CS2MJ3 – Fall 2012. Sample solutions to the assignment 1. Most questions have more than one solution. Total of this assignment is 153pts. 100% of this assignment is 125pts. There are 28 bonus pts. Each assignment is worth 7%.

If you think your solution has been marked wrongly, write a short memo stating where marking in wrong and what you think is right, and resubmit to me during class, office hours, or just slip under the door to my office.

1.[5] Let $M = (Q, \Sigma, \delta, s_0, F)$ be a deterministic finite automaton. Give a deterministic finite state automaton M' such that $L(M') = L(M) - \{\varepsilon\}$.

Solution:

Let $M=(Q,\Sigma,\delta,s_0,F)$ be a DFA such that $L=L(M)$. If $\varepsilon \notin L(M)$, then $M'=M$. If $\varepsilon \in L(M)$ then $s_0 \in F$. Define $M'= (Q\cup\{q_0\},\Sigma,\delta_1,q_0,F)$, where $q_0 \notin Q$, and δ_1 is defined as follows:

$\delta_1(q_0,a) = \delta(s_0,a)$ for each $a\in\Sigma$, $\delta_1(s,a) = \delta(s,a)$ for each $s\in Q$.

Clearly $L(M')=L(M)-\{\varepsilon\}$.

2.[5] Let $L = \{a,b\}$ and $L_1^* = L^*$. Prove that $L \subseteq L_1$.

Solution:

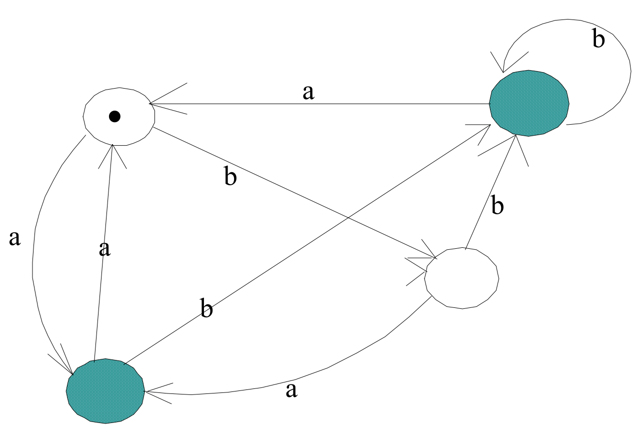
First we prove that $a \in L_1$. Suppose that $a \notin L_1$. This means that $a \notin L_1^*$, a contradiction since $a \in L \subseteq L^* = L_1^*$. Identically we prove $b \in L_1$. Thus $L \subseteq L_1$.

An immediate extension of this result (identical proof!) is:

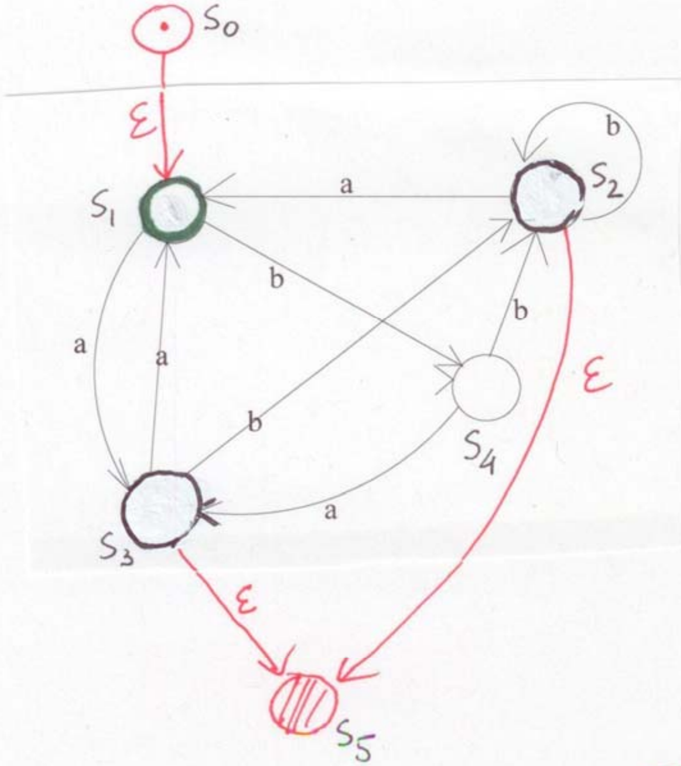
for every *alphabet* Σ and any language L, if $L^* = \Sigma^*$, then $\Sigma \subseteq L$.

3.[10] Consider the automaton below.

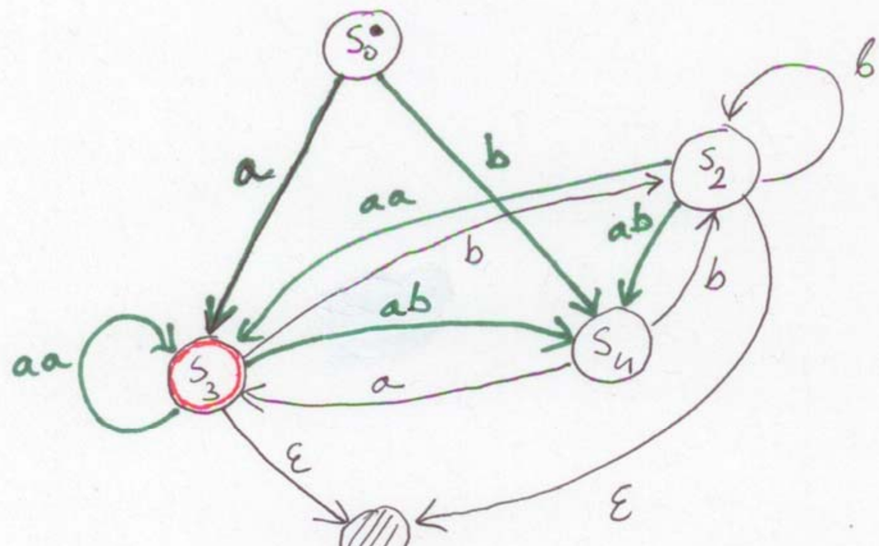
Construct an equivalent regular expression using the idea of generalised nondeterministic automata. Provide all steps.



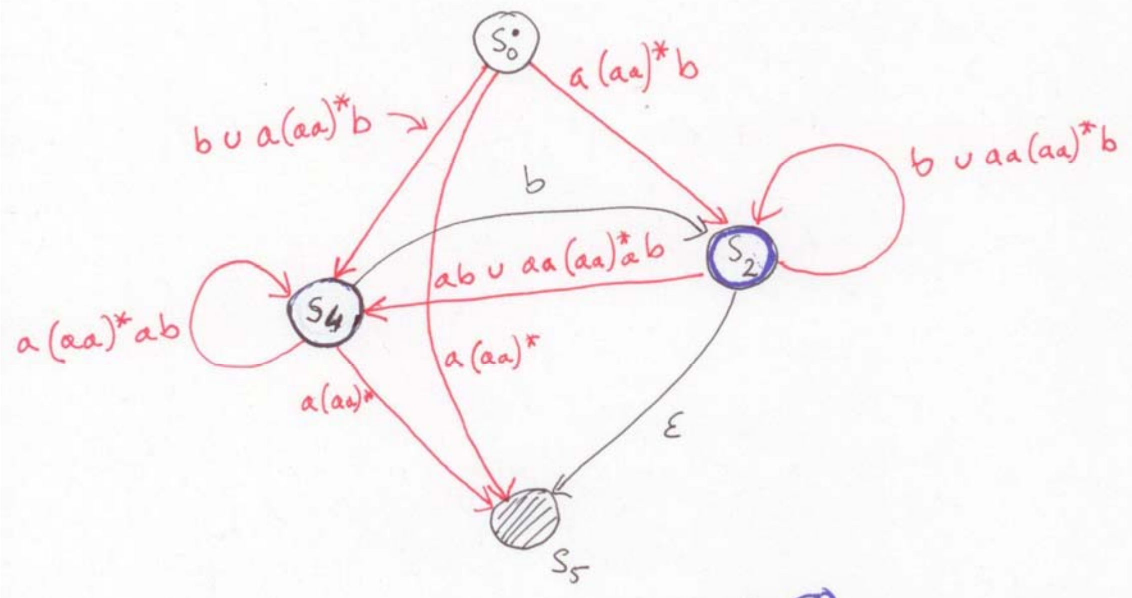
Solution:



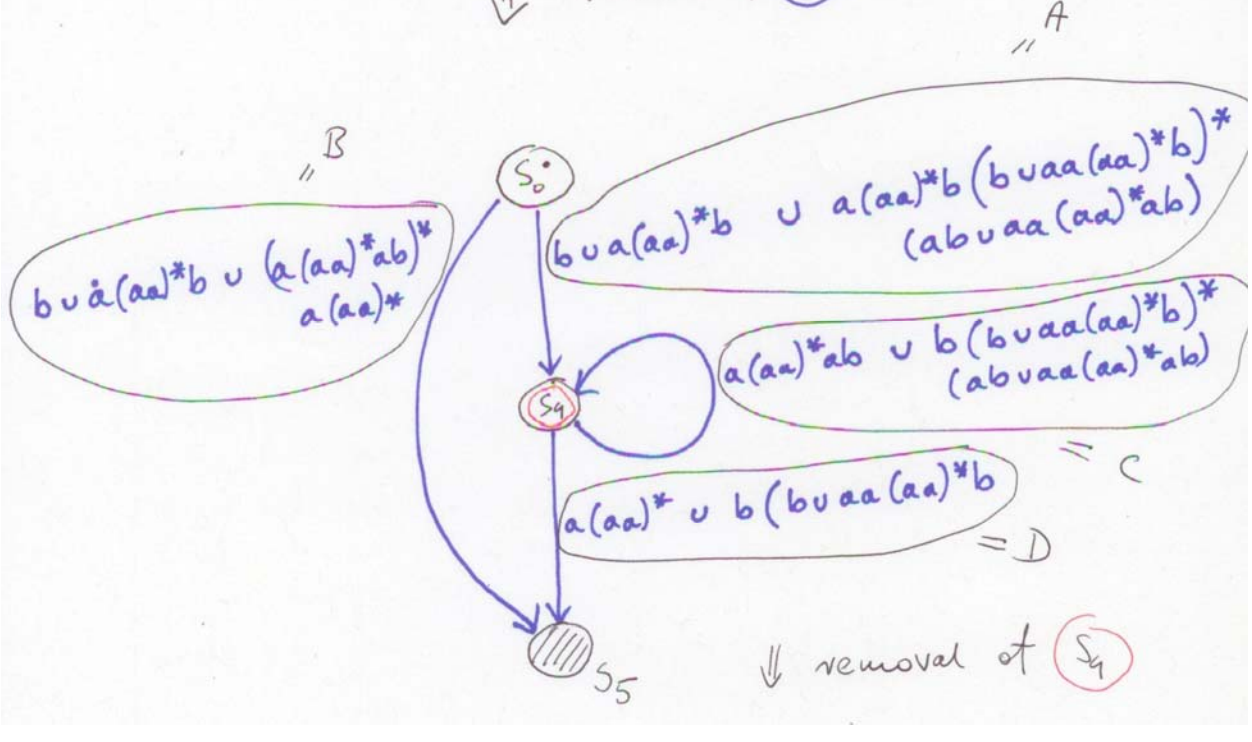
⇓ removal of S1

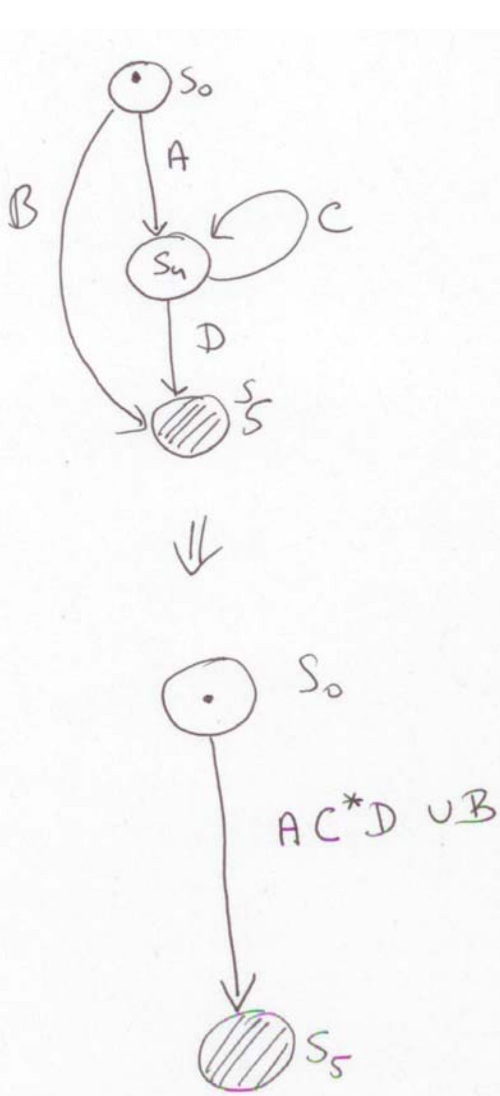


⇓ removal of S3



⇓ removal of s2





This is not the only solution. It may be simplified using rules of regular expressions algebra, but this subject was not discussed in class.

4.[15] Consider the following regular expression: a*b*∪(a*bb∪ab*)a*b.

a. Using a standard construction construct an equivalent non-deterministic finite state automaton.

b. Transform the automaton from 4a into an equivalent deterministic finite state

Solution:

a.[5] Step 1 [5]. We construct a nondeterministic automaton with ε- moves: Q={1,2,3,4,5,6,7,8,9,10,11,12,13,14,15,16,17,18,19,20,21,22}, initial_state=1, F={22}, Σ={a,b}, and δ is defined below

δ	a	b	ε
1			{2,8}
2	{3}		{4}
3			{2}
4		{5}	{22}
5			{4}
6	{7}		{9}
7			{6}
8			{6,11}
9		{10}	
10			{13}
11	{12}		
12			{14}
13		{15}	
14		{16}	{17}
15			{17}
16			{14}
17			{18}
18	{19}		{20}
19			{18}
20		{21}	
21			{22}
22			

Step 2[5]. Now we get rid of ε- moves:
Q={1,2,3,4,5,6,7,8,9}, initial_state=1, F={1,2,4,6,9}, and δ is defined below

δ	a	b
1	{2,7}	{3}
2	{2}	{3,4}
3		{5}
4		{4}
5	{5}	{6}
6		
7	{8}	{7,9}
8	{8}	{9}
9		

Step 3[5]. We now transform non deterministic automaton into deterministic one.

Q={1,2,3,4,5,6,7,8,9,10,11,12,13,14,15,16,17,18,19}, initial_state=1,
F={1,2,4,5,6,9,10,12,13,14,15,17,18}, Σ={a,b}, garbage=19 and δ is defined below

δ	a	b
1	2	3
2	5	4
3	19	8
4	7	6
5	5	12
6	11	10
7	7	9
8	8	16
9	19	19
10	7	13
11	11	14
12	19	15
11	11	14
13	7	13
14	19	19
15	8	17
16	19	19
17	19	18
18	19	18
19	19	19

5.[19] a. Is a language $\{ a^n b^k c^{n+k} \mid n \geq 0, k \geq 0 \}$ regular? Prove your answer.

b. Is the language $\{a^{2m+3n+2} \mid m,n \geq 0\}$ regular? Prove your answer.

c. Is the language $\{a^{n!} \mid n \geq 0\}$ regular? Prove your answer.

Solution:

a.[5] L is not regular.
Suppose it is regular, so Pumping Lemma hold. Let n the number from Pumping Lemma, and let $x = a^n b^k c^{n+k}$. Clearly $|x| \geq n$. Let $x = uvw$, $|uv| \leq n$, $|v| \geq 1$. Since $|uv| \leq n$, $v = a^i$, for some i, where $1 \leq i \leq n$.
Hence, $y = uv^i w = a^{n+i} b^k c^{n+k} \in L$, a contradiction, $n+i+k > n+k$.

b.[5] L is regular as L=(aa)*(aaa)*aa.

c.[5] L is not regular.
Suppose it is regular, so Pumping Lemma hold. Let n the number from Pumping Lemma, and let $x = a^{n!}$. Clearly $|x| = n! \geq n$. Let $x = uvw$, $|uv| \leq n$, $|v| \geq 1$. Let $y = uv^2w$.
Since $|uv| \leq n$ then $|v| \leq n$. Hence $|y| = |uvw| + |v| \leq n! + n$.
However the shortest string in L, which is longer than x is $a^{(n+1)!}$, and
$|a^{(n+1)!}| = (n+1)! = (n+1)n! > n! + n$, so $y \notin L$, a contradiction.

6[5]. Eliminate all useless productions from the below grammar. Give all the steps.

$S \rightarrow a \mid aA \mid B \mid C$
$A \rightarrow aB \mid \varepsilon$
$B \rightarrow Aa$
$C \rightarrow cCD$
$D \rightarrow ddd$

What language does this grammar generate?

Solution:

Step 1. There is only one type 1 useless symbol, namely: C
After erasing it we get:

$S \rightarrow a \mid aA \mid B$
$A \rightarrow aB \mid \varepsilon$
$B \rightarrow Aa$
$D \rightarrow ddd$

Step 2. There is only one type 2 useless symbol, namely: D
After erasing it we get:

$S \rightarrow a \mid aA \mid B$
$A \rightarrow aB \mid \varepsilon$
$B \rightarrow Aa$

The grammar generates the language with odd number of a's, i.e. $\{ a^n \mid n \text{ is odd} \}$

7[5]. Eliminate all ε-productions from

$S \rightarrow ABaC$
$A \rightarrow BC$
$B \rightarrow b \mid \varepsilon$
$C \rightarrow D \mid \varepsilon$
$D \rightarrow d$

Give all the steps.

Solution:

In the first step we conclude that B, C can eventually generate ε, *so they are nullable*.
Then, since $A \rightarrow BC$ we conclude that A is also nullable. It is easy to observe that only A,B,C are nullable. Hence:

$S \rightarrow ABaC \mid BaC \mid AaC \mid ABa \mid aC \mid Ba \mid Aa \mid a$
$A \rightarrow BC \mid B \mid C$
$B \rightarrow b$
$C \rightarrow D$
$D \rightarrow d$

8.[5] Eliminate all unit-productions from

$S \rightarrow Aa \mid N$
$A \rightarrow a \mid bc \mid B$
$B \rightarrow A \mid bb$

Give all the steps.

Solution:

$S \rightarrow N$ is useless. We have $A \rightarrow B$ and $B \rightarrow A$, so actually A is the same as B. Therefore we can change:

$A \rightarrow a \mid bc \mid B$
$B \rightarrow A \mid bb$

into just

$A \rightarrow a \mid bc \mid bb$

Hence the final grammar is:

$S \rightarrow Aa$
$A \rightarrow a \mid bc \mid bb$

9. [10] Convert the grammar

$S \rightarrow abAB$

$A \rightarrow bAB \mid \varepsilon$

$B \rightarrow BAa \mid A \mid \varepsilon$

into Chomsky normal form. Give all the steps.

Solution:

First we have to get rid of all ε-productions. The obtained grammar is:

$S \rightarrow abAB \mid abA \mid abB \mid ab$

$A \rightarrow bAB \mid bA \mid bB \mid b$

$B \rightarrow BAa \mid Aa \mid Ba \mid a \mid A$

Next we have to get rid of all unit-productions. The result is:

$S \rightarrow abAB \mid abA \mid abB \mid ab$

$A \rightarrow bAB \mid bA \mid bB \mid b$

$B \rightarrow BAa \mid Aa \mid Ba \mid a \mid bAB \mid bA \mid bB \mid b$

Now, we may produce a Chomsky normal form:

$S \rightarrow X_aX_{bAB} \mid X_aX_{bA} \mid X_aX_{bB} \mid X_aX_b$

$X_a \rightarrow a$

$X_b \rightarrow b$

$X_{bAB} \rightarrow X_aX_{AB}$

$X_{bA} \rightarrow X_bA$

$X_{bB} \rightarrow X_bB$

$X_{AB} \rightarrow AB$

$A \rightarrow X_BX_{AB} \mid X_bA \mid X_bB \mid b$

$B \rightarrow BAX_a \mid AX_a \mid BX_a \mid a \mid X_BX_{AB} \mid X_aA \mid X_bB \mid b$

10.[10] Convert the grammar

$S \rightarrow ABb \mid a$

$A \rightarrow aaA \mid b$

$B \rightarrow bAb \mid bBB \mid ABB \mid a$

into Greibach normal form.

Solution:

First substitute A by $aaA \mid b$ in $B \rightarrow ABB$. The result is

$B \rightarrow bAb \mid bBB \mid a \mid aaABB \mid bBB$,

i.e. $B \rightarrow bAb \mid bBB \mid a \mid aaABB$

Now we substitute A by $aaA \mid b$ in $S \rightarrow ABb$. The result is

$S \rightarrow aaABb \mid bAAb \mid a$

Hence now we have:

$S \rightarrow aaABb \mid bAAb \mid a$
$A \rightarrow aaA \mid b$
$B \rightarrow bAb \mid bBB \mid a \mid aaABB$

which is almost Graibach. We just need to add $C_a \rightarrow a$, $C_b \rightarrow b$, and replace a and b that are not first by C_a and C_b, respectively.

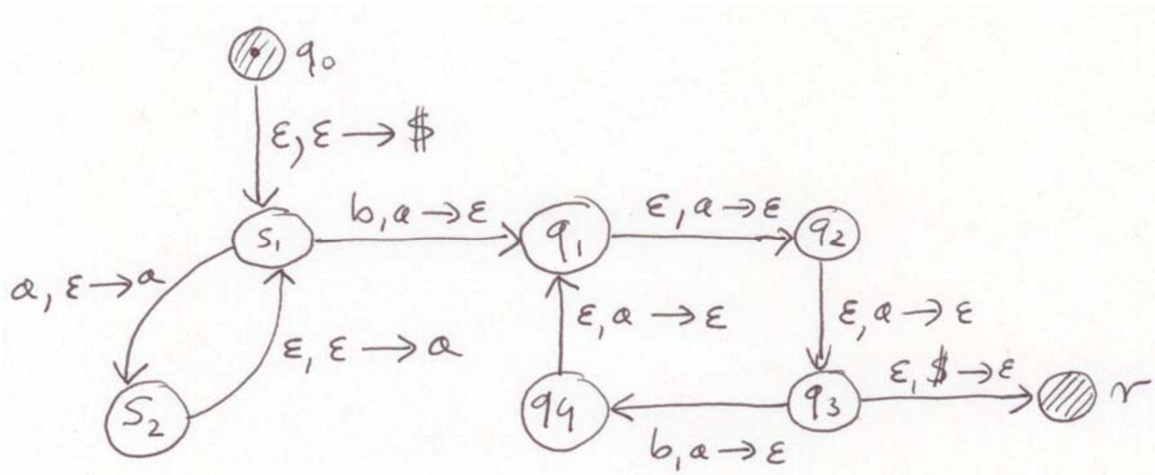
Final result:

$S \rightarrow aC_aABC_b \mid bAAC_b \mid a$
$A \rightarrow aC_aA \mid b$
$B \rightarrow bAC_b \mid bBB \mid a \mid aC_aABB$
$C_a \rightarrow a$
$C_b \rightarrow b$

11.[10]Construct a pushdown automaton accepting: $\{ a^n b^m \mid 2n = 3m, n \geq 0, m \geq 0 \}$.

Solution:

For a string $a^i b^j$ to be in $\{ a^n b^m \mid 2n = 3m, n \geq 0, m \geq 0 \}$, it must be true that i=3k and j=2k for some $k \geq 0$. Thus, for each symbol 'a' of the input, we push two copies of symbol 'a' into the stack (this can be done in two steps, using the states s_1, s_2). Then, for each symbol 'b' we read from the input, we match it with three a's in the stack (this has to be done in three steps, using the states q_1, q_2, q_3). The automaton looks as follows:



For an alternative definition of pushdown automaton (more popular, red stuff in class notes), we have:

$A = (Q, \Sigma, \Gamma, s, \$, \delta, F)$, where

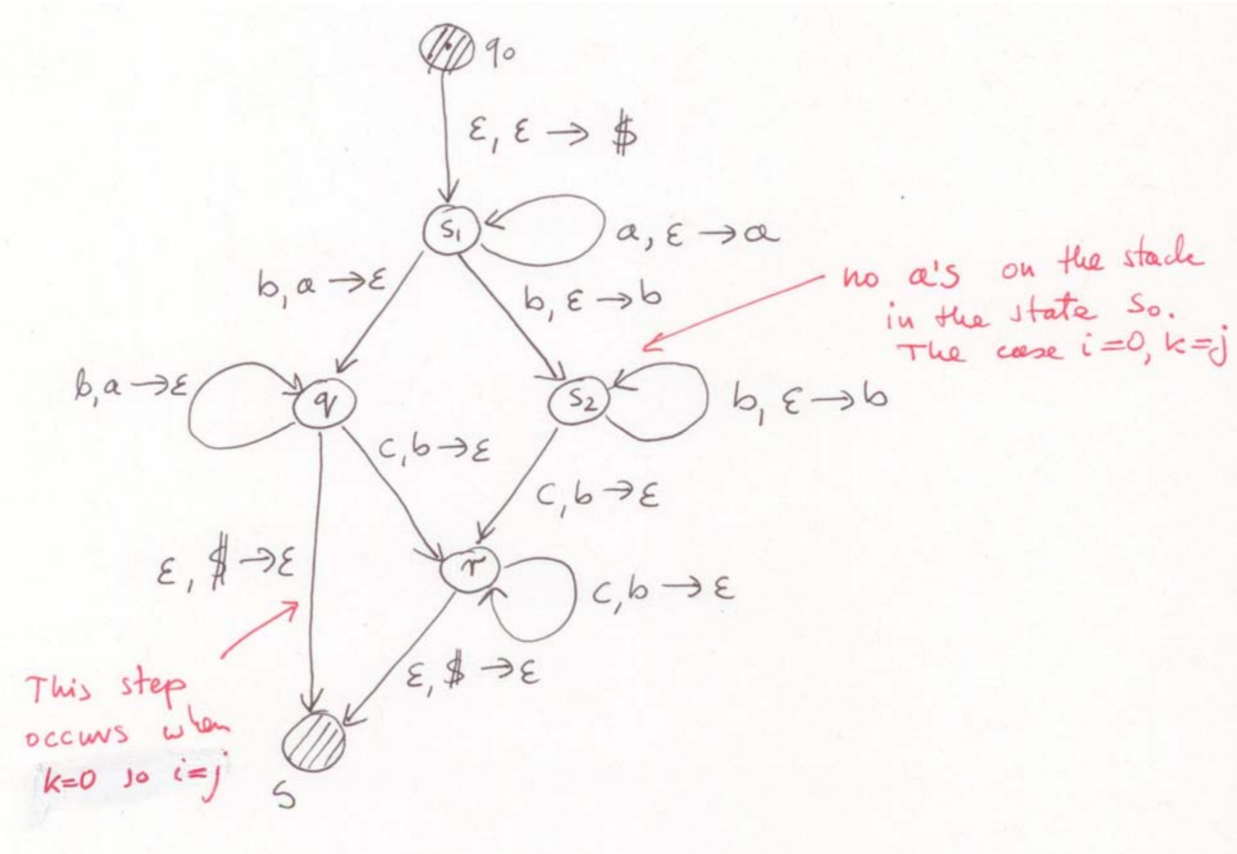
$Q = \{s, q_1, q_2, q_3, q_4, r\}$, $\Sigma=\{a,b\}$, $\Gamma= \{a,b,\$\}$, $F = \{r\}$ and δ is defined as follows:

$\delta(s,\varepsilon,\$) = \{ (r,\varepsilon)\}$ (just to accept empty string)
$\delta(s,a,\$) = \{(s,aa\$)\}$ (we assume here that the top of the stack is the first symbol on the left, so 'a' in this case)
$\delta(s,a,a) = \{(s,aa)\}$
$\delta(s,b,a) = \{(q_1, \varepsilon)\}$
$\delta(q_1,\varepsilon,a) = \{ (q_2,\varepsilon)\}$
$\delta(q_2,\varepsilon,a) = \{ (q_3,\varepsilon)\}$
$\delta(q_3,b,a) = \{ (q_4,\varepsilon)\}$
$\delta(q_3,\varepsilon,\$) = \{ (r,\varepsilon)\}$
$\delta(q_4,b,a) = \{ (q_1,\varepsilon)\}$

12.[10] Construct a pushdown automaton accepting: $\{ a^ib^jc^k \mid i,j,k\geq 0, i+k=j \}$.

Solution:

The basic idea for this problem is to use different states to enforce the order of the occurrences of the letters a, b and c, and use the stack to meet the requirement of i+k=j.



For an alternative definition of pushdown automaton (more popular, red stuff in class notes), we have:

$A = (Q, \Sigma, \Gamma, s, \$, \delta, F)$, where

$Q = \{s, q, p, r\}$, $\Sigma=\{a,b,c\}$, $\Gamma= \{a,b,c,\$\}$, $F =\{r\}$ and δ is defined as follows:

$\delta(s,a,\$) = \{(s,a\$)\}$
$\delta(s,b,\$) = \{(q,b\$)\}$
$\delta(s,a,a) = \{(s,aa)\}$
$\delta(s,b,a) = \{(q,\varepsilon)\}$
$\delta(q,b,a) = \{(q,\varepsilon)\}$
$\delta(q,b,\$) = \{(q,b\$)\}$
$\delta(q,b,b) = \{(p,bb)\}$
$\delta(p,c,b) = \{(p,\varepsilon)\}$
$\delta(p,\varepsilon,\$) = \{(r,\varepsilon)\}$

13.[10] Construct a pushdown automaton that accept the language generated by the following grammar:

$$S \rightarrow aS \mid aSbS \mid \varepsilon$$

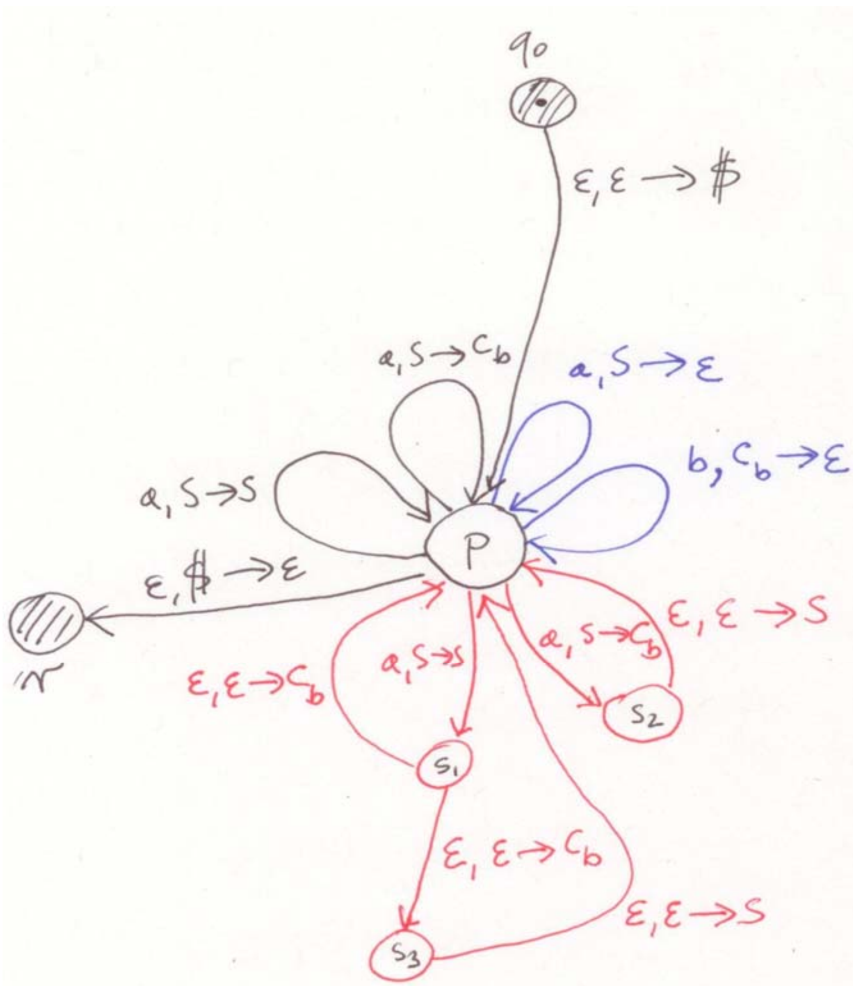
You may use either the construction from the textbook or the one presented in class.

Solution:

Let L be the language generated by a given grammar and let L' = L - {ε}.
We may use the simpler construction presented in class.
After removing ε-productions and transforming the obtained result into the Graibach normal form, we may get the following grammar (construction is not unique):

$$S \rightarrow aS \mid aSC_bS \mid a \mid aC_bS \mid aSC_b \mid aC_b$$
$$C_b \rightarrow b$$

Now, the construction from the class results in (colours in automata correspond to colours in grammar, ε treated separately – q_0 is final):



For an alternative definition of pushdown automaton (more popular, red stuff in class notes), we have:

$A = (\{p,r\},\{a,b\},\{S, C_b\}, p,S ,\delta, F)$, where $F =\{r\}$ and δ is the following:

$\delta(p,a,S) = \{(p,SC_bS), (p,S), (p,C_bS), (p,SC_b), (p,C_b),(p,\varepsilon)\}$
$\delta(p,b,C_b) = \{ (p,\varepsilon)\}$
$\delta(p,\varepsilon,S) = \{(r,\varepsilon)\}$.

14.[10] Show that the following language is not context-free

$$\{ a^i b^j \mid i \leq j^2 \}.$$

Solution:

Assume $L = \{ a^n b^j \mid n \leq j^2 \}$ is context free, then pumping lemma hold.
Let p be the number in pumping lemma. Let $z = a^q b^p$ and $q=p^2$. Clearly, $|z| \geq p$ and $z \in L$.
Therefore, z=uvwxy and $|vx| \geq 1$, $|vwx| \leq p$, and $(\forall i \geq 0)\ uv^i wx^i y \in L$. Let us consider two cases,
Case 1) $vx = a^k$, i.e. containing no b's. Then, $z' = uv^2wx^2y = a^{q+k}b^p$, and $k \geq 1$.
Therefore, $z' \notin L$, contradiction.
Case 2) $vx = a^j b^k$, and $k \geq 1$, i.e. containing at least one b. Then, consider $z' = uv^0wx^0y$.
$\#a(z') = q-j = p^2-j$, and $1 \leq j \leq p-1$ (#a(z') is the number of 'a' in z')
$\#b(z') = p-k$, and $1 \leq k \leq p$.
So, $(\#b(z'))^2 = (p-k)^2 = p^2 -2kp +k^2$
Clearly $kp-j > k^2 - kp$, since $kp-j \geq p-(p-1)=1$, but $k^2 - kp = k(k-p) \leq 0$. Now, it is easy to verify that, $p^2-j > p^2 -2kp +k^2$. Therefore, $z' \notin L$, contradiction.

Bonus questions.

15.[8] Transform the pushdown automaton from Example 2.14 (page 115, Figure 2.15 of the textbook) into an equivalent context free grammar using the construction from Lemma 27 of the textbook.

Solution:

To simplify notation we will write 'i' instead of 'q_i', so Q=[1,2,3,4}. We start with productions that are in fact automaton independent (steps 2 and 3 from the top of page 122):

For all i,j = 1,2,3:

$$A_{ij} \rightarrow A_{i1} A_{1j} \mid A_{i2} A_{2j} \mid A_{i3} A_{3j}$$

For all i=1,2,3:

$A_{ii} \to \varepsilon$

We have $\Gamma = \{0,\$\}$. For u=$, we have:

$\delta(1, \varepsilon, \varepsilon) = \{(2, \$)\}$ and $\delta(3, \varepsilon, \$) = \{(4, \varepsilon)\}$, so we add $A_{14} \to A_{23}$

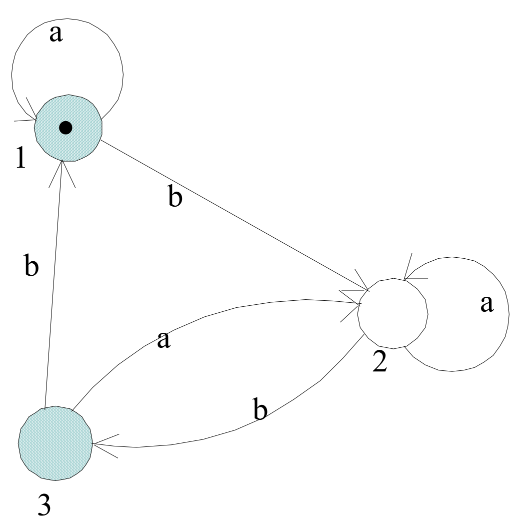
For u=0, we have :

$\delta(2, 0, \varepsilon) = \{(2, 0)\}$ and $\delta(2, 1, 0) = \{(3, \varepsilon)\}$, so we add $A_{13} \to 0A_{22}1$

and

$\delta(2, 0, \varepsilon) = \{(2, 0)\}$ and $\delta(3, 1, 0) = \{(3, \varepsilon)\}$, so we add $A_{23} \to 0A_{23}1$

16.[10] Read pages 3-9 (27-29 on the right top) from Lecture Notes 3, or the similar pages from any other appropriate textbook, and using this method (that does not use the idea of 'generalized automaton') find a regular expression that generates the same language as the automaton below.



	k=0	k=1	k=2	k=3
r_{11}^k	$a \cup \varepsilon$	a^*	a^*	$a^* \cup a^*ba^*b(\varepsilon \cup (a \cup ba^*b) a^*b)^* ba^*$
r_{12}^k	b	a^*b	a^*ba^*	$a^*ba^* \cup a^*ba^*b(\varepsilon \cup (a \cup ba^*b) a^*b)^* (a \cup ba^*b) a^*$
r_{13}^k	$\emptyset$	$\emptyset$	a^*ba^*b	$a^*ba^*b \cup a^*ba^*b(\varepsilon \cup (a \cup ba^*b) a^*b)^*$
r_{21}^k	$\emptyset$	$\emptyset$	$\emptyset$	$a^*b(\varepsilon \cup (a \cup ba^*b) a^*b)^* ba^*$
r_{22}^k	$a \cup \varepsilon$	a^*	a^*	$a^* \cup a^*b(\varepsilon \cup (a \cup ba^*b) a^*b)^* (a \cup ba^*b) a^*$
r_{23}^k	b	b	a^*b	$a^*b \cup a^*b(\varepsilon \cup (a \cup ba^*b) a^*b)^*$
r_{31}^k	b	ba^*	ba^*	$ba^* \cup (\varepsilon \cup (a \cup ba^*b) a^*b) (\varepsilon \cup (a \cup ba^*b) a^*b)^*ba^*$
r_{32}^k	a	$a \cup ba^*b$	$(a \cup ba^*b) a^*$	$(a \cup ba^*b) a^* \cup (\varepsilon \cup (a \cup ba^*b) a^*b) (\varepsilon \cup (a \cup ba^*b) a^*b)^* (a \cup ba^*b) a^*$
r_{33}^k	ε	ε	$\varepsilon \cup (a \cup ba^*b) a^*b$	$(\varepsilon \cup (a \cup ba^*b) a^*b)^*$

$L(M) = L(r_{11}^3 \cup r_{13}^3) =$
$L(a^* \cup a^*ba^*b(\varepsilon \cup (a \cup ba^*b) a^*b)^* ba^* \cup a^*ba^*b \cup a^*ba^*b(\varepsilon \cup (a \cup ba^*b) a^*b)^*)$

Or, by popular 'abuse of notation':

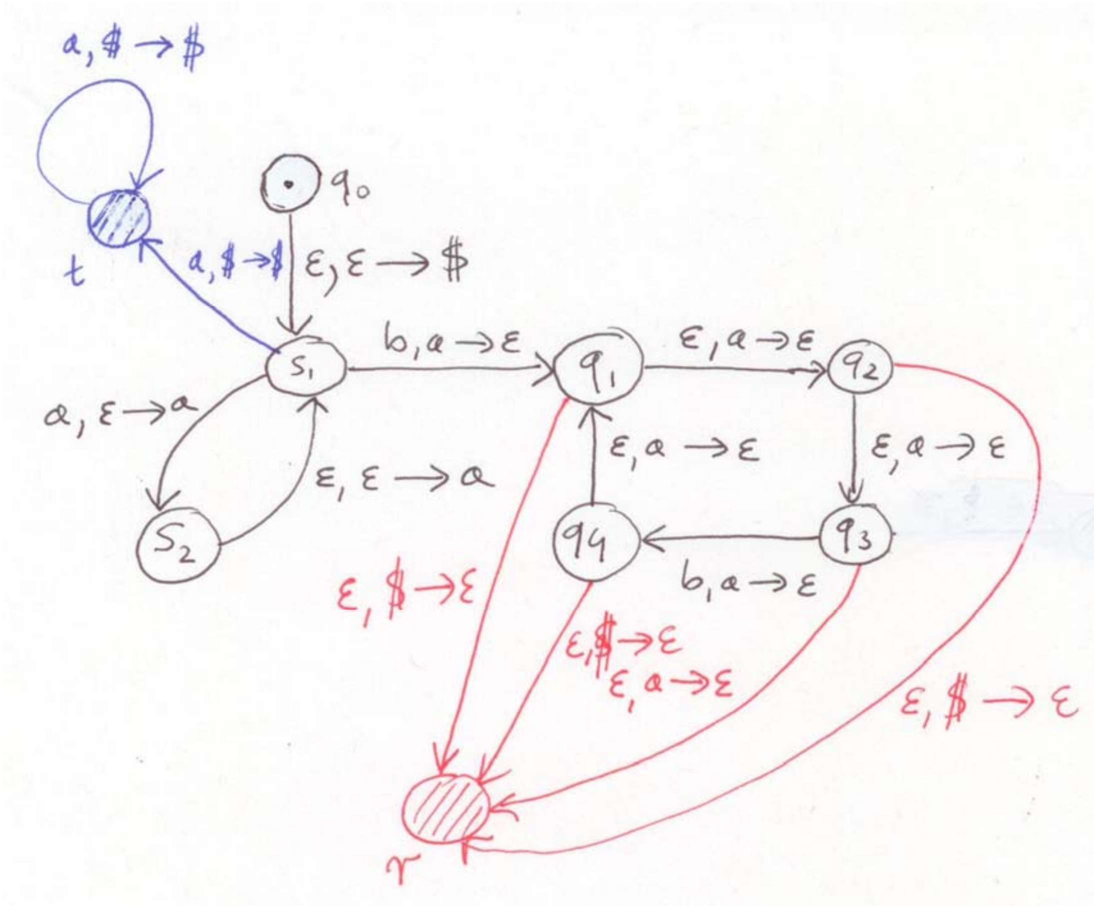
$L(M) = a^* \cup a^*ba^*b(\varepsilon \cup (a \cup ba^*b) a^*b)^* ba^* \cup a^*ba^*b \cup a^*ba^*b(\varepsilon \cup (a \cup ba^*b) a^*b)^*$

17.[10] Construct a pushdown automaton accepting: $\{ a^nb^m \mid 2n \neq 3m, n \geq 0, m \geq 0 \}$.

Solution:

The idea here is to use the pushdown automaton for the language $\{ a^nb^m \mid 2n = 3m, n \geq 0, m \geq 0\}$ from Question 11, to match every 2 b's with 3 a's. First observe that if the state q_1 has been rached them m>0. Moreover if the top of the stack is \$ in any of the states q_1, q_2, or q_4, then 2n>3m and if the top of the stack in q_3 is 'a' and there is no symbol to read, then 2n<3m. The case of m=0 (which implies n>0) needs to treated separately.

This lead us to the following automaton, where the black part is from the automaton accepting the language $\{ a^nb^m \mid 2n = 3m, n \geq 0, m \geq 0\}$, red part deals with m>0 and blue part with m=0.



.

For an alternative definition of pushdown automaton (more popular, red stuff in class notes), we have:

A = (Q, Σ, Γ, s, $, δ, F), where

Q = {s, q_1, q_2, q_3, q_4, r, t}, Σ={a,b}, Γ= {a,b,$}, F = {r,t} and δ is defined as follows:

(the part from acceptance of { $a^n b^m$ | 2n≠3m, n≥0, m≥0 }, and beginning of the case m=0).

δ(s,a,$) = {(s,aa$), (t,$)} (we assume here that the top of the stack is the first symbol on the left, so 'a' in this case)

δ(s,a,a) = {(s,aa)}

δ(s,b,a) = {(q_1, ε)}

δ(q_1,ε,a) = { (q_2,ε)}

δ(q_2,ε,a) = { (q_3,ε)}

δ(q_3,b,a) = { (q_4,ε)}

δ(q_4,b,a) = { (q_1,ε)}

(the part that deals with 2n<3m, m>0)

δ(q_1,ε,$) = { (r,ε)}

δ(q_2,ε,$) = { (r,ε)}

δ(q_4,ε,$) = { (r,ε)}

(the part that deals with 2n>3m, m>0)

$\delta(q_3, \varepsilon, a) = \{ (r, \varepsilon)\}$

(the remaining part that of m=0)

$\delta(t, a, \$) = \{(t, \$)\}$