CSE291 Projec2

Shanchuan You

Abstract

In this report, I will discuss the Conditional Random Field applied on the Bi-LSTM RNN. That is, every time we obtain features from LSTM, instead of directly computing the cross entropy between predicted sequence and the ground true tag, we introduce a higher level idea that it no longer treats each word in the sequence independently, it also considers each word's role in the whole sequence.

1 Model Description

In general, the model will be divided into three main parts. The **first** part will be the initialization of everything and adding sentence into the LSTM model to compute and extract features. The **second** part is to compute the negative loss likelihood, which we obtain the score of a provided tag sequence and then minus the partition function, which is simply the proportionality constant (conditioned on input words) to ensure that the total probability over all tags Y correctly adds up to 1. The main purpose of this model is to **minimize** the negative log likelihood. The **last** part is to perform the viterbi algorithm to select the highest probability of final predicted tags.

2 Initialization and LSTM

In general, the initialization is basically the same as the, except that we no longer input the length of the tags, but a dictionary $tagtoix$ instead, where $tagtoix$ = "O": 0, "I-PER": 1, "I-ORG": 2, 'I-LOC': 3, 'I-MISC': 4, 'B-MISC': 5,'B-ORG': 6,'B-LOC': 7, START TAG: 8, STOP TAG: 9. The dictionary is used to build our transition matrix where each entry of the transition matrix, namely, Tij, the ith row and the jth coloumn of the matrix, represents the weight from jth tag to ith tag.

Initially, each entry of the transition matrix will be set as number between -1 and 1 selected by normal distribution, as the initial value does not matter since computing negative loss likelihood force the matrix to be learnt, and the values will be adjusted later.

3 Partition function and the Score of the sentence

In this part, We calculate the partition function, forward score, and the score of the sentence with tags provided, which is the gold score. Then we perform forward score - gold score. Since both scores are calculated by the log. The subtraction is used to calculate the negative log likelihood, namely, NLL.

3.1 Score of the sentence

This part computes the score of an input sentence when the tag sequence is given. Note that we are using log with plus/add operation, this is the same that log(a*b) = loga + logb.

The core part of this function is that the input sentence is actually a 2D array, namely, the emission feature computed by the LSTM model, which has dimension [n, 8], where n is the sentence length, and 8 is the probabilities of tag with respect to the current words. In other word, for each word in the sentence, we get a probability list of 8 potential tags.

Then, the computation will would be: for each word in the sentence, score = score + emission of the given tag on this word + weight from previous tag to the current tag.

Finally, we need to add the score of a special transition, that is, from the last tag to the $< Stop >$ tag.

3.2 Partition function

In general, this function is pretty much the same as computing score of a sentence. However, this function compute the complete all possible scores. In other words, it compute the sums of all possible scores, where for each word in the sentence, the function considers all its potential tags, and finally return a ”Total” score which adds up all potential tags score together.

Initially, we calculate transition from the $START$ tag to every potential tag of the first word in the sentence. Then, for each potential tag of the current word, we add up all the previous potential tags, and the emission score of current word with ”this” tag. When the scores of all the tags of current word are computed, we move forward with the next word in the sentence. Finally, at the last word, the function computes the transition score from each of the previous potential tags to the $STOP$ tag. Adding them together, and it returns the final overall scores. Now that we obtain the score of the sentence and the partition score, we can easily compute the negative log likelihood and we use it RNN to minimize the NLL so that our trained model approach to the training set.

4 Viterbi Algorithm

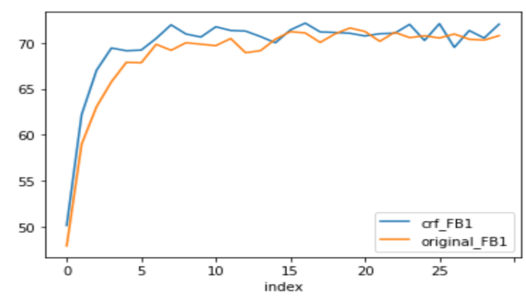
Compared with the greedy algorithm, viterbi algorithms considers more that it no longer consider the possibility of a tag on a specific word. Instead, it considers the overall possibility. For a very simple example, suppose we have a sequence of word that needs to be tagged, with only 2 tags and the word sequence length is 3, [I, love, CS]. Let say I has possibility of tag A is 0.4, and tag B is 0.6. Love has A 0.3, B 0.7. And CS A 0.9, B 0.1. Then for a greedy view, the result should be BBA. However, for viterbi algorithm, not only tag possibility will be considered, but also its previous tag should be considered. That is, given tag sequence, [B,B], we want to compute the possibility of next tag is A. Combine with A's emission, viterbi, in general, at each word, considers its previous tag and a potential current tag possibility, then it times current word emission when choosing this potential tag. Thus we have the following formula, at each word, we have **transition[from prev, to curr] * emission(word | curr)**. Finally, when we compute the whole sentence's possibility, we choose the overall largest possibility as our final decision.

The viterbi algorithm uses a dynamic program as it only record the largest possibility in the previous steps. For example, at time t with many predictions of tag O, such as IIO OOO IOO, etc. That is, II has an arrow points to O, OO has an arrow points to O, and IO also has an arrow points to O. The probabilities are 0.3 0.4 0.5, respectively. Then starting from current state O, no matter what tag we choose for it, the pre probability is already defined. That is, IIOI and OOOI and IOOI is predifined as 0.3 * porb of choosing I, 0.4 * prob of choosing I, and 0.5 * prob of choosing I. Since the post prob of these 3 candidates are the same, thus we can assert that 0.5 * prob of choosing I is the highest prob. As a result, there is no need to expand IIO and OOO. As a result, only IOO needs to be expanded at this time. As a result, in our hw case, each time, the previous step will provide 8 possible choices, then each choice expands 8 possibilities. Now that at time t, we have 8 possibilities to choose tag1, 8 possibilities to choose tag2 ... 8 possibilities to choose tag8, we keep the best choice of each one, and thus 8 choices remain. One for each, and we move them into the next step. Thus, the complexity would be the length of the sequence * the number of choices at each time $->$ 64*n, which becomes linear time in our case. This is what we called, dynamic programming, as it contains the best choices at each step and the other choices are killed, which saves lots of computations.

The Viterbi Algorithm is used to compute the highest score of the give sentence and return the argmax tags, respectively.

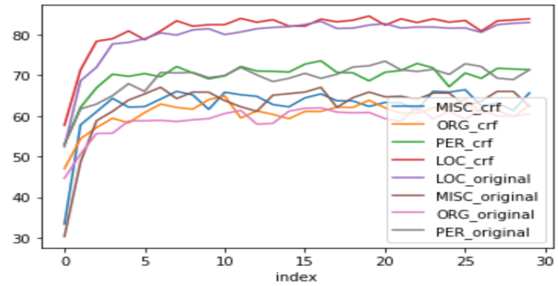
5 Results

The graph below is the comparison between LSTM with crf and the starter code.



Where the Blue line is the performance of crf. In fact, it does have some performance boost, from average around 70 to around 72.

Below is all the other comparisons among crf and the orignal tags.



We can find that tags with crf, in general, performs a little bit better than the original method.

6 Conclusion

In conclusion, crf's overall performance and each tag's performance are all slightly better than than the original method. However, without vectorization of crf, the it cost me around 2 hrs to perform the whole 30 epoches. Even with vectorization, it cost me around 30 mins. No matter which one, the original method only takes 2-3 mins. So, we are sacrificing too much time to exchange for a little bit performance increasing. Maybe we need to find other ways to increase the performance.

CSE 291 Assignment 2 BiLSTM CRF

Download Data/Eval Script

In [1]:
```
!wget https://raw.githubusercontent.com/sighsmile/conlleval/master/conlleval.py
!wget https://raw.githubusercontent.com/tberg12/cse291spr21/main/assignment2/train.data.qu
!wget https://raw.githubusercontent.com/tberg12/cse291spr21/main/assignment2/dev.data.quad
```

```
'wget' is not recognized as an internal or external command,
operable program or batch file.
'wget' is not recognized as an internal or external command,
operable program or batch file.
'wget' is not recognized as an internal or external command,
operable program or batch file.
```

In [1]:
```
import conlleval
from tqdm import tqdm
import numpy as np
from collections import defaultdict, Counter
import torch
import torch.autograd as autograd
import torch.nn as nn
import torch.optim as optim
from torchtext.vocab import Vocab
from torch.nn.utils.rnn import pack_padded_sequence, pad_packed_sequence

torch.manual_seed(291)

device = torch.device('cuda' if torch.cuda.is_available() else 'cpu')
print(device)
```

```
cuda
```

Data Preparation

In [2]:
```
TRAIN_DATA = 'train.data.quad'
VALID_DATA = 'dev.data.quad'
UNK = '<unk>'
PAD = '<pad>'
START_TAG = "<start>"  # you can add this explicitly or use it implicitly in your CRF laye
STOP_TAG = "<stop>"    # you can add this explicitly or use it implicitly in your CRF laye


def read_conll_sentence(path):
    """ Read a CONLL-format sentence into vocab objects
    Args:
        :param path: path to CONLL-format data file
        :param word_vocab: Vocabulary object for source
        :param label_vocab: Vocabulary object for target
    """
    sent = [[], []]
    with open(path) as f:
        for line in f:
            line = line.strip().split()
            if line:
                # replace numbers with 0000
                word = line[0]
                word = '0000' if word.isnumeric() else word
```

```
                sent[0].append(word)
                sent[1].append(line[3])
            else:
                yield sent[0], sent[1]
                sent = [[], []]


def prepare_dataset(dataset, word_vocab, label_vocab):
    dataset = [
      [
        torch.tensor([word_vocab.stoi[word] for word in sent[0]], dtype=torch.long),
        torch.tensor([label_vocab.stoi[label] for label in sent[1]], dtype=torch.long),
      ]
      for sent in dataset
    ]
    return dataset


# load a list of sentences, where each word in the list is a tuple containing the word anc
train_data = list(read_conll_sentence(TRAIN_DATA))
train_word_counter = Counter([word for sent in train_data for word in sent[0]])
train_label_counter = Counter([label for sent in train_data for label in sent[1]])
word_vocab = Vocab(train_word_counter, specials=(UNK, PAD), min_freq=2)
label_vocab = Vocab(train_label_counter, specials=(), min_freq=1)
train_data = prepare_dataset(train_data, word_vocab, label_vocab)
print('Train word vocab:', len(word_vocab), 'symbols.')
print('Train label vocab:', len(label_vocab), f'symbols: {list(label_vocab.stoi.keys())}')
valid_data = list(read_conll_sentence(VALID_DATA))
valid_data = prepare_dataset(valid_data, word_vocab, label_vocab)
print('Train data:', len(train_data), 'sentences.')
print('Valid data:', len(valid_data))

print(' '.join([word_vocab.itos[i.item()] for i in train_data[0][0]]))
print(' '.join([label_vocab.itos[i.item()] for i in train_data[0][1]]))

print(' '.join([word_vocab.itos[i.item()] for i in valid_data[1][0]]))
print(' '.join([label_vocab.itos[i.item()] for i in valid_data[1][1]]))
tag_to_ix = {"O": 0, "I-PER": 1, "I-ORG": 2, 'I-LOC': 3, 'I-MISC': 4, 'B-MISC': 5,'B-ORG':
print(len(tag_to_ix))
```

```
Train word vocab: 3947 symbols.
Train label vocab: 8 symbols: ['O', 'I-PER', 'I-ORG', 'I-LOC', 'I-MISC', 'B-MISC', 'B-OR
G', 'B-LOC']
Train data: 3420 sentences.
Valid data: 800
Pusan 0000 0000 0000 0000 0000 0000
I-ORG O O O O O O
Earlier this month , <unk> denied a Kabul government statement that the two sides had agre
ed to a ceasefire in the north .
O O O O I-PER O O I-LOC O O O O O O O O O O O O O O
10
```

BiLSTMTagger

In [3]:

```
# Starter code implementing a BiLSTM Tagger
# which makes locally normalized, independent
# tag classifications at each time step

class BiLSTMTagger(nn.Module):
    def __init__(self, vocab_size, tag_vocab_size, embedding_dim, hidden_dim, dropout=0.3)
        super(BiLSTMTagger, self).__init__()

        self.embedding_dim = embedding_dim
        self.hidden_dim = hidden_dim
```

```
        self.vocab_size = vocab_size
        self.tagset_size = tag_vocab_size
        self.word_embeds = nn.Embedding(vocab_size, embedding_dim).to(device)
        self.bilstm = nn.LSTM(embedding_dim, hidden_dim // 2,
                              num_layers=1, bidirectional=True, batch_first=True).to(device)
        # logistic regression
        print("hidden_dim: ", hidden_dim)
        print("target_size: ", tag_vocab_size)
        self.tag_projection_layer = nn.Linear(hidden_dim, self.tagset_size).to(device)
        self.dropout = nn.Dropout(p=dropout)

    def init_hidden(self):
        return (torch.randn(2, 1, self.hidden_dim // 2).to(device),
                torch.randn(2, 1, self.hidden_dim // 2).to(device))

    def compute_lstm_emission_features(self, sentence):
        hidden = self.init_hidden()
        embeds = self.word_embeds(sentence)
        #print(embeds.shape)
        #print(embeds)
        embeds = self.dropout(embeds)

        bilstm_out, hidden = self.bilstm(embeds, hidden)
        #print('before dropout: ', bilstm_out.shape)
        bilstm_out = self.dropout(bilstm_out)
        #print('after dropout: ', bilstm_out.shape)
        #print(bilstm_out)
        bilstm_out = bilstm_out
        bilstm_feats = self.tag_projection_layer(bilstm_out)
        return bilstm_feats

    def forward(self, sentence):
        #input a sequence of tag, perform lstm and then logistic regression
        bilstm_feats = self.compute_lstm_emission_features(sentence)
        #print('forward: ',bilstm_feats.shape)
        #print('second: ', bilstm_feats.argmax(-1))
        return bilstm_feats.max(-1)[0].sum(), bilstm_feats.argmax(-1)

    # compute the NLL -> negative likelyhood loss
    def loss(self, sentence, tags):
        #print('sentence length: ', sentence.shape)
        #print('tags length: ', tags.shape)
        bilstm_feats = self.compute_lstm_emission_features(sentence)
        #print('loss: ', bilstm_feats.shape )
        # transform predictions to (n_examples, n_classes) and ground truth to (n_examples
        return torch.nn.functional.cross_entropy(
               bilstm_feats.view(-1, self.tagset_size),
               tags.view(-1),
               reduction='sum'
           )
```

Train / Eval loop

In [4]:

```
def argmax(vec):
    # return the argmax as a python int
    _, idx = torch.max(vec, 1)
    return idx.item()


def prepare_sequence(seq, to_ix):
    idxs = [to_ix[w] for w in seq]
    return torch.tensor(idxs, dtype=torch.long)
```

```python
# Compute log sum exp in a numerically stable way for the forward algorithm
def log_sum_exp(vec):
    max_score = vec[0, argmax(vec)]
    max_score_broadcast = max_score.view(1, -1).expand(1, vec.size()[1])
    return max_score + \
        torch.log(torch.sum(torch.exp(vec - max_score_broadcast)))
```

In [7]:

```python
class BiLSTMTagger3(nn.Module):
    def __init__(self, vocab_size, tag_to_ix, embedding_dim, hidden_dim, dropout=0.3):
        super(BiLSTMTagger3, self).__init__()

        self.embedding_dim = embedding_dim
        self.hidden_dim = hidden_dim
        self.vocab_size = vocab_size
        self.tag_to_ix = tag_to_ix
        self.tagset_size = len(tag_to_ix)

        self.word_embeds = nn.Embedding(vocab_size, embedding_dim).to(device)
        self.lstm = nn.LSTM(embedding_dim, hidden_dim // 2,
                            num_layers=1, bidirectional=True, batch_first=True).to(device)

        self.hidden2tag = nn.Linear(hidden_dim, self.tagset_size).to(device)
        self.dropout = nn.Dropout(p=dropout)

        self.transitions = nn.Parameter(
            torch.randn(self.tagset_size, self.tagset_size))

        self.transitions.data[tag_to_ix[START_TAG], :] = -10000
        self.transitions.data[:, tag_to_ix[STOP_TAG]] = -10000

        self.hidden = self.init_hidden()

    def init_hidden(self):
        return (torch.randn(2, 1, self.hidden_dim // 2).to(device),
                torch.randn(2, 1, self.hidden_dim // 2).to(device))

    def _forward_alg(self, feats):
        # Do the forward algorithm to compute the partition function
        init_alphas = torch.full((1, self.tagset_size), -10000.).to(device)
        # START_TAG has all of the score.
        init_alphas[0][self.tag_to_ix[START_TAG]] = 0.

        # Wrap in a variable so that we will get automatic backprop
        forward_var = init_alphas
        #print(feats.shape)
        # Iterate through the sentence
        for feat in feats:
            alphas_t = []  # The forward tensors at this timestep
            for next_tag in range(self.tagset_size):
                # broadcast the emission score: it is the same regardless of
                # the previous tag
                #print(next_tag)
                #print('fests shape', feats.shape)
                #print(feats)

                emit_score = feat[next_tag].view(
                    1, -1).expand(1, self.tagset_size)
                # the ith entry of trans_score is the score of transitioning to
                # next_tag from i
                trans_score = self.transitions[next_tag].view(1, -1)
                # The ith entry of next_tag_var is the value for the
                # edge (i -> next_tag) before we do log-sum-exp
                next_tag_var = forward_var + trans_score + emit_score
```

```
                # The forward variable for this tag is log-sum-exp of all the
                # scores.
                alphas_t.append(log_sum_exp(next_tag_var).view(1))
            forward_var = torch.cat(alphas_t).view(1, -1)
        terminal_var = forward_var + self.transitions[self.tag_to_ix[STOP_TAG]]
        alpha = log_sum_exp(terminal_var)
        return alpha

    def _get_lstm_features(self, sentence):
        self.hidden = self.init_hidden()
        #embeds = self.word_embeds(sentence).view(len(sentence), 1, -1)
        embeds = self.word_embeds(sentence)
        embeds = self.dropout(embeds)
        #print(embeds.shape)
        #print(embeds)
        lstm_out, self.hidden = self.lstm(embeds, self.hidden)
        #print('lstm_out', lstm_out.shape)
        lstm_out = lstm_out.squeeze(0)
        #print('lstm_out', lstm_out.shape)
        #lstm_out = lstm_out.view(len(sentence), self.hidden_dim)
        lstm_feats = self.hidden2tag(lstm_out)
        #print(lstm_feats.shape)

        return lstm_feats

    def _score_sentence(self, feats, tags):
        # Gives the score of a provided tag sequence
        score = torch.zeros(1).to(device)
        #print(tags.shape)
        temp_tags = tags.squeeze(0)
        #print(temp_tags.shape)
        temp_tags = torch.cat([torch.tensor([self.tag_to_ix[START_TAG]],
                                            dtype=torch.long).to(device), temp_tags])
        #print(tags.shape)
        for i, feat in enumerate(feats):
            #print(i)
            score = score + \
                self.transitions[temp_tags[i + 1],
                                 temp_tags[i]] + feat[temp_tags[i + 1]]
        score = score + self.transitions[self.tag_to_ix[STOP_TAG],
                                         temp_tags[-1]]
        return score

    def _viterbi_decode(self, feats):
        backpointers = []

        # Initialize the viterbi variables in log space
        init_vvars = torch.full((1,
                                 self.tagset_size), -10000.).to(device)
        init_vvars[0][self.tag_to_ix[START_TAG]] = 0

        # forward_var at step i holds the viterbi variables for step i-1
        forward_var = init_vvars
        for feat in feats:
            bptrs_t = []  # holds the backpointers for this step
            viterbivars_t = []  # holds the viterbi variables for this step

            for next_tag in range(self.tagset_size):
                # next_tag_var[i] holds the viterbi variable for tag i at the
                # previous step, plus the score of transitioning
                # from tag i to next_tag.
                # We don't include the emission scores here because the max
                # does not depend on them (we add them in below)
                next_tag_var = forward_var + self.transitions[next_tag]
                best_tag_id = argmax(next_tag_var)
                bptrs_t.append(best_tag_id)
```

```python
                viterbivars_t.append(next_tag_var[0][best_tag_id].view(1))
            # Now add in the emission scores, and assign forward_var to the set
            # of viterbi variables we just computed
            forward_var = (torch.cat(viterbivars_t).to(device) + feat).view(1, -1)
            backpointers.append(bptrs_t)

        # Transition to STOP_TAG
        terminal_var = forward_var + self.transitions[self.tag_to_ix[STOP_TAG]]
        best_tag_id = argmax(terminal_var)
        path_score = terminal_var[0][best_tag_id]

        # Follow the back pointers to decode the best path.
        best_path = [best_tag_id]
        for bptrs_t in reversed(backpointers):
            best_tag_id = bptrs_t[best_tag_id]
            best_path.append(best_tag_id)
        # Pop off the start tag (we dont want to return that to the caller)
        start = best_path.pop()
        assert start == self.tag_to_ix[START_TAG]  # Sanity check
        best_path.reverse()
        return path_score, best_path

    def neg_log_likelihood(self, sentence, tags):
        feats = self._get_lstm_features(sentence)
        forward_score = self._forward_alg(feats)
        gold_score = self._score_sentence(feats, tags)
        return forward_score - gold_score

    def forward(self, sentence):  # dont confuse this with _forward_alg above.
        # Get the emission scores from the BiLSTM
        lstm_feats = self._get_lstm_features(sentence)

        # Find the best path, given the features.
        score, tag_seq = self._viterbi_decode(lstm_feats)
        return score, tag_seq
```

In [12]:

```python
def train(model, train_data, valid_data, word_vocab, label_vocab, epochs, log_interval=25)
    losses_per_epoch = []
    for epoch in range(epochs):
        print(f'--- EPOCH {epoch} ---')
        model.train()
        losses_per_epoch.append([])
        for i, (sent, tags) in enumerate(train_data):
            model.zero_grad()
            sent, tags = sent.to(device), tags.to(device)
            # that is, for each sentence and its corresponding tag sequences
            sent = sent.unsqueeze(0)
            tags = tags.unsqueeze(0)
            #print(sent)
            loss = model.neg_log_likelihood(sent, tags)
            #loss = model.loss(sent, tags)
            loss.backward()
            optimizer.step()
            #print(i)
            losses_per_epoch[-1].append(loss.detach().cpu().item())
            if i > 0 and i % log_interval == 0:
                print(f'Avg loss over last {log_interval} updates: {np.mean(losses_per_epo

        evaluate(model, valid_data, word_vocab, label_vocab)


def evaluate(model, dataset, word_vocab, label_vocab):
    model.eval()
    losses = []
```

```
    scores = []
    true_tags = []
    pred_tags = []
    sents = []
    for i, (sent, tags) in enumerate(dataset):
        with torch.no_grad():
            sent, tags = sent.to(device), tags.to(device)
            sent = sent.unsqueeze(0)
            tags = tags.unsqueeze(0)
            losses.append(model.neg_log_likelihood(sent, tags).cpu().detach().item())
            #losses.append(model.loss(sent, tags).cpu().detach().item())
            score, pred_tag_seq = model(sent)
            scores.append(score.cpu().detach().numpy())
            temp = []
            temp.append(pred_tag_seq)
            pred_tag_seq = temp
            true_tags.append([label_vocab.itos[i] for i in tags.tolist()[0]])
            pred_tags.append([label_vocab.itos[i] for i in pred_tag_seq[0]])
            sents.append([word_vocab.itos[i] for i in sent[0]])

    print('Avg evaluation loss:', np.mean(losses))
    print(conlleval.evaluate([tag for tags in true_tags for tag in tags],
                             [tag for tags in pred_tags for tag in tags], verbose=True))
    print('\n5 random evaluation samples:')
    for i in np.random.randint(0, len(sents), size=5):
        print('SENT:', ' '.join(sents[i]))
        print('TRUE:', ' '.join(true_tags[i]))
        print('PRED:', ' '.join(pred_tags[i]))
    return sents, true_tags, pred_tags
```

Training

In [ ]:

In [13]:

```
%%time
# Train BiLSTM Tagger Baseline
model = BiLSTMTagger3(len(word_vocab), tag_to_ix, 128, 256).to(device)
#model = BiLSTMTagger(len(word_vocab), 8, 128, 256).to(device)
optimizer = optim.Adam(model.parameters(), lr=1e-3)
train(model, train_data, valid_data, word_vocab, label_vocab, epochs=30, log_interval=500)
```

```
--- EPOCH 0 ---
Avg loss over last 500 updates: 9.324693586349488
Avg loss over last 500 updates: 7.752459403514862
Avg loss over last 500 updates: 6.06698727941513
Avg loss over last 500 updates: 5.391325209379196
Avg loss over last 500 updates: 4.470411400556564
Avg loss over last 500 updates: 4.540101306200027
Avg evaluation loss: 4.3108238990604875
processed 11170 tokens with 1231 phrases; found: 792 phrases; correct: 507.
accuracy:  49.17%; (non-O)
accuracy:  89.46%; precision:  64.02%; recall:  41.19%; FB1:  50.12
              LOC: precision:  82.56%; recall:  44.35%; FB1:  57.71  195
             MISC: precision:  54.76%; recall:  23.96%; FB1:  33.33  84
              ORG: precision:  59.80%; recall:  38.76%; FB1:  47.04  199
              PER: precision:  57.64%; recall:  49.05%; FB1:  53.00  314
(64.01515151515152, 41.18602761982129, 50.12357884330203)

5 random evaluation samples:
SENT: There is nothing left for us but to be <unk> to <unk> for <unk> <unk> , " <unk> <unk
> said .
TRUE: O O O O O O O O O O O O O O O O O O I-ORG I-ORG O O
```

```
PRED: O O O O O O O O O O O O O O O O O O I-PER I-PER O O
SENT: Results of South Korean
TRUE: O O I-MISC I-MISC
PRED: O O I-MISC O
SENT: Extras ( <unk> <unk> <unk> ) 0000
TRUE: O O O O O O O
PRED: O O I-PER I-PER I-PER O O
SENT: 0000 - <unk> <unk>
TRUE: O O O O
PRED: O O O O
SENT: <unk> <unk> --- --- --- --- <unk> up <unk>
TRUE: I-LOC B-LOC O O O O O O O
PRED: O O O O O O O O O
--- EPOCH 1 ---
Avg loss over last 500 updates: 3.8126311416625978
Avg loss over last 500 updates: 4.14901714015007
Avg loss over last 500 updates: 3.3279778904914856
Avg loss over last 500 updates: 3.3117338461875914
Avg loss over last 500 updates: 2.650013731479645
Avg loss over last 500 updates: 2.8121376523971557
Avg evaluation loss: 3.237040534466505
processed 11170 tokens with 1231 phrases; found: 959 phrases; correct: 681.
accuracy:  62.22%; (non-O)
accuracy:  91.67%; precision:  71.01%; recall:  55.32%; FB1:  62.19
              LOC: precision:  84.53%; recall:  61.71%; FB1:  71.34  265
             MISC: precision:  79.82%; recall:  45.31%; FB1:  57.81  109
              ORG: precision:  60.32%; recall:  49.51%; FB1:  54.38  252
              PER: precision:  65.47%; recall:  59.08%; FB1:  62.11  333
(71.01147028154328, 55.320877335499596, 62.19178082191782)

5 random evaluation samples:
SENT: When we <unk> the shares we <unk> ...
TRUE: O O O O O O O O
PRED: O O O O O O O O
SENT: <unk> Bremen 0000 <unk> <unk> 0000
TRUE: I-ORG I-ORG O I-ORG I-ORG O
PRED: I-ORG I-ORG O I-ORG I-ORG O
SENT: Reuters has not verified these stories and does not vouch for their accuracy .
TRUE: I-ORG O O O O O O O O O O O O O
PRED: I-ORG O O O O O O O O O O O O O
SENT: SOCCER - <unk> <unk> IN <unk> FOR <unk> <unk> .
TRUE: O O I-LOC O O O O O O O
PRED: O O I-PER I-PER O O O O O O
SENT: It was the second arms <unk> this week .
TRUE: O O O O O O O O O
PRED: O O O O O O O O O
--- EPOCH 2 ---
Avg loss over last 500 updates: 2.5364945862293244
Avg loss over last 500 updates: 2.701589867115021
Avg loss over last 500 updates: 2.272822791337967
Avg loss over last 500 updates: 2.232509036540985
Avg loss over last 500 updates: 1.885722356081009
Avg loss over last 500 updates: 2.038311313152313
Avg evaluation loss: 2.9671986613422634
processed 11170 tokens with 1231 phrases; found: 1019 phrases; correct: 754.
accuracy:  66.57%; (non-O)
accuracy:  92.57%; precision:  73.99%; recall:  61.25%; FB1:  67.02
              LOC: precision:  87.46%; recall:  71.07%; FB1:  78.42  295
             MISC: precision:  81.03%; recall:  48.96%; FB1:  61.04  116
              ORG: precision:  62.31%; recall:  52.77%; FB1:  57.14  260
              PER: precision:  68.97%; recall:  65.04%; FB1:  66.95  348
(73.99411187438666, 61.25101543460602, 67.02222222222223)

5 random evaluation samples:
SENT: He would continue working on various <unk> and might meet " one state <unk> or anoth
er " .
```

```
TRUE: O O O O O O O O O O O O O O O O O O O O O O O O
PRED: O O O O O O O O O O O O O O O O O O O O O O O O
SENT: Johnson said that <unk> interest rates have already been <unk> by the election .
TRUE: I-PER O O O O O O O O O O O O O
PRED: I-PER O O O O O O O O O O O O O
SENT: for , against , points ) :
TRUE: O O O O O O O
PRED: O O O O O O O
SENT: " When I saw the new draw I did n't have to change my <unk> , " <unk> said . "
TRUE: O O O O O O O O O O O O O O O O O I-PER O O O
PRED: O O O O O O O O O O O O O O O O O I-PER O O O
SENT: 9. <unk> <unk> ( China ) <unk>
TRUE: O I-PER I-PER O I-LOC O O
PRED: O I-PER I-PER O I-LOC O O
--- EPOCH 3 ---
Avg loss over last 500 updates: 1.919104497909546
Avg loss over last 500 updates: 2.109152201652527
Avg loss over last 500 updates: 1.7381675925254823
Avg loss over last 500 updates: 1.759016581773758
Avg loss over last 500 updates: 1.5086341819763183
Avg loss over last 500 updates: 1.660850250005722
Avg evaluation loss: 2.836272454857826
processed 11170 tokens with 1231 phrases; found: 998 phrases; correct: 774.
accuracy:  68.34%; (non-O)
accuracy:  93.54%; precision:  77.56%; recall:  62.88%; FB1:  69.45
              LOC: precision:  87.33%; recall:  72.18%; FB1:  79.03  300
             MISC: precision:  82.79%; recall:  52.60%; FB1:  64.33  122
              ORG: precision:  69.74%; recall:  51.79%; FB1:  59.44  228
              PER: precision:  72.41%; recall:  68.29%; FB1:  70.29  348
(77.55511022044088, 62.875710804224205, 69.44818304172274)

5 random evaluation samples:
SENT: <unk> <unk> ( 11th ) , <unk> <unk> ( <unk> and 51st ) , <unk> <unk>
TRUE: I-PER I-PER O O O O I-PER I-PER O O O O O O I-PER I-PER
PRED: I-PER I-PER O O O O I-PER I-PER O O O O O O I-PER I-PER
SENT: PRESS DIGEST - <unk> - AUG 0000 .
TRUE: O O O I-LOC O O O O
PRED: O O O O O O O O
SENT: 0000 GMT
TRUE: O I-MISC
PRED: O I-MISC
SENT: <unk> <unk> <unk> ( <unk> ) <unk> <unk> <unk> US / UK / <unk>
TRUE: I-ORG O O O O O O O O O I-LOC O I-LOC O I-LOC
PRED: O I-PER I-PER O O O O O O I-ORG I-ORG I-ORG O O
SENT: According to a new release from the governor , Wisconsin submitted a plan to the U.
S. Department of <unk> and Human <unk> for administration of the new block <unk> system fo
r welfare just minutes after President Bill Clinton signed the <unk> into law Thursday .
TRUE: O O O O O O O O O I-LOC O O O O O I-ORG I-ORG I-ORG I-ORG I-ORG I-ORG I-ORG O O O O
O O O O O O O O O O I-PER I-PER O O O O O O O
PRED: O O O O O O O O O O O O O O O I-LOC I-LOC I-LOC O O O O O O O O O O O O O O O O O O
I-PER I-PER O O O O O O O
--- EPOCH 4 ---
Avg loss over last 500 updates: 1.5034452352523804
Avg loss over last 500 updates: 1.678650174856186
Avg loss over last 500 updates: 1.4139041152000427
Avg loss over last 500 updates: 1.360803474187851
Avg loss over last 500 updates: 1.212214565515518
Avg loss over last 500 updates: 1.384278891801834
Avg evaluation loss: 2.74378400683403
processed 11170 tokens with 1231 phrases; found: 1022 phrases; correct: 779.
accuracy:  69.33%; (non-O)
accuracy:  93.52%; precision:  76.22%; recall:  63.28%; FB1:  69.15
              LOC: precision:  88.82%; recall:  74.38%; FB1:  80.96  304
             MISC: precision:  80.83%; recall:  50.52%; FB1:  62.18  120
              ORG: precision:  63.71%; recall:  53.75%; FB1:  58.30  259
              PER: precision:  72.86%; recall:  66.94%; FB1:  69.77  339
```

```
(76.22309197651663, 63.28188464662876, 69.15224145583666)

5 random evaluation samples:
SENT: -DOCSTART-
TRUE: O
PRED: O
SENT: 4. <unk> <unk> ( Kenya ) <unk>
TRUE: O I-PER I-PER O I-LOC O O
PRED: O I-PER I-PER O I-LOC O O
SENT: The Greek <unk> party 's executive bureau gave the <unk> light to Prime Minister <un
k> <unk> to call <unk> elections , its general secretary <unk> <unk> told reporters .
TRUE: O I-MISC O O O O O O O O O O O O I-PER I-PER O O O O O O O O I-PER I-PER O O O
PRED: O I-MISC I-ORG I-ORG O O O O O O O O O O I-PER I-PER O O O O O O O O I-PER I-PER O O
O
SENT: <unk> <unk> <unk> ( Netherlands ) <unk> <unk>
TRUE: O I-PER I-PER O I-LOC O I-ORG O
PRED: I-PER I-PER I-PER O I-LOC O O O
SENT: -DOCSTART-
TRUE: O
PRED: O
--- EPOCH 5 ---
Avg loss over last 500 updates: 1.2795283374786377
Avg loss over last 500 updates: 1.375453113079071
Avg loss over last 500 updates: 1.215940806388855
Avg loss over last 500 updates: 1.3309865157604217
Avg loss over last 500 updates: 1.002288314819336
Avg loss over last 500 updates: 1.19175581741333
Avg evaluation loss: 2.8216698206961155
processed 11170 tokens with 1231 phrases; found: 1069 phrases; correct: 796.
accuracy:  71.53%; (non-O)
accuracy:  93.50%; precision:  74.46%; recall:  64.66%; FB1:  69.22
              LOC: precision:  87.54%; recall:  71.63%; FB1:  78.79  297
             MISC: precision:  77.52%; recall:  52.08%; FB1:  62.31  129
              ORG: precision:  62.85%; recall:  58.96%; FB1:  60.84  288
              PER: precision:  71.83%; recall:  69.11%; FB1:  70.44  355
(74.4621141253508, 64.66287571080423, 69.21739130434783)

5 random evaluation samples:
SENT: First Union National Bank of <unk> <unk> suit .
TRUE: I-ORG I-ORG I-ORG I-ORG I-ORG I-ORG O O O
PRED: I-ORG I-ORG I-ORG I-ORG I-ORG I-ORG I-ORG O O
SENT: He spent a year under house arrest and was tried but <unk> last year on charges of o
rdering the murder of four <unk> in 0000 .
TRUE: O O O O O O O O O O O O O O O O O O O O O O O O O O
PRED: O O O O O O O O O O O O O O O O O O O O O O O O O O
SENT: Waqar Younis <unk> <unk> b Harris 0000
TRUE: I-PER I-PER O I-PER O I-PER O
PRED: I-PER I-PER I-PER I-PER O I-PER O
SENT: The <unk> are July <unk> , and the final September <unk> 0000 .
TRUE: O O O O O O O O O O O O O
PRED: O O O O O O O O O O I-ORG O O
SENT: Rangers 0000 <unk> United 0000
TRUE: I-ORG O I-ORG I-ORG O
PRED: I-ORG O I-ORG I-ORG O
--- EPOCH 6 ---
Avg loss over last 500 updates: 1.0786786432266235
Avg loss over last 500 updates: 1.2349101711511612
Avg loss over last 500 updates: 1.1458699822425842
Avg loss over last 500 updates: 1.0566541259288789
Avg loss over last 500 updates: 0.9335016398429871
Avg loss over last 500 updates: 1.1129399638175965
Avg evaluation loss: 2.930451305359602
processed 11170 tokens with 1231 phrases; found: 1090 phrases; correct: 818.
accuracy:  71.59%; (non-O)
accuracy:  93.78%; precision:  75.05%; recall:  66.45%; FB1:  70.49
              LOC: precision:  88.82%; recall:  74.38%; FB1:  80.96  304
```

```
             MISC: precision:  79.84%; recall:  53.65%; FB1:  64.17  129
              ORG: precision:  67.92%; recall:  58.63%; FB1:  62.94  265
              PER: precision:  67.60%; recall:  71.82%; FB1:  69.65  392
(75.04587155963303, 66.45004061738425, 70.48685911245154)

5 random evaluation samples:
SENT: 5. <unk> <unk> ( Italy ) <unk> 0000
TRUE: O I-PER I-PER O I-LOC O I-ORG O
PRED: O I-PER I-PER O I-LOC O O O
SENT: Sri Lanka
TRUE: I-LOC I-LOC
PRED: I-LOC I-LOC
SENT: MILWAUKEE AT CHICAGO
TRUE: I-ORG O I-LOC
PRED: I-ORG O I-LOC
SENT: <unk> <unk> 0000 0000 0000 0000 0000 0000 0000
TRUE: I-ORG I-ORG O O O O O O O
PRED: I-ORG I-ORG O O O O O O O
SENT: <unk> <unk>
TRUE: I-PER I-PER
PRED: I-LOC O
--- EPOCH 7 ---
Avg loss over last 500 updates: 0.9669645204544067
Avg loss over last 500 updates: 1.055318928003311
Avg loss over last 500 updates: 0.8524149534702301
Avg loss over last 500 updates: 0.9359124011993408
Avg loss over last 500 updates: 0.8417094810009003
Avg loss over last 500 updates: 1.0308143017292022
Avg evaluation loss: 2.8696382036805153
processed 11170 tokens with 1231 phrases; found: 1131 phrases; correct: 850.
accuracy:  74.94%; (non-O)
accuracy:  94.31%; precision:  75.15%; recall:  69.05%; FB1:  71.97
              LOC: precision:  90.91%; recall:  77.13%; FB1:  83.46  308
             MISC: precision:  78.99%; recall:  56.77%; FB1:  66.06  138
              ORG: precision:  63.70%; recall:  60.59%; FB1:  62.10  292
              PER: precision:  69.97%; recall:  74.53%; FB1:  72.18  393
(75.15473032714411, 69.04955320877335, 71.97290431837425)

5 random evaluation samples:
SENT: The <unk> of his other <unk> <unk> as " In the Year of January " ( 0000 ) , " The <u
nk> " ( 0000 ) , " <unk> <unk> " ( 0000 ) , " The <unk> <unk> " ( 0000 ) and " A World of
<unk> " ( 0000 ) , followed by " The <unk> , " " <unk> <unk> " and , most recently , " A <
unk> . "
TRUE: O O O O O O O O O I-MISC I-MISC I-MISC I-MISC I-MISC O O O O O O I-MISC I-MISC O O O
O O O I-MISC I-MISC O O O O O O I-MISC I-MISC I-MISC O O O O O O I-MISC I-MISC I-MISC I-MI
SC O O O O O O O O O I-MISC I-MISC O O O I-MISC I-MISC O O O O O O O O I-MISC I-MISC O O
PRED: O O O O O O O O O O O O O O O O O O O O O O O O O O O I-ORG I-ORG O O O O O O O O I-
ORG I-ORG O O O O O O O O I-MISC I-MISC I-MISC O O O O O O O O O O O O O O I-PER I-PER O O O O
O O O O O O O
SENT: <unk> 1996-08-24
TRUE: I-LOC O
PRED: O O
SENT: <unk> : 0000 days <unk> <unk> :
TRUE: O O O O O O O
PRED: O O O O O O O
SENT: South African company results expected next week include the
TRUE: I-MISC I-MISC O O O O O O O
PRED: I-MISC I-MISC O O O O O O O
SENT: Italy <unk> <unk> Cuttitta
TRUE: I-LOC O I-PER I-PER
PRED: I-LOC O O I-PER
--- EPOCH 8 ---
Avg loss over last 500 updates: 0.8937310128211975
Avg loss over last 500 updates: 1.10304132270813
Avg loss over last 500 updates: 0.868907723903656
Avg loss over last 500 updates: 0.836537968158722
```

```
Avg loss over last 500 updates: 0.6583041189908981
Avg loss over last 500 updates: 0.8643292055130005
Avg evaluation loss: 3.1224144876003264
processed 11170 tokens with 1231 phrases; found: 1077 phrases; correct: 819.
accuracy:  71.81%; (non-O)
accuracy:  93.95%; precision:  76.04%; recall:  66.53%; FB1:  70.97
              LOC: precision:  88.54%; recall:  76.58%; FB1:  82.13  314
             MISC: precision:  81.89%; recall:  54.17%; FB1:  65.20  127
              ORG: precision:  64.64%; recall:  58.96%; FB1:  61.67  280
              PER: precision:  71.91%; recall:  69.38%; FB1:  70.62  356
(76.04456824512535, 66.53127538586516, 70.97053726169842)

5 random evaluation samples:
SENT: Martin <unk> , <unk> <unk> ( Sweden )
TRUE: I-PER I-PER O I-PER I-PER O I-LOC O
PRED: I-PER I-PER O I-PER I-PER O I-LOC O
SENT: The court <unk> <unk> 's <unk> that <unk> 's <unk> from Denmark , where he was arres
ted in March last year at the request of German authorities , was illegal .
TRUE: O O O I-PER O O O I-PER O O O I-LOC O O O O O O O O O O O O O I-MISC O O O O O
PRED: O O O I-PER O O O I-ORG O O O I-LOC O O O O O O O O O O O O O I-MISC O O O O O
SENT: <unk>
TRUE: O
PRED: O
SENT: -DOCSTART-
TRUE: O
PRED: O
SENT: An <unk> of <unk> has killed five people in the central Senegal town of <unk> , wher
e health authorities have <unk> 0000 cases since August 0000 , a medical official said on
Thursday .
TRUE: O O O O O O O O O O I-LOC O O I-LOC O O O O O O O O O O O O O O O O O O
PRED: O O O O O O O O O O I-LOC O O I-LOC O O O O O O O O O O O O O O O O O O
--- EPOCH 9 ---
Avg loss over last 500 updates: 0.7717373546361923
Avg loss over last 500 updates: 0.929089684009552
Avg loss over last 500 updates: 0.8032551293373108
Avg loss over last 500 updates: 0.8174836456775665
Avg loss over last 500 updates: 0.612529159784317
Avg loss over last 500 updates: 0.8102035324573517
Avg evaluation loss: 3.0575619511306287
processed 11170 tokens with 1231 phrases; found: 1133 phrases; correct: 835.
accuracy:  74.61%; (non-O)
accuracy:  93.97%; precision:  73.70%; recall:  67.83%; FB1:  70.64
              LOC: precision:  89.39%; recall:  76.58%; FB1:  82.49  311
             MISC: precision:  73.38%; recall:  53.12%; FB1:  61.63  139
              ORG: precision:  67.88%; recall:  60.59%; FB1:  64.03  274
              PER: precision:  65.77%; recall:  72.90%; FB1:  69.15  409
(73.6981465136805, 67.83103168155971, 70.64297800338409)

5 random evaluation samples:
SENT: Standings ( tabulated under played , won , drawn , lost , goals
TRUE: O O O O O O O O O O O O O
PRED: O O O O O O O O O O O O O
SENT: PARIS 1996-08-23
TRUE: I-LOC O
PRED: I-LOC O
SENT: The PKK also <unk> into Turkey from <unk> in the <unk> of northern Iraq .
TRUE: O I-ORG O O O I-LOC O O O O O O O I-LOC O
PRED: O I-ORG O O O I-LOC O O O O O O O I-LOC O
SENT: <unk> - <unk> <unk> figures <unk> to value of orders on books at end of period .
TRUE: O O O O O O O O O O O O O O O O O
PRED: O O I-PER O O O O O O O O O O O O O O
SENT: 5. Steve Brown ( U.S. ) <unk>
TRUE: O I-PER I-PER O I-LOC O O
PRED: O I-PER I-PER O I-LOC O O
--- EPOCH 10 ---
Avg loss over last 500 updates: 0.6969414043426514
```

```
Avg loss over last 500 updates: 0.8385431489944458
Avg loss over last 500 updates: 0.7882229256629943
Avg loss over last 500 updates: 0.7522817523479461
Avg loss over last 500 updates: 0.6692605974674225
Avg loss over last 500 updates: 0.6750851354598999
Avg evaluation loss: 3.185546517819166
processed 11170 tokens with 1231 phrases; found: 1149 phrases; correct: 854.
accuracy:  74.83%; (non-O)
accuracy:  94.08%; precision:  74.33%; recall:  69.37%; FB1:  71.76
              LOC: precision:  87.62%; recall:  77.96%; FB1:  82.51  323
             MISC: precision:  80.45%; recall:  55.73%; FB1:  65.85  133
              ORG: precision:  68.73%; recall:  61.56%; FB1:  64.95  275
              PER: precision:  65.79%; recall:  74.53%; FB1:  69.89  418
(74.325500435161, 69.37449228269699, 71.76470588235294)

5 random evaluation samples:
SENT: SOCCER - <unk> <unk> <unk> <unk> <unk> MANCHESTER <unk> .
TRUE: O O I-PER O O O O I-LOC I-LOC O
PRED: O O I-ORG I-ORG I-ORG I-ORG I-ORG O O O
SENT: India <unk> <unk> ' 0000 sales , output up .
TRUE: I-LOC I-ORG O O O O O O O O
PRED: I-LOC O I-PER O O O O O O O
SENT: Newmont , in fact , will not <unk> from the Santa Fe acquisition on an earnings basi
s for at least two years , which also <unk> its <unk> to <unk> its offer .
TRUE: I-ORG O O O O O O O O O I-ORG I-ORG O O O O O O O O O O O O O O O O O O O O O
PRED: I-PER O O O O O O O O O I-ORG I-ORG O O O O O O O O O O O O O O O O O O O O O
SENT: -DOCSTART-
TRUE: O
PRED: O
SENT: Fox said the British government wanted an end to the alleged <unk> of its <unk> at D
haka airport by customs officials .
TRUE: I-PER O O I-MISC O O O O O O O O O O O O I-LOC O O O O O
PRED: I-PER O O I-MISC O O O O O O O O O O O O I-LOC O O O O O
--- EPOCH 11 ---
Avg loss over last 500 updates: 0.6967372241020202
Avg loss over last 500 updates: 0.7907397847175598
Avg loss over last 500 updates: 0.6878003034591674
Avg loss over last 500 updates: 0.657532648563385
Avg loss over last 500 updates: 0.556335406780243
Avg loss over last 500 updates: 0.6988385837078095
Avg evaluation loss: 3.116869171112776
processed 11170 tokens with 1231 phrases; found: 1140 phrases; correct: 846.
accuracy:  74.78%; (non-O)
accuracy:  94.00%; precision:  74.21%; recall:  68.72%; FB1:  71.36
              LOC: precision:  89.69%; recall:  79.06%; FB1:  84.04  320
             MISC: precision:  79.70%; recall:  55.21%; FB1:  65.23  133
              ORG: precision:  59.42%; recall:  59.61%; FB1:  59.51  308
              PER: precision:  71.24%; recall:  73.17%; FB1:  72.19  379
(74.21052631578947, 68.72461413484972, 71.3622943905525)

5 random evaluation samples:
SENT: 5. <unk> <unk> ( Russia ) <unk>
TRUE: O I-PER I-PER O I-LOC O O
PRED: O I-PER I-PER O I-LOC O O
SENT: Hong Kong 's Tsang sees growth , <unk> <unk> .
TRUE: I-LOC I-LOC O I-PER O O O O O O
PRED: I-LOC I-LOC O O O O O O O O
SENT: Moslem <unk> <unk> 0000 <unk> - agency .
TRUE: I-MISC O O O I-MISC O O O
PRED: I-MISC O O O O O O O
SENT: A hijacked Sudan Airways plane with 0000 passengers and crew on board was expected t
o land at London 's Stansted airport later on Tuesday morning , a police spokeswoman said
.
TRUE: O O I-ORG I-ORG O O O O O O O O O O O O O I-LOC O I-LOC O O O O O O O O O O O
PRED: O O I-ORG I-ORG O O O O O O O O O O O O O I-LOC O I-LOC O O O O O O O O O O O
SENT: 6. Michael <unk> ( U.S. ) <unk>
```

```
TRUE: O I-PER I-PER O I-LOC O O
PRED: O I-PER I-PER O I-LOC O O
--- EPOCH 12 ---
Avg loss over last 500 updates: 0.5625160057544708
Avg loss over last 500 updates: 0.688736927986145
Avg loss over last 500 updates: 0.658545393705368
Avg loss over last 500 updates: 0.624279503583908
Avg loss over last 500 updates: 0.5952160804271698
Avg loss over last 500 updates: 0.6750965361595154
Avg evaluation loss: 3.2722255070507527
processed 11170 tokens with 1231 phrases; found: 1120 phrases; correct: 838.
accuracy:  72.85%; (non-O)
accuracy:  94.06%; precision:  74.82%; recall:  68.07%; FB1:  71.29
              LOC: precision:  88.24%; recall:  78.51%; FB1:  83.09  323
             MISC: precision:  77.54%; recall:  55.73%; FB1:  64.85  138
              ORG: precision:  65.31%; recall:  57.65%; FB1:  61.25  271
              PER: precision:  69.33%; recall:  72.90%; FB1:  71.07  388
(74.82142857142857, 68.07473598700243, 71.28881327094852)

5 random evaluation samples:
SENT: <unk> <unk>
TRUE: I-ORG I-ORG
PRED: O O
SENT: He has since named a prime minister for the first time since early in his rule and o
rdered a crackdown on <unk> .
TRUE: O O O O O O O O O O O O O O O O O O O O O O O
PRED: O O O O O O O O O O O O O O O O O O O O O O O
SENT: <unk> ( <unk> <unk> ) <unk> 0000
TRUE: O O O O O O O
PRED: I-ORG O I-ORG I-ORG O I-ORG O
SENT: U.S. debt futures finished a <unk> <unk> session sharply lower , as the markets were
<unk> by a stronger than expected rise in the August National Association of <unk> <unk> (
<unk> ) index for the Chicago area , traders and analysts said .
TRUE: I-LOC O O O O O O O O O O O O O O O O O O O O O O O O I-ORG I-ORG I-ORG I-ORG I-ORG
O I-ORG O O O O I-LOC O O O O O O O
PRED: I-LOC O O O O O O O O O O O O O O O O O O O O O O O O I-ORG I-ORG I-ORG I-ORG I-ORG
O I-ORG O O O O I-LOC O O O O O O O
SENT: The PKK also <unk> into Turkey from <unk> in the <unk> of northern Iraq .
TRUE: O I-ORG O O O I-LOC O O O O O O O I-LOC O
PRED: O I-ORG O O O I-LOC O O O O O O O I-LOC O
--- EPOCH 13 ---
Avg loss over last 500 updates: 0.5624189565181732
Avg loss over last 500 updates: 0.6365453751087189
Avg loss over last 500 updates: 0.5754836969375611
Avg loss over last 500 updates: 0.603867014169693
Avg loss over last 500 updates: 0.4931268653869629
Avg loss over last 500 updates: 0.6459541335105896
Avg evaluation loss: 3.4159583427011966
processed 11170 tokens with 1231 phrases; found: 1191 phrases; correct: 856.
accuracy:  75.77%; (non-O)
accuracy:  93.77%; precision:  71.87%; recall:  69.54%; FB1:  70.69
              LOC: precision:  90.42%; recall:  77.96%; FB1:  83.73  313
             MISC: precision:  70.32%; recall:  56.77%; FB1:  62.82  155
              ORG: precision:  60.19%; recall:  60.59%; FB1:  60.39  309
              PER: precision:  67.15%; recall:  75.34%; FB1:  71.01  414
(71.87237615449203, 69.53696181965881, 70.68538398018167)

5 random evaluation samples:
SENT: 6. Michael <unk> ( U.S. ) <unk>
TRUE: O I-PER I-PER O I-LOC O O
PRED: O I-PER I-PER O I-LOC O O
SENT: NEW YORK 1996-08-23
TRUE: I-LOC I-LOC O
PRED: I-LOC I-LOC O
SENT: -DOCSTART-
TRUE: O
```

```
PRED: O
SENT: Roe said he was <unk> by his forecast of a <unk> billion <unk> net for 0000 .
TRUE: I-PER O O O O O O O O O O O O O O O O
PRED: I-PER O O O O O O O O O O O O O O O O
SENT: <unk>
TRUE: I-LOC
PRED: O
--- EPOCH 14 ---
Avg loss over last 500 updates: 0.5657826948165894
Avg loss over last 500 updates: 0.7204091582298279
Avg loss over last 500 updates: 0.5510184633731842
Avg loss over last 500 updates: 0.4845707702636719
Avg loss over last 500 updates: 0.49096420764923093
Avg loss over last 500 updates: 0.5778690807819367
Avg evaluation loss: 3.416489884406328
processed 11170 tokens with 1231 phrases; found: 1194 phrases; correct: 849.
accuracy:  73.84%; (non-O)
accuracy:  93.62%; precision:  71.11%; recall:  68.97%; FB1:  70.02
              LOC: precision:  84.55%; recall:  79.89%; FB1:  82.15  343
             MISC: precision:  74.10%; recall:  53.65%; FB1:  62.24  139
              ORG: precision:  59.03%; recall:  59.61%; FB1:  59.32  310
              PER: precision:  67.91%; recall:  73.98%; FB1:  70.82  402
(71.10552763819096, 68.96831844029244, 70.02061855670104)

5 random evaluation samples:
SENT: <unk> Rubin leave , Wall Street would worry that he might take his <unk> policy with
him .
TRUE: O I-PER O O I-LOC I-LOC O O O O O O O O O O O O
PRED: O I-PER O O I-LOC I-LOC O O O O O O O O O O O O
SENT: <unk> has infrastructure in place for <unk> - Tsang .
TRUE: I-LOC O O O O O O O I-PER O
PRED: I-PER O O O O O O O I-PER O
SENT: 5. Wales <unk>
TRUE: O I-LOC O
PRED: O I-LOC O
SENT: <unk> <unk>
TRUE: I-ORG I-ORG
PRED: I-ORG I-ORG
SENT: <unk> interest in crude did not have enough <unk> to <unk> it much higher since many
players had left early to start the Labor <unk> holiday weekend , traders said .
TRUE: O O O O O O O O O O O O O O O O O O O O O O O I-MISC I-MISC O O O O O O
PRED: O O O O O O O O O O O O O O O O O O O O O O O I-ORG I-ORG O O O O O O
--- EPOCH 15 ---
Avg loss over last 500 updates: 0.5596058015823364
Avg loss over last 500 updates: 0.6265998647212982
Avg loss over last 500 updates: 0.5567566292285919
Avg loss over last 500 updates: 0.47660572576522825
Avg loss over last 500 updates: 0.48517633509635927
Avg loss over last 500 updates: 0.5443272039890289
Avg evaluation loss: 3.4912644217908384
processed 11170 tokens with 1231 phrases; found: 1172 phrases; correct: 858.
accuracy:  74.94%; (non-O)
accuracy:  94.10%; precision:  73.21%; recall:  69.70%; FB1:  71.41
              LOC: precision:  86.10%; recall:  78.51%; FB1:  82.13  331
             MISC: precision:  76.43%; recall:  55.73%; FB1:  64.46  140
              ORG: precision:  62.37%; recall:  59.93%; FB1:  61.13  295
              PER: precision:  69.46%; recall:  76.42%; FB1:  72.77  406
(73.20819112627987, 69.69943135662064, 71.41073657927592)

5 random evaluation samples:
SENT: 8. <unk> <unk> ( Belgium ) <unk>
TRUE: O I-PER I-PER O I-LOC O O
PRED: O I-PER I-PER O I-LOC O O
SENT: <unk> has <unk> political leaders to a meeting on Thursday to discuss the final form
of the <unk> which both <unk> and <unk> <unk> <unk> for different <unk> .
TRUE: I-PER O O O O O O O O O O O O O O O O O O O O O I-MISC O I-MISC O O O O O O
```

```
PRED: I-PER O O O O O O O O O O O O O O O O O O O O O O O O O O O O
SENT: BEIJING 1996-12-06
TRUE: I-LOC O
PRED: I-LOC O
SENT: <unk>
TRUE: O
PRED: O
SENT: Tottenham 0000 0000 0000 0000 0000 0000 0000
TRUE: I-ORG O O O O O O O
PRED: I-ORG O O O O O O O
--- EPOCH 16 ---
Avg loss over last 500 updates: 0.5316154316663743
Avg loss over last 500 updates: 0.5638617334365845
Avg loss over last 500 updates: 0.6231105732917785
Avg loss over last 500 updates: 0.4700705800056458
Avg loss over last 500 updates: 0.37194175958633424
Avg loss over last 500 updates: 0.5451498885154724
Avg evaluation loss: 3.425596902668476
processed 11170 tokens with 1231 phrases; found: 1161 phrases; correct: 863.
accuracy:  75.22%; (non-O)
accuracy:  94.22%; precision:  74.33%; recall:  70.11%; FB1:  72.16
              LOC: precision:  89.66%; recall:  78.79%; FB1:  83.87  319
             MISC: precision:  79.26%; recall:  55.73%; FB1:  65.44  135
              ORG: precision:  60.06%; recall:  62.21%; FB1:  61.12  318
              PER: precision:  71.72%; recall:  75.61%; FB1:  73.61  389
(74.33247200689061, 70.10560519902518, 72.1571906354515)

5 random evaluation samples:
SENT: SOCCER - PSV BEAT <unk> <unk> TO <unk> <unk> FROM <unk> .
TRUE: O O I-ORG O I-ORG O O O O O I-ORG O
PRED: O O I-ORG O I-ORG I-ORG O I-ORG O O O O
SENT: " The plant is <unk> as <unk> , " Jose <unk> <unk> , director of <unk> , told Spanis
h state television .
TRUE: O O O O O O O O O I-PER I-PER I-PER O O O O O O I-MISC O O O
PRED: O O O O O O O O O I-PER I-PER I-PER O O O I-ORG O O I-MISC O O O
SENT: -DOCSTART-
TRUE: O
PRED: O
SENT: <unk> sources said feed <unk> demand was keeping <unk> with <unk> production and dri
ving prices higher .
TRUE: I-LOC O O O O O O O O O O O O O O O O O O
PRED: O O O O O O O O O O O O O O O O O O O O
SENT: CRICKET - ENGLAND <unk> <unk> FOR ONE-DAY <unk> .
TRUE: O O I-LOC O O O O O O
PRED: O O I-LOC O I-ORG O O O O
--- EPOCH 17 ---
Avg loss over last 500 updates: 0.4674039533138275
Avg loss over last 500 updates: 0.5557696592807769
Avg loss over last 500 updates: 0.48278678131103514
Avg loss over last 500 updates: 0.48371106362342836
Avg loss over last 500 updates: 0.4126667063236237
Avg loss over last 500 updates: 0.5237756986618042
Avg evaluation loss: 3.558134785890579
processed 11170 tokens with 1231 phrases; found: 1193 phrases; correct: 863.
accuracy:  74.67%; (non-O)
accuracy:  93.91%; precision:  72.34%; recall:  70.11%; FB1:  71.20
              LOC: precision:  86.14%; recall:  80.44%; FB1:  83.19  339
             MISC: precision:  72.67%; recall:  56.77%; FB1:  63.74  150
              ORG: precision:  62.79%; recall:  61.56%; FB1:  62.17  301
              PER: precision:  67.74%; recall:  73.98%; FB1:  70.73  403
(72.33864207879296, 70.10560519902518, 71.2046204620462)

5 random evaluation samples:
SENT: June <unk> v British <unk> ( at Oxford , three days )
TRUE: O O O I-ORG I-ORG O O I-LOC O O O O
PRED: O I-ORG O I-ORG I-ORG O O I-LOC O O O O
```

```
SENT: <unk> 0000 0000 0000 0000 0000 0000 0000
TRUE: I-ORG O O O O O O O O
PRED: I-ORG O O O O O O O O
SENT: From central Texas north to Kansas , <unk> throughout July and August have <unk> mos
t of the <unk> conditions that <unk> the region earlier this year .
TRUE: O O I-LOC O O I-LOC O O O O O O O O O O O O O O O O O O O O O
PRED: O O I-LOC O O I-LOC O O O O O O O O O O O O O O O O O O O O O
SENT: There is nothing left for us but to be <unk> to <unk> for <unk> <unk> , " <unk> <unk
> said .
TRUE: O O O O O O O O O O O O O O O O O I-ORG I-ORG O O
PRED: O O O O O O O O O O O O O O O O O I-PER I-PER O O
SENT: Pakistani bourse to use new <unk> index .
TRUE: I-MISC O O O O O O O
PRED: I-MISC O O O O I-MISC O O
--- EPOCH 18 ---
Avg loss over last 500 updates: 0.4740509853363037
Avg loss over last 500 updates: 0.5172397384643554
Avg loss over last 500 updates: 0.5227785496711731
Avg loss over last 500 updates: 0.4226785435676575
Avg loss over last 500 updates: 0.39808293783664706
Avg loss over last 500 updates: 0.5073284511566162
Avg evaluation loss: 3.69926515892148
processed 11170 tokens with 1231 phrases; found: 1215 phrases; correct: 870.
accuracy:  76.10%; (non-O)
accuracy:  93.84%; precision:  71.60%; recall:  70.67%; FB1:  71.14
              LOC: precision:  88.34%; recall:  79.34%; FB1:  83.60  326
             MISC: precision:  74.31%; recall:  55.73%; FB1:  63.69  144
              ORG: precision:  60.75%; recall:  63.52%; FB1:  62.10  321
              PER: precision:  66.04%; recall:  75.88%; FB1:  70.62  424
(71.60493827160494, 70.67424857839156, 71.13654946852003)

5 random evaluation samples:
SENT: The <unk> parliament failed for a third and final time to elect a new state presiden
t on Tuesday , refusing a second mandate for <unk> <unk> Meri .
TRUE: O I-MISC O O O O O O O O O O O O O O O O O O O O O O O I-PER I-PER O
PRED: O O O O O O O O O O O O O O O O O O O O O O O O O O I-PER O
SENT: Russian and rebel military <unk> finally met in Chechnya on Tuesday for delayed talk
s aimed at <unk> a ceasefire <unk> last week by President <unk> Yeltsin 's <unk> Alexander
Lebed .
TRUE: I-MISC O O O O O O O I-LOC O O O O O O O O O O O O O O O O O O I-PER I-PER O O I-PER I-PER
O
PRED: I-MISC O O O O O O O I-LOC O O O O O O O O O O O O O O O O O O I-PER I-PER O I-PER I-PER I
-PER O
SENT: -- New York <unk> <unk> <unk> 0000 0000 0000
TRUE: O I-ORG I-ORG I-ORG I-ORG O O O O
PRED: O I-ORG I-ORG I-ORG I-ORG I-ORG O O O
SENT: -DOCSTART-
TRUE: O
PRED: O
SENT: <unk>
TRUE: O
PRED: O
--- EPOCH 19 ---
Avg loss over last 500 updates: 0.38547558069229126
Avg loss over last 500 updates: 0.47908888721466064
Avg loss over last 500 updates: 0.434932984828949
Avg loss over last 500 updates: 0.4043343040943146
Avg loss over last 500 updates: 0.32861188769340516
Avg loss over last 500 updates: 0.38310386300086974
Avg evaluation loss: 3.961290597617626
processed 11170 tokens with 1231 phrases; found: 1235 phrases; correct: 876.
accuracy:  76.98%; (non-O)
accuracy:  93.79%; precision:  70.93%; recall:  71.16%; FB1:  71.05
              LOC: precision:  90.31%; recall:  79.61%; FB1:  84.63  320
             MISC: precision:  71.62%; recall:  55.21%; FB1:  62.35  148
              ORG: precision:  65.31%; recall:  62.54%; FB1:  63.89  294
```

```
              PER: precision:  61.10%; recall:  78.32%; FB1:  68.65  473
(70.93117408906883, 71.16165718927701, 71.04622871046229)

5 random evaluation samples:
SENT: The <unk> of his other <unk> <unk> as " In the Year of January " ( 0000 ) , " The <u
nk> " ( 0000 ) , " <unk> <unk> " ( 0000 ) , " The <unk> <unk> " ( 0000 ) and " A World of
<unk> " ( 0000 ) , followed by " The <unk> , " " <unk> <unk> " and , most recently , " A <
unk> . "
TRUE: O O O O O O O O O I-MISC I-MISC I-MISC I-MISC I-MISC O O O O O O I-MISC I-MISC O O O
O O O I-MISC I-MISC O O O O O O I-MISC I-MISC I-MISC O O O O O O I-MISC I-MISC I-MISC I-MI
SC O O O O O O O O I-MISC I-MISC O O O I-MISC I-MISC O O O O O O O I-MISC I-MISC O O
PRED: O O O O O O O O O O O O O O O O O O O O O O O O O O O O I-PER I-PER O O O O O O O I-
ORG I-ORG O O O O O O O O O I-MISC O O O O O O O O O O O O O I-MISC I-MISC O O O O O O O O
O O O
SENT: Summary of game played in the Spanish first division on Saturday : <unk> <unk> 0000
( <unk> <unk> , <unk> minute ) Real Madrid 0000 ( <unk> Carlos <unk> ) .
TRUE: O O O O O O I-MISC O I-MISC O O O I-ORG I-ORG O O I-PER I-PER O O O O I-ORG I-ORG O
O I-PER I-PER O O O
PRED: O O O O O O I-MISC O O O O O I-PER I-PER O O I-PER I-PER O I-PER O O I-ORG I-ORG O O
I-PER I-PER O O O
SENT: Brazil - Giovanni ( <unk> ) , <unk> <unk> ( <unk> )
TRUE: I-LOC O I-PER O O O O I-PER I-PER O O O
PRED: I-LOC O I-PER O O O O I-PER I-PER O I-ORG O
SENT: <unk> 0000 0000 0000 0000 0000 0000 0000
TRUE: I-ORG O O O O O O O
PRED: I-ORG O O O O O O O
SENT: CRICKET - SRI LANKA BEAT <unk> <unk> 0000 <unk> IN ONE-DAY MATCH .
TRUE: O O I-LOC I-LOC O I-LOC O O O O O O O
PRED: O O I-LOC I-LOC O I-ORG I-ORG O O O O O O
--- EPOCH 20 ---
Avg loss over last 500 updates: 0.48135097217559814
Avg loss over last 500 updates: 0.4950535833835602
Avg loss over last 500 updates: 0.5244938030242919
Avg loss over last 500 updates: 0.378475266456604
Avg loss over last 500 updates: 0.3657068989276886
Avg loss over last 500 updates: 0.47909418869018555
Avg evaluation loss: 3.67012115880847
processed 11170 tokens with 1231 phrases; found: 1194 phrases; correct: 858.
accuracy:  75.11%; (non-O)
accuracy:  93.75%; precision:  71.86%; recall:  69.70%; FB1:  70.76
              LOC: precision:  86.63%; recall:  78.51%; FB1:  82.37  329
             MISC: precision:  72.48%; recall:  56.25%; FB1:  63.34  149
              ORG: precision:  60.56%; recall:  63.52%; FB1:  62.00  322
              PER: precision:  68.53%; recall:  73.17%; FB1:  70.77  394
(71.85929648241206, 69.69943135662064, 70.76288659793815)

5 random evaluation samples:
SENT: A South African boy is writing back to an American girl whose message in a <unk> he
found <unk> up on President Nelson <unk> 's <unk> prison island .
TRUE: O I-MISC I-MISC O O O O O O I-MISC O O O O O O O O O O O O I-PER I-PER O O O O O
PRED: O I-MISC I-MISC O O O O O O I-ORG O O O O O O O O O O O O I-PER I-PER O I-MISC O O O
SENT: <unk> <unk> 0000 - <unk> 0000 : <unk> of <unk> at $ <unk> , 10,000 of <unk> at $ <un
k> , <unk> of <unk> at $ <unk> .
TRUE: O O O O O O O O O O O O O O O O O O O O O O O O O O O O
PRED: I-PER I-PER O O I-PER O O O O O O O O O O O O O O O O O O O O O O O
SENT: In an interview following its first-half results , which included a less <unk> forec
ast for the second half of this year than it had made in the past , Sir Colin <unk> said <
unk> had taken defensive action to <unk> it from <unk> markets .
TRUE: O O O O O O O O O O O O O O O O O O O O O O O O O O O O O O I-PER I-PER O I-ORG O O
O O O O O O O O O
PRED: O O O O O O O O O O O O O O O O O O O O O O O O O O O O O O I-PER I-PER O I-ORG O O
O O O O O O O O O
SENT: <unk> 's autonomy was <unk> in 0000 and Serb police forces <unk> down on Albanian pr
otests .
TRUE: I-LOC O O O O O O O I-MISC O O O O O I-MISC O O
PRED: I-ORG O O O O O O O I-MISC O O O O O I-MISC O O
```

```
SENT: - <unk> 0000 down <unk>
TRUE: O I-LOC O O O
PRED: O I-PER O O O
--- EPOCH 21 ---
Avg loss over last 500 updates: 0.438204106092453
Avg loss over last 500 updates: 0.45517674469947816
Avg loss over last 500 updates: 0.400373046875
Avg loss over last 500 updates: 0.3606664307117462
Avg loss over last 500 updates: 0.3681542270183563
Avg loss over last 500 updates: 0.3900785427093506
Avg evaluation loss: 3.8932757551968096
processed 11170 tokens with 1231 phrases; found: 1214 phrases; correct: 868.
accuracy:  76.49%; (non-O)
accuracy:  93.97%; precision:  71.50%; recall:  70.51%; FB1:  71.00
              LOC: precision:  88.41%; recall:  79.89%; FB1:  83.94  328
             MISC: precision:  75.91%; recall:  54.17%; FB1:  63.22  137
              ORG: precision:  58.66%; recall:  62.87%; FB1:  60.69  329
              PER: precision:  66.90%; recall:  76.15%; FB1:  71.23  420
(71.49917627677101, 70.51177904142973, 71.00204498977506)

5 random evaluation samples:
SENT: * <unk> <unk> lost <unk> percent to 0000 francs after a morning trading <unk> during
which it said it had approved plans to buy out its <unk> percent owned transport unit <unk
> <unk> <unk> ( <unk> ) and <unk> shareholders to <unk> their shares .
TRUE: O O I-ORG O O O O O O O O O O O O O O O O O O O O O O O O O O O O O O I-ORG I-ORG I-ORG
O I-ORG O O O O O O O O O O
PRED: O I-ORG I-ORG O O O O O O O O O O O O O O O O O O O O O O O O O O O O O I-ORG I-ORG I-
ORG O I-ORG O O I-ORG O O O O O O
SENT: 1. <unk> 0000
TRUE: O I-PER O
PRED: O I-PER O
SENT: In the <unk> , between 0000 to 0000 percent of the muscle <unk> in one group of <unk
> produced <unk> for two weeks before <unk> .
TRUE: O O O O O O O O O O O O O O O O O O O O O O O O O O
PRED: O O O O O O O O O O O O O O O O O I-ORG O O O O O O O O
SENT: <unk> <unk> ( <unk> , <unk> , <unk> ) , <unk> <unk> ( <unk> ) , Petr
TRUE: I-PER I-PER O O O O O O O O I-PER I-PER O O O O I-PER
PRED: I-PER I-PER O O O O O O O O I-PER I-PER O I-ORG O O I-PER
SENT: What was its right ( to the money ) <unk> do not know , " Rosati told a news confere
nce .
TRUE: O O O O O O O O O O O O O O O I-PER O O O O O
PRED: O O O O O O O O O O O O O O O I-PER O O O O O
--- EPOCH 22 ---
Avg loss over last 500 updates: 0.4492936780452728
Avg loss over last 500 updates: 0.3838615689277649
Avg loss over last 500 updates: 0.42850667905807494
Avg loss over last 500 updates: 0.3253185479640961
Avg loss over last 500 updates: 0.2941946313381195
Avg loss over last 500 updates: 0.4442888958454132
Avg evaluation loss: 3.8677443878352644
processed 11170 tokens with 1231 phrases; found: 1195 phrases; correct: 862.
accuracy:  75.55%; (non-O)
accuracy:  94.03%; precision:  72.13%; recall:  70.02%; FB1:  71.06
              LOC: precision:  86.09%; recall:  80.17%; FB1:  83.02  338
             MISC: precision:  72.86%; recall:  53.12%; FB1:  61.45  140
              ORG: precision:  58.43%; recall:  63.19%; FB1:  60.72  332
              PER: precision:  71.43%; recall:  74.53%; FB1:  72.94  385
(72.13389121338912, 70.02437043054427, 71.06347897774114)

5 random evaluation samples:
SENT: <unk> Gold Inc was up C$ <unk> to C$ <unk> in trading of <unk> shares , while <unk>
Gold Corp gained C$ <unk> to C$ 0000 in volume of <unk> shares .
TRUE: I-ORG I-ORG I-ORG O O I-MISC O O I-MISC O O O O O O O O I-ORG I-ORG I-ORG O I-MISC O
O I-MISC O O O O O O O
PRED: I-ORG I-ORG I-ORG O O I-MISC O O I-MISC O O O O O O O O I-ORG I-ORG I-ORG O I-MISC O
O I-MISC O O O O O O O
```

```
SENT: Chesterfield 0000 <unk> 0000
TRUE: I-ORG O I-ORG O
PRED: I-ORG O I-ORG O
SENT: -DOCSTART-
TRUE: O
PRED: O
SENT: U.S. debt futures end lower , <unk> by Chicago <unk> .
TRUE: I-LOC O O O O O O O O I-LOC I-ORG O
PRED: I-LOC O O O O O O O O I-LOC O O
SENT: - Pakistan <unk> <unk> tonnes of <unk> <unk> yellow <unk> from <unk> Inc for $ <unk>
per tonne , <unk> U.S. Gulf , agents for the <unk> said .
TRUE: O I-LOC O O O O O O O O O I-ORG I-ORG O O O O O O O I-LOC I-LOC O O O O O O O
PRED: O I-LOC O O O O I-ORG I-ORG O O O I-ORG I-ORG O O O O O O O I-LOC I-LOC O O O O I-OR
G O O
--- EPOCH 23 ---
Avg loss over last 500 updates: 0.37144859838485716
Avg loss over last 500 updates: 0.38163819003105165
Avg loss over last 500 updates: 0.4294299356937408
Avg loss over last 500 updates: 0.2975582957267761
Avg loss over last 500 updates: 0.28728287744522096
Avg loss over last 500 updates: 0.3545259051322937
Avg evaluation loss: 4.057561903223395
processed 11170 tokens with 1231 phrases; found: 1185 phrases; correct: 870.
accuracy:  75.61%; (non-O)
accuracy:  94.06%; precision:  73.42%; recall:  70.67%; FB1:  72.02
              LOC: precision:  89.41%; recall:  79.06%; FB1:  83.92  321
             MISC: precision:  76.19%; recall:  58.33%; FB1:  66.08  147
              ORG: precision:  64.11%; recall:  59.93%; FB1:  61.95  287
              PER: precision:  66.74%; recall:  77.78%; FB1:  71.84  430
(73.41772151898735, 70.67424857839156, 72.01986754966889)

5 random evaluation samples:
SENT: Poland 's Foreign Minister <unk> Rosati will visit Yugoslavia on September 0000 and
0000 to revive a <unk> between the two governments which was <unk> <unk> in 0000 , <unk> n
ews agency reported on Friday .
TRUE: I-LOC O O O I-PER I-PER O O I-LOC O O O O O O O O O O O O O O O O O O O O I-ORG O O
O O O O
PRED: I-LOC O O O I-PER I-PER O O I-LOC O O O O O O O O O O O O O O O O O O O O I-ORG O O
O O O O
SENT: But about 0000 university students were still <unk> outside the <unk> of their <unk>
, <unk> said .
TRUE: O O O O O O O O O O O O O O O O O O
PRED: O O O O O O O O O O O O O O O I-PER O O
SENT: <unk> , <unk> .
TRUE: I-LOC O I-LOC O
PRED: O O I-PER O
SENT: Australia at South Africa
TRUE: I-LOC O I-LOC I-LOC
PRED: I-LOC O I-LOC I-LOC
SENT: They said the index could also rise towards <unk> if the <unk> share prices <unk> bu
yers .
TRUE: O O O O O O O O O O O O O O O O O
PRED: O O O O O O O O O O O O O O O O O
--- EPOCH 24 ---
Avg loss over last 500 updates: 0.36597043061256407
Avg loss over last 500 updates: 0.45555487847328185
Avg loss over last 500 updates: 0.40572158312797546
Avg loss over last 500 updates: 0.3514133563041687
Avg loss over last 500 updates: 0.3342776668071747
Avg loss over last 500 updates: 0.3343775763511658
Avg evaluation loss: 4.276873261034488
processed 11170 tokens with 1231 phrases; found: 1245 phrases; correct: 870.
accuracy:  75.99%; (non-O)
accuracy:  93.63%; precision:  69.88%; recall:  70.67%; FB1:  70.27
              LOC: precision:  87.27%; recall:  79.34%; FB1:  83.12  330
             MISC: precision:  75.68%; recall:  58.33%; FB1:  65.88  148
```

```
              ORG: precision:  63.61%; recall:  60.91%; FB1:  62.23  294
              PER: precision:  59.83%; recall:  76.69%; FB1:  67.22  473
(69.87951807228916, 70.67424857839156, 70.27463651050081)

5 random evaluation samples:
SENT: But <unk> said it would be " <unk> " against the other <unk> if only Balkan planes w
ere <unk> in .
TRUE: O I-PER O O O O O O O O O O O O O I-LOC O O O O O
PRED: O I-ORG O O O O O O O O O O O O O O O O O O O
SENT: NEW YORK 1996-08-22
TRUE: I-LOC I-LOC O
PRED: I-LOC I-LOC O
SENT: <unk> newspapers said the men had been <unk> against " strategic <unk> and <unk> <un
k> " but gave no other details .
TRUE: I-ORG O O O O O O O O O O O O O O O O O O O O O
PRED: O O O O O O O O O O O O O O O O O O O O O O
SENT: " Now many Russian banks are strong and can make various <unk> of money <unk> , whil
e <unk> traders are being ousted by more <unk> ones .
TRUE: O O O I-MISC O O O O O O O O O O O O O O O O O O O O O O O
PRED: O O O I-MISC O O O O O O O O O O O O O I-ORG O O O O O O O O O
SENT: Jose <unk> <unk> , who had drawn up the bill .
TRUE: I-PER I-PER I-PER O O O O O O O O
PRED: I-PER I-PER I-PER O O O O O O O O
--- EPOCH 25 ---
Avg loss over last 500 updates: 0.39553758096694946
Avg loss over last 500 updates: 0.3998786451816559
Avg loss over last 500 updates: 0.40556255626678467
Avg loss over last 500 updates: 0.32037803483009336
Avg loss over last 500 updates: 0.29176287245750426
Avg loss over last 500 updates: 0.35602904987335204
Avg evaluation loss: 4.165012877732515
processed 11170 tokens with 1231 phrases; found: 1210 phrases; correct: 880.
accuracy:  76.05%; (non-O)
accuracy:  94.15%; precision:  72.73%; recall:  71.49%; FB1:  72.10
              LOC: precision:  87.84%; recall:  79.61%; FB1:  83.53  329
             MISC: precision:  79.14%; recall:  57.29%; FB1:  66.47  139
              ORG: precision:  65.10%; recall:  63.19%; FB1:  64.13  298
              PER: precision:  64.64%; recall:  77.78%; FB1:  70.60  444
(72.72727272727273, 71.48659626320065, 72.10159770585827)

5 random evaluation samples:
SENT: Roe said he was <unk> by his forecast of a <unk> billion <unk> net for 0000 .
TRUE: I-PER O O O O O O O O O O O O O O O O
PRED: I-PER O O O O O O O O O O O O O O O O
SENT: <unk> <unk> 0000 0000 0000 0000 0000 0000 0000
TRUE: I-ORG I-ORG O O O O O O O
PRED: I-ORG I-ORG O O O O O O O
SENT: The <unk> of his other <unk> <unk> as " In the Year of January " ( 0000 ) , " The <u
nk> " ( 0000 ) , " <unk> <unk> " ( 0000 ) , " The <unk> <unk> " ( 0000 ) and " A World of
<unk> " ( 0000 ) , followed by " The <unk> , " " <unk> <unk> " and , most recently , " A <
unk> . "
TRUE: O O O O O O O O O I-MISC I-MISC I-MISC I-MISC I-MISC O O O O O O I-MISC I-MISC O O O
O O O I-MISC I-MISC O O O O O O I-MISC I-MISC I-MISC O O O O O O I-MISC I-MISC I-MISC I-MI
SC O O O O O O O O I-MISC I-MISC O O O I-MISC I-MISC O O O O O O O I-MISC I-MISC O O
PRED: O O O O O O O O O O O O O O O O O O O O O O O O O O O O I-PER I-PER O O O O O O O I-
ORG I-ORG O O O O O O O O O O O O O O O O O O O O O O O I-MISC I-MISC O O O O O O O O O O
O
SENT: <unk> 0000 0000 0000 0000 0000 0000 0000
TRUE: I-ORG O O O O O O O
PRED: I-ORG O O O O O O O
SENT: It was the second arms <unk> this week .
TRUE: O O O O O O O O O
PRED: O O O O O O O O O
--- EPOCH 26 ---
Avg loss over last 500 updates: 0.3498471896648407
Avg loss over last 500 updates: 0.37959309267997743
```

```
Avg loss over last 500 updates: 0.3461668291091919
Avg loss over last 500 updates: 0.32090911626815793
Avg loss over last 500 updates: 0.262311635017395
Avg loss over last 500 updates: 0.3424121060371399
Avg evaluation loss: 4.212019420862198
processed 11170 tokens with 1231 phrases; found: 1243 phrases; correct: 860.
accuracy:  75.72%; (non-O)
accuracy:  93.69%; precision:  69.19%; recall:  69.86%; FB1:  69.52
              LOC: precision:  82.52%; recall:  79.34%; FB1:  80.90  349
             MISC: precision:  68.75%; recall:  57.29%; FB1:  62.50  160
              ORG: precision:  60.78%; recall:  60.59%; FB1:  60.69  306
              PER: precision:  64.49%; recall:  74.80%; FB1:  69.26  428
(69.18744971842317, 69.86190089358246, 69.52303961196442)

5 random evaluation samples:
SENT: -DOCSTART-
TRUE: O
PRED: O
SENT: As a result of <unk> we are looking at <unk> <unk> more <unk> to see if the profit m
argins are <unk> enough to <unk> things like <unk> security . "
TRUE: O O O O O O O O O O O O O O O O O O O O O O O O O O O O O O
PRED: O O O O O O O O O O O O O O O O O O O O O O O O O O O O O O
SENT: No arrests had been made , a police spokesman said .
TRUE: O O O O O O O O O O O
PRED: O O O O O O O O O O O
SENT: <unk> Bremen 0000 <unk> <unk> 0000
TRUE: I-ORG I-ORG O I-ORG I-ORG O
PRED: I-ORG I-ORG O I-ORG I-ORG O
SENT: AUG 0000 <unk> 'S <unk> 0000 <unk> 0000 <unk>
TRUE: O O O O O O O O O
PRED: O O I-ORG I-ORG I-ORG O I-ORG O O
--- EPOCH 27 ---
Avg loss over last 500 updates: 0.3462287621498108
Avg loss over last 500 updates: 0.3938676681518555
Avg loss over last 500 updates: 0.3202833204269409
Avg loss over last 500 updates: 0.28946461629867554
Avg loss over last 500 updates: 0.3295359447002411
Avg loss over last 500 updates: 0.2931487522125244
Avg evaluation loss: 4.178279275149107
processed 11170 tokens with 1231 phrases; found: 1188 phrases; correct: 863.
accuracy:  75.11%; (non-O)
accuracy:  94.15%; precision:  72.64%; recall:  70.11%; FB1:  71.35
              LOC: precision:  86.65%; recall:  80.44%; FB1:  83.43  337
             MISC: precision:  71.24%; recall:  56.77%; FB1:  63.19  153
              ORG: precision:  63.01%; recall:  59.93%; FB1:  61.44  292
              PER: precision:  68.47%; recall:  75.34%; FB1:  71.74  406
(72.64309764309765, 70.10560519902518, 71.35179826374535)

5 random evaluation samples:
SENT: <unk> 0000 <unk> 0000
TRUE: I-ORG O I-ORG O
PRED: I-ORG O I-ORG O
SENT: <unk>
TRUE: O
PRED: O
SENT: <unk> AT SAN FRANCISCO
TRUE: I-ORG O I-LOC I-LOC
PRED: I-ORG O I-LOC I-LOC
SENT: ( Aug 0000 ) ( Jul 0000 )
TRUE: O O O O O O O O
PRED: O O O O O O O O
SENT: " We <unk> to start international telephone business as soon as possible , " a compa
ny official told Reuters .
TRUE: O O O O O O O O O O O O O O O O O O I-ORG O
PRED: O O O O O O O O O O O O O O O O O O I-ORG O
--- EPOCH 28 ---
```

```
Avg loss over last 500 updates: 0.31965131902694705
Avg loss over last 500 updates: 0.3676064219474792
Avg loss over last 500 updates: 0.32074091863632204
Avg loss over last 500 updates: 0.3590176103115082
Avg loss over last 500 updates: 0.28661622190475466
Avg loss over last 500 updates: 0.29312648105621336
Avg evaluation loss: 4.250497308522463
processed 11170 tokens with 1231 phrases; found: 1219 phrases; correct: 864.
accuracy:  75.39%; (non-O)
accuracy:  93.84%; precision:  70.88%; recall:  70.19%; FB1:  70.53
              LOC: precision:  87.16%; recall:  80.44%; FB1:  83.67  335
             MISC: precision:  68.15%; recall:  55.73%; FB1:  61.32  157
              ORG: precision:  58.39%; recall:  61.24%; FB1:  59.78  322
              PER: precision:  68.40%; recall:  75.07%; FB1:  71.58  405
(70.87776866283839, 70.18683996750609, 70.53061224489797)

5 random evaluation samples:
SENT: .
TRUE: O
PRED: O
SENT: From central Texas north to Kansas , <unk> throughout July and August have <unk> mos
t of the <unk> conditions that <unk> the region earlier this year .
TRUE: O O I-LOC O O I-LOC O O O O O O O O O O O O O O O O O O O O O
PRED: O O I-LOC I-LOC O I-LOC O O O O O O O O O O O O O O O O O O O O O
SENT: -DOCSTART-
TRUE: O
PRED: O
SENT: The <unk> parliament failed for a third and final time to elect a new state presiden
t on Tuesday , refusing a second mandate for <unk> <unk> Meri .
TRUE: O I-MISC O O O O O O O O O O O O O O O O O O O O O O O I-PER I-PER O
PRED: O O O O O O O O O O O O O O O O O O O O O O O O O O I-PER O
SENT: In July , the average rate on a <unk> <unk> rate mortgage was <unk> percent , below
June 's <unk> percent but still higher than February 's <unk> percent , the report <unk> .
TRUE: O O O O O O O O O O O O O O O O O O O O O O O O O O O O O O O O O O
PRED: O O O O O O O O O O O O O O O O O O O O O O O O O O O O O O O O O O
--- EPOCH 29 ---
Avg loss over last 500 updates: 0.26736956453323363
Avg loss over last 500 updates: 0.3369770126342773
Avg loss over last 500 updates: 0.40757529354095456
Avg loss over last 500 updates: 0.29582179522514346
Avg loss over last 500 updates: 0.2658300094604492
Avg loss over last 500 updates: 0.31865103435516356
Avg evaluation loss: 4.538428762853146
processed 11170 tokens with 1231 phrases; found: 1176 phrases; correct: 867.
accuracy:  75.39%; (non-O)
accuracy:  94.17%; precision:  73.72%; recall:  70.43%; FB1:  72.04
              LOC: precision:  88.41%; recall:  79.89%; FB1:  83.94  328
             MISC: precision:  76.03%; recall:  57.81%; FB1:  65.68  146
              ORG: precision:  65.14%; recall:  60.26%; FB1:  62.61  284
              PER: precision:  67.22%; recall:  76.15%; FB1:  71.41  418
(73.72448979591837, 70.43054427294882, 72.03988367262153)

5 random evaluation samples:
SENT: <unk> 1996-08-27
TRUE: I-LOC O
PRED: I-LOC O
SENT: It is an <unk> of Clinton 's strategic planning as he <unk> into the <unk> drive for
the Nov. 0000 presidential election .
TRUE: O O O O O I-PER O O O O O O O O O O O O O O O O O
PRED: O O O O O I-PER O O O O O O O O O O O O O O O O O
SENT: LOS ANGELES 0000 0000 <unk> 0000
TRUE: I-ORG I-ORG O O O O
PRED: I-ORG I-ORG O O O O
SENT: After I won I <unk> I could give them a little <unk> . "
TRUE: O O O O O O O O O O O O O O
PRED: O O O O O O O O O O O O O O
```

```
SENT: GLASGOW 1996-12-07
TRUE: I-LOC O
PRED: I-LOC O
Wall time: 2h 43s
```

In [12]:

```python
df = pd.DataFrame(columns = ['LOC_crf', 'MISC_crf', 'ORG_crf','PER_crf'])
```

In [45]:

```python
df = df.append({'LOC' : 83.94, 'MISC_crf' : 65.68, 'ORG_crf' : 62.61 , 'PER_crf':71.41 },
                ignore_index = True)
```

In [53]:

```python
df
```

Out[53]:

	MISC_crf	ORG_crf	PER_crf	LOC_crf
0	33.33	47.04	53.00	57.71
1	57.81	54.38	62.11	71.34
2	61.04	57.14	66.95	78.42
3	64.33	59.44	70.29	79.03
4	62.18	58.30	69.77	80.96
5	62.31	60.84	70.44	78.79
6	64.17	62.94	69.65	80.96
7	66.06	62.10	72.18	83.46
8	65.20	61.67	70.62	82.13
9	61.63	64.03	69.15	82.49
10	65.85	64.95	69.89	82.51
11	65.23	59.51	72.19	84.04
12	64.85	61.25	71.07	83.09
13	62.82	60.39	71.01	83.73
14	62.24	59.32	70.82	82.15
15	64.46	61.13	72.77	82.13
16	65.44	61.12	73.61	83.87
17	63.74	62.17	70.73	83.19
18	63.69	62.10	70.62	83.60
19	62.35	63.89	68.65	84.63
20	63.34	62.00	70.77	82.37
21	63.22	60.69	71.23	83.94
22	61.45	60.72	72.94	83.02
23	66.08	61.95	71.84	83.92
24	65.88	62.23	67.22	83.12
25	66.47	64.13	70.60	83.53
26	62.50	60.69	69.26	80.90

	MISC_crf	ORG_crf	PER_crf	LOC_crf
27	63.19	61.44	71.74	83.43
28	61.32	59.78	71.58	83.67
29	65.68	62.61	71.41	83.94

In [195…

```
%%time
# Train BiLSTM Tagger Baseline
#model = BiLSTMTagger3(len(word_vocab), tag_to_ix, 128, 256).to(device)
model2 = BiLSTMTagger(len(word_vocab), 8, 128, 256).to(device)
optimizer = optim.Adam(model2.parameters(), lr=1e-3)
train(model2, train_data, valid_data, word_vocab, label_vocab, epochs=30, log_interval=50(
```

```
hidden_dim:  256
target_size:  8
--- EPOCH 0 ---
Avg loss over last 500 updates: 9.207625520944596
Avg loss over last 500 updates: 8.203925667714328
Avg loss over last 500 updates: 6.6126301279459145
Avg loss over last 500 updates: 5.7041344931358475
Avg loss over last 500 updates: 4.952978737355676
Avg loss over last 500 updates: 5.1556796167618595
Avg evaluation loss: 4.763876983722221
processed 11170 tokens with 1231 phrases; found: 736 phrases; correct: 471.
accuracy:  43.17%; (non-O)
accuracy:  89.62%; precision:  63.99%; recall:  38.26%; FB1:  47.89
              LOC: precision:  79.33%; recall:  39.12%; FB1:  52.40  179
             MISC: precision:  64.41%; recall:  19.79%; FB1:  30.28  59
              ORG: precision:  63.10%; recall:  34.53%; FB1:  44.63  168
              PER: precision:  56.06%; recall:  50.14%; FB1:  52.93  330
(63.99456521739131, 38.26157595450853, 47.89018810371123)

5 random evaluation samples:
SENT: ( <unk> ) 7-6 ( 7-4 )
TRUE: O O O O O O O
PRED: O O O O O O O
SENT: 10. <unk> <unk> ( Italy ) <unk> <unk> <unk>
TRUE: O I-PER I-PER O I-LOC O I-MISC I-MISC O
PRED: O I-PER I-PER O I-LOC O O O O
SENT: <unk> tested positive for the <unk> <unk> after the fifth stage of the Tour , in whi
ch he finished third overall .
TRUE: I-PER O O O O O O O O O O O I-MISC O O O O O O O O
PRED: O O O O O O O O O O O O O O O O O O O O O
SENT: <unk> league matches on Sunday :
TRUE: O O O O O O
PRED: O O O O O O
SENT: Hong Kong 's Tsang sees growth , <unk> <unk> .
TRUE: I-LOC I-LOC O I-PER O O O O O O
PRED: I-LOC I-LOC O O O O O O O O
--- EPOCH 1 ---
Avg loss over last 500 updates: 4.386172959215066
Avg loss over last 500 updates: 4.940935809938179
Avg loss over last 500 updates: 4.182841504598487
Avg loss over last 500 updates: 3.832591070348799
Avg loss over last 500 updates: 3.3322007677918517
Avg loss over last 500 updates: 3.5130282022433237
Avg evaluation loss: 3.758712802360669
processed 11170 tokens with 1231 phrases; found: 953 phrases; correct: 644.
accuracy:  58.87%; (non-O)
accuracy:  91.95%; precision:  67.58%; recall:  52.32%; FB1:  58.97
              LOC: precision:  86.25%; recall:  57.02%; FB1:  68.66  240
             MISC: precision:  71.00%; recall:  36.98%; FB1:  48.63  100
```

```
              ORG: precision:  54.51%; recall:  47.23%; FB1:  50.61  266
              PER: precision:  63.69%; recall:  59.89%; FB1:  61.73  347
(67.57607555089193, 52.31519090170593, 58.97435897435898)

5 random evaluation samples:
SENT: <unk> 0000 <unk> 0000
TRUE: I-ORG O I-ORG O
PRED: I-ORG O I-ORG O
SENT: 0000 - <unk> <unk> ( South Africa ) beat Todd <unk>
TRUE: O O I-PER I-PER O I-LOC I-LOC O O I-PER I-PER
PRED: O O I-PER I-PER O I-LOC I-LOC O O I-PER I-PER
SENT: <unk> win 2-1 on aggregate .
TRUE: I-ORG O O O O O
PRED: O O O O O O
SENT: NEW YORK 1996-08-22
TRUE: I-LOC I-LOC O
PRED: I-LOC I-LOC O
SENT: <unk> <unk> ISS <unk> <unk> PAY DATE <unk>
TRUE: O O O O O O O O
PRED: O O O O O O O O
--- EPOCH 2 ---
Avg loss over last 500 updates: 3.224115241336611
Avg loss over last 500 updates: 3.6249761232665914
Avg loss over last 500 updates: 3.0132237215502156
Avg loss over last 500 updates: 3.0386997171622117
Avg loss over last 500 updates: 2.5953576887795027
Avg loss over last 500 updates: 2.795695305746527
Avg evaluation loss: 3.416785683650573
processed 11170 tokens with 1231 phrases; found: 1034 phrases; correct: 714.
accuracy:  64.54%; (non-O)
accuracy:  92.85%; precision:  69.05%; recall:  58.00%; FB1:  63.05
              LOC: precision:  83.88%; recall:  63.09%; FB1:  72.01  273
             MISC: precision:  78.95%; recall:  46.88%; FB1:  58.82  114
              ORG: precision:  58.16%; recall:  53.42%; FB1:  55.69  282
              PER: precision:  63.29%; recall:  62.60%; FB1:  62.94  365
(69.0522243713733, 58.001624695369614, 63.046357615894046)

5 random evaluation samples:
SENT: -DOCSTART-
TRUE: O
PRED: O
SENT: <unk> police shot dead six <unk> suspects as they tried to escape from <unk> in the
northern city of <unk> , the national news agency reported on Friday .
TRUE: I-MISC O O O O O O O O O O O O O O O O O O O O I-LOC O O O O O O O O O O O
PRED: O O O O O O O O O O O O O O O O O O O O I-LOC O O O O O O O O O O O
SENT: At Leicester : <unk> drawn .
TRUE: O I-LOC O O O O
PRED: O O O O O O
SENT: The defeat put the <unk> out of the <unk> Cup .
TRUE: O O O O I-MISC O O O I-MISC I-MISC O
PRED: O O O O O O O O O I-MISC O
SENT: He walked three and struck out three in winning for the 10th time in his last 0000 <
unk> .
TRUE: O O O O O O O O O O O O O O O O O O O
PRED: O O O O O O O O O O O O O O O O O O O
--- EPOCH 3 ---
Avg loss over last 500 updates: 2.446811728320189
Avg loss over last 500 updates: 2.8625406725368676
Avg loss over last 500 updates: 2.3376253051262608
Avg loss over last 500 updates: 2.487354991577225
Avg loss over last 500 updates: 2.083202376675463
Avg loss over last 500 updates: 2.253242397779996
Avg evaluation loss: 3.172191244728601
processed 11170 tokens with 1231 phrases; found: 1062 phrases; correct: 754.
accuracy:  67.68%; (non-O)
accuracy:  93.28%; precision:  71.00%; recall:  61.25%; FB1:  65.77
```

```
              LOC: precision:  86.24%; recall:  70.80%; FB1:  77.76  298
             MISC: precision:  78.69%; recall:  50.00%; FB1:  61.15  122
              ORG: precision:  57.89%; recall:  53.75%; FB1:  55.74  285
              PER: precision:  66.11%; recall:  63.96%; FB1:  65.01  357
(70.99811676082862, 61.25101543460602, 65.76537287396424)

5 random evaluation samples:
SENT: <unk> <unk> <unk>
TRUE: O O O
PRED: I-PER O O
SENT: Iran <unk> <unk> <unk> Iran
TRUE: I-MISC I-MISC O O I-LOC
PRED: I-LOC O O O I-LOC
SENT: for , against , points ) :
TRUE: O O O O O O O
PRED: O O O O O O O
SENT: <unk> , 0000 , had complained of lower back <unk> since a trip to Taiwan in May , wh
en <unk> forced her to go to <unk> <unk> <unk> for an <unk> .
TRUE: I-PER O O O O O O O O O O O O O I-LOC O O O O O O O O O O I-LOC I-LOC I-LOC O O O O
PRED: I-PER O O O O O O O O O O O O O O O O O O O O O O O O O O O O O O O
SENT: Russian and rebel military <unk> finally met in Chechnya on Tuesday for delayed talk
s aimed at <unk> a ceasefire <unk> last week by President <unk> Yeltsin 's <unk> Alexander
Lebed .
TRUE: I-MISC O O O O O O O I-LOC O O O O O O O O O O O O O O O O I-PER I-PER O O I-PER I-PER
O
PRED: I-MISC O O O O O O O I-LOC O O O O O O O O O O O O O O O O I-PER I-PER O O I-PER I-PER
O
--- EPOCH 4 ---
Avg loss over last 500 updates: 2.055506019696447
Avg loss over last 500 updates: 2.437038330635055
Avg loss over last 500 updates: 1.9619857788445643
Avg loss over last 500 updates: 2.0284820533829637
Avg loss over last 500 updates: 1.7362450845651989
Avg loss over last 500 updates: 1.9579637778124224
Avg evaluation loss: 3.126044884925497
processed 11170 tokens with 1231 phrases; found: 1070 phrases; correct: 781.
accuracy:  69.93%; (non-O)
accuracy:  93.72%; precision:  72.99%; recall:  63.44%; FB1:  67.88
              LOC: precision:  87.12%; recall:  70.80%; FB1:  78.12  295
             MISC: precision:  81.45%; recall:  52.60%; FB1:  63.92  124
              ORG: precision:  59.53%; recall:  57.98%; FB1:  58.75  299
              PER: precision:  69.60%; recall:  66.40%; FB1:  67.96  352
(72.99065420560747, 63.444354183590576, 67.88352890047807)

5 random evaluation samples:
SENT: The Civil War <unk> <unk> is mostly <unk> joined by students during the school year
.
TRUE: O I-ORG I-ORG I-ORG I-ORG O O O O O O O O O O
PRED: O I-MISC O I-ORG I-ORG O O O O O O O O O O
SENT: He said at one point during a press conference : " I have seen my <unk> ( party mana
ger ) for next week which , of course , does n't mean very much to me now . "
TRUE: O O O O O O O O O O O O O O O O O O O O O O O O O O O O O O O O O O O O O O
PRED: O O O O O O O O O O O O O O O O O O O O O O O O O O O O O O O O O O O O O O
SENT: Leeds v Wimbledon
TRUE: I-ORG O I-ORG
PRED: I-ORG O I-ORG
SENT: <unk> <unk> <unk> <unk>
TRUE: O O O I-MISC
PRED: O O O O
SENT: JOHANNESBURG 1996-08-26
TRUE: I-LOC O
PRED: I-LOC O
--- EPOCH 5 ---
Avg loss over last 500 updates: 1.7986285788665484
Avg loss over last 500 updates: 2.064853647189681
Avg loss over last 500 updates: 1.7076912592592957
```

```
Avg loss over last 500 updates: 1.7689334122504206
Avg loss over last 500 updates: 1.4995551668749731
Avg loss over last 500 updates: 1.7255591286968157
Avg evaluation loss: 3.187378549123059
processed 11170 tokens with 1231 phrases; found: 1104 phrases; correct: 792.
accuracy:  70.43%; (non-O)
accuracy:  93.66%; precision:  71.74%; recall:  64.34%; FB1:  67.84
              LOC: precision:  86.56%; recall:  72.73%; FB1:  79.04  305
             MISC: precision:  80.30%; recall:  55.21%; FB1:  65.43  132
              ORG: precision:  60.76%; recall:  57.00%; FB1:  58.82  288
              PER: precision:  65.17%; recall:  66.94%; FB1:  66.04  379
(71.73913043478261, 64.33793663688058, 67.8372591006424)

5 random evaluation samples:
SENT: <unk> <unk> ( U.S. ) beat <unk> <unk> ( Netherlands ) 5-7 6-3 6-3
TRUE: I-PER I-PER O I-LOC O O I-PER I-PER O I-LOC O O O O
PRED: I-PER I-PER O I-LOC O O I-PER I-PER O I-LOC O O O O
SENT: <unk> 1996-08-25
TRUE: I-LOC O
PRED: O O
SENT: <unk> 0000 0000 0000 0000 0000 0000 0000
TRUE: I-ORG O O O O O O O
PRED: I-ORG O O O O O O O
SENT: Following are some of the main factors likely to <unk> Indonesian stocks on Thursday
:
TRUE: O O O O O O O O O O I-MISC O O O O
PRED: O O O O O O O O O O I-MISC O O O O
SENT: -DOCSTART-
TRUE: O
PRED: O
--- EPOCH 6 ---
Avg loss over last 500 updates: 1.516061978531084
Avg loss over last 500 updates: 1.6888382547295686
Avg loss over last 500 updates: 1.511420607591895
Avg loss over last 500 updates: 1.6242827976404206
Avg loss over last 500 updates: 1.3553802845773169
Avg loss over last 500 updates: 1.6584229665811536
Avg evaluation loss: 3.0778924888317274
processed 11170 tokens with 1231 phrases; found: 1114 phrases; correct: 819.
accuracy:  72.08%; (non-O)
accuracy:  94.04%; precision:  73.52%; recall:  66.53%; FB1:  69.85
              LOC: precision:  87.66%; recall:  74.38%; FB1:  80.48  308
             MISC: precision:  83.08%; recall:  56.25%; FB1:  67.08  130
              ORG: precision:  57.63%; recall:  60.26%; FB1:  58.92  321
              PER: precision:  72.11%; recall:  69.38%; FB1:  70.72  355
(73.51885098743267, 66.53127538586516, 69.8507462686567)

5 random evaluation samples:
SENT: <unk> 0000 COLORADO 0000
TRUE: I-ORG O I-ORG O
PRED: I-ORG O I-ORG O
SENT: " The Foreign Ministry is trying to find out from the Greek embassy why Albanian ref
ugees have been <unk> from Greece , " <unk> told Reuters .
TRUE: O O I-ORG I-ORG O O O O O O O I-MISC O O I-MISC O O O O O I-LOC O O I-PER O I-ORG O
PRED: O O O I-ORG O O O O O O O I-MISC O O I-MISC O O O O O I-LOC O O I-PER O I-ORG O
SENT: But he said a power struggle in Russia 's ruling <unk> could not be ruled out , whic
h could <unk> further <unk> .
TRUE: O O O O O O O I-LOC O O O O O O O O O O O O O O O O
PRED: O O O O O O O I-LOC O O O O O O O O O O O O O O O O
SENT: <unk> <unk> <unk> ( Netherlands ) <unk> <unk>
TRUE: O I-PER I-PER O I-LOC O I-ORG O
PRED: O I-PER I-PER O I-LOC O O O
SENT: Halftime <unk> .
TRUE: O O O
PRED: O O O
--- EPOCH 7 ---
```

```
Avg loss over last 500 updates: 1.3779640918101819
Avg loss over last 500 updates: 1.5834947387945444
Avg loss over last 500 updates: 1.2839705977744806
Avg loss over last 500 updates: 1.4413288185103965
Avg loss over last 500 updates: 1.2351025580366757
Avg loss over last 500 updates: 1.4268885190182325
Avg evaluation loss: 3.0831734830603508
processed 11170 tokens with 1231 phrases; found: 1162 phrases; correct: 828.
accuracy:  73.02%; (non-O)
accuracy:  93.95%; precision:  71.26%; recall:  67.26%; FB1:  69.20
              LOC: precision:  86.03%; recall:  74.66%; FB1:  79.94  315
             MISC: precision:  75.00%; recall:  56.25%; FB1:  64.29  144
              ORG: precision:  57.10%; recall:  60.26%; FB1:  58.64  324
              PER: precision:  69.66%; recall:  71.54%; FB1:  70.59  379
(71.25645438898451, 67.26238830219334, 69.20183869619724)

5 random evaluation samples:
SENT: With their fifth straight win , the Dodgers moved a <unk> ahead of the <unk> at the
top of the <unk> <unk> <unk> behind <unk> ( <unk> ) , who allowed six hits and walked four
with six <unk> .
TRUE: O O O O O O O I-ORG O O O O O O I-ORG O O O O O O O O O I-PER O O O O O O O O O O O
O O O O
PRED: O O O O O O O I-ORG O O O O O O O O O O O O I-ORG I-ORG I-ORG O O O I-ORG O O O O O
O O O O O O O O
SENT: CHICAGO 0000 0000 .526 0000
TRUE: I-ORG O O O O
PRED: I-ORG O O O O
SENT: " <unk> seen <unk> <unk> his goals ...
TRUE: O O O I-PER O O O O
PRED: O O O O O O O O
SENT: Canada 's <unk> <unk> finished second in his Williams
TRUE: I-LOC O I-PER I-PER O O O O I-ORG
PRED: I-LOC O I-PER I-PER O O O O O
SENT: NEW YORK 1996-12-07
TRUE: I-LOC I-LOC O
PRED: I-LOC I-LOC O
--- EPOCH 8 ---
Avg loss over last 500 updates: 1.357581618527333
Avg loss over last 500 updates: 1.4985661709114633
Avg loss over last 500 updates: 1.2745562377289934
Avg loss over last 500 updates: 1.2496031929099594
Avg loss over last 500 updates: 1.154334063734751
Avg loss over last 500 updates: 1.3127732463782726
Avg evaluation loss: 3.1528406326202454
processed 11170 tokens with 1231 phrases; found: 1171 phrases; correct: 841.
accuracy:  73.07%; (non-O)
accuracy:  93.93%; precision:  71.82%; recall:  68.32%; FB1:  70.02
              LOC: precision:  84.73%; recall:  77.96%; FB1:  81.21  334
             MISC: precision:  75.68%; recall:  58.33%; FB1:  65.88  148
              ORG: precision:  57.59%; recall:  60.59%; FB1:  59.05  323
              PER: precision:  71.04%; recall:  70.46%; FB1:  70.75  366
(71.81895815542272, 68.31844029244517, 70.02497918401332)

5 random evaluation samples:
SENT: -DOCSTART-
TRUE: O
PRED: O
SENT: <unk> <unk> <unk> of South Africa defeated Tim <unk> of Britain 6-4 6-4 after a <unk
> evening rain <unk> and <unk> Thomas <unk> of Sweden won his <unk> match , <unk> Petr <un
k> of the Czech Republic 6-3 6-4 .
TRUE: O I-PER I-PER O I-LOC I-LOC O I-PER I-PER O I-LOC O O O O O O O O O O I-PER I-PER O
I-LOC O O O O O O I-PER I-PER O O I-LOC I-LOC O O O
PRED: O O O O I-LOC I-LOC O I-PER I-PER O I-LOC O O O O O O O O O O I-PER I-PER O I-LOC O
O O O O O I-PER O O O I-LOC I-LOC O O O
SENT: ATLANTA AT PITTSBURGH
TRUE: I-ORG O I-LOC
```

```
PRED: I-ORG O I-LOC
SENT: Suu Kyi , who <unk> a campaign for sanctions on <unk> 's government , was under hous
e arrest for six years without being tried before being released in July 0000 .
TRUE: I-PER I-PER O O O O O O O O I-LOC O O O O O O O O O O O O O O O O O O O O
PRED: I-PER I-PER O O O O O O O O I-ORG O O O O O O O O O O O O O O O O O O O O
SENT: <unk> <unk>
TRUE: I-PER I-PER
PRED: I-PER O
--- EPOCH 9 ---
Avg loss over last 500 updates: 1.1769165323205653
Avg loss over last 500 updates: 1.37241407262937
Avg loss over last 500 updates: 1.1872848890557424
Avg loss over last 500 updates: 1.062428315886572
Avg loss over last 500 updates: 1.0239464682876303
Avg loss over last 500 updates: 1.1727701357061724
Avg evaluation loss: 3.2805274964013518
processed 11170 tokens with 1231 phrases; found: 1148 phrases; correct: 831.
accuracy:  73.35%; (non-O)
accuracy:  94.14%; precision:  72.39%; recall:  67.51%; FB1:  69.86
              LOC: precision:  85.71%; recall:  77.69%; FB1:  81.50  329
             MISC: precision:  78.42%; recall:  56.77%; FB1:  65.86  139
              ORG: precision:  59.74%; recall:  58.96%; FB1:  59.34  303
              PER: precision:  68.70%; recall:  70.19%; FB1:  69.44  377
(72.38675958188153, 67.50609260763608, 69.86128625472887)

5 random evaluation samples:
SENT: Reading 0000 0000 0000 0000 0000 0000 0000
TRUE: I-ORG O O O O O O O
PRED: I-ORG O O O O O O O
SENT: The cars will be <unk> <unk> by <unk> and <unk> <unk> <unk> <unk> , <unk> by a son o
f President <unk> , which plans next year to start <unk> the <unk> in Indonesia .
TRUE: O O O O O O O I-ORG O I-ORG I-ORG I-ORG I-ORG O O O O O O O I-PER O O O O O O O O O
O O I-LOC O
PRED: O O O O O O O O O O O I-PER O O O O O O O O I-PER O O O O O O O O O O O I-LOC O
SENT: SOCCER - <unk> <unk> <unk> <unk> <unk> <unk> .
TRUE: O O I-PER O I-ORG O O O O
PRED: O O O O O O O O O
SENT: <unk> <unk> 0000 0000 0000 0000 0000 0000 0000
TRUE: I-ORG I-ORG O O O O O O O
PRED: I-ORG I-ORG O O O O O O O
SENT: <unk> said <unk> the meeting carried <unk> for Singapore , " however , this is <unk>
<unk> as the <unk> <unk> may not <unk> lead to any additional investment and trade <unk> t
o this region . "
TRUE: I-PER O O O O O O O I-LOC O O O O O O O O O O O O O O O O O O O O O O O O O O O O
PRED: I-PER O O O O O O O I-LOC O O O O O O O O O O O O O O O O O O O O O O O O O O O O
--- EPOCH 10 ---
Avg loss over last 500 updates: 1.126772257849591
Avg loss over last 500 updates: 1.3339204299108196
Avg loss over last 500 updates: 1.1484673207211082
Avg loss over last 500 updates: 1.0757857147213168
Avg loss over last 500 updates: 0.9359214804280749
Avg loss over last 500 updates: 1.0518884640757156
Avg evaluation loss: 3.441275830005336
processed 11170 tokens with 1231 phrases; found: 1122 phrases; correct: 820.
accuracy:  72.41%; (non-O)
accuracy:  94.19%; precision:  73.08%; recall:  66.61%; FB1:  69.70
              LOC: precision:  84.88%; recall:  75.76%; FB1:  80.06  324
             MISC: precision:  78.63%; recall:  53.65%; FB1:  63.78  131
              ORG: precision:  62.72%; recall:  58.63%; FB1:  60.61  287
              PER: precision:  68.95%; recall:  71.00%; FB1:  69.96  380
(73.0837789661319, 66.61251015434605, 69.69825754356141)

5 random evaluation samples:
SENT: <unk> of French first division
TRUE: O O I-MISC O O
PRED: O O I-MISC O O
```

```
SENT: PRESS DIGEST - <unk> - AUG 0000 .
TRUE: O O O I-LOC O O O O
PRED: O O O O O O O O
SENT: 6. <unk> <unk> ( Belgium ) Honda
TRUE: O I-PER I-PER O I-LOC O I-ORG
PRED: O I-PER I-PER O I-LOC O I-ORG
SENT: <unk> government newspapers have <unk> Britain for allowing Islamists , whom they <u
nk> as " <unk> " , to hold their conference , saying the meeting will be a chance for dang
erous Moslem <unk> to <unk> against their countries of <unk> .
TRUE: I-MISC O O O O I-LOC O O I-MISC O O O O O O O O O O O O O O O O O O O O O O O I-MISC
O O O O O O O O O
PRED: O O O O O I-LOC O O I-MISC O O O O O O O O O O O O O O O O O O O O O O O I-MISC O O
O O O O O O O
SENT: Scottish Cup first round
TRUE: I-MISC I-MISC O O
PRED: I-MISC I-MISC O O
--- EPOCH 11 ---
Avg loss over last 500 updates: 0.9937420989043426
Avg loss over last 500 updates: 1.085926796061788
Avg loss over last 500 updates: 1.0147756261950014
Avg loss over last 500 updates: 0.9733271421422574
Avg loss over last 500 updates: 0.913839882302878
Avg loss over last 500 updates: 1.1085896791702925
Avg evaluation loss: 3.3456211852904274
processed 11170 tokens with 1231 phrases; found: 1130 phrases; correct: 832.
accuracy:  73.51%; (non-O)
accuracy:  94.25%; precision:  73.63%; recall:  67.59%; FB1:  70.48
              LOC: precision:  85.98%; recall:  76.03%; FB1:  80.70  321
             MISC: precision:  73.24%; recall:  54.17%; FB1:  62.28  142
              ORG: precision:  63.96%; recall:  58.96%; FB1:  61.36  283
              PER: precision:  70.57%; recall:  73.44%; FB1:  71.98  384
(73.6283185840708, 67.58732737611697, 70.47861075815331)

5 random evaluation samples:
SENT: Hapoel <unk> 0000 Maccabi Tel Aviv 0000
TRUE: I-ORG I-ORG O O I-ORG I-ORG O
PRED: I-ORG I-ORG O I-ORG I-ORG I-ORG O
SENT: - Pakistan <unk> <unk> tonnes of <unk> <unk> yellow <unk> from <unk> Inc for $ <unk>
per tonne , <unk> U.S. Gulf , agents for the <unk> said .
TRUE: O I-LOC O O O O O O O O O I-ORG I-ORG O O O O O O O I-LOC I-LOC O O O O O O O
PRED: O I-LOC O O O O I-ORG O O O O I-ORG I-ORG O O O O O O O I-LOC I-LOC O O O O O O O
SENT: A human rights <unk> said on Wednesday he had been released after more than two week
s in <unk> that followed his call for an <unk> into the death of a Gaza man <unk> by Pales
tinian police .
TRUE: O O O O O O O O O O O O O O O O O O O O O O O O O O O O O O I-LOC O O O I-MISC O O
PRED: O O O O O O O O O O O O O O O O O O O O O O O O O O O O O O I-LOC O O O I-MISC O O
SENT: A <unk> <unk> in U.S. <unk> after a <unk> rise in the Chicago <unk> pulled <unk> low
er , but traders said the market was <unk> <unk> ahead of August <unk> data and the <unk>
<unk> due on Monday .
TRUE: O O O O I-ORG I-ORG O O O O O O I-MISC I-MISC O O O O O O O O O O O O O O O O O O O
O O O O O O
PRED: O O O O I-LOC O O O O O O O I-LOC O O O O O O O O O O O O O O O O O O O O O O O O O
O
SENT: No arrests had been made , a police spokesman said .
TRUE: O O O O O O O O O O O
PRED: O O O O O O O O O O O
--- EPOCH 12 ---
Avg loss over last 500 updates: 0.891786853779171
Avg loss over last 500 updates: 1.015059345187378
Avg loss over last 500 updates: 1.0023076614066495
Avg loss over last 500 updates: 0.931654472417199
Avg loss over last 500 updates: 0.787670547499122
Avg loss over last 500 updates: 0.9972805048931024
Avg evaluation loss: 3.4089158575289127
processed 11170 tokens with 1231 phrases; found: 1212 phrases; correct: 842.
accuracy:  74.78%; (non-O)
```

```
accuracy:  93.85%; precision:  69.47%; recall:  68.40%; FB1:  68.93
              LOC: precision:  85.71%; recall:  77.69%; FB1:  81.50  329
             MISC: precision:  67.72%; recall:  55.73%; FB1:  61.14  158
              ORG: precision:  56.48%; recall:  59.61%; FB1:  58.00  324
              PER: precision:  67.33%; recall:  73.17%; FB1:  70.13  401
(69.47194719471948, 68.3996750609260, 68.9316414244781)

5 random evaluation samples:
SENT: 1. <unk> <unk> ( U.S. ) 0000 minutes <unk> seconds
TRUE: O I-PER I-PER O I-LOC O O O O O
PRED: O I-PER I-PER O I-LOC O O O O O
SENT: Bank of New Zealand said on Thursday it was <unk> its <unk> home lending rates .
TRUE: I-ORG I-ORG I-ORG I-ORG O O O O O O O O O O O O
PRED: I-ORG I-ORG I-LOC I-LOC O O O O O O O O O O O O
SENT: A government statement , <unk> <unk> by state radio , said the two days of <unk> wer
e " for the dead , for peace and <unk> in Guinea , the victory of the new government and t
he health of the head of state " .
TRUE: O O O O O O O O O O O O O O O O O O O O O O O O O O O I-LOC O O O O O O O O O O O O
O O O O O
PRED: O O O O O O O O O O O O O O O O O O O O O O O O O O O I-LOC O O O O O O O O O O O O
O O O O O
SENT: When we <unk> the shares we <unk> ...
TRUE: O O O O O O O O
PRED: O O O O O O O O
SENT: 6. Michael <unk> ( U.S. ) <unk>
TRUE: O I-PER I-PER O I-LOC O O
PRED: O I-PER I-PER O I-LOC O O
--- EPOCH 13 ---
Avg loss over last 500 updates: 0.8148250419716885
Avg loss over last 500 updates: 1.009307216465378
Avg loss over last 500 updates: 0.867149907768798
Avg loss over last 500 updates: 0.918188081821459
Avg loss over last 500 updates: 0.761379256395828
Avg loss over last 500 updates: 0.9401957673084848
Avg evaluation loss: 3.6750318801613546
processed 11170 tokens with 1231 phrases; found: 1178 phrases; correct: 833.
accuracy:  73.90%; (non-O)
accuracy:  94.02%; precision:  70.71%; recall:  67.67%; FB1:  69.16
              LOC: precision:  87.46%; recall:  76.86%; FB1:  81.82  319
             MISC: precision:  73.68%; recall:  58.33%; FB1:  65.12  152
              ORG: precision:  58.36%; recall:  57.98%; FB1:  58.17  305
              PER: precision:  65.67%; recall:  71.54%; FB1:  68.48  402
(70.71307300509338, 67.66856214459789, 69.15732669157326)

5 random evaluation samples:
SENT: Bank of New Zealand said on Thursday it was <unk> its <unk> home lending rates .
TRUE: I-ORG I-ORG I-ORG I-ORG O O O O O O O O O O O O
PRED: I-ORG I-ORG I-LOC I-LOC O O O O O O O O O O O O
SENT: Officials said the <unk> vote removed all existing <unk> hurdles in the way of <unk>
's <unk> .
TRUE: O O O I-ORG O O O O O O O O O O I-ORG O O O
PRED: O O O O O O O O O O O O O O O O O O
SENT: <unk> Bremen 0000 <unk> <unk> 0000
TRUE: I-ORG I-ORG O I-ORG I-ORG O
PRED: I-ORG I-ORG O I-ORG I-ORG O
SENT: Pace <unk> three senior finalists -- <unk> <unk> defensive end <unk> Brown , <unk> S
tate <unk> tackle Juan <unk> and defensive end <unk> <unk> of <unk> .
TRUE: I-PER O O O O O I-ORG I-ORG O O I-PER I-PER O I-ORG I-ORG O O I-PER I-PER O O O I-PE
R I-PER O I-ORG O
PRED: I-PER O O O O O I-PER I-PER O O O I-PER O I-ORG O O O I-PER I-PER O O O O O O O O
SENT: <unk> made <unk> in the <unk> minute when he headed in a <unk> to score the <unk> .
TRUE: I-PER O O O O O O O O O O O O O O O O O
PRED: I-PER O O O O O O O O O O O O O O O O O
--- EPOCH 14 ---
Avg loss over last 500 updates: 0.7740641272167711
Avg loss over last 500 updates: 0.8967057252129684
```

```
Avg loss over last 500 updates: 0.8234615896720595
Avg loss over last 500 updates: 0.8462402475484294
Avg loss over last 500 updates: 0.8342506089115492
Avg loss over last 500 updates: 0.8577411838049119
Avg evaluation loss: 3.5461925919802244
processed 11170 tokens with 1231 phrases; found: 1192 phrases; correct: 853.
accuracy:  74.17%; (non-O)
accuracy:  94.13%; precision:  71.56%; recall:  69.29%; FB1:  70.41
              LOC: precision:  85.84%; recall:  78.51%; FB1:  82.01  332
             MISC: precision:  74.67%; recall:  58.33%; FB1:  65.50  150
              ORG: precision:  63.19%; recall:  59.28%; FB1:  61.18  288
              PER: precision:  64.93%; recall:  74.25%; FB1:  69.28  422
(71.56040268456377, 69.2932575142161, 70.40858439950475)

5 random evaluation samples:
SENT: <unk> Sudan plane expected at London 's Stansted .
TRUE: O I-LOC O O O I-LOC O I-LOC O
PRED: O I-LOC O O O I-LOC O I-LOC O
SENT: Net foreign currency - <unk> - <unk>
TRUE: O O O O O O O
PRED: O O O O O O I-PER
SENT: <unk> <unk> 0000 <unk> <unk> 0000
TRUE: I-ORG I-ORG O I-ORG I-ORG O
PRED: I-ORG I-ORG O I-ORG I-ORG O
SENT: <unk> AT SAN FRANCISCO
TRUE: I-ORG O I-LOC I-LOC
PRED: I-ORG O I-LOC I-LOC
SENT: They said they wanted to <unk> independent unions on university <unk> and <unk> that
details of the punishment of <unk> who allegedly <unk> some students at the October <unk>
be published in newspapers .
TRUE: O O O O O O O O O O O O O O O O O O O O O O O O O O O O O O O O O O
PRED: O O O O O O O O O O O O O O O O O O O I-LOC O O O O O O O O O O O O O O
--- EPOCH 15 ---
Avg loss over last 500 updates: 0.7322419075766545
Avg loss over last 500 updates: 0.9132976816976268
Avg loss over last 500 updates: 0.8170230439607609
Avg loss over last 500 updates: 0.8095421005745819
Avg loss over last 500 updates: 0.7308510470163437
Avg loss over last 500 updates: 0.9245831879022764
Avg evaluation loss: 3.634887895364175
processed 11170 tokens with 1231 phrases; found: 1181 phrases; correct: 859.
accuracy:  74.67%; (non-O)
accuracy:  94.23%; precision:  72.73%; recall:  69.78%; FB1:  71.23
              LOC: precision:  83.81%; recall:  81.27%; FB1:  82.52  352
             MISC: precision:  76.55%; recall:  57.81%; FB1:  65.88  145
              ORG: precision:  61.89%; recall:  61.89%; FB1:  61.89  307
              PER: precision:  69.76%; recall:  71.27%; FB1:  70.51  377
(72.73497036409822, 69.78066612510155, 71.22719734660033)

5 random evaluation samples:
SENT: German <unk> <unk> of motor <unk> jumped <unk> percent in July this year from the <u
nk> period , the Federal office for motor <unk> said on Thursday .
TRUE: I-MISC O O O O O O O O O O O O O O O O O O I-ORG I-ORG I-ORG I-ORG I-ORG O O O O
PRED: I-MISC I-MISC O O O O O O O O O O O O O O O O O I-ORG O O O I-ORG O O O O
SENT: <unk> 0000 0000 0000 0000 0000 0000 0000
TRUE: I-ORG O O O O O O O
PRED: I-ORG O O O O O O O
SENT: -DOCSTART-
TRUE: O
PRED: O
SENT: From Gencor 's <unk> we are taking the position that it is not on , " <unk> said .
TRUE: O I-ORG O O O O O O O O O O O O O O I-PER O O
PRED: O I-ORG O O O O O O O O O O O O O O I-PER O O
SENT: <unk> director <unk> Jordan says he never lost more <unk> over a film than over " Mi
chael <unk> " , his <unk> <unk> about the <unk> which has its <unk> on Saturday at the <un
k> <unk> <unk> .
```

```
TRUE: O O I-PER I-PER O O O O O O O O O O O O I-MISC I-MISC O O O O O O O I-ORG O O O O O
O O O I-MISC I-MISC I-MISC O
PRED: I-ORG O O I-PER O O O O O O O O O O O O I-PER I-PER O O O O O O O I-ORG O O O O O O
O O O O O O
--- EPOCH 16 ---
Avg loss over last 500 updates: 0.6532472014020982
Avg loss over last 500 updates: 0.9010892660501117
Avg loss over last 500 updates: 0.8033524641626472
Avg loss over last 500 updates: 0.6415689004928716
Avg loss over last 500 updates: 0.615256977588911
Avg loss over last 500 updates: 0.7365913449302555
Avg evaluation loss: 3.8705170074968653
processed 11170 tokens with 1231 phrases; found: 1194 phrases; correct: 862.
accuracy:  75.61%; (non-O)
accuracy:  94.29%; precision:  72.19%; recall:  70.02%; FB1:  71.09
              LOC: precision:  88.79%; recall:  78.51%; FB1:  83.33  321
             MISC: precision:  77.93%; recall:  58.85%; FB1:  67.06  145
              ORG: precision:  62.09%; recall:  61.89%; FB1:  61.99  306
              PER: precision:  64.93%; recall:  74.25%; FB1:  69.28  422
(72.19430485762143, 70.02437043054427, 71.09278350515464)

5 random evaluation samples:
SENT: They said they wanted to <unk> independent unions on university <unk> and <unk> that
details of the punishment of <unk> who allegedly <unk> some students at the October <unk>
be published in newspapers .
TRUE: O O O O O O O O O O O O O O O O O O O O O O O O O O O O O O O O O O
PRED: O O O O O O O O O O O O O O O O O O O I-PER O O O O O O O O O O O O O O
SENT: James is a <unk> for a <unk> act of <unk> , while those who have done nothing for th
e cause are free , " she <unk> .
TRUE: I-PER O O O O O O O O O O O O O O O O O O O O O O O O O O
PRED: I-ORG O O O O O O O O O O O O O O O O O O O O O O O O O O
SENT: But <unk> <unk> , one of the refugees who signed the letter of <unk> , told Reuters
:
TRUE: O I-PER I-PER O O O O O O O O O O O O O I-ORG O
PRED: O I-PER I-PER O O O O O O O O O O O O O I-ORG O
SENT: 2-2 ( halftime 0-0 ) in a friendly soccer international on
TRUE: O O O O O O O O O O O
PRED: O O O O O O O O O O O
SENT: Last season 's league and Cup winners Manchester United host 0000 champions Blackbur
n on Sunday .
TRUE: O O O O O I-MISC O I-ORG I-ORG O O O I-ORG O O O
PRED: O O O O O I-MISC O I-ORG I-ORG O O O I-ORG O O O
--- EPOCH 17 ---
Avg loss over last 500 updates: 0.6165716926998883
Avg loss over last 500 updates: 0.7082875396478351
Avg loss over last 500 updates: 0.7745582393392559
Avg loss over last 500 updates: 0.7264612795272467
Avg loss over last 500 updates: 0.5709444740643335
Avg loss over last 500 updates: 0.7498370851717168
Avg evaluation loss: 3.906349148949287
processed 11170 tokens with 1231 phrases; found: 1187 phrases; correct: 847.
accuracy:  74.39%; (non-O)
accuracy:  94.14%; precision:  71.36%; recall:  68.81%; FB1:  70.06
              LOC: precision:  84.82%; recall:  78.51%; FB1:  81.55  336
             MISC: precision:  71.92%; recall:  54.69%; FB1:  62.13  146
              ORG: precision:  61.67%; recall:  60.26%; FB1:  60.96  300
              PER: precision:  67.16%; recall:  73.71%; FB1:  70.28  405
(71.35636057287279, 68.80584890333063, 70.05789909015715)

5 random evaluation samples:
SENT: <unk> 0000 0000 0000 0000 0000 0000 0000
TRUE: I-ORG O O O O O O O
PRED: I-ORG O O O O O O O
SENT: * President Clinton <unk> <unk> plan to <unk> up toxic waste <unk> .
TRUE: O O I-PER O O O O O O O O O O
PRED: O O I-PER I-PER O O O O O O O O O
```

```
SENT: <unk> <unk>
TRUE: I-PER I-PER
PRED: I-PER I-PER
SENT: <unk> <unk> ISS <unk> <unk> PAY DATE <unk>
TRUE: O O O O O O O O
PRED: O O O O O O O O
SENT: Russian and rebel military <unk> finally met in Chechnya on Tuesday for delayed talk
s aimed at <unk> a ceasefire <unk> last week by President <unk> Yeltsin 's <unk> Alexander
Lebed .
TRUE: I-MISC O O O O O O O I-LOC O O O O O O O O O O O O O O O I-PER I-PER O O I-PER I-PER
O
PRED: I-MISC O O O O O O O I-LOC O O O O O O O O O O O O O O O I-PER I-PER O O I-PER I-PER
O
--- EPOCH 18 ---
Avg loss over last 500 updates: 0.7942087516420665
Avg loss over last 500 updates: 0.8223073366666718
Avg loss over last 500 updates: 0.7427986781571476
Avg loss over last 500 updates: 0.7233215832193992
Avg loss over last 500 updates: 0.6209400304548369
Avg loss over last 500 updates: 0.7522316498110118
Avg evaluation loss: 3.7720433978722823
processed 11170 tokens with 1231 phrases; found: 1189 phrases; correct: 859.
accuracy:  74.83%; (non-O)
accuracy:  94.20%; precision:  72.25%; recall:  69.78%; FB1:  70.99
              LOC: precision:  84.41%; recall:  79.06%; FB1:  81.65  340
             MISC: precision:  73.83%; recall:  57.29%; FB1:  64.52  149
              ORG: precision:  62.33%; recall:  59.28%; FB1:  60.77  292
              PER: precision:  68.63%; recall:  75.88%; FB1:  72.07  408
(72.24558452481077, 69.78066612510155, 70.99173553719007)

5 random evaluation samples:
SENT: <unk> <unk> 0000 0000 0000 0000 0000 0000 0000
TRUE: I-ORG I-ORG O O O O O O O
PRED: I-ORG I-ORG O O O O O O O
SENT: <unk> had been working for the <unk> which provides food to civilians for only a few
weeks before he was <unk> .
TRUE: I-PER O O O O O O O O O O O O O O O O O O O O O
PRED: O O O O O O O O O O O O O O O O O O O O O O
SENT: -DOCSTART-
TRUE: O
PRED: O
SENT: -DOCSTART-
TRUE: O
PRED: O
SENT: <unk> :
TRUE: O O
PRED: O O
--- EPOCH 19 ---
Avg loss over last 500 updates: 0.7102352405928112
Avg loss over last 500 updates: 0.650632812975827
Avg loss over last 500 updates: 0.692608732817354
Avg loss over last 500 updates: 0.6503745693749318
Avg loss over last 500 updates: 0.5121360552306327
Avg loss over last 500 updates: 0.697548881384634
Avg evaluation loss: 4.16100994757011
processed 11170 tokens with 1231 phrases; found: 1201 phrases; correct: 871.
accuracy:  75.33%; (non-O)
accuracy:  94.28%; precision:  72.52%; recall:  70.76%; FB1:  71.63
              LOC: precision:  83.57%; recall:  81.27%; FB1:  82.40  353
             MISC: precision:  74.83%; recall:  58.85%; FB1:  65.89  151
              ORG: precision:  63.70%; recall:  58.31%; FB1:  60.88  281
              PER: precision:  68.27%; recall:  76.96%; FB1:  72.36  416
(72.52289758534555, 70.75548334687247, 71.6282894736842)

5 random evaluation samples:
SENT: 3. <unk> <unk> ( Italy ) <unk>
```

```
TRUE: O I-PER I-PER O I-LOC O O
PRED: O I-PER I-PER O I-LOC O O
SENT: <unk> <unk> ( Belarus ) beat <unk> <unk> ( Spain ) 0000
TRUE: I-PER I-PER O I-LOC O O I-PER I-PER O I-LOC O O
PRED: I-PER I-PER O I-LOC O O I-PER I-PER O I-LOC O O
SENT: Sri Lanka
TRUE: I-LOC I-LOC
PRED: I-LOC I-LOC
SENT: The Civil War <unk> <unk> is mostly <unk> joined by students during the school year
.
TRUE: O I-ORG I-ORG I-ORG I-ORG O O O O O O O O O O O
PRED: O I-MISC I-MISC O I-PER O O O O O O O O O O O
SENT: <unk> <unk> 0000 <unk> <unk> 0000
TRUE: I-ORG I-ORG O I-ORG I-ORG O
PRED: I-ORG I-ORG O I-ORG I-ORG O
--- EPOCH 20 ---
Avg loss over last 500 updates: 0.6154825951090147
Avg loss over last 500 updates: 0.684232204180006
Avg loss over last 500 updates: 0.694574514479792
Avg loss over last 500 updates: 0.640440738757681
Avg loss over last 500 updates: 0.5266944110689162
Avg loss over last 500 updates: 0.686997306616179
Avg evaluation loss: 4.066602269197315
processed 11170 tokens with 1231 phrases; found: 1186 phrases; correct: 861.
accuracy:  75.22%; (non-O)
accuracy:  94.23%; precision:  72.60%; recall:  69.94%; FB1:  71.25
              LOC: precision:  86.49%; recall:  79.34%; FB1:  82.76  333
             MISC: precision:  73.51%; recall:  57.81%; FB1:  64.72  151
              ORG: precision:  58.13%; recall:  60.59%; FB1:  59.33  320
              PER: precision:  72.25%; recall:  74.80%; FB1:  73.50  382
(72.59696458684655, 69.94313566206336, 71.2453454695904)

5 random evaluation samples:
SENT: He is not as <unk> as he used to be be but was too good for me in the end .
TRUE: O O O O O O O O O O O O O O O O O O O O O
PRED: O O O O O O O O O O O O O O O O O O O O O
SENT: 7. <unk> 0000
TRUE: O I-ORG O
PRED: O I-ORG O
SENT: Cologne 0000 0000 0000 0000 0000 0000 0000
TRUE: I-ORG O O O O O O O
PRED: I-ORG O O O O O O O
SENT: <unk> sources said feed <unk> demand was keeping <unk> with <unk> production and dri
ving prices higher .
TRUE: I-LOC O O O O O O O O O O O O O O O O O
PRED: O O O O O O O O O O O O O O O O O O
SENT: <unk> - <unk> <unk> figures <unk> to value of orders on books at end of period .
TRUE: O O O O O O O O O O O O O O O O O O
PRED: O O O O O O O O O O O O O O O O O O
--- EPOCH 21 ---
Avg loss over last 500 updates: 0.6009327363215686
Avg loss over last 500 updates: 0.697571519225759
Avg loss over last 500 updates: 0.6478157823512293
Avg loss over last 500 updates: 0.5654472256586893
Avg loss over last 500 updates: 0.5337016331360191
Avg loss over last 500 updates: 0.6954160323206039
Avg evaluation loss: 4.221641649038902
processed 11170 tokens with 1231 phrases; found: 1214 phrases; correct: 858.
accuracy:  75.28%; (non-O)
accuracy:  94.09%; precision:  70.68%; recall:  69.70%; FB1:  70.18
              LOC: precision:  85.76%; recall:  77.96%; FB1:  81.67  330
             MISC: precision:  76.60%; recall:  56.25%; FB1:  64.86  141
              ORG: precision:  57.91%; recall:  59.61%; FB1:  58.75  316
              PER: precision:  66.51%; recall:  76.96%; FB1:  71.36  427
(70.67545304777595, 69.69943135662064, 70.1840490797546)
```

```
5 random evaluation samples:
SENT: " The Foreign Ministry is trying to find out from the Greek embassy why Albanian ref
ugees have been <unk> from Greece , " <unk> told Reuters .
TRUE: O O I-ORG I-ORG O O O O O O O I-MISC O O I-MISC O O O O O I-LOC O O I-PER O I-ORG O
PRED: O O O I-ORG O O O O O O O I-MISC O O I-MISC O O O O O I-LOC O O I-PER O I-ORG O
SENT: <unk> <unk> <unk> ( Netherlands ) <unk> <unk>
TRUE: O I-PER I-PER O I-LOC O I-ORG O
PRED: I-PER I-PER I-PER O I-LOC O O O
SENT: <unk> 1996-12-07
TRUE: I-LOC O
PRED: I-LOC O
SENT: <unk> Mushtaq b Harris 0000
TRUE: I-PER I-PER O I-PER O
PRED: I-PER I-PER O I-PER O
SENT: <unk> <unk> 1996-08-25
TRUE: I-LOC I-LOC O
PRED: I-LOC I-LOC O
--- EPOCH 22 ---
Avg loss over last 500 updates: 0.6030482251726085
Avg loss over last 500 updates: 0.5843819178212584
Avg loss over last 500 updates: 0.5746028549436517
Avg loss over last 500 updates: 0.575509608385968
Avg loss over last 500 updates: 0.4392929700463437
Avg loss over last 500 updates: 0.6482446090732723
Avg evaluation loss: 4.1000374059156055
processed 11170 tokens with 1231 phrases; found: 1203 phrases; correct: 866.
accuracy:  75.33%; (non-O)
accuracy:  94.26%; precision:  71.99%; recall:  70.35%; FB1:  71.16
              LOC: precision:  83.43%; recall:  80.44%; FB1:  81.91  350
             MISC: precision:  74.48%; recall:  56.25%; FB1:  64.09  145
              ORG: precision:  66.06%; recall:  58.96%; FB1:  62.31  274
              PER: precision:  65.67%; recall:  77.24%; FB1:  70.98  434
(71.98669991687449, 70.34930950446791, 71.15858668857847)

5 random evaluation samples:
SENT: 6. <unk> <unk> ( Netherlands ) Rabobank 0000
TRUE: O I-PER I-PER O I-LOC O I-ORG O
PRED: O I-PER I-PER O I-LOC O I-ORG O
SENT: <unk> 's autonomy was <unk> in 0000 and Serb police forces <unk> down on Albanian pr
otests .
TRUE: I-LOC O O O O O O O I-MISC O O O O O I-MISC O O
PRED: I-ORG O O O O O O O I-MISC O O O O O I-MISC O O
SENT: <unk> <unk> ( Sweden ) 0000 0000 , <unk> <unk> 0000 0000 , <unk>
TRUE: I-PER I-PER O I-LOC O O O O I-PER I-PER O O O I-PER
PRED: I-PER I-PER O I-LOC O O O O I-PER I-PER O O O I-PER
SENT: 0000 Robert <unk> 0000 0000 , David Williams 0000 0000 , Thomas <unk>
TRUE: O I-PER I-PER O O O I-PER I-PER O O O I-PER I-PER
PRED: O I-PER I-PER O O O I-PER I-PER O O O I-PER I-PER
SENT: <unk> said he knew of no plans to return the man to Cuba .
TRUE: I-PER O O O O O O O O O O O I-LOC O
PRED: I-PER O O O O O O O O O O O I-LOC O
--- EPOCH 23 ---
Avg loss over last 500 updates: 0.5599351446596423
Avg loss over last 500 updates: 0.6122765303015385
Avg loss over last 500 updates: 0.6597101653421901
Avg loss over last 500 updates: 0.5065431149324071
Avg loss over last 500 updates: 0.4721722274444788
Avg loss over last 500 updates: 0.577534871560387
Avg evaluation loss: 4.098952841927502
processed 11170 tokens with 1231 phrases; found: 1220 phrases; correct: 865.
accuracy:  75.77%; (non-O)
accuracy:  94.21%; precision:  70.90%; recall:  70.27%; FB1:  70.58
              LOC: precision:  83.43%; recall:  80.44%; FB1:  81.91  350
             MISC: precision:  77.86%; recall:  56.77%; FB1:  65.66  140
              ORG: precision:  57.72%; recall:  60.91%; FB1:  59.27  324
              PER: precision:  68.23%; recall:  75.07%; FB1:  71.48  406
```

```
(70.90163934426229, 70.26807473598701, 70.58343533251734)

5 random evaluation samples:
SENT: 0000 - <unk> <unk>
TRUE: O O O O
PRED: O O I-PER I-PER
SENT: England
TRUE: I-LOC
PRED: I-LOC
SENT: As Glenn <unk> , a <unk> Florida minister who is the victim 's father , looked on ,
<unk> was <unk> dead at <unk> a.m. <unk> ( 0000 GMT ) for the murder of <unk> <unk> .
TRUE: O I-PER I-PER O O O I-LOC O O O O O O O O O O O I-PER O O O O O O O O O I-MISC O O O
O O I-PER I-PER O
PRED: O I-PER I-PER O O O I-ORG O O O O O O O O O O O I-PER O O O O O O O O O I-MISC O O O
O O O O O
SENT: <unk> Munich 0000 0000 0000 0000 0000 0000 0000
TRUE: I-ORG I-ORG O O O O O O O
PRED: I-ORG I-ORG O O O O O O O
SENT: <unk> <unk> ( U.S. ) beat <unk> <unk> ( Netherlands ) 5-7 6-3 6-3
TRUE: I-PER I-PER O I-LOC O O I-PER I-PER O I-LOC O O O O
PRED: I-PER I-PER O I-LOC O O I-PER I-PER O I-LOC O O O O
--- EPOCH 24 ---
Avg loss over last 500 updates: 0.5085958704439554
Avg loss over last 500 updates: 0.6501610889473707
Avg loss over last 500 updates: 0.5304056844841569
Avg loss over last 500 updates: 0.5681731101949625
Avg loss over last 500 updates: 0.4932107595173675
Avg loss over last 500 updates: 0.6053295722414498
Avg evaluation loss: 4.554911585001087
processed 11170 tokens with 1231 phrases; found: 1185 phrases; correct: 855.
accuracy:  74.17%; (non-O)
accuracy:  94.20%; precision:  72.15%; recall:  69.46%; FB1:  70.78
              LOC: precision:  83.14%; recall:  80.17%; FB1:  81.63  350
             MISC: precision:  77.86%; recall:  56.77%; FB1:  65.66  140
              ORG: precision:  64.06%; recall:  58.63%; FB1:  61.22  281
              PER: precision:  66.43%; recall:  74.53%; FB1:  70.24  414
(72.15189873417721, 69.4557270511779, 70.77814569536423)

5 random evaluation samples:
SENT: ATLANTA AT PITTSBURGH
TRUE: I-ORG O I-LOC
PRED: I-ORG O I-LOC
SENT: Pires scored first with a <unk> shot in the 35th minute before <unk> again from clos
e range just before the break .
TRUE: I-PER O O O O O O O O O O O O O O O O O O O O O
PRED: I-PER O O O O O O O O O O O O O O O O O O O O O
SENT: <unk> , Belgium 1996-08-25
TRUE: I-LOC O I-LOC O
PRED: I-LOC O I-LOC O
SENT: Australia at South Africa
TRUE: I-LOC O I-LOC I-LOC
PRED: I-LOC O I-LOC I-LOC
SENT: -DOCSTART-
TRUE: O
PRED: O
--- EPOCH 25 ---
Avg loss over last 500 updates: 0.5236864984488501
Avg loss over last 500 updates: 0.655418227541318
Avg loss over last 500 updates: 0.5824074144754928
Avg loss over last 500 updates: 0.5544339723955486
Avg loss over last 500 updates: 0.4802982270438341
Avg loss over last 500 updates: 0.6096444674673303
Avg evaluation loss: 4.349667527265353
processed 11170 tokens with 1231 phrases; found: 1196 phrases; correct: 856.
accuracy:  74.28%; (non-O)
accuracy:  94.17%; precision:  71.57%; recall:  69.54%; FB1:  70.54
```

```
              LOC: precision:  83.57%; recall:  79.89%; FB1:  81.69  347
             MISC: precision:  69.43%; recall:  56.77%; FB1:  62.46  157
              ORG: precision:  60.07%; recall:  58.31%; FB1:  59.17  298
              PER: precision:  70.56%; recall:  75.34%; FB1:  72.87  394
(71.57190635451505, 69.53696181965881, 70.53976102183766)

5 random evaluation samples:
SENT: 0000 - <unk> Khan ( Pakistan ) beat Simon <unk> ( Germany ) 15-12 15-7 <unk> 15-10
TRUE: O O I-PER I-PER O I-LOC O O I-PER I-PER O I-LOC O O O O O
PRED: O O I-PER I-PER O I-LOC O O I-PER I-PER O I-LOC O O O O O
SENT: <unk> profit <unk> vs <unk>
TRUE: O O O O O
PRED: O O O O O
SENT: " There was no <unk> today to make a <unk> run , " said <unk> <unk> , the Canadian m
en 's national coach , <unk> too much new <unk> and poor <unk> .
TRUE: O O O O O O O O O O O O O O I-PER I-PER O O I-MISC O O O O O O O O O O O O O O
PRED: O O O O O O O O O O O O O O I-PER I-PER O O I-MISC O O O O O O O O O O O O O O
SENT: 4. <unk> <unk> ( Kenya ) <unk>
TRUE: O I-PER I-PER O I-LOC O O
PRED: O I-PER I-PER O I-LOC O O
SENT: " <unk> is getting stronger and stronger , and it already has <unk> of 0000 <unk> (
0000 kph ) , " said <unk> <unk> <unk> <unk> .
TRUE: O I-MISC O O O O O O O O O O O O O O O O O O O O O O O I-PER I-PER O
PRED: O O O O O O O O O O O O O O O O O O O O O O O I-PER I-PER I-PER I-PER O
--- EPOCH 26 ---
Avg loss over last 500 updates: 0.4892712708635235
Avg loss over last 500 updates: 0.5606451400599992
Avg loss over last 500 updates: 0.49115067909560084
Avg loss over last 500 updates: 0.567455058201582
Avg loss over last 500 updates: 0.45613947528275833
Avg loss over last 500 updates: 0.6365499658510251
Avg evaluation loss: 4.34903436519975
processed 11170 tokens with 1231 phrases; found: 1184 phrases; correct: 857.
accuracy:  74.56%; (non-O)
accuracy:  94.31%; precision:  72.38%; recall:  69.62%; FB1:  70.97
              LOC: precision:  80.49%; recall:  80.72%; FB1:  80.61  364
             MISC: precision:  76.09%; recall:  54.69%; FB1:  63.64  138
              ORG: precision:  63.67%; recall:  59.93%; FB1:  61.74  289
              PER: precision:  69.97%; recall:  74.53%; FB1:  72.18  393
(72.38175675675676, 69.61819658813972, 70.97308488612836)

5 random evaluation samples:
SENT: 2. <unk> Sang ( Kenya ) <unk>
TRUE: O I-PER I-PER O I-LOC O O
PRED: O I-PER I-PER O I-LOC O O
SENT: Zimbabwe on Monday :
TRUE: I-LOC O O O
PRED: I-LOC O O O
SENT: <unk> <unk> 0000 <unk> Munich 0000 ( <unk> <unk> , <unk> <unk> and
TRUE: I-ORG I-ORG O I-ORG I-ORG O O I-PER O O I-PER O O
PRED: I-ORG I-ORG O I-ORG I-ORG O O I-PER I-PER O I-PER I-PER O
SENT: The teams meet each other once in each
TRUE: O O O O O O O O
PRED: O O O O O O O O
SENT: The defeat put the <unk> out of the <unk> Cup .
TRUE: O O O O I-MISC O O O I-MISC I-MISC O
PRED: O O O O O O O O I-MISC I-MISC O
--- EPOCH 27 ---
Avg loss over last 500 updates: 0.4986109310818142
Avg loss over last 500 updates: 0.5406353076365065
Avg loss over last 500 updates: 0.5493270318661887
Avg loss over last 500 updates: 0.5004222199366561
Avg loss over last 500 updates: 0.433571810602062
Avg loss over last 500 updates: 0.6027956119507611
Avg evaluation loss: 4.683567622896375
processed 11170 tokens with 1231 phrases; found: 1178 phrases; correct: 848.
```

```
accuracy:  73.90%; (non-O)
accuracy:  94.14%; precision:  71.99%; recall:  68.89%; FB1:  70.40
              LOC: precision:  85.29%; recall:  79.89%; FB1:  82.50  340
             MISC: precision:  78.99%; recall:  56.77%; FB1:  66.06  138
              ORG: precision:  61.72%; recall:  58.31%; FB1:  59.97  290
              PER: precision:  65.85%; recall:  73.17%; FB1:  69.32  410
(71.98641765704585, 68.88708367181154, 70.40265670402657)

5 random evaluation samples:
SENT: <unk> 1996-08-26
TRUE: I-LOC O
PRED: I-LOC O
SENT: - Pakistan <unk> <unk> tonnes of <unk> <unk> yellow <unk> from <unk> Inc for $ <unk>
per tonne , <unk> U.S. Gulf , agents for the <unk> said .
TRUE: O I-LOC O O O O O O O O O I-ORG I-ORG O O O O O O O I-LOC I-LOC O O O O O O O
PRED: O I-LOC O O O O O O O O O I-ORG I-ORG O O O O O O O I-LOC I-LOC O O O O O O O
SENT: <unk> 0000 0000 0000 0000 0000 0000 0000
TRUE: I-ORG O O O O O O O
PRED: I-ORG O O O O O O O
SENT: India fishermen say forced to carry Tamil refugees .
TRUE: I-LOC O O O O O I-MISC O O
PRED: I-LOC O O O O O I-MISC O O
SENT: -DOCSTART-
TRUE: O
PRED: O
--- EPOCH 28 ---
Avg loss over last 500 updates: 0.5034467078847457
Avg loss over last 500 updates: 0.546407786855606
Avg loss over last 500 updates: 0.5520131686133827
Avg loss over last 500 updates: 0.5029821045477086
Avg loss over last 500 updates: 0.43543156519198895
Avg loss over last 500 updates: 0.5134126868891331
Avg evaluation loss: 4.6581245195716825
processed 11170 tokens with 1231 phrases; found: 1201 phrases; correct: 855.
accuracy:  75.06%; (non-O)
accuracy:  94.21%; precision:  71.19%; recall:  69.46%; FB1:  70.31
              LOC: precision:  85.38%; recall:  80.44%; FB1:  82.84  342
             MISC: precision:  78.99%; recall:  56.77%; FB1:  66.06  138
              ORG: precision:  59.93%; recall:  59.93%; FB1:  59.93  307
              PER: precision:  65.22%; recall:  73.17%; FB1:  68.97  414
(71.19067443796835, 69.4557270511779, 70.3125)

5 random evaluation samples:
SENT: Extras ( <unk> <unk> <unk> ) 0000
TRUE: O O O O O O O
PRED: O O O O O O O
SENT: 4. <unk> <unk> ( Kenya ) <unk>
TRUE: O I-PER I-PER O I-LOC O O
PRED: O I-PER I-PER O I-LOC O O
SENT: <unk> is <unk> by the level of competition in the <unk> .
TRUE: I-PER O O O O O O O O O O O
PRED: O O O O O O O O O O O O
SENT: 3. <unk> <unk> ( Germany ) <unk>
TRUE: O I-PER I-PER O I-LOC O O
PRED: O I-PER I-PER O I-LOC O O
SENT: France on Friday expelled another African man seized in a police <unk> on a Paris ch
urch as about 0000 Air France workers <unk> " <unk> of <unk> " used to fly illegal <unk> h
ome .
TRUE: I-LOC O O O O I-MISC O O O O O O O O I-LOC O O O O I-ORG I-ORG O O O O O O O O O O
O O O
PRED: I-LOC O O O O I-MISC O O O O O O O O I-LOC O O O O O O I-LOC O I-MISC O I-PER O O O O
O O O O O O
--- EPOCH 29 ---
Avg loss over last 500 updates: 0.42709097764811577
Avg loss over last 500 updates: 0.5115384472019388
Avg loss over last 500 updates: 0.5003525798341107
```

```
Avg loss over last 500 updates: 0.47634904861663413
Avg loss over last 500 updates: 0.4095727051005166
Avg loss over last 500 updates: 0.45489802008140456
Avg evaluation loss: 4.690180216010472
processed 11170 tokens with 1231 phrases; found: 1224 phrases; correct: 869.
accuracy:  75.11%; (non-O)
accuracy:  94.25%; precision:  71.00%; recall:  70.59%; FB1:  70.79
              LOC: precision:  85.22%; recall:  80.99%; FB1:  83.05  345
             MISC: precision:  67.07%; recall:  58.33%; FB1:  62.40  167
              ORG: precision:  62.50%; recall:  58.63%; FB1:  60.50  288
              PER: precision:  66.75%; recall:  76.69%; FB1:  71.37  424
(70.99673202614379, 70.59301380991064, 70.79429735234216)

5 random evaluation samples:
SENT: A <unk> <unk> in U.S. <unk> after a <unk> rise in the Chicago <unk> pulled <unk> low
er , but traders said the market was <unk> <unk> ahead of August <unk> data and the <unk>
<unk> due on Monday .
TRUE: O O O O I-ORG I-ORG O O O O O O I-MISC I-MISC O O O O O O O O O O O O O O O O O O O
O O O O O O
PRED: O O O O I-LOC O O O O O O O I-LOC O O O O O O O O O O O O O O O O O O O O O O O O
O
SENT: The <unk> <unk> of Nigeria ( <unk> ) quoted police spokesman <unk> <unk> as saying t
he six were killed on Wednesday .
TRUE: O I-ORG I-ORG I-ORG I-ORG O I-ORG O O O O I-PER I-PER O O O O O O O O O
PRED: O I-ORG O O I-LOC O O O O O O I-PER I-PER O O O O O O O O O
SENT: <unk> health began to <unk> in 0000 when she was <unk> with a heart <unk> .
TRUE: O O O O O O O O O O O O O O O O
PRED: O O O O O O O O O O O O O O O O
SENT: <unk> 1996-12-07
TRUE: I-LOC O
PRED: I-LOC O
SENT: <unk> times set on Friday
TRUE: O O O O O
PRED: O O O O O
Wall time: 4min 33s
```

In [61]:

```
df = df.drop(columns = ['ORG'])
```

In [1]:

```
crf = {'0': 50.12, '1': 62.19, '2': 67.02, '3': 69.45, '4': 69.15, '5': 69.22, '6': 70.49,
       '10': 71.76, '11': 71.36, '12': 71.29, '13': 70.69, '14': 70.02, '15': 71.41, '16':
       '20': 70.76, '21': 71.00, '22': 71.06, '23': 72.02, '24': 70.27, '25':72.10, '26':
```

In [2]:

```
original = {'0': 47.89, '1':58.97, '2':63.05, '3':65.77, '4':67.88, '5':67.84, '6':69.85,
            '10':69.70, '11':70.48, '12':68.93, '13':69.16, '14':70.41, '15':71.23, '16':
            '20':71.25, '21':70.18, '22':71.16, '23':70.58, '24':70.78, '25':70.54, '26':
```

In [3]:

```
import pandas as pd
```

In [4]:

```
table1 = pd.DataFrame.from_dict(crf, orient='index', columns = ['crf_FB1'])
table2 = pd.DataFrame.from_dict(original, orient='index', columns = ['original_FB1'])
```

In [5]:

```
table1['original_FB1'] = table2['original_FB1']
```

In [6]:

```
table1
```

Out[6]:

	crf_FB1	original_FB1

	crf_FB1	original_FB1
0	50.12	47.89
1	62.19	58.97
2	67.02	63.05
3	69.45	65.77
4	69.15	67.88
5	69.22	67.84
6	70.49	69.85
7	71.97	69.20
8	70.97	70.02
9	70.64	69.86
10	71.76	69.70
11	71.36	70.48
12	71.29	68.93
13	70.69	69.16
14	70.02	70.41
15	71.41	71.23
16	72.16	71.09
17	71.20	70.06
18	71.14	70.99
19	71.05	71.63
20	70.76	71.25
21	71.00	70.18
22	71.06	71.16
23	72.02	70.58
24	70.27	70.78
25	72.10	70.54
26	69.52	70.97
27	71.35	70.40
28	70.53	70.31
29	72.04	70.79

In [7]:

```
table1.sort_values('crf_FB1', ascending = False)
```

Out[7]:

	crf_FB1	original_FB1
16	72.16	71.09
25	72.10	70.54
29	72.04	70.79

	crf_FB1	original_FB1
23	72.02	70.58
7	71.97	69.20
10	71.76	69.70
15	71.41	71.23
11	71.36	70.48
27	71.35	70.40
12	71.29	68.93
17	71.20	70.06
18	71.14	70.99
22	71.06	71.16
19	71.05	71.63
21	71.00	70.18
8	70.97	70.02
20	70.76	71.25
13	70.69	69.16
9	70.64	69.86
28	70.53	70.31
6	70.49	69.85
24	70.27	70.78
14	70.02	70.41
26	69.52	70.97
3	69.45	65.77
5	69.22	67.84
4	69.15	67.88
2	67.02	63.05
1	62.19	58.97
0	50.12	47.89

In [8]:

```
        import matplotlib.pyplot as plt
```

In [9]:

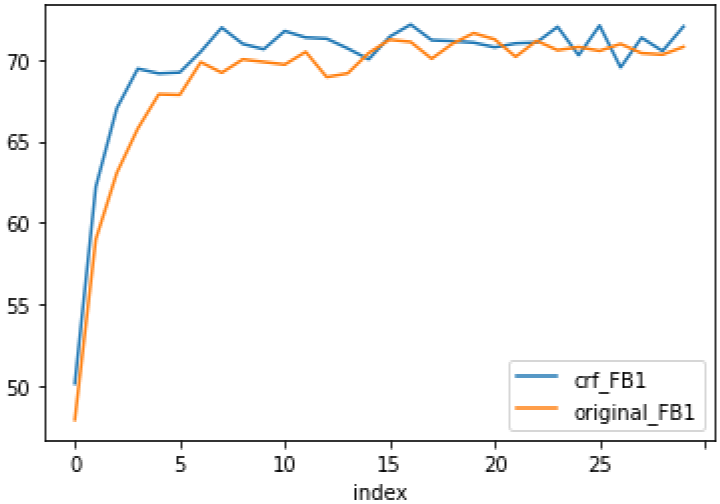
```
table1 = table1.reset_index()
```

In [10]:

```
table1.plot(x = 'index', y = ['crf_FB1', 'original_FB1'])
```

Out[10]:

```
<AxesSubplot:xlabel='index'>
```

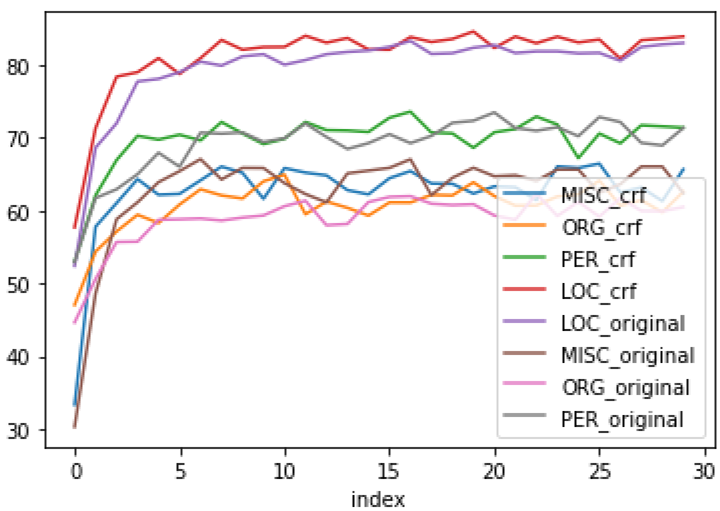


In [65]:

```
df.plot(x = 'index', y = ['MISC_crf', 'ORG_crf', 'PER_crf', 'LOC_crf', 'LOC_original', 'MI
```

Out[65]:

```
<AxesSubplot:xlabel='index'>
```



In [64]:

```
df
```

Out[64]:

	index	MISC_crf	ORG_crf	PER_crf	LOC_crf	LOC_original	MISC_original	ORG_original	PER_original
0	0	33.33	47.04	53.00	57.71	52.40	30.28	44.63	52.93
1	1	57.81	54.38	62.11	71.34	68.66	48.63	50.61	61.73
2	2	61.04	57.14	66.95	78.42	72.01	58.82	55.69	62.94
3	3	64.33	59.44	70.29	79.03	77.76	61.15	55.74	65.01
4	4	62.18	58.30	69.77	80.96	78.12	63.92	58.75	67.96
5	5	62.31	60.84	70.44	78.79	79.04	65.43	58.82	66.04
6	6	64.17	62.94	69.65	80.96	80.48	67.08	58.92	70.72
7	7	66.06	62.10	72.18	83.46	79.94	64.29	58.64	70.59
8	8	65.20	61.67	70.62	82.13	81.21	65.88	59.05	70.75
9	9	61.63	64.03	69.15	82.49	81.50	65.86	59.34	69.44
10	10	65.85	64.95	69.89	82.51	80.06	63.78	60.61	69.96
11	11	65.23	59.51	72.19	84.04	80.70	62.28	61.36	71.98

	index	MISC_crf	ORG_crf	PER_crf	LOC_crf	LOC_original	MISC_original	ORG_original	PER_original
12	12	64.85	61.25	71.07	83.09	81.50	61.14	58.00	70.13
13	13	62.82	60.39	71.01	83.73	81.82	65.12	58.17	68.48
14	14	62.24	59.32	70.82	82.15	82.01	65.50	61.18	69.28
15	15	64.46	61.13	72.77	82.13	82.52	65.88	61.89	70.51
16	16	65.44	61.12	73.61	83.87	83.33	67.06	61.99	69.28
17	17	63.74	62.17	70.73	83.19	81.55	62.13	60.96	70.28
18	18	63.69	62.10	70.62	83.60	81.65	64.52	60.77	72.07
19	19	62.35	63.89	68.65	84.63	82.40	65.89	60.88	72.36
20	20	63.34	62.00	70.77	82.37	82.76	64.72	59.33	73.50
21	21	63.22	60.69	71.23	83.94	81.67	64.86	58.75	71.36
22	22	61.45	60.72	72.94	83.02	81.91	64.09	62.31	70.98
23	23	66.08	61.95	71.84	83.92	81.91	65.66	59.27	71.48
24	24	65.88	62.23	67.22	83.12	81.63	65.66	61.22	70.24
25	25	66.47	64.13	70.60	83.53	81.69	62.46	59.17	72.87
26	26	62.50	60.69	69.26	80.90	80.61	63.64	61.74	72.18
27	27	63.19	61.44	71.74	83.43	82.50	66.06	59.97	69.32
28	28	61.32	59.78	71.58	83.67	82.84	66.06	59.93	68.97
29	29	65.68	62.61	71.41	83.94	83.05	62.40	60.50	71.37

In [ ]: