Midterm Exam

Instructions: The exam has **six** pages, including this cover page, printed out two-sided (no more wasted paper). Please read the questions carefully, then answer carefully. Be mathematically precise, and precise and grammatical with your English, too; you *know* what an ogre that crazy professor is.

Good luck,
phil rogaway

Name:

Notation:

CFG = context free grammar
CFL = context free language
CNF = Chomsky normal form
DFA = deterministic finite automaton
NFA = nondeterministic finite automaton
PDA = pushdown automaton
reasonably efficient = polynomial running time

On page	you got
2	
3	
4	
5	
6	
Σ	

1 True or False

Indicate if the following statements are **true** or **false**, by **filling in** (darkening) the correct box. Some of the questions will be familiar, but do be careful. Do **not** provide any justification. If in doubt, guess; missing answers will be treated as.

---

1. There is an efficient algorithm to decide if a multivariate polynomial over the integers has an integer root. **True** **False**

2. $L = \emptyset$ is a language. **True** **False**

3. All finite languages are regular. **True** **False**

4. An infinite language can have an infinite complement. **True** **False**

5. All infinite languages have infinite complements. **True** **False**

6. The union of infinitely many regular languages is regular. **True** **False**

7. If L is regular then so is $\{xx : x \in L\}$. **True** **False**

8. If L is regular then so is $\{xy : x, y \in L\}$. **True** **False**

9. Let $A = \{1^{2^p} : p \text{ is prime}\}$. Then A^* is regular. **True** **False**

10. The pumping lemma is a useful tool to show that a language is regular. **True** **False**

11. Language $L = \{w \in \{0,1\}^* : \ w \text{ has an equal number of 01's and 10's}\}$ is regular. **True** **False**

12. For every number n, the language $L_n = \{0^n 1^n\}$ is regular. **True** **False**

13. If there's a 10-state DFA that accepts L then there's a 20-state DFA that accepts L. **True** **False**

14. Given a DFA M, there is a reasonably efficient. procedure to find a *smallest* DFA for $L(M)$. **True** **False**

15. Given an NFA M, we know a reasonably efficient procedure to find a *smallest* NFA for $L(M)$. **True** **False**

16. If M is an NFA then $(L(M))^*$ can be accepted by an NFA. **True** **False**

---

17. If L^* is regular then L is regular. **True** **False**

---

18. If M is an NFA and there is one `abb`-labeled path from the start state to a final state, and there is another `abb`-labeled path from the start state to a non-final state, then M is *invalid*: it neither accepts nor rejects `abb`. **True** **False**

---

19. If there's a 10-state NFA that accepts L then there's a 100-state DFA that accepts L. **True** **False**

---

20. Given $h\colon \Sigma \to \{\mathtt{0},\mathtt{1}\}^*$, define $h(L) = \{h(a_1)\cdots h(a_n) : a_1\cdots a_n \in L\}$. Then $h(L)$ is regular if L is regular. **True** **False**

---

21. If L^* is context free then L is context free. **True** **False**

---

22. All DFA-acceptable languages are context free. **True** **False**

---

23. You can always convert a PDA into an NFA for the same language. **True** **False**

---

24. Every subset of a context free language is regular. **True** **False**

---

25. There's a reasonably efficient procedure to decide if a string w is in the language of a CNF CFG G. **True** **False**

---

26. If a CFG G is in CNF, then G is *not* ambiguous. **True** **False**

---

27. Some PDAs need infinite-length descriptions, as the rules (productions) can be an arbitrary subset of $V \times (V \cup \Sigma)^*$. **True** **False**

---

28. The context-free languages are closed under union. **True** **False**

---

29. The context-free languages are closed under complement. **True** **False**

---

30. If A and B are regular then $\{xy \| x \in A \text{ and } y \in B \text{ and } |x| = |y|\}$ is context free. **True** **False**

---

2 Short Answer

1. Complete the following sentence, being mathematically precise and following the conventions of your text: *A **DFA** is a five-tuple $M = (Q, \Sigma, \delta, q_0, F)$ where Q is a finite set, Σ is an alphabet, $q_0 \in Q$, $F \subseteq Q$, and δ is a function with domain* [ ] *and range* [ ] .

2. Carefully explain what it **means** if I say: "the context-free languages are closed under intersection." Don't indicate if the statement is true or false—just provide a precise mathematical translation of the meaning of the claim.

3. List the first five strings, in lexicographic order, of the language

$$L = \{x \neq y : \ x, y \in \{\mathtt{0}, \mathtt{1}\}^* \text{ are unequal strings}\}$$

   Here "$\neq$" is a formal symbol, just like 0 and 1. Assume that characters are ordered $0 < 1 <\neq$.

4. Sketch, briefly and informally, how a PDA for

$$L = \{x \neq y : \ x, y \in \{\mathtt{0}, \mathtt{1}\}^* \text{ are unequal strings}\}$$

   would work.

5. Give an example of a claim that we proved using the **product construction**. Don't prove the claim—just make a precise claim that was proven with the product construction.

6. Let $M = (Q, \Sigma, \delta, q_0, F)$ be a DFA. In class and in the online notes we described an algorithm to minimize the number of states in a DFA for $L(M)$. The algorithm worked by defining an equivalence relation $\sim$ on pairs of states of Q. In particular, we said that

   $p \sim q$ if for all $x \in \Sigma^*$, $\boxed{}$.

   You will wish to use the notation $\delta^*(q, x)$, as we did in class, to indicate the state you are in if you start in state q and consume the string x.

7. You are given the regular expression $\alpha = (000)^* \cup (111)^*$. Composing the constructions given in class and in your text (do not "simplify" anything), imagine converting α into a **DFA** M for which $L(M) = L(\alpha)$. How many states will M have?

8. Specify a CFG for the language $L = \{a^n b^m : m > n\}$. Make your CFG as simple as possible.

9. Complete the definition, being precise with any quantifiers and not using any form of the word "ambiguous" in your definition:

   A CFL L is ***inherently ambiguous*** if:

10. Carefully state the ***pumping lemma*** for **context free** languages. Don't use the word "pumps" and be careful with any quantifiers.

11. How many states are in a *smallest* DFA for the language $L_n = (a^n)^*$ if the underlying alphabet is $\Sigma = \{a, b\}$? (Note: $b^n \notin L_n$.)

12. Continuing the last problem: how would you prove that, for all n, there is no *smaller* DFA for L_n? (You don't need to provide such a proof, but name the main "tool" and sketch the main idea.)

13. A *regular grammar* is a context-free grammar $G = (V, \Sigma, R, S)$ in which every rule is of the form $A \to \varepsilon$ or $A \to aB$, where a is a terminal and A and B are variables. Suppose that L is regular, say $L = L(M)$ for a DFA $M = (Q, \Sigma, \delta, q_0, F)$. Define a CFG $G = (V, \Sigma, R, S)$ such that $L = L(G)$ by saying:

14. A *prefix* of a string y is a string x such that $y = xx'$ for some x'. A prefix is *proper* if it is not the empty string. For any language L, let $\mathcal{G}(L) = \{w \in L \parallel \text{ no proper prefix of } w \text{ is in } L\}$. List the elements of $\mathcal{G}(\{\varepsilon, 00, 01, 110, 0100, 0110, 1110, 1111\})$.

15. In a paragraph of 2–5 clear and grammatical English sentences, answer the following question: what is the *scientific value* of having multiple characterizations of a class of languages, such as the regular languages or the CFLs?