第三讲　非正则语言与泵引理

罗　欢
Email：hluo@fzu.edu.cn

主要内容

- 正则表达式与自动机等价性
- 非正则语言
- 泵引理
- 泵引理应用

主要内容

- 正则表达式与自动机等价性
- 非正则语言
- 泵引理
- 泵引理应用

回顾：正则表达式

正则表达式的形式定义

- 定义2.26: R是正则表达式
  当且仅当R是
  - a, a$\in\Sigma$;　　/* 表示 {a} */
  - ε;　　/* 表示 {ε} */
  - $\emptyset$;　　/* 表示 $\emptyset$ */
  - $(R_1 \cup R_2)$, R_1和R_2都是正则表达式;
  - (R_1R_2), R_1和R_2都是正则表达式;
  - $(R_1{}^*)$, R_1是正则表达式.
- L(R): R表示的语言

正则表达式与自动机等价性

- 定理2.28: 一个语言是正则的当且仅当可用正则表达式描述这个语言.
- 引理2.29:正则表达式描述正则语言.
- 引理2.32:正则语言可用正则表达式描述.

引理2.29

- 引理2.29:正则表达式描述正则语言.

引理2.29

- 引理2.29:正则表达式描述正则语言.
- 证明思路: 把正则表达式转化成 等价NFA

引理2.29

- 引理2.29:正则表达式描述正则语言.
- 证明: 1) $R=a,\ a\in\Sigma$.

$L(R)=\{a\}$,

$N=(\{q_1,q_2\},\Sigma,\delta,q_1,\{q_2\})$,

$\delta(q_1,a)=\{q_2\}$,

$\delta(r,b)=\varnothing$, 若$r\neq q_1$或$b\neq a$.

N

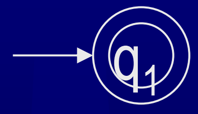
引理2.29

- 引理2.29:正则表达式描述正则语言.
- 证明: 2) $R=\varepsilon$.

$L(R)=\{\varepsilon\}$,

$N=(\{q_1\},\Sigma,\delta,q_1,\{q_1\})$,

$\forall r,\forall b,\ \delta(r,b)=\varnothing$.



N

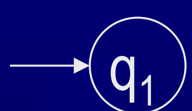
引理2.29

- 引理2.29:正则表达式描述正则语言.
- 证明: 3) $R=\varnothing$.

$L(R)=\varnothing$,

$N=(\{q_1\},\Sigma,\delta,q_1,\varnothing)$,

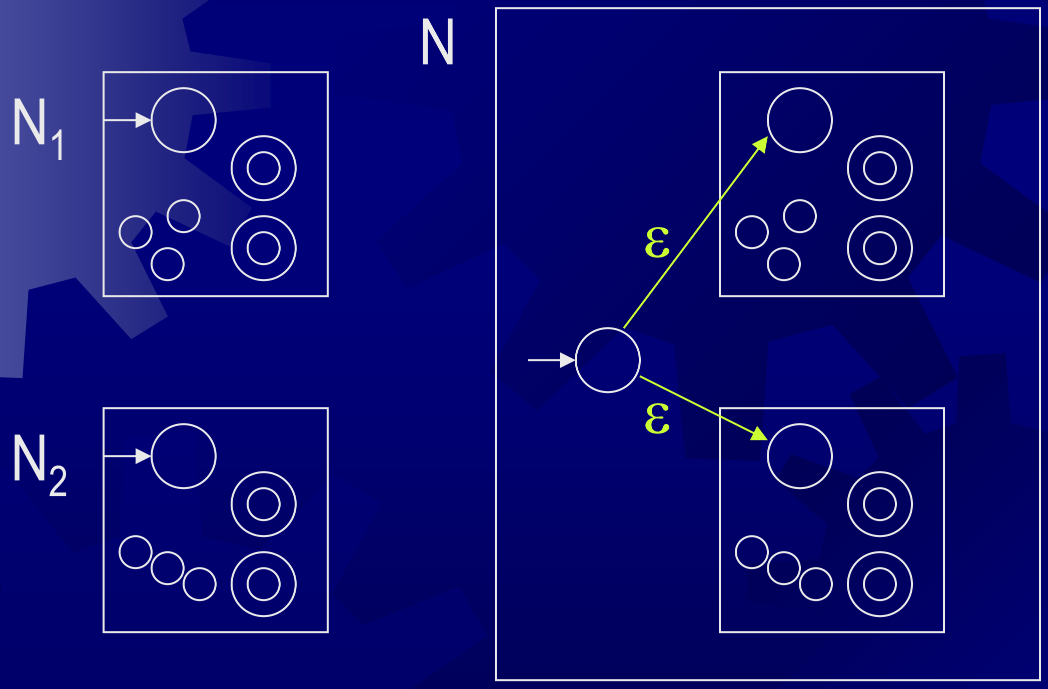
$\forall r,\forall b,\ \delta(r,b)=\varnothing$.



N

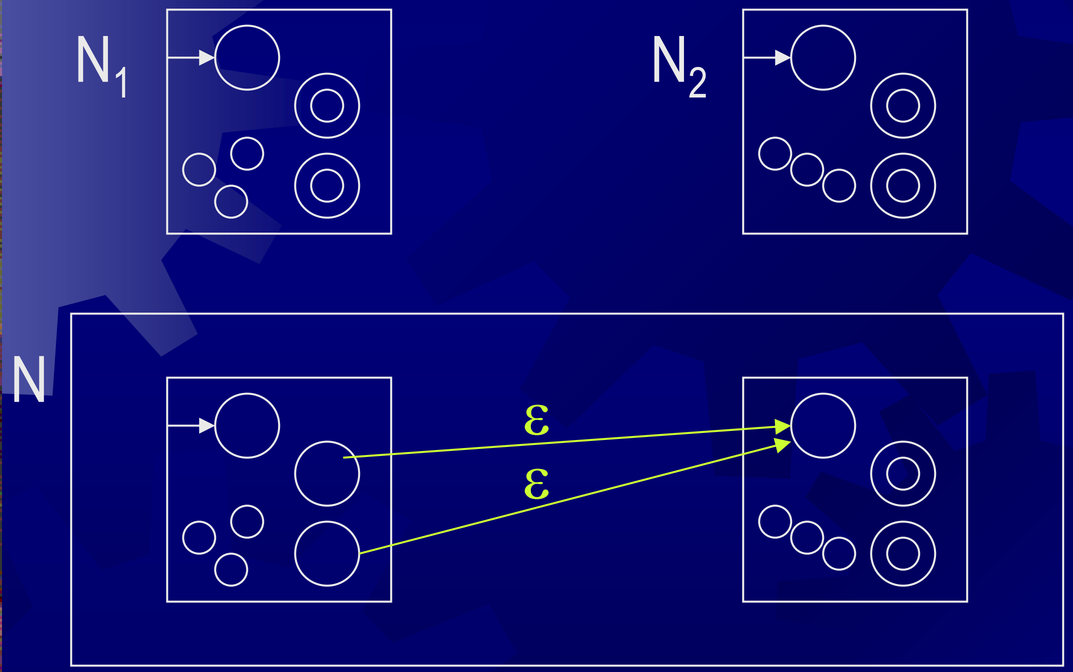
引理2.29

- 引理2.29:正则表达式描述正则语言.
- 证明: 4) $R=(R_1 \cup R_2)$,



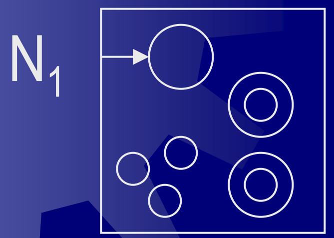
引理2.29

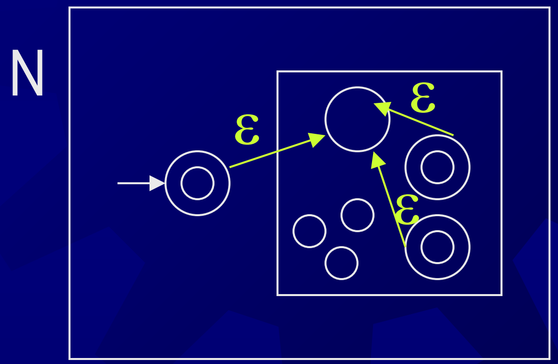
- 引理2.29:正则表达式描述正则语言.
- 证明: 5) $R=(R_1R_2)$,



引理2.29

- 引理2.29:正则表达式描述正则语言.
- 证明: 6) $R=(R_1{}^*)$,





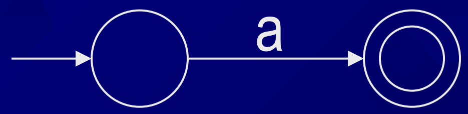
#

例2.30

- $(ab \cup a)^*$

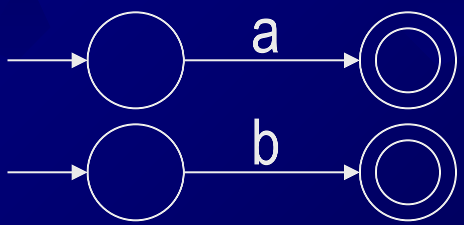
例2.30

- (ab∪a)*
- a



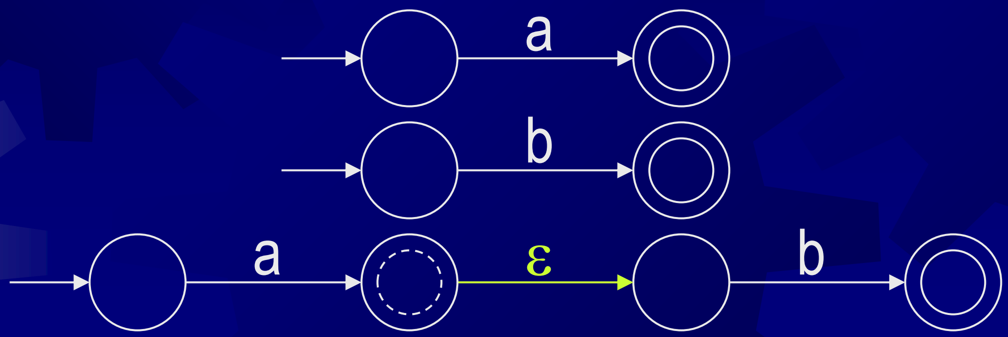
例2.30

- (ab∪a)*
- a
- b



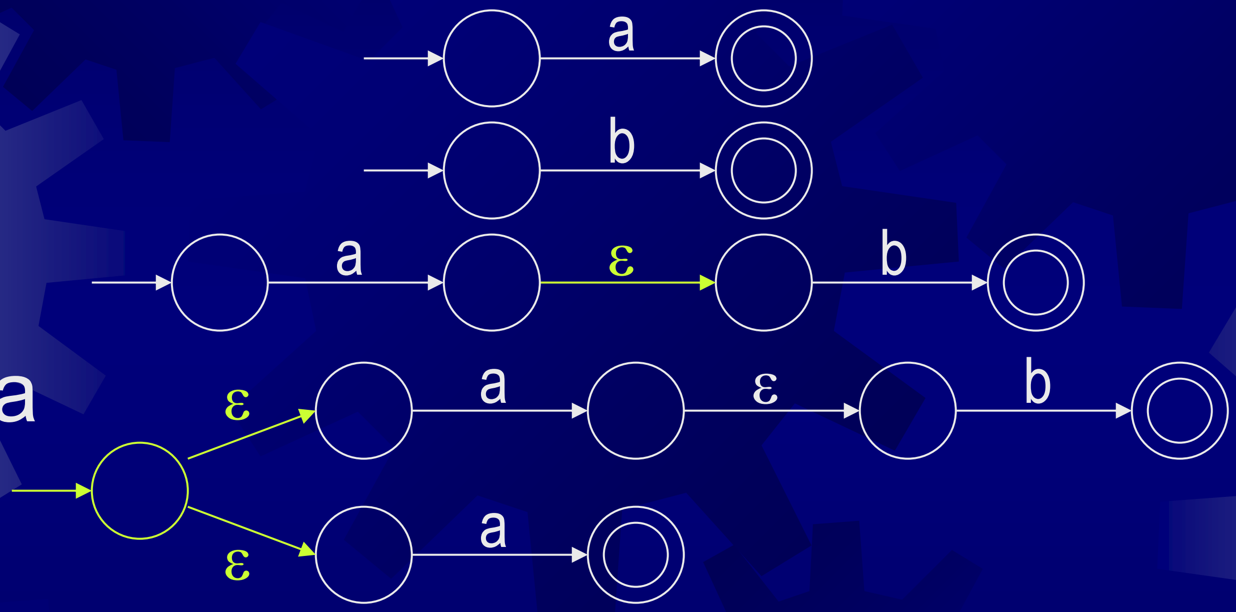
例2.30

- (ab∪a)*
- a
- b
- ab



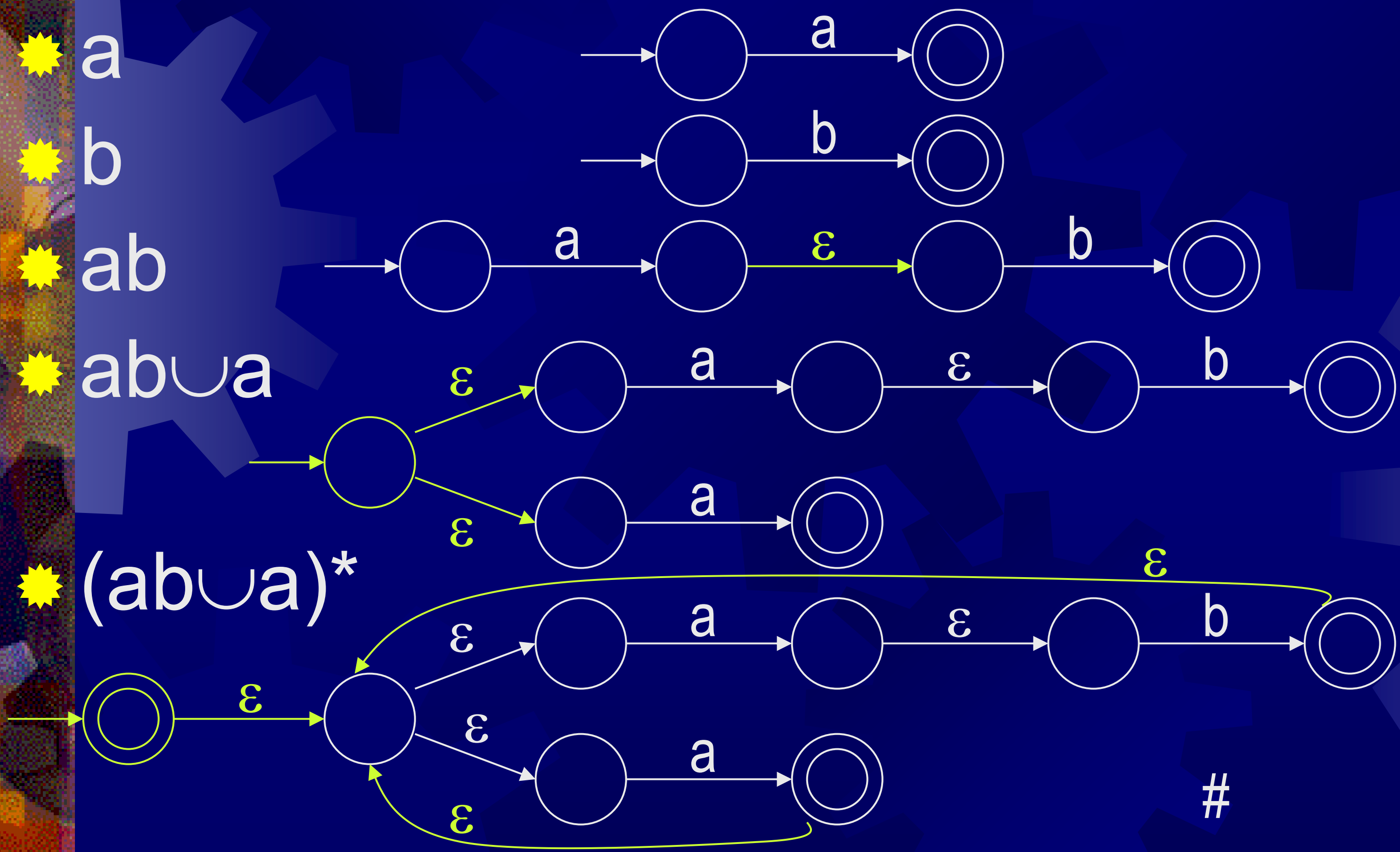
例2.30

- (ab∪a)*
- a
- b
- ab
- ab∪a



例2.30

- (ab∪a)*
- a
- b
- ab
- ab∪a
- (ab∪a)*



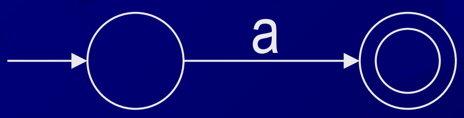
#

例2.31

- $(a \cup b)^*aba$

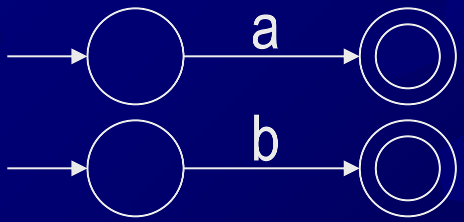
例2.31

- (a∪b)*aba
- a



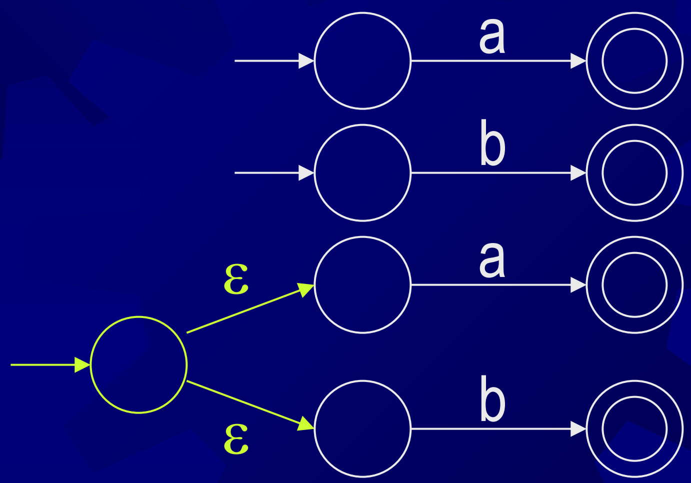
例2.31

- (a∪b)*aba
- a
- b



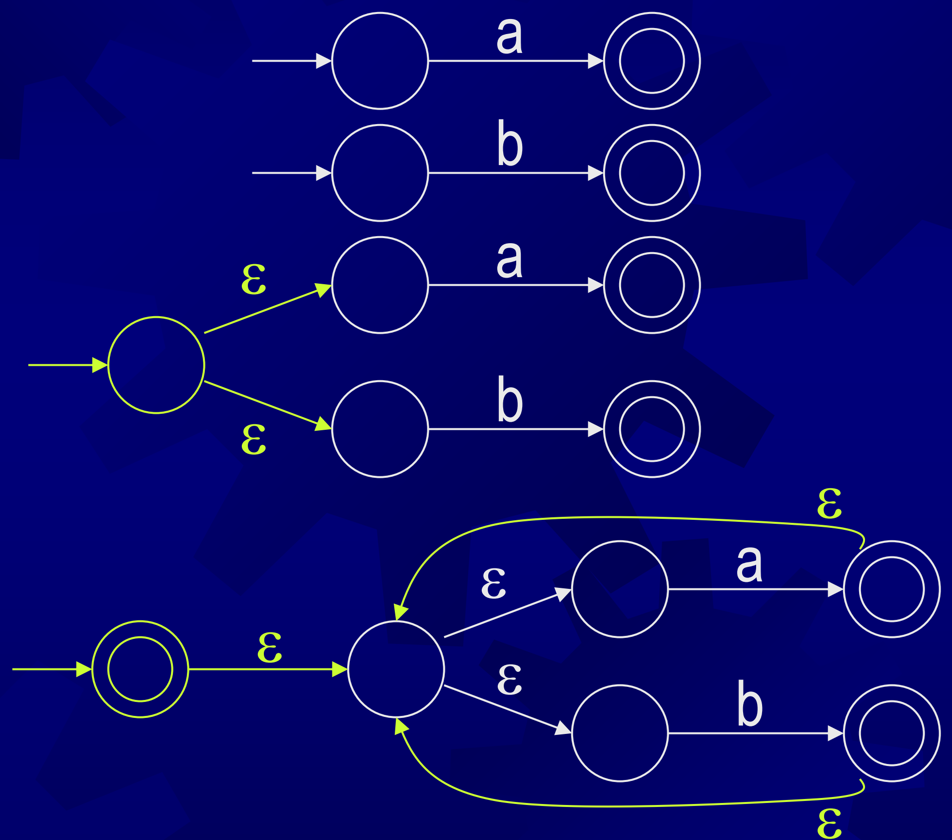
例2.31

- (a∪b)*aba
- a
- b
- a∪b



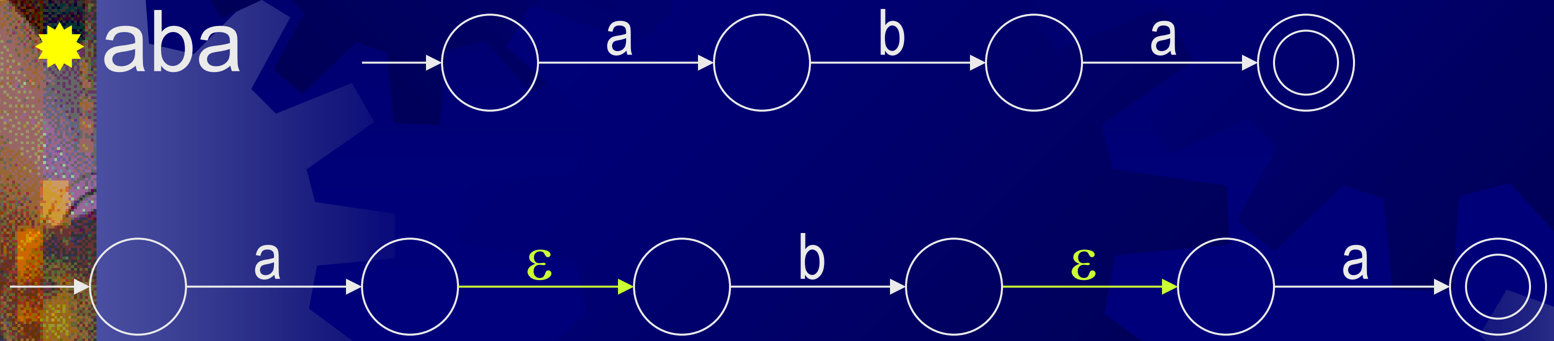
例2.31

- (a∪b)*aba
- a
- b
- a∪b
- (a∪b)*



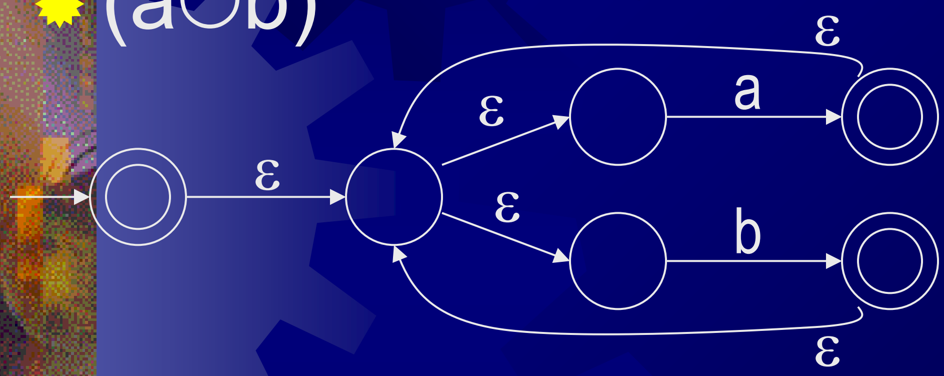
例2.31

- $(a\cup b)^*aba$
- aba

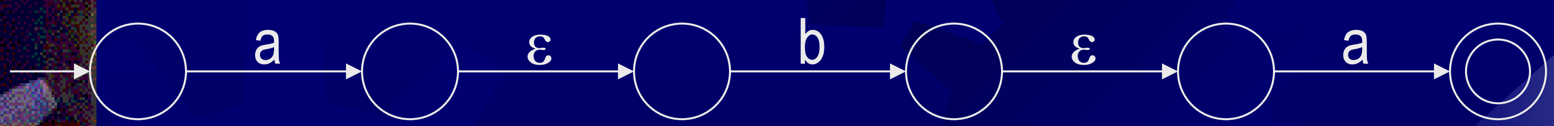


例2.31

- (a∪b)*aba
- (a∪b)*

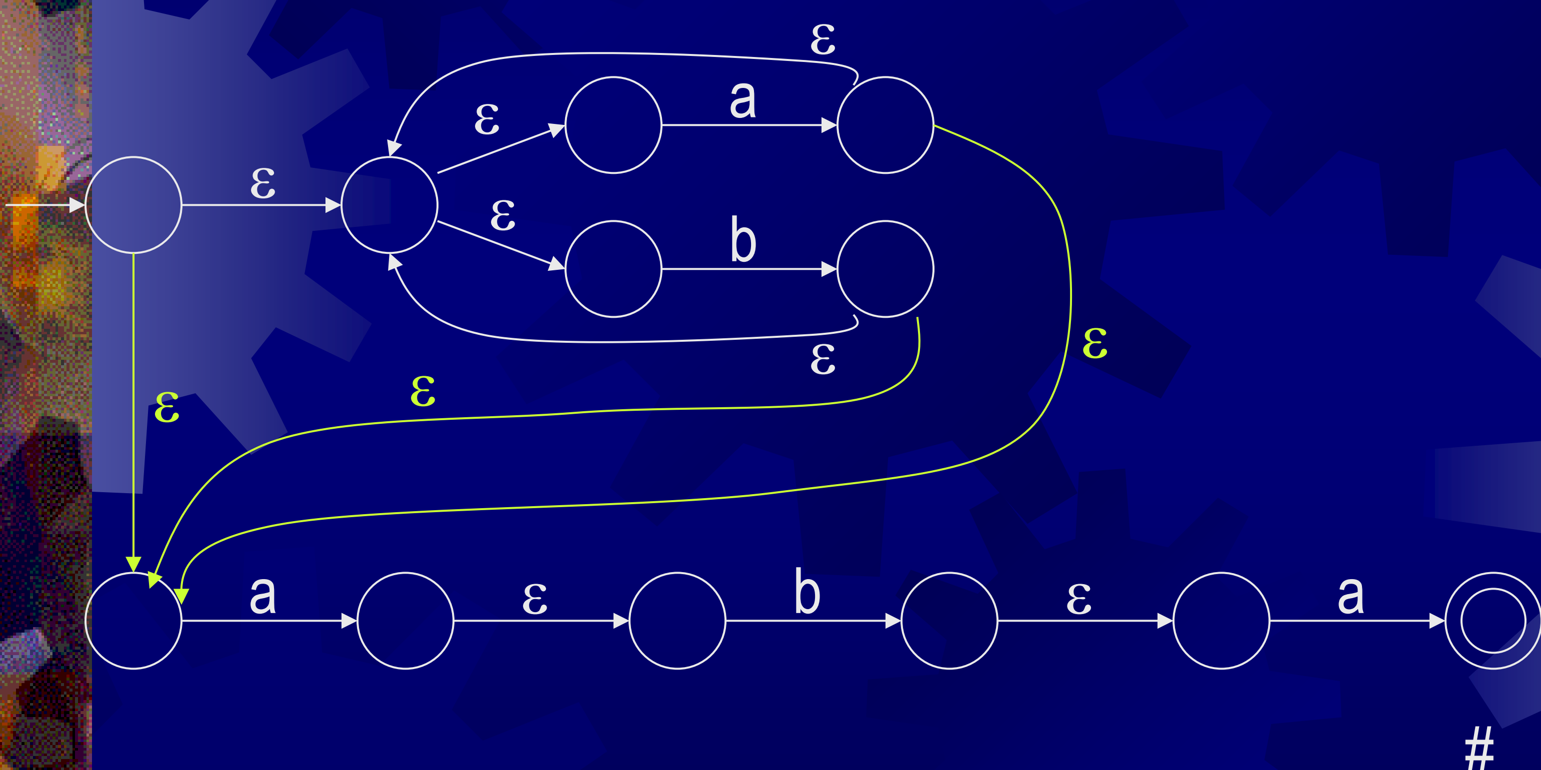


- aba



例2.31

- $(a\cup b)^*aba$



引理2.32

- 引理2.32: 正则语言可用正则表达式描述.

引理2.32

- 引理2.32: 正则语言可用正则表达式描述.
- 证明思路
  - 给定DFA或NFA,构造等价正则表达式

引理2.32

- 引理2.32: 正则语言可用正则表达式描述.
- 证明思路
  - 给定DFA或NFA,构造等价正则表达式
  - 刚才从正则表达式构造等价NFA, 逆过程?

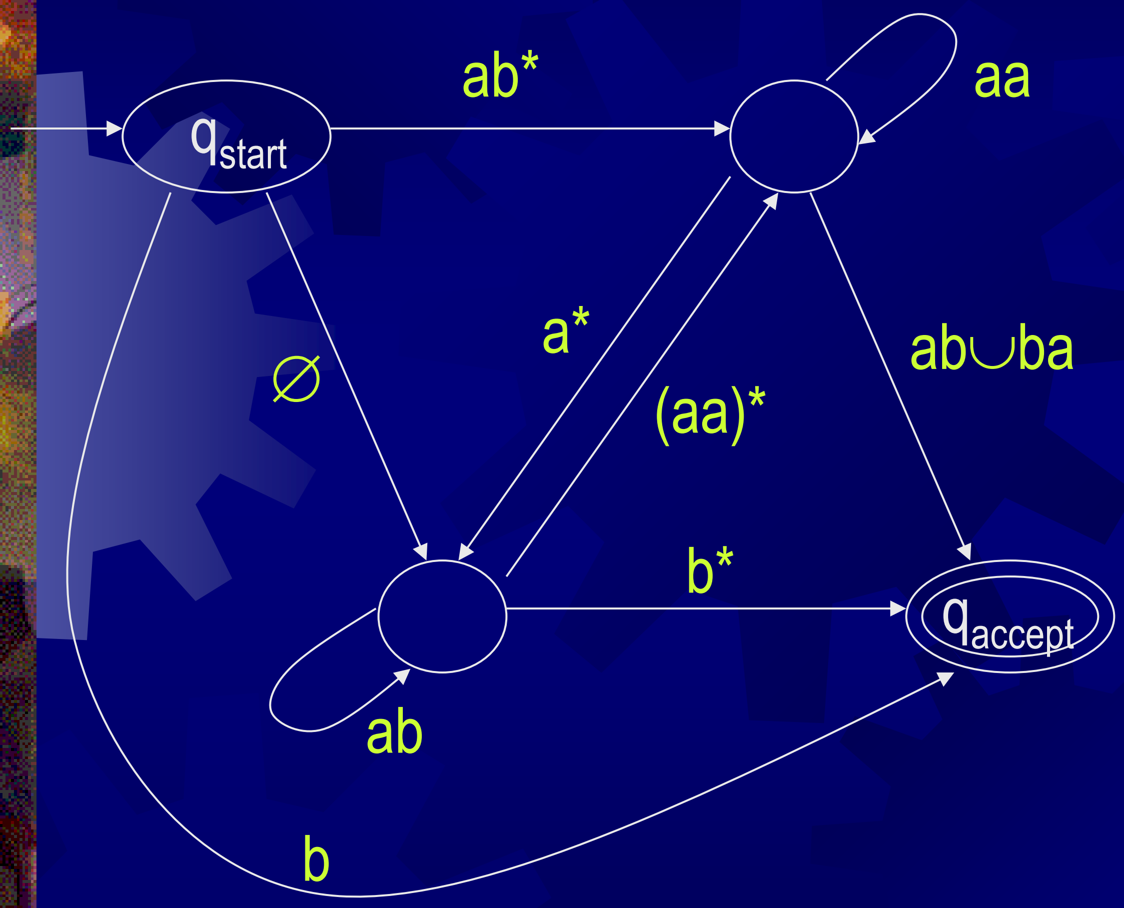
引理2.32

- 引理2.32: 正则语言可用正则表达式描述.
- 证明思路
  - 给定DFA或NFA,构造等价正则表达式
  - 刚才从正则表达式构造等价NFA, 逆过程?
  - 刚才所构造NFA具有特殊形式

引理2.32

- 引理2.32: 正则语言可用正则表达式描述.
- 证明思路
  - 广义非确定型有穷自动机(GNFA)
  - 从DFA构造等价的GNFA
  - 从GNFA构造等价的正则表达式

GNFA: 箭头标号是正则表达式

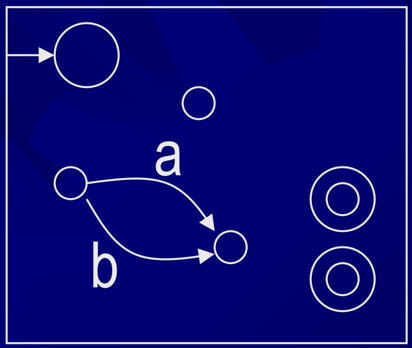


GNFA的特殊形式

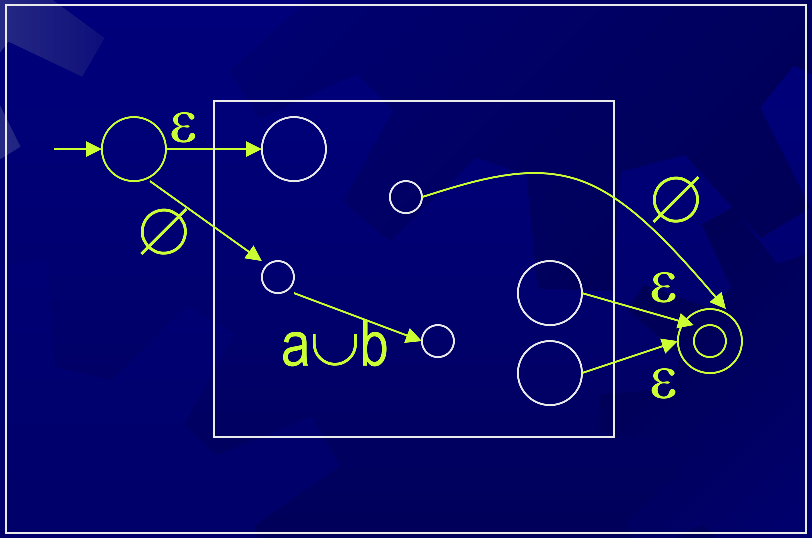
- 初始状态
  - 有到所有其他状态的箭头
  - 所有其他状态都没有到它的箭头
- 接受状态
  - 唯一，且与初始状态不同
  - 没有到任何其他状态的箭头
  - 所有其他状态都有到它的箭头
- 其他每个状态
  - 都有到自身和其他状态的箭头

从DFA到GNFA

DFA

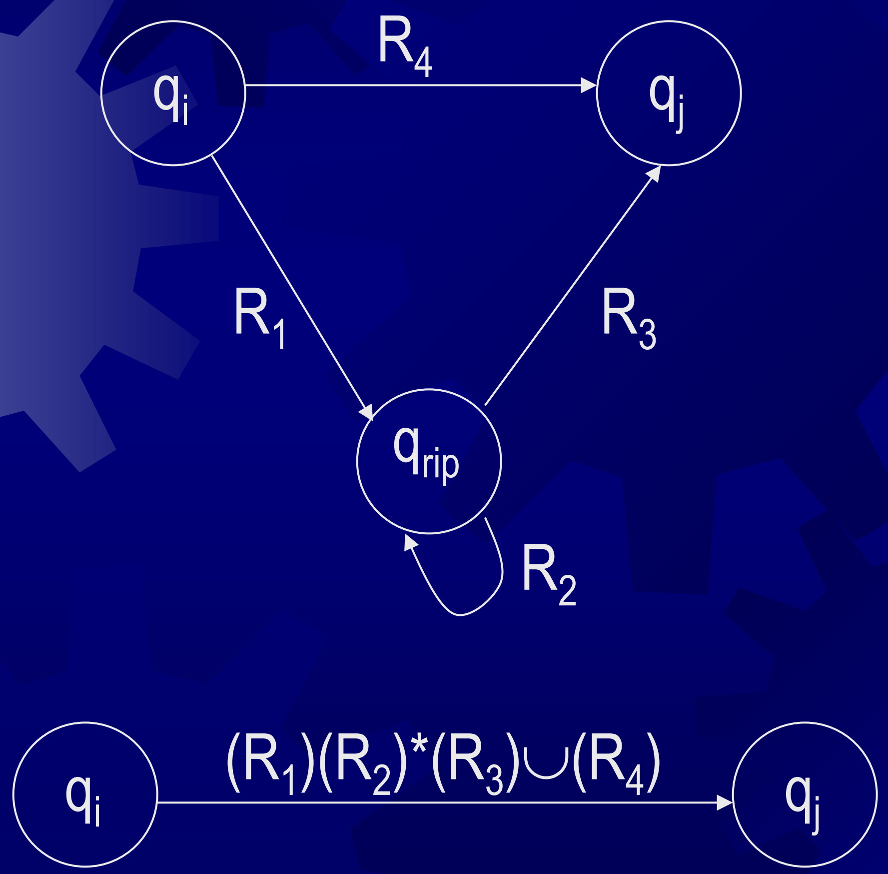


GNFA



从GNFA到正则表达式

- 依次把GNFA状态数减少1个



引理2.32

- 引理2.32: 正则语言可用正则表达式描述.

引理2.32

- 引理2.32: 正则语言可用正则表达式描述.
- 证明:
  - 设正则语言A被DFA M识别.

引理2.32

- 引理2.32: 正则语言可用正则表达式描述.
- 证明:
  - 设正则语言A被DFA M识别.
  - 把M转换成等价的GNFA G.

引理2.32

- 引理2.32: 正则语言可用正则表达式描述.
- 证明:
  - 设正则语言A被DFA M识别.
  - 把M转换成等价的GNFA G.
  - 利用过程CONVERT(G)把G转换成等价的正则表达式.

引理2.32

- 引理2.32: 正则语言可用正则表达式描述.
- 证明:

下面分别给出:

- GNFA的形式定义
- 过程CONVERT(G)
- CONVERT(G)的正确性证明. #

GNFA的形式定义

- GNFA是5元组$(Q,\Sigma,\delta,q_{start},q_{accept})$
  - Q是有穷状态集
  - Σ是输入字母表
  - $\delta:(Q-\{q_{accept}\})\times(Q-\{q_{start}\})\rightarrow R$ 是转移函数 (R是字母表Σ上全体正则表达式的集合)
  - q_{start}是初始状态
  - q_{accept}是接受状态

主要内容

- 正则表达式与自动机等价性
- **非正则语言**
- 泵引理
- 泵引理证明

非正则语言

- 为 $B = \{ 0^n1^n \mid n \geq 0 \}$ 设计DFA？

非正则语言

- 为 $B=\{0^n1^n \mid n\geq 0\}$ 设计DFA？
  - $B=\{0^n1^n \mid n\geq 0\}$是不是正则的？

非正则语言

- 设 $\Sigma=\{0,1\}$
  - $B=\{0^n1^n \mid n\geq 0\}$?

非正则语言

- 设 $\Sigma=\{0,1\}$
  - $B=\{0^n1^n \mid n\geq 0\}$ ？
  - $C=\{w \mid w$中0和1一样多$\}$ ？

非正则语言

- 设 $\Sigma=\{0,1\}$
  - $B=\{0^n1^n \mid n\geq 0\}$ ？
  - $C=\{w \mid w$中0和1一样多$\}$ ？
    - $B=C\cap 0^*1^*$

非正则语言

- 设 $\Sigma=\{0,1\}$
  - $B=\{0^n1^n \mid n\geq 0\}$ ？
  - $C=\{w \mid w$中0和1一样多$\}$ ？
    - $B=C\cap 0^*1^*$
  - $D=\{w \mid w$中子串01和10一样多$\}$ ？

非正则语言

- 设 $\Sigma=\{0,1\}$
  - $B=\{0^n1^n \mid n\geq 0\}$ ？
  - C={w | w中0和1一样多} ？
    - $B=C\cap 0^*1^*$
  - D={w | w中子串01和10一样多 } ？
    - 0100000011111100101110111

非正则语言

- 设 Σ={0,1}
  - $B=\{0^n1^n \mid n\geq 0\}$ ？
  - C={w | w中0和1一样多} ？
    - $B=C\cap 0^*1^*$
  - D={w | w中子串01和10一样多 } ？
    - 0100000011111100101110111
- D是正则的, B和C不是

主要内容

- 正则表达式与自动机等价性
- 非正则语言
- 泵引理
- 泵引理证明

泵引理

- 定理2.37(泵引理): 设A是正则语言,

泵引理

- 定理2.37(泵引理): 设A是正则语言,
则存在常数p(称为泵长度),

泵引理

- 定理2.37(泵引理): 设A是正则语言, 则存在常数p(称为泵长度), 使得若 $s \in A$ 且 $|s| \geq p$, 则

泵引理

- 定理2.37(泵引理): 设A是正则语言, 则存在常数p(称为泵长度), 使得若$s \in A$且$|s| \geq p$, 则$s=xyz$, 并且满足下述条件:

泵引理

- 定理2.37(泵引理): 设A是正则语言, 则存在常数p(称为泵长度), 使得若$s \in A$且$|s| \geq p$, 则$s=xyz$, 并且满足下述条件:

1) $\forall i \geq 0$, $xy^iz \in A$;

泵引理

- 定理2.37(泵引理): 设A是正则语言，则存在常数p(称为泵长度)，使得若$s \in A$且$|s| \geq p$，则$s=xyz$，并且满足下述条件:

1) $\forall i \geq 0$, $xy^iz \in A$;

2) $|y|>0$;

泵引理

- 定理2.37(泵引理): 设A是正则语言, 则存在常数p(称为泵长度), 使得若$s \in A$且$|s| \geq p$, 则$s=xyz$, 并且满足下述条件:

1) $\forall i \geq 0$, $xy^iz \in A$;

2) $|y|>0$;

3) $|xy| \leq p$.

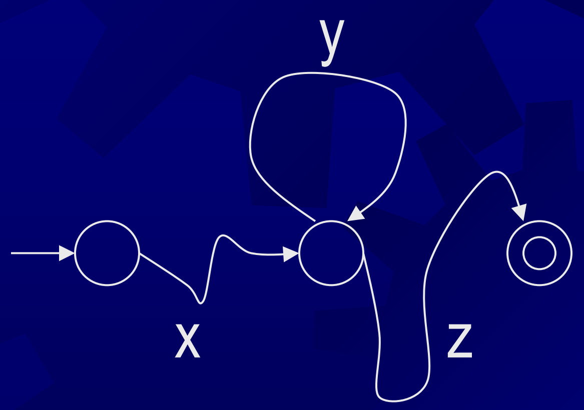
泵引理

- 定理2.37(泵引理): 设A是正则语言, 则存在常数p(称为泵长度), 使得若 $s\in A$ 且 $|s|\geq p$, 则 $s=xyz$, 并且满足下述条件:

1) $\forall i\geq 0,\ xy^iz\in A$;

2) $|y|>0$;

3) $|xy|\leq p$.



主要内容

- 正则表达式与自动机等价性
- 非正则语言
- 泵引理
- 泵引理应用

泵引理用法

- 证明语言$B=\{0^n1^n \mid n\geq 0\}$不是正则的.

泵引理用法

- 证明语言B=$\{0^n1^n \mid n\geq 0\}$不是正则的.
- (反证)
  - 假设B是正则的, 设p是泵长度.

泵引理用法

- 证明语言$B=\{0^n1^n \mid n\geq 0\}$不是正则的.
- (反证)
  - 假设B是正则的, 设p是泵长度.
  - 取$s\in B$, $|s|\geq p$.

泵引理用法

- 证明语言$B=\{0^n1^n \mid n\geq 0\}$不是正则的.
- (反证)
  - 假设B是正则的, 设p是泵长度.
  - 取$s\in B$, $|s|\geq p$.
  - 根据泵引理, $s=xyz$,
    - $\forall i\geq 0$, $xy^iz\in B$.
    - $|xy|\leq p$,
    - $|y|>0$,

泵引理用法

- 证明语言$B=\{0^n1^n \mid n\geq 0\}$不是正则的.
- (反证)
  - 假设B是正则的, 设p是泵长度.
  - 取$s\in B$, $|s|\geq p$.
  - 根据泵引理, $s=xyz$,
    - $\forall i\geq 0$, $xy^iz\in B$.
    - $|xy|\leq p$,
    - $|y|>0$,
  - 证明$\exists i\geq 0$, 使得$xy^iz\notin B$, 矛盾.

例2.38

- $B = \{ 0^n1^n \mid n \geq 0 \}$不是正则的.

例2.38

- $B = \{ 0^n1^n \mid n \geq 0 \}$不是正则的.
- 证明: (反证) 假设B是正则的, 设p是泵长度.

例2.38

- $B = \{ 0^n1^n \mid n \geq 0 \}$不是正则的.
- 证明: (反证) 假设B是正则的, 设p是泵长度. 令$s=0^p1^p$,

例2.38

- $B = \{ 0^n1^n \mid n \geq 0 \}$不是正则的.
- 证明: (反证) 假设B是正则的, 设p是泵长度. 令$s=0^p1^p$, 则$s=xyz$, 对任意$\forall i \geq 0$, $xy^iz \in B$.

例2.38

- $B = \{ 0^n1^n \mid n\geq 0 \}$不是正则的.
- 证明: (反证) 假设B是正则的, 设p是泵长度. 令$s=0^p1^p$, 则 $s=xyz$, 对任意$\forall i\geq 0$, $xy^iz\in B$. 下面分3种情况:

例2.38

- $B = \{ 0^n1^n \mid n\geq 0 \}$不是正则的.
- 证明: (反证) 假设B是正则的, 设p是泵长度. 令$s=0^p1^p$, 则 $s=xyz$, 对任意$\forall i\geq 0$, $xy^iz\in B$. 下面分3种情况:

a) $y\in 0^*0$, 则xy^2z中0比1多, 矛盾;

例2.38

- $B = \{ 0^n1^n \mid n \ge 0 \}$不是正则的.
- 证明: (反证) 假设B是正则的, 设p是泵长度. 令$s=0^p1^p$, 则 $s=xyz$, 对任意$\forall i \ge 0$, $xy^iz \in B$. 下面分3种情况:
  a) $y \in 0^*0$, 则xy^2z中0比1多, 矛盾;
  b) $y \in 1^*1$, 则xy^2z中1比0多, 矛盾;

例2.38

- $B = \{ 0^n1^n \mid n \geq 0 \}$不是正则的.
- 证明: (反证) 假设B是正则的, 设p是泵长度. 令$s=0^p1^p$, 则 $s=xyz$, 对任意$\forall i \geq 0$, $xy^iz \in B$. 下面分3种情况:
  a) $y \in 0^*0$, 则xy^2z中0比1多, 矛盾;
  b) $y \in 1^*1$, 则xy^2z中1比0多, 矛盾;
  c) $y \in 0^*011^*$, 则$xy^2z \notin 0^*1^*$, 矛盾.

  #

例2.39

- C={ w|w中0和1个数相等 }不是正则的.

例2.39

- C={ w|w中0和1个数相等 }不是正则的.
- 证明:(反证) 假设C是正则的.

设p是C的泵长度,令 s=

例2.39

- C={ w|w中0和1个数相等 }不是正则的.
- 证明:(反证) 假设C是正则的.

设p是C的泵长度,令 $s=0^p1^p$.

例2.39

- C={ w|w中0和1个数相等 }不是正则的.
- 证明:(反证) 假设C是正则的.

设p是C的泵长度,令 $s=0^p1^p$.

于是 $s=xyz$, $|xy|\le p$, $y\in 0^*0$, 且对任意i, xy^iz 属于C.

例2.39

- C={ w|w中0和1个数相等 }不是正则的.
- 证明:(反证) 假设C是正则的.

设p是C的泵长度,令 $s=0^p1^p$.

于是 $s=xyz$, $|xy|\le p$, $y\in 0^*0$, 且对任意i, xy^iz 属于C. 但 xy^2z 中 0 的个数比1的个数多,

例2.39

- C={ w|w中0和1个数相等 }不是正则的.
- 证明:(反证) 假设C是正则的.

设p是C的泵长度,令 $s=0^p1^p$.

于是 $s=xyz$, $|xy|\le p$, $y\in 0^*0$, 且对任意i, xy^iz 属于C. 但 xy^2z 中 0 的个数比1的个数多, 所以 xy^2z 不属于C, 矛盾.　#

例2.40

- $F=\{ ww \mid w\in\{0,1\}^* \}$不是正则的.

例2.40

- $F=\{ ww \mid w\in\{0,1\}^* \}$不是正则的.
- 证明:(反证) 假设F是正则的.

  设p是F的泵长度,令s=

例2.40

- $F=\{ ww \mid w\in\{0,1\}^* \}$不是正则的.
- 证明:(反证) 假设F是正则的.

  设p是F的泵长度,令$s=0^p10^p1$.

例2.40

- $F=\{ ww \mid w\in\{0,1\}^* \}$不是正则的.
- 证明:(反证) 假设F是正则的.

  设p是F的泵长度,令$s=0^p10^p1$.

  于是 $s=xyz$, $|xy|\le p$, $y\in 0^*0$,

  且对任意i,xy^iz属于F.

例2.40

- $F=\{ ww \mid w \in \{0,1\}^* \}$不是正则的.
- 证明:(反证) 假设F是正则的.

  设p是F的泵长度,令$s=0^p10^p1$.

  于是 $s=xyz$, $|xy|\le p$, $y\in 0^*0$,

  且对任意i,xy^iz属于F.

  但 xy^2z 不属于F, 矛盾.   #

作业

- 课本 1 .28，1.29（a）（b）