ACCURACY -0.001:

Iter= 1000, Average Loss= 13.620213, Average Accuracy= 0.50%

['spoke', '.', 'then'] - [the] vs [was]

Iter= 2000, Average Loss= 13.756237, Average Accuracy= 0.90%

['?', 'the', 'mice'] - [looked] vs [another]

Iter= 3000, Average Loss= 13.697954, Average Accuracy= 0.70%

['all', 'very', 'well'] - [,] vs [was]

Iter= 4000, Average Loss= 13.804327, Average Accuracy= 0.70%

['applause', ',', 'until'] - [an] vs [was]

Iter= 5000, Average Loss= 13.824120, Average Accuracy= 1.20%

['we', 'should', 'always'] - [know] vs [was]

Iter= 6000, Average Loss= 13.818810, Average Accuracy= 1.10%

['therefore', ',', 'to'] - [propose] vs [until]

Iter= 7000, Average Loss= 13.657680, Average Accuracy= 0.70%

['receive', 'some', 'signal'] - [of] vs [receive]

Iter= 8000, Average Loss= 13.723888, Average Accuracy= 0.90%

['in', 'which', 'the'] - [enemy] vs [another]

Iter= 9000, Average Loss= 13.844065, Average Accuracy= 0.90%

['the', 'case', '.'] - [you] vs [until]

Iter= 10000, Average Loss= 13.792329, Average Accuracy= 0.70%

['but', 'at', 'last'] - [a] vs [another]

Optimization Finished!

Elapsed time: 18.905889749526978 sec

ACCURACY -0.005:

Iter= 1000, Average Loss= 15.282271, Average Accuracy= 1.00%

['?', 'the', 'mice'] - [looked] vs [bell]

Iter= 2000, Average Loss= 15.440006, Average Accuracy= 0.70%

['mouse', 'got', 'up'] - [and] vs [young]

Iter= 3000, Average Loss= 15.452259, Average Accuracy= 0.30%

['and', 'could', 'easily'] - [retire] vs [bell]

Iter= 4000, Average Loss= 15.493158, Average Accuracy= 0.20%

['we', 'should', 'always'] - [know] vs [bell]

Iter= 5000, Average Loss= 15.569394, Average Accuracy= 0.00%

['by', 'a', 'ribbon'] - [round] vs [bell]

Iter= 6000, Average Loss= 15.298655, Average Accuracy= 0.80%

['of', 'her', 'approach'] - [,] vs [bell]

Iter= 7000, Average Loss= 15.232893, Average Accuracy= 1.20%

['chief', 'danger', 'consists'] - [in] vs [bell]

Iter= 8000, Average Loss= 15.319158, Average Accuracy= 0.60%

['he', 'had', 'a'] - [proposal] vs [young]

Iter= 9000, Average Loss= 15.390923, Average Accuracy= 0.80%

['.', 'some', 'said'] - [this] vs [us]

Iter= 10000, Average Loss= 15.402814, Average Accuracy= 1.00%

['general', 'council', 'to'] - [consider] vs [bell]

Optimization Finished!

Elapsed time: 19.6634202003479 sec

AdamOptimizer

Iter= 1000, Average Loss= 14.589781, Average Accuracy= 0.80%

['nobody', 'spoke', '.'] - [then] vs [to]

Iter= 2000, Average Loss= 14.132907, Average Accuracy= 1.40%

['and', 'said', 'that'] - [is] vs [escape]

Iter= 3000, Average Loss= 14.147736, Average Accuracy= 1.40%

['when', 'she', 'was'] - [about] vs [we]

Iter= 4000, Average Loss= 14.384055, Average Accuracy= 1.20%

['a', 'ribbon', 'round'] - [the] vs [we]

Iter= 5000, Average Loss= 14.226566, Average Accuracy= 1.00%

[',', 'therefore', ','] - [to] vs [escape]

Iter= 6000, Average Loss= 13.884891, Average Accuracy= 1.40%

['which', 'the', 'enemy'] - [approaches] vs [was]

Iter= 7000, Average Loss= 14.281076, Average Accuracy= 1.40%

['the', 'sly', 'and'] - [treacherous] vs [escape]

Iter= 8000, Average Loss= 14.387575, Average Accuracy= 1.00%

['thought', 'would', 'meet'] - [the] vs [we]

Iter= 9000, Average Loss= 14.241139, Average Accuracy= 1.30%

['said', 'that', 'but'] - [at] vs [was]

Iter= 10000, Average Loss= 14.296557, Average Accuracy= 1.00%

['council', 'to', 'consider'] - [what] vs [we]

Optimization Finished!

Elapsed time: 19.200052738189697 sec

Input 3 to 4:

Iter= 1000, Average Loss= 18.987998, Average Accuracy= 0.70%

['up', 'and', 'said', 'he'] - [had] vs [was]

Iter= 2000, Average Loss= 18.890027, Average Accuracy= 0.90%

['will', 'all', 'agree', ','] - [said] vs [outwit]

Iter= 3000, Average Loss= 18.889681, Average Accuracy= 0.90%

['consists', 'in', 'the', 'sly'] - [and] vs [council]

Iter= 4000, Average Loss= 18.557030, Average Accuracy= 1.00%

['approaches', 'us', '.', 'now'] - [,] vs [council]

Iter= 5000, Average Loss= 19.184365, Average Accuracy= 0.70%

['of', 'her', 'approach', ','] - [we] vs [outwit]

Iter= 6000, Average Loss= 18.877320, Average Accuracy= 1.00%

['by', 'a', 'ribbon', 'round'] - [the] vs [council]

Iter= 7000, Average Loss= 18.859470, Average Accuracy= 1.40%

['should', 'always', 'know', 'when'] - [she] vs [council]

Iter= 8000, Average Loss= 19.077231, Average Accuracy= 0.60%

['the', 'cat', '?', 'the'] - [mice] vs [outwit]

Iter= 9000, Average Loss= 18.765513, Average Accuracy= 1.00%

['outwit', 'their', 'common', 'enemy'] - [,] vs [council]

Iter= 10000, Average Loss= 18.794681, Average Accuracy= 1.00%

['this', ',', 'and', 'some'] - [said] vs [was]

Optimization Finished!

Elapsed time: 22.15539860725403 sec

n\_hidden = 512 to 256

Iter= 1000, Average Loss= 14.939347, Average Accuracy= 2.50%

['mouse', 'got', 'up', 'and'] - [said] vs [got]

Iter= 2000, Average Loss= 14.907475, Average Accuracy= 2.40%

['which', 'he', 'thought', 'would'] - [meet] vs [was]

Iter= 3000, Average Loss= 15.058449, Average Accuracy= 1.50%

['sly', 'and', 'treacherous', 'manner'] - [in] vs [was]

Iter= 4000, Average Loss= 15.030212, Average Accuracy= 2.30%

[',', 'we', 'could', 'easily'] - [escape] vs [was]

Iter= 5000, Average Loss= 14.826142, Average Accuracy= 2.50%

['small', 'bell', 'be', 'procured'] - [,] vs [was]

Iter= 6000, Average Loss= 15.063664, Average Accuracy= 1.80%

['of', 'the', 'cat', '.'] - [by] vs [nobody]

Iter= 7000, Average Loss= 15.153597, Average Accuracy= 1.60%

['this', 'proposal', 'met', 'with'] - [general] vs [was]

Iter= 8000, Average Loss= 15.016328, Average Accuracy= 2.10%

['up', 'and', 'said', 'that'] - [is] vs [got]

Iter= 9000, Average Loss= 14.966294, Average Accuracy= 1.70%

['very', 'well', ',', 'but'] - [who] vs [was]

Iter= 10000, Average Loss= 15.020157, Average Accuracy= 2.10%

['easy', 'to', 'propose', 'impossible'] - [remedies] vs [was]

Optimization Finished!

Elapsed time: 11.342929124832153 sec

one stack LSTM to two stack LSTM :

Iter= 1000, Average Loss= 6.369002, Average Accuracy= 0.70%

['and', 'some', 'said', 'that'] - [but] vs [is]

Iter= 2000, Average Loss= 6.242217, Average Accuracy= 0.60%

['a', 'young', 'mouse', 'got'] - [up] vs [with]

Iter= 3000, Average Loss= 6.427570, Average Accuracy= 0.80%

['you', 'will', 'all', 'agree'] - [,] vs [with]

Iter= 4000, Average Loss= 6.295753, Average Accuracy= 0.80%

['treacherous', 'manner', 'in', 'which'] - [the] vs [with]

Iter= 5000, Average Loss= 6.355538, Average Accuracy= 1.10%

['her', 'approach', ',', 'we'] - [could] vs [with]

Iter= 6000, Average Loss= 6.367167, Average Accuracy= 0.30%

['of', 'the', 'cat', '.'] - [by] vs [is]

Iter= 7000, Average Loss= 6.289504, Average Accuracy= 1.10%

['could', 'easily', 'retire', 'while'] - [she] vs [with]

Iter= 8000, Average Loss= 6.386863, Average Accuracy= 1.10%

['mouse', 'got', 'up', 'and'] - [said] vs [with]

Iter= 9000, Average Loss= 6.245974, Average Accuracy= 0.60%

['.', 'then', 'the', 'old'] - [mouse] vs [with]

Iter= 10000, Average Loss= 6.384537, Average Accuracy= 1.00%

['consider', 'what', 'measures', 'they'] - [could] vs [with]

Optimization Finished!

Elapsed time: 17.53446340560913 sec