

西北工业大学

《编译原理》词法分析实验

学 院： 软件学院

学 号： 2018303081

姓 名： 马泽红

专 业： 软件工程

西北工业大学

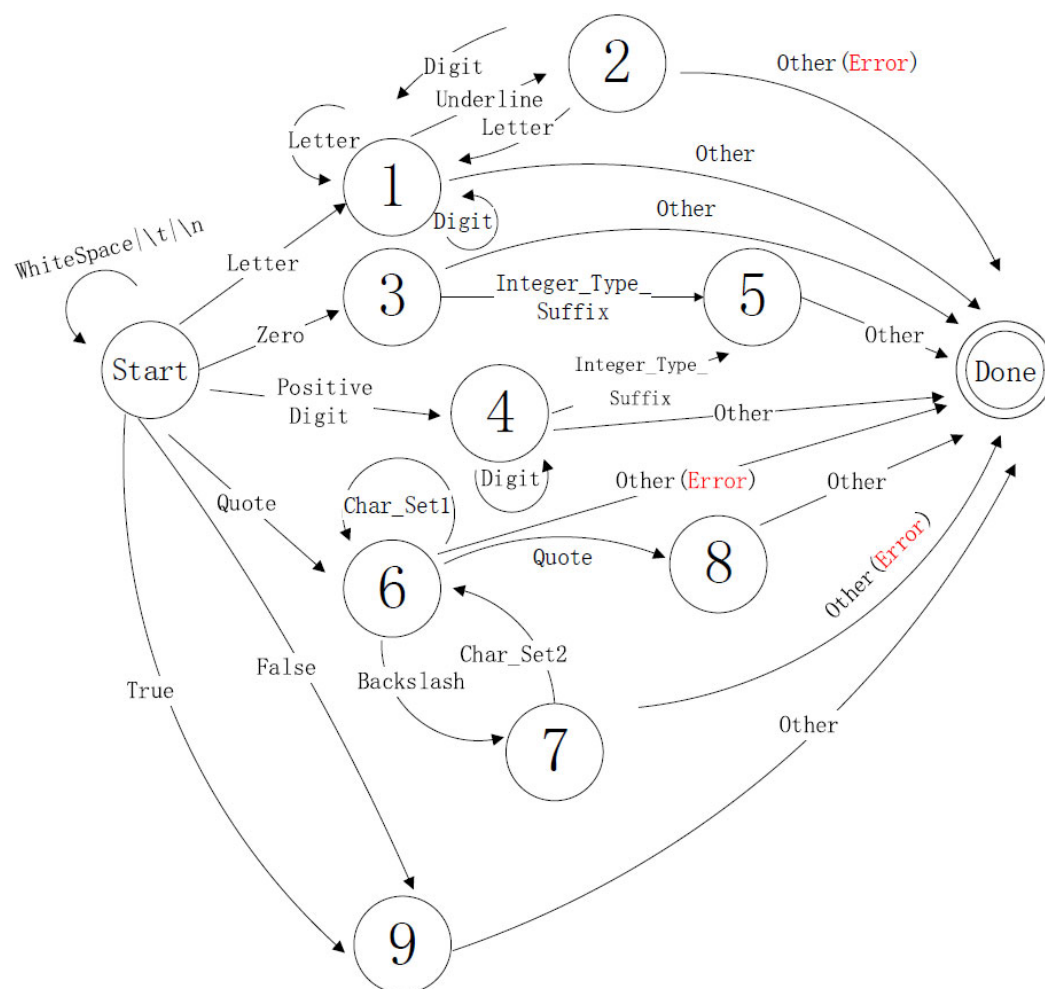
2021 年 4 月

目录

一.状态转换图（DFA）及描述	3
1.1 DFA 图	3
1.2 DFA 说明	3
1.2.1 正则表达式与状态对应关系	3
1.2.1 符号含义	4
二.构造 DFA	5
2.1 构造 NFA	5
2.2 NFA→DFA 映射	5
2.3 DFA 化简	7

一.状态转换图（DFA）及描述

1.1 DFA 图



1.2 DFA 说明

1.2.1 正则表达式与状态对应关系

ID 正则表达式由状态 Start, 1, 2, Done 来表示。

INTEGER 正则表达式由状态 Start, 3, 4, 5, Done 来表示。

STRING 正则表达式由状态 Start, 6, 7, 8, Done 来表示，其中状态 8 是为

了和其他正则表达式共同使用接受状态 Done 而引入的（如果想要精简的话可以删去）。

BOOLEAN 表达式由 START, 9, Done 来表示。

1.2.1 符号含义

Zero: 0

Underline: ' _ '

Positive Digit: [1-9]

Quote: ' " ' '

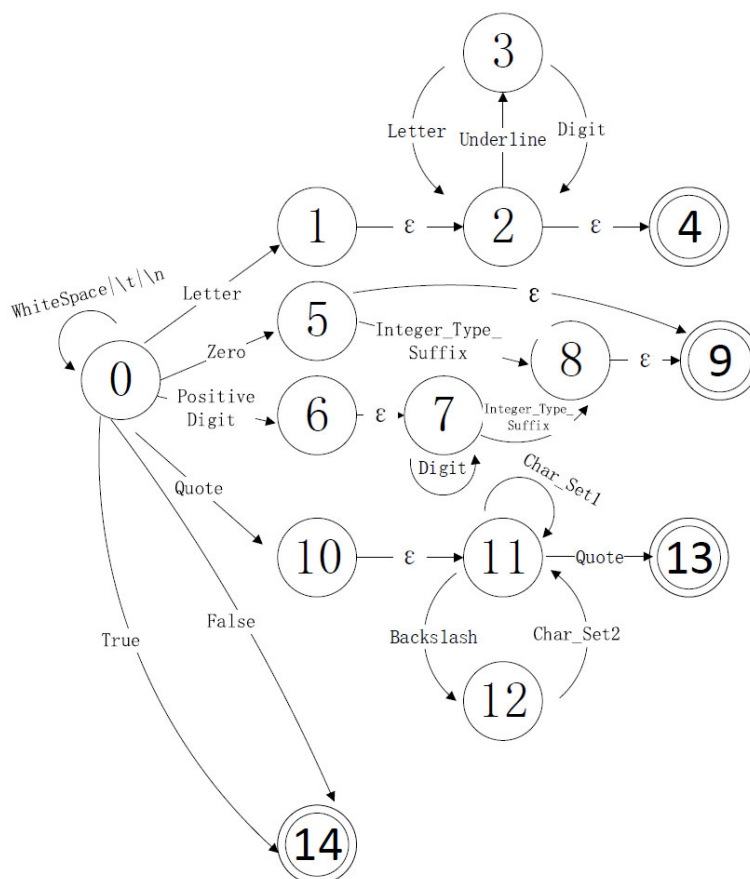
Backslash: ' \ '

Char_Set1: [~\|~"]

Char_Set2: [b | t | n | f | r | " | \]

二.构造 DFA

2.1 构造 NFA



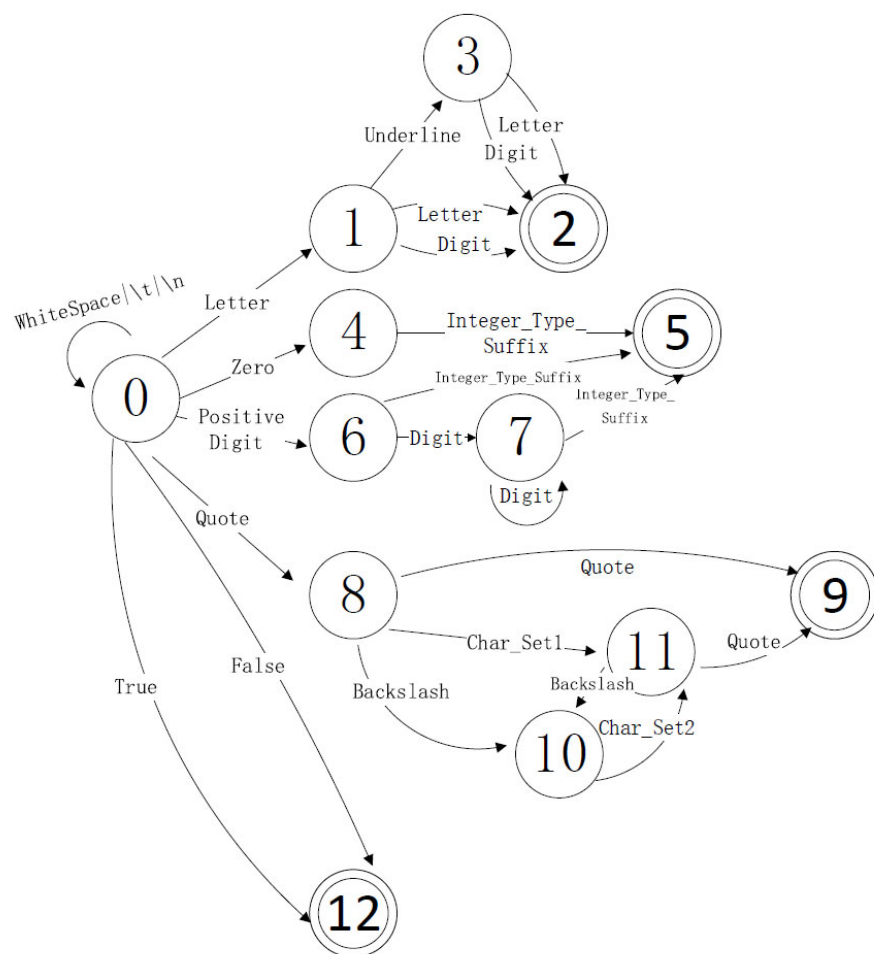
利用正则表达式构造 NFA 如上图，接受状态由于没有引入 Done 状态，所以有多个，每个正则表达式都有自己的接受状态。

2.2 NFA→DFA 映射

利用 NFA→DFA 的映射方法，根据 NFA 状态图，利用等价类闭包，我们可以求出以下表格，其中黄色标注的行表示没有新状态产生，即此支路映射完成。

状态\终结符	Letter	Digit	Underline	Zero	Positive Digit	Integer_Type_Suffix	Quote	Char_Set1	Back Slash	Char_Set2	False	True
0 {0}	{1,2,4}	∅	∅	{5,9}	{6,7,9}		{10,11}	∅	∅	∅	{14}	{14}
1 {1,2,4}	{2,4}	{2,4}	{3}	∅	∅	∅	∅	∅	∅	∅	∅	∅
2 {2,4}	{2,4}	{2,4}	∅	∅	∅		∅	∅	∅	∅	∅	∅
3 {3}	{2,4}	{2,4}	∅	∅	∅		∅	∅	∅	∅	∅	∅
4 {5,9}	∅	∅	∅	∅	∅	{8,9}	∅	∅	∅	∅	∅	∅
5 {8,9}	∅	∅	∅	∅	∅	∅	∅	∅	∅	∅	∅	∅
6 {6,7,9}	∅	{7,9}	∅	∅	∅	{8,9}	∅	∅	∅	∅	∅	∅
7 {7,9}	∅	{7,9}	∅	∅	∅	{8,9}	∅	∅	∅	∅	∅	∅
8 {10,11}	∅	∅	∅	∅	∅	∅	{13}	{11}	{12}	∅	∅	∅
9 {13}	∅	∅	∅	∅	∅	∅	∅	∅	∅	∅	∅	∅
10 {12}	∅	∅	∅	∅	∅	∅	∅	∅	∅	{11}	∅	∅
11 {11}	∅	∅	∅	∅	∅	∅	{13}	{11}	{12}	∅	∅	∅
12 {14}	∅	∅	∅	∅	∅	∅	∅	∅	∅	∅	∅	∅

利用上表，绘制出的 DFA 如下：



2.3 DFA 化简

利用等价类划分，对 2.2 中的 DFA 进行化简，可得到以下的 DFA：

