

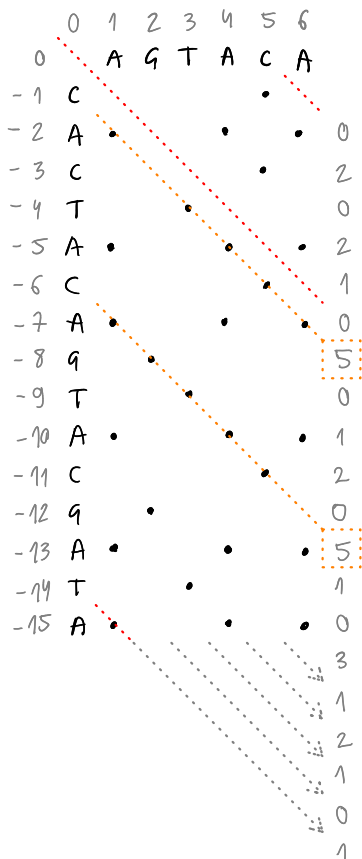
# V A J E 4. - Poravnava zaporedij

1. Poiščite najboljše ujemanje vzorca  $p = \text{AGTACA}$  v besedilu  $w = \text{CACTACAGTACGATA}$  z uporabo točkovne matrike na  $k$ -tericah za

- (a)  $k = 1$
- (b)  $k = 4$

Katera diagonalna je najbolj popularna?

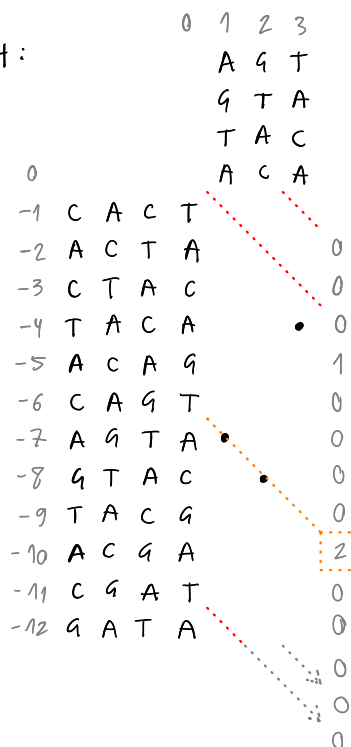
a)  $k = 1$ :



- Narediš pikice kjer je ujemanje ter diagonale na indeksu 0 in zadnjih indeksih zaporedij.
  - Za vsako diagonalo prešteješ pikice.
  - Najpopularnejša diagonalna ima največ pikic.
- Diagonali -1 in -6 (obe imata 5).

$$\# \text{diagonal} = m + n - k, \text{ kjer } m = |p| \text{ in } n = |w| \quad (6 + 15 - 1 = 20)$$

b)  $k = 4$ :



⇒ Diagonalna -6.

## Hidden Markov Model

2. Dana imamo dva kovanca,  $A, B$  in zaporedje izidov metov  $X = \text{MMCCMM}$ .

(a) Kateri kovanec smo z največjo verjetnostjo uporabili za  $X$ , če veljajo naslednje verjetnosti

Prehod:

	A	B
A	0,6	0,4
B	0,4	0,6

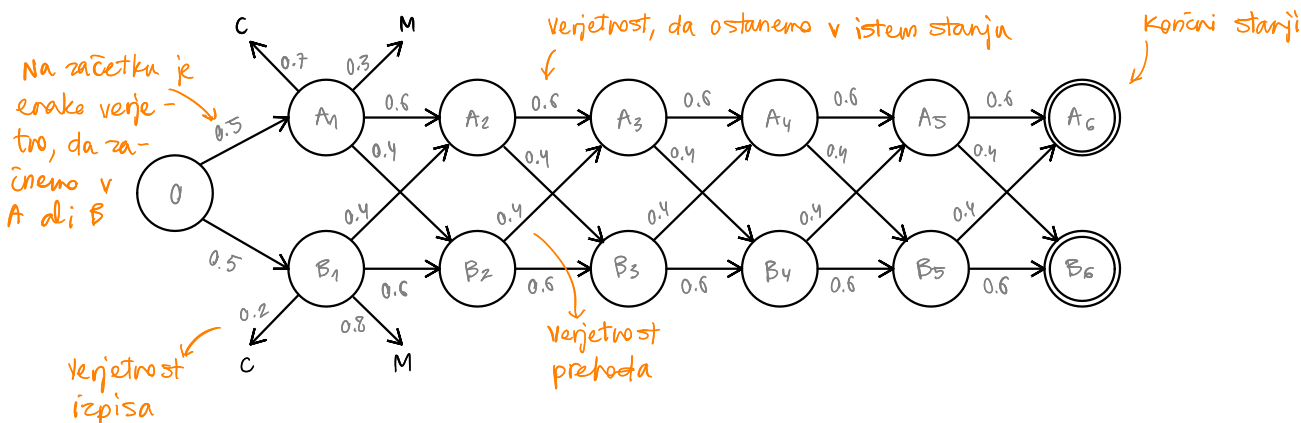
Izpis:

	C	M
A	0,7	0,3
B	0,2	0,8

(b) Kakšna je verjetnost, da smo metali samo kovanec  $A$ ?

(c) Kakšna je verjetnost, da smo metali samo kovanec  $B$ ?

↳ Iznjstil bom (a) del, saj je rekla prof. da ne bo takih dolgih izračunov. Bom pa vseeno narisal Hidden Markov Model.



b) Verjetnost, da smo metali samo A?

$X = M M C M M$   
 $A A A A A$   
 $\hookrightarrow \hookrightarrow \hookrightarrow \hookrightarrow \hookrightarrow$

- 2x vržemo C
- 4x vržemo M
- 5x ostanemo v stanju A

c) Verjetnost, da smo metali samo B?

Analogno:

$$p(B) = (0.2)^2 \times (0.8)^4 \times (0.6)^5$$

$$= 0.0012$$

$$\Rightarrow p(A) = (0.7)^2 \times (0.3)^4 \times (0.6)^5$$

$$= 0.003$$

## Globalna poravnava

3. Obravnavajmo naslednjo poravnavo:

t	C	G	T	G	A	A	T	T	C	A	T	-
s	-	G	-	-	A	C	T	T	-	A	-	C

- Koliko operacij potrebujemo, da pretvorimo s v t?
- Če smo za vsako ujemanje nagrajeni z 1 in imamo za vsako vstavljanje/brisanje kazni 2, ter za vsako zamenjavo kazni 3, kolikšna je vrednost zgornje poravnave?
- Ali je ta vrednost optimalna?
- Kako dobimo optimalno poravnavo? Ali je rešitev enolično določena?
- Določite optimalno lokalno poravnavo.

a)

t	C	G	T	G	A	A	T	T	C	A	T	-	
s	-	G	-	-	A	C	T	T	-	A	-	C	
		V	U	V	V	U	Z	U	U	V	U	V	D

↑  
ujemanje

- 3 možne operacije: brisanje, vstavljanje, zamenjava
- Za  $s \rightarrow t$  moramo vstavljati razaj tam kjer so vrzeli v s
- Če je vrzel v t, izbrisemo v s
- Če je neujemanje, moramo zamenjati v s

skupaj = 7 operacij

b) Ujemanje = +1

Vstavljanje/brisanje = -2

Zamenjava = -3

$$\Rightarrow 5 \times V + 5 \times U + 1 \times D + 1 \times Z$$

$$= 5 \times (-2) + 5 \times (1) + 1 \times (-2) + 1 \times (-3)$$

$$= -10 + 5 - 2 - 3$$

$$= -10$$

c) Ne morem reči dokler ne

naredim globalne poravnave z Needleman-Wunsch algoritmom.

Kazni in nagrade veljajo od primera (b): ujemanje = +1, brisanje/vstavljanje = -2, zamenjava = -3

t \ s	-	C	G	T	G	A	A	T	T	C	A	T
-	0	-2	-4	-6	-8	-10	-12	-14	-16	-18	-20	-22
G	-2	-4	-1	-3	-5	-7	-9	-11	-13	-15	-17	-19
A	-4	-6	-3	-5	-7	-4	-6	-8	-10	-12	-14	-16
C	-6	-3	-5	-7	-9	-6	-8	-10	-12	-9	-11	-13
T	-8	-5	-7	-4	-6	-8	-10	-7	-9	-11	-13	-10
T	-10	-7	-8	-6	-7	-9	-11	-9	-6	-8	-10	-12
A	-12	-9	-10	-8	-9	-6	-8	-10	-8	-9	-7	-9
C	-14	-11	-12	-10	-11	-8	-9	-11	-10	-7	-9	-10

Sledimo puščicam razaj do začetka da najdemo celotno poravnavo.

Ocena poravnave je očitna iz zadnjega kvadratka: -10

Ker sta oceni isti in poravnava z algoritmom je enaka prejšnji (res? nisem preveril), je vrednost iz (b) optimalna.

d) Sledimo puščicam v matriki. Ne, ni enolično določena.

e) Lokalna poravnava ima podoben pristop, le da za min vrednost lahko vzamemo  $\bar{y}$ .

$s \backslash t$	-	C	G	T	G	A	A	T	T	C	A	T
-	0	0	0	0	0	0	0	0	0	0	0	0
G	0	0	1	0	1	0	0	0	0	0	0	0
A	0	0	0	0	0	2	1	0	0	0	1	0
C	0	1	0	0	0	0	0	0	0	1	0	0
T	0	0	0	1	0	0	0	1	1	0	0	1
T	0	0	0	1	0	0	0	1	2	0	0	1
A	0	0	0	0	0	1	1	0	0	0	1	0
C	0	1	0	0	0	0	0	0	0	1	0	0

