CSci 435: Formal Languages and Automata

Instructor: Dr. M. E. Kim Date: October 15th, 2020

**Midterm: 100 points + 20 points (optional)**

**Due: by the end of the day, 10/18 (Sun.)**

Name: \_Grant Haataja \_

1. Your answer should be precise and fully described; any sloppy answer will not get a full point.
2. Your hand writing should be clear and readable.
3. Do not insert a photo copy of your handwriting but draw a figure using a graphic tool.

Mark the followings;

Difficulty:

Very Easy: \_\_\_\_\_\_ Easy: \_\_\_\_\_\_ Moderate: \_\_\_✓\_\_\_ Difficult: \_\_\_\_\_\_ Very Difficulty: \_\_\_\_\_\_

Time Taken:

\_\_\_3 Hours and \_\_\_\_15\_\_\_ Minutes.

Q1. [10] Construct a minimal DFA that accepts L = {w ∈{*a, b*}\* | w has at least one *b* and exactly two *a*’s}.

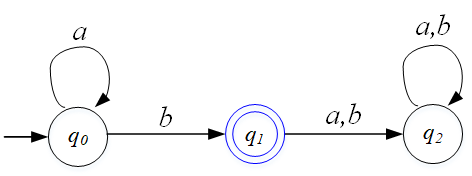
Hint: the number of states are 7.

*Ans.*

Diagram

Description automatically generated

Q2. [10] Let L be the language accepted by the DFA below. Construct a DFA for the language **L2 – L**.

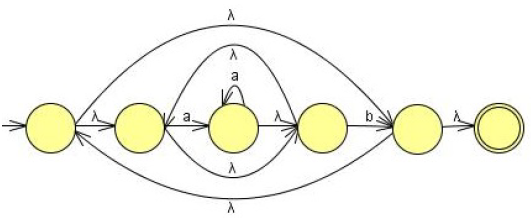


*Ans.* Since L2 = LL, the set of strings in L concatenated with all other strings in L, there are no strings in L that are also in L2. Thus, L2 – L = L2. A DFA for the language is shown below.

A drawing of a person

Description automatically generated

Q3. [15] In the given NFA M of the figure,



1. [5] Give the language **L(M)** that is accepted by M in the simplest ***regular expression***.

*Ans.* L(a\*b)\*

1. [10] Construct a ***minimal DFA*** that is equivalent to the NFA M.

*Ans.*

A drawing of a person

Description automatically generated

Q4. [10] For the language L(*aa*\*(*ab*+*a*)\*)

1. [5] Construct a (minimal) NFA with a single final state that accepts L.

*Ans.*

Diagram

Description automatically generated

1. [5] Construct a right-linear ***regular grammar*** that generates L in the simplest form

Hint: 3 variables and 6 production rules. e.g) A → BC | λ is counted to 2 rules: A → BC or A → λ

*Ans.* S → aA, A → aA | aB | a | λ, B → bA

Q5. [10, optional] The ***chopleft*** operationof a regular language L is removing the leftmost symbol of every string in L:

***chopleft*(L)** = { *w* | *vw* ∈ L, with |*v*| = 1}.

Prove or disprove that the family of regular languages is ***closed*** under the *chopleft* operation.

Hint: If it’s regular, give an idea of constructing an FA that accepts chopleft(L) using an FA M that accepts L.

Otherwise, give a counterexample.

*Ans.* Let D = (Q, Σ, δ, *q0*, F) be a DFA that accepts L. We construct an NFA M = (Q’, Σ, δ’, *q0’*, F’) that accepts chopleft(L) as shown:

Q’ = (Q – {*q0*}) ∪ {*s*} (this indicates removing the leftmost symbol of each string in L, and adding a new state)

*q0*’ = *s*

δ’(*s*, λ) = *q1*

F’ = F

M is obtained by removing the initial state of D, and adding a new initial state *s* connected via λ edges to the original accepting states of D. This effectively removes only the leftmost symbol and then adds a new state with a λ edge to preserve the rest of the structure of D for the new NFA M.

Q6. [10] Prove or disprove that the language L = {*anblak | n=l* or *l* ≠ *k* } is regular.  
If L is regular, give a *regular grammar* that generates L. Otherwise, disprove it by *Pumping Lemma*.

*Ans.* Assume that L is regular, so the pumping lemma must hold. Given an arbitrary positive integer *m*, we pick a string *w* such that |*w*| > *m*.

Case 1: Let *w* = *ambmam*∈L. We split the string *w* into substrings such that *w* = *xyz*, with the constraints |*xy*| ≤ *m*, and ≥ 1. So, we have *x* = *ar, y* = *as*, *z* = *am-r-sbmam*, *r*+*s* ≤ *m*, *s* ≥ 1. Then, by pumping with *i =* 0, we have *w0* = *ar*(*as*)0*am-r-sbmam* = *am-sbmam* which is clearly not in L since the number of leading *a*’s and the number of *b*’s are not equal. This contradicts the pumping lemma.

Case 2: Let *w* = *ambmam+1*∈L. Due to the constraints |*xy*| ≤ *m* and ≥ 1, we have that *y* must be all *b*’s. *y* = *bt*, 1 ≤ *t* ≤ *m*. Then, by pumping with *i* = 2, we obtain the string *w2* = *am+tbta2m+1* which is not in L when *t* = 2*m* + 1. This contradicts the pumping lemma, which proves that L is not regular.

Q7. [10] Show that the L = { *anbmck* | *n=m* or *m* ≠ *k* } where *n, m, k* ≥ 0 is context-free, by giving the Context-Free Grammar that generates it.

*Ans.* We have three cases.

Case 1: Equal number of a’s and b’s, with 0 or more c’s.

Case 2: Less b’s than c’s, with 0 or more a’s.

Case 3: More b’s than c’s, with 0 or more a’s.

S → AB | CD | CE

A → *a*A*b* | *ab*

B → B*c* | *λ*

C → *a*C | *λ*

D → *b*D*c* | D*c* | *c*

E → *b*E*c* | *b*E | *b*

Q8. [15] In the given CFG, G = ( {S}, {*a, b*}, S, P ) with productions S → SS | *a*S*b* | *b*S*a*| λ

1. [5] Give the language, L(G), that is generated by G, in a formal expression.

*Ans.* L = {*w* ∈{*a, b*}\*| *na*(*w*) = *nb*(*w*)}

1. [5] Decide if the G is ***ambiguous*** or not. Justify your answer.

*Ans.* G is not ambiguous, because there is no way to obtain more than one derivation tree for any input string. The productions are very clear, and there is only one way to derive from them.

1. [5] If G is ambiguous, give an unambiguous grammar. Otherwise, show that G is inherently ambiguous.

*Ans.* G is already unambiguous.

Q9. [10] Transform the grammar with the following productions into a ***Chomsky Normal Form***

S → *ba*AB, A → *b*AB | λ, B → BA*a* |A | λ.

*Ans.* First, we eliminate the λ-productions, giving:

S → *ba*AB | *ba*A | *ba*B | *ba*, A → *b*AB| *b*A | *b*B | *b*, B → BA*a* | A | *a*

Next, we eliminate unit-productions, giving:

S → *ba*AB | *ba*A | *ba*B | *ba*, A → *b*AB| *b*A | *b*B | *b*, B → BA*a* | *b*AB| *b*A | *b*B | *a* | *b*

Then, we create new productions so that each body has only a single terminal, or multiple variables.

S → BbA­­aAB | BbAaA | BbAaB | ­BbAa, A → BbAB | BbA | BbB | *b*, B → BAAa | BbAB | BbA | BbB | *a* | *b*,

Aa → *a*, Bb → *b*

Finally, we create new productions so that each body has only a single terminal or **two**variables.

S → CD | CA | CB | BbAa, A → BbD | BbA | BbB | *b*, B → EAa | BbD | BbA | BbB | *a* | *b*,

A­a → *a*, Bb → *b*, C → BbAa, D → AB, E → BA

This grammar is now in Chomsky Normal Form.

Q10. [10] In the grammar G= ({S, A, B}, {*a, b*}, S, P) with the following productions,

S → *aA*B, A → *b*B*b*, B → A | λ

1. [5] Generate the simplified equivalent simplest grammar.

*Ans.* S → *a*AA, A → *b*A*b* | λ

1. [5] Give the language, L(G), that is generated by G, in a formal expression.

*Ans.* L = {*w* ∈{*a, b*}\*| *w* = *ab2n*, *n* ≥ 0}

Q11. [10, optional] The reverse of a string can be defined recursively as:

for all *a*∈Σ, w∈Σ\*.

Prove that **,** for all *u, v* ∈, by Mathematical Induction.

*Ans.*

Base Case: Let *u* be a string of length 0, so *u* = *λ*.

Then (*uv*)R = (*λv*)R = *v*R = *v*R*λ* = *v*R*λ*R = *v*R*u*R

Induction Hypothesis: For all *n* ≥ 0, for any string *u* of length *n*, and for all strings *v* ∈, (*uv*)R = *v*R*u*R.

Induction Step: Let *u* be a string of length *n* ≥ 1. Then *u* = *ax* for some string *x* where |*x*| < *n* and *a* ∈.

Then (*uv*)R = ((*ax*)*v*)R = (*a*(*x*v))R = (*xv*)R*a*R = *v*R*x*R*a*R = *v*R(*ax*)R = *v*R*u*R

Therefore, **,** for all *u, v* ∈.