



正则语言的性
质

姚刚

目录

正则语言的封
闭性质

正则语言的基
本问题

识别非正则语
言

第四章 正则语言的性质

姚 刚

中国科学院信息工程研究所



定理

DFA 正则表达式 正则文法

定理

如果 L_1 和 L_2 是正则语言，那么 $L_1 \cup L_2$ ， $L_1 \cap L_2$ ， $L_1 L_2$ ， $\overline{L_1}$ 和 L_1^* 也都是正则的。于是，我们认为正则语言在并、交、连接、补和星闭包上都是封闭的。

正则语言族在差运算上是封闭的。

定理

正则语言族在逆运算上是封闭的。

正则语言的性质

姚刚

目录

正则语言的封闭性质

正则语言的基本问题

识别非正则语言



例子

Easy: 同态映像直接代换即可

正则语言的性质
姚刚

目录

正则语言的封闭性质

正则语言的基本问题

识别非正则语言

已知 $\Sigma = \{a, b\}$ 和 $\Gamma = \{a, b, c\}$, 定义 h 为

$$h(a) = ab, \quad h(b) = bbc,$$

试给出语言 $L = \{aa, aba\}$ 的同态映像。

已知 $\Sigma = \{a, b\}$ 和 $\Gamma = \{b, c, d\}$, 定义 h 为

$$h(a) = dbcc, \quad h(b) = bbc,$$

正则表达式
表达的
语言
的同态像

试给出语言 L 的同态映像, 其中 L 用正则表达式 $r = (a + b^*)(aa)^*$ 表示。



定理

正则语言的性质

姚刚

目录

正则语言的封闭性质

正则语言的基本问题

识别非正则语言

定理

设 h 是同态函数。如果 L 是正则的，那么它的同态映像 $h(L)$ 也是正则的。因此，正则语言族在任意同态运算上都是封闭的。



右商(right quotient)

正则语言的性质

姚刚

目录

正则语言的封闭性质

正则语言的基本问题

识别非正则语言

定义(右商)

设 L_1 和 L_2 是定义在同一个字母表上的语言, 那么 L_1 和 L_2 的右商定义为

$$L_1/L_2 = \{x : \exists y \in L_2, \text{ 使得 } xy \in L_1\}.$$

如果 L_1 去掉属于 L_2 的后缀还是 L_1
非确定型 \rightarrow 确定型

$$L_1 = L(M_1)$$

$$L_2 = L(M_2)$$

$$L(M) = L_1/L_2$$

$$L_1 = \{a^n b^m : n \geq 1, m \geq 0\} \cup \{ba\},$$

$$L_2 = \{b^m : m \geq 1\},$$

试给出 L_1/L_2 。

路径上和 L_1 一致
转移函数一致
哪些状态是终止状态



定理

正则语言的性质

姚刚

目录

正则语言的封闭性质

正则语言的基本问题

识别非正则语言

定理

如果 L_1 和 L_2 是正则语言，那么 L_1/L_2 也是正则的。于是我们称正则语言族与正则语言在右商运算上是封闭的。

如果

$$L_1 = L(a^*baa^*)$$

$$L_2 = L(ab^*)$$

试给出 L_1/L_2 。

$$L(a^*ba^*)$$

正则表达式

先找有穷接收器转化

为确定型

构造接受商的有接器

四个状态

哪个是终止状态

从某状态出发Mq012找

终止状态



标准表示法

$(aa)^*$ 偶数个a

在我们考虑问题的时候，很多命题说“给定语言……”，我们需要明确其包含的意思，这对于解决问题很重要。我们描述正则语言的方式有：非形式化口头描述、集合表示、有穷接受器、正则表达式和正则文法。

这里，只有后三项是充分定义的，并且在各个定理中使用。我们说一个正则语言是以标准表示法(standard representation)给定的，当且仅当它是由有穷接受器、正则表达式和正则文法表示的。

正则语言的性质

姚刚

目录

正则语言的封闭性质

正则语言的基本问题

识别非正则语言



定理

正则语言的性质

姚刚

目录

正则语言的封闭性质

正则语言的基本问题

识别非正则语言

定理

给定一种标准表示法，表示任何定义在 Σ 上的正则语言 L 和任何 $w \in \Sigma^*$ 。这种标准表示法都有判定 w 是否属于 L 的算法。

定理

对于使用标准表示法给出的任何一个正则语言，都存在算法可以判断这个语言是空的，有穷的，还是无穷的。

存在回路



定理

正则语言的性质

姚刚

目录

正则语言的封闭性质

正则语言的基本问题

识别非正则语言

定理

对于使用标准表示法给定的任意两个正则语言 L_1 和 L_2 ，总存在判断 $L_1 = L_2$ 是否成立的算法。

有穷接收器
正则文法



鸽巢原理

任何一个有穷接受器都不接受它

反证法：假设存在有穷接受器，然后 $L(M) \neq L$

假设确定型，DFA $M = (Q, \sigma, \delta, q_0, F)$ $Q = \{q_0, \dots, q_m\}$

定理 (鸽巢原理)

如果我们把 n 个物品(鸽子)放到 m 个盒子(鸽巢)中，并且 $n > m$ ，那么至少有一个盒子(鸽巢)里面装有一个以上的物品(鸽子)。

语言 $L = \{a^n b^n : n \geq 0\}$ 是正则语言吗？

不是正则语言，因为要保证 ab 的数量一致 只能读一遍，在相同的时候才接受 但是 a 的数量没办法记录，状态头记录的话只能记录有限多个

正则语言的性质

姚刚

目录

正则语言的封闭性质

正则语言的基本问题

识别非正则语言



例子

正则语言的性质

姚刚

目录

正则语言的封闭性质

正则语言的基本问题

识别非正则语言

设 $\Sigma = \{a, b\}$, 证明语言 $L = \{ww^R : w \in \Sigma^*\}$ 不是正则语言。

设 $\Sigma = \{a, b\}$, 证明语言 $L = \{w \in \Sigma^* : n_a(w) < n_b(w)\}$ 不是正则语言。

证明语言 $L = \{(ab)^n a^k : n > k, k \geq 0\}$ 不是正则语言。



例子

正则语言的性质

姚刚

目录

正则语言的封闭性质

正则语言的基本问题

识别非正则语言

证明语言 $L = \{a^{(n!)} : n \geq 0\}$ 不是正则语言。

证明语言 $L = \{a^n b^k c^{n+k} : n \geq 0, k \geq 0\}$ 不是正则语言。

证明语言 $L = \{a^n b^l : n \neq l\}$ 不是正则语言。