**DATA**

The CoNLL-2003 shared task data files contain four columns separated by a single space. Each word has been put on a separate line and there is an empty line after each sentence. The first item on each line is a word, the second a part-of-speech (POS) tag, the third a syntactic chunk tag and the fourth the named entity tag. The chunk tags and the named entity tags have the format I-TYPE which means that the word is inside a phrase of type TYPE. Only if two phrases of the same type immediately follow each other, the first word of the second phrase will have tag B-TYPE to show that it starts a new phrase. A word with tag O is not part of a phrase. Here is an example:

U.N. NNP I-NP I-ORG

official NN I-NP O

Ekeus NNP I-NP I-PER

heads VBZ I-VP O

for IN I-PP O

Baghdad NNP I-NP I-LOC

. . O O

The data consists of three files : one training file and two test files testa and testb. The first test file will be used in the development phase for finding good parameters for the learning system.The second test file will be used for the final evaluation.

**Result**

The implementation achieves a test F1 score of 0.8329 with 10 epochs. Increase the number of epochs to 30 reach an F1 over 0.88.

**Result for 30 epocs**

**F1 Score for testa** :

0.26131 0.64425 0.70428 0.75219 0.74125 0.77936 0.79676 0.81087 0.79782 0.81346 0.83290 0.82388 0.82097 0.83880 0.83896 0.84304 0.84318 0.85113 0.85233 0.84989 0.86489 0.86825 0.85948 0.86395 0.86438 0.86729 0.84989 0.85032 0.86210 0.87503

**F1 Score for testb :**

0.29737 0.68803 0.72796 0.76147 0.77171 0.78682 0.81509 0.82887 0.81169 0.83549 0.84864 0.83872 0.83285 0.85849 0.86246 0.86866 0.85926 0.87176 0.88184 0.88087 0.89176 0.88993 0.88058 0.89545 0.88666 0.89383 0.88173 0.88636 0.90045 0.90657