3.

IBM Model 1

---------------------------

Average AER: 0.630

IBM Model 2

---------------------------

Average AER: 0.642

time: 63.298078 sec

The IBM model2 has a better performance than IBM model1, since it considers the p(i|j,l,m), it as a constrain for the position for each alignment. So it will selected a more accurate result.

4.

|  |  |  |  |
| --- | --- | --- | --- |
| iteration | model 1 | model 2 | time |
| 1 | 0.873 | 0.646 | 39.071674 sec |
| 2 | 0.684 | 0.644 | 48.998142 sec |
| 3 | 0.641 | 0.644 | 59.151355 sec |
| 4 | 0.63 | 0.642 | 76.088925 sec |
| 5 | 0.627 | 0.644 | 76.503074 sec |
| 6 | 0.626 | 0.647 | 96.139349 sec |
| 7 | 0.629 | 0.646 | 99.204378 sec |
| 8 | 0.631 | 0.649 | 123.862151 sec |
| 9 | 0.628 | 0.649 | 122.032705 sec |
| 10 | 0.665 | 0.65 | 118.462113 sec |
| 11 | 0.666 | 0.649 | 138.623361 sec |
| 12 | 0.666 | 0.65 | 202.062393 sec |
| 13 | 0.666 | 0.652 | 195.626067 sec |
| 14 | 0.665 | 0.652 | 178.63248 sec |
| 15 | 0.665 | 0.65 | 196.825193 sec |
| 16 | 0.665 | 0.65 | 201.416533 sec |
| 17 | 0.662 | 0.651 | 232.620126 sec |
| 18 | 0.661 | 0.651 | 241.001096 sec |
| 19 | 0.661 | 0.651 | 232.171709 sec |
| 20 | 0.661 | 0.648 | 243.248479 sec |

1.

IBM Model 1

---------------------------

Average AER: 0.873

IBM Model 2

---------------------------

Average AER: 0.646

time: 39.071674 sec

2. IBM Model 1

---------------------------

Average AER: 0.684

IBM Model 2

---------------------------

Average AER: 0.644

time: 48.998142 sec

3.

IBM Model 1

---------------------------

Average AER: 0.641

IBM Model 2

---------------------------

Average AER: 0.644

time: 59.151355 sec

4.

IBM Model 1

---------------------------

Average AER: 0.630

IBM Model 2

---------------------------

Average AER: 0.642

time: 76.088925 sec

5.

IBM Model 1

---------------------------

Average AER: 0.627

IBM Model 2

---------------------------

Average AER: 0.644

time: 76.503074 sec

6.

IBM Model 1

---------------------------

Average AER: 0.626

IBM Model 2

---------------------------

Average AER: 0.647

time: 96.139349 sec

7.

IBM Model 1

---------------------------

Average AER: 0.629

IBM Model 2

---------------------------

Average AER: 0.646

time: 99.204378 sec

8.

IBM Model 1

---------------------------

Average AER: 0.631

IBM Model 2

---------------------------

Average AER: 0.649

time: 123.862151 sec

9.

IBM Model 1

---------------------------

Average AER: 0.628

IBM Model 2

---------------------------

Average AER: 0.649

time: 122.032705 sec

10.

IBM Model 1

---------------------------

Average AER: 0.665

IBM Model 2

---------------------------

Average AER: 0.650

time: 118.462113 sec

11.

IBM Model 1

---------------------------

Average AER: 0.666

IBM Model 2

---------------------------

Average AER: 0.649

time: 138.623361 sec

12.

IBM Model 1

---------------------------

Average AER: 0.666

IBM Model 2

---------------------------

Average AER: 0.650

time: 202.062393 sec

13.

IBM Model 1

---------------------------

Average AER: 0.666

IBM Model 2

---------------------------

Average AER: 0.652

time: 195.626067 sec

14.

IBM Model 1

---------------------------

Average AER: 0.665

IBM Model 2

---------------------------

Average AER: 0.652

time: 178.63248 sec

15.

IBM Model 1

---------------------------

Average AER: 0.665

IBM Model 2

---------------------------

Average AER: 0.650

time: 196.825193 sec

16.

IBM Model 1

---------------------------

Average AER: 0.665

IBM Model 2

---------------------------

Average AER: 0.650

time: 201.416533 sec

17.

IBM Model 1

---------------------------

Average AER: 0.662

IBM Model 2

---------------------------

Average AER: 0.651

time: 232.620126 sec

18.

IBM Model 1

---------------------------

Average AER: 0.661

IBM Model 2

---------------------------

Average AER: 0.651

time: 241.001096 sec

19.

IBM Model 1

---------------------------

Average AER: 0.661

IBM Model 2

---------------------------

Average AER: 0.651

time: 232.171709 sec

20.

IBM Model 1

---------------------------

Average AER: 0.661

IBM Model 2

---------------------------

Average AER: 0.648

time: 243.248479 sec

Berkeley Aligner

---------------------------

Average AER: 0.531

Part B time: 25.293862 sec

((u'vorliegt', u','), 1.4604394381929528e-22),

((u'vorliegt', u'Mrs'), 5.081549535272721e-105),

((u'vorliegt', u'have'), 5.430908982890409e-07),

((u'vorliegt', u'when'), 4.039425525491705e-33),

((u'vorliegt', u'look'), 1.6582486702211533e-27),

((u'vorliegt', u'I'), 0.0068116735654250435),

((u'vorliegt', u'your'), 1.9093149753523917e-86),

((u'vorliegt', u'case'), 0.007153833192379151),

((u'vorliegt', u'to'), 3.3496870694925958e-124),

((u'vorliegt', u'the'), 4.589120968457509e-09),

((u'vorliegt', u'Schroedter'), 1.409397298570379e-132), ((u'vorliegt', u'facts'), 7.004465865224898e-111),

((u'vorliegt', u'be'), 2.0799560139995686e-101),

((u'vorliegt', u'into'), 1.008799835472676e-124),

((u'vorliegt', u'pleased'), 4.150520549097627e-30), ((u'vorliegt', u'received'), 1.0032022954457874e-77), ((u'vorliegt', u'of'), 5.78276061883093e-08),

((u'vorliegt', u'Yes'), 6.795706836954967e-91),

((u'vorliegt', u'this'), 2.9222669980970906e-15),

((u'vorliegt', u'shall'), 2.9361019525586247e-40),

((u'vorliegt', u'.'), 5.473795731395001e-67),

((u'vorliegt', u'letter'), 4.1057551806597464e-114)

((u'vorliegt', u'be'), 0.0),

((u'vorliegt', u','), 0.0),

((u'vorliegt', u'Mrs'), 9.306515606470729e-40),

((u'vorliegt', u'into'), 3.03095233988371e-22),

((u'vorliegt', u'pleased'), 1.936035459039568e-130), ((u'vorliegt', u'have'), 0.0),

((u'vorliegt', u'received'), 7.933379137569352e-40), ((u'vorliegt', u'of'), 2.214399457043937e-21),

((u'vorliegt', u'Yes'), 3.344779023048302e-161),

((u'vorliegt', u'when'), 1.4149303664479086e-304),

((u'vorliegt', u'look'), 3.614237755331764e-08),

((u'vorliegt', u'Schroedter'), 0.0),

((u'vorliegt', u'your'), 2.4924105150634397e-218),

((u'vorliegt', u'this'), 3.0948082888807853e-22),

((u'vorliegt', u'shall'), 0.0),

((u'vorliegt', u'I'), 0.0),

((u'vorliegt', u'.'), 0.0),

((u'vorliegt', u'case'), 0.0),

((u'vorliegt', u'to'), 0.0),

((u'vorliegt', u'the'), 2.9549712751468876e-23),

((u'vorliegt', u'letter'), 1.2085806760625622e-44), ((u'vornehmen', u'harmonisation'), 1.4300559337244884e-233),

B.py

* In your report, give an example of a sentence pair that the Berkeley Aligner performs better on than the IBM models, and explain why you think this is the case.

[u'Madam', u'President', u',', u'on', u'a', u'point', u'of', u'order', u'.']

[u'Frau', u'Pr\xe4sidentin', u',', u'zur', u'Gesch\xe4ftsordnung', u'.']

This sentence in aligned\_sents is a good example. It has 0.571428571 AER in IBM model2 and just 0.2 AER in Berkeley Aligner. After observing it’s alignment “0-0 1-1 2-2 3-3 4-4 4-5 4-6 4-7 5-8”. The English->German alignment shows, each German words will combine with an English words, but German -> English alignment will have multiply English words to one German words. After applying the Berkeley aligner algorithm, the model will got higher accuracy based on the one-to-one correspondence in English🡪German model. After averaging the high performance in English 🡪 German model, the total accuracy raise a lot.

* 6)  (Extra Credit) Think of a way to improve upon the Berkeley Aligner model. Specifically examine the way we quantify agreement between the two models. In our implementation, we computed agreement as the average expected count of the two models. Implement an improved Berkeley Aligner model that computes agreement in a better way. There is skeleton code in EC.py (same as for B.py) Compute the average AER for the first 50 sentences. Compare to the other models. Again, this part is optional but if your implementation is interesting and shows improved performance, you will be eligible for bonus points.