

SVEUČILIŠTE U ZAGREBU  
FAKULTET ELEKTROTEHNIKE I RAČUNARSTVA

SEMINAR

# Poravnanje para sekvenci korištenjem HMM

*Petra Dunja Grujić Ostojić , Tea Čutić*

*Voditelj: izv. prof. dr. sc. Mirjana Domazet-Lošo*

Zagreb, siječanj 2024.

# SADRŽAJ

<b>1. Uvod</b>	<b>1</b>
<b>2. Opis algoritma</b>	<b>2</b>
2.1. Baum-Welch algoritam . . . . .	2
2.2. Viterbi algoritam . . . . .	2
<b>3. Analiza rezultata</b>	<b>3</b>
<b>4. Literatura</b>	<b>4</b>

# 1. Uvod

Jedan od najtežih, a time i najistraživanijih, problema u području bioinformatike je problem poravnanja dviju sekvenci. Poravnanje je ključan korak u uspoređivanju i analizi DNA, RNA i proteinskih sekvenci te omogućuje uvid u zajedničke evolucijske korijene ili indikatore genetskih bolesti.

Ovaj rad bavi se implementacijom algoritma za poravnanje koji koristi skrivene Markovljeve modele (eng. Hidden Markov Models - HMM). Skriveni Markovljevi modeli su probabilistički modeli korišteni za predstavljanje stohastičkih procesa. Ovaj model sastoji se od više različitih stanja koja, u kontekstu modeliranja odnosa dvaju znakovnih nizova, mogu predstavljati stanja podudaranja, umetanja i brisanja. Vezama između tih stanja označene su vjerojatnosti prelaska iz jednog stanja u drugo ili ostanka u istom stanju. U ovom su radu korištena tri algoritma: Baum-Welch algoritam, Viterbijev algoritam te Needleman-Wunsch algoritam. Od navedenih algoritama, za algoritme Baum-Welch i Viterbi pisana je vlastita implementacija dok je algoritam Needleman-Wunsch preuzet iz javno dostupnog izvora te se koristi za procjenu kvalitete poravnanja kojeg dobijemo u našim rezultatima.

Drugo poglavlje rada bavi se detaljnijim opisom algoritama Baum-Welch i Viterbi te jednostavnim primjerom njihovog rada. Treće poglavlje bavi se analizom dobivenih rezultata te njihovom usporedbom s rezultatima dobivenima s pomoću algoritma Needleman-Wunsch.

## 2. Opis algoritma

### 2.1. Baum-Welch algoritam

### 2.2. Viterbi algoritam

```
if (patt[i] = t[v]){  
    Cv' = min({Cu/(u,v)eE}u{i-1})  
} else {  
    Cv' = 1 + min(Cv, min Cu/(u,v)eE)  
}
```

**Listing 2.1:** Ažuriranje vrijednosti Cv

### **3. Analiza rezultata**

## 4. Literatura

- [1] Veli Makinen Mikko Rautiainen i Tobias Marschall. *Bit-parallel sequence-to-graph alignment*. 2019.
- [2] Gonzalo Navarro. *Improved approximate pattern matching on hypertex*. University of Chile, 1999.