## CSE 311: Foundations of Computing

## Lecture 20: Structural Induction, Regular Expressions



## Last Time: Recursive Definitions

- Any recursively defined set can be translated into a Java class
- Any recursively defined function can be translated into a Java function
- some (but not all) can be written more cleanly as loops
- Recursively defined functions and sets are our mathematical models of code and the data it operates on


## Last time: Structural Induction

How to prove $\forall x \in S, P(x)$ is true:

Base Case: Show that $P(u)$ is true for all specific elements $u$ of $S$ mentioned in the Basis step

Inductive Hypothesis: Assume that $P$ is true for some arbitrary values of each of the existing named elements mentioned in the Recursive step

Inductive Step: Prove that $P(w)$ holds for each of the new elements $w$ constructed in the Recursive step using the named elements mentioned in the Inductive Hypothesis

Conclude that $\forall x \in S, P(x)$

## Linked Lists of Integers

- Basis: null $\in$ Lists
- Recursive step:


## If $L \in$ Lists and $v \in \mathbb{Z}$, then $\operatorname{Node}(v, L) \in$ Lists

Examples:

- null
[]
- Node(1, null)
[1]
- Node(1, Node(2, null))
[1, 2]


## Functions on Linked Lists

Set of numbers stored in a list:

- values(null) = $\varnothing$
- values $(\operatorname{Node}(\mathrm{v}, \mathrm{L}))=\{\mathrm{v}\} \cup$ values $(\mathrm{L})$

Example:
values(Node(1, Node(2, null))

$$
\begin{array}{ll}
=\{1\} \cup \text { values(Node(2, null) } & \\
=\{1\} \cup\{2\} \cup \text { Dalues of values } \\
=\{1\} \cup\{2\} \cup \emptyset & \\
& \text { Def of values } \\
=\{1,2\} & \\
\text { Def of values } \\
& \\
\text { Def of } \cup
\end{array}
$$

## Functions on Linked Lists

Remove the numbers that don't satisfy $p(v)$ :

- filter $\left.{ }^{(n u l l}\right)=$ null
- filter $_{p}(\operatorname{Node}(\mathrm{v}, \mathrm{L}))=\operatorname{Node}\left(\mathrm{v}\right.$, filter $\left._{p}(\mathrm{~L})\right) \quad$ if $p(\mathrm{v})$
- filter $_{p}(\operatorname{Node}(v, L))=$ filter $_{p}(\mathrm{~L})$
otherwise

Example: $\mathrm{p}(\mathrm{v}):=\mathrm{v}<2$
filter ${ }_{p}(\operatorname{Node}(1, \operatorname{Node}(2$, null)))
$=\operatorname{Node}\left(1\right.$, filter $_{\mathrm{p}}(\operatorname{Node}(2$, null) $))$
$=$ Node(1, filter ${ }_{p}($ null $)$ )
$=$ Node(1, null)

Def filter ${ }_{p}$
Def filter ${ }_{p}$
Def filter ${ }_{p}$

## Claim: $x \in$ values(filter $\left.{ }_{p}(L)\right)$ iff $p(x) \wedge x \in \operatorname{values}(L)$

## Claim: $x \in$ values(filter $\left.{ }_{p}(L)\right)$ iff $p(x) \wedge x \in \operatorname{values}(L)$

$Q(L):=$ " $x \in$ values $^{(f i l t e r}(\mathrm{L})$ ) iff $p(x) \wedge x \in$ values $(L)$ for all $x \in \mathbb{Z}$ ". We will prove $Q(L)$ for $L \in$ Lists by structural induction.

## Claim: $x \in$ values(filter $\left.{ }_{p}(\mathrm{~L})\right)$ iff $p(\mathrm{x}) \wedge x \in \operatorname{values}(\mathrm{~L})$

$Q(L):=$ " $x \in$ values(filter $(L))$ iff $p(x) \wedge x \in$ values $(L)$ for all $x \in \mathbb{Z}$ ".
We will prove $Q(L)$ for $L \in$ Lists by structural induction.
Base Case: Let $x \in \mathbb{Z}$ be arbitrary.
LHS is $x \in$ values(filter ${ }_{p}($ null))

$$
\begin{array}{ll}
\equiv x \in \text { values(null) } & \text { Def of filter }{ }_{p} \\
\equiv x \in \emptyset & \text { Def of values } \\
\equiv F & \text { Def of } \varnothing
\end{array}
$$

RHS is $p(x) \wedge x \in$ values(null)

$$
\begin{aligned}
& \equiv p(x) \wedge x \in \emptyset \\
& \equiv p(x) \wedge F \\
& \equiv F
\end{aligned}
$$

Def of values
Def of $\varnothing$
Domination
These are equivalent as required (LHS $\equiv \mathrm{F} \equiv \mathrm{RHS}$ ).
Since $x$ was arbitrary, this shows that $Q$ (null) holds.

## Claim: $x \in$ values(filter $\left.{ }_{p}(\mathrm{~L})\right)$ iff $p(x) \wedge x \in \operatorname{values}(\mathrm{~L})$

$Q(L):=$ " $x \in$ values(filter $_{p}(L)$ ) iff $p(x) \wedge x \in$ values $(L)$ for all $x \in \mathbb{Z}$ ". We will prove $Q(L)$ for $L \in$ Lists by structural induction.
Base Case: ... so $Q$ (null) holds.
Inductive Hypothesis: Suppose $Q(L)$ holds for an arbitrary list L,
i.e., we have $x \in$ values(filter $(L))$ iff $p(x) \wedge x \in$ values(L).

Inductive Step: Goal: Prove $Q(\operatorname{Node}(\mathrm{v}, \mathrm{L}))$ for all $v \in \mathbb{Z}$

## Claim: $v \in$ values(filter $\left.{ }_{p}(\mathrm{~L})\right)$ iff $p(x) \wedge x \in$ values $(\mathrm{L})$

$Q(L):=$ " $x \in$ values $^{(\text {filter }}(\mathrm{L}(\mathrm{L})$ ) iff $p(x) \wedge x \in$ values $(L)$ for all $x \in \mathbb{Z}$ ".
We will prove $Q(L)$ for $L \in$ Lists by structural induction.
Base Case: ... so $Q$ (null) holds.
Inductive Hypothesis: Suppose $Q(L)$ holds for an arbitrary list $L$,
i.e., we have $x \in$ values $\left(\right.$ filter $\left._{p}(L)\right)$ iff $p(x) \wedge x \in$ values $(L)$.

Inductive Step: Goal: Prove $Q($ Node( $v, L)$ ) for all $v \in \mathbb{Z}$
Let $v, x \in \mathbb{Z}$ be arbitrary. We go by cases. Suppose $\neg p(v)$.
$x \in \operatorname{values}\left(\right.$ filter $\left._{p}(\operatorname{Node}(v, L))\right)$
$\equiv x \in$ values(filter $\left.{ }_{p}(L)\right) \quad$ Def filter $_{p}$
$\equiv \mathrm{p}(\mathrm{x}) \wedge \mathrm{x} \in \operatorname{values}(\mathrm{L})$
IH
$\equiv \mathrm{p}(\mathrm{x}) \wedge \mathrm{x} \in \operatorname{values}(\operatorname{Node}(\mathrm{v}, \mathrm{L}))$

## Claim: $v \in$ values(filter $\left.{ }_{p}(\mathrm{~L})\right)$ iff $p(x) \wedge x \in$ values $(\mathrm{L})$

$Q(L):=$ " $x \in$ values(filter $(L))$ iff $p(x) \wedge x \in$ values $(L)$ for all $x \in \mathbb{Z}$ ".
We will prove $Q(L)$ for $L \in$ Lists by structural induction.
Base Case: ... so $Q$ (null) holds.
Inductive Hypothesis: Suppose $Q(L)$ holds for an arbitrary list $L$,
i.e., we have $x \in$ values(filter $(\mathrm{L})$ ) iff $p(x) \wedge x \in$ values $(\mathrm{L})$.

Inductive Step: Goal: Prove $Q($ Node( $v, L)$ ) for all $v \in \mathbb{Z}$
Let $v, x \in \mathbb{Z}$ be arbitrary. We go by cases. Suppose $\neg p(v)$.
$x \in$ values(filter $\left.{ }_{p}(\operatorname{Node}(v, L))\right)$
$\equiv x \in$ values(filter $(L)) \quad$ Def filter $_{p}$
$\equiv p(x) \wedge x \in \operatorname{values}(L) \quad I H$
If $\neg p(x)$, then this and $p(x) \wedge x \in \operatorname{values}(\operatorname{Node}(v, L))$ are equivalent as they are both false. So now suppose $p(x)$...

## Claim: $v \in$ values(filter $\left.{ }_{p}(\mathrm{~L})\right)$ iff $p(x) \wedge x \in$ values $(\mathrm{L})$

$Q(L):=$ " $x \in$ values(filter $\left._{p}(L)\right)$ iff $p(x) \wedge x \in$ values(L) for all $x \in \mathbb{Z}$ ".
We will prove $Q(L)$ for $L \in$ Lists by structural induction.
Base Case: ... so Q(null) holds.
Inductive Hypothesis: Suppose $Q(L)$ holds for an arbitrary list $L$,
i.e., we have $x \in$ values $\left(\right.$ filter $\left._{p}(L)\right)$ iff $p(x) \wedge x \in$ values $(L)$.

Inductive Step: Goal: Prove $Q($ Node( $v, L)$ ) for all $v \in \mathbb{Z}$
Let $v, x \in \mathbb{Z}$ be arbitrary. We go by cases. Suppose $\neg p(v)$.
$x \in$ values(filter $\left.{ }_{p}(\operatorname{Node}(v, L))\right)$
$\equiv x \in$ values(filter $(L)) \quad$ Def filter $_{p}$
$\equiv \mathrm{p}(\mathrm{x}) \wedge \mathrm{x} \in \operatorname{values}(\mathrm{L})$
$\equiv \mathrm{p}(\mathrm{x}) \wedge(\mathrm{x} \in\{\mathrm{v}\} \vee \mathrm{x} \in \operatorname{values}(\mathrm{L}))$
$\equiv \mathrm{p}(\mathrm{x}) \wedge(\mathrm{x} \in\{\mathrm{v}\} \cup$ values $(\mathrm{L})) \quad$ Def $\cup$
$\equiv p(x) \wedge(x \in$ values(Node(v, L))) Def values

## Claim: $v \in$ values(filter $\left.{ }_{p}(\mathrm{~L})\right)$ iff $p(x) \wedge x \in$ values $(\mathrm{L})$

$Q(L):=$ " $x \in$ values(filter $\left._{p}(L)\right)$ iff $p(x) \wedge x \in$ values(L) for all $x \in \mathbb{Z}$ ".
We will prove $Q(L)$ for $L \in$ Lists by structural induction.
Base Case: ... so $Q$ (null) holds.
Inductive Hypothesis: Suppose $Q(L)$ holds for an arbitrary list $L$,
i.e., we have $x \in$ values $\left(\right.$ filter $\left._{p}(L)\right)$ iff $p(x) \wedge x \in$ values $(L)$.

Inductive Step: Goal: Prove $\mathrm{Q}($ Node( $v, L)$ ) for all $v \in \mathbb{Z}$
Let $v, x \in \mathbb{Z}$ be arbitrary. We go by cases. Suppose $\neg p(v)$.
$x \in \operatorname{values}\left(\right.$ filter $\left._{p}(\operatorname{Node}(v, L))\right)$
$\equiv x \in$ values(filter ${ }_{p}(\mathrm{~L})$ )
$\equiv \mathrm{p}(\mathrm{x}) \wedge x \in \operatorname{values}(\mathrm{~L})$
$\equiv \mathrm{p}(\mathrm{x}) \wedge(\mathrm{F} \vee \mathrm{x} \in \operatorname{values}(\mathrm{L}))$
$\equiv \mathrm{p}(\mathrm{x}) \wedge(\mathrm{x} \in\{\mathrm{v}\} \vee \mathrm{x} \in \operatorname{values}(\mathrm{L}))$
$\equiv p(x) \wedge(x \in\{v\} \cup$ values $(\mathrm{L}))$
$\equiv p(x) \wedge(x \in \operatorname{values}(\operatorname{Node}(v, L))) \quad$ Def values

Def filter ${ }_{p}$
IH
Identity
suppose $p(x)$...

## Claim: $v \in$ values(filter $\left.{ }_{p}(\mathrm{~L})\right)$ iff $p(x) \wedge x \in$ values $(\mathrm{L})$

$Q(L):=$ " $x \in$ values(filter $\left._{p}(L)\right)$ iff $p(x) \wedge x \in$ values(L) for all $x \in \mathbb{Z}$ ".
We will prove $Q(L)$ for $L \in$ Lists by structural induction.
Base Case: ... so $Q$ (null) holds.
Inductive Hypothesis: Suppose $Q(L)$ holds for an arbitrary list $L$,
i.e., we have $x \in$ values(filter $(\mathrm{L})$ ) iff $p(x) \wedge x \in$ values $(\mathrm{L})$.

Inductive Step: Goal: Prove $\mathrm{Q}($ Node( $v, L)$ ) for all $v \in \mathbb{Z}$
Let $v, x \in \mathbb{Z}$ be arbitrary. We go by cases. Suppose $\neg p(v)$.
$x \in \operatorname{values}\left(\right.$ filter $\left._{p}(\operatorname{Node}(v, L))\right)$
$\equiv x \in$ values(filter ${ }_{p}(\mathrm{~L})$ )
$\equiv \mathrm{p}(\mathrm{x}) \wedge x \in \operatorname{values}(\mathrm{~L})$
$\equiv \mathrm{p}(\mathrm{x}) \wedge(\mathrm{F} \vee \mathrm{x} \in$ values $(\mathrm{L}))$
$\equiv p(x) \wedge(x \in\{v\} \vee x \in \operatorname{values}(\mathrm{~L}))$
$\equiv p(x) \wedge(x \in\{v\} \cup$ values $(\mathrm{L}))$
$\equiv p(x) \wedge(x \in \operatorname{values}(\operatorname{Node}(v, L)))$
Def filter ${ }_{p}$
IH
Identity
$\mathrm{x} \neq \mathrm{v}$ as $\mathrm{p}(\mathrm{x})$ but $\neg \mathrm{p}(\mathrm{v})$
Def $U$
Def values

## Claim: $v \in$ values(filter $\left.{ }_{p}(\mathrm{~L})\right)$ iff $p(x) \wedge x \in$ values $(\mathrm{L})$

$Q(L):=$ " $x \in$ values(filter $(L))$ iff $p(x) \wedge x \in$ values $(L)$ for all $x \in \mathbb{Z}$ ".
We will prove $Q(L)$ for $L \in$ Lists by structural induction.
Base Case: ... so $Q$ (null) holds.
Inductive Hypothesis: Suppose $Q(L)$ holds for an arbitrary list $L$,
i.e., we have $x \in$ values(filter $(\mathrm{L})$ ) iff $p(x) \wedge x \in$ values $(\mathrm{L})$.

Inductive Step: Goal: Prove $Q($ Node( $v, L)$ ) for all $v \in \mathbb{Z}$
Let $v, x \in \mathbb{Z}$ be arbitrary. We go by cases. Suppose $\neg p(v)$.
$x \in \operatorname{values}\left(\right.$ filter $\left._{p}(\operatorname{Node}(v, L))\right)$
$\equiv$...
$\equiv \mathrm{p}(\mathrm{x}) \wedge(\mathrm{x} \in \operatorname{values}(\operatorname{Node}(\mathrm{v}, \mathrm{L})))$
Thus, by cases $(p(x) \& \neg p(x))$, the claimed bicondition holds. Since $x$ was arbitrary, we have shown $Q(\operatorname{Node}(v, L))$.

## Claim: $v \in$ values(filter $\left.{ }_{p}(\mathrm{~L})\right)$ iff $p(x) \wedge x \in$ values $(\mathrm{L})$

$Q(L):=$ " $x \in$ values(filter $(L))$ iff $p(x) \wedge x \in$ values $(L)$ for all $x \in \mathbb{Z}$ ".
We will prove $Q(L)$ for $L \in$ Lists by structural induction.
Base Case: ... so $Q$ (null) holds.
Inductive Hypothesis: Suppose $Q(L)$ holds for an arbitrary list $L$,
i.e., we have $x \in$ values $\left(\right.$ filter $\left._{p}(L)\right)$ iff $p(x) \wedge x \in$ values $(L)$.

Inductive Step: Goal: Prove $Q($ Node( $v, L)$ ) for all $v \in \mathbb{Z}$
Let $v, x \in \mathbb{Z}$ be arbitrary. We go by cases. Suppose $p(v)$.
$x \in \operatorname{values}\left(\right.$ filter $\left._{p}(\operatorname{Node}(v, L))\right)$
$\equiv x \in \operatorname{values}\left(\operatorname{Node}^{\left.\left(v, \text { filter }_{p}(L)\right)\right) \quad \text { Def filter }_{p}}\right.$
$\equiv x \in\{v\} \cup$ values(filter $\left.{ }_{p}(\mathrm{~L})\right)$
$\equiv x \in\{v\} \vee x \in$ values $\left(\right.$ filter $\left._{p}(\mathrm{~L})\right)$
Def values
$\equiv x \in\{v\} \vee(p(x) \wedge x \in \operatorname{values}(L))$
Def U
IH

## Claim: $v \in$ values(filter $\left.{ }_{p}(\mathrm{~L})\right)$ iff $p(x) \wedge x \in \operatorname{values}(\mathrm{~L})$

$Q(L):=$ " $x \in$ values(filter $\left._{p}(L)\right)$ iff $p(x) \wedge x \in$ values(L) for all $x \in \mathbb{Z}$ ".
We will prove $Q(L)$ for $L \in$ Lists by structural induction.
Base Case: ... so $Q$ (null) holds.
Inductive Hypothesis: Suppose $Q(L)$ holds for an arbitrary list L,
i.e., we have $x \in$ values $\left(\right.$ filter $\left._{p}(L)\right)$ iff $p(x) \wedge x \in$ values $(L)$.

Inductive Step: Goal: Prove $Q($ Node(v, L)) for all $v \in \mathbb{Z}$
Let $v, x \in \mathbb{Z}$ be arbitrary. We go by cases. Suppose $p(v)$.
$x \in \operatorname{values}\left(\right.$ filter $\left._{p}(\operatorname{Node}(v, L))\right)$
$\equiv x \in \operatorname{values}\left(\operatorname{Node}^{\left.\left(v, \text { filter }_{p}(L)\right)\right) \quad \text { Def filter }_{p}}\right.$
$\equiv x \in\{v\} \cup$ values(filter $\left.{ }_{p}(\mathrm{~L})\right)$
$\equiv x \in\{v\} \vee x \in$ values(filter $(\mathrm{L})) \quad$ Def $\cup$
$\equiv x \in\{v\} \vee(p(x) \wedge x \in \operatorname{values}(L)) \quad$ IH
$\equiv(x \in\{v\} \vee p(x)) \wedge(x \in\{v\} \vee x \in \operatorname{values}(L))$ Distributivity
$\equiv(x \in\{v\} \vee p(x)) \wedge(x \in \operatorname{values}(\operatorname{Node}(v, L))) \quad$ Def $U$, values

## Claim: $v \in$ values(filter $\left.{ }_{p}(\mathrm{~L})\right)$ iff $p(x) \wedge x \in$ values $(\mathrm{L})$

$Q(L):=$ " $x \in$ values(filter $(L))$ iff $p(x) \wedge x \in$ values $(L)$ for all $x \in \mathbb{Z}$ ".
We will prove $Q(L)$ for $L \in$ Lists by structural induction.
Base Case: ... so $Q$ (null) holds.
Inductive Hypothesis: Suppose $Q(L)$ holds for an arbitrary list $L$,
i.e., we have $x \in$ values $\left(\right.$ filter $\left._{p}(L)\right)$ iff $p(x) \wedge x \in$ values $(L)$.

Inductive Step: Goal: Prove $Q($ Node( $v, L)$ ) for all $v \in \mathbb{Z}$
Let $v, x \in \mathbb{Z}$ be arbitrary. We go by cases. Suppose $p(v)$.
$x \in \operatorname{values}\left(\right.$ filter $\left._{p}(\operatorname{Node}(v, L))\right)$
$\equiv$...
$\equiv(x \in\{v\} \vee p(x)) \wedge(x \in \operatorname{values}(\operatorname{Node}(v, L)))$
If $x \in\{v\}$ is false, then the first part is $F \vee p(x) \equiv p(x)$.
If true, then $x=v$, and first part and $p(x)$ are both true. Thus,

$$
\equiv p(x) \wedge(x \in \operatorname{values}(\operatorname{Node}(v, L)))
$$

## Claim: $v \in$ values(filter $\left.{ }_{p}(\mathrm{~L})\right)$ iff $p(x) \wedge x \in$ values( $(\mathrm{L})$

$Q(L):=$ " $x \in$ values $^{(\text {filter }}(\mathrm{L}(\mathrm{L})$ ) iff $p(x) \wedge x \in$ values $(L)$ for all $x \in \mathbb{Z}$ ".
We will prove $Q(L)$ for $L \in$ Lists by structural induction.
Base Case: ... so $Q$ (null) holds.
Inductive Hypothesis: Suppose $Q(L)$ holds for an arbitrary list $L$,
i.e., we have $x \in$ values(filter $(\mathrm{L})$ ) iff $p(x) \wedge x \in$ values $(\mathrm{L})$.

Inductive Step: Goal: Prove $Q($ Node( $v, L)$ ) for all $v \in \mathbb{Z}$
Let $v, x \in \mathbb{Z}$ be arbitrary. We go by cases. Suppose $p(v)$.
$x \in \operatorname{values}\left(\right.$ filter $\left._{p}(\operatorname{Node}(v, L))\right)$
$\equiv$...
$\equiv \mathrm{p}(\mathrm{x}) \wedge(\mathrm{x} \in \operatorname{values}(\operatorname{Node}(\mathrm{v}, \mathrm{L})))$
Thus, by cases, the claimed bicondition holds.
Since $x$ was arbitrary, we have shown $Q(\operatorname{Node}(v, L))$.
Hence, we have shown $Q(L)$ for all lists by structural induction.

## Theoretical Computer Science

## Languages: Sets of Strings

- Subsets of strings are called languages
- Examples:
$-\Sigma^{*}=$ All strings over alphabet $\Sigma$
- Palindromes over $\Sigma$
- Binary strings that don't have a 0 after a 1
- Binary strings with an equal \# of 0's and 1's
- Legal variable names in Java/C/C++
- Syntactically correct Java/C/C++ programs
- Valid English sentences


## Foreword on Intro to Theory C.S.

- Look at different ways of defining languages
- See which are more expressive than others
- i.e., which can define more languages
- Later: connect ways of defining languages to different types of (restricted) computers
- computers capable of recognizing those languages i.e., distinguishing strings in the language from not
- Consequence: computers that recognize more expressive languages are more powerful


## Regular Expressions

## Regular expressions over $\Sigma$

- Basis:
$\varepsilon$ is a regular expression
(could also include $\varnothing$ )
$a$ is a regular expression for any $a \in \Sigma$
- Recursive step:

If $A$ and $B$ are regular expressions then so are:
$A \cup B$
AB
A*

## Each Regular Expression is a "pattern"

$\varepsilon$ matches only the empty string
a matches only the one-character string $a$
$A \cup B$ matches all strings that either A matches or B matches (or both)
$A B$ matches all strings that have a first part that $A$ matches followed by a second part that B matches
A* matches all strings that have any number of strings (even 0) that A matches, one after another $(\varepsilon \cup \mathbf{A} \cup \mathbf{A A} \cup A A A \cup \ldots)$

## Language of a Regular Expression

The language defined by a regular expression:

$$
\begin{aligned}
& \mathrm{L}(\varepsilon)=\{\varepsilon\} \\
& \mathrm{L}(a)=\{a\} \\
& \mathrm{L}(A \cup B)=L(A) \cup L(B) \\
& \mathrm{L}(A B)=\{x \bullet y \mid x \in L(A), y \in L(B)\} \\
& \mathrm{L}\left(A^{*}\right)=\bigcup_{n=0}^{\infty} A^{n} \\
& \quad A^{n} \text { defined recursively by } \\
& \quad A^{0}=\emptyset \\
& \quad A^{n+1}=A^{n} A
\end{aligned}
$$

## Examples

## 001*

$0 * 1 *$

## Examples

## 001*

$\{00,001,0011,00111, \ldots\}$

0*1*

Any number of 0's followed by any number of 1's

## Examples

## $(0 \cup 1) 0(0 \cup 1) 0$

(0*1*)*

## Examples

$(0 \cup 1) 0(0 \cup 1) 0$
$\{0000,0010,1000,1010\}$
$(0 * 1 *) *$

All binary strings

## Examples

$(0 \cup 1) * 0110(0 \cup 1) *$
$(00 \cup 11) *(01010 \cup 10001)(0 \cup 1) *$

## Examples

$(0 \cup 1) * 0110(0 \cup 1) *$

Binary strings that contain "0110"
$(00 \cup 11) *(01010 \cup 10001)(0 \cup 1) *$
Binary strings that begin with pairs of characters followed by "01010" or "10001"

## Examples

- All binary strings that have an even \# of 1's


## Examples

- All binary strings that have an even \# of 1's

$$
\text { e.g., } 0^{*}\left(10^{*} 10^{*}\right)^{*}
$$

## Examples

- All binary strings that have an even \# of 1's

$$
\text { e.g., } 0^{*}\left(10^{*} 10^{*}\right)^{*}
$$

- All binary strings that don't contain 101


## Examples

- All binary strings that have an even \# of 1's

$$
\text { e.g., } 0^{*}\left(10 * 10^{*}\right)^{*}
$$

- All binary strings that don't contain 101

$$
\begin{aligned}
& \text { e.g., } 0^{*}\left(1 \cup 1000^{*}\right)^{*}\left(0^{*} \cup 10^{*}\right) \\
& \text { at least two 0s between 1s }
\end{aligned}
$$

## Regular Expressions in Practice

- Used to define the "tokens": e.g., legal variable names, keywords in programming languages and compilers
- Used in grep, a program that does pattern matching searches in UNIX/LINUX
- Pattern matching using regular expressions is an essential feature of PHP
- We can use regular expressions in programs to process strings!


## Regular Expressions in Java

- Pattern p = Pattern.compile("a*b");
- Matcher m = p.matcher("aaaaab");
- boolean b = m.matches();
[01] a 0 or a 1 ^ start of string $\$$ end of string
[0-9] any single digit $\backslash$. period <br>, comma \-minus
. any single character
ab a followed by b
(AB)
(a|b) a orb
$(A \cup B)$
a? zero or one of a
$(A \cup \varepsilon)$
a* zero or more of a
A*
a+ one or more of a AA*
- e.g. ^[\-+]?[0-9]*(\. <br>,)?[0-9]+\$

General form of decimal number e.g. 9.12 or $-9,8$ (Europe)

## Limitations of Regular Expressions

- Not all languages can be specified by regular expressions
- Even some easy things like
- Palindromes
- Strings with equal number of 0's and 1's
- But also more complicated structures in programming languages
- Matched parentheses
- Properly formed arithmetic expressions
- etc.


## Context-Free Grammars

- A Context-Free Grammar (CFG) is given by a finite set of substitution rules involving
- A finite set $V$ of variables that can be replaced
- Alphabet $\Sigma$ of terminal symbols that can't be replaced
- One variable, usually $\mathbf{S}$, is called the start symbol
- The substitution rules involving a variable $\mathbf{A}$, written as

$$
A \rightarrow w_{1}\left|w_{2}\right| \cdots \mid w_{k}
$$

where each $w_{i}$ is a string of variables and terminals

- that is $\mathrm{w}_{\mathrm{i}} \in(\mathbf{V} \cup \Sigma)^{*}$


## How CFGs generate strings

- Begin with start symbol S
- If there is some variable $\mathbf{A}$ in the current string you can replace it by one of the w's in the rules for $\mathbf{A}$
$-A \rightarrow w_{1}\left|w_{2}\right| \cdots \mid w_{k}$
- Write this as $\quad x A y \Rightarrow x w y$
- Repeat until no variables left
- The set of strings the CFG describes are all strings, containing no variables, that can be generated in this manner (after a finite number of steps)


## Example Context-Free Grammars

## Example: $\quad \mathbf{S} \rightarrow \mathbf{0 S} 0 \mid 1 \mathbf{S 1 | 0 | 1 | \varepsilon}$

## Example Context-Free Grammars

## Example: $\quad \mathbf{S} \rightarrow \mathbf{0 S O} \mathbf{1 S 1 | 0 | 1 | \varepsilon}$

The set of all binary palindromes

## Example Context-Free Grammars

## Example: $\mathbf{S} \rightarrow \mathbf{O S O} \mid$ 1S1| $0|1| \varepsilon$

The set of all binary palindromes

Example: $\quad \mathbf{S} \rightarrow \mathbf{O S}|\mathbf{S} 1| \varepsilon$

## Example Context-Free Grammars

## Example: $\quad \mathbf{S} \rightarrow \mathbf{0 S O} \mid 1 \mathbf{S 1 | 0 | 1 | \varepsilon}$

## The set of all binary palindromes

Example: $\quad \mathbf{S} \rightarrow \mathbf{O S}|\mathbf{S} 1| \varepsilon$

$$
0^{*} 1^{*}
$$

