1. Recall that a context-free grammar G is ambiguous if there exists a string w that can be derived by G in two different ways. Show that the language CFG-AMBIG consisting of all ambiguous CFGs is undecidable.

Hint: Reduce PCP to CFG-AMBIG.

Solution: Show PCP $\leq_m$ CFG-AMBIG. Given an instance of PCP, the idea is to construct a context-free grammar G as follows. The start symbol of G is S, and G contains the rules:

$$\begin{aligned} S &\to\ *A \\ S &\to B \\ A &\to \varepsilon \\ B &\to \varepsilon \end{aligned}$$

Number the dominoes $D_1, \ldots, D_n$. For each domino $D_i = \binom{x_i}{y_i}$, where $x_i, y_i \in \Sigma^*$, G contains the rules:

$$\begin{aligned} A &\to x_i A \# i \\ B &\to y_i B \# i \\ B &\to\ * y_i B \# i \end{aligned}$$

Suppose that the instance of PCP has a match:

$$x_{i_1} x_{1_2} \cdots x_{i_t} = y_{i_1} y_{i_2} \cdots y_{i_t} = w$$

Then the string $*w\#i_t\#i_{t-1}\#\cdots\#i_1$ has two derivations in G:

$$\begin{aligned} S \to\ *A \to\ *x_{i_1} A \# i_1 \to\ *x_{i_1} x_{i_2} A \# i_2 \# i_1 \to \cdots \to\ *w \# i_t \# \cdots \# i_1 \\ S \to B \to\ *y_{i_1} B \# i_1 \to *y_{i_1} y_{i_2} B \# i_2 \# i_1 \to \cdots \to\ *w \# i_t \# \cdots \# i_1 \end{aligned}$$

On the other hand, suppose that some string w has two derivations in G. Then $w \neq \varepsilon$ (since ε has only one derivation $S \to B \to \varepsilon$), so

$$w = *w' \# i_t \# i_{t-1} \# \cdots \# i_2 \# i_1$$

where each i_j is a number in $\{1, \ldots, n\}$. In this case by the construction of G it follows that $D_{i_1} D_{i_2} \cdots D_{i_t}$ is a matching sequence of dominoes (the top and bottom strings are both w').