REDUCED CFG

$G = (\{S, A, B\}, \{a\}, R, S)$

R: $S \to AB$

$A \to a$

$B \to \lambda$

NULLABLE = {B}

REMOVE λ-RULES

$S \to A \mid AB$

$A \to a$

$B \to$

CHAINS

CHAIN(S) = {S, A}

CHAIN(A) = {A}

CHAIN(B) = {B}

REMOVE UNIT-RULES

$S \to a \mid AB$

$A \to a$

$B \to$

CHAINS

CHAIN(S) = {S}

CHAIN(A) = {A}

CHAIN(B) = {B}

REMOVE UNIT-RULES

$S \to AB$

$A \to a$

$B \to \lambda$

NULLABLE = {B}

REMOVE λ-RULES

$S \to A \mid AB$

$A \to a$

$B \to$

OOPS! ORDER DOES MATTER

DO λ-REMOVAL THEN UNIT-REMOVAL

NOW HAVE $G = (\{S, A, B\}, \{a\}, R, S)$

R: $S \to a \mid AB$
$A \to a$
$B \to$

PRODUCTIVE = $\{S, A\}$
REMOVE UNPRODUCTIVE
$S \to a$
$A \to a$

USEFUL = $\{S, a\}$
REMOVE UNREACHABLE
$S \to a$

USEFUL = $\{S, A, B, a\}$
REMOVE UNREACHABLE
$S \to a \mid AB$
$A \to a$
$B \to$

PRODUCTIVE = $\{S, A\}$
REMOVE UNPRODUCTIVE
$S \to a$
$A \to a$

OOPS

ORDER DOES MATTER

DO UNPRODUCTIVE-REMOVAL
THEN UNREACHABLE-REMOVAL

CNF

START WITH REDUCED CFG

IF $|RHS| = 1$ THEN $RHS \in \Sigma$

$$A \to a$$

SO IN CNF

IF $|RHS| > 1$ THEN CAN BE MIX OF CHARACTERS FROM $V \cup \Sigma$

CHANGE ~~ALL~~ EACH SYMBOL IN Σ, SAY a, TO $\langle a \rangle$ AND ADD $\langle a \rangle$ TO V, PLUS ADD RULE

$$\langle a \rangle \to a$$

NOW IF $|RHS| > 1$, IT IS OVER V^+

ITERATIVELY CHANGE A $|RHS| > 2$ LIKE

$$A \to B_1 B_2 \cdots B_{k+1}$$

TO

$$A \to B_1 \langle B_2 \cdots B_{k+1} \rangle$$

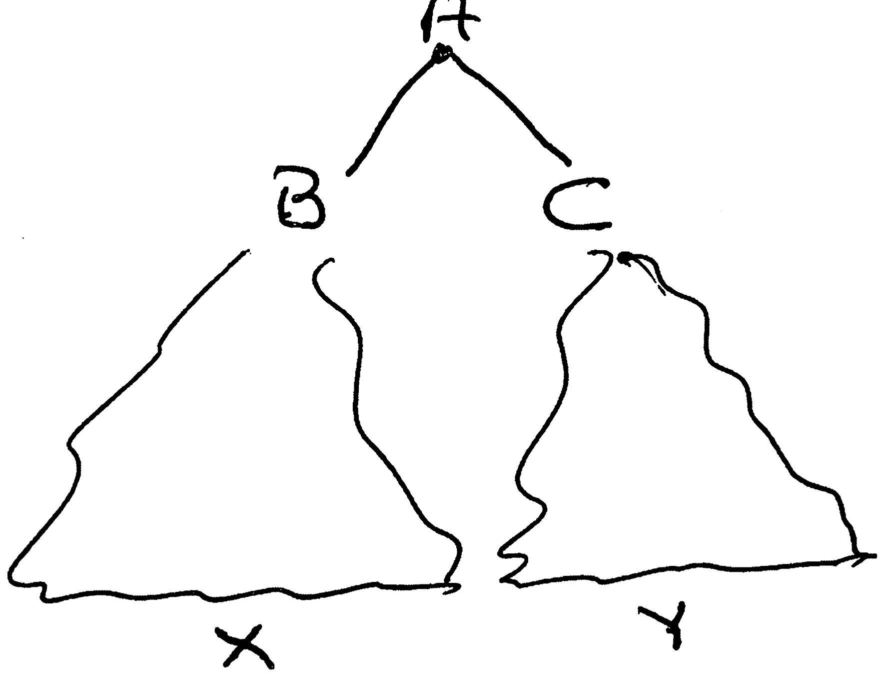
AND ADD

$$\langle B_2 \cdots B_{k+1} \rangle \to B_2 \cdots B_{k+1}$$

ITERATE UNTIL ALL $|RHS| \le 2$

CKY PARSER

$$A \to BC$$



IF $|XY| = k$ THEN HAVE CASES

$|X| = 1 \quad |Y| = k-1$

$\vdots \qquad \vdots$

$|X| = k-1 \quad |Y| = 1$

COCKE-KASAMI-YOUNGER
CKY

$S \to AB \mid AC \mid AD \mid AE$
$\quad \mid BA \mid BF \mid BG \mid BH$

$A \to a \qquad C \to DS \qquad F \to GS$
$B \to b \qquad D \to SB \qquad G \to SA$
$\qquad\qquad E \to BS \qquad H \to AS$

	b	a	b	b	a	a	b	a
1	B	A	B	B	A	A	B	A
2	S	S		S		S	S	
3	E,D	D	E	G	H	H,G		
4		S	S	S				
5	E,C	H,G	E,D	F,G				
6	S	S	S					
7	E,C,D	H,F,G						
8	S							

ACCEPT (VERY AMBIGUOUS)

CKY Example #1

It's easy to see how we can fill in the slots in row 1, since slot (1,j) is just the non-terminals, A, having a rule $A \rightarrow a_j$. Slot (i,j) gets symbol A if there is a rule $A \rightarrow BC$, where B is in slot (k,j), C is in (i-k,j+k). Think about it!!!

Present the CKY recognition matrix for the string a a b a b b assuming the grammar

S →	**A T | B U**		
T →	**b | B S | A V**	**V →**	**T T**
U →	**a | A S | B W**	**W →**	**U U**
A →	**a**	**B →**	**b**

	a	a	b	a	b	b
1	A,U	A,U	B,T	A,U	B,T	B,T
2	W	S	S	S	V	
3	U	U	T	T		
4	W	S	V			
5	U	T				
6	S					

The string can be generated from S, since S appears in the slot associated with a string of length 6, starting at the first position.

CKY Example #2

Present the CKY recognition matrix for the string (x + x) * x assuming the grammar

S →	**S T | x | LU**		
T →	**P S | N S | M S | D S**		
U →	**S R**	**P →**	**+**
L →	**(**	**N →**	**-**
R →	**)**	**M →**	*****
		D →	**/**

	(	x	+	x	)	*	x
1	L	S	P	S	R	M	S
2			T	U		T	
3		S					
4		U					
5	S						
6							
7	S						

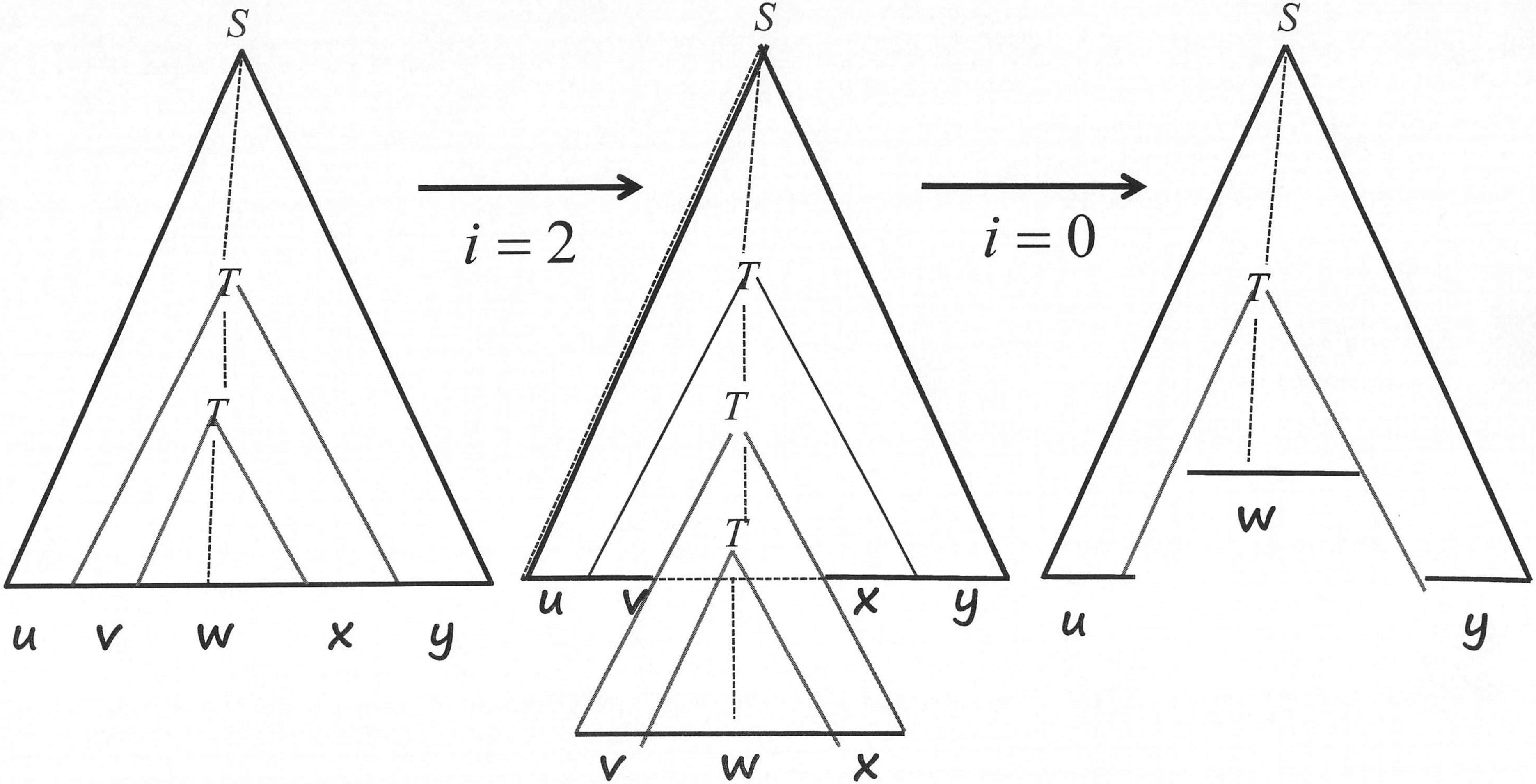
PUMPING LEMMA FOR CFLs

IF L IS A CFL, THEN $\exists N > 0$, SUCH THAT IF $z \in L$ AND $|z| \geq N$ THEN $z = uvwxy$, WHERE $|vwx| \leq N$, $|vx| > 0$ AND $\forall i \geq 0$ $uv^iwx^iy \in L$

NOTE THAT IF $|\Sigma| = 1$ THEN PUMPING LEMMA FOR CFL DEGENERATES TO PUMPING LEMMA FOR REGULAR

THUS, IF $L \subseteq \Sigma^*$, $|\Sigma| = 1$, AND L IS NOT REGULAR THEN L IS NOT A CFL

Visual Support of Proof



Lemma's Adversarial Process

- Assume $L = \{a^n b^n c^n \mid n>0\}$ is a CFL
- P.L.: Provides $N>0$ We CANNOT choose N; that's the P.L.'s job
- Our turn: Choose $a^N b^N c^N \in L$ We get to select a string in L
- P.L.: $a^N b^N c^N = uvwxy$, where $|vwx| \le N$, $|vx|>0$, and for all $i \ge 0$, $uv^iwx^iy \in L$ We CANNOT choose split, but P.L. is constrained by N
- Our turn: Choose $i=0$. We have the power here
- P.L: Two cases:

  (1) vwx contains some a's and maybe some b's. Because $|vwx| \le N$, it cannot contain c's if it has a's. $i=0$ erases some a's but we still have N c's so $uwy \notin L$

  (2) vwx contains no a's. Because $|vx|>0$, vx contains some b's or c's or some of each. $i=0$ erases some b's and/or c's but we still have N a's so $uwy \notin L$
- CONTRADICTION, so L is <u>NOT</u> a CFL

Assignment # 7.1b Sample Key

1. b.) **$L = \{ a^n b^{n!} \mid n>0 \}$**

   PL: Provides N>0

   We: Choose $a^N b^{N!} \in L$

   PL: Splits $a^N b^{N!}$ into uvwxy, $|vwx| \leq N$, $|vx| > 0$, such that $\forall i \geq 0\ uv^i wx^i y \in L$

   We: Choose i=2

   Case 1: vwx contains only b's, then we are increasing the number of b's while leaving the number of a's unchanged. In this case uv^2wx^2y is of form $a^N b^{N!+c}$, c>0 and this is not in L.

   Case 2: vwx contains some a's and maybe some b's. Under this circumstances uv^2wx^2y has at least N+1 a's and at most N!+N-1 b's. But $(N+1)! = N!(N+1) = N!*N+N \geq N! + N > N!+N-1$ and so is not in L.

   Cases 1 and 2 cover all possible situations, so L is not a CFL

$L = \{ ww \mid w \in \{a,b\}^* \}$ IS NOT A CFL

1. ASSUME L IS A CFL
2. PL GIVES YOU N
3. YOU CHOOSE $a^N b^N a^N b^N$
4. PL SPLITS $a^N b^N a^N b^N$ INTO $uvwxy$
   WHERE $|vwx| \le N$, $|vx| > 0$
   AND STATES $\forall i \ge 0 \; uv^i wx^i y \in L$
5. a.) IF vwx IS OVER a'S ONLY
   THEN SINCE $|vwx| \le N$
   vwx CANNOT SPAN FROM ONE OF
   THE a^N SUBSTRINGS TO OTHER,
   AND SO FOR $i=0$, WE GET EITHER
   $a^{N-|vx|} b^N a^N b^N \in L$ OR $a^N b^N a^{N-|vx|} b^N \in L$
   AND BOTH ARE NOT SO.
   b) IF vwx IS OVER b'S ONLY, THE
   SAME ARGUMENTS APPLY
   c) IF vwx SPANS a'S AND b'S THEN
   THIS MUST BE WITHIN ONE OF
   THE $a^N b^N$ OR $b^N a^N$ SUBSTRINGS
   AS $|vwx| \le N$. IF $i=0$ THEN WE
   REDUCE SOME a'S AND SOME b'S, BUT NOT ALL,
   WITHIN JUST ONE OF THE CONSECUTIVE
   a'S AND b'S, BUT NOT THE OTHER
   AND, SO THE RESULTING STRING IS NOT
   IN L ⚔

ANY LANGUAGE OVER A SINGLE LETTER, E.G., $\{a\}$, THAT IS NOT REGULAR IS ALSO NOT A CFL

REASON

$uvwxy$ WHEN OVER A SINGLE LETTER CAN BE REASSOCIATED AS

$w(vx)(uy)$ AND THEN vx CAN BE SEEN AS JUST Z

HERE $|wvx| = |wz| \leq N$

$|vx| = |z| > 0$ NOTE: $v^i x^i = z^i$

RECASTING uy AS t

WE HAVE $\exists N > 0$, SUCH THAT

IF $s \in L$ AND $|s| \geq N$

THEN $s = wzt$, WHERE

$|wz| \leq N$, $|z| > 0$ AND $\forall i \geq 0$

$wz^it \in L$

THAT'S JUST THE PL FOR REGULAR

Non-Closure

- Intersection ($\{ a^n b^n c^n \mid n \geq 0 \}$ is not a CFL)
  $\{ a^n b^n c^n \mid n \geq 0 \} =$
  $\{ a^n b^n c^m \mid n,m \geq 0 \} \cap \{ a^m b^n c^n \mid n,m \geq 0 \}$
  Both of the above are CFLs
- Complement
  If closed under complement then would be closed under Intersection as
  $A \cap B = \sim(\sim A \cup \sim B)$

Max and Min of CFL

- Consider the two operations on languages max and min, where
  - max(L) = $\{ x \mid x \in L$ and, for no non-null y does $xy \in L \}$ and
  - min(L) = $\{ x \mid x \in L$ and, for no proper prefix of x, y, does $y \in L \}$
- Describe the languages produced by max and min. for each of :
  - $L1 = \{ a^i b^j c^k \mid k \le i \text{ or } k \le j \}$ CFL
    - max(L1) = $\{ a^i b^j c^k \mid k = \max(i, j) \}$ Non-CFL
    - min(L1) = $\{ \lambda \}$ (string of length 0) Regular
  - $L2 = \{ a^i b^j c^k \mid k > i \text{ or } k > j \}$ CFL
    - max(L2) = $\{ \}$ (empty) Regular
    - min(L2) = $\{ a^i b^j c^k \mid k = \min(i, j)+1 \}$ Non-CFL
- max(L1) shows CFL not closed under max
- min(L2) shows CFL not closed under min

Solvable CFL Problems

- Let L be an arbitrary CFL generated by CFG G with start symbol S then the following are all decidable
  - Is w in L? — Run CKY; If S in final cell then $w \in L$
  - Is L empty (non-empty)? — Reduce G; If no rules left then empty
  - Is L finite (infinite)? — Reduce G; Run DFS(S); If no loops then finite