

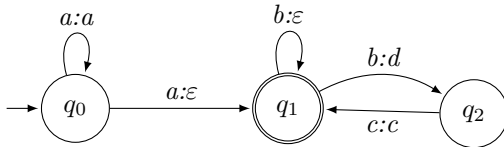
Einführung in die Computerlinguistik

Hausaufgabe zu FST und Morphologie, Abgabe 30.04.2019

Laura Kallmeyer

Sommer 2019, Heinrich-Heine-Universität Düsseldorf

Aufgabe 1 Betrachten Sie folgenden FST:



- Auf welche Strings bildet der FST die folgenden Eingaben ab:
 (a) aa (b) aaabbb (c) aaabbbc (d) aabbbcbcbbbb
- Welche Strings akzeptiert dieser FST und wie transformiert er sie? (Eine Beschreibung der Transformation in Worten genügt.)

Lösung:

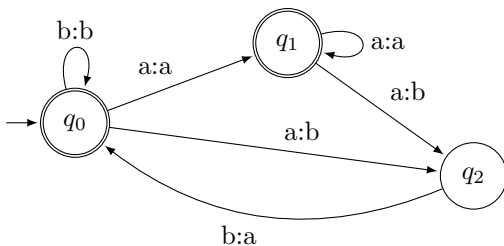
- (a) a (b) aa (c) aadc (d) adcdc
- Akzeptiert wird die von $a^*a(b^*bc)^*b^*$ denotierte Sprache. a 's, die von einem a gefolgt sind, werden in die Ausgabe kopiert, a 's unmittelbar vor einem b werden gelöscht. b 's, die vor einem b oder am Wortende stehen, werden gelöscht. Ein b vor eine c wird durch d ersetzt. Und jedes c wird in die Ausgabe übernommen.

Aufgabe 2 Erstellen Sie einen Finite State Transducer, der $\{a,b\}^*$ akzeptiert und jedes vor einem b stehende a mit dem b vertauscht. Alle anderen Eingabesymbole sollen unverändert in die Ausgabe übernommen werden.

Z.B.

Eingabe	aaa	bbbbaaaa	aabbbbbba	aabbbab	bbbabbbaab
Ausgabe	aaa	bbbbaaaa	ababbbba	ababba	bbbabbaba

Lösung:



Aufgabe 3

Erstellen Sie einen Finite State Transducer, der die Steigerungsformen (Positiv, Komparativ, Superlativ) der Adjektive "alt", "hübsch" und "gut" erkennt und eine entsprechende Analyse ausgibt. Flektionen nach Numerus und Kasus sollen nicht berücksichtigt werden. D.h., es sollen folgende Analysen abgedeckt sein:

Ausgabe: lexikalischen Ebene	Eingabe: Oberflächenebene
alt Pos	alt
alt Komp	älter
alt Sup	ältesten
hübsch Pos	hübsch
hübsch Komp	hübscher
hübsch Sup	hübschesten
gut Pos	gut
gut Komp	besser
gut Sup	besten

Sie dürfen mehrere aufeinander folgende Buchstaben zusammenfassen.

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

