Solutions Exam 170530

Here we only give a brief explanation of the solution. Your solution should in general be more elaborated than these ones.

1. Our property is: $P(n)$: if $S \Rightarrow^n w$ then w starts and ends with an a and any b in the word is immediately preceded and followed by an a.

   We will use course-of-value/strong induction on the length of the derivation (number of steps) $S \Rightarrow^n w$.

   Base case: $S \Rightarrow w$, hence the rule applied should have been $S \to aba$.

   Here, aba starts and ends with a and the only b in the words is clearly placed between two a's.

   Step case: Our IH is: if $S \Rightarrow^k w$ in at most $n > 0$ steps ($1 \leqslant k \leqslant n$) then w starts and ends with an a and any b in the word is immediately preceded and followed by an a

   Let $S \Rightarrow^{n+1} w$ with $n > 0$.

   Since $n > 0$ then the first rule applied should have been $S \to SabaS$ or $S \to aSbSa$.

   In the case the first rule was $S \to SabaS$ then $w = w_1abaw_2$ with $S \Rightarrow^i w_1$, $S \Rightarrow^j w_2$, and $1 \leqslant i, j \leqslant n$. Then the IH applies to both w_1 and w_2. Since by IH w_1 starts with a then so does w, and since by IH w_2 ends with a then so does w. Any b in w is either a b in w_1 or in w_2 and hence by IH the b is immediately preceded and followed by an a, or is the b just added but the first rule in the derivation, which is clearly immediately preceded and followed by an a.

   In the case the first rule was $S \to aSbSa$ then $w = aw_1bw_2a$ with $S \Rightarrow^i w_1$, $S \Rightarrow^j w_2$, and $1 \leqslant i, j \leqslant n$. Then the IH applies to both w_1 and w_2. In this case w clearly starts and ends with an a. Any b in w is either a b in w_1 or in w_2 and hence by IH the b is immediately preceded and followed by an a, or is the b just added but the first rule in the derivation which is located between w_1 and w_2. Since by IH w_1 ends with an a then this b is immediately preceded by an a; since by IH w_2 starts with an a then this b is immediately followed by an a.

2. We define a NFA:

	0	1
$\to^* q_0$	$\emptyset$	$\{q_1\}$
$^*q_1$	$\{q_1, q_2\}$	$\{q_1\}$
q_2	$\emptyset$	$\{q_3\}$
$^*q_3$	$\{q_3\}$	$\{q_1\}$

3.

	0	1
$\to^* q_0$	$q_0q_1q_2$	q_1
$^*q_0q_1q_2$	$q_0q_1q_2q_4$	$q_1q_2q_3q_4$
q_1	$-$	$q_1q_3q_4$
$^*q_0q_1q_2q_4$	$q_0q_1q_2q_4$	$q_1q_2q_3q_4q_5$
$q_1q_2q_3q_4$	$q_1q_3q_4q_5$	$q_1q_2q_3q_4q_5$
$q_1q_3q_4$	$q_1q_3q_4q_5$	$q_1q_3q_4q_5$
$^*q_1q_2q_3q_4q_5$	$q_1q_3q_4q_5$	$q_1q_2q_3q_4q_5$
$^*q_1q_3q_4q_5$	$q_1q_3q_4q_5$	$q_1q_3q_4q_5$

4. I will solve equations:

$$\begin{array}{ll} E_0 = 0E_0 + 1E_1 + 2E_2 & E_0 = 0E_0 + 1E_1 + 20E_3 + 21 \\ E_1 = 0E_2 + 2E_3 & E_1 = 00E_3 + 01 + 2E_3 = (00+2)E_3 + 01 \\ E_2 = 0E_3 + 1E_4 & \\ E_3 = 0E_2 + 1E_4 + 2E_1 & E_3 = 00E_3 + 01 + 1 + 2E_1 E_4 = \epsilon \end{array}$$

$$\begin{array}{l} E_0 = 0E_0 + 1(00+2)E_3 + 101 + 20E_3 + 21 = 0E_0 + (100+12+20)E_3 + 101 + 21 \\ E_3 = 00E_3 + 01 + 1 + 2(00+2)E_3 + 201 = (00+200+22)E_3 + (01+1+201) \end{array}$$

Hence $E_3 = (200+22+00)^*(201+01+1)$ and

$$E_0 = 0^*((100+12+20)E_3 + 101 + 21) = 0^*((100+12+20)(200+22+00)^*(201+01+1) + 101 + 21)$$

5. (a)

	q_0	q_1	q_2	q_3	q_4
q_5		X	X	X	X
q_4	X	X	X		
q_3	X	X	X		
q_2	X				
q_1	X				

The equivalent classes are $\{q_0, q_5\}, \{q_1, q_2\}, \{q_3, q_4\}$.
The resulting automaton is:

	a	b
$\rightarrow^* q_0q_5$	q_1q_2	q_1q_2
q_1q_2	q_3q_4	q_3q_4
q_3q_4	q_0q_5	q_0q_5

(b) $((a+b)(a+b)(a+b))^*$

6. $2(0^*2 + (00)^*01)^*(2+0)^*2$

7. (a) See slide 8 lecture 9.

(b) Let us assume our language $\mathcal{L}$ of palindromes over $\{0,1\}^*$ is a regular language.
Hence the PL should apply.
Let n be the constant given by the PL.
Let $w = 0^n 110^n$. We have that $w \in \mathcal{L}$ and that $|w| \geqslant n$.
Hence $w = xyz$ with $y \neq \epsilon$ and $|xy| \leqslant n$.
So y should contain only 0's and at least one 0.
For any $k > 1$ then xy^kz will contain more than n 0's at the beginning of the word (before the part 11) while only n 0's at the end of the word (after the part 11). Hence this word is not a palindrom and thus $xy^kz \notin \mathcal{L}$ for $k > 1$, which contradicts the PL.
Then, $\mathcal{L}$ cannot be regular.

8. (a)

$$\begin{array}{ll} S \rightarrow LC \mid AR & A \rightarrow \epsilon \mid aA \\ B \rightarrow \epsilon \mid bB & C \rightarrow \epsilon \mid cC \\ L \rightarrow aLb \mid aA \mid bB & R \rightarrow bRc \mid bB \mid cC \end{array}$$

(b) We divide the problem into 2 cases: $i \neq j$ and $j \neq k$.
LC will generate words $a^ib^jc^k$ when $i \neq j$ and AR will generate words $a^ib^jc^k$ when $j \neq k$.
A, B and C generate zero or more a's, b's or c's respectively.

Observe that in the case $i \neq j$ the number of c's is independent of the number of a's and/or b's. So here ($S \to LC$) L will ensure that the number of a's is NOT the same as the number of b's at the beginning of the word, and then C will generate as many c's as desired at the end of the word.

To ensure $i \neq j$, L will first put the same amount of a's than of b's (via the recursive rule $L \to aLb$) and then will finish by either adding one or more a's (by moving to aA) OR one or more b's (by moving to bB).

Similarly, in the case $j \neq k$ the number of a's is independent of the number of b's and/or c's. So here ($S \to AR$) A will generate as many a's as desired at the beginning of the word, and then R will ensure that the number of b's is NOT the same as the number of c's at the end of the word.

To ensure $j \neq k$ R, will first put the same amount of b's than of c's (via the recursive rule $R \to bRc$) and then will finish by either adding one or more b's (by moving to bB) OR one or more c's (by moving to cC).

(c) See slides 14 and 15 in lecture 12.

(d) Yes, words where $i \neq j$ and $j \neq k$, for example, $abbc$ will have 2 parse tree generating them: one using $S \to LC$ first and one using $S \to AR$ first.

(e) Leftmost derivation: $S \Rightarrow^{lm} AR \Rightarrow^{lm} aAR \Rightarrow^{lm} aaAR \Rightarrow^{lm} aaR \Rightarrow^{lm} aabRc \Rightarrow^{lm} aabbBc \Rightarrow^{lm} aabbc$

9. (a) See slide 21 lecture 13.

(b) i. (2pts) A, B, D and H are nullable

$$
\begin{array}{lll}
S \to aAbB \mid abB \mid aAb \mid ab \mid C \mid dD \mid d \mid EF & A \to aA \mid a \mid G & B \to bB \mid b \mid GF \mid D \\
C \to cC \mid CC & D \to dD \mid d & E \to fF \mid Ee \\
F \to eE \mid Ff & G \to gG \mid cC \mid Ee & H \to hH \mid h
\end{array}
$$

ii. (1.5pts) unit productions: $S \to C, A \to G, B \to D$

$$
\begin{array}{lll}
S \to aAbB \mid abB \mid aAb \mid ab \mid cC \mid CC \mid dD \mid d \mid EF & A \to aA \mid a \mid gG \mid cC \mid Ee & B \to bB \mid b \mid GF \mid dD \mid d \\
C \to cC \mid CC & D \to dD \mid d & E \to fF \mid Ee \\
F \to eE \mid Ff & G \to gG \mid cC \mid Ee & H \to hH \mid h
\end{array}
$$

iii. (2pts) non-generating symbols: C, E, F, G

$$
\begin{array}{lll}
S \to aAbB \mid abB \mid aAb \mid ab \mid dD \mid d & A \to aA \mid a & B \to bB \mid b \mid dD \mid d \\
 & D \to dD \mid d & H \to hH \mid h
\end{array}
$$

iv. (0.5pts) non-reachable symbols: H, h

$$
S \to aAbB \mid abB \mid aAb \mid ab \mid dD \mid d \qquad A \to aA \mid a \qquad B \to bB \mid b \mid dD \mid d \qquad D \to dD \mid d
$$

(c) See slide 22 lecture 13.

(d)

$$
\begin{array}{lll}
S \to TR \mid XR \mid TY \mid XY \mid ZD \mid d & T \to XA & R \to YB \\
A \to XA \mid a & B \to YB \mid b \mid ZD \mid d & D \to ZD \mid d \\
X \to a & Y \to b & Z \to d
\end{array}
$$

10.

$$
\begin{array}{|ccccc}
\{S,A\} & & & & \\
\emptyset & \{A\} & & & \\
\emptyset & \{S,B\} & \{S,A\} & & \\
\emptyset & \{S,B\} & \emptyset & \{S,A\} & \\
\{C\} & \{B\} & \{C\} & \{C\} & \{A\} \\
\hline
c & b & c & c & a
\end{array}
$$

S belongs to the upper-most set, which means that the word is generated by the grammar since S is the starting symbol of the grammar.

11. Let $\Sigma = \{0, 1, 2, X\}$.

Let $M = (\{q_0, ..., q_4, q_f\}, \Sigma, \delta, q_0, \Box, \{q_f\})$, with δ is as follows:

$\delta(q_0, 0) = (q_1, X, R)$	we need to read at least one 0;
$\delta(q_1, 0) = (q_1, 0, R)$	there could come even more 0's;
$\delta(q_1, \Box) = (q_2, \Box, L)$	if after the 0's I read a blank then there is nothing in the middle part; we need to check that there was at least a 0 in the end;
$\delta(q_2, 0) = (q_f, 0, R)$	Since when going left after the 0's we didn't read X it means we have at least two 0's which is fine so we accept;
$\delta(q_1, 2) = (q_3, 2, R)$	the middle part starts and we read an odd nr of 2's (fine);
$\delta(q_3, 2) = (q_4, 2, R)$	we read an even nr of 2's (not fine);
$\delta(q_4, 2) = (q_3, 2, R)$	we read an odd nr of 2's (fine);
$\delta(q_3, 1) = (q_5, 1, R)$	after an odd nr of 2's we read a 1 (still fine);
$\delta(q_5, 2) = (q_3, 2, R)$	we get another round of the middle part and we read an odd nr of 2's (fine);
$\delta(q_5, 0) = (q_6, 0, R)$	the 0's in the end are coming;
$\delta(q_6, 0) = (q_6, 0, R)$	more 0's are coming;
$\delta(q_6, \Box) = (q_f, \Box, R)$	we have a right tape so we accept.