



МОСКОВСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ ИМЕНИ М. В. ЛОМОНОСОВА
Факультет вычислительной математики и кибернетики
Кафедра математических методов прогнозирования

Отчет к лабораторной работе №3 по ММАТ

Студент 517 группы:
Оспанов А.М.

Москва, 2017

В данной работе была реализована модель IBM Model 1 с Prior Model. Код написан на языке Python версии 2.

Методы перевода были реализованы по методичке Д. Тэлбота. Ничего особенного данная работа не представляет. Были просто реализованы модель IBM Model 1 в соответствии с подсказками и с помощью методички Д. Тэлбота. Модель Prior Model был сделан по аналогии с Translation Model, только в соответствии с методичкой.

Модель IBM Model 1 работает точно как в задании:

```
> python scripts/word_alignment.py training/en-cs.en.all.tokens.10K
training/en-cs.cs.all.tokens.10K 10 m1
```

```
corpus log likelihood: -1119633.093
corpus log likelihood: -1040102.818
corpus log likelihood: -1003314.528
corpus log likelihood: -982945.436
corpus log likelihood: -969891.372
corpus log likelihood: -960717.282
corpus log likelihood: -953864.349
corpus log likelihood: -948519.649
corpus log likelihood: -944215.264
```

```
> python scripts/eval.py ... m1.dev.wa
recall 0.485; precision 0.506; aer 0.504
```

После реализации Prior Model результат улучшился:

```
> python scripts/word_alignment.py training/en-cs.en.all.tokens.10K
training/en-cs.cs.all.tokens.10K 10 m1
```

```
corpus log likelihood: -1151135.623
corpus log likelihood: -1032545.529
corpus log likelihood: -970212.411
corpus log likelihood: -933371.477
corpus log likelihood: -908916.422
corpus log likelihood: -891431.699
corpus log likelihood: -878269.304
corpus log likelihood: -867974.931
corpus log likelihood: -859683.165
```

```
> python scripts/eval.py ... m1.dev.wa
recall 0.568; precision 0.602; aer 0.415
```