

Einführung in die Computerlinguistik

Hausaufgabe 4, Abgabe 07.05.2012

Laura Kallmeyer

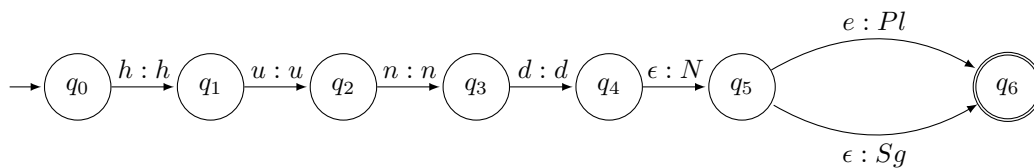
SS 2012, Heinrich-Heine-Universität Düsseldorf

Aufgabe 1 Gegeben ist das folgende Tupel für einen finite state transducer:

- $Q = \{q_0, q_1, q_2, q_3, q_4, q_5, q_6\}$
- $\Sigma = \{h, u, n, d, e\}$
- $\Delta = \{h, u, n, d, N, Sg, Pl\}$
- 1. $\delta(q_0, h, h) = \{q_1\}$ 2. $\delta(q_1, u, u) = \{q_2\}$ 3. $\delta(q_2, n, n) = \{q_3\}$ 4. $\delta(q_3, d, d) = \{q_4\}$
5. $\delta(q_4, \epsilon, N) = \{q_5\}$ 6. $\delta(q_5, \epsilon, Sg) = \{q_6\}$ 7. $\delta(q_5, e, Pl) = \{q_6\}$
- Startzustand ist q_0 .
- $F = \{q_6\}$

1. Geben Sie die graphische Darstellung dieses FST an.
2. Geben Sie für beide Ausdrücke, die der Automat generieren kann, die lexikalische Ebene und die Oberflächenebene an.

Lösung:



1.

Lexikalische Ebene:

...	h	u	n	d	N	Pl	...
-----	---	---	---	---	---	----	-----

Oberfläche:

...	h	u	n	d	e	...	
-----	---	---	---	---	---	-----	--

2.

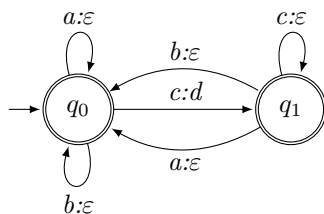
Lexikalische Ebene:

...	h	u	n	d	N	Sg	...
-----	---	---	---	---	---	----	-----

Oberfläche:

...	h	u	n	d	...	
-----	---	---	---	---	-----	--

Aufgabe 2 Betrachten Sie folgenden FST:



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

Lösung:

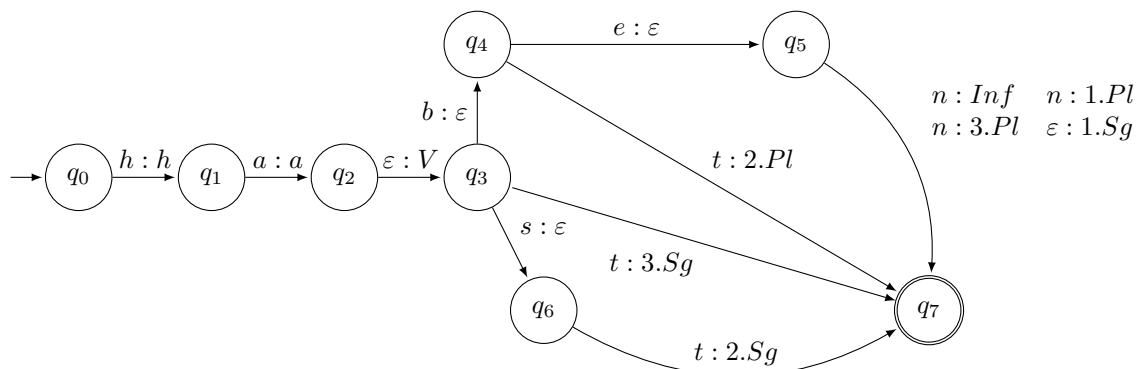
- (a) ddd (b) ε (c) ε (d) d
- Akzeptiert wird $\{a, b, c\}^*$. Ausgabe ist eine Folge von ds , wobei die Anzahl der ds genau die Anzahl der maximal langen Folgen von cs (mindestens ein c enthaltend) im Ursprungsstring sind.

Aufgabe 3

- Erstellen Sie einen Finite State Transducer, der alle flektierten Formen des Verbs "haben" im Präsens erkennt und deren morphologische Analyse ausgibt. Geben Sie das Sixtupel dieses FST an.
- Geben Sie die Oberfläche und die lexikalische Ebene von einer flektierten Form an.

Lösung:

1.



FST:

$\langle \{q_0, q_1, q_2, q_3, q_4, q_5, q_6, q_7\}, \{h, a, b, e, s, t, n\}, \{h, a, V, 1.Sg, 2.Sg, 3.Sg, 1.Pl, 2.Pl, 3.Pl, Inf\}, \delta, q_0, \{q_7\} \rangle$
mit

- (a) $\delta(q_0, h, h) = \{q_1\}$ (b) $\delta(q_1, a, a) = \{q_2\}$ (c) $\delta(q_2, \varepsilon, V) = \{q_3\}$ (d) $\delta(q_3, b, \varepsilon) = \{q_4\}$ (e) $\delta(q_3, s, \varepsilon) = \{q_6\}$ (f) $\delta(q_4, e, \varepsilon) = \{q_5\}$ (g) $\delta(q_4, t, 2.Pl) = \{q_7\}$ (h) $\delta(q_3, t, 3.Sg) = \{q_7\}$ (i) $\delta(q_5, \varepsilon, 1.Sg) = \{q_7\}$ (j) $\delta(q_5, n, Inf) = q_7$ (k) $\delta(q_5, n, 1.Pl) = \{q_7\}$ (l) $\delta(q_5, n, 3.Pl) = \{q_7\}$ (m) $\delta(q_6, t, 2.Sg) = \{q_7\}$

2. Mögliche Lösungen:

Lexikalische Ebene:	...	h	a	V	1.Sg	...
Oberfläche:	...	h	a	b	e	...

Lexikalische Ebene:	...	h	a	V	2.Sg	...
Oberfläche:	...	h	a	s	t	...

Lexikalische Ebene:	...	h	a	V	3.Sg	...
Oberfläche:	...	h	a	t	...	

Lexikalische Ebene:	...	h	a	V	1.Pl	...	
Oberfläche:	...	h	a	b	e	n	...

Lexikalische Ebene:	...	h	a	V	2.Pl	...
Oberfläche:	...	h	a	b	t	...

Lexikalische Ebene:	...	h	a	V	3.Pl	...	
Oberfläche:	...	h	a	b	e	n	...