第二章：词法分析

NFA与自动机的最小化

内容介绍

- 非确定有限自动机（NFA）的定义
- NFA到DFA的转换
- 自动机的最小化
- 自动机的化简

1.1 非确定有限自动机的定义

- 非确定有限自动机NFA为一个五元组(Σ,S,S_0,f,Z)，其中：

Σ是一个有穷字母表，它的每个元素称为一个输入字符；

S是状态的集合，它的每个元素称为一个状态；

$S_0 \subseteq S$，是非确定有限自动机的初始状态集；

f是一个从$S\times(\Sigma\cup\{\varepsilon\})$到S 的子集的映射，即 $S\times(\Sigma\cup\{\varepsilon\})\to 2^S$

$Z\subseteq S$，是一个终止状态集，又称为接受状态集

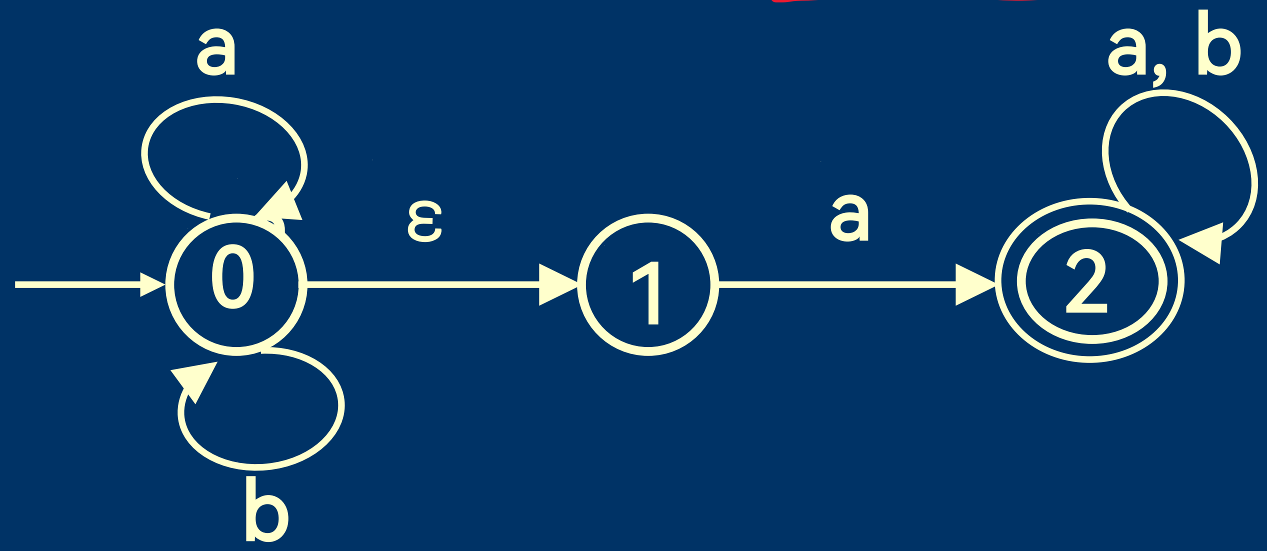
1.2 DFA和NFA的区别

- 总结来看有3点区别

1. 一个状态的不同输出边可以标有相同符号
2. 允许有多个开始状态
3. 允许有空边

1.3 NFA的一些问题

- NFA所能接受的串与DFA的定义是相同的



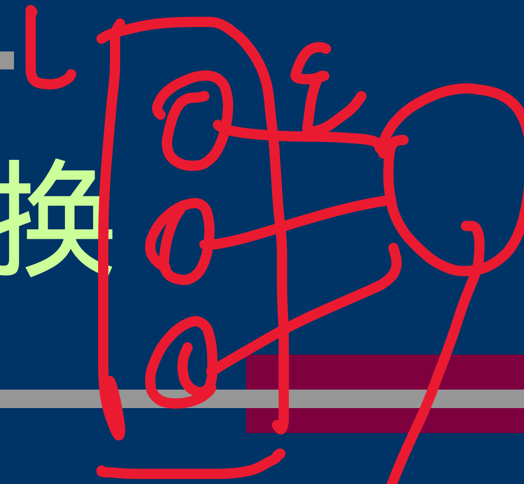
- 实现起来很困难

2.1 自动机等价

- 定义：设A_1和A_2是同一个字母表上的自动机，如果有$L(A_1)=L(A_2)$,则称A_1和A_2等价

- 定理：对于每一个非确定自动机A,存在一个确定自动机A',使得L(A)=L(A')

2.2 NFA到DFA的转换



- 状态集I的ε闭包：设I是NFA M状态集的子集，定义I的ε闭包ε-CLOSURE(I)为：

1. 若q ∈I ,则q ∈ε_CLOSURE(I).
2. 若q ∈I,那么从q出发经任意条ε弧而能到达的任何状态q'都属于 ε-CLOSURE(I).

2.2 NFA到DFA的转换

- 状态集I的a转换：若$I=\{S_1, \ldots, S_m\}$是NFA的状态集的一个子集，$a \in \Sigma$, 则定义:

  $Ia = \varepsilon\text{-CLOSURE}(J)$

其中:

$J = f(S_1,a) \cup f(S_2,a) \ldots \cup f(S_m,a)$

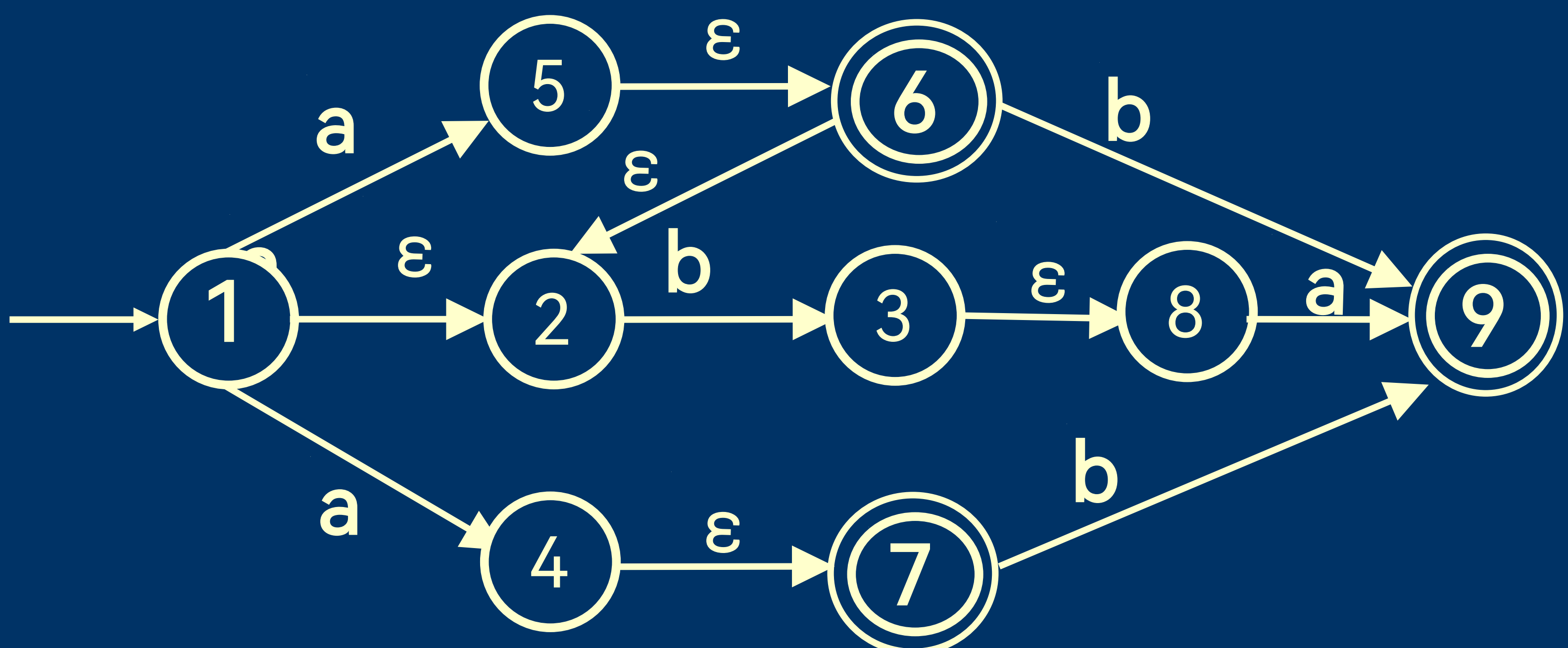
2.2 NFA到DFA的转换

- 已知 A： NFA, 构造 A':DFA

1. 令A'的初始状态为I_0'=ε_CLOSURE({S_1,S_2,…S_k}), 其中S_1…S_k是A的全部初始状态。
2. 若I={S_1,…,S_m}是A'的一个状态，a∈∑，则定义f'(I, a)=Ia将Ia加入S'，重复该过程，直到S'不产生新状态。
3. 若I'={S_1,…,S_n}是A'的一个状态,且存在一个S_i是A的终止状态，则令I'为A'的终止状态。

2.2 NFA到DFA的转换

例子:



2.2 NFA到DFA的转换

输入字 状态	a	b
+{1,2}	{2,4,5,6,7}	{3,8}
-{2,4,5,6,7}	{}	{3,8,9}
{3,8}	{9}	{}
-{3,8,9}	{9}	{}
-{9}	{}	{}

2.2 NFA到DFA的转换

- 转化后的结果

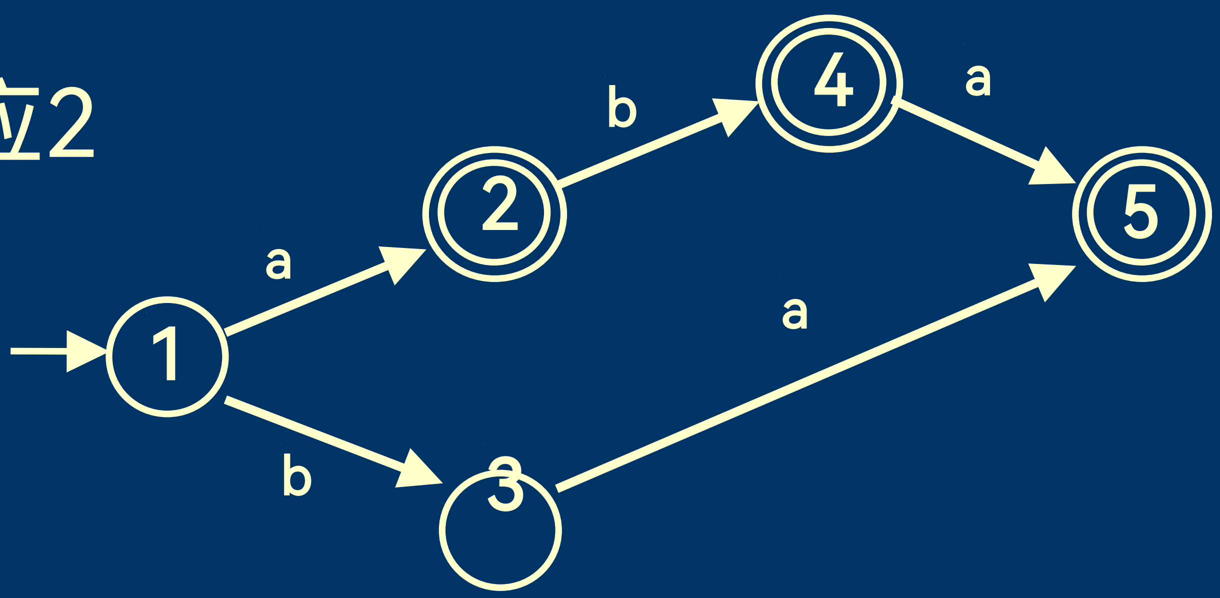
{1,2}对应1

{2,4,5,6,7}对应2

{3,8}对应3

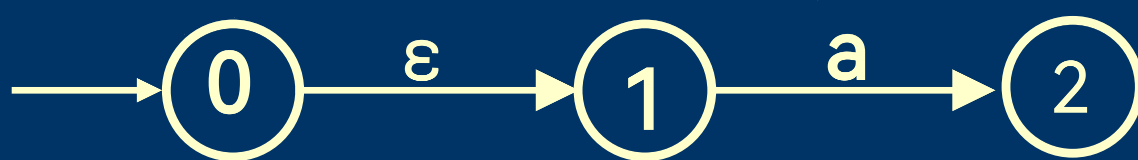
{3,8,9}对应4

{9}对应5



2.2 NFA到DFA的转换

- 另外一种消除空边的转换方式



- 删去空边，增加0到2的边

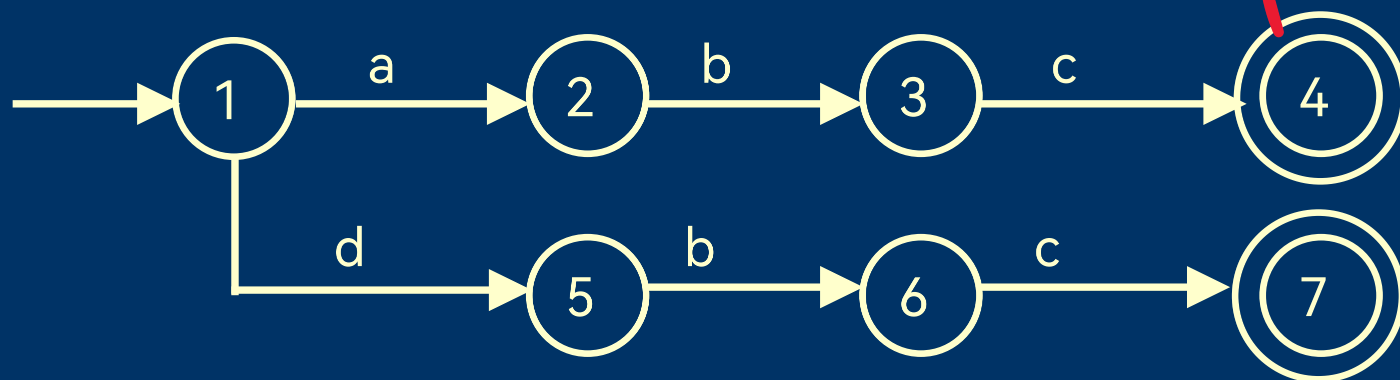
- ε环路的时候，就把这几个状态合成一个

2.3 自动机的最小化

- 定义：　等价状态

　设DFA M 的两个状态S_1和S_2，如果对任意输入的符号串x，从S_1和S_2出发，总是同时到达接受状态或拒绝状态中，则称S_1和S_2是等价的。

结局相同.



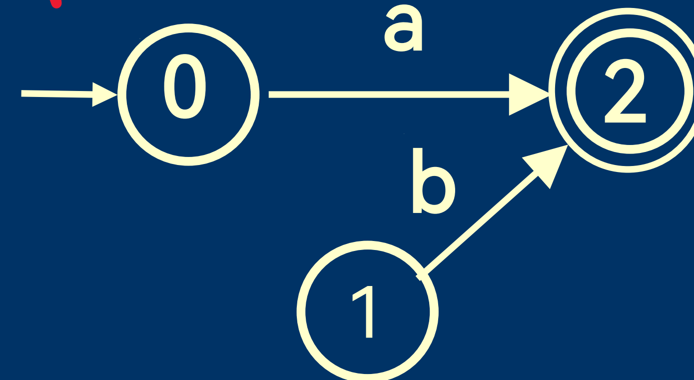
2.3 自动机的最小化

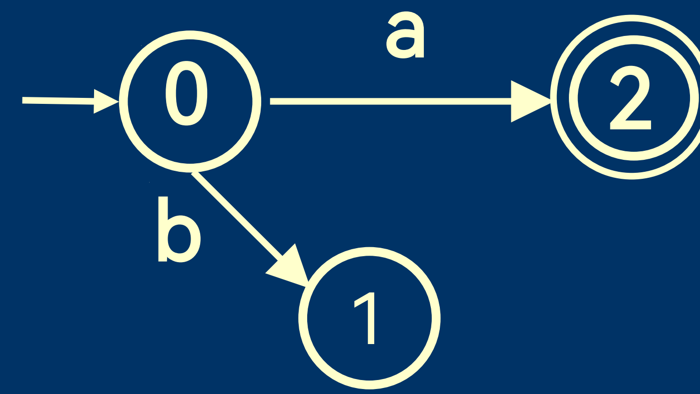
- 定义：无关状态 从来到不了的

  设S是DFA M的一个状态，若：

1. 从开始状态无到S的通路，或
2. S到任意终止状态无通路

则称S为M的无关状态





2.3 自动机的最小化

- 定义：最小(最简)自动机
  如果DFA M 没有无关状态，也没有等价状态，则称M 为最小自动机
- 结论：任一DFA都可以化为最简自动机，即任一DFA M都存在DFA M'，使得 L(M)=L(M')，且M'是最简自动机

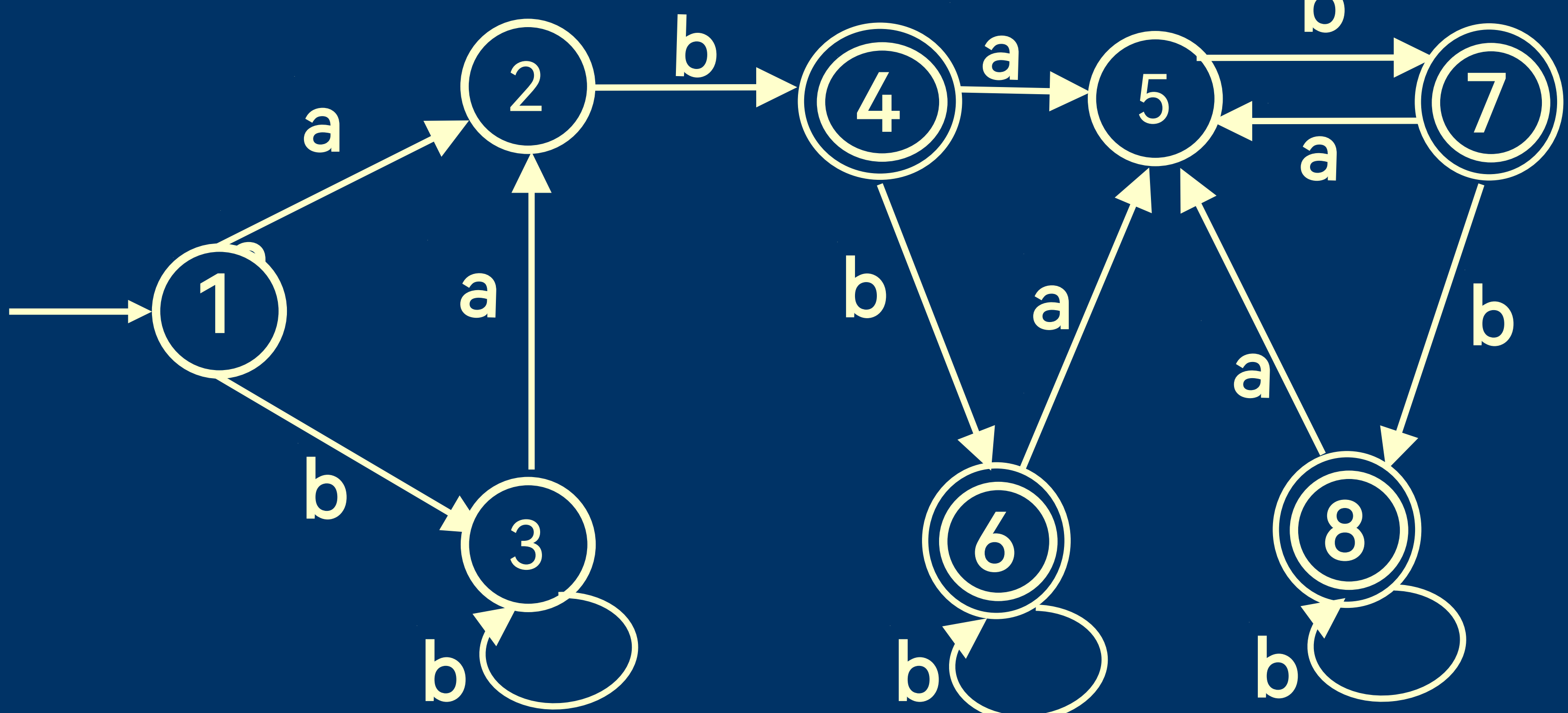
2.4 自动机的化简

- 状态分离法

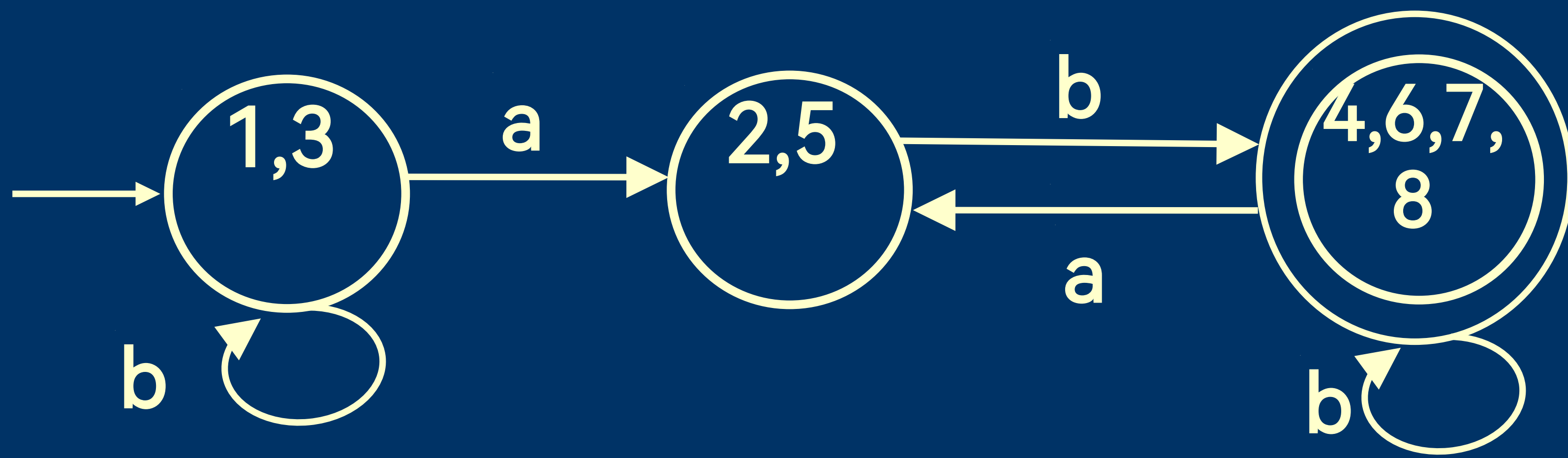
1. 终止状态为一组，非终止状态为一组
2. 对每一组进行分离，若每组中的元素映射到不同的组，则表示他们不等价，就可以划分出来。
3. 重复2，知道没有新组产生，此时每组中的状态都为等价状态。

2.4 自动机的化简

- 例1

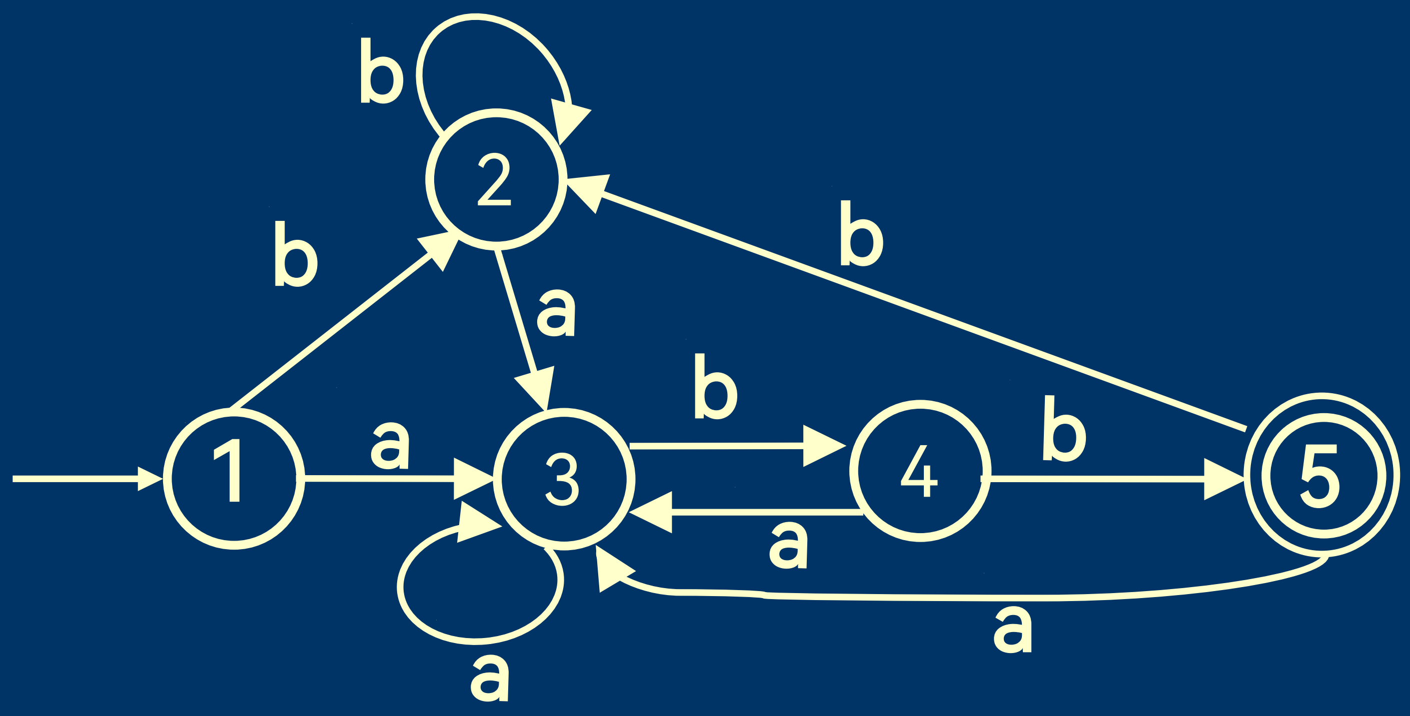


2.4 自动机的化简

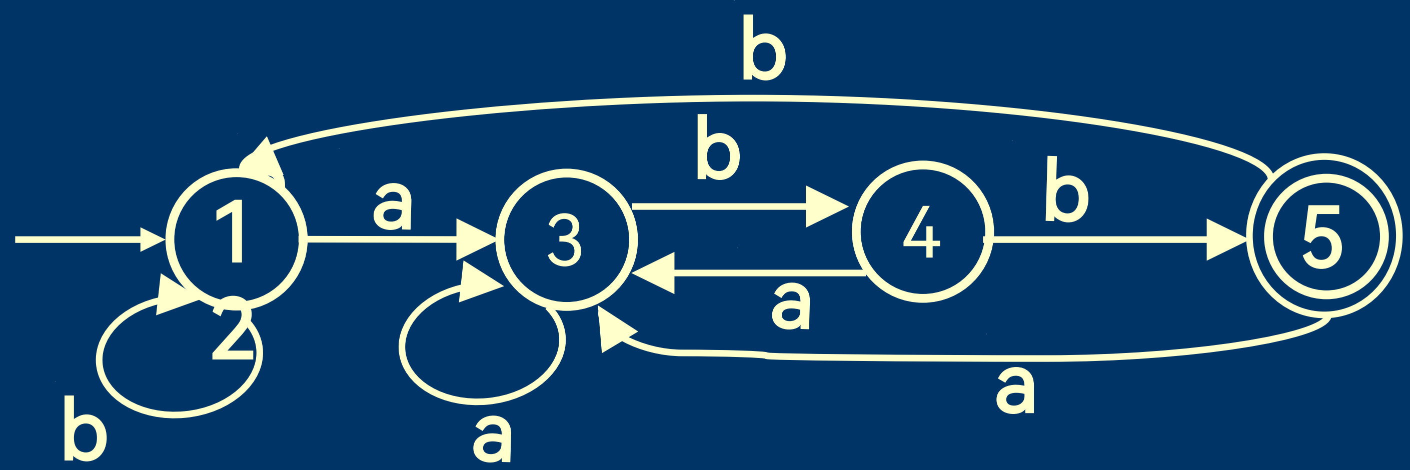


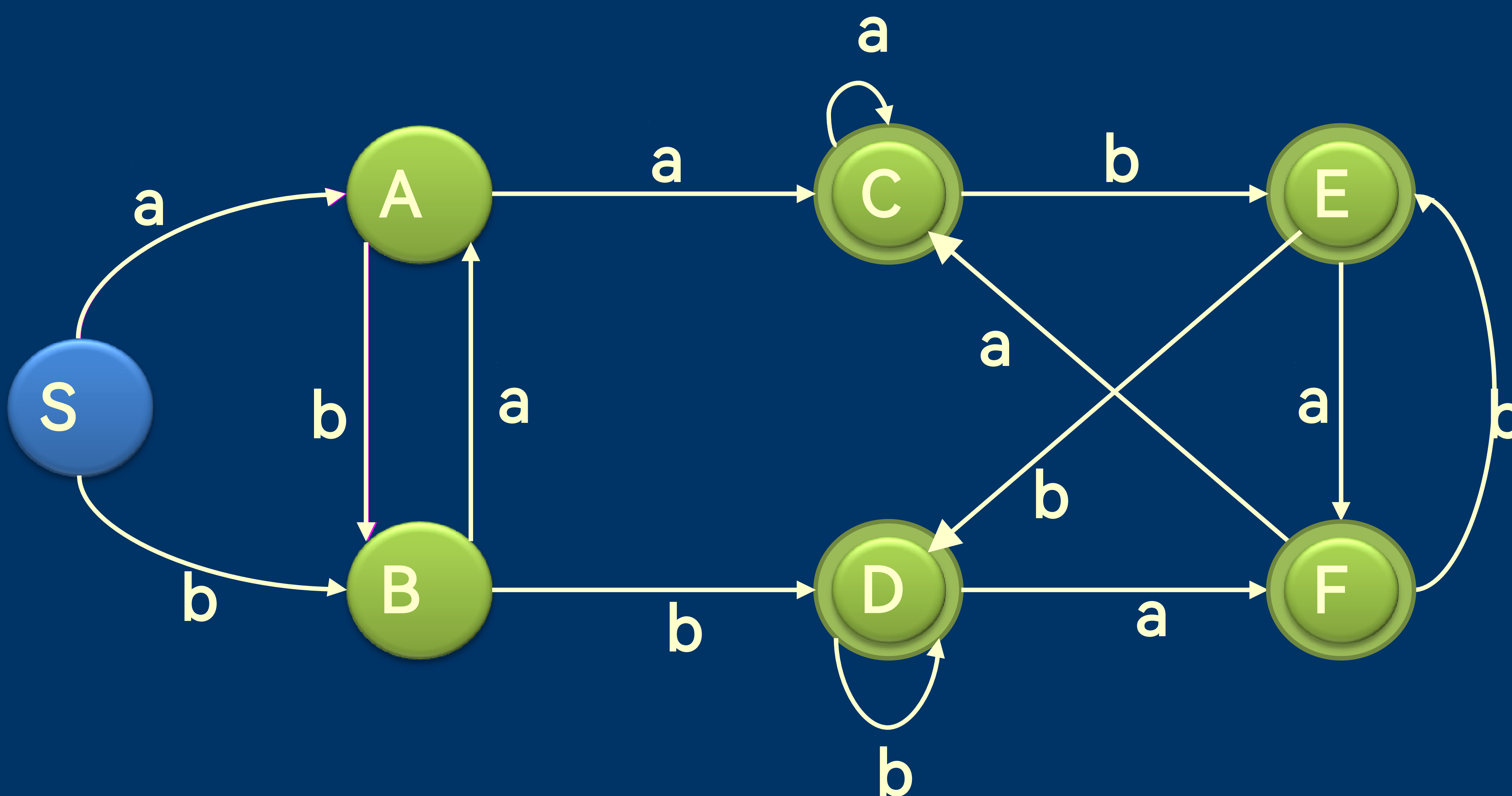
2.4 自动机的化简

- 例2



2.4 自动机的化简





S
A
B
C
D
E
F
a
b
a
b
a
b
a
b
a
a
b
b
a
b