L. H. Quam   12-June-69

A USER MODIFIABLE LISP SCANNER

LISP now uses a table driven scanner, whos table may be
modified by the user for the purpose of implementing scanners
for other languages. For simplicity, the functions for constructing
the scanner table initially give an ALGOL type scanner, that is, the
ALGOL definitions for identifiers, strings and numbers (except for
powers of ten). The ALGOL table may be deviated from using additional
functions to include additional characters in identifiers, and to
specify delimiters for strings.

IDENTIFIER

Syntax:

comment   ::=   (a comment-start followed by any sequence of characters
                          ending in a comment-end)

digit ::= "∅" | "1" | ... | "9"

letter ::= "A" | "B" | ... | "Z" | "a" | "b" ... | "z" | extra-letter

character ::= (any ASCII character other than null, rubout, and comment-start)

delimiter ::= (any character not a letter and not a digit)

identifier ::= letter
            ::= identifier digit
            ::= identifier letter
            ::= slashifier character
            ::= identifier slashifier character

Semantics:

   In the above syntax, underlined symbols can be specified by the user.
For instance, comment-start and comment-end could both be specified to be
~~double quote (")~~.   If it is desirable to include dollar and percent signs
in identifiers, then both of those characters should be extra-letters.
Slashifiers make it possible to include any delimiter in an identifier.
For example, if question mark is the slashifier (?) and it is desirable
to include the character (+) in the identifier V as V+, then one should
slashify the plus: V?+.

STRING

Syntax:

string ::= (a <u>string-start</u> followed any sequence of characters ending in a <u>string-end</u>)

Semantics

String-start and string-end are specifiable.

NUMBERS

<u>Syntax</u>:   Same is specified in SAILON 28.2, except that leading signs (+ or -) in numbers are treated as delimiters.

DESCRIPTIONS OF THE SCANNER MODIFYING FUNCTIONS

In the following descriptions, all characters are specified by their numerical ASCII value. For example, blank is 40_8 and A is 101_8.

(SCANINIT comment-start comment-end string-start string-end slashifier)

SCANINIT modifies the LISP scanner to be an ALGOL-type scanner with the special delimiters for comments and strings. SCANINIT must be called before any of the other scanner functions.

(LETTER x)

LETTER specifies to the scanner that x is an extra-letter, and thus allows x to be in an identifier.

(IGNORE x)

IGNORE specifies to the scanner that x is not to be returned as a delimiter from scan, but instead will be ignored. However, x will still function as a seperator between identifiers and numbers. For example, carriage-return, line-feed , tab and blank are useful characters to ignore as delimiters.

(SCAN)

SCAN reads an atom or delimiter and sets the value of the global variable SCANVAL to the value read, and returns a number corresponding to the syntactic type read as follows:

syntactic type	value of SCAN	value of SCANVAL
identifier	0	the uninterned identifier
string	1	the string
number	2	the value
delimiter	3	the ASCII numerical value of the delimiter.

(SCANSET)

  SCANSET modifies the LISP scanner in READ according to the user specifications. Evaluate (SCANSET) before calling SCAN.

(SCANRESET)

  SCANRESET unmodifies the LISP scanner to its normal state, and must be called before READ will work properly again once SCANSET is used. Use INITFN to call SCANRESET after errors.