Things that make scanning difficult
In PL/1 (an old IBM language) there are no reserved words. The following is a legitimate line of code:

```
IF IF THEN ELSE=THEN ELSE IF= THEN
```

This is a little easier to mentally parse if we write it:

```
IF (IF) THEN
  ELSE = THEN
ELSE
  IF = THEN
```

You have to determine whether a string is a keyword or variable from context.
In some old versions of FORTRAN identifiers could contain white space.

The statement

\texttt{DO 10 I = 1, 5}

means "do line 10 for variable \texttt{I} in the range from 1 to 5.

\texttt{DO 10 I = 1.5}

means variable \texttt{DO 10 I} is assigned value 1.5

You might need to read an arbitrary number of characters past the end of a token to determine what the token is.
Some languages have nested comments. DFAs can't handle nested comments, so a scanner for such languages has to mix other techniques into the usual scanner routines.