

TEST 语言自动机

根据 TEST 语言的如下语法规则，实现基于自动机的词法分析器。

$\langle \text{letter} \rangle \rightarrow a|b|\dots|z|A|B|\dots|Z$

$\langle \text{digit} \rangle \rightarrow 1|2|\dots|9|0$

$\langle \text{number} \rangle \rightarrow \langle \text{digit} \rangle +$

$\langle \text{identifier} \rangle \rightarrow \langle \text{letter} \rangle (\langle \text{letter} \rangle | \langle \text{digit} \rangle)^*$

$\langle \text{singleword} \rangle \rightarrow +|-|*|()|\{ \} |:|,|;$

$\langle \text{division} \rangle \rightarrow /$

$\langle \text{comparison_single} \rangle \rightarrow <|>|!|=$

$\langle \text{comparison_double} \rangle \rightarrow >|=|<|=|!|=|$

$\langle \text{comment first} \rangle \rightarrow /*$

$\langle \text{commentlast} \rangle \rightarrow */$

将 identifier、number、singleword、comparison_single、comparison_double、division、comment 的 DFA 组装在一起，最后整体自动机如图 1。

由 Start 状态出发，以 F 为接受状态，F 状态边上的*表示从输入中多读了一个字符，需要回退到输入中。F_{comment} 边上的#表示注释开始后一直从输入中读取字符，直到*/为止。所有 F 状态还要再回到 Start 状态继续读取，直到读完所有输入字符串。

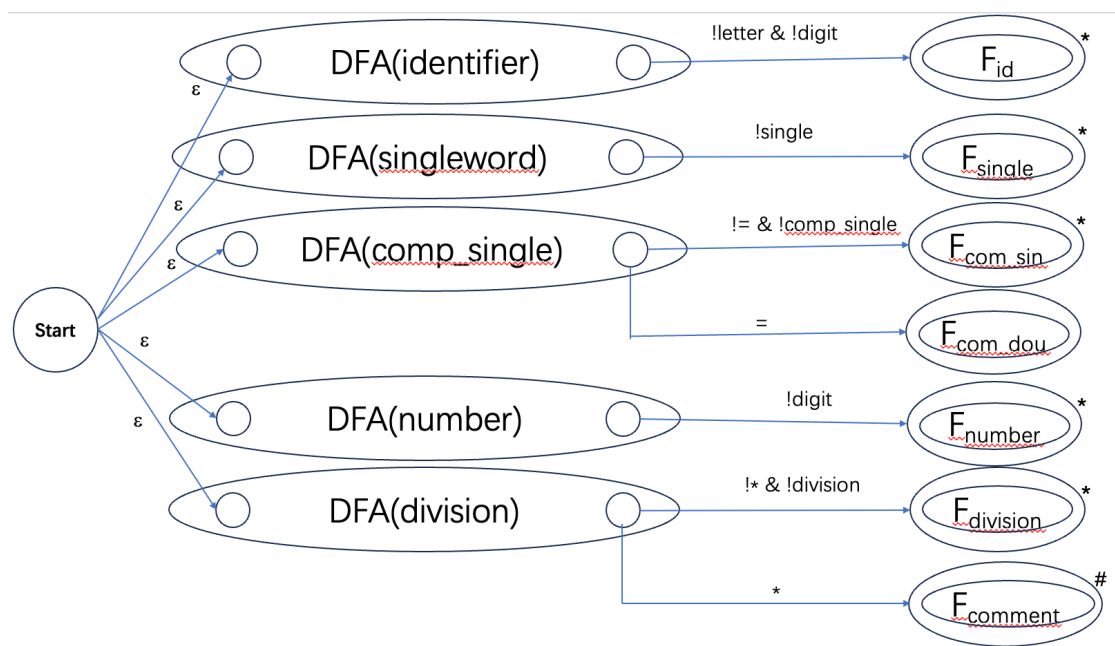


图 1. 组装所有 DFA