RegEx algebra

Igor Engel

1. Algebra exploration

1.1. Primitive regex

$$\bot \text{ - unmatchable}$$

$$\varepsilon \text{ - empty strirng}$$

$$\sigma_x \text{ - symbol } x$$

$$\sigma_. \text{ - wildcard}$$

$$\top \text{ - universal match}$$

1.2. Basic combinators

1.2.1. Concatenation

$$a \diamond b \text{ - concatenation}$$

$$a \diamond (b \diamond c) = (a \diamond b) \diamond c$$

$$\bot \diamond b = a \diamond \bot = \bot$$

$$\top \diamond \top = \top$$

$$\varepsilon \diamond b = b$$

$$a \diamond \varepsilon = a$$

1.2.2. Alternation

$$a \mid b \text{ - alternation}$$

$$a \mid (b \mid c) = (a \mid b) \mid c$$

$$a \mid b = b \mid a$$

$$a \mid a = a$$

$$\{a_1, \ldots, a_n\} = \{b_1, \ldots, b_m\} \implies a_1 \mid \ldots \mid a_n = b_1 \mid \ldots \mid b_m$$

$$a \diamond (b \mid c) = (a \diamond b) \mid (a \diamond c)$$

$$(a \mid b) \diamond c = (a \diamond c) \mid (b \diamond c)$$

$$\bot \,|\, b = b$$

$$a \,|\, \bot = a$$

$$\top \,|\, b = a \,|\, \top = \top$$

$$\top \diamond (\varepsilon \,|\, c) = (\varepsilon \,|\, b) \diamond \top = \top$$

1.2.3. Kleene star

$$a^* \text{ - Kleene star}$$

$$\varepsilon \,|\, a^* = a^* \,|\, \varepsilon = a^*$$

$$a^* = \varepsilon \,|\, (a \diamond a^*)$$

$$a^* \diamond a^* = a^*$$

$$a \,|\, a^* = a^*$$

$$(a^*)^* = a^*$$

$$\bot^* = \varepsilon$$

$$\varepsilon^* = \varepsilon$$

$$\sigma_.^* = \top$$

$$\top^* = \top$$

1.3. Advanced combinators

1.3.1. n-to-m

$$a_{n,k} \text{ - from } n \text{ to } n+k \text{ copies of } a,\ n \in \mathbb{N}, k \in \mathbb{N} \cup \{\infty\}$$

$$a_{n,} = a_{n,\infty} = a_{n,0} \diamond a^*$$

$$a_{,k} = a_{0,k}$$

$$a_{1,0} = a$$

$$a_{0,0} = \varepsilon$$

$$\varepsilon_{n,k} = \varepsilon$$

$$\bot_{0,k} = \varepsilon$$

$$n > 0 \implies \bot_{n,k} = \bot$$

$$n + k > 0 \implies \top_{n,k} = \top$$

$$(a_{0,k})_{n',k'} = a_{0,k(n'+k')}$$

$$(a_{1,k})_{n',k'} = a_{n',(1+k)(n'+k')}$$

$$(a_{0,k})^* = (a^*)_{0,k} = (a_{1,k})^* = (a^*)_{1,k} = a_{0,} = a^*$$

$$(\sigma.)_{0,} = \top$$

$$(\sigma_*)_{n,} = (\sigma_*)_{n,0} \diamond \top$$

$$a^+ = a_{1,} = a \diamond a^*$$

$$a^? = a_{0,1} = \varepsilon \,|\, a$$

$$\left(a^?\right)_{n,k} = a_{0,n+k}$$

$$(a^*)^+ = \left(a^+\right)^* = (a^*)^? = \left(a^?\right)^* = a^*$$

$$\left(a^+\right)^? = \left(a^?\right)^+ = a^*$$

$$a_{0,k+1} = \varepsilon \,|\, a_{1,k}$$

$$a_{n,k} \diamond a_{n',k'} = a_{n+n',k+k'}$$

$$a_{n,k+1} = a_{n,k} \diamond \left(a^?\right)$$

$$a_{n+1,k} = a \diamond a_{n,k} = a_{n,k} \diamond a$$

1.3.2. Character group

$$[x_1, \ldots, x_n] \text{ - any of } x_1, \ldots x_n$$

$$[] = \varepsilon$$

$$[x_1, \ldots, x_n] = \sigma_{x_1} \,|\, \ldots \,|\, \sigma_{x_n}$$

$$[x_1, \ldots, x_n] \,|\, \sigma_y = \sigma_y \,|\, [x_1, \ldots, x_n] = [y, x_1, \ldots, x_n]$$

$$[y, y, x_1, \ldots, x_n] = [y, x_1, \ldots, x_n]$$

$$s \in S_n \implies [x_1, \ldots, x_n] = [x_{s_1}, \ldots, x_{s_n}]$$

$$\{x_1, \ldots, x_n\} = \{y_1, \ldots, y_m\} \iff [x_1, \ldots, x_n] = [y_1, \ldots, y_m]$$

1.3.3. String

$$\overline{\sigma_s} \text{ - string } s$$

$$\overline{\sigma_\varepsilon} = \varepsilon$$

$$\sigma_x = \overline{\sigma_x}$$

$$\overline{\sigma_s} = \sigma_{s_1} \diamond \ldots \diamond \sigma_{s_n}$$

$$\overline{\sigma_s} \diamond \sigma_x = \sigma_{sx}$$

$$\sigma_x \diamond \overline{\sigma_s} = \sigma_{xs}$$

$$\overline{\sigma_s}_{n,0} = \overline{\sigma_{s^n}}$$

$$\overline{\sigma_s}_{n,k} = \overline{\sigma_{s^n}} \diamond \overline{\sigma_s}_{0,k}$$