

# Задачи оценивания значимости выравнивания при помощи скрытых марковских моделей

Власенко Даниил Владимирович, гр.19.Б04-мм

Научный руководитель: к.ф.-м.н. Коробейников А.И.

Санкт-Петербургский государственный университет  
Прикладная математика и информатика  
Вычислительная стохастика и статистические модели

Отчет по производственной практике

Санкт-Петербург, 2022

Оценивание значимости выравнивания

Задачи оценивания значимости выравнивания  
при помощи скрытых марковских моделей

Власенко Даниил Владимирович, гр.19.Б04-мм  
Научный руководитель: к.ф.-м.н. Коробейников А.И.

Санкт-Петербургский государственный университет  
Прикладная математика и информатика  
Вычислительная стохастика и статистические модели

Отчет по производственной практике  
Санкт-Петербург, 2022

Научный руководитель к.ф.-м.н., Коробейников А.И.,  
кафедра статистического моделирования

Пусть дан алфавит символов  $\Sigma$ .

## Определение

Последовательностью длины  $L$  над алфавитом  $\Sigma$  будем называть такой элемент  $X$ , что  $X \in \Sigma^L$ . Последовательностью  $X$  над алфавитом  $\Sigma$  будем называть такой  $X$ , что  $X \in \bigcup_{L=0}^{\infty} \Sigma^L$ .

## Определение

Парное выравнивание последовательностей называется отображение  $Q : (\bigcup_{L_1=0}^{L_1=\infty} \Sigma^{L_1} \times \bigcup_{L_2=0}^{L_2=\infty} \Sigma^{L_2}) \rightarrow (\Sigma^{\max(L_1, L_2)} \times \Sigma^{\max(L_1, L_2)})$ , такое что:

1. Возможны вставки символа — в последовательностях.
2. Вставка — на одинаковых позициях в обеих последовательностях запрещена.
3. Порядок изначальных символов внутри последовательностей сохраняется.

## Оценивание значимости выравнивания

— Введение

### Введение

Пусть дан алфавит символов  $\Sigma$ .

#### Определение

Последовательностью длины  $L$  над алфавитом  $\Sigma$  будем называть такой элемент  $X$ , что  $X \in \Sigma^L$ . Последовательностью  $X$  над алфавитом  $\Sigma$  будем называть такой  $X$ , что  $X \in \bigcup_{L=0}^{\infty} \Sigma^L$ .

#### Определение

Парное выравнивание последовательностей называется отображение  $Q : (\bigcup_{L_1=0}^{L_1=\infty} \Sigma^{L_1} \times \bigcup_{L_2=0}^{L_2=\infty} \Sigma^{L_2}) \rightarrow (\Sigma^{\max(L_1, L_2)} \times \Sigma^{\max(L_1, L_2)})$ , такое что:

1. Возможны вставки символа — в последовательностях.
2. Вставка — на одинаковых позициях в обеих последовательностях запрещена.
3. Порядок изначальных символов внутри последовательностей сохраняется.

Пусть дан алфавит символов  $\Sigma$ .

## Определение

Последовательностью длины  $L$  над алфавитом  $\Sigma$  будем называть такой элемент  $X$ , что  $X \in \Sigma^L$ . Последовательностью  $X$  над алфавитом  $\Sigma$  будем называть такой  $X$ , что  $X \in \bigcup_{L=0}^{\infty} \Sigma^L$ .

## Определение

Парное выравнивание последовательностей называется отображение  $Q : (\bigcup_{L_1=0}^{L_1=\infty} \Sigma^{L_1} \times \bigcup_{L_2=0}^{L_2=\infty} \Sigma^{L_2}) \rightarrow (\Sigma^{\max(L_1, L_2)} \times \Sigma^{\max(L_1, L_2)})$ , такое что:

1. Возможны вставки символа — в последовательностях.
2. Вставка — на одинаковых позициях в обеих последовательностях запрещена.
3. Порядок изначальных символов внутри последовательностей сохраняется.

A	C	E	A	A	F	A	E
C	E	A	F	D	C	E	

  

A	C	E	A	A	F	A	—	E
—	C	E	A	—	F	D	C	E

Рис. 1: Последовательности до и после парного выравнивания.

Примем множество  $\bigcup_{L_1=0}^{L_1=\infty} \Sigma^{L_1} \times \bigcup_{L_2=0}^{L_2=\infty} \Sigma^{L_2}$  за пространство элементарных исходов  $\Omega$ . Область значений выравнивания  $Q$  обозначим как  $\bar{\Omega}$ .

## Определение

Оценкой парного выравнивания называется случайная величина  $s : \bar{\Omega} \rightarrow \mathbb{R}$ .

## Оценивание значимости выравнивания

└ Введение

### Введение

A	C	E	A	A	F	A	E
C	E	A	F	D	C	E	

  

A	C	E	A	A	F	A	—	E
—	C	E	A	—	F	D	C	E

Рис. 1: Последовательности до и после парного выравнивания.

Примем множество  $\bigcup_{L_1=0}^{L_1=\infty} \Sigma^{L_1} \times \bigcup_{L_2=0}^{L_2=\infty} \Sigma^{L_2}$  за пространство элементарных исходов  $\Omega$ . Область значений выравнивания  $Q$  обозначим как  $\bar{\Omega}$ .

### Определение

Оценкой парного выравнивания называется случайная величина  $s : \bar{\Omega} \rightarrow \mathbb{R}$ .

Примем множество  $\bigcup_{L_1=0}^{L_1=\infty} \Sigma^{L_1} \times \bigcup_{L_2=0}^{L_2=\infty} \Sigma^{L_2}$  за пространство элементарных исходов  $\Omega$ . Область значений выравнивания  $Q$  обозначим как  $\bar{\Omega}$ .

## Определение

Оценкой парного выравнивания называется случайная величина  $s : \bar{\Omega} \rightarrow \mathbb{R}$ .