§2.6 从正规式到NFA

一.从正规式构造NFA

\*正规式描述语言的能力与NFA﹑DFA识别语言的能力是一样的.

\*从任一正规式r可以构造识别L(r)的NFA.

1. 算法: 从正规式构造NFA(Thompson构造)

输入: 字母表Σ上的正规式r

输出: 接受L(r)的NFA N

根据正规式r的结构,归纳地构造

1. 对ε构造NFA(见书 P101，P159)

\*正确性证明

(2) 对任意a∈Σ, 构造NFA(见书 P101,P159)

1. 如果N(s)和N(t)是正规式s和t的NFA,
   1. 对s | t , 构造N(s | t) (见书 P101,P160)

\*正确性证明

* 1. 对st , 构造NFA(st) (见书 P101, P160)
  2. 对正规式s\*, 构造NFA(s\*) (见书 P102, P161)
  3. 对正规式( s ), N( s )即是它的NFA

1. 性质:

(1) N(r)的状态数最多是r中符号(operands)和算符(operators)数的两倍;

(2) N(r)只有一个开始状态和一个接受状态,接受状态没有向外的转换边, 开始状态没有进入的转换边.

(3) N(r)的每个状态,除接受状态外,或者有一个Σ中符号标识的向外转换边,或者最多有两个向外的ε转换边.

1. 例子: 正规式r = (a | b)\*abb的NFA

\*r的分析树(parse tree)(见书 P102, P162)

\*由r构造的NFA(见书 P98, P155), 构造过程见书P102―103.

§2.7 DFA的化简

一.定理: 每个正规集都可由一个状态数最小的DFA识别,并且这个DFA唯一.

假设当前的DFA M的每一个状态对每个输入符号都有转换(transition). 必要时,引入死状态(dead state).

二. 极小化的原理

1.我们说串w区别(distinguish)状态s和t,如果DFA M从状态s出发,输入w, 它停在某个接受状态,但从t出发,同样的输入, 它停在非接受状态.

例子: (见书 P99(P96), P156(P151))

\*状态A和B由bb区别;

\*但状态A和C不能由任何串区别.

2.算法思想:

极小化DFA状态数的算法就是把DFA的状态分成一些不相交的子集,每一子集中的状态都是不可区别的,不同子集的状态都是可区别的.每个子集合并成一个状态.

三. 算法: 极小化DFA的状态数

输入:一个DFA M, 它的状态集合是S,输入符号集Σ,转换函数δ: S×Σ→S, 开始状态s0, 接受状态集F .

输出: 一个DFA M’, 它和M接受同样的语言, 且状态数最少.

算法:

1.构造状态集合的初始划分(initial partition)Π, 分成两组,接受状态组F和非接受状态组S―F.

2.应用下面的过程对Π构造新的划分Πnew

for Π中的每个组(group)G do

begin

把G划分成小组,G的两个状态s和t在同一个小组中,当且仅当对所有的输入符号a, s和t的a转换是到Π的同一组中;

在Πnew中,用G的划分代替G;

end;

3.如果Πnew=Π,令Πfinal := Π, 再执行步骤(4),

否则令Π:=Πnew ,转(2)

4.在Πfinal中的每个状态组中选一个状态代表它,这些代表(representative)是简化的DFA M’的状态.如果s是这样一个代表,在DFA M中,s的a转换到t,并且t所在的组的代表是r,那么在M’中,s的a转换到r.包含s0的状态组的代表是M’的开始状态,M’的接受状态是那些在F中的代表.

5.如果M’有死状态,则去掉它.从开始状态不可及的状态也删除.从任何其它状态到死状态的转换都成为无定义(undefined).

四. 例子:上例中DFA最小化. (见书 P116, P183)

\*阅读: 教材 3.7.4, 3.9.6 节

作业3:

1.给定以下NFA N(见黑板),给出N处理输入串ababbab的移动序列, 并将该NFA化为等价的DFA.

2. 构造下列正规式的NFA, DFA和最小DFA

(a | b)\*a(a | b) .

3. 假设L和M都是正规集. 证明: L∪M, L∩M和(补集)也是正规集.(提示: 根据L和M的DFA,构造L∩M和的DFA).