Examination
Formal Languages and Automata Theory
TDDD14 & TDDD85

(Formella Språk och Automatateori)

2016–06–02, 8.00 – 12.00

1. **NOT ALL PROBLEMS ARE FOR BOTH COURSES.**
   Pay attention to "only" comments.
2. Allowed help materials
   - A sheet of notes - 2 sided A5 or 1 sided A4.
     The contents is up to you.
     The notes should be signed in the same way as the exam sheets and returned together with the exam.
   - English dictionary

   Tillåtna hjälpmedel:
   - Ett papper med valfria anteckningar - 2 sidor A5 eller 1 sida A4. Anteckningarna ska signeras på samma sätt som tentamensarken och bifogas tentamen vid inlämnandet.
   - Engelsk ordbok
3. You may answer in Swedish or English.
4. Total number of credits is 33. Limits:

   3: 16 p, 4: 22 p, 5: 28 p.
5. *Jour* (person on duty): Johannes Schmidt, tel. 07 25 72 18 03

GOOD LUCK !

Make sure that you justify your answers! Unexplained answers will be granted 0 points. (For instance, if you are writing a grammar for a given language then you should explain that the grammar indeed generates the language. If you apply some known method then you should explain each step. And so on.)

1. *(3p)* The NFAϵ $N = (Q, \Sigma, \Delta, S, F)$ is defined as follows:

$$Q = \{0,1,2,3,4,5\} \qquad \Sigma = \{a,b\} \qquad S = \{0\} \qquad F = \{5\}$$

with the transition function Δ given by

	ϵ	a	b
$\rightarrow 0$	$\{4\}$	$\{1\}$	$\{1\}$
1	$\{3\}$	$\emptyset$	$\{2,4\}$
2	$\{3,4\}$	$\{2\}$	$\{2,4\}$
3	$\emptyset$	$\emptyset$	$\{4\}$
4	$\emptyset$	$\{5\}$	$\{5\}$
5 F	$\{1\}$	$\emptyset$	$\emptyset$

Using the standard method, construct an equivalent DFA M.

2. *(4p)* The DFA $M = (Q, \Sigma, \delta, s, F)$ is defined as follows:

$$Q = \{1,2,3,4,5\} \qquad \Sigma = \{a,b\} \qquad s = 1 \qquad F = \{5\}$$

with the transition function δ given by

	a	b
$\rightarrow 1$	2	5
2	4	1
3	4	2
4	2	1
5 F	5	4

   (a) Using the standard method, construct an equivalent, minimal DFA M_{min}.

   (b) Let L be the language defined by M_{min} (and thus by M), and consider the relation $R_L \subseteq \Sigma^* \times \Sigma^*$ defined by

$$xR_Ly \Leftrightarrow (\forall z \in \Sigma^*(xz \in L \Leftrightarrow yz \in L))$$

   How many equivalence classes does R_L have? Why? Choose two of the equivalence classes, and give two DFA's defining them.

3. *(4p)* For each of the following two languages answer whether it is regular, context-free but not regular, or not context-free and **prove these facts**. In case of non-regularity or non-context-freeness use the appropriate **Pumping lemma**.

   (a) $\{ xwx^R \mid w, x \in \{a, b\}^*, |w| = 1\}$

   (b) $\{ xwx^R \mid w, x \in \{a, b\}^*, |w| \geq 0 \}$

4. *(5p)* For each of the following languages answer whether it is regular, context-free but not regular, or not context-free. (Here a brief explanation is sufficient).

   (a) $\{ xyx \mid x, y \in \{a, b, c\}^*,\ \#c(y) = |y| \}$

   (b) $\{ xyx \mid x, y \in \{a, b, c\}^*,\ \#c(x) = |x| \}$

   (c) $\{ w \in \{a, b, c\}^* \mid \#a(w)$ is a prime number, $\#a(w) = \#b(w) = |w| \}$

   (d) The image of the language $\{a^n b^n \mid n \geq 0\}$ under the homomorphism h given by $h(a) = aba$, $h(b) = bab$.

   (e) The intersection of the languages from (a) and (c).

5. *(6p)* Which of the following statements are true, which are false? Why?

   (a) Any context-free language can be written as intersection of two regular languages.

   (b) Any regular language can be written as intersection of two context-free languages.

   (c) The intersection of two context-free languages can be recognized by a minimal DFA.

   (d) There is a regular language that contains a non-recursive subset.

   (e) The union of three context-free languages is recursively enumerable.

   (f) The language recognized by a non-deterministic push-down automata can be recognized by a deterministic Turing machine.

6. *(3p)* **Only TDDD14**
   Explain the notions of a recursive language and a recursively enumerable language. Show that the intersection of two recursively enumerable languages is recursively enumerable.

7. *(3p)* **Only TDDD85**
   Explain what it means that a context-free grammar is in Chomsky Normal Form. Write a context-free grammar in Chomsky Normal Form for the language $\{a^n b^n \mid n \geq 0\}$.

8. *(2p)* Can you think of a strategy for an algorithm to determine whether two given regular expressions are equivalent (that is, whether they define the same language)? What can you say about the runnning time of this algorithm? Can this be done within a finite number of steps?

9. *(6p)* In an attempt to construct LR parsers for certain grammars, we applied the standard method of constructing a DFA for the viable prefixes of a grammar. Some fragments of the obtained DFA's are given below.

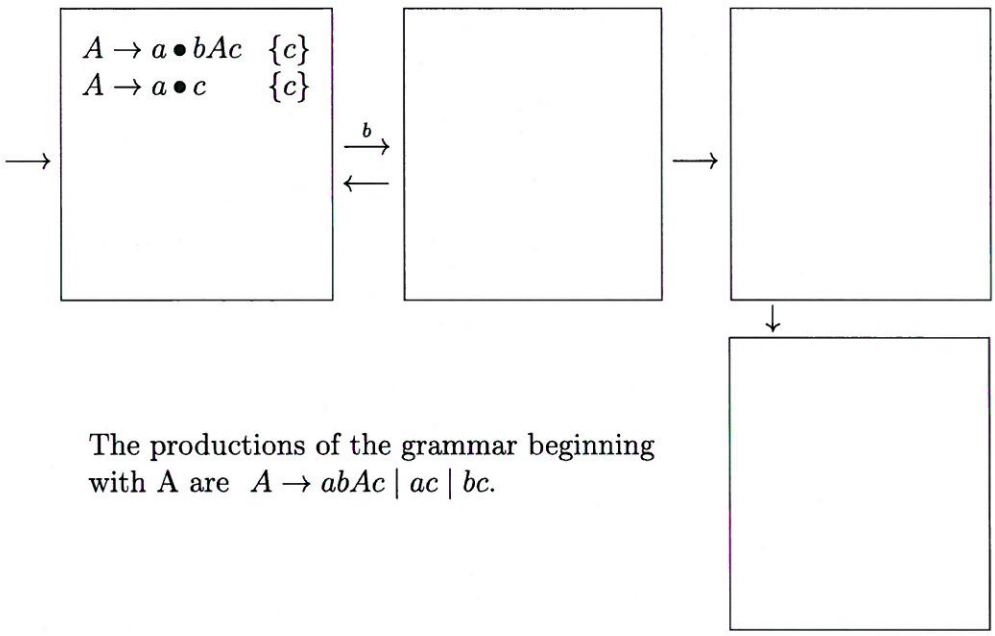
   Complete the missing items in the given states, the missing lookahead sets, and the missing symbols labelling the arrows. In each case, answer the following questions. Justify your answers.

   - Does the fragment of a DFA satisfy the conditions for the grammar to be LR(0)?
   - The same question about the conditions for LR(1).

   You may skip adding missing items or lookahead sets if they are not needed to answer the questions. For instance if you find the items in some state to violate the LR(1) conditions then you do not need to complete the other states. In such a case just make an appropriate comment.

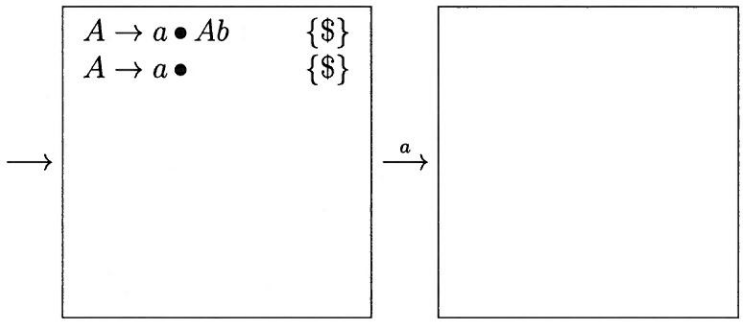
   a, b, c are terminal symbols and S, A, B are nonterminal symbols of the grammars; S is the start symbol.

   (a)



   The productions of the grammar beginning with A are $A \to abAc \mid ac \mid bc$.

(b)



The productions of the grammar beginning with A are $A \to aAb \mid bAc \mid a$.

THE END. Remember the warning from the beginning of page 2. Check that your answers are justified.