Homework 2: due before midterm exam

Question 1 (2 points). Consider the following grammar \( G = (\Sigma, V, R, S) \):
- \( \Sigma = \{a, b\} \).
- \( V = \{S\} \), and \( S \) is the start variable.
- \( R \) consists of the following rules: \( S \rightarrow aS \mid aSbS \mid \epsilon \).
Determine which of the following words are in \( L(G) \).
(a) \( a^2b^3a \), i.e., \( aabbba \).
(b) \( a^4b^2 \), i.e., \( aaaabb \).
(c) \( abba \).
(d) \( a^3b^2a^2b^3 \), i.e., \( aaabbaabbb \).
Please substantiate your claim, i.e., if you claim a word is in \( L(G) \), you should provide its derivation tree. If you claim it is not, then briefly state your reason why.

Question 2 (2 points). Construct the CFG for each of the following languages.
(a) \( L_1 = \{a^mb^n \mid m > n\} \).
(b) \( L_2 = \{w^R \mid w \in \{a, b\}^*\} \).

Here \( L_2 \) is a language over the alphabet \( \{a, b, \$\} \), and \( w^R \) denotes the reverse of \( w \). For example, if \( w = aabbb \), then \( w^R = bbbaa \). If \( w = abababa \), then \( w^R = abababa \), which is the same as \( w \) itself. Likewise, if \( w = \epsilon \), then \( w^R = \epsilon \).
(c) \( L_3 = \{w_1 \$ w_1^R \$ w_2 \$ w_2^R \$ \cdots \$ w_k \$ w_k^R \$ \mid \) each \( w_i \in \{a, b\}^* \) for some \( k \geq 1 \}. \) That is, \( L_3 \) consists of all the words of the form:
\[
u_1 \$ v_1 \$ u_2 \$ v_2 \$ \cdots \$ u_k \$ v_k \$
\]
for some \( k \geq 1 \) and for each \( 1 \leq i \leq k \), \( u_i, v_i \in \{a, b\}^* \) and \( v_i^R = u_i \).
(d) \( L_4 \) is the complement of the language \( \{a^mb^n \mid m \geq n \geq 0\} \) over the alphabet \( \{a, b\} \). More formally, \( L_4 = \Sigma^* \setminus \{a^mb^n \mid m \geq n \geq 0\} \), where \( \Sigma = \{a, b\} \).

Question 3 (2 points). Show that the following languages are not CFL.
(a) \( L_5 = \{ww \mid w \in \{a, b\}^*\} \). That is, a word is in \( L_5 \), if and only if when it is divided in the middle, the first and the last half of the word are the same.
(b) \( L_6 = \{a^n \mid n \text{ is a prime number}\} \).

Question 4 (2 points). Prove or disprove the following.
(a) If \( L \) is CFL and \( K \) is regular, then \( L \cap K \) is regular.
(b) If \( L \) is CFL and \( K \) is regular, then \( L \cup K \) is regular.

Question 5 (2 points). Recall the following definitions from HW 1. For a language \( L \subseteq \Sigma^* \), we define a relation \( \sim_L \) on \( \Sigma^* \) as follows. \( u \sim_L v \), if the following holds: For every \( w \in \Sigma^* \), \( uw \in L \) if and only if \( vw \in L \). The number of equivalence classes of \( \sim_L \) is denoted by \( \#(\sim_L) \), i.e., the index of \( \sim_L \).

Prove or disprove the following. For every CFL language \( L \), the index \( \#(\sim_L) \) is finite.