CAT-centric view of formal languages, or, why is there not a 'category' of 'word concatenation things', what about Chomsky's hierarchy

https://golem.ph.utexas.edu/category/2010/09/categories_and_information_in.html

https://golem.ph.utexas.edu/category/2010/09/categories_and_information_in.html

In 1956, the famous linguist Noam Chomsky described a hierarchy of **formal grammars**: that is, rules for forming strings of symbols. Besides its importance in computer science, this 'Chomsky hierarchy' was a preliminary attempt to size up the problem of how humans can generate grammatical sentences.

Anyone with a drop of category-theoretic blood in their veins need only take one look at the **Chomsky hierarchy**, and they'll wish it was described in the language of category theory.

Indeed, the grammars in the lowest level in the Chomsky hierarchy, the **regular grammars**, are in one-to-one correspondence with **nondeterministic finite-state machines**. The grammar generates exactly the language that the corresponding machine accepts.

Later, Samuel Eilenberg, one of the fathers of category theory, shocked a lot of people by leaving algebraic topology and devoting a years to work on finite-state machines from a category-theoretic perspective:

- Samuel Eilenberg, ***Automata, Languages and Machines***, Vol. A, Academic Press, New York, 1974.
- Samuel Eilenberg, *Automata, Languages and Machines*, Vol. B, Academic Press, New York, 1976.

Then, later, Joachim Lambek did a lot of work on **category theory and grammar**...

... but I've got to go make breakfast now!

Posted by: John Baez on September 12, 2010 5:00 AM | Permalink | Reply to this

Automata, Languages and Machines

Mathematical tier	**Language**	**Chomsky language type**	**Basic Objects**	**Logic (in terms of cardinality of set of possible truth values)**	**Bounded by**
combinatory logic restricted to **I**	same name as tier	regular	**I**	0	1
combinatory logic	same name as tier	context-free	terms (built from **S** and **K** --- **I** is not needed since **I** = **SKK**)	1 (everything is true; there is no negation operator)	countable infinity
ZFC set theory	first order logic	context-sensitive	sets	2 (law of excluded middle)	least inaccessible cardinal
constructive type theory	typed λ-calculus	recursively enumerable (the most general Chomsky language)	functions	3 (true, false and undefined)	(?) some large cardinal X
category theory	extended natural language	(?) (a more general definition of language is needed)	objects and arrows	countable infinity (see subobject classifier of topos)	(?) some cardinal greater than X

A hierarchy of languages, logics, and mathematical theories

'Set Theory' has a wrong and totally misleading name.
It makes it look like 'the' theory of collections, but set theory is laden with retrospective foundational baggage.

On the previous pages and also in Chapter II we gave a number of effective procedures for deciding a number of questions. We did this without defining the notion of an "effective procedure." We simply described the procedure and assumed that the reader will find it "effective" enough for his satisfaction.

This chapter introduces the notions of an automaton and recognized by an automaton. The sets in question are subsets of a free monoid Σ^* where Σ is a finite alphabet. Thus we shall be dealing with sets of words in a finite alphabet. The approach is finitistic and explicit and all the arguments are effective and constructible. To some extent this will be compensated for in Chapter III where the approach will be somewhat more algebraic.

6.5 FO(+1) definability: LTT

This weakness of LT is, simply put, an insensitivity to quantity as opposed to simple occurrence. We can overcome this by adding *quantification* to our logical descriptions, that is, by moving from Propositional logic to a First-Order logic which we call FO(+1). Logical formulae of this sort make (Boolean combinations of) assertions about which symbols occur at which positions ($\sigma(x)$, where σ is a symbol in the vocabulary and x is a variable ranging over positions in a string), about the adjacency of positions ($x \triangleleft y$, which asserts that the position represented by y is the successor of that represented by x) and about the identity of positions ($x \approx y$), with the positions being quantified existentially ($\exists$) or universally ($\forall$). This allows us to distinguish, for example, one occurrence of a B from another:

$$\varphi_{\text{One-}B} = (\exists x)[B(x) \wedge (\neg\exists y)[B(y) \wedge \neg x \approx y]]$$

This FO(+1) formula requires that there is some position in the string (call it x) at which a B occurs and there is no position (y) at which a B occurs that is distinct from that ($\neg x \approx y$).

model theory Branch of logic that deals with the various relations between logical languages and their models. The paradigm is the model theory of first order logic, which was initiated by Tarski's truth definition, and which has reached a high degree of mathematical sophistication (Chang and Keisler, 1990). Some typical questions in model theory are the following. Given a class of models K, can we find some theory T that is true exactly in the models in K? There are several of these *definability theorems* for first-order logic. For example: (a) the class of finite models is not first-order definable; (b) the class of all infinite models is first-order definable; (c) there is no infinite cardinal number κ for which the class of all models of cardinality κ is first-order definable. Result (a) follows from the Compactness Theorem, and (c) from the Löwenheim–Skolem theorem, two famous results in model theory (see **logic**). A general result on definability is Keisler's theorem, which establishes necessary and sufficient conditions under which a class of models is definable: a class of models is first-order definable if and only if both it and its complement are closed under *ultraproducts* and *isomorphisms*.

Another usual type of question in model theory concerns the relation between some syntactic property of a formula (or a theory) and a property of models. Prime example is the Completeness theorem, which establishes the equivalence of derivability and validity (see **logic**). Other examples of this pattern are so called Preservation theorems, for example: a formula is preserved under submodels if and only if it is equivalent to a universal formula.

Abstract model theory is a general reflection on properties of logics and model theoretic results such as the Compactness theorem or Löwenheim–Skolem theorem (Barwise, 1974; Barwise and Feferman, 1985). A famous result here is Lindström's theorem, which states that first-order logic is the only *abstract* logic that has the Compactness and Löwenheim–Skolem properties.