SOLUTIONS HW#3

1. Like in the hint, let us consider $M = (Q, \Sigma, \delta, s, F)$ the DFA for A. Let's give first the informal description of the new NFA for the MiddleThirds in terms of pebbles. We start with 5 pebbles, 1,2,3,4,5. Pebble 1 is at the start state, 2,3 and 4,5 in pairs on some nondeterministically chosen states of M. In each step we move pebble 3 according to the input symbol and we move pebbles 1 and 5 according to some nondeterministically chosen symbol. Pebbles 2 and 4 never move. When the end of y is reached we accept if 1 and 2 will be on the same state, also 3 and 4 should be on the same state and pebble 5 should be in an accept state.

Let's do now formally. We define the NFA $M' = (Q', \Sigma, \Delta, S', F')$ as follows:

$Q' = Q^5$
$S' = \{(s, t, t, u, u) \mid t, u \in Q\}$
$F' = \{(v, v, w, w, f) \mid v, w \in Q, f \in F\}$
$\Delta((p, q, r, t, u), a) = \{(\delta(p, b), q, \delta(r, a), t, \delta(u, c)) \mid b, c \in \Sigma\}$

It is easy to prove now the relation for the strings by induction:

$\Delta(S', y) = \{ (\delta(s, x), q, \delta(q, y), t, \delta(t, z) \mid q, t \in Q, x, z \in \Sigma^{|y|}\}$

Using this we can prove now that L(M')=MiddleThirds L(M):

$y \in L(M') \Leftrightarrow$
$\Delta(S', y) \cap F' \neq \emptyset \Leftrightarrow$
$\{(\delta(s, x), q, \delta(q, y), t, \delta(t, z) \mid q, t \in Q, x, z \in \Sigma^{|y|}\} \cap \{(v, v, w, w, f) \mid v, w \in Q, f \in F\} \neq \emptyset \Leftrightarrow$
$\exists\, x, z \in \Sigma^{|y|}, q, t \in Q$ such that $q = \delta(s, x)$, $t = \delta(q, y)$ and $\delta(t, z) \in F \Leftrightarrow$
$\exists\, x, z \in \Sigma^{|y|}\ \delta(\delta(\delta(s, x), y), z) \in F \Leftrightarrow$
$\exists\, x, z \in \Sigma^{|y|}\ \delta(s, xyz) \in F \Leftrightarrow$
$\exists\, x, z \in \Sigma^{|y|}\ xyz \in L(M) \Leftrightarrow$
$y \in$ MiddleThirds L(M)

This ends the proof of our problem.

2. (a) First automaton: 1, 2, 3, 4, 5, 6 are the accessible states, 7 and 8 are the inaccessible states. For the second automaton all the states are accessible.

(b) For the first automaton we have the following table:

```
1
#  2
#  #  3
#  #  -  4
#  -  #  #  5
-  #  #  #  #  6
```

So looking at the unmarked pairs we get that 3=4, 2=5, 1=6.

For the second automaton we have:

```
1
-  2
#  #  3
#  #  -  4
#  #  #  #  5
#  #  #  #  -  6
#  #  #  #  -  -  7
#  #  -  -  #  #  #  8
```

We get that the equivalence classes are: [1]={1, 2}, [3]={3, 4, 8}, [5]={5, 6, 7}.

(c) The two automata are: (see additional page at the end)

3. Q={1,2,3,4,5,6,7,8}.

$L(M)=a^{Q}_{13}+ a^{Q}_{14}$
$a^{Q}_{13}= a^{Q-\{4\}}_{13}+ a^{Q-\{4\}}_{14}(a^{Q-\{4\}}_{44})^* a^{Q-\{4\}}_{43}$

$a^{Q-\{4\}}_{13}=\emptyset$
$a^{Q-\{4\}}_{14}= a^{Q-\{4,3\}}_{14}+ a^{Q-\{4,3\}}_{13}(a^{Q-\{4,3\}}_{33})^* a^{Q-\{4,3\}}_{34}=a^*b$
$a^{Q-\{4\}}_{44}= a^{Q-\{4,3\}}_{44}+ a^{Q-\{4,3\}}_{43}(a^{Q-\{4,3\}}_{33})^* a^{Q-\{4,3\}}_{34}= (e+ba)+(a+bba^*b)(ba)^*(a+bba^*b)$
$a^{Q-\{4\}}_{43}= a^{Q-\{4,3\}}_{43}+ a^{Q-\{4,3\}}_{43}(a^{Q-\{4,3\}}_{33})^* a^{Q-\{4,3\}}_{33}=(a+bba^*b)+(a+bba^*b)(ba)^*ba$

So $a^{Q}_{13}=\gamma(e +\beta+\delta\beta^*\delta)(\delta+\delta\beta^*\beta)$

$a^{Q}_{14}= a^{Q-\{4\}}_{14}+ a^{Q-\{4\}}_{14}(a^{Q-\{4\}}_{44})^* a^{Q-\{4\}}_{44}$

$a^{Q-\{4\}}_{14}=a^*b$
$a^{Q-\{4\}}_{44}= (e +ba)+(a+bba^*b)(ba)^*(a+bba^*b)$

This gives us: $a^{Q}_{14}=\gamma+\gamma(e+\beta+\delta\beta^*\delta)^*(e+\beta+\delta\beta^*\delta)$

So:

$$L(M)=[\gamma(e+\beta+\delta\beta^*\delta)(\delta+\delta\beta^*\beta)]+[\gamma+\gamma(e+\beta+\delta\beta^*\delta)^*(e+\beta+\delta\beta^*\delta)]$$

4. i) We have three equivalence classes corresponding to [a], [b] and [ab].

[a]={w in {a,b}* | w ends in a}

[b]={w in {a,b}* | w is e or it ends in bb}

[ab]={w in {a,b}* | w ends in ab}

We observe that the three sets are disjoint and their union is {a,b}*. The corresponding automaton is: (see additional page)

ii) The set is nonregular, so the number of equivalence classes is infinite.

If k < m, then a^k not equivalent with a^m since $a^k b^k \in R$ but $a^m b^k \notin R$.

The equivalence classes are:

$G_k=[a^k]=\{a^k\}$, k>=0, and we can append all strings in $\{a^m b^{m+n} | m>=0, n>=k\}$

$H_k=[a^k b]=\{a^{n+k} b^n, n>0\}$, k>=0, and we can append all the strings in $b^k\{b\}^*$

F=[ba]={all the other strings}, and no string we can append to a string in F to obtain a string in R.

iii) The corresponding automaton is: (see additional page)

We observe that we have three states, so if we can find three strings that are not equivalent, that means that this is the minimal automaton and we have exactly three equivalence classes. The three strings are e, a and b.

5. Let us denote by #a(x) and #b(x) the number of a's and b's in the string x.

Let's consider $A=\{x\in\{a,b\}^* | \#a(x) = \#b(x)\}$ and G be the *CFG*:

$S\rightarrow aSb \mid bSa \mid SS \mid e$

We want to show L(G)=A. We will prove this by double inclusion and induction.

We prove first that $L(G)\subseteq A$ by induction on the length of the derivation.

Basis:

It is trivial since the start symbol satisfies the condition #a(x) = #b(x).

Induction step:

We know that if $S\rightarrow a$ then $S\rightarrow\beta\rightarrow a$. From the induction hypotheses we know that $\#a(\beta) = \#b(\beta)$. Also there exist $\beta_1, \beta_2 \in (N\cup\Sigma)^*$ such that $\beta=\beta_1 S\beta_2$ and a is one of the following four expressions: $\beta_1 aSb\beta_2$, $\beta_1 bSa\beta_2$, $\beta_1 SS\beta_2$, $\beta_1\beta_2$. But after we apply this operation, we will still have #a(a) = #b(a) because either both will remain unchanged or both values will increase by 1.

This proves that $L(G) \subseteq A$.

Let's prove now the reverse implication, $A \subseteq L(G)$, also by induction on $|x|$.

Basis:
If $|x|=0$, we have $x=e$ and $S \rightarrow x$ in one step, since $S \rightarrow e$.

Induction step:
Let $|x|>0$ be a string in A. We have two situations:

(a) There exists a proper prefix y of x such that $0<|y|<|x|$ and #a(y) = #b(y). This implies that x=yz, with both y, $z \in A$. By induction, $S \rightarrow y$ and $S \rightarrow z$, so

$$S \rightarrow SS \rightarrow yS \rightarrow yz=x.$$

(b) No such prefix exists, so in this case x=ayb or x=bya for some string $y \in A$. This implies:

$$S \rightarrow aSb \rightarrow ayb=x, \text{ by induction.}$$

So we proved the double inclusion and the proof is done.