形式语言与自动机

李舟军 教授

北京航空航天大学计算机学院

lizj@buaa.edu.cn

访问主页
标题页
第 1 页 共 47 页
返 回
全屏显示
关 闭
退 出

第一章　预备知识

1 预备知识

1.1. 字母表、字和语言

- 自然语言与人工语言

1 预备知识

1.1. 字母表、字和语言

- 自然语言与人工语言
  - 通常使用的语言，如汉语、英语等，多是在人们日常的社会实践中自然地形成的，所以称为自然语言.

1 预备知识

1.1. 字母表、字和语言

- 自然语言与人工语言
  - 通常使用的语言，如汉语、英语等，多是在人们日常的社会实践中自然地形成的，所以称为自然语言.
  - 数理逻辑中的命题演算和谓词演算，电子计算机上的所使用的种种程序设计语言,如Java，C等，称为人工语言.

1 预备知识

1.1. 字母表、字和语言

- 自然语言与人工语言
  - 通常使用的语言，如汉语、英语等，多是在人们日常的社会实践中自然地形成的，所以称为自然语言.
  - 数理逻辑中的命题演算和谓词演算，电子计算机上的所使用的种种程序设计语言,如Java，C等，称为人工语言.
  - 自然语言与人工语言的主要区别：语法和语义.

1 预备知识

1.1. 字母表、字和语言

- 自然语言与人工语言
  - 通常使用的语言，如汉语、英语等，多是在人们日常的社会实践中自然地形成的，所以称为自然语言.
  - 数理逻辑中的命题演算和谓词演算，电子计算机上的所使用的种种程序设计语言,如Java，C等，称为人工语言.
  - 自然语言与人工语言的主要区别：语法和语义.
  - 形式语言的研究目标是要把人工语言定义为一个严格的数学系统，其严格的形式性使得我们能够给出形式语言构成的数学描述，并能进而揭示所描述的语言的结构，特性及其应用。

- 和自然语言一样，形式语言的构成基础是所谓的符号，符号是一个没有严格定义的原始概念，只能给它以直观描述或说明。

- 和自然语言一样，形式语言的构成基础是所谓的符号，符号是一个没有严格定义的原始概念，只能给它以直观描述或说明。
- 当我们把一个记号作为不可分割的整体看待时，就称这个记号为一个符号。

- 和自然语言一样，形式语言的构成基础是所谓的符号，符号是一个没有严格定义的原始概念，只能给它以直观描述或说明。
- 当我们把一个记号作为不可分割的整体看待时，就称这个记号为一个符号。
- 常用的符号：

- 和自然语言一样，形式语言的构成基础是所谓的符号，符号是一个没有严格定义的原始概念，只能给它以直观描述或说明。
- 当我们把一个记号作为不可分割的整体看待时，就称这个记号为一个符号。
- 常用的符号：
  - 数字：0，1，2，...,9

- 和自然语言一样，形式语言的构成基础是所谓的符号，符号是一个没有严格定义的原始概念，只能给它以直观描述或说明。
- 当我们把一个记号作为不可分割的整体看待时，就称这个记号为一个符号。
- 常用的符号：
  - 数字：0，1，2，...,9
  - 小写拉丁字母：a,b,c,...,z

- 和自然语言一样，形式语言的构成基础是所谓的符号，符号是一个没有严格定义的原始概念，只能给它以直观描述或说明。
- 当我们把一个记号作为不可分割的整体看待时，就称这个记号为一个符号。
- 常用的符号：
  - 数字：0，1，2，...,9
  - 小写拉丁字母：a,b,c,...,z
  - 大写拉丁字母：A,B,C,...,Z

- 和自然语言一样，形式语言的构成基础是所谓的符号，符号是一个没有严格定义的原始概念，只能给它以直观描述或说明。
- 当我们把一个记号作为不可分割的整体看待时，就称这个记号为一个符号。
- 常用的符号：
  - 数字：0，1，2，...,9
  - 小写拉丁字母：a,b,c,...,z
  - 大写拉丁字母：A,B,C,...,Z
  - 括号：’,’,(,),[,],{,}

- 和自然语言一样，形式语言的构成基础是所谓的符号，符号是一个没有严格定义的原始概念，只能给它以直观描述或说明。
- 当我们把一个记号作为不可分割的整体看待时，就称这个记号为一个符号。
- 常用的符号：
  - 数字：0，1，2，...,9
  - 小写拉丁字母：a,b,c,...,z
  - 大写拉丁字母：A,B,C,...,Z
  - 括号：',',(,),[,],{,}
  - 运算符：+，-，×,*,/,÷

- 和自然语言一样，形式语言的构成基础是所谓的符号，符号是一个没有严格定义的原始概念，只能给它以直观描述或说明。
- 当我们把一个记号作为不可分割的整体看待时，就称这个记号为一个符号。
- 常用的符号：
  - 数字：0，1，2，...,9
  - 小写拉丁字母：a,b,c,...,z
  - 大写拉丁字母：A,B,C,...,Z
  - 括号：’,’,(,),[,],{,}
  - 运算符：+，-，$\times$,*,/,$\div$
  - 关系符：=，$\neq, >, \geq, \leq, <, \in, \notin$，$\subset$，$\subseteq$

- 和自然语言一样，形式语言的构成基础是所谓的符号，符号是一个没有严格定义的原始概念，只能给它以直观描述或说明。
- 当我们把一个记号作为不可分割的整体看待时，就称这个记号为一个符号。
- 常用的符号：
  - 数字：0，1，2，...,9
  - 小写拉丁字母：a,b,c,...,z
  - 大写拉丁字母：A,B,C,...,Z
  - 括号：',',(,),[,],{,}
  - 运算符：+，-，×,*,/,÷
  - 关系符：=，≠,>,≥,≤,<,∈,∉，⊂，⊆
  - 标点附号：.，，，;，:

- 和自然语言一样，形式语言的构成基础是所谓的符号，符号是一个没有严格定义的原始概念，只能给它以直观描述或说明。
- 当我们把一个记号作为不可分割的整体看待时，就称这个记号为一个符号。
- 常用的符号：
  - 数字：0，1，2，...,9
  - 小写拉丁字母：a,b,c,...,z
  - 大写拉丁字母：A,B,C,...,Z
  - 括号：’,’,(,),[,],{,}
  - 运算符：+，-，$\times$,*,/,$\div$
  - 关系符：=，$\neq,>,\geq,\leq,<,\in,\notin$，$\subset$，$\subseteq$
  - 标点附号：.，，，;，：
  - 汉字

- 和自然语言一样，形式语言的构成基础是所谓的符号，符号是一个没有严格定义的原始概念，只能给它以直观描述或说明。
- 当我们把一个记号作为不可分割的整体看待时，就称这个记号为一个符号。
- 常用的符号：
  - 数字：0，1，2，...,9
  - 小写拉丁字母：a,b,c,...,z
  - 大写拉丁字母：A,B,C,...,Z
  - 括号：’,’,(,),[,],{,}
  - 运算符：+，-，$\times$,*,/,$\div$
  - 关系符：=，$\neq$,>,$\geq$,$\leq$,<,$\in$,$\notin$，$\subset$，$\subseteq$
  - 标点附号：.，，，;，：
  - 汉字

访问主页
标题页
◀◀ ▶▶
◀ ▶
第 4 页 共 47 页
返回
全屏显示
关闭
退出

- 字母表、字、语言

定义1.1.1 字母表：符号的非空集合称为字母表，记为Σ;

- 字母表、字、语言

定义1.1.1 字母表：符号的非空集合称为字母表，记为Σ;

- 字母表Σ中的元素就是符号，常用英文字母中较前面的小写拉丁字母a,b,c,...等表示。

- 字母表、字、语言

定义1.1.1 字母表：符号的非空集合称为字母表，记为Σ;

- 字母表Σ中的元素就是符号，常用英文字母中较前面的小写拉丁字母a,b,c,...等表示。
- 数理逻辑中的命题演算和谓词演算，其字母表都是无穷集;

- 字母表、字、语言

  定义1.1.1 字母表： 符号的非空集合称为字母表，记为Σ;

  - 字母表Σ中的元素就是符号，常用英文字母中较前面的小写拉丁字母a,b,c,...等表示。
  - 数理逻辑中的命题演算和谓词演算，其字母表都是无穷集;
  - 程序设计语言的符号表都是有穷集，它们是较为常见的有穷字母表的例子。

- 字母表、字、语言

例1：下列都是常用的较为简单的字母表：

- 字母表、字、语言

例1： 下列都是常用的较为简单的字母表：

  - $\{a,b\}$
  - $\{0,1,\ldots,9\}$
  - $\{a,b,c,\ldots,z\}$

- 字母表、字、语言

  例1：下列都是常用的较为简单的字母表：

  - {a,b}
  - {0,1,...,9}
  - {a,b,c,...,z}

- **例2**：程序设计语言ALGOL60的字母表由以下十一类符号组成：

  – 拉丁字母 数字 逻辑值 算术运算符 逻辑运算符 关系运算符 顺序运算符 分隔符 括号 说明符 分类符

访问主页
标题页
第 6 页 共 47 页
返回
全屏显示
关闭
退出

- 字母表、字、语言

定义1.1.2：设Σ字母表：

- 字母表、字、语言

定义1.1.2：设Σ字母表：

i 对任意的$n \in \mathbb{N}$及任意的$a_1, a_2, ..., a_n \in \Sigma$,称$a_1, a_2, ..., a_n$并置在一起所构成的符号串$a_1...a_n$ 为Σ上的一个字。

- 字母表、字、语言

定义1.1.2：设Σ字母表：

i 对任意的$n \in \mathbb{N}$及任意的$a_1, a_2, ..., a_n \in \Sigma$,称$a_1, a_2, ..., a_n$并置在一起所构成的符号串$a_1...a_n$ 为Σ上的一个字。

ii 不含任何符号的空字符串(即当$n = 0$时的符号串$a_1...a_n$)，称为空字，并用ϵ表示。

- 字母表、字、语言

定义1.1.2：设Σ字母表：

i 对任意的$n \in \mathbb{N}$及任意的$a_1, a_2, ..., a_n \in \Sigma$,称$a_1, a_2, ..., a_n$并置在一起所构成的符号串$a_1...a_n$ 为Σ上的一个字。

ii 不含任何符号的空字符串(即当$n = 0$时的符号串$a_1...a_n$)，称为空字，并用ϵ表示。

iii 称集合$\{\alpha|\alpha$为Σ上的一个字$\}$为Σ上所有字的集合，用Σ^*表示，并令$\Sigma^+ = \Sigma^* \backslash \{\epsilon\}, \Sigma_\epsilon = \Sigma \cup \{\epsilon\}$。

- 字母表、字、语言

定义1.1.3: 设Σ字母表,对任意的$\alpha, \beta \in \Sigma^*$,不妨设$\alpha = a_1...a_n, \beta = b_1...b_m, a_1, a_2, ..., a_n, b_1, ..., b_m \in \Sigma$,若令$\alpha \cdot \beta = a_1...a_n b_1...b_m$,则称$\alpha \cdot \beta$为字$\alpha$与$\beta$的联结，简记为$\alpha\beta$.

- 字母表、字、语言

定义1.1.3： 设Σ字母表,对任意的$\alpha,\beta\in\Sigma^*$,不妨设$\alpha=a_1...a_n,\beta=b_1...b_m,a_1,a_2,...,a_n,b_1,...,b_m\in\Sigma$,若令$\alpha\cdot\beta=a_1...a_nb_1...b_m$,则称$\alpha\cdot\beta$为字$\alpha$与$\beta$的联结，简记为$\alpha\beta$.

定义1.1.3中引进的$\cdot$显然是集合Σ^*上的一种运算,并且显然运算$\cdot$不具有可交换性。

- 字母表、字、语言

定理1.1.1：设Σ字母表,如果$S \subseteq \Sigma^*$满足：

i $\epsilon \in S$;

- 字母表、字、语言

定理1.1.1：设Σ字母表,如果$S \subseteq \Sigma^*$满足：

i $\epsilon \in S$;

ii 若$\alpha \in S$且$a \in \Sigma$,则$\alpha a \in S$;

- 字母表、字、语言

定理1.1.1：设Σ字母表,如果$S \subseteq \Sigma^*$满足：

i $\epsilon \in S$;

ii 若$\alpha \in S$且$a \in \Sigma$,则$\alpha a \in S$;

则有：$S = \Sigma^*$。

- 字母表、字、语言

定理1.1.2： 设Σ字母表,如果$S \subseteq \Sigma^*$满足：

i $\epsilon \in S$;

ii 若$\alpha \in S$且$a \in \Sigma$,则$a\alpha \in S$;

则有： $S = \Sigma^*$。

访问主页
标题页
第 10 页 共 47 页
返 回
全屏显示
关 闭
退 出

- 字母表、字、语言

定理1.1.3：设Σ为字母表,$a \in \Sigma$且$\alpha, \beta, \gamma \in \Sigma^*$,则有：

i $\alpha \cdot \beta \in \Sigma^*$;

ii $\alpha\epsilon = \alpha = \epsilon\alpha$;

iii $\alpha a \neq \epsilon$且$a\alpha \neq \epsilon$;

iv $(\alpha \cdot \beta) \cdot \gamma = \alpha \cdot (\beta \cdot \gamma)$;

v 若$\alpha\beta = \alpha\gamma$，则$\beta = \gamma$;

vi 若$\alpha\gamma = \beta\gamma$，则$\alpha = \beta$;

- 字母表、字、语言

定理1.1.4: 若Σ为字母表,则$< \Sigma^*, \cdot, \epsilon >$为Monoid(含幺半群).

- 字母表、字、语言

定理1.1.4: 若Σ为字母表,则$< \Sigma^*, \cdot, \epsilon >$为Monoid(含幺半群).

定理1.1.5: 设Σ字母表,若$\alpha \in \Sigma^*$,则:

i $\alpha^{n+1} = \alpha^n \cdot \alpha$;

ii $\alpha^n \cdot \alpha^m = \alpha^{n+m}$;

iii $(\alpha^n)^m = \alpha^{nm}$;

- 字母表、字、语言

定义1.1.4: 设Σ为字母表,$\alpha \in \Sigma^*, L_1 \subseteq \Sigma^*$且$L_2 \subseteq \Sigma^*$,令:

$\alpha \cdot L_1 = \{\alpha x | x \in L_1\};$

$L_1 \cdot \alpha = \{x\alpha | x \in L_1\};$

$L_1 \cdot L_2 = \{xy | x \in L_1$且$y \in L_2\};$

分别把$\alpha \cdot L_1, L_1 \cdot \alpha, L_1 \cdot L_2$简记为$\alpha L_1, L_1\alpha, L_1L_2$.

- 字母表、字、语言

定义1.1.4： 设Σ为字母表,$\alpha \in \Sigma^*, L_1 \subseteq \Sigma^*$且$L_2 \subseteq \Sigma^*$,令：

$\alpha \cdot L_1 = \{\alpha x | x \in L_1\}$;

$L_1 \cdot \alpha = \{x\alpha | x \in L_1\}$;

$L_1 \cdot L_2 = \{xy | x \in L_1$且$y \in L_2\}$;

分别把$\alpha \cdot L_1, L_1 \cdot \alpha, L_1 \cdot L_2$简记为$\alpha L_1, L_1\alpha, L_1 L_2$.

定理1.1.6： 设Σ为字母表,则:

i $< 2^{\Sigma^*}, \cdot, \{\epsilon\} >$为Monoid;

ii 若$L_1 \subseteq \Sigma^*$且$L_2 \subseteq \Sigma^*$,则：$L_1 \cdot L_2 = \cup_{x \in L_1} x L_2 = \cup_{y \in L_2} L_1 y$;

iii 若$\alpha, \beta \in \Sigma^*$且$L \subseteq \Sigma^*$,则：$\alpha(\beta L) = (\alpha\beta)L$;

iv 若$\alpha \in \Sigma^*, L_1 \subseteq \Sigma^*$且$L_2 \subseteq \Sigma^*$,则有：$\alpha(L_1 \cdot L_2) = (\alpha L_1) \cdot L_2$.

- 字母表、字、语言

定理1.1.7: 设Σ为字母表,若$L \subseteq \Sigma^*$,则:

i $L^{n+1} = L^n \cdot L$;

ii $L^n \cdot L^m = L^{n+m}$;

iii $(L^n)^m = L^{nm}$;

- 字母表、字、语言

定义1.1.5：设Σ为字母表

i 若$L \subseteq \Sigma^*$,则称L为Σ上的一个语言;

ii 若$\mathscr{L} \subseteq 2^{\Sigma^*}$,则称$\mathscr{L}$为$\Sigma$上的一个语言族;

iii 设L为Σ上的一个语言，若$L = \emptyset$,则称L为空语言;若$\#L < \infty$,则称L为有穷语言;若$\#L = \infty$,则称L为无穷语言。

iv 设$\mathscr{L}$为Σ上的一个语言族，若$\mathscr{L} = \emptyset$,则称$\mathscr{L}$为空语言族;若$\#\mathscr{L} < \infty$,则称$\mathscr{L}$为有穷语言族; 若$\#\mathscr{L} = \infty$,则称L为无穷语言族。

- 字母表、字、语言

例：设$\Sigma=\{a,b,c,d\}$,若令：

$L_1=\{a,b,d\}$;

$L_2=\{a\omega b|\omega\in\Sigma^*\}$;

$L_3=\{a^P|P\text{为素数}\}$;

$L_4=\{a^nb^mc^nd^m|n,m\in\mathbb{N}\}$;

$L_5=\{a^nb^nc^n|n\in\mathbb{N}\}$;

$L_6=\{\omega\in\Sigma^*|\omega\text{中所含}a,b\text{与}c\text{的个数之和等于}\omega\text{中所含}d\text{的个数}\}$;

则$L_1,L_2,...,L_6$都是Σ上的语言，L_1为有穷语言，而其余语言都为无穷语言。

若令：

$\mathscr{L}_1=\{L_1,L_2,...,L_6\}$;

$\mathscr{L}_2=\{\{a^n\omega|\omega\in\{b,c\}^*\}|n\in\mathbb{N}\}$;

$\mathscr{L}_3=\{\{\omega\in\Sigma^*|\omega\text{中恰含}n\text{个}a\}|n\in\mathbb{N}\}$;

则$\mathscr{L}_1,\mathscr{L}_2,\mathscr{L}_3$都是$\Sigma$上的语言族，$\mathscr{L}_1$为有穷语言族，而$\mathscr{L}_2$和$\mathscr{L}_3$都是无穷语言族。

1.2. 自由Monoid

定义1.2.1设$< S, \cdot, e >$为Monoid，如果存在非空集合$A \subseteq S$满足:

FM_1:$e \notin A$;

FM_2:若$u \in S$且$x \in A$,则$ux \neq e$;

FM_3:若$u, v \in S$和$x, y \in A$使得$ux = vy$,则$u = v$且$x = y$;

FM_4:对任意的$B \subseteq S$,若B满足：

i $e \in B$;

ii 若$\omega \in B$且$x \in A$,则$\omega x \in B$;

则有$B = S$.

我们就称$< S, \cdot, e >$为一个自由Monoid，并称A为它的基。

例1对自然数集合$\mathbb{N}$上的加法运算+,显然$< \mathbb{N},+,0 >$是一个自由Monoid,其基为集合$\{1\}$.

例1对自然数集合$\mathbb{N}$上的加法运算+,显然$< \mathbb{N}, +, 0 >$是一个自由Monoid,其基为集合$\{1\}$.

例2对任意的$n \in I_+$,若$\mathbb{N}$上的模n加法运算$+_n$定义如下：

$$i +_n j = \begin{cases} i + j & ,\text{若} i + j < n \\ i + j - n & ,\text{否则} \end{cases}$$

则$< \mathbb{N}, +_n, 0 >$是一个Monoid,但不是一个自由Monoid.

例1对自然数集合$\mathbb{N}$上的加法运算+,显然$< \mathbb{N}, +, 0 >$是一个自由Monoid,其基为集合$\{1\}$.

例2对任意的$n \in I_+$,若$\mathbb{N}$上的模n加法运算$+_n$定义如下：

$$i +_n j = \begin{cases} i + j & ,\text{若} i + j < n \\ i + j - n & ,\text{否则} \end{cases}$$

则$< \mathbb{N}, +_n, 0 >$是一个Monoid,但不是一个自由Monoid.

例3对自然数集合上$\mathbb{N}$上的乘法运算$\cdot$,显然$< \mathbb{N}, \cdot, 1 >$是一个Monoid,但不是自由Monoid.

定理1.2.1 若Σ是字母表，则$< \Sigma^*, ., \epsilon >$是一个以$\Sigma$为基的自由Monoid.

定理1.2.1 若Σ是字母表，则$< \Sigma^*, ., \epsilon >$是一个以$\Sigma$为基的自由Monoid.

引理1 设A为自由Monoid$< S, ., e >$的基.

i) 若$x, y \in A$,则$xy \neq e$;

定理1.2.1 若Σ是字母表，则$< \Sigma^*, ., \epsilon >$是一个以$\Sigma$为基的自由Monoid.

引理1 设A为自由Monoid$< S, ., e >$的基.

i) 若$x, y \in A$,则$xy \neq e$;

ii) 若$x, y, z \in A$,则$x \neq yz$;

定理1.2.1 若Σ是字母表，则$< \Sigma^*, ., \epsilon >$是一个以$\Sigma$为基的自由Monoid.

引理1 设A为自由Monoid$< S, ., e >$的基.

i) 若$x, y \in A$,则$xy \neq e$;

ii) 若$x, y, z \in A$,则$x \neq yz$;

iii) 若令：

$$\bar{A} = \{x_1 \cdots x_n | n \in \mathbb{N} \text{并且} x_1, ..., x_n \in A\}$$

则$\bar{A} = S$.

引理2 设A_1和A_2为自由Monoid$< s, ., e >$的基，则$A_1 = A_2$.

引理2 设A_1和A_2为自由Monoid$< s, ., e >$的基，则$A_1 = A_2$.

引理3 设A为自由Monoid$< S, ., e >$的基,$w \in S$且$w \neq e$.

i) 有$u \in S$及$x \in A$使得$w = ux$;

引理2 设A_1和A_2为自由Monoid$< s, ., e >$的基，则$A_1 = A_2$.

引理3 设A为自由Monoid$< S, ., e >$的基,$w \in S$且$w \neq e$.

i) 有$u \in S$及$x \in A$使得$w = ux$;

ii) 有$v \in S$及$y \in A$使得$w = yv$;

定义1.2.2 设Σ为自由Monoid$< S, ., e >$的基.

i) 称Σ为一个字母表，并称Σ中的元素为符号.

ii) 若$w \in S$,则称w为Σ上的一个字.

iii) 称e为空字，并改记为ϵ.

iv) 称S为Σ上所有字的集合，记为Σ^*,即$S = \Sigma^*$

定义1.2.2 设Σ为自由Monoid$< S, ., e >$的基.

i) 称Σ为一个字母表，并称Σ中的元素为符号.

ii) 若$w \in S$,则称w为Σ上的一个字.

iii) 称e为空字，并改记为ϵ.

iv) 称S为Σ上所有字的集合，记为Σ^*,即$S = \Sigma^*$

定理1.2.2 设A为自由Monoid$< S, ., e >$的基,M为任意集合且$m_0 \in M$.对每个函数$g : M \times S \times A \to M$,皆有唯一的一个函数$f : S \to M$满足:

i) $f(e) = m_0$.

ii) 若$w \in S$且$x \in A$,则$f(wx) = g(f(w), w, x)$.

定理1.2.3 若A为自由Monoid$< S, ., e >$的基,则存在唯一的一个函数$l : S \to \mathbb{N}$满足:

i) $l(e) = 0$.

ii) 若$w \in S$且$x \in A$,则$l(wx) = l(w) + 1$.

定理1.2.3 若A为自由Monoid$< S, ., e >$的基,则存在唯一的一个函数$l : S \to \mathbb{N}$满足:

i) $l(e) = 0.$

ii) 若$w \in S$且$x \in A$,则$l(wx) = l(w) + 1.$

定理1.2.4 若A为自由Monoid$< s, ., e >$的基,则存在唯一的一个函数$R : S \to S$满足:

i) $R(e) = e.$

ii) 若$w \in S$且$x \in A$,则$R(wx) = xR(w).$

定义1.2.3 设$< s, \cdot, e >$为自由Monoid,$w \in S$且$L \subseteq S$.

i) 称$l(w)$为w的长度，记为$|w|$.

ii) 称$R(w)$为w的逆，记为w^R.

iii) 称$\{w^R | w \in L\}$为L的逆，记为L^R.

定理1.2.5 设A为自由Monoid$< s, ., e >$的基,则:

i) $|e| = 0$.

ii) 若$u \in S$,则$|u| \geq 0$.

iii) 若$a \in A$,则$|a| = 1$.

iv) 若$a_1, ..., a_n \in A$,则$|a_1 \cdots a_n| = n$.

v) 若$u \in S$且$a \in A$,则$|au| = |u| + 1 = |ua|$.

vi) 若$u, v \in S$且$a \in A$,则$|uv| = |u| + |v| = |vu|$.

访问主页
标题页
第 24 页 共 47 页
返回
全屏显示
关闭
退出

定理1.2.5 设A为自由Monoid$< s, ., e >$的基,则:

i) $|e| = 0$.

ii) 若$u \in S$,则$|u| \geq 0$.

iii) 若$a \in A$,则$|a| = 1$.

iv) 若$a_1, ..., a_n \in A$,则$|a_1 \cdots a_n| = n$.

v) 若$u \in S$且$a \in A$,则$|au| = |u| + 1 = |ua|$.

vi) 若$u, v \in S$且$a \in A$,则$|uv| = |u| + |v| = |vu|$.

定理1.2.6 设A为自由Monoid$< s, ., e >$的基,则:

i) $e^R = e$.

ii) 若$a \in A$,则$a^R = a$.

iii) 若$a \in A$且$u \in S$,则$(au)^R = u^R a$且$(ua)^R = au^R$.

iv) 若$u, v \in S$,则$(uv)^R = v^R u^R$.

v) 若$u \in S$,则$(u^R)^R = u$.

vi) 若$a_1, ..., a_n \in A$,则$(a_1 \cdots a_n)^R = a_n \cdots a_1$.

定理1.2.7 设$< s, ., e >$为自由Monoid,则:

i) $\emptyset^R = \emptyset$.

ii) 若$u \in S$且$L \subseteq S$,则$(uL)^R = L^R u^R$且$(Lu)^R = u^R L^R$.

iii) 若$L_1 \subseteq S$且$L_2 \subseteq S$,则$(L_1 L_2)^R = L_2^R L_1^R$.

iv) 若$L \subseteq S$,则$(L^R)^R = L$.

定理1.2.8 若A为自由Monoid$< s, ., e >$的基,$< s', *, e' >$为Monoid,则对每个函数$f_0 : A \to S'$,都存在唯一的一个同态$f : S \to S'$为f_0之延拓.

定理1.2.9 设$< s, ., e >$为Monoid,若存在非空集合$A \subseteq S$,使得对任意Monoid$< S', *, e' >$及任意函数$f_0 : A \rightarrow S'$,皆有唯一的一个同态$f : S \rightarrow S'$为f_0的延拓，则$s, ., e$为一个以A为基的自由Monoid.

定理1.2.10 设A为自由Monoid$< s, ., e >$的基,A'为自由Monoid$< s', *, e' >$的基,若$\#A = \#A'$,则$< s, ., e >$和$< s', *, e' >$同构.

1.3. 字的组合性质

定理1.3.1 设$u, v, x, y \in S$且$uv = xy$.

i) 若$|u| \geq |x|$,则有唯一的$z \in S$使$u = xz$且$y = zv$;

ii) 若$|u| = |x|$,则$u = x$且$v = y$.

1.3. 字的组合性质

定理1.3.1 设$u, v, x, y \in S$且$uv = xy$.

i) 若$|u| \geq |x|$,则有唯一的$z \in S$使$u = xz$且$y = zv$;

ii) 若$|u| = |x|$,则$u = x$且$v = y$.

推论1.3.2 若$x, y \in \Sigma$,$u, v \in S$且$xu = yv$,则:

$x = y$且$u = v$.

定义1.3.1 设$u, v \in S$.

i) 若有$w \in S$使$u = vw$,则称v为u的前缀，记为$v \leq u$;

ii) 若有$w \in S$使$u = wv$,则称v为u的后缀，记为$v \geq u$;.

定义1.3.1 设$u, v \in S$.

i) 若有$w \in S$使$u = vw$,则称v为u的前缀，记为$v \leq u$;

ii) 若有$w \in S$使$u = wv$,则称v为u的后缀，记为$v \geq u$;.

定理1.3.3 设$u, v, x, y \in S$,

i) $v \leq u$当且仅当$v^R \geq u^R$.

ii) 若$uv = xy$,则$u \leq x$或$x \leq u$.

iii) 若$u \leq x$且$v \leq xy$,则$u \leq v$或$v \leq u$.

定理1.3.4 设$x, y, z \in S, x \neq \epsilon$且$z \neq \epsilon$,若$xy = yz$,则有$u, v \in S$及$n \in \mathbb{N}$使:

$$x = uv, z = vu \text{且} y = (uv)^n u = u(vu)^n;$$

定理1.3.4 设$x, y, z \in S, x \neq \epsilon$且$z \neq \epsilon$,若$xy = yz$,则有$u, v \in S$及$n \in \mathbb{N}$使:

$$x = uv, z = vu \text{且} y = (uv)^n u = u(vu)^n;$$

定理1.3.5 若$u, v \in S$,则以下两条件互相等价:

i) 有$w \in S$及$m, n \in \mathbb{N}$使$u = w^m$且$v = w^n$.

ii) 有$p, q \in \mathbb{N}$使u^P和v^q含有长为$|u| + |v| - (|u|, |v|)$的公共前缀，其中(i, j)表示自然数i和j的最大公因数.

定理1.3.4 设$x, y, z \in S, x \neq \epsilon$且$z \neq \epsilon$,若$xy = yz$,则有$u, v \in S$及$n \in \mathbb{N}$使:

$$x = uv, z = vu \text{且} y = (uv)^n u = u(vu)^n;$$

定理1.3.5 若$u, v \in S$,则以下两条件互相等价:

i) 有$w \in S$及$m, n \in \mathbb{N}$使$u = w^m$且$v = w^n$.

ii) 有$p, q \in \mathbb{N}$使u^P和v^q含有长为$|u| + |v| - (|u|, |v|)$的公共前缀,其中(i, j)表示自然数i和j的最大公因数.

定理1.3.6 若$u, v \in S$,若$uv = vu$,则有$w \in S$及$m, n \in \mathbb{N}$使$u = w^m$且$v = w^n$.

1.4. 置换

定义1.4.1 设Σ和Δ为两个字母表，则称函数$f : \Sigma \to 2^{\Delta^*}$为一个从$\Sigma$到$\Delta$的置换.如果对每个$a \in \Sigma$皆有$f(a) \in \Delta^*$,则称置换$f$为同态.

1.4. 置换

定义1.4.1 设Σ和Δ为两个字母表，则称函数$f : \Sigma \to 2^{\Delta^*}$为一个从$\Sigma$到$\Delta$的置换.如果对每个$a \in \Sigma$皆有$f(a) \in \Delta^*$,则称置换$f$为同态.

例1 取$\Sigma\{0,1\}$,$\Delta = \{a,b,c\}$,若令:

$$f_i(0) = a, 1 \leq i \leq .$$

且

$$f_1(1) = \emptyset;$$

$$f_2(1) = \epsilon;$$

$$f_3(1) = \{a^n b^n c^n | n \in \mathbb{N}\text{且}n \geq 1\};$$

$$f_4(1) = \{ww^R | w \in \{b,c\}^*\};$$

则f_1, f_2, f_3和f_4都是从$\{0,1\}$到$\{a,b,c\}$的置换，而且f_2还是同态.

定理1.4.1 若Σ和Δ为两个字母表，f为从Σ到Δ的置换，则必存在唯一的一个函数$F_f : \Sigma^* \to 2^{\Delta^*}$满足：

i) 若$a \in \Sigma$,则$F_f(a) = f(a)$;

ii) 若$u, v \in \Sigma^*$,则$F_f(uv) = F_f(u)F_f(v)$.

定义1.4.2 设Σ和Δ为两个字母表，

i) 若f为从Σ到Δ的置换且$L \subseteq \Sigma^*$，则令：

$$f(L) = \cup_{w \in L} f(w)$$

ii) 若h为从Σ到Δ的同态，$x \in \Delta^*$且$\widetilde{L} \subseteq \Delta^*$，则令：

$$h^{-1}(x) = \{w \in \Sigma^* | h(w) = x\};$$

$$h^{-1}(\widetilde{L}) = \{w \in \Sigma^* | h(w) \in \widetilde{L}\} = \cup_{x \in \widetilde{L}} h^{-1}(x);$$

并称h^{-1}为h的逆同态.

例2 取$\Sigma = \{0, 1\}$, $\Delta = \{a, b\}$,若定义从Σ到Δ的同态h_1如下：

$h_1(0) = aa$; $h_1(1) = aba$;

则对$010, 101 \in \Sigma^*$显然有：

$h_1(010)aaabaaa$且$h_1(101) = abaaaaba$

若取$L_1 = \{01\}^*$及$L_2 = \{ab, ba\}^*\{a\}$,则显然有：

$h_1(L_1) = \{aaaba\}^*$;$h_1^{-1}(L_2) = \{1\}$;

因此$h_1(h_1^{-1}(L_2)) = \{aba\} \neq L_2$.

例2 取$\Sigma = \{0,1\}, \Delta = \{a,b\}$,若定义从$\Sigma$到$\Delta$的同态$h_1$如下：

$$h_1(0) = aa;\ h_1(1) = aba;$$

则对$010, 101 \in \Sigma^*$显然有：

$$h_1(010)aaabaaa\text{且}h_1(101) = abaaaaba$$

若取$L_1 = \{01\}^*$及$L_2 = \{ab, ba\}^*\{a\}$,则显然有：

$$h_1(L_1) = \{aaaba\}^*;h_1^{-1}(L_2) = \{1\};$$

因此$h_1(h_1^{-1}(L_2)) = \{aba\} \neq L_2$.

例3 取$\Sigma = \{0,1\}, \Delta = \{a,b\}$,若定义从$\Sigma$到$\Delta$的同态$h_2$如下：

$$h_2(0) = h_2(1) = \epsilon;$$

则显然有：

$$h_2(\Sigma)\{\epsilon\}\text{且}h_2^{-1}(\Delta) = \emptyset$$

因此

$$h_2^{-1}(h_2(\Sigma)) = \Sigma^*;\ h_2^{-1}(h_2(\Delta)) = \emptyset.$$

定 理1.4.2 设Σ和Δ为两个字母表，f为从Σ到Δ的置换，如果$L, L_1, L_2 \in 2^{\Sigma^*}$,则：

$$f(L_1 \cup L_2) = f(L_1) \cup f(L_2);$$

$$f(L_1 L_2) = f(L_1) f(L_1);$$

$$f(L^*) = f(L)^*;$$

$$f(L_1 \wedge L_2) \subseteq f(L_1) \wedge f(L_2);$$

定理1.4.2 设Σ和Δ为两个字母表，f为从Σ到Δ的置换，如果$L, L_1, L_2 \in 2^{\Sigma^*}$,则：

$$f(L_1 \cup L_2) = f(L_1) \cup f(L_2);$$

$$f(L_1 L_2) = f(L_1) f(L_1);$$

$$f(L^*) = f(L)^*;$$

$$f(L_1 \wedge L_2) \subseteq f(L_1) \wedge f(L_2);$$

定理1.4.3 设Σ和Δ为两个字母表，h为从Σ到Δ的同态，如果$L, L_1, L_2 \in 2^{\Delta^*}$,则：

$$h^{-1}(L_1 \cup L_2) = h^{-1}(L_1) \cup h^{-1}(L_2);$$

$$h^{-1}(L_1 \wedge L_2) = h^{-1}(L_1) \wedge h^{-1}(L_2);$$

$$h^{-1}(\sim L) = \sim h^{-1}(L);$$

$$h(h^{-1}(L)) \subseteq L;$$

若$\tilde{L} \subseteq \Sigma^*$,则$\tilde{L} \subseteq h^{-1}(h(\tilde{L}))$;

定义1.4.3 设Σ和Δ为两个字母表，f为从Σ到Δ的置换：

i) 若对每个$a \in \Sigma$皆有$\epsilon \notin f(a)$,则称f为$\epsilon-$无关的;

ii) 若对每个$a \in \Sigma$皆有$\# f(a) < \infty$,则称f为有穷的;

设$L \subseteq \Sigma^*$, 若对每个$w \in L$, 当$w = xuy$且$|u| > k$时，皆有$f(w) \neq f(x)f(y)$, 则称f对L为$k-$抹去的;

设$\mathscr{L} \subseteq 2^{\Sigma^*}$, 若对每个$L \in \mathscr{L}$, f对L都是$k-$抹去的, 则称f对$\mathscr{L}$为$k-$抹去的;

定义1.4.3 设Σ和Δ为两个字母表，f为从Σ到Δ的置换：

i) 若对每个$a \in \Sigma$皆有$\epsilon \notin f(a)$,则称f为$\epsilon-$无关的;

ii) 若对每个$a \in \Sigma$皆有$\# f(a) < \infty$,则称f为有穷的;

设$L \subseteq \Sigma^*$，若对每个$w \in L$，当$w = xuy$且$|u| > k$时，皆有$f(w) \neq f(x)f(y)$, 则称f对L为$k-$抹去的;

设$\mathscr{L} \subseteq 2^{\Sigma^*}$，若对每个$L \in \mathscr{L}$，$f$对$L$都是$k-$抹去的，则称$f$对$\mathscr{L}$为$k-$抹去的;

定义1.4.4 给定语言类$\mathscr{L}$和置换类$\mathscr{S}$，若对每个$L \in \mathscr{L}$和$f \in \mathscr{S}$,当$a \in_{dom} (f)$时,若$f(a) \in \mathscr{L}$,则必有$f(L) \in \mathscr{L}$.就称$\mathscr{L}$关于$\mathscr{S}$为封闭的.

定义1.4.3 设Σ和Δ为两个字母表，f为从Σ到Δ的置换：

i) 若对每个$a \in \Sigma$皆有$\epsilon \notin f(a)$,则称f为$\epsilon-$无关的;

ii) 若对每个$a \in \Sigma$皆有$\# f(a) < \infty$,则称f为有穷的;

设$L \subseteq \Sigma^*$，若对每个$w \in L$，当$w = xuy$且$|u| > k$时，皆有$f(w) \neq f(x)f(y)$, 则称f对L为$k-$抹去的;

设$\mathscr{L} \subseteq 2^{\Sigma^*}$，若对每个$L \in \mathscr{L}$，$f$对$L$都是$k-$抹去的，则称$f$对$\mathscr{L}$为$k-$抹去的;

定义1.4.4 给定语言类$\mathscr{L}$和置换类$\mathscr{S}$，若对每个$L \in \mathscr{L}$和$f \in \mathscr{S}$,当$a \in_{dom} (f)$时,若$f(a) \in \mathscr{L}$,则必有$f(L) \in \mathscr{L}$.就称$\mathscr{L}$关于$\mathscr{S}$为封闭的.

定义1.4.5 对给定语言类$\mathscr{L}$和同态类$\mathscr{H}$，若对任意的$L \in \mathscr{L}$和$h \in \mathscr{H}$,皆有$h^{-1}(L) \in \mathscr{L}$,则称$\mathscr{L}$关于$\mathscr{H}$中的逆同态为封闭的.

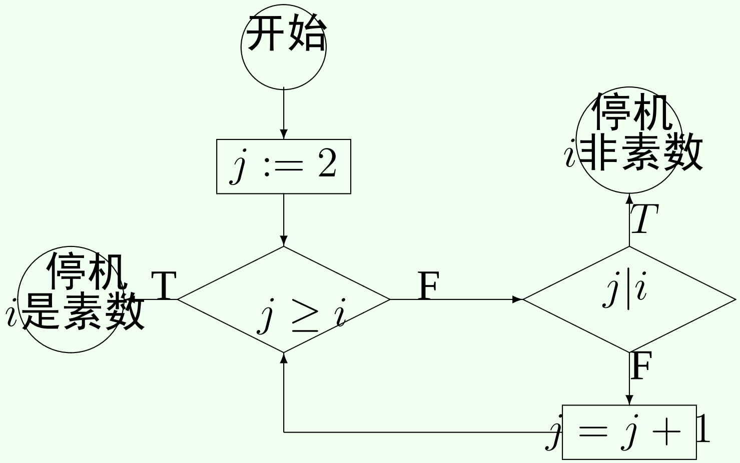
1.5. 过程与算法

过程：就是一个能够被机械地执行的有穷指令序列。例如：一个计算机程序就是一个过程.

1.5. 过程与算法

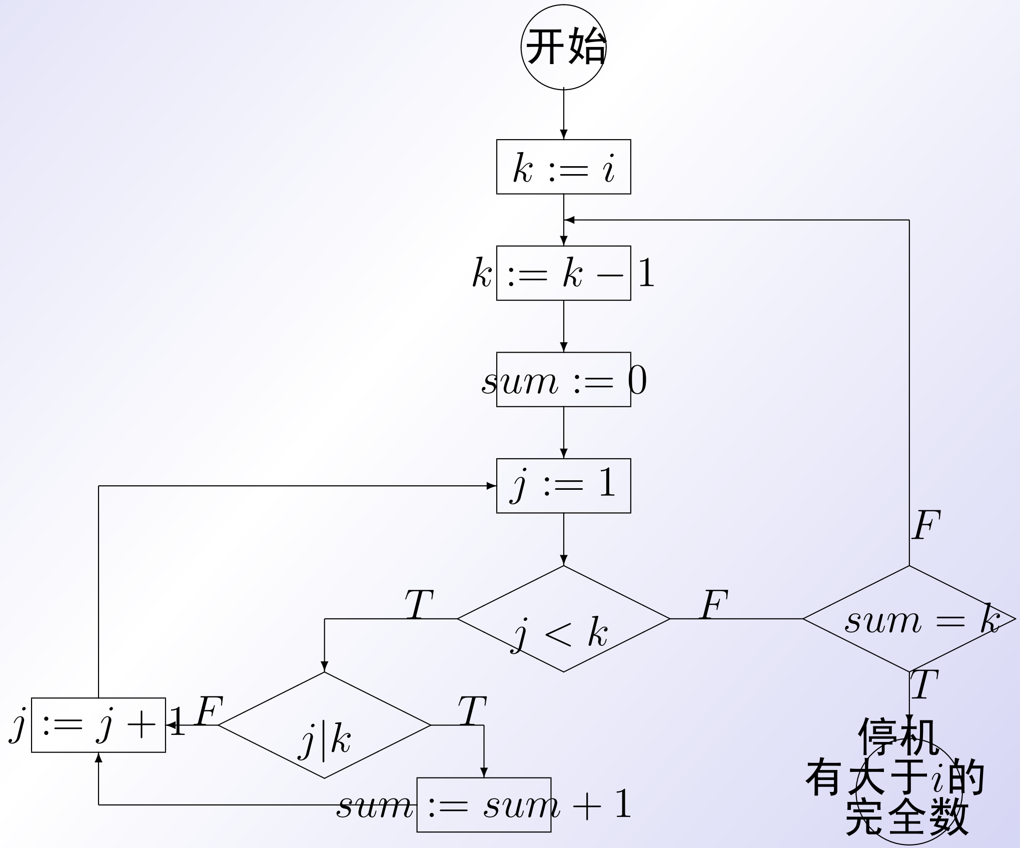
过程： 就是一个能够被机械地执行的有穷指令序列。例如：一个计算机程序就是一个过程.

例1 下图所示是一个判定比1大的整数i是否是素数的过程：



判定$i > 1$是否为素数的过程

例2 下图所示就是一个判定是否存在大于已给整数i的完全数之过程：



判定是否有一个大于i的完全数的过程

算法：称总会停机的过程为算法.

1.6. 语言的表示

大多数语言，如各种程序设计语言，都是无穷语言.因此，我们不能把它的所有句子都一一列举出来.为了更好地研究和掌握它们，就需要寻找它们的有穷表示.

1.6. 语言的表示

大多数语言，如各种程序设计语言，都是无穷语言.因此，我们不能把它的所有句子都一一列举出来.为了更好地研究和掌握它们，就需要寻找它们的有穷表示.

设L为字母表Σ上的一个语言，可以从识别L和产生L这两个方面来分别考虑L的表示法.

从识别方面考虑：$L=\{w|w\in\Sigma^*$且$P_{识}(w)$为$true\}$.

给定一个过程P,对每个$w\in\Sigma^*$,P的执行结果为：

i) 若$w\in L$,则P停机并回答”是”;

ii) 若$w\notin L$,则P不停机，或者P停机并回答”否”;

这时我们就称P为L的一个识别过程.当P为一个算法时，称P为L的一个识别算法。

从产生方面考虑：$L = \{P_{生}(i) | i \in \mathbb{N}\} = Run(P_{生})$.

给定一个过程P,如果P能依某种次序系统地产生出L的每个句子，就称P为L的一个产生过程。当P为一个算法时，称P为L的一个产生算法.

任何无穷语言都不可能有产生算法.

下面来讨论语言这两种表示法之间的关系.

下面来讨论语言这两种表示法之间的关系.

在讨论之前先介绍一个重要的函数$\pi^2:\mathbb{N}^2\rightarrow\mathbb{N}$:

$$\pi^2(x,y)=\frac{(x+y)(x+y+1)}{2}+y$$

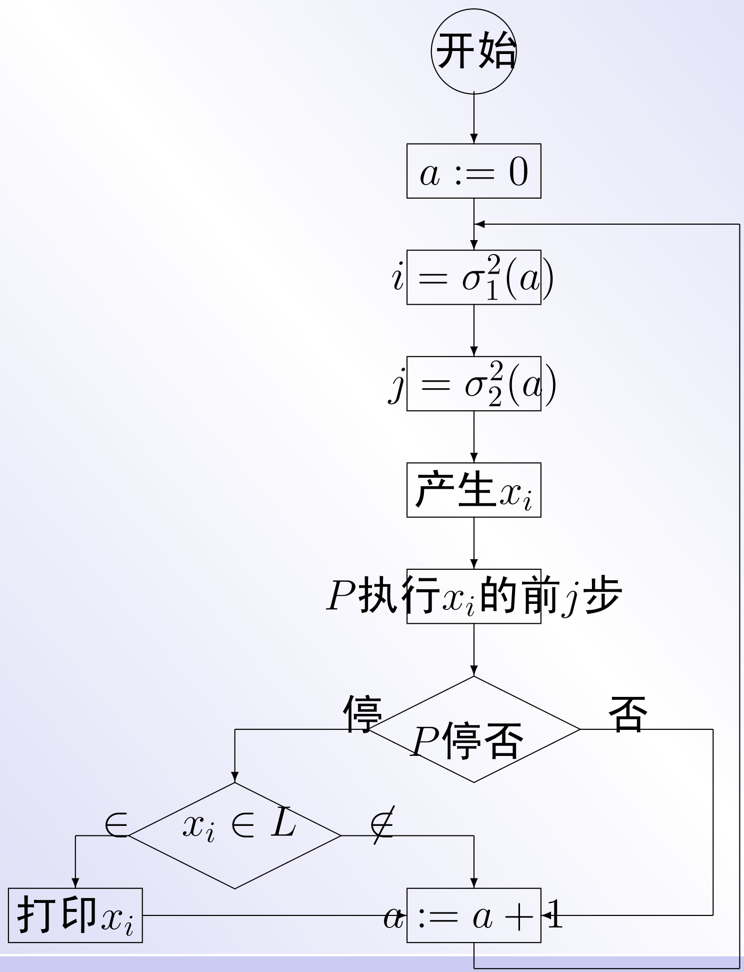
下面来讨论语言这两种表示法之间的关系.

在讨论之前先介绍一个重要的函数$\pi^2 : \mathbb{N}^2 \to \mathbb{N}$:

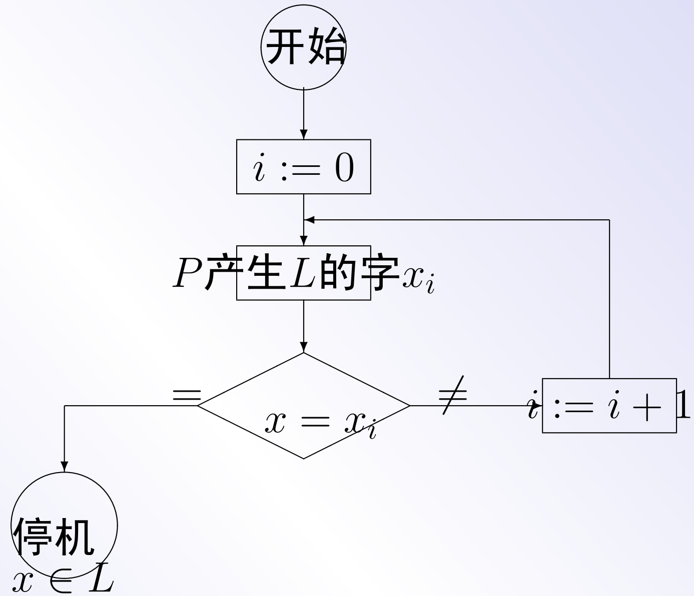
$$\pi^2(x, y) = \frac{(x+y)(x+y+1)}{2} + y$$

引理1： $\pi^2 : \mathbb{N}^2 \to \mathbb{N}$是一个双射.

定理1.6.1： 设L为字母表上的一个语言.若P为L的一个识别过程，则有一个产生L的过程Q.



定理1.6.2： 设L为字母表上的一个语言.若P为L的一个产生过程，则有L的一个识别过程Q.



定义1.6.1： 设L为字母表上的一个语言:

i) 若存在一个产生L的过程，则称L为递归可枚举的;

定义1.6.1：设L为字母表上的一个语言:

i) 若存在一个产生L的过程，则称L为递归可枚举的;

ii) 若存在一个识别L的过程，则称L为递归的;