The University of Sydney

School of Information Technologies

Unit of Study: ______________________________

Assignment/Project Title Page

Individual Assessment

Assignment name: ______________________________

Tutorial time: ______________________________

Tutor name: ______________________________

Declaration

I declare that I have read and understood the *University of Sydney Student Plagiarism: Coursework Policy and Procedure*, and except where specifically acknowledged, the work contained in this assignment/project is my own work, and has not been copied from other sources or been previously submitted for award or assessment.

I understand that understand that failure to comply with the *Student Plagiarism: Coursework Policy and Procedure* can lead to severe penalties as outlined under Chapter 8 of the *University of Sydney By-Law 1999* (as amended). These penalties may be imposed in cases where any significant portion of my submitted work has been copied without proper acknowledgement from other sources, including published works, the internet, existing programs, the work of other students, or work previously submitted for other awards or assessments.

I realise that I may be asked to identify those portions of the work contributed by me and required to demonstrate my knowledge of the relevant material by answering oral questions or by undertaking supplementary work, either written or in the laboratory, in order to arrive at the final assessment mark.

Student Id: ____________________

Student name: ____

Signed: ______________________

Date: ______________________

COMP3310/3610: Theory of Computation

Problem set 1

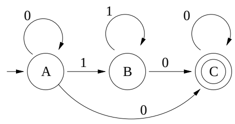
due 10am Tuesday September 12, 2006

Total marks: 100. This problem set is 10% of the final raw mark.
Answer all questions 1-5. All questions common for both COMP3310/3610. Assume that the alphabet is $\Sigma = \{0, 1\}$ unless stated otherwise. Please submit the University "Cover sheet for an individual project/assignment" with your answers.
This is an individual assessment; submit your own write-up of your solutions. **Collaboration** and discussion of the problems in this assignment is **encouraged**, but please write your solutions on your own and list the names of the people you collaborated with for this problem set.
Electronic submissions are much prefered. Please send a pdf of your solutions to `tasos@it.usyd.edu.au`.
Hard copies (**including a signed University "Cover sheet for an individual project/assignment"**) can be submitted at the School of IT building, 2nd floor, reception area, COMP3310 submission box.

Problem set 1 for COMP3310/3610

1. **(drawing DFAs, 15 marks)**
   Draw a deterministic finite state automaton for each of the following languages. Assume the alphabet is $\Sigma = \{0, 1\}$.
   (a) $L_1 = \{$ all strings that contain 0101 $\}$
   (b) $L_2 = \{$ all strings with exactly two 0s and at least two 1s$\}$
   (c) $L_3 = \{$The set of all strings such that every third symbol in the string is the same as the first symbol in the string$\}$

2. **(regular expressions and automata, 10 marks)**
   (a) Convert the following regular expression to a non-deterministic finite state automaton: $(0(0^*10^*)) \cup (1(1^*01^*))$. *(5 marks)*
   (b) Convert the following non-deterministic automaton to a deterministic one. *(5 marks)*



3. **(Properties of regular languages, 25 marks)**
   (a) Use the pumping lemma to prove that the following language is not regular. *(10 marks)*
   $$L_4 = \{1^m 0^n 1^{m+n} | m, n \geq 1\}$$
   (b) Prove the following property: If R_1 and R_2 are regular languages, then the following language is also regular
   $$R_1 - R_2 = \{w | w \in R_1 \text{ and } w \notin R_2\}$$
   Your proof should show how to design the new automaton for $R_1 - R_2$, including a description of its set of states and its transition function. *(5 marks)*
   (c) As an application of the previous question, construct the automaton that accepts the language $L_5 - L_6$ where $L_5 = \{w | w$ has length that is a multiple of $2\}$ and $L_6 = \{w | w$ has length that is a multiple of $4\}$. *(10 marks)*

4. **(Computability, 30 marks)**
   (a) Given a Turing machine M, consider the problem of deciding whether M, starting with a blank tape, will eventually write a non-blank symbol on its tape. Is this problem decidable or undecidable? Prove your claim. To be more precise, we can state this as a language problem. We would like to know whether the following language is decidable or not: $L = \{< M > | M$ is a Turing machine such that, when started on a blank tape, will eventually write a blank symbol on its tape$\}$. *(10 marks)*
   (b) Prove that the Post Correspondence problem is decidable over the unary alphabet $\Sigma = \{1\}$. *(10 marks)*
   (c) A useless state in a Turing machine is a state that is never entered on any input. Consider the problem of determining whether a Turing machine has useless states. Formulate this problem as a language problem and prove that it is undecidable. *(10 marks)*

5. **(Complexity, 20 marks)**

(a) Let G represent an undirected graph and define:

$$SPATH = \{< G, a, b, k > | G \text{ contains a simple path of length at most } k \text{ from } a \text{ to } b\}$$

and

$$LPATH = \{< G, a, b, k > | G \text{ contains a simple path of length at least } k \text{ from } a \text{ to } b\}$$

Prove that $SPATH \in P$ and $LPATH$ is NP-complete. *(15 marks)*

(b) Prove that deciding whether a formula given in disjunctive normal form (DNF) is satisfiable is in P. A DNF formula is an OR of terms $\phi = (x_1 \wedge x_4 \wedge \cdots) \vee (\overline{x}_1 \wedge x_2 \wedge \cdots) \vee \cdots (x_1 \wedge x_3)$. *(5 marks)*