Skriftlig Eksamen
Beregnelighed (DM517)

Institut for Matematik & Datalogi
Syddansk Universitet

Mandag den 7 Januar 2008, kl. 9–13

Alle sædvanlige hjælpemidler (lærebøger, notater etc.) samt brug af lommeregner er tilladt. Eksamenssættet består af 5 opgaver på siderne (2–6). Fuld besvarelse er besvarelse af alle opgaver. De enkelte opgavers vægt ved bedømmelsen er angivet i procent. Der må gerne refereres til resultater fra lærebogen. Specielt må man gerne begrunde en påstand med at henvise til, at den umiddelbart følger fra et resultat i lærebogen (hvis dette altså er sandt!). I må gerne bruge metoder eller udvidelser af sætninger som er udledt i opgaver, der er stillet i løbet af kurset. Bemærk dog, at det ikke er tilladt at besvare et delspørgsmål, udelukkende med en henvisning til, at det følger af en af opgaverne. Henvisninger til andre bøger (ud over lærebogen) accepteres ikke som besvarelse af et spørgsmål! **Husk at begrunde alle dine svar!**

You are allowed to use the text-book, your personal notes and a pocket calculator. The exam contains 5 problems, each on a single page. The danish version can be found on pages 2-6, the english version is on pages 7-11.
A complete solution consists of a solution for all 5 problems. How much a problem weights can be seen from the numbers given in brackets for each problem. In general you may refer to results from the book unless it is explicitely stated that you may not. The same holds for problems solved in the exercise sessions. Of course, if you cite such a result it has to be obvious that your claim really is a straightforward consequence of that result.
It will not be accepted as an answer to refer to other books than the text-book. **Recall that you have to explain all your claims!**

OPGAVE 1 (20%)

Beskriv for hvert af de tre sprog nedenfor deterministiske endelige automater som accepterer præcis disse. Du skal argumentere kort for korrektheden af dine DFA'er.

Bemærk, at $\#_a(w)$ er antallet af a'er i strengen w.

(a) $L(R)$ hvor $R = (a \cup b)^* abba (a \cup b)^*$

(b) $L = \{w \in \{a, b\}^* | \#_a(w) \geq 2 \text{ eller } \#_b(w) \geq 2\}$

(c) $L' = \{w \in \{a, b\}^* | \#_a(w) \geq 2 \text{ og } \#_b(w) \geq 2\}$

OPGAVE 2 (20%)

Lad $L = \{a^n b^k c^m | n, k, m \geq 0 \text{ og } n + k = m\}$

Spørgsmål a:

Gør rede for at L ikke er regulært.

Spørgsmål b:

Er L kontekstfrit?

Du skal argumentere for dit svar. F.eks. kan du, hvis L er kontekstfrit, vise dette ved at angive en kontekstfri grammatik G med $L(G) = L$.

OPGAVE 3 (15%)

Lad $G = (\{a,b,c\},\{S,T\},R,S)$ være den kontekstfrie grammatik med startsymbol S og følgende regler

- $S \to aSbTc$
- $S \to aSb$
- $S \to \epsilon$
- $T \to bTc$
- $T \to \epsilon$

Omdan G til en kontekstfri grammatik G' på Chomsky normal form som opfylder at $L(G') = L(G) - \{\epsilon\}$.

OPGAVE 4 (21%)

Giv beskrivelser i ord og med angivelse af væsentlige skridt (via konfigurationer undervejs i beregningen) af Turing maskiner som beregner hver af funktionerne nedenfor. For alle tre skal der argumenteres i ord for at dine Turing maskiner faktisk beregner de angivne funktioner. Bemærk at du IKKE skal angive Turing maskinerne på diagramform eller som transitionstabeller!

Du må gerne bruge flere bånd og andre udvidede Turing maskine modeller fra bogen og kan antage at du har en Turing maskine $MULT$ som for givne binære tal a, b udregner $a * b$ (a gange b).

Brug notationen fra forelæsningerne til angivelse af båndinhold for Turing maskiner, så du f.eks. altid starter i konfigurationen.

$$\triangleright\underline{\sqcup}w$$

I alle tre spørgsmål kan du antage at input altid er lovligt. Din Turing maskine skal altså ikke checke dette.

1. $f(w) = ww^Rw$ hvor $w \in \{a, b\}^*$.
2. $$f(x, y) = \begin{cases} 1 & \text{hvis } x \leq y \\ 0 & \text{hvis } x > y \end{cases}$$

   Her antages x og y are være binære tal.
3. $f(n) = \lfloor\sqrt{n}\rfloor$ hvor n er et naturligt tal angivet på binær form. Hint: du må gerne bruge den Turing maskine der beregner funktionen $f(x, y)$ ovenfor som en subrutine.

OPGAVE 5 (24%)

Nedenfor er angivet en række sprog som alle er over alfabetet for den universelle Turing maskine. Du skal udvælge præcis fire af disse og for hvert af dem argumentere for om de er afgørlige eller uafgørlige. Det er tilladt at anvende Rice's sætning, men hvis den anvendes, skal du argumentere for at den faktisk kan anvendes på det konkrete sprog.

Vi betegner altid mængden af strenge som Turing maskinen M stopper på med $L(M)$. Desuden betegner vi med $< M >$ koden for M over alfabetet for den universelle Turing maskine.

(1) $L_1 = \{< M > | M$ stopper på to strenge af forskellig længde$\}$.

(2) $L_2 = \{< M >< w > | M$ bruger mindst 3 tilstande på input $w\}$.

(3) $L_3 = \{< M > | L(M)$ er kontekstfrit $\}$.

(4) $L_4 = \{< M_1 >< M_2 > | M_1$ stopper på flere strenge end $M_2\}$.

(5) $L_5 = \{< M_1 >< M_2 >< w > | M_1$ stopper før M_2 på input $w\}$.

(6) $L_6 = \{< M_1 >< M_2 > | \exists w \in L(M_1) \cap L(M_2) : M_1$ stopper før M_2 på input $w\}$.

Problem 1 (20%)

For each of the three languages below describe deterministic finite automata which accept exactly these languages. You must give a short argument for the correctness of each of the DFA's.

Note that $\#_a(w)$ is the number of a's in the string w.

(a) $L(R)$ where $R = (a \cup b)^* abba(a \cup b)^*$

(b) $L = \{w \in \{a, b\}^* | \#_a(w) \geq 2 \text{ or } \#_b(w) \geq 2\}$

(c) $L' = \{w \in \{a, b\}^* | \#_a(w) \geq 2 \text{ and } \#_b(w) \geq 2\}$

Problem 2 (20%)

Let $L = \{a^n b^k c^m | n, k, m \geq 0 \text{ and } n + k = m\}$

Spørgsmål a:

Prove that L is not a regular language.

Spørgsmål b:

Is L a context-free language?

You must give an argument supporting your answer. For example, if L is context-free, you may prove this by giving a context-free grammar G with $L(G) = L$.

Problem 3 (15%)

Let $G = (\{a, b, c\}, \{S, T\}, R, S)$ be the context-free grammar with starting symbol S and the following rules.

- $S \rightarrow aSbTc$
- $S \rightarrow aSb$
- $S \rightarrow \epsilon$
- $T \rightarrow bTc$
- $T \rightarrow \epsilon$

Convert G into a context-free grammar G' in Chomsky normal form such that $L(G') = L(G) - \{\epsilon\}$.

Problem 4 (21%)

Give descriptions in words, including important steps (via configurations during the calculation) of Turing machines which compute each of the functions below. For each of these you must argue why your machines compute precisely the corresponding function. Note that you should NOT descibe the Turing machines as diagrams or using transition tabels!

You may use several tapes and other extended Turing machine models from the book and you may also assume that you have at your disposal a Turing machine $MULT$ which given two binary numbers a, b computes $a * b$ (a times b).

Use the notation from the lectures when you describe the content of the tape(s) of a Turing machine, for example you may assume that the Turing machine always starts in the configuration

$$\triangleright \underline{\sqcup} w$$

In each of the subproblems below you may assume that the input is always legal. Thus your Turing machine does not need to check for errors in the input.

1. $f(w) = ww^R w$ where $w \in \{a, b\}^*$.
2. $$f(x, y) = \begin{cases} 1 & \text{if } x \leq y \\ 0 & \text{if } x > y \end{cases}$$

   Here x and y are binary numbers.
3. $f(n) = \lfloor \sqrt{n} \rfloor$ where n is a natural number given in binary notation. Hint: you may use the Turing machine which calculates the function $f(x, y)$ above as a subroutine.

Problem 5 (24%)

Below there is a collection of languages, all over the alphabet of the universal Turing machine. You must select exactly four of these and for each of these you must argue whether they are decidable or undecidable. You may use Rice's theorem, but if you want to use it, you must argue why it can be applied to the language in question.

We always denote the set of strings on which the Turing machine M will halt by $L(M)$. Furthermore we denote by $< M >$ the coding of M over the alphabet of the universal Turing machine.

(1) $L_1 = \{< M > | M$ halts on two strings of different lengths $\}$.

(2) $L_2 = \{< M >< w > | M$ uses at least three states on input $w\}$.

(3) $L_3 = \{< M > | L(M)$ is context-free $\}$.

(4) $L_4 = \{< M_1 >< M_2 > | M_1$ halts on more strings than $M_2\}$.

(5) $L_5 = \{< M_1 >< M_2 >< w > | M_1$ halts before M_2 on input $w\}$.

(6) $L_6 = \{< M_1 >< M_2 > | \exists w \in L(M_1) \cap L(M_2) : M_1$ halts before M_2 on input $w\}$.