP PP), and so on. It has also been observed that each of these corresponds to a passive form in which one NP is missing (although a PP of the form by NP is optionally possible instead, this is irrelevant for our purpose). Again, if it were just the presence or absence of the object NP that distinguished the two voices, we could use the parenthesis device. However, the form of the verb also changes from active (e.g. sees) to passive (e.g. is seen). Such phenomena, which can be handled with metarules or transformations, also cannot be handled with attribute grammars.

The problem of the definability of constructions is a more complicated one. The results below have been obtained for formulas of the standard first order language. A correct account of recursion

requires a formal language with something like the Kleene star, and require more space than we have in this paper. The idea is roughly this: We have introduced a set of sublanguages of $Lnf(G)$ to avoid spurious categories. Now, we make one more restriction: let $Lnf(NP)$ be the language that contains only those categories except *NP* which appear in the *RHS* of the formulas $[NP, [RHS]]$ which are derivable in (the CFT corresponding to) the grammar. Similarly, for other categories/symbols. The (&-)definability of a construction *X* are defined as for other relations, except that the defining formula must belong to $Lnf(X)$.

Theorem 4. Constructions with two variants which differ by the order or number of constituents are not &-definable in attribute grammars.

Theorem 5. Constructions with two variants which differ by the order or number of constituents together with a difference in some other element are not &-definable in attribute grammars with ID/LP and parentheses.

Now, consider an extension of attribute grammars which parametrizes the presence/absence of constituents. Thus, we write rules like, where *ont* is a parameter controlling the appearance or absence of an element (i.e., $+ont(X)$ means *X* appears, $-ont(X)$ that it does not),

$$VP \rightarrow \left[\begin{array}{c} V \\ \alpha \text{active} \end{array} \right] \alpha \text{ont}(NP)$$

Now we can &-define the different kinds of transitive constructions, by using *ont* to control whether the object NP is realized (in the active) or null (in the passive).

Now, the use of this device allows us to &-define the three different kinds of transitive constructions, but not the passive construction, which still requires disjunction (for the same reason that the active requires disjunction). This is exactly the same as with metarules and transformations (as we will see below). The formalism provides no way of making the verb class attributes (trans, ditrans, trans_prep) agree with the number and kind of constituent to the right of the verb (NP vs. NP NP vs. NP PP). Some linguists don't mind this, but we will show below how that can also be done (what is required is a way of making the verb class attribute agree with the number and kind of constituent following the verb). In order to &-define passive, we would need a more powerful

kind of parameter, which can control the number and kind of constituents, which we call *sel*

$$VP \rightarrow \left[\begin{array}{c} V \\ X \text{ sel} \end{array} \right] X$$

And combining the two (*ont* and *sel*), we can describe both the transitive and the passive by a rule like:

$$VP \rightarrow \left[\begin{array}{c} V \\ NP.X \text{ sel} \\ y \text{ voice} \end{array} \right] y \text{ ont}(NP) X$$

However, historically such devices as *conc*, *ont*, and *sel* have been unavailable, and instead, transformations and metarules have been used to obtain essentially the same effect. Hence, we proceed to show how the power of these models can be represented

4. Derivability, TGs and M-grammars

We will consider very simple kinds of TGs and MGs, which operate on (sub)trees of depth one. This is enough to capture GPSG use of metarules, but not the full power of conventional TG. The more general model will be discussed briefly, but for our purposes it is more convenient sometimes to consider special cases, which make the demonstrations simpler.

The TGs and grammars with metarules (M-Gs) deal not only with strings but also with trees. To compare them we have to use a common formalism. Let T be a collection of trees (over some alphabet with terminals, non-terminals, and perhaps other symbols), where each tree is a pair $[Node, Sons]$, where *Sons* is a list of trees. Then each rule of a context free grammar can be represented as a tree of depth one (the definition of depth being obvious), e.g. $[S(x,y), [NP(x), VP(y)]]$, or $[N(dog), [\emptyset]]$.

We need now to establish the following interpretation:

- Trees, as defined above, will be interpreted as formulas;
- The rules of proof will be expressed in the Gentzen style: $Tree1, Tree2 // Tree3$;
- One of the rules of proof will be substitution, as in

$$\frac{[S(x,y), [NP(x), VP(y)]] \quad [NP(u,w), [ADJ(u), NP(w)]]}{[S(u.w.y), [[NP(u.w), [ADJ(u), NP(w)]], VP(y)]]}$$

Clearly, with this rule we can prove about a string that it is an S if it is generated by the corresponding context free grammar. In a natural way we can extend this definition of an inference rule to cover attribute grammars: attributes can be treated simply as constraints.

- A metarule in an M-grammar such as $VP \rightarrow X NP // VP[PAS] \rightarrow X (PP)$ which relates passive and active, can be treated as an inference rule.
$$\frac{[VP(x,y), [X(x), NP(y)]]}{[VP(x.z)(PAS), [X(x), PP(by.z)]]}$$
- A transformation in a TG can be understood exactly the same way, as a rule $Tree1 // Tree2$.

The difference between TG and grammars with metarules (M-Gs) can be now expressed in the definition of a proof.

- In the case of TGs a proof of a formula F_n is a sequence (P, Q) where
 1. P is a sequence of formulas F_1, \dots, F_k such that each F_i is a formula corresponding to a context free rule of the grammar or is obtained from F_j , ($j < i$) by the rule of substitution, and F_k is $[S(\dots), [RHS]]$, where RHS contains only the terminal symbols (i.e. F_k represents a fully expanded tree).
 2. Q is a sequence F_k, \dots, F_n , where each formula is obtained from the previous one by a rule corresponding to a transformation.
- For M-Gs a proof of a formula F_n is a sequence $(P1, Q, P)$ where
 1. $P1$ is a collection of formulas of depth one;
 2. Q -- formulas/trees obtained by applying metarules;
 3. P -- formulas/trees obtained by applying the rule of substitution.

These definitions allow us to show that a variety of constructions not allowable by CFG's or attribute grammars are definable by MG's and TG's. Also, we obtain a rather neat characterization of the similarities and differences between TG's and MG's.

5. Other Notions of Definability

The notion of &-definability of categories and constructions is not the only one that we could have employed. One alternative would be to look at what relations are definable. For instance, Sells (p.93) remarks that it is impossible to express Subject-Subject-Raising as a metarule. From our perspective this of course still holds true, but we can also see that one can easily define a relation $SSR(t_1, t_2)$ which holds only if the first tree has been obtained by subject raising from the second tree:

$$SSR(t_1, t_2) \leftarrow \\ t_1 = [S(\text{seem}.y.z), [\text{seem}, NP(y), VP(z)]] \\ \& t_2 = [S(y.\text{seems}.z), [NP(y), \text{seems}, VP(z)]]$$

("seem" stands for all verbs of this class).

Another extension would be to consider definability of categories, constructions, or relations not for a single grammar but across an (infinite) class of grammars. For example, we have seen that in a language of CFT a definition of agreement can be given only by a disjunctive formula, but it can be given, so long as we confine ourselves to a single grammar.

However, in linguistics there is precedent for (informal) arguments that some notion, while definable for each grammar of a class, cannot be defined for the whole class. This idea leads to a new notion of definability:

Definition. A property Pr is definable over/across a class of theories C , if there is a formula $F(x)$ s.t. for any T in C , and any term t of the language, we have $Pr(t)$ iff $T + F(x)$ proves $P(t)$.

Similarly, one defines &-definability across C . ($Pr(t)$ means that it holds in any model of T).

Let's concentrate on the agreement wrt gender. Notice that the assumption that agreement exists and that it should be somehow expressible in a grammar (using disjunctions or not) is an empirical statement. We could express it formally by augmenting our language with a (higher order)

device detecting the presence of a substring 'fem' in a predicate as in

$$S(x.y) \leftarrow NP - \text{sing} - \text{fem}(x) \& VP - \text{sing} - \text{fem}(y).$$

We will be interested in a subclass of CFT that allow formulas of the sort:

$$S(x.y) \leftarrow NP - at_1 - at_2 - \dots - at_N - \text{fem}(x) \& \\ VP - at_1 - \dots - at_N - \text{fem}(y)$$

Assuming that we can talk about agreement formally, we can formulate and prove the following fact.

Theorem 6. Agreement with respect to an attribute is not definable across the class CFT of context free theories.

We now see that much stronger results can be achieved across languages than for one language. For a single CFT, we can only that agreement is not &-definable, but for the CFT's as a class we have just seen that it is not definable at all.

6. Conclusions

We proved a number of results about the expressive power of a number of grammatical formalisms, not just CFGs, but also others that resemble more closely what linguists actually work with. More important, we have proposed a method which can be extended to any class of grammars for characterizing precisely what relations, constructions, and categories this kind of grammar can "capture". In the process, we have clarified the notion of category, defined for the first time the notion of construction, and proposed a number of grammatical devices that have not been considered before, and cast a new light on the problem of the relation of metarules, transformations, and extensions of phrase structure, such as parametrizing categories (i.e., using attributes) and parametrizing constructions. We believe that the crucial step was dealing not directly with grammars but with corresponding logical theories for such grammars, and this will continue to prove fruitful in the future.