

سوال ۲: ادفا

	۱	۲	۳	۴	۵	۶
C	۲	۱	۲	۲	۲	۱
T	۱	۲	۱	۱	۲	۱
A	۲	۳	۱	۲	۱	۱
G	۱	۱	۳	۱	۱	۳

background prop
= 0.25

	۱	۲	۳	۴	۵	۶
C	0, ۲	—	0, ۲	0, ۲	0, ۲	—
T	0, ۲	0, ۲	—	0, ۲	0, ۲	0, ۲
A	0, ۲	0, ۲	—	0, ۲	—	0, ۲
G	—	—	0, ۲	۱	0, ۲	0, ۲

background prop
= 0.25

	۱	۲	۳	۴	۵	۶
C	1, ۲	—	1, ۲	1, ۲	1, ۲	—
T	0, ۱	1, ۲	—	0, ۱	1, ۲	0, ۱
A	1, ۲	۲, ۲	—	1, ۲	—	0, ۱
G	—	—	۲, ۲	—	0, ۱	۲, ۲

Q2

	۱	۲	۳	۴	۵	۶
C	0, ۲, ۱	—	0, ۲, ۱	0, ۲, ۱	0, ۲, ۱	—
T	—, 0, ۲, ۱	0, ۲, ۱	—, 0, ۲, ۱	0, ۲, ۱	0, ۲, ۱	—, 0, ۲, ۱
A	0, ۲, ۱	1, ۲, ۳	—, 0, ۲, ۱	0, ۲, ۱	—, 0, ۲, ۱	—, 0, ۲, ۱
G	—	—	1, ۲, ۳	—	—	1, ۲, ۳

سوال ۲ :-

$$H_u = - \sum_i f_{u,i} \log_2 f_{u,i} \quad I_u = \log_2 Z_0 - H_u$$

$$H_1 = - (2 \times \log_2 2 + 1 \times \log_2 1 + 2 \times \log_2 2) = -4$$

$$I_1 = \log_2 Z_0 - H_u = 4,332 + 4 = 8,332$$

$$\rightarrow A: 2 \times 8,332 = 16,664 \quad C: 2 \times 8,332 = 16,664$$

$$G: 0$$

$$T = -1 \times 8,332 = -8,332$$

سوال ۳ الف)

	۱	۲	۳	۴	۵	۶	۷	۸
A	0,۳۳۳	0,۲۲۲	0,۲۲۲	0,۲۲۲	0,۱۱۱	0,۱۱۱	0,۱۱۱	0,۳۳۳
T	0,۲۲۲	0,۳۳۳	0,۱۱۱	0,۱۱۱	0,۲۲۲	0,۱۱۱	0,۲۲۲	0,۲۲۲
C	0,۲۲۲	0,۱۱۱	0,۱۱۱	0,۲۲۲	0,۲۲۲	0,۳۳۳	0,۳۳۳	0,۱۱۱
G	0,۱۱۱	0,۱۱۱	0,۲۲۲	0,۳۳۳	0,۱۱۱	0,۲۲۲	0,۲۲۲	0,۲۲۲
-	0,۱۱۱	0,۲۲۲	0,۳۳۳	0,۱۱۱	0,۳۳۳	0,۲۲۲	0,۱۱۱	0,۱۱۱

background
Prob

$$A = 1/4 \quad B = 1/4 \quad - = 1/4$$

$$G = 1/4 \quad T = 1/4$$

	۱	۲	۳	۴	۵	۶	۷	۸
A	1,۵۲۷	1,۱۰۱۸	1,۱۰۱۸	1,۱۰۱۸	0,۱۵۰۹	0,۱۵۰۹	0,۱۵۰۹	1,۵۲۷
T	1,۱۰۱۸	1,۷۸۰	0,۱۵۹۳	0,۱۵۹۳	1,۱۱۷	0,۱۵۹۳	1,۱۱۷	1,۱۱۷
C	1,۱۰۱۸	0,۱۵۰۹	0,۱۵۰۹	1,۱۰۱۸	1,۱۰۱۸	1,۵۲۷	1,۵۲۷	0,۱۵۰۹
G	0,۱۵۹۳	0,۱۵۹۳	1,۱۱۷	1,۷۸۰	0,۱۵۹۳	1,۱۱۷	1,۱۱۷	1,۱۱۷
-	0,۱۵۹۳	1,۱۱۷	1,۷۸۰	0,۱۵۹۳	1,۷۸۰	1,۱۱۷	0,۱۵۹۳	0,۱۵۹۳

log₂ →

A	0,۴۱	0,۱۰۲۵	0,۱۰۲۵	0,۱۰۲۵	-0,۹۷۴	-0,۹۷۴	-0,۹۷۴	0,۴۱
T	0,۱۵۹۳	-0,۹۷۴	-0,۱۷۵۳	-0,۱۷۵۳	0,۲۴۷	-0,۱۷۵۳	0,۲۴۷	0,۲۴۷
C	0,۱۰۲۵	-0,۹۷۴	-0,۹۷۴	0,۱۰۲۵	0,۱۰۲۵	0,۴۱	0,۴۱	-0,۹۷۴
G	-0,۱۷۵۳	-0,۱۷۵۳	0,۲۴۷	0,۱۷۵۳	-0,۱۷۵۳	0,۲۴۷	0,۲۴۷	0,۲۴۷
-	-0,۱۷۵۳	0,۲۴۷	0,۱۷۵۳	-0,۱۷۵۳	0,۱۷۵۳	0,۲۴۷	-0,۱۷۵۳	*

AA-CTCTG

سوال ۳: -

اعداد هر کارکتر را در

درستونش جمع کرده

$$0,41 + 0,025 + 0,187 + 0,025 + 0,247$$

$$+ 0,41 + 0,247 + 0,187 = 2,182$$

$$2,182 = 7,175$$

نسبت احتمالات

سوال ۳: ج

در داخل هرستون کاراکتری با max مقدار، اسی با هم

که به دو جواب می رسیم

AT-G-CCG

AT-G-CCT

سوال ۱۴ الف)

$$\alpha_1(i) = \pi_1 b_i(x_1) \quad 1 \leq i \leq N$$

$$\alpha_t(i) = \left[\sum_{j=1}^N \alpha_{t-1}(j) a_{ji} b_i(x_t) \right] \quad \begin{matrix} 2 \leq t \leq T \\ 1 \leq i \leq N \end{matrix}$$

A T C G

S₁ 0/14 0/014 0/00414 0/00000144

S₂ 0/104 0/0344 0/00444 0/00141001

S₃ 0/104 0/0444 0/004444 0/000001444

$$P(ATCG) = \sum_{i=1}^N \alpha_G(i) = 0/0040141$$

(- : Focus)

C G A T

S₁ 0/11 0/0014 0/00141 0/0004044

S₂ 0/104 0/044 0/00141 0/0004444

S₃ 0/104 0/0044 0/00444 0/0014044

$$\delta_1(i) = \pi_i b_i(x_1) \quad 1 \leq i \leq N$$

$$\delta_t(j) = \max_{1 \leq i \leq N} [\delta_{t-1}(i) a_{ij}] b_j(x_t) \quad \begin{matrix} 2 \leq t \leq T \\ 1 \leq j \leq N \end{matrix}$$

S₁ → S₂ → S₃ → S₃

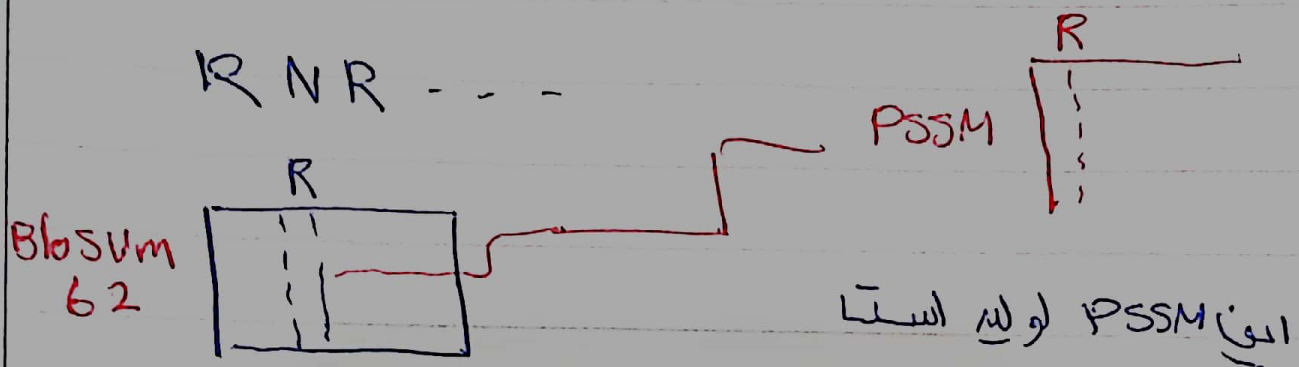
1 ≤ j ≤ N

سوال ۱: الف)

با استفاده از BLOSUM62 وون بایلی بی میلی

است ما به ازای هر ک، اکثر از این ماتریس BLOSUM

سئون آن ک، اکثر اید داشته و در PSSM می گذاریم



سوال ۱: ب) Profile drift این است که ما دنباله ای را

یاد دنباله هایی که همولوگ به استباه گرفتیم در حلقه وارد می شود

و این استباه ها یتمیع می شود و باعث خرابی نتیجه و کجیم

خطای شود

سوال ۱: ج) تعداد این False Pos_{Active} ها افزایش می یابد و در نتیجه
این مشکلات روی هم افزوده شده و خطای تمیع می شود پس بدتر است