

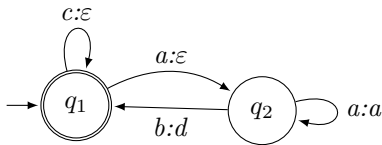
Einführung in die Computerlinguistik

Hausaufgabe zu FST und Morphologie, Abgabe 11.05.2020

Laura Kallmeyer

Sommer 2020, Heinrich-Heine-Universität Düsseldorf

Aufgabe 1 Betrachten Sie folgenden FST:



1. Auf welche Strings bildet der FST die folgenden Eingaben ab:
 (a) cc (b) aaab (c) aaabccc (d) ccababcaab
2. Welche Strings akzeptiert dieser FST und wie transformiert er sie? (Eine Beschreibung der Transformation in Worten genügt.)

Lösung:

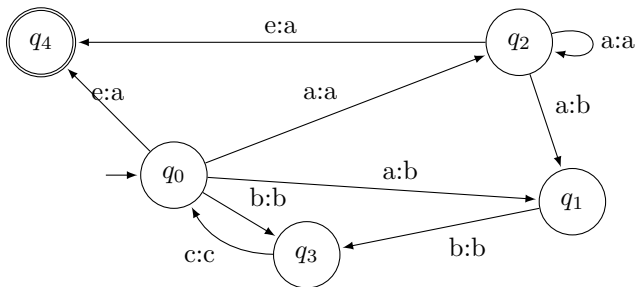
1. (a) ε (b) aad (c) aad (d) ddad
2. Akzeptiert wird die von $(c|a^+b)^*$ denotierte Sprache. Jedes c wird gelöscht, bei jeder Folge von a 's wird das erste gelöscht, die anderen werden in die Ausgabe kopiert, und jedes b wird d ersetzt.

Aufgabe 2 Erstellen Sie einen Finite State Transducer, der die von $(a|bc)^*e$ denotierte Sprache akzeptiert, und der jedes vor einem b stehende a durch ein b ersetzt. b 's und c werden ansonsten unverändert in die Ausgabe übernommen, und ein e wird durch ein a ersetzt.

Z.B.

Eingabe	e	$aaaae$	$aabcbce$	$bcbaae$	$abcabce$
Ausgabe	a	$aaaa$	$abbcbca$	$bcbaaa$	$bbcbca$

Lösung:



Aufgabe 3

Erstellen Sie einen Finite State Transducer, der die Formen der französischen Adjektive "bleu" (= blau), "jaune" (= gelb) und "gris" (= grau) jeweils im Genus F oder M und im Numerus Sg oder Pl erkennt und eine entsprechende Analyse ausgibt. D.h., es sollen folgende Analysen abgedeckt sein:

Ausgabe: lexikalischen Ebene	Eingabe: Oberflächenebene
bleu F Sg	bleue
bleu M Sg	bleu
bleu F Pl	bleues
bleu M Pl	bleus
jaune F Sg	jaune
jaune M Sg	jaune
jaune F Pl	jaunes
jaune M Pl	jaunes
gris F Sg	grise
gris M Sg	gris
gris F Pl	grises
gris M Pl	gris

Sie dürfen mehrere aufeinander folgende Buchstaben zusammenfassen.

Lösung: (es gibt natürlich mehrere, die Lösung sollte in einem gewissen Maße generalisieren)

