(1) \( (9 + 6 + 6 = 21 \text{ pts.}) \)

Let \( A = \{ a^n b^n : n \geq 1 \} \). Define \( E \) to be the language of strings that differ in at most one place from a string in \( A \). An example of a string in \( E \) is \( aaba \), since changing the last \( a \) to \( b \) gives a string in \( A \). Note that \( E \) contains \( A \), and that the strings in \( E \) have the same lengths as strings in \( A \).

Define \( G \) to be the context-free grammar (\( \{ S, T, U \}, \{ a, b \}, R, S \)), where the rules in \( R \) are:

\[
S \rightarrow aSb \mid aTU \mid UTb \\
T \rightarrow aTb \mid \epsilon \\
U \rightarrow a \mid b.
\]

(a) For each of the following strings, say whether it belongs to \( E \), and if so, give a derivation for it (need not be a leftmost derivation): (i) \( \epsilon \), (ii) \( bb \), (iii) \( aaabb \), (iv) \( aabbb \).

(b) Find an ambiguous string and draw two different parse trees for it.

(c) Show the step of removing \( \epsilon \)-rules from the grammar without changing the language it derives. (Do not convert all the way into Chomsky normal form. The grammar is repeated overleaf.)

This was not covered (yet) in 2019.
(2) \(18 + 6 = 24\) pts.

This problem refers to the grammar \(G\) and languages \(A\) and \(E\) of problem (1) (the period is just punctuation):

\[
\begin{align*}
S & \rightarrow aSb \mid aTU \mid UTb \\
T & \rightarrow aTb \mid \epsilon \\
U & \rightarrow a \mid b.
\end{align*}
\]

(a) Prove by the structural induction technique that \(L(G) \subseteq E\). (You may speak in terms of \(E\) “allowing up to one error.” As usual, “reasonable proof shortcuts” are OK.)

(b) Is \(L(G) = E\)? Justify your answer briefly by referring to your parsing strategies in problem (1), but you need not give a formal proof.

The actual exam paper had two extra blank pages, one after here and one at the end, which are not used in this sample.
Please circle the unique best answer or use the fill-in blanks at bottom. Justifications may help for partial credit but are not needed.

1. The intersection of two DCFLs:
   (a) Is always decidable.
   (b) Might not be a CFL.
   (c) Could be a regular language.
   (d) All of the above.

2. If \( L(G) = T \) and \( G' \) is a grammar obtained by deleting one rule from \( G \), then:
   (a) \( L(G') \supseteq T \).
   (b) \( L(G') \subseteq T \).
   (c) \( L(G') = T \).
   (d) \( L(G') \) could be either a subset or superset of \( T \), it depends.

3. Let \( G = (V, \Sigma, R, S) \) be a context-free grammar that generates a language \( A \). Suppose you add the rule \( S \rightarrow SS \) to make a new grammar \( G' \). Then \( L(G') \) equals:
   (a) \( A \cdot A \).
   (b) \( A^* \).
   (c) \( A^+ \).
   (d) \( \emptyset \).

4. Let \( G_1 = (V_1, \Sigma, R_1, S_1) \) and \( G_2 = (V_2, \Sigma, R_2, S_2) \) be two CFGs with the same \( \Sigma \). Suppose you add the rule \( S_1 \rightarrow S_2 \) and make \( S_1 \) the start symbol of a combined grammar \( G' = (V', \Sigma, R', S_1) \) [namely, with \( V' = V_1 \cup V_2 \) and \( R' = R_1 \cup R_2 \cup \{S_1 \rightarrow S_2\} \)]. Then \( L(G') \) equals:
   (a) \( L(G_2)^+ \).
   (b) \( L(G_1) \cdot L(G_2) \).
   (c) \( L(G_1) \cup L(G_2) \).
   (d) None of the above.

5. If a deterministic Turing machine never moves any of its tape heads left, then its language is:
   (a) Always regular.
   (b) Always a CFL but possibly not regular.
   (c) Always decidable but possibly not a CFL.
   (d) Always empty.
With $\Sigma = \{a, b, c\}$, define
\[ L = \{ xcy : x, y \in \{a, b\}^* \land \#b(x) = \#b(y) \land |x| = |y| \}. \]

(a) Prove that $L$ is not a CFL via the CFL Pumping Lemma. (Hint: Try test strings $x$ that have four “regions,” where each “region” is $a^n$ or $b^n$, plus the middle $c$.)

(b) Describe in English words a 2-tape deterministic TM $M$ such that $L(M) = L$. You need not give an arc-node diagram but wording similar to the text’s. The words should describe each left-to-right and/or right-to-left “pass” made by $M$ on a general input $x$ and must give enough detail to pinpoint where and why $M$ is not a PDA.

End of Exam