Albert-Ludwigs-Universität
Institut für Informatik
Prof. Dr. F. Kuhn

Exam Theoretical Computer Science - Bridging Course

Tuesday, February 21, 2017, 9:00-10:30

Name: ....

Matriculation Nr.: ....

Signature: ....

Do not open or turn until told so by the supervisor!

- Write your **name** and **matriculation number** on this page and sign the document!
- Write your name on **all sheets**!
- Your **signature** confirms that you have answered all exam questions without any help, and that you have notified exam supervision of any interference.
- Write legibly and only use a pen (ink or ball point). Do **not use red**! Do **not use a pencil**!
- This is an **open book exam** therefore printed or hand-written material is allowed.
- However, **no electronic devices** are allowed.
- There are **seven tasks** (with several sub-tasks each) and there is a **total of 90 points**.
- **40 points are sufficient** in order to pass the exam.
- Only **one solution per task** is considered! Make sure to strike out alternative solutions, otherwise the one yielding the minimal number of points is considered.
- **Detailed steps** might help you to get more points in case your final result is incorrect.
- The keywords **Show...** or **Prove...** indicate that you need to prove or explain your answer carefully.
- The keywords **Give...** or **State...** indicate that you only need to provide a plain answer.
- **Read each task thoroughly** and make sure you understood what is required of you.

Task	Max. Points	Achieved
1	11	
2	14	
3	10	
4	13	
5	15	
6	12	
7	15	
Σ	90	

Task 1: Regular Languages (11 Points)

Let $L = \{ab(ab)^n b^m \mid n, m \in \mathbb{N} \cup \{0\}\}$ be a language over the alphabet $\Sigma = \{a, b\}$.

(a) Give a **regular expression** that generates L *(1 point)*.

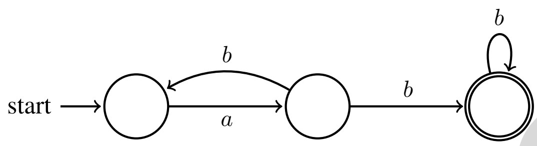
(b) Give a **nondeterministic** finite automaton that recognizes L and has **at most three** states *(5 points)*.

(c) Give a **deterministic** finite automaton that recognizes L *(5 points)*.
Remark: You can directly give a deterministic automaton. No intermediate steps required.

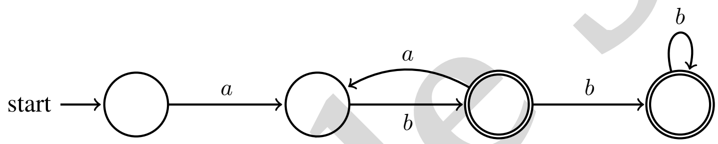
Sample Solution

(a) $ab(ab)^*b^*$

(b)



(c)



Task 2: Context-Free Languages (14 Points)

Consider the language $L = \{(ab)^n(cd)^n \mid n \in \mathbb{N}\}$ over alphabet $\Sigma = \{a, b, c, d\}$. Note that $0 \notin \mathbb{N}$!

(a) Give a **context-free grammar** that generates L *(2 points)*.

(b) Give a grammar in **Chomsky Normal Form** (CNF) that generates L *(5 points)*.
Remark: You can directly give a grammar in CNF for L, no intermediate steps required.

(c) Prove that L is **not** a regular language. Use the Pumping Lemma *(7 points)*.

Sample Solution

(a) Let $G = (V, \Sigma, R, S)$ with $V := \{S, X, Y\}$ and

$$R := \{S \to XY | XSY, X \to ab, Y \to cd\}.$$

(b) The Grammar $G = (V, \Sigma, R, S)$ with $V := \{S, X, Y, Z, A, B, C, D\}$ and

$$R := \{S \to XY | XZ, Z \to SY, X \to AB, Y \to CD, A \to a, B \to b, C \to c, D \to d\}$$

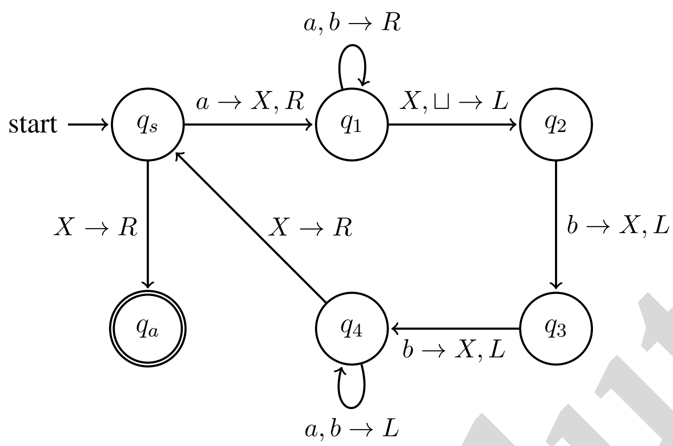
generates L and is in Chomsky Normal Form.

(c) We prove that L is not regular by showing that L can not always be pumped. Let p be the pumping length. We investigate the string $s = (ab)^p(cd)^p \in L$ which is obviously longer than p.

Consider a partition $xyz = s$ with $|y| \geq 1$ and $|xy| \leq p$. It follows that y consists of at least one symbol a or b but contains no other symbols. Then xy^2z has more a's or more b's than c's (d's) thus $xy^2z \notin L$.

Task 3: Turing Machines (10 Points)

Consider the Turing machine M over the alphabet $\Sigma = \{a, b\}$, which is given via the following state diagram. *Note: The blank symbol $\sqcup \in \Gamma$ represents an empty cell on the tape.*



(a) Simulate M with input $s_1 = ab$ on its tape until it halts. Give **all configurations** that M passes through. State whether $s_1 \in L$ or not *(3 points)*.

(b) Simulate M with input $s_2 = abb$ on its tape until it halts. Give **all configurations** that M passes through. State whether $s_2 \in L$ or not *(4 points)*.

(c) Give a description of the language $L(M)$ that M recognizes in the **form of a set** *(3 points)*.

Sample Solution

(a) M run on input s_1 yields the following configurations:

$\sqcup$	q_s	a	b	$\sqcup$
$\sqcup$	X	q_1	b	$\sqcup$
$\sqcup$	X	b	q_1	$\sqcup$
$\sqcup$	X	q_2	b	$\sqcup$
$\sqcup$	q_3	X	X	$\sqcup$

(b) M run on input s_2 yields the following configurations:

$\sqcup$	q_s	a	b	b	$\sqcup$
$\sqcup$	X	q_1	b	b	$\sqcup$
$\sqcup$	X	b	q_1	b	$\sqcup$
$\sqcup$	X	b	b	q_1	$\sqcup$
$\sqcup$	X	b	q_2	b	$\sqcup$
$\sqcup$	X	q_3	b	X	$\sqcup$
$\sqcup$	q_4	X	X	X	$\sqcup$
$\sqcup$	X	q_s	X	X	$\sqcup$
$\sqcup$	X	X	q_a	X	$\sqcup$

(c) $L(M) = \{a^n b^{2n} \mid n \in \mathbb{N}\}$.

Task 4: $\mathcal{O}$ - Notation (13 Points)

State whether the following claims are true or false *(1 point each)*. Then **prove or disprove** the claim *(5 points for (a) and 6 points for (b))*. Use the definition of the $\mathcal{O}$-notation.

(a) $\log_2(2^n \cdot n^3) \in \mathcal{O}(n)$. ***Hint****: You may use that* $\log_2 n \leq n$ *for all* $n \in \mathbb{N}$.

(b) $\sqrt[3]{n^2} \in \mathcal{O}(\sqrt{n})$.

Sample Solution

(a) The claim is true. We give $c > 0$ and $M \in \mathbb{N}$ such that for all $n \geq M : \log_2(2^n \cdot n^3) \leq c \cdot n$.

$$\begin{aligned} &\log_2(2^n \cdot n^3) \\ &= \log_2(2^n) + \log_2(n^3) \\ &= n + 3 \cdot \log_2(n) \\ &\leq n + 3n = 4n. \end{aligned}$$

Thus $\log_2(2^n \cdot n^3) \leq c \cdot n$ for $n \geq M := 1$ and $c := 4$.

(b) The claim is false. We disprove it by showing a contradiction. Assume there is a constant $c > 0$ and a $M \in \mathbb{N}$ such that for all $n > M$ the inequality $\sqrt[3]{n^2} \leq c \cdot \sqrt{n}$ holds.

$$\begin{aligned} \iff\quad & \sqrt[3]{n^2} \leq c \cdot \sqrt{n} \\ \iff\quad & n^{2/3} \leq c \cdot n^{1/2} \\ \iff\quad & \frac{n^{2/3}}{n^{1/2}} \leq c \\ \iff\quad & n^{2/3-1/2} \leq c \\ \iff\quad & n^{1/6} \leq c \\ \iff\quad & n \leq c^6 \end{aligned}$$

Thus the above statement is violated for all $n > \max\{c^6, M\}$ which is a contradiction.

Task 5: Decidability (15 Points)

(a) Consider the problem VERTEXCOVER:

$$\text{VERTEXCOVER} := \{\langle G, k\rangle \,|\, \text{undirected, simple graph } G \text{ has a } \textbf{vertex cover} \text{ of size } k\}.$$

A **vertex cover** of size k of $G = (V, E)$ is a subset $C \subseteq V$ of nodes, such that $|C| = k$ and for all $\{u, v\} \in E$ it holds that $u \in C$ or $v \in C$.

Show that VERTEXCOVER is **decidable** by giving an algorithm (abstract description or pseudo-code) that decides whether a graph has a vertex cover of size k or not *(6 points)*.

Explain why your algorithm accepts **exactly** the instances $\langle G, k\rangle$ which have a vertex cover of size k *(1 point)* and why it always halts *(1 points)*.

(b) Let $c \in \mathbb{N}$ be a fixed **constant** and let Σ be a **finite** alphabet.

Consider the **length-restricted** Halting problem defined over Σ.

$$H_c := \{\langle M, s\rangle \mid \text{Turing machine } M \text{ halts on input } s \text{ and } |\langle M, s\rangle| \leq c\},$$

H_c is similar to the usual Halting problem but restricted to strings $\langle M, s\rangle$ shorter than c.

State whether H_c is decidable or not *(2 points)*. Proof your claim *(5 points)*.

Sample Solution

(a) Given an input graph $G = (V, E)$ with input number k, we test for all of the $\binom{|V|}{k}$ possible subsets $C \subseteq V$ of size k, whether they cover all edges *(Remark: The exact number of subsets does not matter, it is sufficient to state that it is finite)*.

To check if $C \subseteq V$ covers all edges we have to check for each edge $\{u, v\}$ whether $u \in C$ or $v \in C$ which takes at most $\mathcal{O}(|E||C|)$ *(i.e., finite)* time. If we find a vertex cover (all edges covered for one subset C) the algorithm halts and accepts. After all subsets were tested negatively the algorithm halts and rejects.

In total this takes at most $\mathcal{O}(\binom{|V|}{k}|E||C|)$ time *(again an actual time estimation is not required, it is sufficient to argue that it is finite)* and thence the algorithm finishes eventually. Since we examine all subsets of size k of V, a vertex cover of size k is found iff one exists.

(b) The length-restricted Halting problem H_c is decidable. This is because H_c is a *finite* language since it contains at most $|\Sigma|^c$ different strings *(Remark: This is sufficient to prove the decidability and receive the 5 points)*.

Therefore we can construct a decider which accepts an input $\langle M, s\rangle$ if it equals one of the at most $|\Sigma|^c$ different strings in H_c else rejects. Since there will be at most $|\Sigma|^c$ comparisons this process terminates in finite time.

Task 6: Complexity Theory (12 Points)

Consider the following problems

$$\textsc{IndependentSet} := \{\langle G, k\rangle \mid \text{undirected, simple graph } G \text{ has an } \textbf{independent set} \text{ of size } k\}.$$

An **independent set** of size k of $G = (V, E)$ is a subset $I \subseteq V$ of nodes, such that $|I| = k$ and $\{u, v\} \notin E$ for all $u, v \in I, u \neq v$.

$$\textsc{Clique} := \{\langle G, k\rangle \mid \text{undirected, simple graph } G \text{ has a } \textbf{clique} \text{ of size } k\}.$$

A **clique** of size k of $G = (V, E)$ is a subset $C \subseteq V$ of nodes, such that $|C| = k$ and $\{u, v\} \in E$ for all $u, v \in C, u \neq v$.

Use that CLIQUE is a **known** $\mathcal{NP}$-complete problem to **show** that INDEPENDENTSET is $\mathcal{NP}$-complete.

Remark: Document the steps of your proof carefully.

Sample Solution

First we prove INDEPENDENTSET $\in \mathcal{NP}$ by giving a non-deterministic algorithm that solves the problem in polynomial time.

Guess: For input $\langle G, k\rangle$ with $G = (V, E)$ the non-deterministic part of the algorithm guesses a subset $I \subseteq V$.

Check: First check if $|I| \geq k$ and if not, reject *(Remark: One can also guess a subset of size k in the first place, then one does not have to check that)*. Then check for all $\binom{|I|}{2} \in \mathcal{O}(|V|^2)$ pairs of nodes $u, v \in I$, if $\{u, v\} \notin E$. If that is the case for all pairs then accept, else reject. The adjacency tests take at most $\mathcal{O}(|V|^2 \cdot |E|)$ time in total, counting I is possible in $\mathcal{O}(|V|)$ time.

Second we prove that INDEPENDENTSET is $\mathcal{NP}$-hard by showing the polynomial reduction CLIQUE $\leq_p$ INDEPENDENTSET. Since we already know that CLIQUE is $\mathcal{NP}$-hard, we then know the same for INDEPENDENTSET.

For instance $\langle G, k\rangle$ with $G = (V, E)$ of CLIQUE we define $f(\langle G, k\rangle) := \langle \tilde{G}, k\rangle$ where $\tilde{G} = (V, \tilde{E})$ has the same nodes as G and $\tilde{E} := \{\{u, v\} \mid u, v \in V, u \neq v, \{u, v\} \notin E\}$ is the 'inverted' set of edges, i.e., $\tilde{G}$ has an edge where G has none and vice versa.

The computation of $\tilde{G}$ can be done by checking for the at most $\mathcal{O}(|V|^2)$ possibilities whether $\{u, v\} \notin E$, in which case we add $\{u, v\}$ to $\tilde{E}$. Each check takes at most $\mathcal{O}(|E|)$ time which makes a total of $\mathcal{O}(|V|^2 \cdot |E|)$ time and thus we require polynomial time in the input size.

Finally we have to show $\langle G, k\rangle \in$ CLIQUE $\iff f(\langle G, k\rangle) \in$ INDEPENDENTSET:

$$\begin{aligned}
& G = (V, E) \text{ has a clique } C \subseteq V \text{ of size } k \\
\iff & \{u, v\} \in E \text{ for all } u, v \in C, u \neq v \\
\iff & \{u, v\} \notin \tilde{E} \text{ for all } u, v \in C, u \neq v \\
\iff & \tilde{G} = (V, \tilde{E}) \text{ has an independent set } C \subseteq V \text{ of size } k.
\end{aligned}$$

Task 7: Logic (15 Points)

(a) Consider the following **propositional logical** entailment

$$\{p \lor r \lor s, p \lor \neg r, \neg p \lor q, \neg q\} \models s \land \neg r$$

(i) Use known equivalencies to convert the above entailment into the form $KB \models \bot$, where KB is in **Conjunctive Normal Form** (3 points).

(ii) Use the **Resolution Calculus** to prove this logical entailment (5 points).

(b) Consider the following **first order logical** formula

$$\varphi := [\forall x \forall y\, f(s(x,y)) \doteq s(f(x), f(y))] \land [\exists x \exists y\, \neg(\, f(x) \doteq f(y)\,)]$$

where x, y are variable symbols and f, s are function symbols.

(i) Give an interpretation which is **no model** of φ *(3 points)*.

(ii) Give a **model** of φ *(4 points)*.

Remark: No proof required. Mark which interpretation is a model and which is not.

Sample Solution

(a) (i) With the known equivalency $KB \models \varphi \Leftrightarrow KB \cup \neg\varphi \models \bot$ and DeMorgan we obtain

$$\begin{aligned} & \{p \lor r \lor s, p \lor \neg r, \neg p \lor q, \neg q\} \models s \land \neg r \\ \Longleftrightarrow\ & \{p \lor r \lor s, p \lor \neg r, \neg p \lor q, \neg q, \neg s \lor r\} \models \bot \\ \Longleftrightarrow\ & \{\{p, r, s\}, \{p, \neg r\}, \{\neg p, q\}, \{\neg q\}, \{\neg s, r\}\} \models \bot \end{aligned}$$

which is in conjunctive normal form.

(ii) We use the resolution inference rule to derive an unsatisfiable formula

$$\begin{aligned} \{p, r, s\}, \{p, \neg r\} &\vdash_{\mathbf{R}} \{p, s\} \\ \{p, s\}, \{\neg p, q\} &\vdash_{\mathbf{R}} \{q, s\} \\ \{q, s\}, \{\neg q\} &\vdash_{\mathbf{R}} \{s\} \\ \{s\}, \{\neg s, r\} &\vdash_{\mathbf{R}} \{r\} \\ \{r\}, \{p, \neg r\} &\vdash_{\mathbf{R}} \{p\} \\ \{p\}, \{\neg p, q\} &\vdash_{\mathbf{R}} \{q\} \\ \{q\}, \{\neg q\} &\vdash_{\mathbf{R}} \Box \end{aligned}$$

(b) (i) We define $I := (\mathbb{R}, \cdot^I)$ with

$$f^I : \mathbb{R} \to \mathbb{R},\ x \mapsto 0$$
$$s^I : \mathbb{R}^2 \to \mathbb{R},\ (x, y) \mapsto x + y.$$

Since f is a constant function this already violates the second condition thus I can be no model. The function s is irrelevant to this fact and we can pick anything.

Remark: No explanation required to receive full points!

(ii) We define $I := (\mathbb{R}, \cdot^I)$ and

$$f^I : \mathbb{R} \to \mathbb{R},\ x \mapsto x$$
$$s^I : \mathbb{R}^2 \to \mathbb{R},\ (x, y) \mapsto x + y.$$

Since f is no constant function the second condition is fulfilled. In fact f is a linear function, thus it respects the linearity condition $f(x + y) = f(x) + f(y)$ implied by the first statement.

Remark: No explanation required to receive full points!