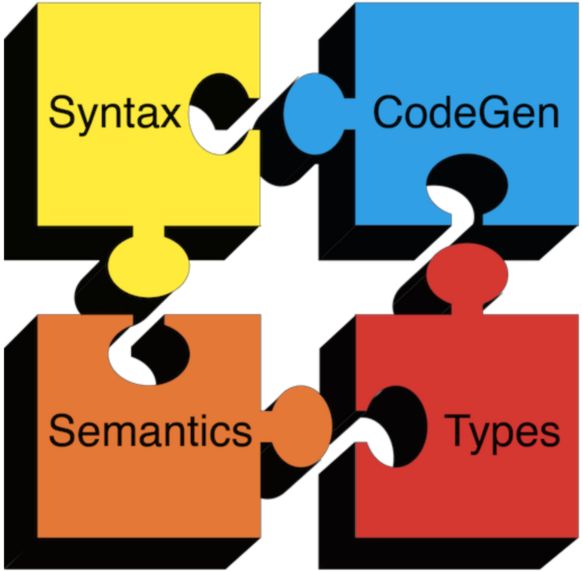

Compilers

Regular Languages

Regular Languages

- Lexical structure = token classes

- We must say what set of strings is in a token class
  - Use *regular languages*

Regular Languages

Regular Expressions

- Single character

$$\underline{'c'} \quad = \quad \underline{\{"c"\}}$$

- Epsilon

$$\begin{array}{ccc}\varepsilon & = & \underbrace{\{""\}} \\ \neq & & \\ \emptyset & & \end{array}$$

Regular Languages

- Union

$$A + B = \{a \mid a \in A\} \cup \{b \mid b \in B\}$$

- Concatenation

$$\underline{AB} = \{\underline{ab} \mid \underline{a \in A} \wedge \underline{b \in B}\}$$

- Iteration

$$\underline{A^*} = \bigcup_{i \geq 0} A^i \qquad A^i = \underbrace{A \ldots A}_{i \text{ times}}$$

$$A^0 = \epsilon$$

Regular Languages

- **Def.** The *regular expressions over* Σ are the smallest set of expressions including

$$\left[\begin{aligned} \underline{R} \;&=\; \epsilon \\ &\mid\; \text{'c'} \qquad\qquad c \in \underline{\Sigma} \\ &\mid\; \underline{R} + \underline{R} \\ &\mid\; R\,R \\ &\mid\; R^* \end{aligned}\right]$$

<u>grammar</u>

Regular Languages

$\Sigma = \{0,1\}$

$$1^* = \bigcup_{i\ge 0} 1^i = \overset{"" + 1 + 11 + 111 + 1111 + \cdots}{"" + \underbrace{1 \cdots\cdots 1}_{i} + \cdots} = \text{all strings of 1's}$$

$1^* = 1^* + 1$

$$(1+0)1 = \{ab \mid a \in 1+0 \wedge b \in 1\} = \{11, 10\}$$

$11 + 10$

$$0^* + 1^* = \{0^i \mid i \ge 0\} \cup \{1^i \mid i \ge 0\}$$

$$(0+1)^* = \bigcup_{i\ge 0} (0+1)^i =$$

$$"",\ 0+1,\ (0+1)(0+1),\ \dots\ \underbrace{(0+1)\cdots(0+1)}_{i \text{ times}} =$$

all strings of 0's and 1's

$= \Sigma^*$

Regular Languages

Choose the regular languages that are equivalent to the given regular language: (0 + 1)*1(0 + 1)*

$\Sigma = \{ 0, 1 \}$

- □ (01 + 11)*(0 + 1)*
- □ (0 + 1)*(10 + 11 + 1)(0 + 1)*
- □ (1 + 0)*1(1 + 0)*
- □ (0+ 1)*(0 + 1)(0 + 1)*

Regular Languages

- Regular expressions specify regular languages

syntax set of strings

- Five constructs
  - Two base cases
    - empty and 1-character strings

ϵ

  - Three compound expressions
    - union, concatenation, iteration