同濟大學
TONGJI UNIVERSITY
地址：中国上海市四平路1239号　邮编：200092
1239 SIPING ROAD SHANGHAI CHINA 200092
电话（TEL）：+86 21-　传真（FAX）：+86 21-
网址（WEB）：www.tongji.edu.cn

重点题型：

✓1° 基本的文法，FA的构造　Ch2
✓2° NFA转DFA　Ch3
✓3° RE，RL，FA的等价转换构造　Ch4　（重点，大概率要考）
✓4° FA与RG互转构造　Ch5　（重点，大概率要考）
✓5° 正则语言泵引理　Ch5
✓6° 自动机的等价性与最小化算法　Ch5　（M-N，填表算法）
✓7° 上下文无关文法 语法树构造　Ch6　（看作业）
✓8° 下推自动机的2种构造方法　Ch6
✓9° CFG转PDA算法　Ch6　（重点！必说要考）
✓10° 上下文无关文法化简、GNF构造　Ch7　（重点！必说要考）
✓11° 上下文无关文法泵引理　Ch7

复习策略：

① 重点题型，总结归纳解题步骤，配合作业题目，务必搞懂！！！
② 概念中的重点内容，在打印课件中荧光标注，反复记忆
务必熟背上述内容

要求：

6.21 24:00 前完成题型归纳整理
6.22 10:00 考前，强化记忆反复复习

地址：中国上海市四平路1239号　　邮编：200092
1239 SIPING ROAD SHANGHAI CHINA 200092
电话（TEL）：+86 21-　　　　传真（FAX）：+86 21-
网址（WEB）：www.tongji.edu.cn

10° 上下文无关文法化简方法

顺序
- 消除 ε-产生式
- 消除单一产生式
- 消除不可产生符号
- 消除不可达符号

这2步不能换序

Step 1. 消除 ε-产生式方法

* 形如 $A \to \varepsilon$，是 ε-产生式
* 文法 $G=(V,T,P,S)$ 中任意变量 A，若有 $A \Rightarrow^* \varepsilon$，则 A 为可空变量
* 例. 若有 $A \Rightarrow CBC \Rightarrow BC \Rightarrow C \Rightarrow \varepsilon$，则 ABC 全是可空变量

消除 ε-产生式的方法.

① 求可空变量集 U

② 对于每一产生式 $A \to H_1H_2H_3 \cdots H_m$：

* 删去 $H_i$ 中属于 U 的变量，代替 $A \to H_1H_2 \cdots H_m$

若 $H_1 \sim H_m$ 全属于 U，不动，保留 $A \to H_1H_2 \cdots H_m$

Step 2 消除单一产生式

① 确定单元对、若有 $A \to B$，则 [A,B] 是单元对，

若 [A,B] [B,C] 是单元对，[A,C] 是单元对

② 删除全部形如 $A \to B$ 的单元产生式

对于每一单元对 [A,B]，将 B 复制给 A

eg. $\begin{cases} S \to A \mid B \mid 0S1 \\ A \to 0A \mid 0 \\ B \to 1B \mid 1 \end{cases} \xrightarrow[{[S,B]}]{[S,A]} \begin{cases} S \to 0A \mid 0 \mid 1 \mid 0S1 \mid 1B \\ A \to 0A \mid 0 \\ B \to 1B \mid 1 \end{cases}$

同濟大學
TONGJI UNIVERSITY
地址：中国上海市四平路1239号　邮编：200092
1239 SIPING ROAD SHANGHAI CHINA 200092
电话（TEL）：+86 21-　传真（FAX）：+86 21-
网址（WEB）：www.tongji.edu.cn

Step 3. 消除非可产生符号

~~能从 $S \to \alpha$ 推出的 $\alpha$ 中所有符号可产生~~

~~eg. $\begin{cases} S \to AB \mid a \\ A \to b \end{cases} \Longrightarrow$~~

① 每个终结符可产生

② 若 $A \to \alpha \in P$ 且 $\alpha$ 中符号都可产生，则 $A$ 可产生

eg. $\begin{cases} S \to AB \mid a \\ A \to b \end{cases} \xRightarrow{B\text{不可产生}} \begin{cases} S \to a \\ A \to b \end{cases}$

Step 4. 可达符号集，消除非可达

① 起始变元 S 可达

② 若 $A \to \alpha$ 且 $A$ 可达，$\alpha$ 中所有符号可达

eg. $\begin{cases} S \to AB \mid a \\ A \to b \end{cases} \xrightarrow{\text{消除非可产生}} \begin{cases} S \to a \\ A \to b \end{cases} \xrightarrow{\text{消除非可达}} S \to a$

同濟大學
TONGJI UNIVERSITY
地址：中国上海市四平路1239号　邮编：200092
1239 SIPING ROAD SHANGHAI CHINA 200092
电话（TEL）：+86 21-　传真（FAX）：+86 21-
网址（WEB）：www.tongji.edu.cn

在上述文法化简基础上求CNF：

引入新变量 $X \to \alpha_1 | \alpha_2 | \cdots | \alpha_n$ （$\alpha_i$ 是终结字符）

用X代替多元变量，例 $X \to AB$

最终使文法仅有只有 $A \to BC$，$A \to a$ 2种格式

同濟大學
TONGJI UNIVERSITY
地址：中国上海市四平路1239号　邮编：200092
1239 SIPING ROAD SHANGHAI CHINA 200092
电话（TEL）：+86 21-　传真（FAX）：+86 21-
网址（WEB）：www.tongji.edu.cn

格雷德巴赫范式GNF：

若CFG $G=(V,T,P,S)$ 中所有产生式满足 $A\to a\alpha$
$(A\in V,\ \alpha\in V^*,\ a\in T)$

则称G为格雷德巴赫范式文法，GNF

$\Rightarrow$ { 右线性文法是特殊的GNF
GNF派生长度n的串要n步
略

CFG转GNF方法：

课件不全，实际方法很复杂很变态。看书 P113~115

① 消除间接左递归
② 消除直接左递归
③ 重新代换

菜鸡wsy总结：

① 消除间接递归

用 $A_1, A_2, \cdots, A_m$ 代替m个变元后：

1° 对于所有 $A_i \to A_j\alpha$，通过不断替换，使 $i \le j$

eg：

$$\begin{cases} A \to BC \\ B \to CA \mid b \\ C \to AB \mid a \end{cases} \xrightarrow[A=A_1]{B=A_2,\ C=A_3} \begin{cases} A_1 \to A_2A_3 \\ A_2 \to A_3A_1 \mid b \\ \underline{A_3 \to A_1A_2 \mid a} \end{cases}$$

$A_3 \to A_1A_2$ 不满足

$$A_3 \to A_1A_2 \to A_2A_3A_2 \xrightarrow{\text{代入所有}} A_3A_1A_3A_2 \mid bA_3A_2 \qquad (j=i)$$

② 消除直接左递归 $(A \to A\alpha)$

若 $j=i$，再增加一个状态

$$A_3 \to A_3A_1A_3A_2 \mid bA_3A_2$$

引入规则：

$$\begin{cases} A \to A\alpha_1 \mid A\alpha_2 \mid \cdots \mid A\alpha_r \\ A \to \beta_1 \mid \beta_2 \mid \cdots \mid \beta_s \end{cases}$$

背！

$$\Leftrightarrow \begin{cases} A \to \beta_i \\ A \to \beta_i B & (i=1,2,\cdots,s) \\ B \to \alpha_j \\ B \to \alpha_j B & (j=1,2,\cdots,r) \end{cases}$$

同濟大學
TONGJI UNIVERSITY
地址：中国上海市四平路1239号 邮编：200092
1239 SIPING ROAD SHANGHAI CHINA 200092
电话（TEL）：+86 21- 传真（FAX）：+86 21-
网址（WEB）：www.tongji.edu.cn

~~(2) 消除直接递归~~

(3) 变量名称换回来

略

地址：中国上海市四平路1239号　　邮编：200092
1239 SIPING ROAD SHANGHAI CHINA 200092
电话（TEL）：+86 21-　　　　传真（FAX）：+86 21-
网址（WEB）：www.tongji.edu.cn

11、上下文无关文法泵引理

对于任意的上下文无关语言L，存在泵长p，满足：

$\forall\ s\in L,\ |s|\geq p,\ \exists u,v,x,y,z\ \ s.t.\ \ s=uvxyz$

且 $\begin{cases} uv^ixy^iz\in L\ (i\geq 0) \\ |vy|\geq 1 \\ |vxy|\leq p \end{cases}$　$\Downarrow$ $uxz\in L$

看课件、书本、作业中的证明思路

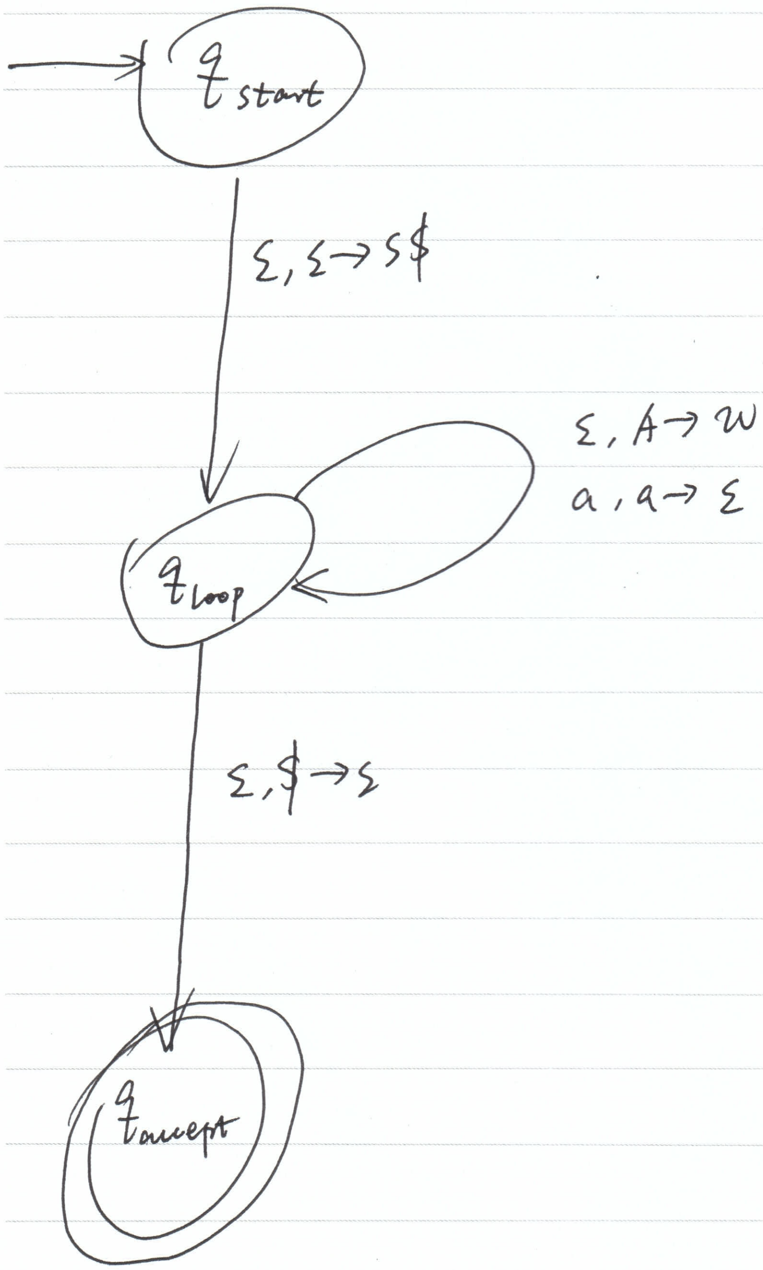
主要是利用泵长p，分段，讨论 v x y 三部分构成情况，

寻找矛盾！（主要是指数之长度规律）

考前看2本黑皮书章节题相关题型

同濟大學
TONGJI UNIVERSITY
地址：中国上海市四平路1239号　　邮编：200092
1239 SIPING ROAD SHANGHAI CHINA 200092
电话（TEL）：+86 21-　　　　　传真（FAX）：+86 21-
网址（WEB）：www.tongji.edu.cn

9° CFG转PDA




for $A \to w$　没必要细分了

for terminal $a$

8° 下推自动机构造方法

下推自动机 $P=(Q,\Sigma,q_0,\delta,F,T,Z_0)$

$T$：栈字符集，$F$：接受状态集，$Z_0$：初始栈底符号

$\delta: Q\times(\Sigma\cup\{\varepsilon\})\times T \to 2^{Q\times T^*}$ 转移函数

~~终态接受PDA~~

~~下推自动机 $P=(Q,\Sigma,T,\delta,q_0,F,Z_0)$~~

~~空栈接受PDA~~

终态接受PDA：PDA $P=(Q,\Sigma,q_0,\delta,F,T,Z_0)$

语言 $L(P)$ 定义：$w\in L(P)$ 当且仅当 $\exists q\in F, \alpha\in T^*$ s.t. $(q_0,w,Z_0)\vdash^*(q,\varepsilon,\alpha)$

空栈接受PDA：PDA $P=(Q,\Sigma,q_0,\delta,F,T,Z_0)$

语言 $N(P)$ 定义：$w\in N(P)$ 当且仅当 $\exists q\in Q$, s.t. $(q_0,w,Z_0)\vdash^*(q,\varepsilon,\varepsilon)$

看书本和作业例子

地址：中国上海市四平路1239号　邮编：200092
1239 SIPING ROAD SHANGHAI CHINA 200092
电话（TEL）：+86 21-　传真（FAX）：+86 21-
网址（WEB）：www.tongji.edu.cn

1° 上下文无关文法语法树构造

掌握叶子边缘

最左/右推导(派生)

结合作业复习即可，不难。略

同濟大學
TONGJI UNIVERSITY
地址：中国上海市四平路1239号　　邮编：200092
1239 SIPING ROAD SHANGHAI CHINA 200092
电话（TEL）：+86 21-　　传真（FAX）：+86 21-
网址（WEB）：www.tongji.edu.cn

1° 基本文法. FA构造

略. 直接复习作业

2° NFA转DFA

法一. 定理法

通过从$S_0$开始计算~~[illegible]~~，至状态数不再增加为止，见Ch2 / P41
很好理解

法二. 子集构造法

看作业答案

本质上与由法一一样（结果一样）

同濟大學
TONGJI UNIVERSITY
地址：中国上海市四平路1239号 邮编：200092
1239 SIPING ROAD SHANGHAI CHINA 200092
电话（TEL）：+86 21- 传真（FAX）：+86 21-
网址（WEB）：www.tongji.edu.cn

3° RE，RL，FA的等价转换构造

(1) ① 构造正则表达式的一般规则

① 语言只含字符串本身

② 语言含多个字符串连接得到的串（连接）

③ 语言含零次或多次出现的串（闭包）

④ 语言中的串或子串有多种可能形式（+，并）

⑤ 语言~~中的子~~的串中含有子符子元的子串：用ε和该子串"+"后放进总串

(2) 正则表达式转NFA：

步步分解，详细看懂 Ch4 / P24 例题

(3) FA转RE的2种方法

A. GNFA法

① 加首尾 $q_{start} \xrightarrow{\varepsilon} q_1$ $\quad q_j \xrightarrow{\varepsilon} q_{accept}$（所有终态都加）

② 补空边 $q_i \xrightarrow{\emptyset} q_j$

③ 并括号 $q_i \overset{1}{\underset{0}{\rightleftarrows}} q_j \Longrightarrow q_i \xrightarrow{0+1} q_j$

④ 移除中间态：

移除 $q_{rip} = Q - \{q_{start}\} - \{q_{accept}\}$

求新的转移表达式：

$$\delta'(q_i, q_j) = \delta(q_i, q_j) + \delta(q_i, q_{rip})(\delta(rip, rip))^* \delta(q_{rip}, q_j)$$

对每一 $q_i \in Q - \{q_{rip}\} - \{q_{accept}\}$

$q_j \in Q - \{q_{rip}\} - \{q_{start}\}$

同济大学
TONGJI UNIVERSITY
地址：中国上海市四平路1239号　邮编：200092
1239 SIPING ROAD SHANGHAI CHINA 200092
电话（TEL）：+86 21-　传真（FAX）：+86 21-
网址（WEB）：www.tongji.edu.cn

B. $R_{ij}^{k}$ 迭代法

设 $R_{ij}^{k} = \{ x \mid \delta(q_i, x) = q_j$，且中间不经过序号大于 $k$ 的状态，但 $i, j$ 可以大于 $k$，$x \in \Sigma^* \}$

由递归可得：

$$\begin{cases} R_{ij}^{k} = R_{ij}^{k-1} + R_{ik}^{k-1} (R_{kk}^{k-1})^* R_{kj}^{k-1} & k \geq 1 \\ R_{ij}^{0} = \{ a \mid a \in \Sigma, \delta(q_i, a) = q_j \} & i \neq j \\ R_{ij}^{0} = \{ a \mid a \in \Sigma, \delta(q_i, a) = q_j \} \underline{\underline{\cup \{\varepsilon\}}} & i = j \end{cases}$$

① 对于有 n 个态的 M，

$$L(M) = \bigcup_{q_f \in F} R_{1f}^{n}$$

（从一到所有终态的最大标号不超过 n 的并）

② 列表求 $r_{ij}^{k}$ 即可

4° FA与RG的互转构造

定理：

DFA $M=(Q,\Sigma,\delta,q_0,F)$

有 RG $G=(Q,\Sigma,P,q_0)$

$$P=\{q\to ap \mid \delta(q,a)=p\}\cup\{q\to a \mid \delta(q,a)=p,\ p\in F\}$$

例题：课件 Ch4 /P7　(DFA 转 RG)

Ch4 /P12, P13　(RG 转 FA)

~~不~~含ε-转移的情况

5° 正则语言的泵引理

泵引理：对于每一正则语言L，存在泵长 $p$

$\forall s \in L, |s| \geq p$. 则 $\exists x, y, z$, s.t $s = xyz$ 且

$$\begin{cases} xy^iz \in L & (i \geq 0) \\ |y| > 0 \\ |xy| \leq p \end{cases}$$

扩充泵引理：

设L是一正则语言，则：

存在只依赖于L的正整数 $k$，$\forall x, y, z$ $(xyz \in L)$

只要 $|y| \geq k$，就有 $y = uvw$ $(v \neq \varepsilon, |uv| \leq k)$

使得 $xuv^iwz \in L$ 对 $i \geq 0$ 都成立

同濟大學
TONGJI UNIVERSITY
地址：中国上海市四平路1239号　邮编：200092
1239 SIPING ROAD SHANGHAI CHINA 200092
电话（TEL）：+86 21-　传真（FAX）：+86 21-
网址（WEB）：www.tongji.edu.cn

6° DFA的等价性与最小化

右不变等价关系、指数、等价类的划分　　概念见课件

M—N定理：

(1) $L \subseteq \Sigma^*$ 是RL

(2) L是$\Sigma^*$上某一个具有有穷指数的右不变等价关系R的某些等价类划分

(3) $R_L$具有有穷指数

$$\left|\Sigma^*/R_{L(M)}\right| \le \left|\Sigma^*/R_M\right| \le |Q|$$

同构关系，DFA最小化唯一

最小化DFA的算法：

定义：可区分的状态对

DFA $M=(Q,\Sigma,\delta,q_0,F)$

若对 $x\in\Sigma^*$，对于Q中两状态q, p

有 $\delta(q,x)$ 和 $\delta(p,x)$ 有且只有一个属于F

则称p, q可区分，~~记作 p≡q~~ 否则，称p q等价，记作 $p\equiv q$

步骤：

① 为所有(p,q) (p,q∈Q)画表，格子空白

② 对所有 $p\in F, q\notin F$ 的状态对(p,q)，其格子打叉，表示p q可区分

③ 重复以下操作，直到表中内容不再改变为止：

对于一个未被标记的状态对(p,q)，若存在 $a\in\Sigma$

$(r=\delta(p,a),\ s=\delta(q,a))$ 已经被标为×，则(p,q)也可以标叉

④ 完成后 所有未被标记的状态对(p,q)都是等价的，p,q可以合并