What will we do in CPS 140?

Questions

- Can you write a program to determine if a string is an integer?
  
  9998.89 8abab 789342

- Can you do this if your machine had no additional memory other than the program? (can’t store any values and look at them again)

- Can you write a program to determine if the following are arithmetic expressions?
  
  $\left((34 + 7 \times \left(18 / 6\right))\right)$

  $\left(((((a + b) + c) \times d(e + f)))\right)$

- Can you do this if your machine had no additional memory other than the program?

- Can you write a program to determine the value of the following expression?
  
  $\left((34 + 7 \times \left(18 / 6\right))\right)$

- Can you write a program to determine if a file is a valid Java program?

- Can you write a program to determine if a Java program given as input will ever halt?

Language Hierarchy
### Power of Machines

<table>
<thead>
<tr>
<th>Automata</th>
<th>Can do?</th>
<th>Can’t do?</th>
</tr>
</thead>
<tbody>
<tr>
<td>Finite automata (FA) (no memory)</td>
<td>integers</td>
<td>arith expr</td>
</tr>
<tr>
<td>Pushdown automata (PDA) (only memory is stack)</td>
<td>arith expr</td>
<td>compute expr</td>
</tr>
<tr>
<td>Turing machines (TM) (infinite memory)</td>
<td>compute expr</td>
<td>decide if halts</td>
</tr>
</tbody>
</table>

### Application

**Compiler**

- Question: Given a Java program - is it valid?
- Question: language L, program P - is P valid?

![Diagram of compiler stages](image)

**Stages of a Compiler**

1. **C++ program**
2. **lexical analysis**
3. **tokens**
4. **syntax analysis**
5. **parse tree**
6. **code generation**
7. **assembly language program**
8. **executable**
9. **prog.exe**
L-Systems - Model the Growth of Plants

Set Theory - Read Chapter 1 Linz.

A Set is a collection of elements.

A={1,4,6,8}, B={2,4,8}, C={3,6,9,12,...}, D={4,8,12,16,...}

- (union) A∪B=
- (intersection) A∩B=
- C∩D=
- (member of) 42 ∈ C?
- (subset) B⊂C?
- B∩A ⊆ D?
- (product) A×B=
- |B|=
- |A×B|=
- 0 ∈ B∩C?
- (powerset) 2^B=

Example What are all the subsets of {3,5}?  

How many subsets does a set S have?

<table>
<thead>
<tr>
<th></th>
<th>number of subsets</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>3</td>
<td>4</td>
</tr>
</tbody>
</table>

How do you prove? Set S has 2^|S| subsets.
Technique: Proof by Induction

1. Basis: P(1)? Prove smallest instance is true.

2. Induction Hypothesis - I.H.
   Assume P(n) is true for 1,2,...,n

3. Induction Step - I.S.
   Show P(n+1) is true (using I.H.)

Proof of Example:

1. Basis:
2. I.H. Assume
3. I.S. Show

Ch. 1: 3 Major Concepts

- languages
- grammars
- automata

Languages

- \( \Sigma \) - set of symbols, alphabet
- string - finite sequence of symbols
- language - set of strings defined over \( \Sigma \)

Examples

- \( \Sigma = \{0, 1, 2, 3, 4, 5, 6, 7, 8, 9\} \)
  \( L=\{0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, \ldots \} \)
- \( \Sigma=\{a, b, c\} \)
  \( L=\{ab, ac, cabb\} \)
- \( \Sigma=\{a, b\} \)
  \( L=\{a^n b^n \mid n > 0\} \)
Notation

• symbols in alphabet: a, b, c, d, ...
• string names: u, v, w, ...

Definition of concatenation

Let \( w = a_1a_2 \ldots a_n \) and \( v = b_1b_2 \ldots b_m \)

Then \( w \circ v \) OR \( wv = \)

See book for formal definitions of other operations.

String Operations

strings: \( w = abbc, v = ab, u = c \)

• size of string
  \( |w| + |v| = \)
• concatenation
  \( v^3 = vvv = vovov = \)
• \( v^0 = \)
• \( w^R = \)
• \( |vv^Rw| = \)
• \( ab \circ \lambda = \)

Definition

\( \Sigma^* = \) set of strings obtained by concatenating 0 or more symbols from \( \Sigma \)

Example

\( \Sigma = \{a, b\} \)

\( \Sigma^* = \)

\( \Sigma^+ = \)

Examples

\( \Sigma = \{a, b, c\}, L_1 = \{ab, bc, aba\}, L_2 = \{c, bc, bcc\} \)

• \( L_1 \cup L_2 = \)
• \( L_1 \cap L_2 = \)
• \( L_1^c = \)
• \( L_1 \cap L_2 = \)
• \( L_1 \circ L_2 = \{xy \mid x \in L_1 \text{ and } y \in L_2\} = \)
**Definition**

\[ L^0 = \{ \lambda \} \]
\[ L^2 = L \circ L \]
\[ L^3 = L \circ L \circ L \]
\[ L^* = L^0 \cup L^1 \cup L^2 \cup L^3 \ldots \]
\[ L^+ = L^1 \cup L^2 \cup L^3 \ldots \]

**Grammars**

grammars for english

\[
\begin{align*}
\text{<sentence>} & \rightarrow \text{<subject><verb><d.o.>} \\
\text{<subject>} & \rightarrow \text{<noun> | <article><noun>} \\
\text{<verb>} & \rightarrow \text{hit | ran | ate} \\
\text{<d.o.>} & \rightarrow \text{<article><noun> | <noun>} \\
\text{<noun>} & \rightarrow \text{Fritz | ball} \\
\text{<article>} & \rightarrow \text{the | an | a}
\end{align*}
\]

**Examples** (derive a sentence)

Fritz hit the ball.

\[
\begin{align*}
\text{<sentence>} & \rightarrow \text{<subject><verb><d.o>} \\
& \rightarrow \text{<noun><verb><d.o.>} \\
& \rightarrow \text{Fritz <verb><d.o.>} \\
& \rightarrow \text{Fritz hit <d.o.>} \\
& \rightarrow \text{Fritz hit <article><noun>} \\
& \rightarrow \text{Fritz hit the <noun>} \\
& \rightarrow \text{Fritz hit the ball}
\end{align*}
\]

Can we also derive the sentences?

The ball hit Fritz.

The ball ate the ball

Syntactically correct?

Semantically correct?
Grammar

G=(V,T,S,P) where

- V - variables (or nonterminals)
- T - terminals
- S - start variable (S \in V)
- P - productions (rules)
  \( x \rightarrow y \) “means” replace \( x \) by \( y \)
  \( x \in (V \cup T)^+ \), \( y \in (V \cup T)^* \)
  where \( V, T, \) and \( P \) are finite sets.

Definition

- \( w \Rightarrow z \) \( w \) derives \( z \)
- \( w \Rightarrow^* z \) derives in 0 or more steps
- \( w \Rightarrow^+ z \) derives in 1 or more steps

Definition

\[ L(G) = \{ w \in T^* | S \Rightarrow^* w \} \]

Example

\( G=(\{S\}, \{a,b\}, S, P) \)
\( P \{S \rightarrow aaS, S \rightarrow b\} \)
\( L(G)= \)

Example

\( L(G) = \{ a^nccb^n | n > 0 \} \)
\( G = \)

Example

\( G=(\{S\}, \{a,b\}, S, P) \)
\( P \{S \rightarrow aSb, S \rightarrow SS, S \rightarrow ab\} \)

Which of these strings \( aabb, abab, abba, babab \) can be generated by this grammar? Show the derivations.
\( L(G) = \)
Automata Abstract model of a digital computer

![Automata Diagram]