

编译原理 - 作业(1) : 词法分析

Due: Mar 16, 2021 at the beginning of the class (14:20)

Exercise 3.1.2 (P114): Tagged languages like HTML or XML are different from conventional programming languages in that the punctuation (tags) are either very numerous (as in HTML) or a user-defined set (as in XML). Further, tags can often have parameters. Suggest how to divide the following HTML document:

Here is a photo of my house:

<P>

See More Pictures if you liked that one.<P>

into appropriate lexemes. Which lexemes should get associated lexical values, and what should those values be?

Exercise 3.3.5 (P125): Write regular definitions for the following languages:

- b). All strings of lowercase letters in which the letters are in ascending lexicographic order (词典递增序排列).
- c). Comments, consisting of a string surrounded by /* and */, without an intervening */, unless it is inside double-quotes (").
- h). All strings of a's and b's that do not contain the substring abb.

Exercise 3.7.3 (P166): Convert the following regular expressions to DFAs:

- d). $(a|b)^*abb(a|b)^*$