Name: __________________________________________

Put all your answers on the test itself. Be sure to put your name above.

1. Using the method from class (this is the same method as the one in Theorem 1.25 on page 46 of the text), give the state diagram of a DFA that recognizes the union of the language recognized by the first DFA below with the language recognized by the second DFA below: (You must use the method of Theorem 1.25. If you use the method of Theorem 1.45, which only gives an NFA and not a DFA, you will not receive credit.)

\begin{align*}
\text{Solution:} \\
\end{align*}

[22 points]
2. (a) Let $M$ be the NFA given below.

If you want to transform $M$ into an equivalent regular expression, you have to first modify $M$ into a GNFA $M'$ before you start eliminating states. Give the state diagram of this GNFA $M'$. (You are not being asked to eliminate any states. Just give the GNFA $M'$ that you first transform $M$ into. Do not give any transitions labeled with $\emptyset$.)

Solution:
(b) Suppose that at some point while transforming an NFA into a regular expression you have the following GNFA.

Show the GNFA you would get from this one by eliminating state $p$. (You are not being asked to convert the GNFA into a regular expression. Just eliminate the state $p$.)

[22 points]
3. Let $G$ be the grammar

\[
\begin{align*}
S & \rightarrow WaWaT \\
T & \rightarrow aUbT|bWaT|\varepsilon \\
U & \rightarrow aUbU|\varepsilon \\
W & \rightarrow bWaW|\varepsilon
\end{align*}
\]

(It may help you to know that this grammar generates the set of strings in \{a, b\}∗ that have exactly two more $a$’s than $b$’s, the variable $T$ generates the strings with the same number of $a$’s as $b$’s, the variable $U$ generates the strings with the same number of $a$’s as $b$’s such that every prefix of the string contains at least as many $a$’s as $b$’s, and the variable $W$ generates the set of strings in \{a, b\}∗ such that the string has the same numbers of $a$’s as $b$’s, and every prefix of the string has at least as many $b$’s as $a$’s.)

(a) Give a leftmost derivation in $G$ for the string $baabaaab$. (Show each step of the derivation. Do not combine several steps into one.)

**Solution:**

\[
\begin{align*}
S & \Rightarrow WaWaT \Rightarrow bWaWaWaT \Rightarrow baWaWaT \Rightarrow baaWaT \Rightarrow baabWaWaT \Rightarrow baabaWaT \Rightarrow baabaaWaT \Rightarrow baabaaUbT \Rightarrow baabaaabT \Rightarrow baabaaab
\end{align*}
\]
(b) Using the method from class, give a PDA that recognizes the language generated by $G$. (Your PDA can push several symbols onto the stack in one move.)

\[\begin{align*}
\epsilon, S & \rightarrow WaWaT \\
\epsilon, T & \rightarrow aUbT \\
\epsilon, T & \rightarrow bWaT \\
\epsilon, T & \rightarrow \epsilon \\
\epsilon, U & \rightarrow aUbU \\
\epsilon, U & \rightarrow \epsilon \\
\epsilon, W & \rightarrow bWaW \\
\epsilon, W & \rightarrow \epsilon
\end{align*}\]

\[\begin{align*}
\epsilon, \epsilon & \rightarrow S$ \\
\epsilon, S & \rightarrow S$ \\
\epsilon, \epsilon & \rightarrow S$ \\
\epsilon, S & \rightarrow S$ \\
\epsilon, S & \rightarrow S$
\end{align*}\]
4. Classify each of the following languages as being i) regular, or ii) context-free but not regular, or iii) not context-free. Justify your answers.

(a) \(\{a^k b^a a^k | k \leq n\}\)
   Classification: iii
   Explanation:
   We use the Pumping Lemma for context-free languages. Given \(p \geq 1\), choose \(s = a^p b^p a^p\). Then, \(s\) is in the language and \(|s| = 3p \geq p\).
   Given \(u, v, x, y, z\) with \(s = uvxyz\), \(|vy| > 0\) and \(|vxy| \leq p\), we choose \(i = 0\). We show that \(uv^i xy^i z = uvy\) is not in the language.
   Case 1: Either \(v\) or \(y\) contains one or more \(a\)'s. Since \(|vxy| \leq p\), \(v\) and \(y\) together only contain \(a\)'s to the left of the \(b\)'s, or only contain \(a\)'s to the right of the \(b\)'s (plus possibly some \(b\)'s). Thus, in \(uxz\), the number of \(a\)'s to the left of the \(b\)'s is not equal to the number of \(a\)'s to the right of the \(b\)'s, so \(uxz\) is not in the language.
   Case 2: Neither \(v\) nor \(y\) contains \(a\)'s. Then, \(uxz = a^p b^p - |vy| a^p\). Since \(|vy| > 0\), \(p - |vy| < p\), so \(uxz\) is not in the language.
(b) $\{a^k b^n c^m d^k \mid n \leq m \}$

Classification: \( \mathbb{II} \)

Explanation:

The language is context-free because it is generated by the following grammar:

$$
S \rightarrow aSdT
\hspace{2cm} T \rightarrow bTcTc\varepsilon
$$

We use the Pumping Lemma to show that the language is not regular. Given $p \geq 1$, choose $s = a^pd^p$. Then, $s$ is in the language and $|s| = 2p \geq p$. Given $x, y, z$ with $s = xyz$, $|y| > 0$ and $|xy| \leq p$, we must have $y = a^k$ for some $k$ with $1 \leq k \leq p$. Choose $i = 2$. Then, $xy^iz = a^{p+k}d^p$, which is not in the language.
(c) \( \{ w \in \{0,1\}^* | w \text{ has even length and contains at least two 0's} \} \)

Classification: \( \hat{1} \)

Explanation:

Let

\[
\begin{align*}
L_1 &= \{ w \in \{0,1\}^* | w \text{ has even length} \} \\
L_2 &= \{ w \in \{0,1\}^* | w \text{ has at least two 0's} \}
\end{align*}
\]

The language given in the problem is \( L_1 \cap L_2 \). Since the regular languages are closed under intersection, if \( L_1 \) and \( L_2 \) are regular, then the given language is regular. \( L_1 \) is given by the regular expression \( (0 \cup 1)(0 \cup 1)^* \), so it is regular, and \( L_2 \) is given by the regular expression \( 1^*01^*0(0 \cup 1)^* \), so it is regular.

(Another proof would be to give a DFA for the language. The DFA can be constructed with six states.)

[24 points]
5. Let $A$ be a regular language and $w$ be a minimal length string in $A$. (This means that $w$ is in $A$, and if $w'$ is a string of length less than the length of $w$, then $w'$ is not in $A$. For instance, a minimal length string in $a(a \cup b)(a \cup b)^*$ is $ab$.)

Prove that if $M$ is a DFA that recognizes $A$, then the number of states in $M$ is greater than the length of $w$.

Solution: Since $w \in A$, $M$ accepts $w$. During the computation of $M$ on $w$, $|w| + 1$ states appear. If any of these states are the same and $w'$ is obtained from $w$ by removing the part of $w$ between the repeated states, the $w'$ is also accepted by $M$ and thus $w'$ would be in $A$. This would contradict the fact that $w$ is a minimal length string in $A$. Thus, the $|w| + 1$ states that appear in the computation of $M$ on $w$ are all different. This means that $M$ contains at least $|w| + 1$ states. [10 points]