

Computer language and category theory

Source: <http://sci.tech-archive.net/Archive/sci.math/2004-11/3730.html>

From: Jon Haugsand (jonhaug_at_ifi.uio.no)

Date: 11/18/04

Date: 18 Nov 2004 16:40:33 +0100

Is there any work done on computer languages and category theory? To clarify, what I am looking into is something like the following:

Given the language

$S ::= A \mid B$

$A ::= aa \mid ba \mid A-aa \mid B-ba$

$B ::= ab \mid bb \mid A-ab \mid B-bb$

Thus, the set of sentences in S is sequences

$\{a,b\}^* \{a,b\} \{a,b\}^* \{a,b\} \{a,b\}^* \dots$ such that the letter in front of the dash (-) is the same as the letter behind it.

I think this language can be described as a category as follows:

$C = \langle \text{Obj}, \text{Morph}, * \rangle$

$\text{Obj} = \{A, B\}$

$\text{Morph} = \text{Hom}(A, A) \cup \text{Hom}(A, B) \cup \text{Hom}(B, A) \cup \text{Hom}(B, B)$

$\text{Hom}(x, y) =$ any sentence s in S such that it starts with an x and ends with an y .

E.g. $bb-ba, ba, ba-aa-ab-ba \in \text{Hom}(B, A)$

The operator $*$ is defined as follows:

Given $x \in \text{Hom}(X, Y)$ and $y \in \text{Hom}(Y, Z)$ such that neither x nor y is a unit morphism. Then $x*y = x-y$

The unit elements of $\text{Hom}(A, A)$ and $\text{Hom}(B, B)$ is just an empty sentence handled specially:

$x * e = e * x = x$

(a) Does this look sensible?

However, I don't have such a simple language, but rather a slightly more complicated one:

$T ::= A \mid B$
 $A ::= aa \mid ba \mid A-aa \mid B-ba$
 $B ::= ab \mid bb \mid A-ab \mid B-bb \mid A-K$
 $K ::= (A)-B \mid (B)-B$

The K , introduced in order to be able to create a tree-like structure, makes the elements of T unsuited to form a set of morphisms, as each morphism in a category have exactly one source object and one destination object.

(b) What kind of mathematical tools should I study to handle language constructs like the language T ?

In reality, our objects are a little different, but not in essence. My colleague wants to define something like a "partial parametrized monoid", an animal I never before have encountered. Partiality follows from the fact that you cannot take two arbitrary elements of S (or T) and form a new element of S (or T). Parametrization does occur in some sense in the language T , where we must choose whether to insert an A sentence into the $(A)-B$ or a B into the same meta sentence. (Was this understandable?)

However, I don't like such animals. So I wonder if there is some use of category theory or something else that have been used to model language like constructs.

--

Jon Haugsand

Dept. of Informatics, Univ. of Oslo, Norway, <mailto:jonhaug@ifi.uio.no>
<http://www.ifi.uio.no/~jonhaug/>, Phone: +47 22 85 24 92