# Assignment 4

**Pin-Jie Lin**

pili00001@stud.uni-saarland.de
https://github.com/pjlintw/CL20/tree/main/assignment4

## 1. Introduction

In the report, we demonstrate the comparison between our IBM model 1 and other implementations, namely another IBM model 1 and the second version `fast_align` . You can run it with the code:

```
python run_aligner.py –n 10000 > myIBM–1k
```

For all the evalution, we use the script `score–alignments` .

```
python score–alignments < myIBM–1k
```

## 2. Compare with other Implemenation

We compare our implementation with other on 1000, 3000 and 500k sentence pairs. It gives similar results between the two implementations.

### Comparison with other IBM Model 1 Implementation

| Models | Precision | Recall | AER |
|---|---|---|---|
| Another IBM-Model1(1k) | 25 | 23 | 76 |
| Another IBM-Model1(3k) | 26 | 24 | 74 |
| Another IBM-Model1(5k) | 27 | 25 | 74 |
| Our IBM-Model1(1k) | 24 | 22 | 76 |
| Our IBM-Model1(3k) | 25 | 24 | 75 |
| Our IBM-Model1(5k) | 26 | 25 | 74 |

# 3. Compare with `fast_align`

We evaluates our IBM model 1 with the second version of model `fast_align` on different scale datasets. We set the iteration 1 and train them on 10000, 30000, 50000, 80000 sentence pairs.

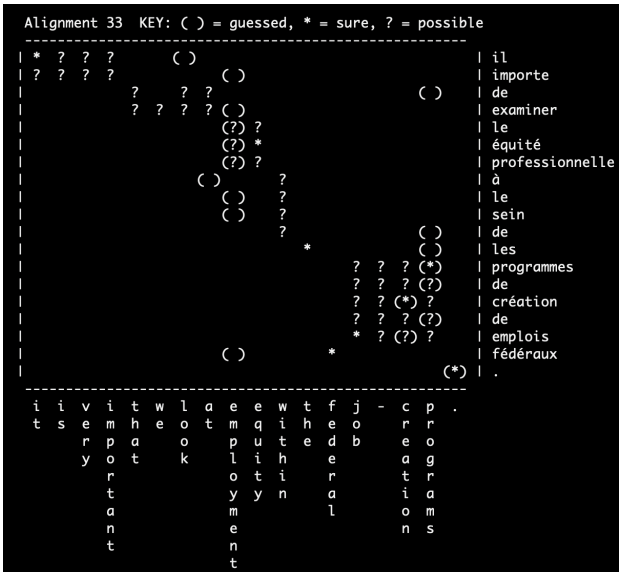
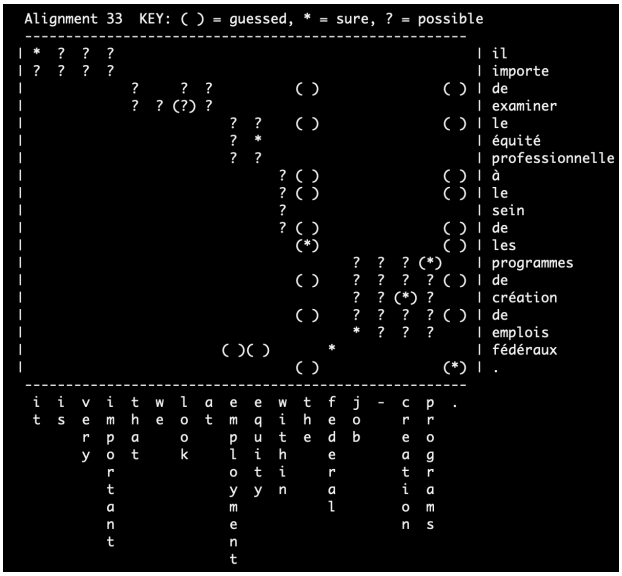
Both models perform better when increasing training examples. The `fast_align` can achieve the 28 accuracy with only 1000 examples but our IBM model requires 8x examples. In the recall score, the `fast_align` on 1k, 3k, 5k and 8k are not as good as its precision.

### Performances of IBM Model 1 and fast_align

| Models | IBM Model1(precision) | IBM Model1(recall) | IBM Model1(AER) | fast_align(precision) | fast_align(recall) | fast_align(AER) |
|---|---|---|---|---|---|---|
| 1,000 | 24 | 22 | 76 | 28 | 21 | 75 |
| 3,000 | 26 | 24 | 75 | 28 | 21 | 75 |
| 5,000 | 26 | 25 | 74 | 28 | 21 | 74 |
| 8,000 | 28 | 28 | 72 | 29 | 22 | 73 |

# 4. Visualizations

We visualize the a aligments example from our implemenation and the baseline.

| Alignments (our) | Alignment (other) |
|---|---|
|  |  |