But the Lord came down to see the city and the tower the people were building. The Lord said, "If as one people speaking the same language they have begun to do this, then nothing they plan to do will be impossible for them. Come, let us go down and confuse their language so they will not understand each other."

— Genesis 11:6–7 (New International Version)

Soyez réglé dans votre vie et ordinaire comme un bourgeois,
afin d'être violent et original dans vos œuvres.
[Be regular and orderly in your life like a bourgeois,
so that you may be violent and original in your work.]

— Gustave Flaubert, in a letter to Gertrude Tennant (December 25, 1876)

Some people, when confronted with a problem, think "I know, I'll use regular expressions." Now they have two problems.

— Jamie Zawinski, alt.religion.emacs (August 12, 1997)

As far as I am aware this pronunciation is incorrect in all known languages.

— Kenneth Kleene, describing his father Stephen's pronunciation of his last name

Formal languange = set of strings

$\varnothing = \{\ \}$

Σ^*

$\{\varepsilon\}$

ε is not a language

$L = \{w \in \{0,1\}^* \mid w \text{ has even \# of 1s}\}$

$L = \{\text{BMO}\}$

$L = \{\text{FINN, JAKE, ICEKING}\}$

$L = A \cup B \qquad L = A \cap B \qquad L = A \setminus B = A \cap \overline{B}$

$L = \overline{A} = \Sigma^* \setminus A$

$L = A \bullet B = AB = \{w \mid w = xy \text{ for some } x \in A \text{ and } y \in B\}$

$\{\text{ABRA, HOCUS}\} \bullet \{\text{CADABRA, POCUS}\}$ has 4 strings

$\varnothing \bullet L = \varnothing \qquad \{\varepsilon\} \bullet L = L$

$L^* = \{\varepsilon\} \cup L \cup L \cdot L \cup L \cdot L \cdot L \cup \cdots$ Kleene closure

$w \in L^*$ iff either $w = \varepsilon$
or $w = x \cdot y$ for some $x \in L$, $y \in L^*$

$\{0,1\}^*$

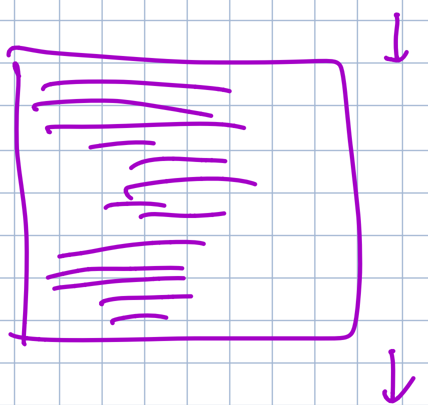
Is L^* always infinite? NO! $\emptyset^* = \{\varepsilon\} = \{\varepsilon\}^*$

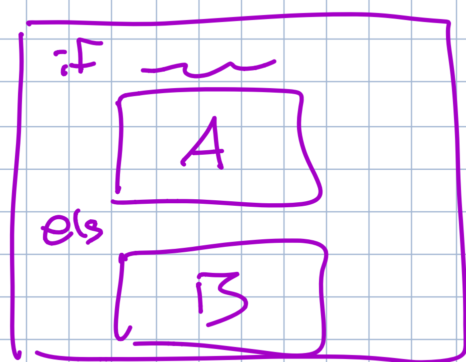
$(0+11)^*$

$(\{0\} \cup \{11\})^*$ — string of 0's and 1's
every run of 1s has even length.

Regular language is either

- $\emptyset$
- $\{w\}$ for some string w ← line of code
- $A \cup B$ where A, B regular ← conditional
- $A \bullet B$ " " ← sequencing
- A^* where A is regular ← loops





$A \cup B$

Not $\{0^n 1^n \mid n \geq 0\}$

Reg expression $\emptyset$, w, $A+B$, AB, A^*
with parens to disambiguate

$$10^* \neq (10)^* \neq (1+0)^* = (0^* 1^*)^*$$

$$(1+0)^* = \{0,1\}^* = (\{0\} \cup \{1\})^*$$

$$\underbrace{\varepsilon}\,\underbrace{0011}\,\underbrace{01}\,\underbrace{\varepsilon}$$

$00000^* = 0000(0)^*$

$(00000)^*$

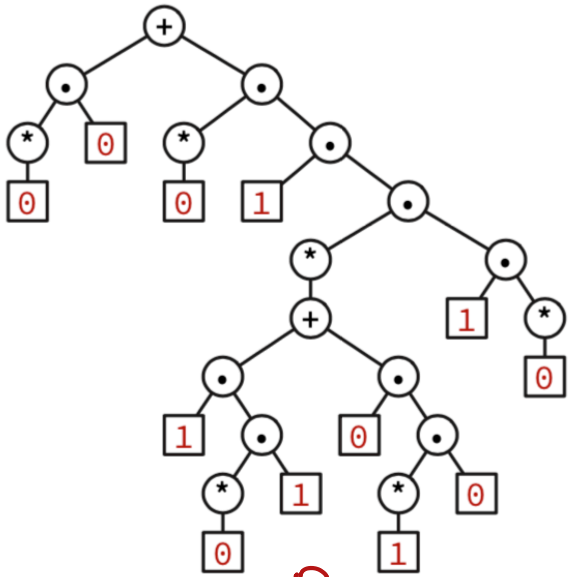
Alternating 0s and 1s

$$(01)^*0 + (01)^* + (10)^* + (10)^*1$$

$$= (01)^*(0+\varepsilon) + (10)^*(1+\varepsilon)$$

$$= (1+\varepsilon)(01)^*(0+\varepsilon)$$

01001



A regular expression tree for $00^* + 0^*1(10^*1 + 01^*0)^*10^*$

every regular expression smaller than R

Proof: Let R be an arbitrary regular expression.
Assume that **every proper subexpression of R** is perfectly cromulent.
There are five cases to consider.

- Suppose $R = \varnothing$.

  Prove: $\varnothing$ is p.c.

  Therefore, R is perfectly cromulent.
- Suppose R is a single string.

  w is p.c.

  Therefore, R is perfectly cromulent.
- Suppose $R = S + T$ for some regular expressions S and T.
  The induction hypothesis implies that S and T are perfectly cromulent.

  $S+T$ is p.c.

  Therefore, R is perfectly cromulent.
- Suppose $R = S \bullet T$ for some regular expressions S and T.
  The induction hypothesis implies that S and T are perfectly cromulent.

  $S \bullet T$ is p.c.

  Therefore, R is perfectly cromulent.
- Suppose $R = S^*$ for some regular expression. S.
  The induction hypothesis implies that S is perfectly cromulent.

  S^* is p.c.

  Therefore, R is perfectly cromulent.

In all cases, we conclude that R is perfectly cromulent. □

Theorem: Every regular expression without $\emptyset$ represents a non-empty language.

Proof: Let R be an arbitrary reg. expression without $\emptyset$

Assume every proper subexpression of R represents a non-empty language.

There are five cases

- $R = \emptyset$ $\not\leftarrow$ can't happen.
- $R = w$ $\longrightarrow L(R) = \{w\} \neq \emptyset$
- $R = A + B$ Both A, B have no $\emptyset$

  IH $\Rightarrow L(A) \neq \emptyset$ Pick $w \in L(A)$

  $w \in L(A+B)$ ✓
- $R = A \cdot B$ IH $\Rightarrow L(A) \neq \emptyset$ $L(B) \neq \emptyset$

  Pick $w \in L(A)$ $x \in L(B)$

  $wx \in L(A \cdot B)$
- $R = A^*$ $\varepsilon \in L(A^*)$ ✓

Therefore $L(R)$ is not empty.