

Matematik för språkteknologer

Övningsuppgifter: Automater och Markov-modeller

2009-12-14

Uppgift 1

- (a) Förklara följande språktekniska begrepp:
- (i) Alfabet
 - (ii) Strängar
 - (iii) Språk
- (b) Förklara skillnaden mellan en *deterministisk* finit automat och en *icke-deterministisk* finit automat.
- (c) Vad kännetecknar en Mealyautomat respektive en Mooreautomat?
- (d) Beskriv förhållandet mellan finita automater, reguljära uttryck och reguljära språk.

Uppgift 2

- (a) Hur ser tillståndsdiagrammet ut för automaten $M = (Q, \Sigma, \delta, q_0, F)$, där $Q = \{s_0, s_1, s_2, s_3\}$, $\Sigma = \{a, b\}$, $q_0 = s_0$, $F = \{s_3\}$, och δ ges av tabellen nedan?

	a	b
s_0	s_0	s_1
s_1	s_1	s_2
s_2	s_2	s_3
s_3	s_3	s_3

- (b) Uttryck i ord vilka strängar automaten accepterar.

Uppgift 3

- (a) Konstruera en automat som känner igen språket av alla binära strängar som börjar och slutar på 0.
- (b) Konstruera en automat som känner igen språket av alla binära strängar som innehåller delsträngen 101.

Uppgift 4

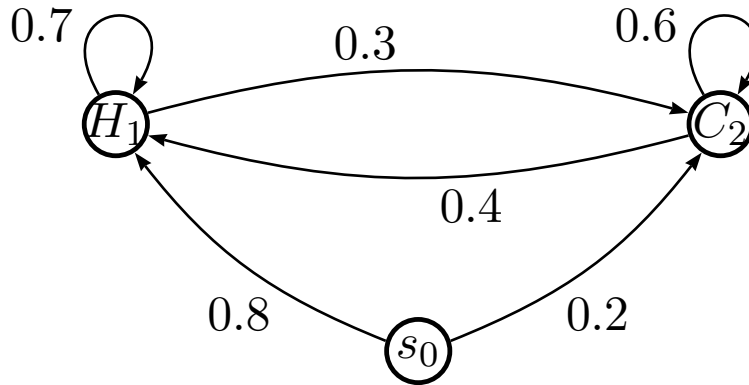
Konstruera automater som känner igen följande reguljära språk:

(i) $(ab)^*abc(ab)^*$

(ii) $(b^*)|(b^*)a(b^*)$

Uppgift 5

Automaten nedan tillsammans med dess emissionsmatris återger Markov-modellen i Jurafsky & Martin (sid 212). S_0 är starttillståndet. H_1 (Hot) och C_2 (Cold) är sluttillstånd.



	1	2	3
S_0	–	–	–
H_1	0.2	0.4	0.4
C_2	0.5	0.4	0.1

- Beräkna den förenade sannolikheten för händelserna $3\ 1\ 3$ och den dolda tillståndssekvensen *hot cold hot*.
- Välj någon annan möjlig tillståndsekvens för händelserna $3\ 1\ 3$ (*cold cold cold*, *cold cold hot*, ...) och beräkna den förenade sannolikheten för observationssekvensen $3\ 1\ 3$ och den tillståndsekvens du valt.
- Hur räknar vi ut den totala sannolikheten för observationssekvensen $3\ 1\ 3$ givet modellen? Diskutera.
- Hur räknar vi ut den mest sannolika dolda tillståndsekvensen för observationssekvensen $3\ 1\ 3$? Diskutera.