MA/CSSE 474 Exam 2 Winter 2009-10 Name________________________

Closed book and notes, except for three 8.5 x 11 sheets of paper (can be 2-sided).

No electronic devices, especially ones with headphones.

Scores:

Problem	Possible	Score
1	48	
2	15	
3	10	
4	10	
5a	10	
5b	10	
5c	10	
6	7	
Total	120	

1. (45 points) For each of the following statements, circle T or F to indicate whether it is *True* or *False*.
If it is sometimes False, you should choose False.
You do not have to give proofs or counterexamples.
For each part, you get 1 point for circling IDK (I don't know), 3 for circling the correct answer, and 0 for circling the incorrect answer. Reason: When you don't know something, knowing that you don't know counts for something.

a) T **F** IDK (a ∪ b) a* = (ε ∪ b) a* (equal in the sense that they define the same language).
b) T **F** IDK $L(\varepsilon^*) \cap L(\varnothing^*) = \varnothing$.
c) **T** F IDK $\{a^n b^* a^n : n \geq 0\} \cap \{b^n a^* b^n : n \geq 0\}$ is a regular language.
d) **T** F IDK $\{w \in \{a, b, c\}^* : (|w| \text{ is even}) \rightarrow (w \text{ contains an even number of a's}\}$ is a regular language.
e) T **F** IDK $\{w = xy, x \in a^*b, y \in a^*b, |x| = |y|\}$ is a regular language.
f) T **F** IDK $\{(abc)^n a^n : n \geq 0\}$ is a regular language.
g) **T** F IDK $\{x\#y: x, y \in \{a, b\}^*, |x| + |y| \geq 4\}$ is a regular language.
h) **T** F IDK Let $L = \{a^p : p \text{ is a prime integer}\}$. L* is a regular language.
i) T **F** IDK If L is regular and L ∩ M is regular, then M must be regular.
j) **T** F IDK If L is regular and L ∩ M is not regular, then M cannot be regular.
k) **T** F IDK If L and M are regular, then $N = \{x : x \in L \text{ and } x^R \in M\}$ must be regular.
l) **T** F IDK The number of regular languages over the alphabet Σ ={a, b} is countable.
m) T **F** IDK Let *M* be a DFSM such that $|K_M| = 100$, $\Sigma_M = \{a, b\}$, and $L(M)$ is finite.
We cannot tell whether $w = a^{100}b^{100} \in L(M)$ without running *M* on *w*.
n) **T** F IDK The nonregular languages are closed under complement.
o) T **F** IDK Let *L* be such that, for each $w \in L$, there exists some DFSM that accepts *w*. Then *L* must be regular.
p) T **F** IDK There is an infinite number of uncountable languages.

Explanations are on the next two pages.

a) T **F** IDK $(a \cup b)\, a^* = (\varepsilon \cup b)\, a^*$ (equal in the sense that they define the same language).

The empty string is not in the first language, but it is in the second language.

b) T **F** IDK $L(\varepsilon^*) \cap L(\varnothing^*) = \varnothing$.

The intersection contains ε.

c) **T** F IDK $\{a^n b^* a^n : n \geq 0\} \cap \{b^n a^* b^n : n \geq 0\}$ is a regular language.

Note that:

- $\{a^n b^* a^n : n \geq 0\} = (aa)^* \cup b^* \cup$ {some strings that start with a and contain both a's and b's}.
- $\{b^n a^* b^n : n \geq 0\} = a^* \cup (bb)^* \cup$ {some strings that start with b and contain both a's and b's}.

L contains all strings that are in both of those languages. So $L = (aa)^* \cup (bb)^*$.

d) **T** F IDK $\{w \in \{a, b, c\}^* : (|w| \text{ is even}) \rightarrow (w$ contains an even number of a's$\}$ is a regular language.

Regular. $L = \{w \in \{a, b, c\}^* : |w|$ is odd or w contains an even number of a's$\}$. Build an FSM for $\{w \in \{a, b, c\}^* : |w|$ is odd$\}$ and one for $\{w \in \{a, b, c\}^* : w$ contains an even number of a's$\}$. The regular languages are closed under union.

e) T **F** IDK $\{w = xy, x \in a^*b, y \in a^*b, |x| = |y|\}$ is a regular language.

Not regular. $L = \{a^n b a^n b, n \geq 0\}$, which we show is not regular by pumping. Let $w = a^k b a^k b$. y must occur in the first a region and be equal to a^p for some nonzero p. Let $q = 2$. The resulting string is $a^{k+p} b a^k b$, which is not in L.

f) T **F** IDK $\{(abc)^n a^n : n \geq 0\}$ is a regular language.

Not regular. It is possible to do this directly using the Pumping Theorem, but it is tedious. Instead, note that if L were regular, then L^R would also be regular. $L^R = \{a^n (cba)^n : n \geq 0\}$. We show that this is not regular by using the Pumping Theorem. Let $w = a^k (cba)^k$. y must occur in the initial a region and be equal to a^p for some nonzero p. Let $q = 0$ (i.e., pump out). The resulting string is: $a^{k-p}(cba)^k$. This string is not in L because the number of initial a's no longer equals the number of occurrences of cba.

g) **T** F IDK $\{x\#y: x, y \in \{a, b\}^*, |x| + |y| \geq 4\}$ is a regular language.

Regular. One way to prove this is:

Let $L_1 = (a \cup b)^* \# (a \cup b)^*$.

Let $L_2 = \{w \in \{a, b, \#\}^* : |w| < 5\}$. L_2 is regular because it is finite.

Then we observe that $L = L_1 - L_2$. So L must be regular (the set of regular languages is closed under set difference).

h) **T** F IDK Let $L = \{a^p : p$ a is prime integer$\}$. L* is a regular language.

Yes. $L = \{aa, aaa, aaaaa, \ldots\}$. So L^* contains ε plus every string that consists of only a's except for the single string a. L^* can thus be described the regular expression $\varepsilon \cup aaa^*$.

i) T **F** IDK If L is regular and L ∩ M is regular, then M must be regular.

No Let M be any non-regular language, and L be ∅. Then L ∩ M = ∅, which is regular.

j) **T** F IDK If L is regular and L ∩ M is not regular, then M cannot be regular.

If L and M are both regular, so is L ∩ M.

k) **T** F IDK If L and M are regular, then N = $\{x : x \in L \text{ and } x^R \in M\}$ must be regular.

The language is $L \cap M^R$. The regular languages are closed under both reverse and intersection.

l) **T** F IDK There is a countable number of regular languages over the alphabet Σ ={a, b}.

True. The number of regular languages over Σ is at least countably infinite because it includes all of {a}, {aa}, {aaa}, {aaaa}, … There can be no more than a countably infinite number of regular languages over Σ since every regular language can be described by some regular expression. The number of regular expressions, given the alphabet Σ, is countably infinite because it is possible to enumerate them lexicographically.

m) T **F** IDK Let M be a DFSM such that $|K_M| = 100$, $\Sigma_M = \{a, b\}$, and $L(M)$ is finite. We cannot tell whether $w = a^{100}b^{100} \in M$ without running M on w.

False. Since $L(M)$ is finite, we know that M does not contain any loops that are on a path to an accepting state. So, since $|K_M| = 100$, $L(M)$ can contain no strings of length greater than 99. So we know that $w \notin L(M)$.

n) **T** F IDK The nonregular languages are closed under complement.

By contradiction. Suppose L is not-regular and ¬L is regular. Then ¬¬L must be regular, but ¬¬L = L.

o) T **F** IDK Let L be such that, for each $w \in L$, there exists some DFSM that accepts w. Then L must be regular.

False. Given any individual string w, there is a simple DFSM that accepts it. But there's only a DFSM that accepts L (thus making L regular) if there's a single DFSM that accepts **all** of the strings in L. If this claim were true, **every** language would be regular. While any union of a finite languages is regular, the union of an infinite set of regular languages is not necessarily regular.

p) T **F** IDK There is an infinite number of uncountable languages.

There are no uncountable languages, since each language's alphabet is finite, and each string is finite.

2. (15 points) Let *mid* be a function on strings, defined as follows:
For any string *s* in some language *L* over Σ:

- If $|s| \leq 2$ then $mid(s) = \varepsilon$.
- If $|s| > 2$ then let *x* and *z* be single characters in Σ.
Then we can rewrite *s* as *xwz* for some $w \in \Sigma^*$. $mid(s) = w$.

For any language *L* over Σ, we can define the function *midL*(*L*) as follows:
$midL(L) = \{t \in \Sigma^* : t = mid(s) \text{ for some } s \in L\}$

a) (2) What is *midL*(a*ba*)?
L(a*ba*∪ a*) ([if you "take the ends off of" ba^n, you get a^{n-1}]

b) (3) Is FIN (the set of finite languages) closed under *midL*? YES NO
The length of the longest string in *midL(L)* is less than the length of the longest string in L. An infinite language must contain arbitrarily long strings.

c) (5) Is INF (the set of infinite languages) closed under *midL*? YES NO
Claim: For any N>0 midL(L) contains a string whose length is at least N. Why? L contains a string w whose length is N+2. Then |mid(w)| = N+2-2 = N. Thus midL(L) is infinite.

d) (5) Are the regular languages closed under *midL*? YES NO
Yes. Note that *midL*(*L*) contains all strings that can be derived by taking some string in *L* and erasing the first and last characters. If *L* is regular, then it is accepted by some DFSM *M* = (*K*, Σ, δ, *s*, *A*). From *M*, we construct a new FSM *M*′ that accepts *midL*(*L*). Initially, let *M*′ be *M*. Create a new start state *s*′ and a new accepting state *a*′. Make *a*′ the only accepting state of *M*′. For every transition ((*s*, *c*), *q*) in δ, add to *M*′ the transition ((*s*′, ε), *q*). (So, for every first move that *M* can make on some input character, *M*′ can make the same move without that character.) Similarly, for every transition ((*q*, *c*), *a*) in δ, where $a \in A$, add to *M*′ the transition ((*q*, ε), *a*′). (So, for every last move that *M* can make on some input character, *M*′ can make the same move without that character.)

3. (10 points) Give a decision procedure for the following problem:
Let Σ = {a, b}. Given an FSM *M*, does *L*(*M*) contain at least one string that starts with ab?

   1. Build an FSM *M** that accepts the language (ab)(a ∪ b)*.
   2. Build an FSM *M*** that accepts *L*(*M*) ∩ *L*(*M**).
   3. If *L*(*M***) is empty, return *False*, else return *True*.

4. (10 points) Show a context-free grammar that generates $\{a^n b^m : m \geq n, m\text{-}n \text{ is odd}\}$:
$S \rightarrow aSb \mid S \rightarrow Sbb \mid b$

5. (30 points) For each part of this problem, choose one of the following parts of problem 1 (circle its letter here).
Then circle the answer that you gave for that part.
Finally, give a proof of the answer that you gave. Your proof does not have to be formal, but it must be convincing.
You will not get credit for all three parts of this problem if you do not choose at least one part for which you answered T and at least one part for which you answered F.
No credit for a part of this problem if your answer for the corresponding part in #1 is incorrect.
I may give a few extra-credit points for any part if you choose one of the harder ones and explain it very well.
I consider parts c, f, and g to be the hard ones that may be worth extra credit.

See #1 for the answers

6. (7 points) The table below shows the DFSM and my solution to a problem from HW7 (calculating a regular expression that defines the language that is recognized by a given DFSM). Your job is to use these results to compute the value of r_{223}. Be sure that your answer demonstrates that you use the recursive formula. You are not required to simplify

$r_{223} = r_{222} \cup r_{232} r_{332}^* r_{322} = (10)^* \cup (10)^*11(\varepsilon \cup 0(01)^*1)^* 00(10)^*$

	k=0	k=1	k=2
r_{11k}	ε	ε	$\varepsilon \cup 0(10)^*1 = (01)^*$
r_{12k}	0	0	$0 \cup 0(\varepsilon \cup 10)^* (\varepsilon \cup 10) = 0(10)^*$
r_{13k}	1	1	$1 \cup 0(\varepsilon \cup 10)^*11 = 1 \cup (01)^+1 = (01)^*1$
r_{14k}	$\varnothing$	$\varnothing$	$0(10)^*0$ Here and in later rows, I skip the $(\varepsilon \cup 10)^* = (10)^*$ step
r_{21k}	1	1	$(10)^*1$
r_{22k}	ε	$\varepsilon \cup 10$	$(10)^*$
r_{23k}	$\varnothing$	11	$11 \cup (\varepsilon \cup 10)(10)^*11 = (10)^*11$
r_{24k}	0	0	0
r_{31k}	0	0	$0 \cup 00(10)^*1 = 0(\varepsilon \cup 0(10)^*1) = 0(01)^*$
r_{32k}	$\varnothing$	00	$00 \cup 00(10)^* (\varepsilon \cup 10) = 00(10)^*$
r_{33k}	ε	$\varepsilon \cup 01$	$\varepsilon \cup 01 \cup 00(01)^*11 = \varepsilon \cup 0(\varepsilon \cup 0(01)^*1)1 = \varepsilon \cup 0(01)^*1$
r_{34k}	1	1	1
r_{41k}	$\varnothing$	$\varnothing$	$\varnothing$
r_{42k}	$\varnothing$	$\varnothing$	$\varnothing$
r_{43k}	$\varnothing$	$\varnothing$	$\varnothing$
r_{44k}	$\varepsilon \cup 0 \cup 1$	$\varepsilon \cup 0 \cup 1$	$\varepsilon \cup 0 \cup 1$