

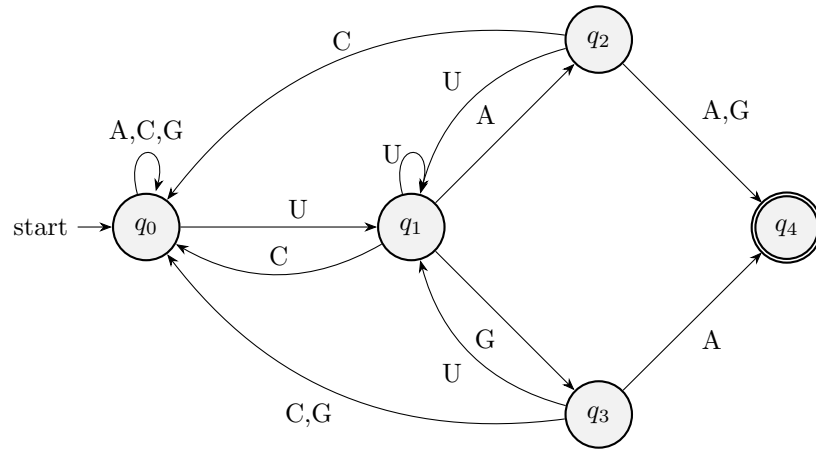
# Algoritmi v bioinformatiki - 1. Domača naloga

Jan Panjan

March 25, 2025

1. *Konstruirajte deterministični končni avtomat, ki v mRNK materialu prepozna zaključne kodone.*

(a) Grafično:



(b) S formalnim opisom peterike  $[\Sigma, Q, q_0, F, \delta]$ :

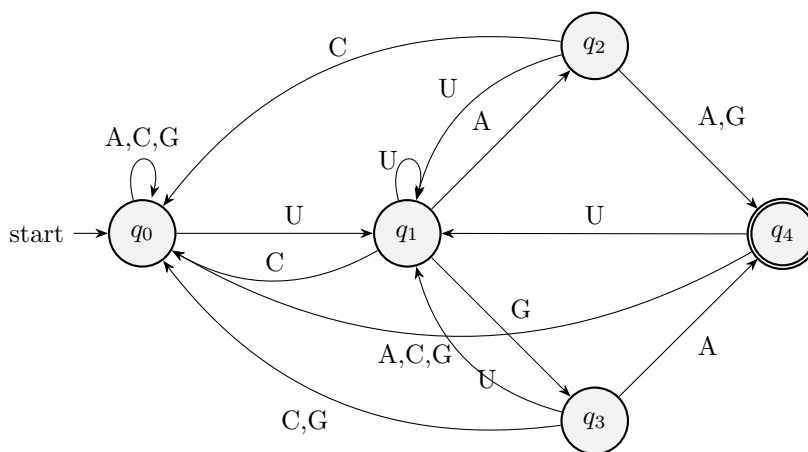
- $\Sigma = \{A, U, C, G\}$
- $Q = \{q_0, q_1, q_2, q_3, q_4\}$
- $q_0 = q_0$
- $F = \{q_4\}$
- stanja so pisana samo s številko in pot ki pripelje do končnega stanja je označena z rdečo barvo, da je bolj berljivo.

$\delta$	A	U	C	G
0	0	1	0	0
1	2	1	0	3
2	4	1	0	4
3	4	1	0	0
4	/	/	/	/

2. Kako se rešitev 1. naloge spremeni, če želimo s pomočjo končnega avtomata poiskati vse pojavitve zaključnih kodonov? Zapišite algoritem in ponazorite njegovo delovanje na delu mRNK AUAUAAUGCUUGA. Koliko zaključnih kodonov vsebuje dani mRNK?

Njegovo končno stanje se spremeni, tako da ponovno začne iskati vzorec, ko pride enkrat do končnega stanja. To je vidno grafično kot povezava od  $q_4$  do  $q_0$  in spremenjene vrednosti v zadnji vrstici  $\delta$ -tabele.

(a) Grafično:



(b) S formalnim opisom peterike  $[\Sigma, Q, q_0, F, \delta]$ :

- $\Sigma = \{A, U, C, G\}$
- $Q = \{q_1, q_2, q_3, q_4\}$
- $q_0 = q_0$
- $F = \{q_4\}$

$\delta$	A	U	C	G
0	0	1	0	0
1	2	1	0	3
2	4	1	0	4
3	4	1	0	0
4	0	1	0	0

**Algoritem za iskanje STOP kodonov v mRNA vzorcu:** algoritem za izgradnjo  $\delta$ -tabele smo podali na vajah in ga ne bom ponovno napisal. Predpostavljam torej, da je tabela za naš končni avtomat že izgrajena. Iskanje vzorca v besedilu AUAUAAUGCUUGA poteka tako:

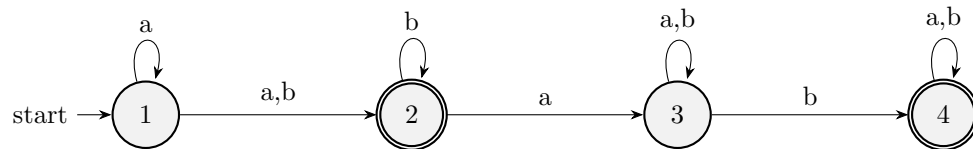
kaj je mišljeno kot algoritem tu..?

Dani vzorec mRNA vsebuje dva stop kodona: AUAUAAAUGCUUGA

3. *Konstruirajte determinističen končni avtomat za naslednji nedeterminističen končni avtomat.*

Deterministični končni avtomat iz nedeterminističnega izgradimo s pomočjo  $\delta$ -tabele tako da zapišemo vse neprazne podmnožice stanj. Za stanje npr. 1, 2 naredimo unijo sledečih stanj za  $a$  in  $b$ , t.j. 1, 2, 3 in 2., t.j. 1, 2, 3 in 2. Da se izognemo pisanju nepotrebnih stanj, naredimo tabelo samo za dosegljiva stanja.

**Nedeterminističen končni avtomat:**



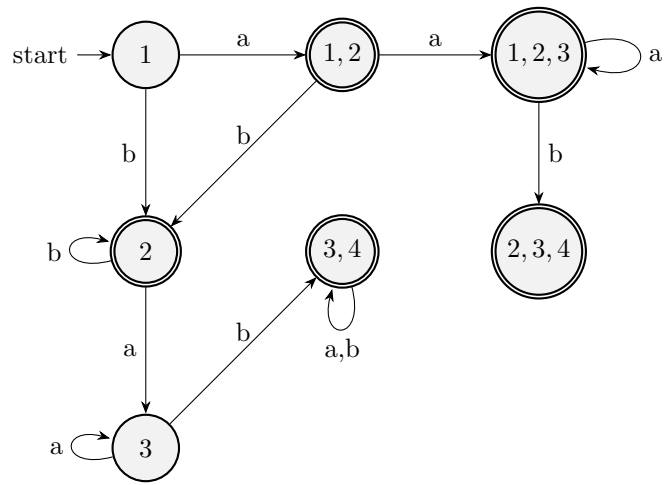
$\delta$ -tabela:

$\delta$	a	b
1	1,2	2
2	3	2
3	3	3,4
4	4	4
1,2	1,2,3	2
3,4	3,4	3,4
1,2,3	1,2,3	2,3,4
2,3,4	3,4	2,3,4

Končna stanja sta 2 in 4, zato bo imel novi končni avtomat za končna stanja vsa stanja, ki vsebujejo tako 2 ali 4.

Zgodi se, da stanje 4 odpade, saj nobeno stanje ne vodi vanj.

**Determinističen končni avtomat:**



4. haha