

Statističko i vjerojatnosno modeliranje DNA nizova kroz primjer na palindromima

Ivo Ugrina

Palindromi su većini poznati kao nizovi slova koji se čitaju ili izgovaraju isto od početka ili kraja. U DNA nizovima palindrome definiramo na sličan način kao i u prirodnim jezicima. Isto čitanje s oba kraja je nužno, ali uz prethodnu operaciju komplementiranja znakova. Recimo, za niz A C G T bi definiranjem komplementarnosti putem prirodnog uparivanja baza $A \sim T$ i $C \sim G$ definirali komplementaran niz sa T G C A te bi rekli da je početni niz palindrom jer se čita od početka isto kao i njemu komplementarni od kraja.

Potreba za definiranjem palindroma u DNA nizovima, i proučavanje istih, može se činiti čudnom. Međutim, ako primijetimo da palindromi u DNA nizu opisuju spajanje između dva dijela DNA niza onda nam motivacija postaje bliža. Budući da su palindromi značajni za funkcioniranje genoma, korisno je imati nekakav oblik statističkog testa koji bi odavao imali palindroma određene duljine značajno više ili manje od očekivanog broja unutar nekog npr. roda koji se proučava. Iskaz rezultata koji omogućavaju konstrukciju takvog testa bit će prezentiran u ovom predavanju.