

University of Waterloo
CS 462 — Formal Languages and Parsing
Winter 2011
Problem Set 8

Distributed Wednesday, March 2 2011.

Due Wednesday, March 9 2011, in class.

All answers should be accompanied by proofs.

1. [10 marks] Consider the language L consisting of all subwords of the Thue-Morse word. Show that L is not context-free. Hint: the Thue-Morse word avoids cubes.
2. [10 marks] Consider, for each $a \in \Sigma$, the new letter a' . The set of all such letters is Σ' . For a word $x \in (\Sigma \cup \Sigma')^*$, define the reduced word $r(x)$ obtained by treating a' as an inverse to a satisfying $aa' = a'a = \epsilon$, and applying the transformations $aa' \rightarrow \epsilon$ and $a'a \rightarrow \epsilon$ repeatedly until no letter is next to its inverse. For example, $r(cb'aa'babb') = ca$, because $cb'aa'babb' \rightarrow cb'babb' \rightarrow cabb' \rightarrow ca$.

Show that if $L \subseteq (\Sigma \cup \Sigma')^*$ is context-free, then $r(L)$ need not be context-free. Hint: show that that we can define the quotient of two languages in terms of the operation r and other operations that preserve context-freeness.

3. [10 marks] Let k be a fixed integer ≥ 2 . Given a fixed alphabet Σ , define $L_k = \{w^k : w \in \Sigma^*\}$ to be the set of all k 'th powers of words in Σ^* . Show that the complement $\Sigma^* - L_k$ is always context-free. This is the language of “non- k th-powers”.

Hint: on input x , use counters on the stack to keep track of some quantity that depends on both the length of x and the distance between two symbols in x that are different. Yes, I know this is a vague hint.