The answers must be the original work of the author. While discussion with others is permitted and encouraged, the final work should be done individually. You are not allowed to work in groups. You are allowed to build on the material supplied in the class. Any other source must be specified clearly.

1. (5+5 points) Consider the following grammar $G$. Note that the grammar does not contain λ-rules except at $S$.

   \[
   \begin{align*}
   S & \rightarrow aSb | DEF | D | \lambda \\
   D & \rightarrow E | EF | abEF \\
   E & \rightarrow eEFf | a | F \\
   F & \rightarrow fFfE | a
   \end{align*}
   \]

   a. Use algorithm 4.3.1 to construct the CHAIN sets for the variables in $V$.

   b. Construct an equivalent grammar $G_c$ that does not contain chain rules.

2. (10+10 points) Consider the following grammar $G$:

   \[
   \begin{align*}
   S & \rightarrow ACH | BB \\
   A & \rightarrow aA | aF \\
   B & \rightarrow CFH | b \\
   C & \rightarrow aC | DH \\
   D & \rightarrow aD | BD | Ca \\
   F & \rightarrow bB | b \\
   H & \rightarrow dH | d \\
   \end{align*}
   \]

   a. Construct the TERM set for $G$ and construct the equivalent grammar $G_T$ that does not contain variables that do not generate strings of terminals.

   b. Construct the REACH set for $G_T$ and construct the equivalent grammar $G_U$ that does not contain useless variables.

3. (10 points) Show that all the symbols of the grammar

   \[
   \begin{align*}
   S & \rightarrow AB | C \\
   A & \rightarrow aA | a \\
   B & \rightarrow bB | b \\
   C & \rightarrow D | E \\
   D & \rightarrow aDb | c \\
   E & \rightarrow bEa | c
   \end{align*}
   \]

   are useful. Construct an equivalent grammar $G_C$ by removing the chain rules from $G$. Show that $G_C$ contains useless symbols.

Please turn the page over.
4. (10 points) Convert the following grammar $G$ into Chomsky normal form. Show your steps clearly. Note that $G$ already satisfies the conditions on the start symbol $S$, $\lambda$-rules, useless symbols, and chain rules.

\[
S \rightarrow aBA \mid dd \\
A \rightarrow aAa \mid e \\
B \rightarrow fAeAf \mid bbb
\]

5. (10 points) Let $G$ be a grammar in Chomsky normal form, and $w \in L(G)$ be a string of length $n \ (n \geq 1)$. How many steps are needed to derive $w$ in $G$?

6. (10 points) Consider the grammar $G$ from Example 4.5.2:

\[
S \rightarrow AT \mid AB \\
T \rightarrow XB \\
X \rightarrow AT \mid AB \\
A \rightarrow a \\
B \rightarrow b
\]

Give the upper diagonal matrix produced by the CYK algorithm when run with $G$ and the input string $abbb$.

7. (10+5+10+5 points) Consider the Chomsky normal grammar $G$:

\[
S \rightarrow AB \\
A \rightarrow aA \mid d \\
B \rightarrow bBb \mid d
\]

The above grammar in Chomsky normal form is as follows:

\[
S \rightarrow AB \\
A \rightarrow CA \mid d \\
B \rightarrow DT \mid d \\
T \rightarrow BD \\
C \rightarrow a \\
D \rightarrow b
\]

a. Give the upper diagonal matrix produced by the CYK algorithm when run with $G$ and the input string $dbdb$. Show all your work.

b. Is $dbdb \in L(G)$? Why? Provide the reason based on the upper diagonal matrix you constructed.

c. Give the upper diagonal matrix produced by the CYK algorithm when run with $G$ and the input string $aadd$. Show all your work.

d. Is $aadd \in L(G)$? Why? Provide the reason based on the upper diagonal matrix you constructed.