**אלגוריתמים בביולוגיה חישובית – תרגיל 1**

1. נראה כי ישנם אופציות לעימוד שרשראות S,T באורך n.

נגדיר , אזי, לכדי לבחור את העימוד בעמדה הראשונה יש לנו את האופציות הבאות:

|  |  |  |  |
| --- | --- | --- | --- |
|  |  |  |  |
|  |  |  |  |

|  |  |  |  |
| --- | --- | --- | --- |
|  |  |  |  |
|  |  |  |  |

|  |  |  |  |
| --- | --- | --- | --- |
|  |  |  |  |
|  |  |  |  |

כאשר אופציה של לעמד גאפ מול גאפ היא טריוויאלית ולא נתחשב בה.  
כעת, ראינו שיש לנו 3 אופציות לעימוד הראשון, נמשיך את ההוכחה באינדוקציה, שאת הצעד הראשון הצגנו כרגע.

הנחת האינדוקציה: נניח נכונות עבור רצפים באורך , שלהם לפחות עימודים, ונראה עבור רצפים באורך n.

*צעד האינדוקציה:  
עבור שני רצפים באורך n ש- n-1 הנקלאוטידים הראשונים מועמדים בהם, נגיע לשתי האותיות האחרונות, כאשר הן יכולות להיות כמו בצעד הבסיס, או שתיהן אות, או שאחד הרצפים ביניהם נגמר ויש להציב גאפ מול האות הנותרת, כלומר יש 3 אפשרויות, כאשר שלושתן תקיפות לכל עימוד באורך לכן ישנן אופציות לעימוד.*

1. כאשר אנו משתמשים בשיטת affine gap penalty על-מנת להפיק את העימוד האופטימאלי בין שני רצפים, כאשר אנו נמצאים במיקום (i , j) עלינו "לזכור" את התוצאה הטובה ביותר אם אנו לא משתמשים ב-gap וכן את התוצאה הטובה ביותר אילו השתמשנו ב-gap.

על-מנת לעשות זאת נצטרך להתבונן בכמה מונים שונים עבור העמדת אל מול :

F(i , j) – תוצאה אם עומד מול

G(i , j) – תוצאה אם עומד אל מול gap אחרי הופעת

H(i , j) – תוצאה אם עומד אל מול gap אחרי

נגדיר את V(i , j) להיות התוצאה הטובה ביותר עבור העמדת אל מול .

נגדיר את Ptr(i , j) להיות העימוד האופטימאלי עבור העמדת אל מול .

פסואודו-קוד עבור אלגוריתם Needleman-Wunsch בשיטת affine gap penalty:(תחת ההנחה כי d > e)

אתחול:

V(0 , 0) = 0

V(i , 0) = d + (i-1)\*e

V(0 , j) = d + (j-1)\*e

איטרציה:

For each :

For each :

V(i , j) = max{F(i , j), G(i , j), H(i , j)}

F(i , j) = V(i-1 , j-1) + s(, )

G(i , j) = max

H(i , j) = max

Ptr(i , j) =

סיום ואופן החזרת העימוד האופטימאלי:

V(M,N) הינה התוצאה הטובה ביותר של העימוד האופטימאלי.

Ptr(i , j) הינו העימוד האופטימאלי.