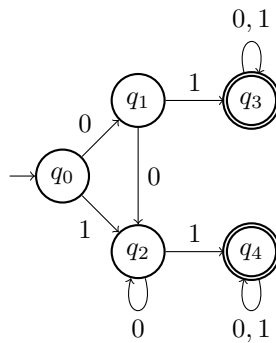


# Assignment 3

Due date: February 17, 2014

## Mandatory exercises

1) Minimize the following DFA. Explain your steps.



2) Consider the grammar  $G$  defined as follows.

$$\begin{aligned} S &\rightarrow AB \\ A &\rightarrow BB \mid a \\ B &\rightarrow AB \mid b \end{aligned}$$

Use the CYK algorithm to determine whether  $aabb$ ,  $aabba$ , and  $abbbb$  belong to  $L(G)$ . In particular, include the CYK tables in your answer.

## Voluntary exercises (for higher grades than 3)

3) Imagine that you are a spammer who wants to collect email addresses from web pages. You want to do this by constructing a set  $\{r_1, \dots, r_n\}$  of regular expressions such that  $L(r_1) \cup \dots \cup L(r_n)$  capture as many email addresses as possible and as few other pieces of text as possible.

When constructing your regular expressions, keep in mind that those who construct web pages often try to foil people like you by obfuscating their email addresses, e.g. as follows:

- h e n r i k b @ c s . u m u . s e
- henrikb AT cs DOT umu DOT se
- henrikb @ cs umu se
- henrikb (at) cs (dot) umu (dot) se
- etc.