1993 Paper 5 Question 12

Formal Languages and Automata

Explain what is meant by a *regular expression* over an alphabet Σ, and by the language $L(r)$ denoted by such a regular expression r. [5 marks]

For any regular expressions r, s, t, show that if $L(r)$ contains $L(t|sr)$ then it also contains $L(s*t)$. [5 marks]

Assuming that the empty string ε is not in $L(s)$, show that if $L(r) = L(t|sr)$ then $L(r) = L(s*t)$. Hint: argue by induction on the length of strings in $L(r)$. [5 marks]

Give an example to show that the above assumption $\varepsilon \notin L(s)$ is necessary. [3 marks]

Deduce that when $\varepsilon \notin L(s)$, r and t|sr denote the same language if and only if r and s*t denote the same language. [2 marks]