Adgrams: Categories and Linguistics

Marco Benini

M.Benini@leeds.ac.uk

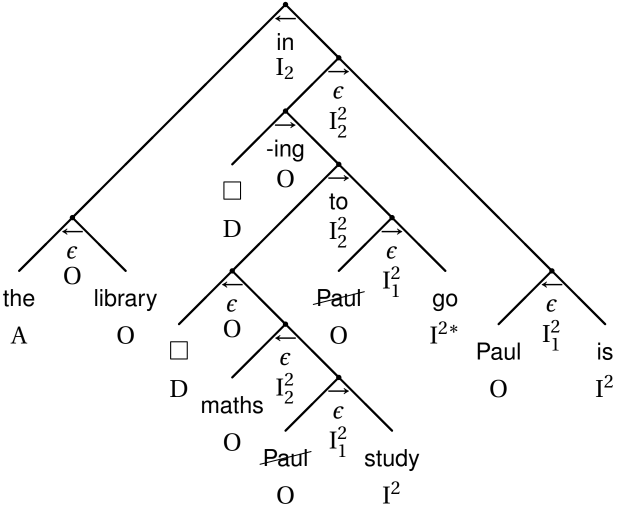
Department of Pure Mathematics
University of Leeds

26th October 2011

UNIVERSITY OF LEEDS

An example

"Paul is going to study maths in the library".



Categories as languages

To interpret the objects of a small category as the collection of expressions of some language, the category must have enough "structure" to encode the grammar constructions in its arrows.

The first part of the formal model, what we are going to describe in this talk, is devoted to describe the required structure.

Although the motivation for our definitions can be traced back to what happens in natural languages, we will limit ourselves to the example just shown, as it suffices for the purposes of this talk.

Grammar character

A grammar character is a collection of expressions in the language that can be used to construct new expressions. Informally, “nouns”, “adjectives”, “verbs” are grammar characters (but beware that they do not properly work as you may expect!)

Definition 1 (Grammar Character)

Given a small category $\mathbb{C}$, we say that G is a set of *grammar characters* on $\mathbb{C}$, indexed by GC, if $G = \{G_i\}_{i \in GC}$ is a partition on $\mathrm{Obj}\,\mathbb{C}$ such that every class is inhabited.

In the example, the grammar characters' names are A, I^2, I^2_1, I^2_2, O and D.

Grammar product

Definition 2 (Grammar Product)

Fixed a small category $\mathbb{C}$, a set $G = \{G_i\}_{i \in GC}$ of grammar characters on $\mathbb{C}$ and $W \in G$ (the *adpositions*), any pair $\langle Y_1 \times \cdots \times Y_n, (z_1, \ldots, z_n) \rangle$ is a *grammar product* on $\mathbb{C}$, G, W when $n \in \mathbb{N}$ (the *degree* of the product) and, for every i in $1 \ldots n$, $Y_i \in G \setminus \{W\}$ and $z_i \in W$.

In the example, every part of a branch starting with a right-most leaf is an instance of a grammar product.

Since most of the relevant information about a construction is in the grammar product, it is convenient to give a separate definition for this auxiliary concept.

Grammar construction I

Definition 3 (Grammar construction)

Given a small category $\mathbb{C}$, a set $G = \{G_i\}_{i \in GC}$ of grammar characters on $\mathbb{C}$ and $W \in G$, a *grammar construction* $\theta_{P,x}$ over P from x is an indexed collection $\{f_j\}_{j \in P'}$ where

- $P = \langle P', \alpha \rangle$ is a grammar product over $\mathbb{C}$, G, W;
- x is an object of $\mathbb{C}$, called the *governor* of the construction;
- there is $V \in G \setminus \{W\}$ such that, for every $j \in P'$, $f_j \in \mathrm{Hom}_{\mathbb{C}}(x, y)$ with $y \in V$. Every component of j is called a *dependent* of the construction.

We denote V as $\mathrm{cod}_{\theta_{P,x}}$, x as $\mathrm{dom}_{\theta_{P,x}}$ and sometimes we write $\theta_{j,x}$ for f_j. A family $\{\theta_{P,x}\}_{x \in U}$ of grammar constructions over P from x varying in the U grammar character forms a *grammar construction* $\theta_{P,U}$ over P from U when its elements share the same codomain.

Grammar construction II

In the example, we see instances of

- the verbal constructions with product $\langle \mathrm{O} \times \mathrm{O}, (\epsilon, \epsilon) \rangle$, $\langle \mathrm{O} \times \mathrm{O}, (\epsilon, \mathrm{to}) \rangle$;
- the infinitive construction with product $\langle \mathrm{D}, (\epsilon) \rangle$;
- the gerund construction with product $\langle \mathrm{D}, (\epsilon) \rangle$;
- the adjunctive construction with product $\langle \mathrm{A}, (\epsilon) \rangle$;
- the circumstantial construction with product $\langle \mathrm{O}, (\mathrm{in}) \rangle$.

Conjugate construction

Definition 4 (Conjugate construction)

Fixed a small category $\mathbb{C}$, a set $G = \{G_i\}_{i \in GC}$ of grammar characters on $\mathbb{C}$, with $W \in G$, and a grammar construction $\theta = \left\{ \{f_j : x \to c(x,j)\}_{j \in P'} \right\}_{x \in V}$ over $P = \langle P' = P_1 \times \cdots \times P_n, (a_1, \ldots, a_n) \rangle$ from the V grammar character, we say that a grammar construction σ over Q from U is *conjugate* to θ if, for some k in $1 \ldots n$, $U = P_k$,

$$Q = \langle P_1 \times \cdots \times P_{k-1} \times V \times P_{k+1} \times \cdots \times P_n, (a_k, a_2, \ldots, a_{k-1}, a_1, a_{k+1}, \ldots, a_n) \rangle \ ,$$

and, for all $(x, j_1, \ldots, j_n) \in V \times P'$, $\operatorname{cod} \theta_{(j_1, \ldots, j_n), x} = \operatorname{cod} \sigma_{(j_1, \ldots, j_{k-1}, x, j_{k+1}, \ldots, j_n), j_k}$.

Essentially, a conjugate construction is just the “normal” construction where the governor is exchanged with one dependent. It models a different, but equivalent, point of view of what matters in the construction.

Composition of constructions I

Definition 5 (Composition)

Given a small category $\mathbb{C}$, a set $G = \{G_i\}_{i \in GC}$ of grammar characters on $\mathbb{C}$, with $W \in G$, and grammar constructions

$$\eta = \left\{ \{ f_{(y_1,\ldots,y_n)} : x \to c(x, y_1, \ldots, y_n) \}_{y_1 \in P_1, \ldots, y_n \in P_n} \right\}_{x \in V}$$

over $P = \langle P_1 \times \cdots \times P_n, (a_1, \ldots, a_n) \rangle$ from V, and

$$\theta = \left\{ \{ g_{(z_1,\ldots,z_m)} : x \to d(x, z_1, \ldots, z_m) \}_{z_1 \in Q_1, \ldots, z_m \in Q_m} \right\}_{x \in U}$$

over $Q = \langle Q_1 \times \cdots \times Q_m, (b_1, \ldots, b_m) \rangle$ from U, the *composition* $\theta \circ \eta$ is defined when $\operatorname{dom} \theta = U = \operatorname{cod} \eta$. $\hookrightarrow$

Composition of constructions II

$\hookrightarrow$ (Composition)

In this case,

$$\theta \circ \eta = \left\{ \{ g_{(z_1,\ldots,z_m)} \circ f_{(y_1,\ldots,y_n)} \}_{y_1 \in P_1,\ldots,y_n \in P_n, z_1 \in Q_1,\ldots,z_m \in Q_m} \right\}_{x \in V}$$

so $\operatorname{dom} \theta \circ \eta = \operatorname{dom} \eta = V$, $\operatorname{cod} \theta \circ \eta = \operatorname{cod} \theta$, and the product is $\langle P_1 \times \cdots \times P_n \times Q_1 \times \cdots \times Q_m, (a_1, \ldots, a_n, b_1, \ldots, b_m) \rangle$.

Despite the apparent complexity, composition is what one naturally expects to be.

Grammar category I

Definition 6 (Grammar category)

Consider a small category $\mathbb{C}$, a set $G = \{G_i\}_{i \in GC}$ of grammar characters on $\mathbb{C}$, with $W \in G$, and a set C of grammar constructions over products in $\mathbb{C}$, G, W from grammar characters in G. Then, the structure $\langle \mathbb{C}, G, W, C \rangle$ is said to be a *grammar category* when

1. (well-foundedness) There is $M \subseteq \text{Obj}\,\mathbb{C}$, the set of *atoms*, such that, for each object x, there is an arrow $f \in \text{Hom}_{\mathbb{C}}(m, x)$ for some $m \in M$. Moreover, if $x \in M$, $f = 1_x$.
2. (covering) For each arrow $f \in \text{Hom}_{\mathbb{C}}(a, b)$, there is $\eta_{P,V} \in C$ such that $a \in V$ and $f \in \eta_{P,a}$.
3. (1-degree regularity) Every construction in C is the composition of a finite sequence of constructions in C each one having degree 1.
4. (conjugate regularity) Each construction in C has exactly one conjugate construction for every element in its product.

Grammar category II

Apart the technicalities in the definition, a small category can be interpreted as a grammar category, i.e., a language with a formal adpositional grammar, when we can find a set of grammar characters, a set of adpositions, a set of atoms and a set of constructions in it such that

- every object can be constructed starting from the atoms;
- every arrow is part of some construction;
- every construction can be decomposed in a sequence of elementary (one governor, one dependent) constructions;
- conjugate constructions depend only on the product.

It is always possible to satisfy these requirements in a trivial way. The thing becomes interesting, as usual, in the non-trivial case.

And yes, English, Italian, Chinese, ... can be described in this way!

Adtrees

Given a grammar category, we can write every arrow in a construction

$$\left\{f_{(y_1,\dots,y_n)} \colon x \to c(x, y_1, \dots, y_n)\right\}_{y_1 \in Y_1, \dots, y_n \in Y_n}$$

over the product $\langle Y_1 \times \cdots \times Y_n, (a_1, \dots, a_n)\rangle$ from x as

$$x \xrightarrow{f_{(y_1,\dots,y_n)}} a_n(y_n, a_{n-1}(y_{n-1}, \dots, a_1(x, y_1)\dots)) = c(x, y_1, \dots, y_n) \quad .$$

If we omit the f, which is normally clear from the context, and we write the target term as a tree, we get a graphical representation like our example. This representation is called *adtree*.

When every object can be represented in a unique way as an adtree (modulo conjugate constructions), a grammar category is said to be *pure*.

Parsing

Theorem 7
Let $\mathbb{C}$ be a grammar category. Then, every object of $\mathbb{C}$ is denoted by an adtree X, and there is an adtree Y, denoting the same object, such that Y is linear (its governor and its dependents are atoms) and Y is equivalent to X modulo conjugation.

When the collections of objects and arrows of $\mathbb{C}$ are small enough, then the proof of the theorem provides a (very inefficient) algorithm to reconstruct X from Y, and Y is (almost) the usual way we write expressions.

Moreover, we can limit ourselves to observe adtrees: if a grammar category is pure, up to conjugates, every adtree represents (“is”) a unique object; in the general case, a grammar category is a quotient of the category $\mathbf{Adtree}(\mathbb{C})$ having adtrees as objects and instances of constructions as arrows.

Transformations

Definition 8 (Transformation)

Given a grammar category $\mathbb{C}$, a *transformation* on $\mathbb{C}$ is an endofunctor in the category $\mathbf{Adtree}(\mathbb{C})$ of adtrees on $\mathbb{C}$.

Defining constructions is easy when we deal with the elementary aspects of a natural language. But, when dealing with sophisticated concepts, it is easier to define a construction as the result of a transformation over (a specific subset of) adtrees.

In our example, the verbal construction and the circumstantial construction are elementary, while the infinitive and the gerund are best understood when considering them as the result of appropriate transformations mapping the “direct” sentence into the “abstract” tense.

In general, one may think to transformations as “translations” of expressions into a context.

Language generation I

In the case of a pure grammar category $\mathbb{C}$, the category $\textbf{Adtree}(\mathbb{C})$ is isomorphic to $\mathbb{C}$, so the language can be generated by constructing every possible adtree. This is possible because the grammar category is well-founded.

In the case of English, it is immediate to understand that every correct adtree is generated, but also incorrect ones are.

So, we have to recognize the correct adtrees among the collection of all the generated adtrees. A closer analysis reveals that the “wrong” adtrees have the “right” structure but they fail to agree on the person-verb (go , goes), on the plural-singular, etc.

Language generation II

All these details are called *redundancies* in general linguistics, as they are deputed to “repeat” pieces of information already present in an expression, so to unambiguously determine its meaning.

To cope with redundancies we need to deal with the morphemic structure, i.e., to the way words are formed: I will not deepen this aspect in this talk. It suffices to say that the morphemic structure is captured by grammar categories whose adtrees represent “words” and idiomatic expressions.

Also, we can always build a grammar category over another one, whose adtrees become the atoms of the new category. This construction is beyond the scope of this talk, but it suffices to say that the linguistic level, which is a grammar category, is built upon the morphemic level, another grammar category, and it is possible to recover the morphemic structure in a canonical way from the linguistic level, when the morphemic category is known.

Language generation III

It is always possible to imagine a transformation that maps an adtree X into another one where a redundancy aspect is corrected if it was wrong in X. By the way, this is the "usual" way used to explain redundancies (except that no one speaks of transformations...).

So, let's suppose to have a collection $\mathbb{R}$ of transformations deputed to correct redundancies. Evidently, an adtree X is correct with respect to redundancies if, and only if, $X = \tau(X)$ for every $\tau \in \mathbb{R}$. That is, X is correct if and only if X is a fixed point for every element of $\mathbb{R}$.

Hence, the language generated by a grammar category with respect to a set of redundancy-transformations $\mathbb{R}$ is the collection of adtrees which are fixed points for $\mathbb{R}$.

Semantics and pragmatics

We can develop semantics and pragmatics inside this framework by considering “narratives”: the meaning and the intention of an expression are given by relations with the world. And, in turn, the world is a huge description (a *narrative*) which is presumed by the speaker.

The world’s narrative can be coded by some adtree λ and the expression X, together with its narrative, becomes the enriched adtree $\epsilon(\lambda, X)$.

Semantics arise when we precisely define how X relates to λ. Pragmatics arise when we study as a sequence of expression develops in time, in particular, in which way the λ’s are modified.

In our approach, semantics has been partially developed, although it has not been completed. On the other side, pragmatics has been developed according to the ideas of Searle, and it has been applied to study the interaction between a therapist and a patient in psycoanalytical practice.

Links with Leeds I

Building a theory in Mathematics is also a linguistic act!

We define a formal language, we declare axioms, we provide definitions, we prove theorems, referring to an implicit, but well described world.

To what extent we can use the illustrated linguistic approach to infer properties of mathematical theories?

Being constructive, or being predicative, is perhaps a “linguistic” property of a theory?

Links with Leeds II

I believe that we cannot expect to gain deep results on syntax or semantics from the illustrated approach.
But, I also believe that we can, and we should, learn more about our theories when we consider also the pragmatics of development of mathematical theories.

In particular, I think that being constructive or predicative are mainly linguistic properties of a theory. And this is an “attack point” for the research project I’ll work on in my period in the School of Mathematics.

More information

- Federico Gobbo and Marco Benini, *Constructive Adpositional Grammars: Foundations of Constructive Linguistics*, Cambridge Scholar Publishing, Newcastle upon Tyne, UK (2011)
- Ask the second author!

The end

Questions?