Formal Languages

languages = sets of strings

ex $\{$ string, array $\}$

$\{0,1\}^*$

The set of all Python programs that print "Hello, world!"

	language?	length/size?
$\emptyset$	yes	$\lvert\emptyset\rvert = 0$
ε	no	$\lvert\varepsilon\rvert = 0$
$\{\varepsilon\}$	yes	$\lvert\{\varepsilon\}\rvert = 1$

[ ] empty list

"" empty string (in Python)

$L = \{ w \in \{0,1\}^* : w$ has an even number of 1's $\}$

$00 \in L$
$0101 \in L$

$\varepsilon \in L$

$1 \notin L$

Familiar set operations:

$L = A \cup B$

$L = A \cap B$

$L = A \setminus B$

Some new operators:

Concatenation:

$L = A \circ B = \{ x \circ y : x \in A \text{ and } y \in B \}$

$A = \{ \text{black}, \text{blue} \}$
$B = \{ \text{berry}, \text{fish}, \text{top} \}$

$A \circ B = \{$ blackberry, blackfish, blueberry, bluefish, blacktop, bluetop $\}$

$|A \circ B| = |A| \cdot |B|$

what is $\emptyset \circ L$ $\forall$ languages L? $\emptyset$

Check your answer: what is

$|\emptyset \circ L| \overset{?}{=} 0 \cdot |L| = 0$

what should we concatenate with L to get L?

$X \cdot L = L$ what is X?

$X = \{\varepsilon\}$

Kleene Star

$L = A^* =$ concatenation of any strings in A

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

$A = \{ba, bar\}$ how big is A^*?

$barbaba \in A^*$

why? $\varepsilon \in A^*$

$$ba = \underset{\text{in } A}{ba} \cdot \underset{\text{in } A^*}{\varepsilon} \in A^*$$

$$baba = \underset{\text{in } A}{ba} \cdot \underset{\text{in } A^*}{ba} \in A^*$$

$$barbaba = bar \cdot baba \in A^*$$

Let L be a language. Is L^* always infinite?

if $L = \{\varepsilon\}$, then $L^* = \{\varepsilon\}$

consider $L = \{a\}$ what is L^*?

$\{\varepsilon, a, aa, aaa, aaaa, \ldots\}$

if $L = \emptyset$, then $L^* = \{\varepsilon\}$

Regular Languages

Language L is regular $\Longleftrightarrow$

$$\begin{cases} L = \emptyset & \\ L = \{w\} & \text{for some } w \in \Sigma^* \\ L = A \cup B & \text{for reg. languages } A, B \quad \text{(branching)} \\ L = A \circ B & \text{for reg. lang. } A, B \quad \text{(sequencing)} \\ L = A^* & \text{for reg. lang. } A \quad \text{(repetition)} \end{cases}$$

while

w

if

$A \cup B$

else

$A \circ B$

Are all languages regular?

Some new notation for regular expression:

$L = \emptyset$	$\emptyset$
$L = \{w\}$	w
$L = A \cup B$	$A + B$
$L = A \circ B$	AB

note: $*$ has highest precedence

$L = A^*$ $\quad$ A^*

ex $\quad 0 + 10^* = \{0\} \cup (\{1\} \cdot \{0\}^*)$

$= \{0, 1, 100, 10, 1000, \ldots\}$

1 followed by any # of 0's and 0

L = alternating 0's & 1's = never have 00 or 11

$\boxed{\text{what is a regular expression for } L\text{?}}$

strings in L?

1010 1, ε, 101, 101010, 1, 0, ...

strings not in L?

11, 00, 01010101010100, ...

look for patterns that are repeated in long strings

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

0	1
010	101
01	1010
01010	
0101	
⋮	

Describe in English:

00000^*

$(00000)^*$ $\quad$ 0000000000 ϵ

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

operator precedence:

$*$

$\cdot$ (concatenation)

$+$

Give a regular expression for:

The set of all binary strings that contain the substring 0000.

$$(1+0)^*(0000)(1+0)^*$$

000000

Proofs about regular languages

Theorem: every regular expression is perfectly cromulent.

Proof: Let R be an arbitrary regular expression.
Assume that **every regular expression smaller than R** is perfectly cromulent. ← IH
There are five cases to consider.

- Suppose $R = \varnothing$.

  prove R is P.C

  Therefore, R is perfectly cromulent.

- Suppose R is a single string.

  prove that R 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.

  prove that R 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.

  prove that R 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.

  prove that R is P.C.

  Therefore, R is perfectly cromulent.

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