

# MAT-41186 Formal Languages Examination 11.12.2009

**NB** This is a closed-book exam, no material is allowed. Calculators are allowed.

1.
  - a) Show that the finite code  $a + ab$  cannot be recognized by any deterministic finite automaton with only one terminal state.
  - b) Show that, on the other hand, every prefix code (finite or infinite) which is a regular language can be recognized by some deterministic finite automaton with exactly one terminal state.
2. Which of the following implications are correct and which are false? Explain your reasoning!
  - a) If  $L_1$  is a regular language and  $L_2$  is a CF-language, then  $L_1 - L_2$  is a CF-language.
  - b) If  $L_1$  is a regular language and  $L_2$  is a CF-language, then  $L_2 - L_1$  is a CF-language.
  - c) If  $L_1$  is a regular language and  $L_2$  is a CF-language and  $L_1 \cup L_2$  is a regular language, then  $L_1 - L_2$  is a CF-language.
  - d) If  $L_1$  is a regular language and  $L_2$  is a CF-language and  $L_1 \cap L_2$  is a regular language, then  $L_1 - L_2$  is a CF-language.

(Here, as usual,  $L_1 - L_2$  denotes the difference of  $L_1$  and  $L_2$ , i.e., the language  $L_1 \cap \bar{L}_2$ .)
3. Explain briefly what is **a)** unambiguity of a CF-grammar, **b)** an inherently ambiguous CF-language, **c)** a deterministic CF-language.
4. CS-languages, their definition and basic properties.
5. The weights are  $P_i = C2^{n-i}$  ( $i = 1, \dots, n$ ) where  $C = 1/(2^n - 1)$ . Use Huffman's algorithm to find an optimal prefix code over the alphabet  $\{a, b, c\}$  for **a)**  $n = 4$ , **b)**  $n = 5$ , and **c)** a general value of  $n$ .

**Huom!** Mukana ei saa olla kirjallisuutta, tietokoneita eikä taulukoita. Laskuvälineet ovat sallittuja.

1.
  - a) Totea, että äärellistä koodia  $a + ab$  ei voi tunnistaa millään deterministisellä äärellisellä automaatilla, jossa on vain yksi lopputila.
  - b) Totea, että toisaalta jokainen prefiksikoodi (äärellinen tai ääretön), joka on säännöllinen kieli, voidaan tunnistaa jollain deterministisellä äärellisellä automaatilla, jossa on tarkalleen yksi lopputila.
  
2. Mitkä seuraavista päättelyistä ovat oikeita ja mitkä väriä? Perustelut mukaan!
  - a) Jos  $L_1$  on säännöllinen kieli ja  $L_2$  on CF-kieli, niin  $L_1 - L_2$  on CF-kieli.
  - b) Jos  $L_1$  on säännöllinen kieli ja  $L_2$  on CF-kieli, niin  $L_2 - L_1$  on CF-kieli.
  - c) Jos  $L_1$  on säännöllinen kieli ja  $L_2$  on CF-kieli ja  $L_1 \cup L_2$  on säännöllinen kieli, niin  $L_1 - L_2$  on CF-kieli.
  - d) Jos  $L_1$  on säännöllinen kieli ja  $L_2$  on CF-kieli ja  $L_1 \cap L_2$  on säännöllinen kieli, niin  $L_1 - L_2$  on CF-kieli.

(Tässä tavan mukaan merkitään  $L_1 - L_2$ :llä kielten  $L_1$  ja  $L_2$  erotusta eli kieltä  $L_1 \cap \bar{L}_2$ .)
  
3. Selosta lyhyesti, mitä ovat **a)** CF-kieliopin yksiselitteisyys, **b)** luontaisesti moniselitteinen CF-kieli ja **c)** deterministinen CF-kieli.
  
4. CS-kielet, niiden määrittely ja perusominaisuudet.
  
5. Painot ovat  $P_i = C2^{n-i}$  ( $i = 1, \dots, n$ ), missä  $C = 1/(2^n - 1)$ . Etsi Huffmanin algoritmilla aakkoston  $\{a, b, c\}$  optimaalinen prefiksikoodi, kun **a)**  $n = 4$ , **b)**  $n = 5$ , ja **c)** yleiselle  $n:n$  arvolle.