

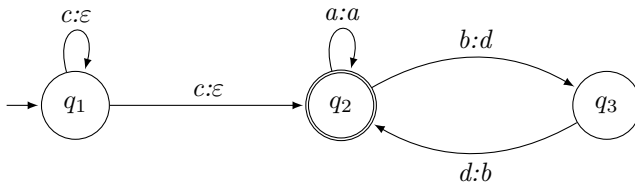
Einführung in die Computerlinguistik

Hausaufgabe zu FST und Morphologie, Abgabe 10.05.2021

Laura Kallmeyer

Sommer 2021, Heinrich-Heine-Universität Düsseldorf

Aufgabe 1 Betrachten Sie folgenden FST:



1. Auf welche Strings bildet der FST die folgenden Eingaben ab? Akzeptiert er dabei die jeweilige Eingabe?
 (a) *cc* (b) *cab* (c) *ccaabdabd* (d) *cbdbdabd*
2. Welche Strings akzeptiert dieser FST und wie transformiert er sie? (Eine Beschreibung der Transformation in Worten genügt.)

Lösung:

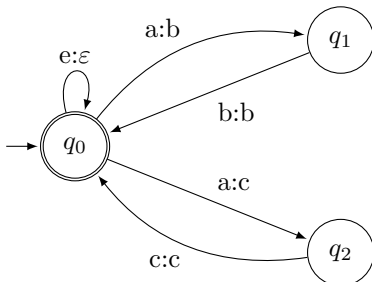
1. (a) ε , Akzeptanz
 (b) *ad*, keine Akzeptanz, da kein Endzustand erreicht wird
 (c) *aadbadb*, Akzeptanz
 (d) *dbdbadb*, Akzeptanz
2. Akzeptiert wird die von $c^+(a|bd)^*$ denotierte Sprache. Jedes *c* wird gelöscht, jedes *a* in die Ausgabe kopiert, jedes *b* durch *d* ersetzt und jedes *d* durch *b* ersetzt.

Aufgabe 2 Erstellen Sie einen Finite State Transducer, der die von $(ab|ac|e)^*$ denotierte Sprache akzeptiert, und der jedes *a*, das vor einem *b* steht, durch *b* ersetzt, jedes *a*, das vor einem *c* steht, durch *c* ersetzt, und jedes *b* oder *c* in der Eingabe in die Ausgabe kopiert. Jedes *e* in der Eingabe soll gelöscht werden.

Z.B.

Eingabe	<i>e</i>	<i>eeeabab</i>	<i>ababeeaceab</i>
Ausgabe	ε	<i>bbbb</i>	<i>bbbcbcb</i>

Lösung:



Aufgabe 3

Erstellen Sie einen Finite State Transducer, der die finiten Formen im Präsens Indikativ der französischen Verben "venir" (= kommen), "tenir" (= halten) und "aller" (= gehen) jeweils in allen drei Personen un im Numerus Sg oder Pl erkennt und eine entsprechende Analyse ausgibt. D.h., es sollen folgende Analysen abgedeckt sein:

Ausgabe: lexikalischen Ebene	Eingabe: Oberflächenebene
venir 1 Sg	viens
venir 2 Sg	viens
venir 3 Sg	vient
venir 1 Pl	venons
venir 2 Pl	venez
venir 3 Pl	vienment
tenir 1 Sg	tiens
tenir 2 Sg	tiens
tenir 3 Sg	tient
tenir 1 Pl	tenons
tenir 2 Pl	tenez
tenir 3 Pl	tiennent
aller 1 Sg	vais
aller 2 Sg	vas
aller 3 Sg	va
aller 1 Pl	allons
aller 2 Pl	allez
aller 3 Pl	vont

Sie dürfen mehrere aufeinander folgende Buchstaben zusammenfassen.

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

