CS 107 (Automata Theory)
Homework Supplement

Numbers enclosed by :	**parentheses**	**brackets**
are exercises from :	Introduction to Automata Theory, Languages and Computation”	Introduction to the Theory of Computation
by authors :	John Hopcroft and Jeffrey Ullman, First Edition	Michael Sipser, Second Edition

Chapter 1 : Fundamentals

1-1 (1.2) Prove by induction on n that

a) $\sum_{i=0}^{n} i = \frac{n(n+1)}{2}$ b) $\sum_{i=0}^{n} i^3 = \left(\sum_{i=0}^{n} i\right)^2$

1-2 (1.3) A palindrome can be defined as a string that reads the same forward and backward, or by the following definition

a) ε is a palindrome
b) If a is any symbol, then the string a is a palindrome
c) If a is any symbol and x is a palindrome, then axa is a palindrome
d) Nothing is a palindrome unless it follows from a), b) and c)

Prove by induction, on the length |w| of the string w, that the two definitions are equivalent.

1-3 (1.7) Find the transitive closure, the reflexive and transitive closure, and the symmetric closure of the relation {(1,2), (2,3), (3,4), (5,4)}.

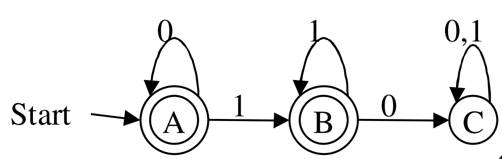
1-4 (1.10) Prove that any subset of a countably infinite set is either finite or countably infinite.

1-5 a) [0.5] b) [0.9] c) [0.10] d) [0.11]

Chapter 2 : Deterministic Finite Automata

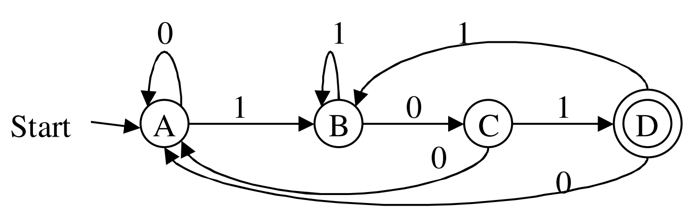
2-1 (2.6) Describe in English the sets accepted by the finite automata whose transition diagrams are given below:

a)

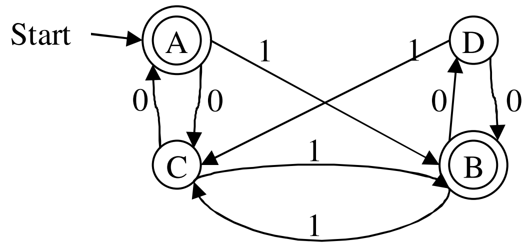


this is discussed in class.

b)



this looks easy but is a bit tricky.

c) 

this is very difficult to describe.

2-2 Consider the language

$L = \{w \in \{a, b\}^* \mid$ every a in w is followed immediately by the string bb$\}$

Draw the state diagram of a deterministic finite automaton that recognizes L.

2-3 Let $\Sigma = \{0, 1\}$. Consider the language L consisting of words that contain 010.

a) Draw a DFA A for L.
b) For each state q of the DFA A in part (a), precisely describe the set of words that lead the DFA from its initial state to q.

2-4 (2.5) Give deterministic finite automata accepting the following languages over the alphabet $\Sigma=\{0,1\}$.

a) The set of all strings ending in 00.
b) The set of all strings with three consecutive 0's.
c*) The set of all strings such that every substring of length three contains exactly one 0.
d) The set of all strings beginning with a 1 which, interpreted as the binary representation of an integer, is congruent to zero modulo 5. (Hint: A left shift in binary is equivalent to multiplying the number by two. Observe how different numbers multiplied by 2 behave modulo 5.)
e*) The set of all strings such that the 3rd symbol from the right end is 1.
f) The set of all strings excluding 011 and 10.

2-5 True or False: Give a short justification (one or two sentences).

The language of a DFA is empty if and only if the set of its final states is empty.

2-6 (2.4) Suppose δ is the transition function of a DFA. Prove that for any input strings x and y, $\delta(q, xy)=\delta(\delta(q, x), y)$. Hint: Use induction on $|y|$.

2-7 a) [1.6c,d,i,l]

Chapter 3 : Nondeterministic Finite Automata

3-1 For $\Sigma = \{ a, b, c, d \}$, consider the set L of strings w such that the last symbol of w has not appeared before. For example, the strings ε, c, adcdb, and ddbbca are in L, but the strings aba and dbcdc are not in L. Give a nondeterministic finite automaton that accepts L. Try to take advantage of nondeterminism as much as possible. Explain your answer.

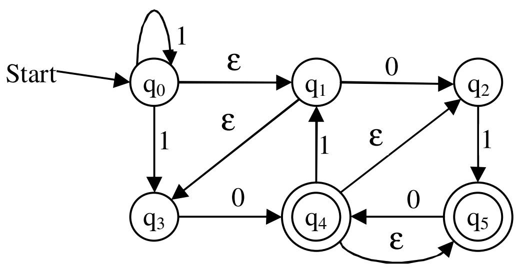
3-2 (2.8) Give nondeterministic finite automata accepting the following languages :

a) The set of strings in $\{0,1\}^*$ such that some two 0's are separated by a string whose length is 4i, for some integer $i\geq 0$.
c*) The set of all strings of 0's and 1's such that the 3rd symbol from the right end is a 1. How does your answer compare with the DFA of exercise 2.5e* (i.e. 2-4e* above) ?

3-3 For two words y and w, we say that y is a subword of w, if w = xyz for some words x and z. For example, the word 011 has the following subwords: ε, 0, 1, 01, 11, and 011. For a language L, the set of its subwords, denoted subwords(L), contains all the subwords of the words in L. That is, a word y belongs to subwords(L) precisely when there exists a word w in L such that y is a subword of w.

Let L be a language, and let A be a DFA for L with states Q, initial state q_0, final states F, and transition function δ. Define an automaton B that accepts subwords(L). The automaton B can be a DFA or an NFA or an ε-NFA. Ensure that B accepts all strings in subwords(L), and all strings accepted by B are in subwords(L).

3-4 For the following NFA N, produce an equivalent DFA D. Draw D's state diagram. Show your solution, not just the final answer.



3-5 (2.9) Construct DFA's equivalent to the NFA's

a) $(\{p,q,r,s\},\{0,1\},\delta_1,p,\{s\})$

b) $(\{p,q,r,s\},\{0,1\},\delta_2,p,\{q,s\})$

where δ_1 and δ_2 are given below :

δ_1	0	1
p	p,q	p
q	r	r
r	s	-
s	s	s

δ_2	0	1
p	q,s	q
q	r	q,r
r	s	p
s	-	p

3-6 a) [1.7b,c] b) [1.16a,b]

Chapter 4 : Regular Expressions

4-1 (2.11) Describe in English the sets denoted by the following regular expressions :

a) (11+0)*(00+1)*

b*) (ε+1+11)(0+01+011)*

c) [00+11+(01+10)(00+11)*(01+10)]*

4-2 (2.10) Write regular expressions for each of the following languages over the alphabet {0,1}. Provide justification that your regular expression is correct.

a) The set of all strings with at most one pair of consecutive 0's and at most one pair of consecutive 1's.

b) The set of all strings in which every pair of adjacent 0's appears before any pair of adjacent 1's.

c) The set of all strings not containing 101 as a substring.

4-3 Let $\Sigma = \{0, 1\}$. Consider the language L consisting of words that contain 010. Write a regular expression for L.

4-4 Given the language $L = \{\ xyx^R \mid x \in \{a, b\}^+, y \in \{a, b\}^*$, and x^R is x written in reverse $\}$

a) Describe the strings in L in English.
b) Show that L is regular by giving a regular expression for L.

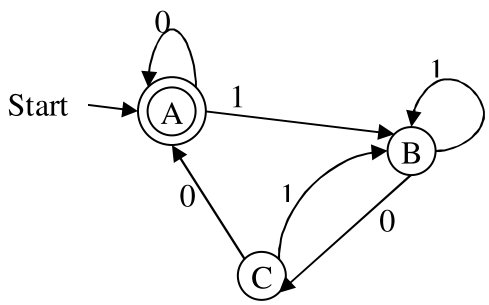
4-5 Referring to exercise 3-3 , if $L = 0^+10^*$, write a regular expression that captures subwords(L).

4-6 (2.12) Construct finite automata equivalent to the following regular expressions :

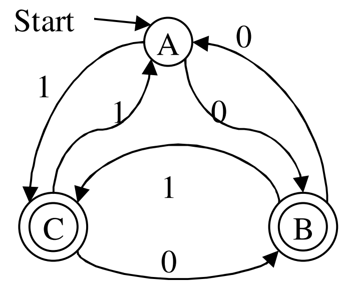
a) 10+(0+11)0*1
b) 01[((10)*+111)*+0]*1
c) ((0+1)(0+1))*+((0+1)(0+1)(0+1))*

4-7 (2.13) Construct regular expressions corresponding to the state diagrams given below :

a)



b)



4-8 Referring to exercise 2-2 , write a regular expression that describes L.

4-9 (2.24) Give Mealy and Moore machines for the following processes :

a) For input from (0+1)*, if the input ends in 101, output A; if the input ends in 110, output B; otherwise, output C
b) For input from (0+1+2)* print the remainder modulo 5 of the input treated as a ternary (base 3, with digits 0,1,2) number.

4-10 a) [1.7e] b) [1.8b] c) [1.9a] d) [1.10b] e) [1.12] f) [1.18a,b] g) [1.21a,b]

Chapter 5 : Pumping Lemma for Regular Languages

5-1 A word w is a palindrome if it is its own reverse (that is, $w = w^R$). For example, 010 is a palindrome, but 001 is not. Let $\Sigma = \{0, 1\}$. Prove that the set of palindromes $P = \{ w \mid w \text{ in } \Sigma^* \text{ and } w = w^R\}$ is not regular using the Pumping Lemma.

5-2 Show that the language $L = \{\ wa^na^nw^R \mid w \in \{a, b\}^* \text{ and } n>0\ \}$ is not regular using the Pumping Lemma.

5-3 (3.1) Which of the following languages are regular sets? Prove your answer.

a) $\{ 0^{2n} \mid n \geq 1 \}$
b) $\{ 0^m1^n0^{m+n} \mid m \geq 1 \text{ and } n \geq 1 \}$
c) $\{ 0^n \mid n \text{ is a prime} \}$
d) the set of all strings that do not have three consecutive 0's.

g) $\{xwx^R \mid x, w \text{ in } (0+1)^+\}$. x^R is the string reversal of x.

5-4 a) [1.29b] b) [1.30]

Chapter 6 : Closure and Decision Problems for Regular Languages

6-1 (3.6) Show that $L=\{ 0^i1^j \mid \gcd(i,j)=1 \}$ is not regular. Hint: Consider $\overline{L} \cap 0^*1^*$.

6-2 (3.4) Let L be a regular set. Which of the following sets are regular? Justify your answers.

a) $\{ a_1a_3a_5 \ldots a_{2n-1} \mid a_1a_2a_3a_4\ldots a_{2n} \text{ is in } L \}$

b) $\{ a_2a_1a_4a_3a_6a_5\ldots a_{2n}a_{2n-1} \mid a_1a_2a_3a_4\ldots a_{2n} \text{ is in } L \}$

c) CYCLE(L) = $\{ x_1x_2 \mid x_2x_1 \text{ is in L for strings } x_1 \text{ and } x_2 \}$

d) MAX(L) = { x in L | for no y other than ε is xy in L }

e) MIN(L) = { x in L | no proper prefix of x is in L }

f) INIT(L) = { x | for some y, xy is in L }

g) $L^R = \{ x \mid x^R \text{ is in } L \}$ (This is an example in the lecture. Try to do this by NFA.)

h) $\{x \mid xx^R \text{ is in } L \}$

6-3 Are the following languages regular or not regular ? Prove your answer.

a) $L_1 = \{ 0^m1^n \mid m \neq 2n \}$ Hint: Consider $\overline{L} \cap 0^*1^*$.

b) $L_2 = \{ w \mid w \in \{a, b\}^* \text{ and } w \neq w^R \}$ Hint: Consider $\overline{L}$.

c) (3.1h) $L_3 = \{xx^Rw \mid x, w \text{ in } (0+1)^+ \}$. x^R is the string reversal of x. (Note: very tricky.)

Hint: Consider $L_3 \cap 10^*110^*11$

6-4 True or False: Give a short justification (one or two sentences).

Regular languages are closed under the operation of set difference.

6-5 Let $\Sigma = \{a, b\}$. Let L1 be the set of words w that contain an even number of a's. Let L2 be the set of words w that end with b. Let L3 = L1 $\cap$ L2.

a) Draw a DFA A for L1 with 2 states.

b) Draw a DFA B for L2 with 2 states.

c) Draw a DFA C for L3 with 4 states using the product construction.

6-6 a) [1.4a,c,g] b) [1.5c,e,g]

Chapter 7 : DFA Minimization

7-1 Continuing exercise 6-5 ,

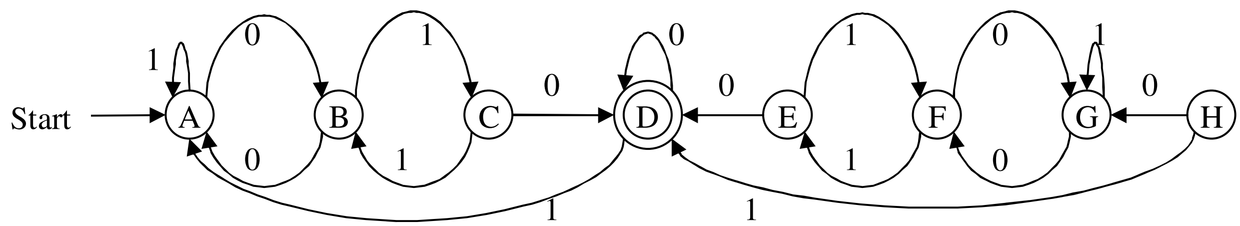
d) Run the DFA minimization algorithm to determine the distinguishable states of DFA C.

e) Draw a DFA D for L3 with the minimal number of states.

7-2 True or False: Give a short justification (one or two sentences).

If A and B are DFAs such that L(A) = L(B) then A and B must have the same number of states.

7-3 (3.25) a) Find the minimum-state finite automaton equivalent to the transition diagram below :



b*) Repeat part a) assuming H is the start state instead.

7-4 (3.26)

a) What are the equivalence classes of R_L in the Myhill-Nerode theorem for

$L = \{ 0^n1^n \mid n \geq 1 \}$?

b) Use your answer in (a) to show $L = \{ 0^n1^n \mid n \geq 1 \}$ is not regular.

c) Repeat (a) for { x | x has an equal number of 0's and 1's }.

Chapter 8 : Context-Free Grammars

8-1 (4.1) Give context-free grammars generating the following sets :

c) The set of all strings over alphabet {a,b} with exactly twice as many a's as b's.

e) The set of all strings over alphabet {a,b} not of the form ww for some string w.

f) $\{ a^ib^jc^k \mid i \neq j \text{ or } j \neq k \}$

8-2 (4.4) Construct a CFG generating the set

$\{ w\#w^R\# \mid w \text{ in } (0+1)^+ \}^*$.

8-3 Show that the language $L = \{ wa^na^nw^R \mid w \in \{a, b\}^* \text{ and } n>0 \}$ is context free by giving a context free grammar for L.

8-4 (4.8) Let G be the grammar

$S \rightarrow aB \mid bA$
$A \rightarrow a \mid aS \mid bAA$
$B \rightarrow b \mid bS \mid aBB$

For the string aaabbabbba find a

a) leftmost derivation b) rightmost derivation c) parse tree

8-5 (4.9) Is the grammar in exercise (4.8) unambiguous ?

8-6 Consider the following CFG G whose only variable E :

$E \rightarrow E \wedge E \mid E \vee E \mid (E) \mid p \mid q$

Show two different parse trees for the string $p \wedge q \vee p$

8-7 (4.5) The grammar

$E \rightarrow E + E \mid E * E \mid (E) \mid id$

generates the set of arithmetic expressions with +, *, parentheses and id. The grammar is ambiguous since id + id * id can be generated by two distinct leftmost derivations.

a) Construct an equivalent unambiguous grammar.

b) Construct an unambiguous grammar for all arithmetic expressions with no redundant parentheses. A set of parentheses is redundant if its removal does not change the expression, e.g., the parentheses are redundant in id + (id * id) but not in (id + id) * id.

8-8 Consider the grammar G with a single variable R, which is also the start variable, over the alphabet { 0, 1, +, ∗, (,) } :

$$R \rightarrow 0 \mid 1 \mid R* \mid R + R \mid (R)$$

a) Show a left-most derivation of the word 0+(1∗+0)
b) Show a parse tree for the word (0+1)∗+1
c) Is G ambiguous? Justify your answer in one or two sentences, no proof is needed.
d) If L(G) regular? Justify your answer in one or two sentences, no proof is needed.

8-9 Construct a context free grammar equivalent to the regular expression (a+bab)*(b+a)

8-10 a) [2.1c,d] b) [2.4b,c,f] c) [2.10]

Chapter 9 : CFG Normal Forms

9-1 (4.10) Find a CFG and with no useless symbols equivalent to

$S \rightarrow AB \mid CA$ $\quad B \rightarrow BC \mid AB$
$A \rightarrow a$ $\quad C \rightarrow aB \mid b$

9-2 Referring to the grammar G described in exercise 8-8 , is G in Chomsky Normal Form?

9-3 Convert the following grammar to Chomsky Normal Form.

$S \rightarrow A \mid B$ $\quad C \rightarrow pqrW$ $\quad T \rightarrow t \mid \varepsilon$
$A \rightarrow aA \mid aC$ $\quad W \rightarrow TV$ $\quad V \rightarrow v \mid \varepsilon$
$B \rightarrow bB \mid bC$

9-4 (4.11) Suppose G is a CFG and w, of length p, is in L(G). How long is a derivation of w in G if
a) G is in CNF,
b) G is in GNF.

9-5 (4.12) Let G be the CFG generating well-formed formulas of propositional calculus with predicates p and q :

$S \rightarrow \sim S \mid [S \supset S] \mid p \mid q$

The terminals are p, q, ~, [,], and ⊃. Find a Chomsky normal-form grammar generating L(G).

9-6 (4.14) Find a Greibach normal-form grammar equivalent to the following CFG :

$S \rightarrow AA \mid 0$
$A \rightarrow SS \mid 1$

9-7 a) [2.15] b) [2.26]

Chapter 10 : Pushdown Automata

10-1 Construct a pushdown automaton for each of the following languages :

a) The set of palindromes $\{w \in \{a, b\}^* \mid w = w^R\}$

b) The set of all strings of balanced and properly nested parentheses.

c) The set of all strings over alphabet {a,b} with exactly twice as many a's as b's.

d) $\{ a^i b^j c^k \mid i \neq j \text{ or } j \neq k \}$

10-2 (5.2) Construct a PDA equivalent to the following grammar :

$S \rightarrow aAA \qquad A \rightarrow aS \mid bS \mid a$

10-3 Consider the following CFG G whose only variable E :

$E \rightarrow E \wedge E \mid E \vee E \mid (E) \mid p \mid q$

Convert the CFG above into a PDA that accepts L(G)

10-4 Give a PDA for the following language:

$L = \{ \mathbf{0^i 1 0^j 1 0^k} \mid k < i + j \}$

Your PDA may use acceptance by final state or acceptance by empty stack, make sure to state your choice clearly. There is no need to give a proof that your PDA indeed accepts the above language, but explain how your PDA works in a few sentences.

10-5 Consider the language

$L = \{ w \in \{a, b\}^* \mid |w| \geq 1$ and the first, middle and last symbols in w are the same$\}$

a) Give a state diagram of a pushdown automaton that recognizes L. Do not use the conversion algorithm to convert your CFG in b) to a PDA.

b) Define a context-free grammar that generates L.

10-6 (5.6) Give a grammar for the language N(M) where

$M = (\{q_0, q_1\}, \{0, 1\}, \{\$, X\}, \delta, q_0, \$, \varnothing)$

and δ is given by

$\delta(q_0, 1, \$) = \{(q_0, X\$)\}$, $\quad \delta(q_0, \varepsilon, \$) = \{(q_0, \varepsilon)\}$,

$\delta(q_0, 1, X) = \{(q_0, XX)\}$, $\quad \delta(q_1, 1, X) = \{(q_1, \varepsilon)\}$,

$\delta(q_0, 0, X) = \{(q_1, X)\}$, $\quad \delta(q_1, 0, \$) = \{(q_0, \$)\}$

10-7 a) [2.5b,c,f] b) [2.11] c) [2.12]

Chapter 11 : Pumping Lemma for CFLs

11-1 (6.1) Show that the following are not context-free languages:

a) $\{ a^i b^j c^k \mid i<j<k \}$

b) $\{ a^i b^j \mid j = i^2 \}$

c) $\{ a^i \mid i \text{ is a prime} \}$

d) the set of strings of a's, b's and c's with an equal number of each

e) $\{ a^n b^n c^m \mid n \leq m \leq 2n \}$

11-2 Prove that the following languages are not CFLs using the pumping lemma for CFLs :

a) $L_1 = \{ \mathbf{a^i b^j c^k} \mid i = k \text{ and } j > i \}$

b) $L_2 = \{ \mathbf{1^k} \mid k=n^2 \text{ and } n \geq 0 \}$

c) $L_3 = \{ \mathbf{a^i\, b^j\, c^i\, d^j} \mid i, j \geq 0 \}$

11-3 (6.3) Prove that the following are not CFL's :

a) $\{ a^i b^j a^k \mid j = \max \{i,k\} \}$

b) $\{ a^n b^n c^i \mid i \neq n \}$

[Hint: Use Ogden's Lemma on a string of the form $a^n b^n c^{n!}$]

11-4 a) [2.30a] b) [2.31]

Chapter 12 : Closure and Decision Problems for CFLs

12-1 Prove that the context-free languages are closed

a) under union and

b) under concatenation.

Don't just copy your class notes - this should be good practice for proving a closure property on the exam.

12-2 (6.2) Which of the following are CFL's ?

a) $\{ a^i b^j \mid i \neq j \text{ and } i \neq 2j \}$

c) $\{ ww^R w \mid w \text{ is in } (a + b)^* \}$

d) $\{ b_i \# b_{i+1} \mid b_i \text{ is } i \text{ in binary}, i \geq 1 \}$

e) $\{ wxw \mid w \text{ and } x \text{ are in } (a+b)^* \}$

12-3 For each of the following languages, determine whether it is (I) regular, (II) not regular but context-free, or (III) not context-free. Prove your answer.

a) $L_1 = \{ \mathbf{a^i b^j c^k} \mid i < j < k \}$

b) $L_2 = \{ \mathbf{a^i\, b^i\, c^j\, d^j} \mid i, j \geq 0 \}$

c) $L_3 = \{ \mathbf{w} \mid \mathbf{w} = \mathbf{a^i\, b^i\, c^j\, d^j}$ where $i, j \geq 0$ and
$\mathbf{w}$ does not contain the substring $\mathbf{bccd} \}$

d) $L_4 = \{ \mathbf{a^i\, b^j} \mid 2i \neq 3j \}$

e) $L_5 = \{ \mathbf{0^n\, 0^{2n}\, 0^{3n}} \mid n \geq 0 \}$

f) $L_6 = \{ \mathbf{xy} \mid \mathbf{x, y} \in \{0, 1\}^* \text{ and } \mathbf{x} \neq \mathbf{y} \}$

12-4 Prove your answer to the following questions :

a) If L_1 is a CFL and $L_2 \subseteq L_1$, is it necessarily true that L_2 is a CFL ?

b) If L_1 is a CFL and $L_2 \subseteq L_1$, can L_2 be a regular language ?

12-5 (6.17) Use the CYK algorithm to determine whether

a) aaaaa b) aaaaaa

are in the grammar of Example 6.7

$S \rightarrow AB \mid BC \qquad B \rightarrow CC \mid b$

$A \rightarrow BA \mid a \qquad C \rightarrow AB \mid a$

12-6 a) [2.6b] b) [2.35]

Chapter 13 : Turing Machines

13-1 (7.1) Design Turing machines to recognize the following languages :

a) { $0^n1^n0^n$ | n≥1 }

b) { ww^R | w is in (0+1)*}

c) The set of strings with an equal number of 0's and 1's.

13-2 Consider the Turing machine with states $\{q_0, q_1, q_2, q_f\}$, input alphabet {0, 1}, tape alphabet {0, 1, B}, initial state q_0, final state q_f, blank symbol B, and transition function given by:

$\delta(q_0, 0) = (q_1, 1, R)$;
$\delta(q_1, 1) = (q_2, 0, L)$;
$\delta(q_2, 1) = (q_0, 1, R)$;
$\delta(q_1, B) = (q_f, B, R)$

Clearly describe the language accepted by this machine.

13-3 Let $\Sigma = \{0, 1\}$. Consider the language

$L = \{ \mathbf{0^n 10^{2n} 10^{3n}} \mid n \geq 0 \}$

Design a standard (deterministic, single-tape) Turing Machine M that accepts L by drawing the transition diagram with accompanying description of how the machine works.

13-4 Describe a standard (deterministic, single-tape) Turing machine that accepts the following language:

$L = \{ \mathbf{0^p} \mid p \text{ is not a prime} \}$

13-5 a) [3.2b,d] b) [3.8b]

Chapter 14 : Turing Machine Variants

14-1 (7.2) Design Turing machines to compute the following functions

a) $\lceil \log_2 n \rceil$ b) n! c) n^2

14-2 (7.3) Show that if L is accepted by a k-tape, p-dimensional, nondeterministic TM with m heads per tape, then L is accepted by a deterministic TM with one semi-infinite tape and one tape head.

14-3 (7.8) Design a Turing machine to enumerate { 0^n1^n | n≥1 }

14-4 A k-head (deterministic) Turing machine has k heads reading cells of a single tape. A move depends on the current state and the symbols read by the heads. In one move, the machine can update the state, write a new symbol on each of cells scanned by the heads, and move each head left, right, or keep it stationary.

Since several heads may be scanning the same cell, assume that the heads are numbered 1 through k, and the symbol written by the highest numbered head scanning a given cell is the one that is actually written in that cell. Prove that the languages accepted by k-head TMs are the same as those accepted by standard TMs.

14-5 a) [3.9]

Chapter 15 : Undecidable Languages

15-1 (8.1) Suppose the tape alphabets of all Turing machines are selected from some infinite set of symbols a_1, a_2, a_3, … . Show how each TM may be encoded as a binary string.

15-2 (8.2) Which of the following properties of recursively enumerable sets are themselves recursively enumerable ?

a) L contains at least two strings

b) L is infinite

c) L is a context-free language

d) $L = L^R$

15-3 (8.3) Show that it is undecidable whether a TM halts on all inputs.

15-4 (8.7) Show that the following problems about programs in a real programming language are undecidable :

a) Whether a given program can loop forever on some input

b) Whether a given program ever produces an output

c) Whether two programs produce the same output on all inputs

15-5 a) [3.15] b) [3.16] c) [4.3] d) [4.4] e) [4.10] f) [5.1] g) [5.4] h) [5.9]