CFLs GOOD FOR DESCRIBING INFINITE SET OF STRINGS IN FINITE WAY

SYNTAX OF PROGRAMMING LANGUAGE
WELL FORMED ARITHMETIC EXPRESSIONS
WELL-NESTED BEGIN-END BLOCKS
STRINGS OF BALANCED PARENTHESES, ETC.

ALL REGULAR SETS ARE CFLs.

NON REGULAR EXAMPLES

$\{0^n 1^n \mid n \geq 0\}$

{PALINDROMES}

{BALANCED STRING OF PARENTHESES}

NOT ALL SETS ARE CFLs $\{a^n b^n a^n \mid n \geq 0\}$

CONTEXT-FREE GRAMMAR CFG

$$G = (\overset{V}{N}, \Sigma, \overset{R}{P}, S)$$ quadruple.

WHERE

N IS A FINITE SET (NON TERMINAL SYMBOLS) — VARIABLES

Σ IS A FINITE SET (TERMINAL SYMBOLS) DISJOINT FROM N

$$N \cap \Sigma = \emptyset$$

P IS A FINITE SUBSET OF $\underset{\text{VARIABLE}}{N} \times \underset{\text{STRING OF VARIABLES AND TERMINALS}}{(N \cup \Sigma)^*}$ (PRODUCTIONS) — RULES

$S \in N$ START SYMBOL (START VARIABLE

$A, B, C, \ldots$ FOR NONTERMINALS

$a, b, c, \ldots$ FOR TERMINALS

$\alpha, \beta, \gamma, \ldots$ STRINGS IN $(N \cup \Sigma)^*$

$A \to \alpha$ FOR (A, α)

$A \to \alpha_1$

$A \to \alpha_2$

$A \to \alpha_3$

$$A \to \alpha_1 \mid \alpha_2 \mid \alpha_3$$

IF $\alpha, \beta \in (N \cup \Sigma)^*$

$\alpha \xrightarrow[G]{1} \beta$ $\qquad\qquad$ $\alpha \Rightarrow \beta$ YIELDS

AND β DERIVABLE FROM α IN ONE STEP. $\longrightarrow$ EXTRA 'INSERT' PAGE

THIS IS A RELATION $\subseteq (N \cup \Sigma)^* \times (N \cup \Sigma)^*$

$\xrightarrow[G]{*}$ REFLEXIVE TRANSITIVE CLOSURE OF $\xrightarrow[G]{1}$

$\alpha \overset{*}{\Rightarrow} \beta$

THAT IS DEFINE

$\alpha \xrightarrow[G]{0} \alpha$ FOR ALL α

$\alpha \xrightarrow[G]{m+1} \beta$ IF THERE EXISTS γ S.T. $\alpha \xrightarrow[G]{m} \gamma$ AND $\gamma \xrightarrow[G]{1} \beta$

$\alpha \xrightarrow[G]{*} \beta$ IF $\alpha \xrightarrow[G]{m} \beta$ FOR SOME $m \geq 0$

STRING DERIVABLE FROM START SYMBOL

IF $S \xrightarrow[G]{*} x$ $\quad$ x IS A SENTENTIAL FORM

SENTENCE IF ONLY CONSISTS OF TERMINAL SYMBOLS

I.E. IF IT IS IN Σ^*

$L(G) = \{ x \in \Sigma^* \mid S \xrightarrow[G]{*} x \}$ LANGUAGE GENERATED BY G.

SUBSET $A \subseteq \Sigma^*$ IS A CONTEXT-FREE LANG. CFL

IF $A = L(G)$ FOR SOME CFG G.

"INSERT"

β CAN BE OBTAINED FROM α BY REPLACING SOME OCCURRENCE OF A NON-TERMINAL ~~AND~~ SAY A IN α WITH γ, WHERE $A \to \gamma \in P$. THAT IS IF THERE EXISTS $\alpha_1, \alpha_2 \in (N \cup \Sigma)^*$ AND PRODUCTION $A \to \gamma$ SUCH THAT $\alpha = \alpha_1 A \alpha_2$ $\beta = \alpha_1 \gamma \alpha_2$.

ABREVIATED FORM FOR CFGs

WRITE ONLY THE RULES / PRODUCTIONS

THE VARIABLES/NONTERMINALS ARE THE THINGS TO THE LEFT OF THE $\rightarrow$

THE REMAINING SYMBOLS ARE THE TERMINALS

THE START VARIABLE/NONTERMINAL IS THE LEFT SYMBOL OF THE FIRST RULE.

CFG EXAMPLES

SET OF EVEN LENGTH STRINGS WITH TWO MIDDLE SYMBOLS EQUAL.

$S \to aSa \mid aSb \mid bSa \mid bSb \mid aa \mid bb$

SET OF ODD LENGTH STRINGS WHOSE FIRST, MIDDLE AND LAST SYMBOLS ARE THE SAME.

$S \to aTa \mid bUb$

$T \to aTa \mid aTb \mid bTa \mid bTb \mid a$ ODD $\bar{c}$ a IN MID.

$U \to aUa \mid aUb \mid bUa \mid bUb \mid b$ ODD $\bar{c}$ b IN MID

SET OF STRINGS WITH EQUAL # OF a's & b's

$S \to \epsilon \mid aSb \mid bSa \mid SS$

ALSO

$S \to aB \mid bA \mid \epsilon$

$A \to aS \mid bAA$

$B \to bS \mid aBB$

Σ^* GENERATED BY CFG

$S \rightarrow \epsilon \mid 0 \mid 1 \mid 0S \mid 1S$

ALSO

$S \rightarrow \epsilon \mid L \mid LS$

$L \rightarrow 0 \mid 1$

ALSO

$S \rightarrow \epsilon \mid 0S \mid 1S$ SIMPLEST.

CFG EXAMPLES

EVEN LENGTH PALINDROMES OVER $\{a, b\}$

$$S \to aSa \mid bSb \mid \epsilon$$

ODD LENGTH PALINDROMES OVER $\{a, b\}$

$$S \to aSa \mid bSb \mid a \mid b$$

THE SET OF STRINGS OVER $\{a, b\}$ SUCH THAT x^r IS x WITH THE as AS bs AND THE bs AS as

$$S \to aSb \mid bSa \mid \epsilon$$

THE SET OF EVEN LENGTH STRINGS OVER $\{a, b\}$

$$S \to aT \mid bT \mid \epsilon$$

$$T \to aS \mid bS$$

THE SET OF ODD LENGTH STRINGS WITH MIDDLE SYMBOL a

$$S \to aSa \mid aSb \mid bSa \mid bSb \mid a$$

EXAMPLES

$\{a^n b^n \mid n \geq 0\}$ IS A CFL

$$S \to aSb \mid \epsilon$$

$G = (N, \Sigma, P, S\}$ WHERE

$N = \{S\}$

$\Sigma = \{a, b\}$

$P = \{S \to aSb, \; S \to \epsilon\}$

$S = S$

DERIVATION OF $a^3 b^3$ IN G

$$S \xrightarrow{1} aSb \xrightarrow{1} aaSbb \xrightarrow{1} aaaSbbb \xrightarrow{1} aaabbb$$

$$S \xrightarrow{4} aaabbb$$

PALINDROMES

$$\{ x \in \{a,b\}^* \mid x = x^R \}$$

$$S \rightarrow aSa \mid bSb \mid a \mid b \mid \epsilon$$

BALANCED PARENTHESES

$$S \rightarrow [S] \mid SS \mid \epsilon$$

EXAMPLE $\{ a^n b^n \mid n \geq 0 \}$ IS A CFL

$G = (N, \Sigma, P, S)$ WHERE

$N = \{ S \}$

$\Sigma = \{ a, b \}$

$P = \{ S \rightarrow aSb,\ S \rightarrow \epsilon \}$

ABREVIATED FORM

$S \rightarrow aSb \mid \epsilon$

DERIVATION OF $a^3 b^3$ IN G.

$$S \xrightarrow[G]{1} aSb \xrightarrow[G]{1} aaSbb \xrightarrow[G]{1} aaaSbbb \xrightarrow[G]{1} aaabbb$$

SO $S \xrightarrow[G]{4} aaabbb$ DERIVABLE FROM S IN 4 STEPS.

CAN SHOW BY INDUCTION ON n THAT

?<~~DON'T NEED INDUCTION~~>

$$S \xrightarrow[G]{n+1} a^n b^n$$

SO ALL STRINGS OF THE FORM $a^n b^n$ ARE IN $L(G)$.

CLOSURE PROPERTIES OF CFLs. (MORE DETAIL IN FOLLOWING PAGES.)

CFLs CLOSED UNDER UNION

GIVEN G_1 G_2 G' s.t. $L(G') = L(G_1) \cup L(G_2)$

$S' \to \cancel{G_1}S_1$

$S' \to \cancel{G_2}S_2$

$R' = R_1 \cup R_2 \cup \{ S' \to S_1 , S' \to S_2 \}$

RENAME ANY DUPLICATE ~~SYMBOL~~ VARIABLES.

CFLs CLOSED UNDER CONCATENATION

~~SAME~~

$R' = R_1 \cup R_2 \cup \{ S' \to S_1 S_2 \}$

CFLs CLOSED UNDER KLEENE STAR.

GIVEN G_1

$R' = R_1 \cup \{ S' \to S_1 S' , S' \to \epsilon \}$

SOME CLOSURE PROPERTIES

$G_1 = (N_1, \Sigma, P_1, S_1)$ $L_1 = L(G_1)$

$G_2 = (N_2, \Sigma, P_2, S_2)$ $L_2 = L(G_2)$

UNION

$G_{L_1 \cup L_2} = (N_{L_1 \cup L_2}, \Sigma, P_{L_1 \cup L_2}, S_{L_1 \cup L_2})$

$L_1 \cup L_2 = L(G_{L_1 \cup L_2})$

ASSUME $N_1 \cap N_2 = \emptyset$

$N_{L_1 \cup L_2} = N_1 \cup N_2 \cup \{S_{L_1 \cup L_2}\}$

$P_{L_1 \cup L_2} = P_1 \cup P_2 \cup \{S_{L_1 \cup L_2} \rightarrow S_1, S_{L_1 \cup L_2} \rightarrow S_2\}$

CONCATENATION

$$G_{L_1L_2} = (N_{L_1L_2}, \Sigma, P_{L_1L_2}, S_{L_1L_2})$$

$$L_1L_2 = L(G_{L_1L_2})$$

$$N_{L_1L_2} = N_1 \cup N_2 \cup \{S_{L_1L_2}\}$$

$$P_{L_1L_2} = P_1 \cup P_2 \cup \{S_{L_1L_2} \to S_1 S_2\}$$

2.2a

CFLs NOT CLOSED UNDER INTERSECTION

$$\underbrace{\{a^m b^m c^n \mid m,n \geq 0\}}_{\text{CFL}} \cap \underbrace{\{a^m b^n c^n \mid m,n \geq 0\}}_{\text{CFL}} = \underbrace{\{a^n b^n c^n \mid n \geq 0\}}_{\text{NOT CFL}}$$

S = AB
A = aAb | ε
B = cB | ε

CFL ∩ RL is CFL

PROOF: A PRODUCT LIKE CONSTRUCTION BETWEEN NPDA AND DFA

$$\underbrace{a^m b^m}_{A}\ \underbrace{c^n}_{B}$$

$$\underbrace{a^m}_{A}\ \underbrace{b^n c^n}_{B}$$

S = AB
A = aA | ε
B = bBc | ε

KLEENE STAR

$$G_{L_1^*} = (N_{L_1^*} , \Sigma , P_{L_1^*} , S_{L_1^*})$$

$$L_1^* = L(G_{L_1^*})$$

$$N_{L_1^*} = N_1 \cup \{ S_{L_1^*} \}$$

$$P_{L_1^*} = P_1 \cup \{ S_{L_1^*} \rightarrow S_1 S_{L_1^*} , S_{L_1^*} \rightarrow \epsilon \}$$

SPECIAL TYPES OF CFGs

PRODUCTIONS RESTRICTED TO

RIGHT LINEAR $\quad A \to xB$ OR $A \to x$

STRONGLY RIGHT LINEAR $\quad A \to aB$ OR $A \to \varepsilon$

LEFT LINEAR $\quad A \to Bx$ OR $A \to x$

STRONGLY LEFT LINEAR $\quad A \to Ba$ OR $A \to \varepsilon$

$x \in \Sigma^*$ $\quad a \in \Sigma$

EACH GENERATES EXACTLY THE REGULAR SETS.

GIVEN DFA CAN CONSTRUCT A STRONGLY RIGHT LINEAR GRAMMAR THAT WILL GENERATE THE LANGUAGE THE DFA RECOGNIZES

$$V = \{ V_i \mid q_i \in Q \}$$

$$\Sigma = \Sigma$$

FOR EACH

$\delta(q_i, a) \rightarrow q_j$ WHERE $a \in \Sigma$

MAKE RULE / PRODUCTION

$$V_i \rightarrow a V_j$$

FOR EACH

$q_i \in F$

MAKE RULE / PRODUCTION

$$V_i \rightarrow \varepsilon$$

$S = V_0$ WHERE q_0 = START STATE.

DFA TO CFG

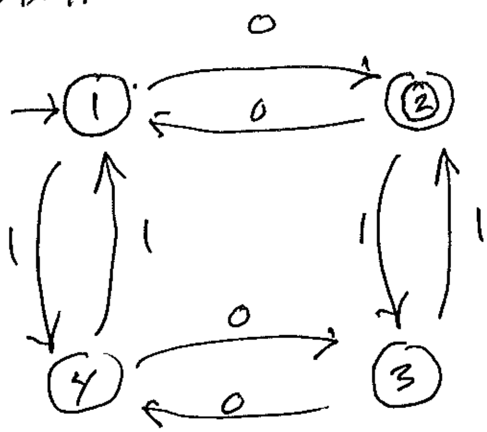
V_i FOR EACH q_i OF DFA

RULE $V_i \rightarrow a V_j$ IF $\delta(q_i, a) \rightarrow q_j$

RULE $V_i \rightarrow \epsilon$ IF $q_i \in F$

MAKE V_0 START VARIABLE WHERE q_0 IS START STATE.

EXAMPLE



ODD # 0s

EVEN # 1s

$V_1 \to 0 V_2$

$V_1 \to 1 V_4$

$V_2 \to 0 V_1$

$V_2 \to 1 V_3$

$V_3 \to 0 V_4$

$V_3 \to 1 V_2$

$V_4 \to 0 V_3$

$V_4 \to 1 V_1$

$V_2 \to \epsilon$

$V_1 \to 0 V_2 \mid 1 V_4$

$V_2 \to 0 V_1 \mid 1 V_3 \mid \epsilon$

$V_3 \to 0 V_4 \mid 1 V_2$

$V_4 \to 0 V_3 \mid 1 V_1$

GIVEN A STRONGLY RIGHT LINEAR GRAMMAR G CAN CONSTRUCT A DFA THAT WILL ~~RECOGNIZE~~ ACCEPT ANY STRING G CAN GENERATE.

N.F.A.

~~WRITTEN~~

$Q = V$

$\Sigma = \Sigma$

$\delta : Q \times \Sigma \to Q$

$\delta(V_i, a) \to \{V_j\}$ iff $V_i \to aV_j \in P$

$s = S$

$F = \{V_i \in Q \mid V_i \to \epsilon \in P\}$

CHOMSKY NORMAL FORM

ALL PRODUCTIONS ARE OF THE FORM

$$A \rightarrow BC \quad \text{OR} \quad A \rightarrow a$$

WHERE $A, B, C \in V$ AND $a \in \Sigma$

GREIBACH NORMAL FORM

ALL PRODUCTIONS ARE OF THE FORM

$$A \rightarrow a B_1 B_2 B_3 \ldots B_k \quad \text{FOR SOME } k \geq 0$$

WHERE $A, B_1, B_2, B_3, \ldots, B_k \in V$ AND $a \in \Sigma$

EXAMPLES: STRINGS OF BALANCED PARENTHESES

$$S \rightarrow [S] \mid SS \mid \varepsilon$$

CNF $\quad S \rightarrow AB \mid AC \mid SS, \; C \rightarrow SB, \; A \rightarrow [, \; B \rightarrow]$

GNF $\quad S \rightarrow [B \mid [SB \mid [BS \mid [SBS, \quad B \rightarrow]$

THM FOR ANY CFG G, THERE IS A CFG G' IN CNF AND A CFG G'' IN GNF SUCH THAT

$$L(G'') = L(G') = L(G) - \{\varepsilon\}$$

NOTE NO CNF OR GNF G CAN GENERATE ε.

ELIMINATING ε AND UNIT PRODUCTIONS FROM CFGs.

ε PRODUCTIONS $\quad A \to \varepsilon$ FOR SOME $A \in V$

UNIT PRODUCTIONS $\quad A \to B$ FOR SOME $A, B \in V$

THM FOR ANY CFG $G = (V, \Sigma, R, S)$, THERE IS A CFG G' WITH NO ε OR UNIT PRODUCTIONS SUCH THAT $L(G') = L(G) - \{\varepsilon\}$

PROOF LET $\hat{R}$ BE THE SMALLEST SET OF PRODUCTIONS CONTAINING R THAT IS CLOSED UNDER

1. IF $A \to \alpha B \beta$ AND $B \to \varepsilon$ ARE IN $\hat{R}$, THEN $A \to \alpha\beta$ IS IN $\hat{R}$.
2. IF $A \to B$ AND $B \to \gamma$ ARE IN $\hat{R}$, THEN $A \to \gamma$ IS IN $\hat{R}$.

CAN CONSTRUCT $\hat{R}$ RECURSIVELY FROM R

NOTE EACH ADDITIONS RIGHT SIDE IS A SUBSTRING OF A RIGHT SIDE IN R $\therefore$ ONLY FINITE NUMBER OF RULES WILL BE ADDED.

NOW LET $\hat{G}$ BE THE GRAMMAR

$$\hat{G} = (V, \Sigma, \hat{R}, S)$$

THEN EVERY DERIVATION OF G IS A DERIVATION OF $\hat{G}$ SINCE $R \subseteq \hat{R}$.

THUS $L(G) \subseteq L(\hat{G})$

BUT $L(\hat{G}) \subseteq L(G)$ SINCE ANY RULE IN $\hat{G}$ THAT IS NOT IN G CAN BE SIMULATED BY TWO RULES IN G.

THEREFORE $L(\hat{G}) = L(G)$

FOR ALL $x \neq \epsilon$ $x \in \Sigma^*$ A DERIVATION $S \underset{\hat{G}}{\overset{*}{\Rightarrow}} x$ IS OF MINIMAL LENGTH IMPLIES IT DOES NOT USE ANY ϵ OR UNIT PRODUCTIONS/RULES.

FIRST NO ϵ PRODUCTIONS

TOWARD CONTRADICTION ASSUME $x \neq \epsilon$ AND $S \underset{\hat{G}}{\overset{*}{\Rightarrow}} x$ IS OF MINIMAL LENGTH AND USES SOME ϵ PRODUCTION, I.E. A RULE OF FORM $B \rightarrow \epsilon$ IS USED AT SOME POINT. SAY

$$S \underset{\hat{G}}{\overset{*}{\Rightarrow}} \gamma B \delta \underset{\hat{G}}{\overset{1}{\Rightarrow}} \gamma\delta \underset{\hat{G}}{\overset{*}{\Rightarrow}} x$$

NOTE $x \neq \epsilon$ IMPLIES AT LEAST ONE OF γ OR $\delta \neq \epsilon$.

AND B OCCURED IN SOME PRODUCTION $A \rightarrow \alpha B \beta$

SO WE HAVE FOR SOME $m, M, K \geq 0$

$$S \underset{\hat{G}}{\overset{m}{\Rightarrow}} \eta A \theta \underset{\hat{G}}{\overset{1}{\Rightarrow}} \underbrace{\eta \alpha B \beta \theta \underset{\hat{G}}{\overset{M}{\Rightarrow}} \gamma B \delta}\ \underset{\hat{G}}{\overset{1}{\Rightarrow}} \gamma\delta \underset{\hat{G}}{\overset{K}{\Rightarrow}} x \qquad m+1+M+1+K$$

BUT BY CLOSURE PROPERTY 1, $A \rightarrow \alpha\beta$ IS ALSO IN $\hat{R}$

$$S \underset{\hat{G}}{\overset{m}{\Rightarrow}} \eta A \theta \underset{\hat{G}}{\overset{1}{\Rightarrow}} \underbrace{\eta \alpha \beta \theta \underset{\hat{G}}{\overset{M}{\Rightarrow}} \gamma\delta}\ \underset{\hat{G}}{\overset{K}{\Rightarrow}} x$$

$m+1+M+K$ STRICKLY LESS THAN

CONTRADICTION $\frac{1}{2}$ Q.E.D.

SECOND NO UNIT PRODUCTIONS.

TOWARD A CONTRADICTION ASSUME $x \neq \varepsilon$ AND $S \overset{*}{\underset{\hat{G}}{\Rightarrow}} x$ IS OF MINIMAL LENGTH AND USES SOME UNIT PRODUCTION.

SAY $A \rightarrow B$ IS USED AT SOME POINT.

$$S \overset{*}{\underset{\hat{G}}{\Rightarrow}} \alpha A \beta \overset{1}{\underset{\hat{G}}{\Rightarrow}} \alpha B \beta \overset{*}{\underset{\hat{G}}{\Rightarrow}} x$$

SOME PLACE LATTER MUST HAVE $B \rightarrow \gamma$

$$S \overset{m}{\underset{\hat{G}}{\Rightarrow}} \alpha A \beta \overset{1}{\underset{\hat{G}}{\Rightarrow}} \underbrace{\alpha B \beta \overset{n}{\underset{\hat{G}}{\Rightarrow}} \eta B \theta} \overset{1}{\underset{\hat{G}}{\Rightarrow}} \eta \gamma \theta \overset{k}{\underset{\hat{G}}{\Rightarrow}} x \qquad m+1+n+1+k$$

BUT BY CLOSURE PROPERTY 2. $A \rightarrow \gamma$ IS ALSO IN $\hat{R}$.

$$S \overset{m}{\underset{\hat{G}}{\Rightarrow}} \alpha A \beta \overset{1}{\underset{\hat{G}}{\Rightarrow}} \underbrace{\alpha \gamma \beta \overset{n}{\underset{\hat{G}}{\Rightarrow}} \eta \gamma \theta} \overset{k}{\underset{\hat{G}}{\Rightarrow}} x \qquad m+1+n+k$$

STRICKLY LESS THAN

CONTRADICTION.

Q.E.D.

SO WE CAN GENERATE G' WITHOUT ε OR unit PRODUCTIONS FROM G s.t. $L(G') = L(G) - \varepsilon$

GETTING TO CHOMSKY NORMAL FORM AFTER REMOVAL OF ϵ AND UNIT PRODUCTIONS.

FOR EACH TERMINAL $a \in \Sigma$

ADD NEW VARIABLE V_a

ADD NEW RULE $V_a \to a$

REPLACE ALL a's ~~on~~ IN OLD RULES, EXCEPT OF THE FORM $B \to a$, WITH V_a

NOW ALL PRODUCTIONS /RULES ARE OF THE FORM

$$A \to a \quad \text{OR} \quad A \to B_1 B_2 B_3 \ldots B_k \quad \geq 2$$

AND THE B_i ARE ALL VARIABLES.

SAME SET OF TERMINAL STRINGS ARE GENERATED BUT TAKES ONE MORE STEP.

NOW DO THE FOLLOWING UNTIL ALL PRODUCTIONS HAVE RIGHT HAND SIDES OF LENGTH 2 OR LESS.

$A \to B_1 B_2 \ldots B_k$ replace with

$A \to B_1 C$, $C \to B_2 B_3 \ldots B_k$

EXAMPLES

$\{a^n b^n \mid n \geq 0\} - \{\varepsilon\}$ $\quad\quad$ $\{a^n b^n \mid n \geq 1\}$

$S \to aSb \mid \varepsilon$

ADD $S \to ab$ RULE 1.
REMOVE ε PRODUCTION

$S \to aSb \mid ab$

FOR a AND b ADD VARIABLES A, B

AND PRODUCTIONS $A \to a$ $\quad B \to b$

$S \to A\underbrace{SB} \mid AB$, $A \to a$, $B \to b$

ADD $C \to SB$

$S \to AC \mid AB$, $C \to SB$, $A \to a$, $B \to b$

EXAMPLE: STRINGS OF BALANCED PARENTHESES

$$S \to [S] \mid SS \mid \varepsilon$$

ADD $S \to [\,]$ RULE 1.

REMOVING ε PRODUCTION

$$S \to [S] \mid SS \mid [\,]$$

FOR [AND] ADD $A \to [$ AND $B \to]$

$$S \to ASB \mid SS \mid AB, \quad A \to [, \quad B \to]$$

ADD $C \to SB$

$$S \to AC \mid SS \mid AB, \quad C \to SB, \quad A \to [, \quad B \to]$$

Example 2.10 in Sipser (2cd).

$$S \to ASA|aB \tag{1}$$
$$A \to B|S \tag{2}$$
$$B \to b|\epsilon \tag{3}$$

$A \to B$ and $B \to b$, so add $A \to b$.

$$S \to ASA|aB \tag{4}$$
$$A \to B|S|b \tag{5}$$
$$B \to b|\epsilon \tag{6}$$

$A \to S$ and $S \to ASA$, so add $A \to ASA$.

$$S \to ASA|aB \tag{7}$$
$$A \to B|S|b|ASA \tag{8}$$
$$B \to b|\epsilon \tag{9}$$

$A \to S$ and $S \to aB$, so add $A \to aB$.

$$S \to ASA|aB \tag{10}$$
$$A \to B|S|b|ASA|aB \tag{11}$$
$$B \to b|\epsilon \tag{12}$$

$A \to B$ and $B \to \epsilon$, so add $A \to \epsilon$.

$$S \to ASA|aB \tag{13}$$
$$A \to B|S|b|ASA|aB|\epsilon \tag{14}$$
$$B \to b|\epsilon \tag{15}$$

$S \to ASA$ and $A \to \epsilon$, so add $S \to AS|SA$.

$$S \to ASA|aB|AS|SA \tag{16}$$
$$A \to B|S|b|ASA|aB|\epsilon \tag{17}$$
$$B \to b|\epsilon \tag{18}$$

$S \to aB$ and $B \to \epsilon$, so add $S \to a$.

$$S \to ASA|aB|AS|SA|a \tag{19}$$
$$A \to B|S|b|ASA|aB|\epsilon \tag{20}$$
$$B \to b|\epsilon \tag{21}$$

$A \to ASA$ and $A \to \epsilon$, so add $A \to AS|SA$.

$$S \to ASA|aB|AS|SA|a \tag{22}$$
$$A \to B|S|b|ASA|aB|\epsilon|AS|SA \tag{23}$$
$$B \to b|\epsilon \tag{24}$$

$A \to aB$ and $B \to \epsilon$, so add $A \to a$.

$$S \to ASA|aB|AS|SA|a \tag{25}$$
$$A \to B|S|b|ASA|aB|\epsilon|A\underset{S}{\cancel{A}}|SA|a \tag{26}$$
$$B \to b|\epsilon \tag{27}$$

Now closed so drop all ϵ and unit productions.

$$S \to ASA|aB|AS|SA|a \tag{28}$$
$$A \to b|ASA|aB|AS|SA|a \tag{29}$$
$$B \to b \tag{30}$$

Add variables for terminals in strings of two or more symbols.

$$S \to ASA|V_aB|AS|SA|a \tag{31}$$
$$A \to b|ASA|V_aB|AS|SA|a \tag{32}$$
$$B \to b \tag{33}$$
$$V_a \to a \tag{34}$$

Reduce remaining productions to no more than two variables each.

$$S \to AC|V_aB|AS|SA|a \tag{35}$$
$$A \to b|AC|V_aB|AS|SA|a \tag{36}$$
$$B \to b \tag{37}$$
$$V_a \to a \tag{38}$$
$$C \to SA \tag{39}$$