COT 4210 Fall 2017 **Midterm 2** **Name:** KEY

Total Points Available 72 Your Raw Score _______ **Grade: M2: ____ Combined: ______**

4 **1.** Write a Context Free Grammar for the language **L**, where
L = { $a^i b^j c^k$ | i, k > 0 , j = i + k }, that is #b's = #a's + #c's

$$S \rightarrow \langle AB\rangle \langle BC\rangle$$
$$\langle AB\rangle \rightarrow a\langle AB\rangle b \mid ab$$
$$\langle BC\rangle \rightarrow b\langle BC\rangle c \mid bc$$

8 **2.** Use the Pumping Lemma for CFLs to show that the following language **L** is not Context Free.
L = { $a^i b^j c^k$ | i, k >0, j = i * k }, that is #b's = #a's * #c's
Be explicit as to why each case you analyze fails to be in **L**. I will do the first two steps.

Me: Assume L is Context Free
PL: Provides a whole number N>0 that is the value associated with L based on the Pumping Lemma for CFLs

ME: $z = a^N b^{N^2} c^N$

PL: $z = uvwxy$, $|vwx| \le N$, $|vx| > 0$ AND $\forall i \ge 0$ $uv^iwx^iy \in L$

ME: CASE 1: ASSUME vwx CONTAINS ONLY b'S AND LET $i=0$
#b'S IN uwy IS AT MOST N^2-1 AND #a'S AND #b'S ARE STILL N. AS $N^2-1 < N^2$, $uwy \notin L$

CASE 2: ASSUME vwx CONTAINS AT LEAST ONE a OR AT LEAST ONE c (IT CANNOT CONTAIN BOTH AS THAT WOULD REQUIRE $|vwx| > N$). LET $i = 2$.
uv^2wx^2y MUST THEN CONTAIN AT LEAST N+1 a'S OR c'S AND JUST N OF THE OTHER. THAT IS, N a'S AND N+1 c'S, AT A MINIMUM, N+1 a'S AND N c'S OR
WE THEN NEED AT LEAST $N(N+1) = N^2+N$ b'S,
BUT uv^2wx^2y CAN CONTAIN AT MOST N^2+N-1 b'S, SINCE AT LEAST ONE OF THE AT MOST N CHARACTERS IN vwx IS AN a OR A c. BUT $N^2+N-1 < N^2+N$ AND SO $uv^2wx^2y \notin L$

THIS COVERS ALL CASES AND SO PL SAYS L IS NOT A CFL.

12 **3.** Consider some language **L**. For each of (a) and (b), and for each of the three possible complexities of **L**, indicate whether this is possible (**Y** or **N**) and present evidence, if you answered **Y**, by providing an example language **A**, with a homomorphism **σ** for **(b)**, and the resulting **L**, or, if you answered **N**, state some known closure property that reflects a bound on the complexity of **L**.

a.) **L = min(A)** where **A** is context-free, not regular.
Can **L** be Regular? Circle **Y** or **N**.
If yes, show **A** and argue **min(A)** is Regular; if no, why not?

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

$L = \text{MIN}(A) = \{\lambda\}$ WHICH IS REGULAR

Can **L** be a non-regular CFL? Circle **Y** or **N**.
If yes, show **A** and argue **min(A)** is a CFL; if no, why not?

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

$L = \text{MIN}(A) = \{a^n b^n \mid n > 0\}$ WHICH IS A NON-REG CFL

Can **L** be more complex that a CFL? Circle **Y** or **N**.
If yes, show **A** and argue **min(A)** is not a CFL; if no, why not?

$A = \{a^i b^j c^k \mid k > i \text{ OR } k > j\}$ IS A CFL

$L = \text{MIN}(A) = \{a^i b^j c^k \mid k = \text{MIN}(i,j)+1\}$ IS NO A CFL

b.) Let **σ** be a homomorphism from **Σ** into regular languages, such that, for each **a∈Σ**, **σ(a) = w_a**, for some string **w_a**. Let **A** be a context free, non-regular language and let **L = σ(A)**.
Can **L** be Regular? Circle **Y** or **N**.
If yes, show **A** and **σ**, and argue **σ(A)** is Regular; if no, why not?

$A = \{a^n b^n \mid n > 0\}$ $\sigma(a) = \lambda$ $\sigma(b) = \lambda$

$L = \sigma(A) = \{\lambda\}$ WHICH IS REGULAR

Can **L** be a non-regular CFL? Circle **Y** or **N**.
If yes, show **A** and **σ**, and argue **σ(A)** is a CFL; if no, why not?

$A = \{a^n b^n \mid n > 0\}$ $\sigma(a) = a$ $\sigma(b) = b$

$L = \sigma(A) = \{a^n b^n \mid n > 0\}$ WHICH IS A NON-REG. CFL

Can **L** be more complex that a CFL? Circle **Y** or **N**.
If yes, show **A** and **σ**, and argue **σ(A)** is not a CFL; if no, why not?

MUST BE A CFL AS THEY ARE CLOSED UNDER HOMOMORPHISM

10 **4.** Present the **CKY** recognition matrix for the string **a-(a)+a** assuming the Chomsky Normal Form grammar, **G** = ({ **E, L, P, R, T, Y, Z** }, { **a, +, -, (,)** }, **R, E**), specified by the rules **R**:.

E → a | EY | LZ | PE
Z → ER
Y → PT
T → a | LZ
P → + | -
L → (
R →)
Note: Do not be surprised if many of the cells are blank when you complete this.

	a	-	(	a	)	+	a
1	E,T	P	L	E,T	R	P	E,T
2				Z		E,Y	
3			E,T				
4		E,Y					
5	E		E				
6		E					
7	E						

A

A little help

Non-Terminal	**First Symbol in Rules**	**Second Symbol in Rules**
E	**E → EY; Z → ER**	**E → PE**
Z	**None**	**E → LZ; T → LZ**
Y	**None**	**E → EY**
T	**None**	**Y → PT**
P	**E → PE; Y → PT**	**None**
L	**E → LZ; T → LZ**	**None**
R	None	**Z → ER**

Is **a-(a)+a** in **L(G)**? Y

What is the order of execution of this approach to determine if some **w**, **|w| = N**, is in **L**? $O(N^3)$

What is the algorithmic strategy, e.g., greedy, divide and conquer, dynamic programming, backtracking or randomized, associated with this **CKY** algorithm called? DYNAMIC PROGRAMMING

8 5. Prove that **Context-Free Languages** are closed under **double quotient with Regular Languages**. Double Quotient of a CFL **L** and a Regular Language **R**, both of which are over the alphabet Σ, is denoted **DQ(L, R)**), and defined as $\mathbf{DQ(L, R) = \{ x \mid xyz \in L \text{ and } y, z \in R \}}$.
That is, we select a prefix **x** of **xyz** in **L**, provided **y** and **z** are both in the Regular Language **R**.
You may assume substitution $\mathbf{f(a) = \{a, a'\}}$, and homomorphisms $\mathbf{g(a) = a'}$ and $\mathbf{h(a) = a, h(a') = \lambda}$.
Here $\mathbf{a \in \Sigma}$ and **a'** is a distinct new character associated with each $\mathbf{a \in \Sigma}$.

DQ(L, R) = $h(f(L) \cap \Sigma^* g(R) g(R))$

You must now be very explicit as to why each language you produce by your various transformations (substitutions and other operations) is a CFL or a Regular language. Do this for **all** parts of the expression you write, with **appropriate justifications** based on known closure properties. You do **not** need to describe what is produced by each such transformation

$g(R)$ REG AS REG. CLOSED UNDER HOMOMORPHISM

$\Sigma^* g(R) g(R)$ REG AS REG CLOSED UNDER * AND CONCAT.

$f(L)$ CFL AS CFL CLOSED UNDER SUBSTITUTION

$f(L) \cap \Sigma^* g(R) g(R)$ CFL AS CFL CLOSED UNDER INTERSECTION WITH REGULAR

$h(f(L) \cap \Sigma^* g(R) g(R))$ CFL AS CFL CLOSED UNDER HOMOMORPHISM

THUS, DQ(L,R) IS A CFL

5 6. Let **L1, L2** be Non-Regular CFLs; **R1, R2** be Regular; Answer is about S and there should be just one cell per row that has a check mark, or **X** if your prefer.

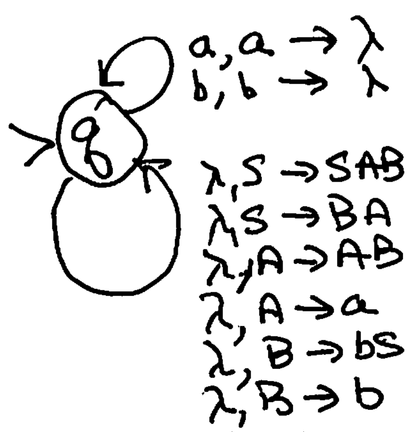
Definition of S / Characterization of S	*Always Regular*	*At worst CFL*	*Might not be CFL*
$S \subseteq L1 \cup L2$			X
$S = L1 \cap R1$		X	
$S = R1 - R2$	X		
$S \supseteq R1$			X
$S = L1 \cup R1$		X	

12 7. Consider the CFG **G = ({ S, A, B } , { a , b } , R , S)** where **R** is:

S → SAB | BA
A → AB | a
B → bS | b

In the PDAs below, you may (and are encouraged) to use a transition diagram where transitions have arcs with labels of form **a, α → β** where **a** ∈ Σ∪{λ}, α, β ∈ Γ*. Note: This just means that you can use extended stack operations that push or pop arbitrary length strings.

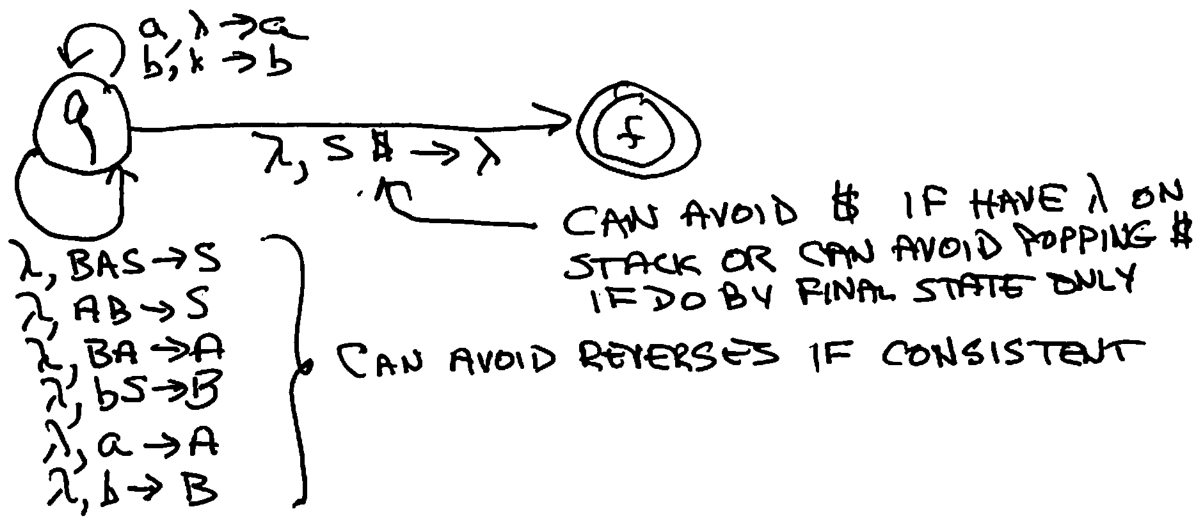
a.) Present a pushdown automaton that parses the language **L(G)** using a top down strategy.
INITIAL CONTENTS OF STACK = S



b.) Now, using the notation of **IDs** (Instantaneous Descriptions, **[q, x, z]**), describe how your PDA in (a) accepts strings generated by **G**.

$$[q, w, S] \vdash^* [q, \lambda, \lambda]$$

c.) Present a pushdown automaton that parses the language **L(G)** using a bottom up strategy. Note: I am fine with your showing strings that are on top of the stack in either reversed or non-reversed form.
INITIAL CONTENTS OF STACK = #



d) Now, using the notation of **IDs** (Instantaneous Descriptions, **[q, x, z]**), describe how your PDA in (c) accepts strings generated by **G**.

$$[q, w, \#] \vdash^* [f, \lambda, \lambda]$$

8. Consider the context-free grammar **G = ({ S, A, B, C, D, E }, { 0, 1 }, R, S)**, where **R** is:

$S \rightarrow A \mid B$
$A \rightarrow 1CD$
$B \rightarrow 1CE$
$C \rightarrow 1C1 \mid 01 \mid \lambda$
$D \rightarrow 0C$
$E \rightarrow 0B \mid A$

3 **a.)** Remove λ-rules, creating an equivalent grammar G'. Show **all** rules.
Nullable = { C }

$S \rightarrow A \mid B$
$A \rightarrow 1CD \mid 1D$
$B \rightarrow 1CE \mid 1E$
$C \rightarrow 1C1 \mid 11 \mid 01$
$D \rightarrow 0C \mid 0$
$E \rightarrow 0B \mid A$

3 **b.)** Remove all **unit** rules, creating an equivalent grammar G''. Show **all** rules.
Unit(S) = {S, A, B}; *Unit(A)* = { A }; *Unit(B)* = { B };
Unit(C) = { C }; *Unit(D)* = { D }; *Unit(E)* = { E, A }

$S \rightarrow 1CD \mid 1D \mid 1CE \mid 1E$
$A \rightarrow 1CD \mid 1D$
$B \rightarrow 1CE \mid 1E$
$C \rightarrow 1C1 \mid 11 \mid 01$
$D \rightarrow 0C \mid 0$
$E \rightarrow 0B \mid 1CD \mid 1D$

2 **c.)** Remove all unproductive symbols, creating an equivalent grammar G'''. Show **all** rules.
Productive = {S, A, B, C, D, E}; *Unproductive* = { }

SAME AS (B)

2 **d.)** Remove all unreachable symbols, creating an equivalent grammar G^{iv}. Show **all** rules.
Unreachable = { A }

$S \rightarrow 1CD \mid 1D \mid 1CE \mid 1E$
$B \rightarrow 1CE \mid 1E$
$C \rightarrow 1C1 \mid 11 \mid 01$
$D \rightarrow 0C \mid 0$
$E \rightarrow 0B \mid 1CD \mid 1D$

3 **e.)** Convert grammar G^{iv} to its **Chomsky Normal Form** equivalent, G^{V}. Show **all** rules.

$S \rightarrow \langle 1C \rangle D \mid \langle 1 \rangle D \mid \langle 1C \rangle E \mid \langle 1 \rangle E$
$B \rightarrow \langle 1C \rangle E \mid \langle 1 \rangle E$
$C \rightarrow \langle 1C \rangle \langle 1 \rangle \mid \langle 1 \rangle \langle 1 \rangle \mid \langle 0 \rangle \langle 1 \rangle$
$D \rightarrow \langle 0 \rangle C \mid 0$
$E \rightarrow \langle 0 \rangle B \mid \langle 1C \rangle D \mid \langle 1 \rangle D$
$\langle 0 \rangle \rightarrow 0$
$\langle 1 \rangle \rightarrow 1$
$\langle 1C \rangle \rightarrow \langle 1 \rangle C$