Announcements:

- This is a math course with applications. Prereq: Compsci 230 or equivalent, CompSci 201.
- Course web page:
www.cs.duke.edu/courses /
spring21/compsci334
Familiarize yourself with all parts of the web page.
- Read Chapter 1 in the Linz book for next time.
- Complete the reading quiz on Sakai.
- Course bulletin board: Piazza

What will we do in Compsci 334 ?
Questions

- Can you write a program to determine if a string is an integer? $9998.89 \quad 8 \mathrm{abab} \quad 789342$
- 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 correct arithmetic expressions?

$$
\begin{gathered}
((34+7 *(18 / 6))) \\
(((((((a+b)+c) * d(e+f)))))
\end{gathered}
$$

- 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?

$$
((34+7 *(18 / 6)))
$$

- 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

## Grammars

## Automata



# Power of Machines 

## automata <br> Can do? <br> Can't do?

FA integers arith expr
(no memory)
PDA
arith expr
compute expr (stack)

TM
compute expr decide if halts (infinite)

Application
Compiler

- Our focus - Question: Given a program in some language (say Java or $\mathbf{C}++$ ) - is it valid?
- Question: language L, program P is P valid?
- Other things to consider, how is the compilation process different for different programming languages? (Java vs C ++ ?)


## Stages of a Compiler


assembly language program

## L-Systems - Model the Growth of Plants



Chapter 1 - Set Theory
A Set is a collection of elements.
$\mathrm{A}=\{1,4,6,8\}, \mathrm{B}=\{2,4,8\}$,
$\mathrm{C}=\{3,6,9,12, \ldots\}, \mathrm{D}=\{4,8,12,16, \ldots\}$

- (union) $\mathrm{A} \cup \mathrm{B}=$
- (intersection) $\mathbf{A} \cap \mathrm{B}=$
- $\mathrm{C} \cap \mathrm{D}=$
- (member of) $42 \in \mathrm{C}$ ?
- (subset) $\mathrm{B} \subset \mathrm{C}$ ?
- $\mathrm{B} \cap \mathrm{A} \subseteq \mathrm{D}$ ?
- $|\mathbf{B}|=$
- (product) $\mathbf{A} \times \mathbf{B}=$
- $|\mathbf{A} \times \mathbf{B}|=$
- $\emptyset \in \mathrm{B} \cap \mathrm{C}$ ?
- (powerset) $2^{B}=$

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

How many subsets does a set S have?
$|S|$ number of subsets
0
1
2
3
4

How do you prove? Set $\mathbf{S}$ has $2^{|S|}$ subsets.

## Technique: Proof by Induction

1. Basis: $P(1)$ ?
2. I.H.

Assume $P(n)$ is true for $1,2, \ldots, n$ 3. 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 $\Sigma$
alphabet $\Sigma$
Examples
- $\Sigma=\{0,1,2,3,4,5,6,7,8,9\}$ $\mathbf{L}=\{0,1,2, \ldots, 12,13,14, \ldots\}$
- $\Sigma=\{a, b, c\}$

$$
\mathbf{L}=\{a b, a c, c a b b\}
$$

- $\Sigma=\{a, b\}$

$$
\mathbf{L}=\left\{a^{n} b^{n} \mid n>0\right\}
$$

## Notation

- symbols in alphabet: a, b, c, d, ... - string names: $u, v, w, \ldots$


## Definition of concatenation

Let $\mathbf{w}=a_{1} a_{2} \ldots a_{n}$ and $\mathbf{v}=b_{1} b_{2} \ldots b_{m}$
Then $w \circ v$ OR $\mathbf{w v}=$

## String Operations

strings: $\mathbf{w}=\mathrm{abbc}, \mathrm{v}=\mathrm{ab}, \mathrm{u}=\mathbf{c}$

- size of string
$|w|+|v|=$
- concatenation
$v^{3}=\mathbf{v v v}=\mathbf{v} \circ \mathbf{v} \circ \mathbf{v}=$
- $v^{0}=$
- $w^{R}=$
- $\left|v v^{R} w\right|=$
- ab $\circ \lambda=$


## Definition

$\Sigma^{*}$ concatenate 0 or more
Example
$\Sigma=\{a, b\}$
$\Sigma^{*}=$
$\Sigma^{+}=$

## Examples

$\Sigma=\{a, b, c\}, L_{1}=\{a b, b c, a b a\}$,
$L_{2}=\{c, b c, b c c\}$

- $L_{1} \cup L_{2}=$
- $L_{1} \cap L_{2}=$
- $\overline{L_{1}}=$
- $\overline{L_{1} \cap L_{2}}=$
- $L_{1} \circ L_{2}=\left\{x y \mid x \in L_{1}\right.$ and $\left.y \in L_{2}\right\}=$


## Definition

$$
\begin{aligned}
& L^{0}=\{\lambda\} \\
& L^{2}=L \circ L \\
& L^{3}=L \circ L \circ L \\
& L^{*}=L^{0} \cup L^{1} \cup L^{2} \cup L^{3} \ldots \\
& L^{+}=L^{1} \cup L^{2} \cup L^{3} \ldots
\end{aligned}
$$

Grammars
Grammar for english
$<$ sentence $>\rightarrow$
<subject><verb><d.o.>
$<$ subject $>\rightarrow<$ noun $>\mid$
$<$ article $><$ noun $>$
$<$ verb $>\rightarrow$ hit $\mid$ ran $\mid$ ate
$<$ d.o. $>\rightarrow<$ article $><$ noun $>\mid<$ noun $>$
$<$ noun $>\rightarrow$ Fritz $\mid$ ball
$<$ article $>\rightarrow$ the $\mid$ an $\mid$ a

Examples (derive a sentence)
Fritz hit the ball.
<sentence> -> <subject><verb><d.o>
-> <noun><verb><d.o>
-> Fritz <verb><d.o.>
-> Fritz hit <d.o.>
-> Fritz hit <article><noun>
-> Fritz hit the <noun>
-> Fritz hit the ball

# 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 $(\mathbf{S} \in \mathrm{V})$
- P - productions (rules)
$\mathbf{x} \rightarrow \mathbf{y} \in(\mathbf{V} \cup \mathbf{T})^{+}, \mathbf{y} \in(\mathbf{V} \cup \mathbf{T})^{*}$
Definition
$\mathrm{w} \Rightarrow \mathrm{z} \quad \mathrm{w}$ derives z
$\mathbf{w} \xrightarrow{*} \mathbf{z}$ derives in 0 or more steps
$w \stackrel{ \pm}{\Rightarrow} \mathbf{z}$ derives in 1 or more steps

Definition
$\mathbf{G}=(\mathbf{V}, \mathbf{T}, \mathbf{S}, \mathbf{P})$
$\mathbf{L}(\mathbf{G})=\left\{\mathbf{w} \in T^{*} \mid \mathbf{S} \stackrel{*}{\Rightarrow} \mathbf{w}\right\}$

## Example

$\mathbf{G}=(\{\mathbf{S}\},\{\mathbf{a}, \mathbf{b}\}, \mathbf{S}, \mathbf{P})$
$\mathbf{P}=\{\mathbf{S} \rightarrow \mathbf{a} \mathbf{a}, \mathbf{S} \rightarrow \mathbf{b}\}$
$\mathrm{L}(\mathrm{G})=$

## Example

$\mathbf{L}(\mathbf{G})=\left\{a^{n} c c b^{n} \mid n>0\right\}$
$\mathrm{G}=$

Example
$\mathbf{G}=(\{\mathbf{S}\},\{\mathbf{a}, \mathbf{b}\}, \mathbf{S}, \mathbf{P})$
$\mathbf{P}=\{\mathbf{S} \rightarrow \mathbf{a S b}, \mathbf{S} \rightarrow \mathbf{S S}, \mathbf{S} \rightarrow \mathbf{a b}\}$
Which of these strings
$a a b b, a b a b, a b b a, b a b a b$ can be generated by this grammar? Show the derivations.
$\mathrm{L}(\mathrm{G})=$

## Automata

## Abstract model of a digital computer



