CSE 440 - Compiler Construction I

Instructor: Dr. Rida Bazzi
Notes written by Brett Hansen

Contents

1 Week of August 14th, 2016

1.1 Compiler Components

1. Syntax
   (a) Lexical Analysis
   (b) Parsing
2. Semantics
   (a) Type Checking *(Project 1)*
   (b) Code Generation
      i. Intermediate Representation *(Project 2)*
      ii. Analysis/Optimization *(Project 2)*
      iii. Low Level Code Generation & Optimization *(Project 3 & 4)*

1.2 Language Components

1. Character Set
2. Tokens: sequence of characters
   (a) identifiers
   (b) keywords
   (c) separators
   (d) operators
   (e) built-in types
   (f) modifiers

   Tokens are defined using regular expressions.

1.3 Regular Expressions

Regular Expression	**String Representation**
$\emptyset$	*empty set*
ϵ	*empty string*
a	{a}
$R_1 \vert R_2$	$L(R_1) \cup L(R_2)$
$R_1 \cdot R_2$	$L(R_1) \cdot L(R_2)$
R^*	$\{\epsilon\} \cup L(R) \cup L(R) \cdot L(R) \cup \ldots$
(R)	*grouping*

$L(R)$	The set of strings that R represents, also known as the *language of R*.
Precedence	$* \quad \rightarrow \quad \cdot \quad \rightarrow \quad \vert$

Regular Expression Example:

ID	*letter followed by 0 or more letters or digits*
letter	$a \mid b \mid \cdots \mid z \mid A \mid B \mid \cdots \mid Z$
digit	$0 \mid 1 \mid \cdots \mid 9$
ID	$\mathit{letter} \cdot (\mathit{letter} \mid \mathit{digit})^*$

From regular expressions $\longrightarrow$ non-deterministic finite state automata
$\hookrightarrow$ deterministic finite state automata
$\hookrightarrow$ program

2 Week of August 21st, 2016

2.1 Finite State Machines

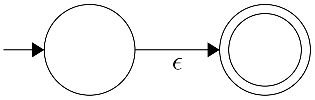
See Paper Notes

2.1.1 DFA and NFA Defintions

Definition 1 (Deterministic Finite State Automata)**.** *A deterministic finite state automata is a 5-tuple, $(Q, q_0, \Sigma, F, \delta)$ where:*
Q: finite set of states
q_0: initial state
Σ: alphabet
F: set of final states
δ: transition function, and
$\delta: \ Q \times \Sigma \longrightarrow Q$
$\delta^*: \ \delta^*(q, \epsilon) = q$
$\delta^*(q, aw) = \delta^*(\delta(q, a), w)$ where $a \in \Sigma$ and w is a string of elements of Σ
The string w is only accepted by the DFA if $\delta * (q_0, w) \in F$.

Definition 2 (Non-deterministic Finite State Automata)**.** *A non-deterministic finite state automata is a 5-tuple, $(Q, q_0, \Sigma, F, \delta)$ where:*
Q: finite set of states
q_0: initial state
Σ: alphabet
F: set of final states
δ: transition function, and
$\delta: \ Q \times \Sigma \cup \{\epsilon\} \longrightarrow 2^Q$ or $\mathbb{P}(Q)$

2.1.2 Regular Expressions to NFA

Regular Expression	**Equivalent NFA**
$\emptyset$	
ϵ	
a	
$R_1 \vert R_2$	
$R_1 \cdot R_2$	
R^*	
(R)	

3 Week of August 28th, 2016

4 Week of September 4th, 2016

5 Week of September 11th, 2016

6 Week of September 18th, 2016

7 Week of September 25th, 2016

8 Week of October 2nd, 2016

9 Week of October 9th, 2016

10 Week of October 16th, 2016

11 Week of October 23rd, 2016

12 Week of October 30th, 2016

13 Week of November 6th, 2016

14 Week of November 13th, 2016

15 Week of November 20th, 2016

16 Week of November 27th, 2016