Σ^*和Σ^+

如果w是符号串，那么w^n表示重复w这个符号串n次。我们定义：对于所有的w，都有$w^0 = \varepsilon$。

如果Σ是字母表，那么Σ^*表示连接Σ中的零个或者多个符号获得的所有符号串的集合。集合Σ^*总是包含ε，为了把空串排除在外，定义$\Sigma^+ = \Sigma^* - \{\varepsilon\}$。

计算理论导引
姚刚
目录
数学预备知识和表示
三个基本概念
课程内容简介

幂集和笛卡儿积

一个给定的集合通常有很多子集。集合S的所有子集组成的集合称为集合S的幂集(powerset)，记为2^S。

如果一个集合的元素是其他集合元素的有序排列，这个集合称为其他集合的笛卡儿积(Cartesian product)。一般有

$$S_1 \times S_2 \times \cdots \times S_n = \{(x_1, x_2, \cdots, x_n) : x_i \in S_i\}。$$

计算理论导引
姚刚
目录
数学预备知识和表示
三个基本概念
课程内容简介

语言(language)

一种语言通常定义为Σ^*上的子集。

一种语言L中的一个符号串称为这个语言L的一个句子(sentence)。字母表Σ上的符号串的任意集合都可以看成是一种语言。

计算理论导引
姚刚
目录
数学预备知识和表示
三个基本概念
课程内容简介

连接和逆

如果$w = a_1a_2\cdots a_n$，$v = b_1b_2\cdots b_m$。符号w和v的连接(concatenation)是把v添加到w的右端。我们用wv表示w和v的连接，有$wv = a_1a_2\cdots a_nb_1b_2\cdots b_m$。

符号串的逆(reverse)是把符号串中的符号按照相反的顺序列出。符号串w的逆用w^R来表示，有$w^R = a_n\cdots a_2a_1$。

闭包

一种语言L的星闭包(star-closure)定义为$L^* = L^0 \cup L^1 \cup L^2 \cup \cdots$。

一种语言L的正闭包(positive closure)定义为$L^+ = L^1 \cup L^2 \cup \cdots$。

语言的运算

计算理论导引
姚刚
目录
数学预备知识和表示
三个基本概念
课程内容简介

语言是集合，可以定义两种语言的并、交和差。

在全集Σ^*上定义语言L的补集为$\overline{L} = \Sigma^* - L$。

一种语言的逆指的是所有的符号串的逆所构成的集合，即$L^R = \{w^R : w \in L\}$。

例子

计算理论导引
姚刚
目录
数学预备知识和表示
三个基本概念
课程内容简介

设$\Sigma = \{a, b\}$，则$\Sigma^* = \{\varepsilon, a, b, aa, ab, ba, bb, \cdots\}$。

$L = \{a^n b^n : n \geq 0\}$是一个语言，
$L^2 = \{a^n b^n a^m b^m : n \geq 0, m \geq 0\}$，
$L^R = \{b^n a^n : n \geq 0\}$。

语言的连接

计算理论导引
姚刚
目录
数学预备知识和表示
三个基本概念
课程内容简介

两种语言L_1和L_2的连接指的是L_1中的任意元素和L_2中的任意元素通过连接形成的所有符号串的集合。具体表示为$L_1L_2 = \{xy : x \in L_1, y \in L_2\}$。

我们定义L^n为L自身连接n次，特殊地，对任意语言L，都有$L^0 = \{\varepsilon\}$，$L^1 = L$。

例子

给出生成语言$L(G) = \{a^n b^{n+1} : n \geq 0\}$的一个文法。

设$\Sigma = \{a, b\}$，$n_a(w)$和$n_b(w)$分别表示符号串w中a和b的个数。

文法G的产生式定义为$S \to SS$，$S \to \varepsilon$，$S \to aSb$和$S \to bSa$，则它生成的语言是$L = \{w : n_a(w) = n_b(w)\}$。

计算理论导引
姚刚
目录
数学预备知识和表示
三个基本概念
课程内容简介

文法生成的语言

定义

如果$w \in L(G)$，那么序列$S \Rightarrow w_1 \Rightarrow w_2 \Rightarrow \cdots \Rightarrow w_n \Rightarrow w$ 就是句子w的一个推导(*derivation*)，包含变量和终结符的符号串$S, w_1, w_2, \cdots, w_n$ 称为推导的句型(*sentence form*)。

设$G = (\{S\}, \{a, b\}, S, P)$，其中$P$定义为$S \to \varepsilon$和$S \to aSb$，则$L(G) = \{a^n b^n : n \geq 0\}$。

计算理论导引
姚刚
目录
数学预备知识和表示
三个基本概念
课程内容简介

正则语言性质

- 正则语言对并、交、补、连接等运算封闭。
- 可以确定有限自动机接受的语言是空的、有限的、无限的。
- 可以判定一个元素是否属于某个正则语言。
- 存在算法，能够判定两个正则语言是否等价。

计算理论导引
姚刚
目录
数学预备知识和表示
三个基本概念
课程内容简介

语言由文法生成

为了证明某个语言L是由文法G生成的，我们必须证明：

1. 每个$w \in L$都可以使用G中的产生式由S推导出；
2. 每个这样推导出的符号串都是语言L中的句子。

上下文无关语言

上下文无关文法化简：消除无用产生式、消除空产生式、消除单位产生式。

范式：乔姆斯基范式和格里巴克范式。

性质：可以判定一个元素是否属于上下文无关语言；可以确定上下文无关语言是否是空的、有限的、无限的；上下文无关语言对并和连接运算封闭。

下推自动机

下推自动机可以用一个七元组$(Q, \Sigma, \Gamma, \delta, q_0, z_0, F)$表示，其中$\Gamma$是一个栈字母表。

如果转移函数是确定的，则称为确定型下推自动机。确定型下推自动机和非确定型下推自动机不等价。

图灵机

标准图灵机的定义。图灵机的机制是非常简单的，但是足以解决非常复杂的过程。

被图灵机接受的语言称为递归可枚举语言。

图灵机的变种：带有不动选择的图灵机、多道图灵机；单向无穷带的图灵机；多带图灵机、多维图灵机；非确定型图灵机等。

上下文无关文法

设G是一个文法，如果它的产生式都有形式$A \to B$，其中A是单变量，B是$(V \cup T)^*$中的元素，则称为上下文无关文法。

一个语言是上下文无关语言当且仅当存在一个下推自动机识别它。

线性有界自动机

线性有界自动机包含两个特殊的符号：[和]，分别是左端标志和右端标志，并且在[处没有向左移动的动作，在]处没有向右移动的动作。

线性有界自动机接受的语言称为上下文相关语言。

计算理论导引
姚刚
目录
数学预备知识和表示
三个基本概念
课程内容简介

乔姆斯基层次

图灵机⇔递归可枚举语言

线性有界自动机⇔上下文相关语言

下推自动机⇔上下文无关语言

有限自动机⇔正则语言

计算理论导引
姚刚
目录
数学预备知识和表示
三个基本概念
课程内容简介

其他问题

- 图灵机不能解决的问题
- 递归可枚举语言和上下文无关语言的不可判定问题
- 计算模型
- 计算复杂性