Technische Hogeschool, Eindhoven.
Onderafdeling der Wiskunde.
Notitie 11 , d.d. 10 december 1970.

Coding system for AUT-QE.

by N.G. de Bruijn.

The expressions and categories to be stored are all of the form EXPRESSION, as defined in the following syntax. The notion is a slight extension of those explained in [2] and [3].

The basic symbols are

type | genre | , | [|] | { | } | (|)

and, furthermore, the elements of the sets <variable>, <constant> and <dummy variable>. These three sets are disjoint; <variable> and <constant> contain positive integers only; <dummy variable> contains integers < -1000 only.

The notions <EXPRESSION> and <EXPRESSION string> are defined by:

<EXPRESSION string> ::= <EXPRESSION> | <EXPRESSION string> , <EXPRESSION>

<EXPRESSION> ::= type | genre | <constant> | <variable> | <dummy variable> |
<constant> (<EXPRESSION string>) | { <EXPRESSION>}<EXPRESSION> |
[<dummy variable> , <EXPRESSION>] <EXPRESSION>

There are three arrays in which the information about EXPRESSIONs and EXPRESSION strings is stored: list1[1:P], list2[1:P], list3[1:P].

Every integer k ($1 \le k \le P$) refers to an EXPRESSION string. In our present discussion we shall denote this string by Ω_k (metalingual symbol). If Ω_k has the form Ω_h, Λ (where Λ is an EXPRESSION) then we have list 1[k] = h; if Ω_k has the form Λ, where Λ is an EXPRESSION, we have list 1[k] = 0. The information about Λ is stored in list 2[k] and list 3[k].

If Λ = type then list 2[k] = 0, list 3[k]= -1000.

If Λ = genre then list 2[k] = 0, list 3[k] = -2000.

If Λ = c, where c $\in$ <constant>, then list 2[k] = c, list 3[k] = 0.

If Λ = x, where x $\in$ <variable>, or x $\in$ <dummy variable>, then list2[k] = x, list3[k] = -5000 or -4000.

The entry -4000 should not be used if Ω_k is not an indicator string (Ω_k is certainly no indicator string if x is a dummy variable).

If Λ has the form c(<EXPRESSION string>), and if that EXPRESSION string is Ω_h, then

list2[k] = c, list3[k] = h.

If Λ has the form $\{\Lambda_1\}\ \Lambda_2$, and if Ω_h is the EXPRESSION string Λ_1, Λ_2 (this string consists of just two expressions), then

list2[k] = -12, list3[k] = h.

If Λ has the form $[t, \Lambda_1]\ \Lambda_2$, and if Ω_h is the expression string Λ_1, Λ_2, then

list2[k] = t, list3[k] = h.

Note that the above system is obtained from the one in [1] for expressions of the form <constant> (<expression string>) if we add the following conventions:

<u>type</u>	is considered as			$0(\Omega_{-1000})$
<u>genre</u>	"	"	"	$0(\Omega_{-2000})$
c	"	"	"	$c(\Omega_0)$
x	"	"	"	$x(\Omega_{-4000})$ or $x(\Omega_{-5000})$
t	"	"	"	$t(\Omega_{-5000})$
$\{\Lambda_1\}\ \Lambda_2$	"	"	"	$-12(\Lambda_1, \Lambda_2)$
$[t, \Lambda_1]\ \Lambda_2$	"	"	"	$t(\Lambda_1, \Lambda_2)$

We did not put the empty string into our syntax. Nevertheless we consider the empty string occasionally, and we give it list number 0, i.e. Ω_0 represents the empty string.

We remind the reader of the definition of indicator string. An indicator string is either the empty string or a string of variables (satisfying the condition that the indicator string of the last variable is obtained by taking that last entry away). In the non-empty case it can, of course, be considered as an EXPRESSION string and will be stored as such.

The contents of a book are stored in three arrays: indstr[1:m], middle[1:m], cat[1:m].

If $1 \leq n \leq m$, and if the indicator string of the n-th line of the book is Ω_k, then indstr[n] = k.

If the middle part of the n-th line is an EXPRESSION Λ, and if Ω_k is the string consisting of the single entry Λ, then middle[n] = k. (Note that list1[k] = 0 in this case.)

If the middle part of the n-th line is PN, then middle[n] = -1.

If the middle part of the n-th line is EB, and if Ω_k is the extended indicator string of that line (i.e. the indicator string followed by n) then middle[n] = -100 -k.

If the middle part of the n-th line is not EB, and if cat[n] = k, then Ω_k is the EXPRESSION string consisting of just one entry, viz. the category part of the n-th line. (Whence list1[k] = 0 in this case.) If, however, the middle part is EB, then Ω_k is the <u>category string</u> of the extended indicator string of

that line. (If $x_1,\ldots,x_j,n$ is the extended indicator string, then this category string is $\Gamma_1,\ldots,\Gamma_j,\Gamma_{j+1}$, forming the categories of $x_1,\ldots,x_j,n$, respectively.)

Note that this difference between EB or non-EB applies to list1[cat[n]] only.

References.

[1] N.G. de Bruijn, A Processor for PAL. Internal report, Notitie 30,(26 maart 1970), Technological University Eindhoven.

[2] ———————— The syntax of PAL and AUTOMATH, Technological University Eindhoven, Internal Report, Notitie 32,(9 April 1970).

[3] ———————— On the use of bound variables in AUTOMATH, Technological University Eindhoven, Internal Report, Notitie 9, (26 November 1970).