8 Regular Languages and Finite Automata (AMP)

If $r$ and $s$ are regular expressions, write $r \preceq s$ to mean that the language of strings matching $r$ is contained in the language of strings matching $s$.

(a) Show that if $r_1 \preceq s_1$ and $r_2 \preceq s_2$, then $r_1 r_2 \preceq s_1 s_2$. [2 marks]

(b) Show that if $r \preceq s$, then $r^* \preceq s^*$. [2 marks]

(c) Suppose $s \preceq t$ and $rt \preceq t$. Prove by induction that $r^n s \preceq t$ holds for all $n \geq 0$; deduce that $r^* s \preceq t$. [3 marks]

(d) Which of the following instances of the $\preceq$ relation are valid? In each case either give a proof, or specific examples of $r$ and $s$ for which the relation fails to hold. [Hint: You may find part (c) helpful for some of the proofs.]

(i) $r^* | s^* \preceq (r | s)^*$ [1 mark]

(ii) $(r | s)^* \preceq r^* | s^*$ [1 mark]

(iii) $(r^* s^*)^* \preceq (r | s)^*$ [2 marks]

(iv) $(r | s)^* \preceq (r^* s^*)^*$ [2 marks]

(v) $(rs | r)^* r \preceq r(s r | r)^*$ [2 marks]

(e) Briefly explain why there exists an algorithm for deciding whether or not $r \preceq s$ holds for any given regular expressions $r$ and $s$ (over some fixed alphabet). [5 marks]