

1 正規表現

記号列 (文字列・語): アルファベット V (記号の有限集合) から取り出した記号からなる「列」。

空列: 長さ 0 の記号列。 ϕ や λ , ε と表記する。

V から作られる記号列をすべて含む集合を V^* (ヴィ・スター) と書く。

$$V = \{a, b\} \text{ のとき}$$
$$V^* = \{\phi, a, b, aa, ab, ba, bb, aaa, \dots\}$$

注意

- V^* には空列が含まれる。
- V^* の要素は中かっこで囲わない (集合ではなく文字列)

1.1 正規表現 (regular expression)

記号列の集合 (V^*) を表記する方法の 1 つ。 $[A-Z][0-9]^+[a-z]^*[1-5|9]^?$ この正規表現には "E1333akashi9" などが該当する。

正規表現とオートマトン

- 正規表現があれば、それに対応する有限オートマトンがある。
- 有限オートマトンがあれば、それに対応する正規表現がある。

2 形式言語

言語には「自然言語 (日本語など)」と「形式言語 (C 言語など)」がある。

生成文法: 形式言語を生成する規則

オートマトン: 形式言語を認識する状態機械

BNF (バックス・ナウア記法)