

5. LOW LEVEL SPECIFICATIONS

We now develop a set of specifications that appear independent of Pascal grammar. Whereas the previous section left undefined the behaviour of indenting programs when invalid constructs of Pascal are given as input, this section specifies what transformation is to be done for an arbitrary input string, and hence an arbitrary sequence of tokens. This latter transformation is designed to coincide with that given above for all syntactically valid constructs of Pascal. An outline of a proof of this fact is given in the next section.

5.1. Lexical analysis

Lexical analysis is a process that breaks up strings into sequences of 'words', more widely known as tokens. We say a character string w is a *word* if $\text{TKN}(w)$ is not undefined, where TKN is a partial function that maps character strings to tokens as elaborated below.

Definition of TKN

For a given string w , $\text{TKN}(w)$ is defined as t if there is a pair $\langle w, t \rangle$ in one of the following sets; otherwise $\text{TKN}(w)$ is undefined.

1. Let w be free of delimiters, namely the following characters: blank, tab, end-of-line, end-of-file, parentheses, braces, semicolon, colon, asterisk, quote and period. (Other conventional delimiters do not concern us.)

{ <code>"procedure"</code>	, PROCEDURE},
<code>"function"</code>	, FUNCTION},
<code>"program"</code>	, PROGRAM},
<code>"forward"</code>	, FORWARD}
<code>"repeat"</code>	, REPEAT},
<code>"record"</code>	, RECORD},
<code>"extern"</code>	, EXTERN},
<code>"while"</code>	, WHILE},
<code>"until"</code>	, UNTIL},
<code>"label"</code>	, LABEL},
<code>"const"</code>	, CONST},
<code>"begin"</code>	, BEGIN},
<code>"with"</code>	, WITH},
<code>"type"</code>	, TYPE},
<code>"then"</code>	, THEN},
<code>"goto"</code>	, GOTO},
<code>"else"</code>	, ELSE},
<code>"case"</code>	, CASE},
<code>"var"</code>	, VAR},
<code>"for"</code>	, FOR},
<code>"end"</code>	, END},
<code>"of"</code>	, OF},
<code>"if"</code>	, IF},
<code>"do"</code>	, DO},
<code>other w</code>	, ORDINARY} }