

# AADatasets\_roberta

August 11, 2022

## 0.1 Install packages

```
[ ]: !git clone https://github.com/megagonlabs/ditto
      %cd ditto
      !pip install -r requirements.txt
      !git clone https://github.com/rit-git/Snipnext_public
      %cd Snipnext_public
      !pip install -r requirements.txt
      %cd ..

      import nltk
      nltk.download('stopwords')
```

```
Cloning into 'ditto'...
remote: Enumerating objects: 291, done.
remote: Total 291 (delta 0), reused 0 (delta 0), pack-reused 291
Receiving objects: 100% (291/291), 26.87 MiB | 16.59 MiB/s, done.
Resolving deltas: 100% (142/142), done.
/content/ditto
Looking in indexes: https://pypi.org/simple, https://us-python.pkg.dev/colab-
wheels/public/simple/
Collecting gensim==3.8.1
  Downloading gensim-3.8.1-cp37-cp37m-manylinux1_x86_64.whl (24.2 MB)
    |                               | 24.2 MB 6.3 MB/s
Collecting numpy==1.19.2
  Downloading numpy-1.19.2-cp37-cp37m-manylinux2010_x86_64.whl (14.5 MB)
    |                               | 14.5 MB 61.7 MB/s
Collecting regex==2019.12.20
  Downloading regex-2019.12.20-cp37-cp37m-manylinux2010_x86_64.whl (689 kB)
    |                               | 689 kB 64.5 MB/s
Collecting scipy==1.3.2
  Downloading scipy-1.3.2-cp37-cp37m-manylinux1_x86_64.whl (25.2 MB)
    |                               | 25.2 MB 82.6 MB/s
Collecting sentencepiece==0.1.85
  Downloading sentencepiece-0.1.85-cp37-cp37m-manylinux1_x86_64.whl (1.0 MB)
    |                               | 1.0 MB 53.8 MB/s
Requirement already satisfied: sklearn==0.0 in
/usr/local/lib/python3.7/dist-packages (from -r requirements.txt (line 6)) (0.0)
```

```

Collecting spacy==3.1
  Downloading
spacy-3.1.0-cp37-cp37m-manylinux_2_17_x86_64.manylinux2014_x86_64.whl (6.4 MB)
|                                     | 6.4 MB 54.9 MB/s
ERROR: Could not find a version that satisfies the requirement
torch==1.9.0+cu111 (from versions: 1.0.0, 1.0.1, 1.0.1.post2, 1.1.0, 1.2.0,
1.3.0, 1.3.1, 1.4.0, 1.5.0, 1.5.1, 1.6.0, 1.7.0, 1.7.1, 1.8.0, 1.8.1, 1.9.0,
1.9.1, 1.10.0, 1.10.1, 1.10.2, 1.11.0, 1.12.0, 1.12.1)
ERROR: No matching distribution found for torch==1.9.0+cu111
Cloning into 'Snippext_public'...
remote: Enumerating objects: 413, done.
remote: Counting objects: 100% (22/22), done.
remote: Compressing objects: 100% (8/8), done.
remote: Total 413 (delta 15), reused 14 (delta 14), pack-reused 391
Receiving objects: 100% (413/413), 21.06 MiB | 15.70 MiB/s, done.
Resolving deltas: 100% (196/196), done.
/content/ditto/Snippext_public
Looking in indexes: https://pypi.org/simple, https://us-python.pkg.dev/colab-
wheels/public/simple/
Collecting gensim==3.8.1
  Using cached gensim-3.8.1-cp37-cp37m-manylinux1_x86_64.whl (24.2 MB)
Collecting numpy==1.19.2
  Using cached numpy-1.19.2-cp37-cp37m-manylinux2010_x86_64.whl (14.5 MB)
Collecting regex==2019.12.20
  Using cached regex-2019.12.20-cp37-cp37m-manylinux2010_x86_64.whl (689 kB)
Collecting spacy==2.2.3
  Downloading spacy-2.2.3-cp37-cp37m-manylinux1_x86_64.whl (10.4 MB)
|                                     | 10.4 MB 35.1 MB/s
Collecting sentencepiece==0.1.85
  Using cached sentencepiece-0.1.85-cp37-cp37m-manylinux1_x86_64.whl (1.0 MB)
Requirement already satisfied: sklearn==0.0 in /usr/local/lib/python3.7/dist-
packages (from -r requirements.txt (line 6)) (0.0)
Collecting tensorboardX==2.0
  Downloading tensorboardX-2.0-py2.py3-none-any.whl (195 kB)
|                                     | 195 kB 66.6 MB/s
Collecting torch==1.4.0
  Downloading torch-1.4.0-cp37-cp37m-manylinux1_x86_64.whl (753.4 MB)
|                                     | 753.4 MB 7.1 kB/s
Collecting tqdm==4.41.0
  Downloading tqdm-4.41.0-py2.py3-none-any.whl (56 kB)
|                                     | 56 kB 5.2 MB/s
Collecting transformers==3.1.0
  Downloading transformers-3.1.0-py3-none-any.whl (884 kB)
|                                     | 884 kB 68.2 MB/s
Collecting jsonlines==1.2.0
  Downloading jsonlines-1.2.0-py2.py3-none-any.whl (7.6 kB)

```

```

Collecting nltk==3.4.5
  Downloading nltk-3.4.5.zip (1.5 MB)
    | 1.5 MB 41.0 MB/s
Requirement already satisfied: six>=1.5.0 in
/usr/local/lib/python3.7/dist-packages (from gensim==3.8.1->-r requirements.txt
(line 1)) (1.15.0)
Requirement already satisfied: smart-open>=1.8.1 in
/usr/local/lib/python3.7/dist-packages (from gensim==3.8.1->-r requirements.txt
(line 1)) (5.2.1)
Requirement already satisfied: scipy>=0.18.1 in /usr/local/lib/python3.7/dist-
packages (from gensim==3.8.1->-r requirements.txt (line 1)) (1.7.3)
Collecting blis<0.5.0,>=0.4.0
  Downloading blis-0.4.1-cp37-cp37m-manylinux1_x86_64.whl (3.7 MB)
    | 3.7 MB 57.3 MB/s
Collecting catalogue<1.1.0,>=0.0.7
  Downloading catalogue-1.0.0-py2.py3-none-any.whl (7.7 kB)
Collecting thinc<7.4.0,>=7.3.0
  Downloading thinc-7.3.1-cp37-cp37m-manylinux1_x86_64.whl (2.2 MB)
    | 2.2 MB 50.9 MB/s
Requirement already satisfied: wasabi<1.1.0,>=0.4.0 in
/usr/local/lib/python3.7/dist-packages (from spacy==2.2.3->-r requirements.txt
(line 4)) (0.10.1)
Requirement already satisfied: cymem<2.1.0,>=2.0.2 in
/usr/local/lib/python3.7/dist-packages (from spacy==2.2.3->-r requirements.txt
(line 4)) (2.0.6)
Requirement already satisfied: murmurhash<1.1.0,>=0.28.0 in
/usr/local/lib/python3.7/dist-packages (from spacy==2.2.3->-r requirements.txt
(line 4)) (1.0.7)
Collecting plac<1.2.0,>=0.9.6
  Downloading plac-1.1.3-py2.py3-none-any.whl (20 kB)
Requirement already satisfied: requests<3.0.0,>=2.13.0 in
/usr/local/lib/python3.7/dist-packages (from spacy==2.2.3->-r requirements.txt
(line 4)) (2.23.0)
Requirement already satisfied: preshed<3.1.0,>=3.0.2 in
/usr/local/lib/python3.7/dist-packages (from spacy==2.2.3->-r requirements.txt
(line 4)) (3.0.6)
Collecting srsly<1.1.0,>=0.1.0
  Downloading srsly-1.0.5-cp37-cp37m-manylinux2014_x86_64.whl (184 kB)
    | 184 kB 68.8 MB/s
Requirement already satisfied: setuptools in
/usr/local/lib/python3.7/dist-packages (from spacy==2.2.3->-r requirements.txt
(line 4)) (57.4.0)
Requirement already satisfied: scikit-learn in /usr/local/lib/python3.7/dist-
packages (from sklearn==0.0->-r requirements.txt (line 6)) (1.0.2)
Requirement already satisfied: protobuf>=3.8.0 in /usr/local/lib/python3.7/dist-
packages (from tensorboardX==2.0->-r requirements.txt (line 8)) (3.17.3)
Collecting sacremoses
  Downloading sacremoses-0.0.53.tar.gz (880 kB)

```

```

| 880 kB 66.0 MB/s
Requirement already satisfied: packaging in /usr/local/lib/python3.7/dist-
packages (from transformers==3.1.0->-r requirements.txt (line 11)) (21.3)
Collecting tokenizers==0.8.1.rc2
  Downloading tokenizers-0.8.1rc2-cp37-cp37m-manylinux1_x86_64.whl (3.0 MB)
| 3.0 MB 52.5 MB/s
Requirement already satisfied: filelock in /usr/local/lib/python3.7/dist-
packages (from transformers==3.1.0->-r requirements.txt (line 11)) (3.7.1)
Requirement already satisfied: importlib-metadata>=0.20 in
/usr/local/lib/python3.7/dist-packages (from
catalogue<1.1.0,>=0.0.7->spacy==2.2.3->-r requirements.txt (line 4)) (4.12.0)
Requirement already satisfied: typing-extensions>=3.6.4 in
/usr/local/lib/python3.7/dist-packages (from importlib-
metadata>=0.20->catalogue<1.1.0,>=0.0.7->spacy==2.2.3->-r requirements.txt (line
4)) (4.1.1)
Requirement already satisfied: zipp>=0.5 in /usr/local/lib/python3.7/dist-
packages (from importlib-
metadata>=0.20->catalogue<1.1.0,>=0.0.7->spacy==2.2.3->-r requirements.txt (line
4)) (3.8.1)
Requirement already satisfied: chardet<4,>=3.0.2 in
/usr/local/lib/python3.7/dist-packages (from
requests<3.0.0,>=2.13.0->spacy==2.2.3->-r requirements.txt (line 4)) (3.0.4)
Requirement already satisfied: certifi>=2017.4.17 in
/usr/local/lib/python3.7/dist-packages (from
requests<3.0.0,>=2.13.0->spacy==2.2.3->-r requirements.txt (line 4)) (2022.6.15)
Requirement already satisfied: idna<3,>=2.5 in /usr/local/lib/python3.7/dist-
packages (from requests<3.0.0,>=2.13.0->spacy==2.2.3->-r requirements.txt (line
4)) (2.10)
Requirement already satisfied: urllib3!=1.25.0,!1.25.1,<1.26,>=1.21.1 in
/usr/local/lib/python3.7/dist-packages (from
requests<3.0.0,>=2.13.0->spacy==2.2.3->-r requirements.txt (line 4)) (1.24.3)
Requirement already satisfied: pyparsing!=3.0.5,>=2.0.2 in
/usr/local/lib/python3.7/dist-packages (from packaging->transformers==3.1.0->-r
requirements.txt (line 11)) (3.0.9)
Requirement already satisfied: click in /usr/local/lib/python3.7/dist-packages
(from sacremoses->transformers==3.1.0->-r requirements.txt (line 11)) (7.1.2)
Requirement already satisfied: joblib in /usr/local/lib/python3.7/dist-packages
(from sacremoses->transformers==3.1.0->-r requirements.txt (line 11)) (1.1.0)
Requirement already satisfied: threadpoolctl>=2.0.0 in
/usr/local/lib/python3.7/dist-packages (from scikit-learn->sklearn==0.0->-r
requirements.txt (line 6)) (3.1.0)
Building wheels for collected packages: nltk, sacremoses
  Building wheel for nltk (setup.py) ... done
  Created wheel for nltk: filename=nltk-3.4.5-py3-none-any.whl size=1449925
sha256=b8f3743bf72b9bcde5911fd29233433e2d2c82ee596c2a44ee55ef5a9124ed83
  Stored in directory: /root/.cache/pip/wheels/48/8b/7f/473521e0c731c6566d631b28
1f323842bbda9bd819eb9a3ead
  Building wheel for sacremoses (setup.py) ... done

```

```

Created wheel for sacremoses: filename=sacremoses-0.0.53-py3-none-any.whl
size=895260
sha256=b0d015fbe95bf0a66069f29c16ace6ab80c35ccf252f4359724dac74f90c3a13
Stored in directory: /root/.cache/pip/wheels/87/39/dd/a83eeef36d0bf98e7a4d1933
a4ad2d660295a40613079bafc9
Successfully built nltk sacremoses
Installing collected packages: numpy, tqdm, srsly, regex, plac, blis,
tokenizers, thinc, sentencepiece, sacremoses, catalogue, transformers, torch,
tensorboardX, spacy, nltk, jsonlines, gensim
Attempting uninstall: numpy
  Found existing installation: numpy 1.21.6
  Uninstalling numpy-1.21.6:
    Successfully uninstalled numpy-1.21.6
Attempting uninstall: tqdm
  Found existing installation: tqdm 4.64.0
  Uninstalling tqdm-4.64.0:
    Successfully uninstalled tqdm-4.64.0
Attempting uninstall: srsly
  Found existing installation: srsly 2.4.4
  Uninstalling srsly-2.4.4:
    Successfully uninstalled srsly-2.4.4
Attempting uninstall: regex
  Found existing installation: regex 2022.6.2
  Uninstalling regex-2022.6.2:
    Successfully uninstalled regex-2022.6.2
Attempting uninstall: blis
  Found existing installation: blis 0.7.8
  Uninstalling blis-0.7.8:
    Successfully uninstalled blis-0.7.8
Attempting uninstall: thinc
  Found existing installation: thinc 8.1.0
  Uninstalling thinc-8.1.0:
    Successfully uninstalled thinc-8.1.0
Attempting uninstall: catalogue
  Found existing installation: catalogue 2.0.8
  Uninstalling catalogue-2.0.8:
    Successfully uninstalled catalogue-2.0.8
Attempting uninstall: torch
  Found existing installation: torch 1.12.0+cu113
  Uninstalling torch-1.12.0+cu113:
    Successfully uninstalled torch-1.12.0+cu113
Attempting uninstall: spacy
  Found existing installation: spacy 3.4.1
  Uninstalling spacy-3.4.1:
    Successfully uninstalled spacy-3.4.1
Attempting uninstall: nltk
  Found existing installation: nltk 3.7
  Uninstalling nltk-3.7:

```

```

    Successfully uninstalled nltk-3.7
Attempting uninstall: gensim
    Found existing installation: gensim 3.6.0
    Uninstalling gensim-3.6.0:
        Successfully uninstalled gensim-3.6.0
ERROR: pip's dependency resolver does not currently take into account all
the packages that are installed. This behaviour is the source of the following
dependency conflicts.

xarray-einstats 0.2.2 requires numpy>=1.21, but you have numpy 1.19.2 which is
incompatible.

torchvision 0.13.0+cu113 requires torch==1.12.0, but you have torch 1.4.0 which
is incompatible.

torchtext 0.13.0 requires torch==1.12.0, but you have torch 1.4.0 which is
incompatible.

torchaudio 0.12.0+cu113 requires torch==1.12.0, but you have torch 1.4.0 which
is incompatible.

tensorflow 2.8.2+zzzcolab20220719082949 requires numpy>=1.20, but you have numpy
1.19.2 which is incompatible.

panel 0.12.1 requires tqdm>=4.48.0, but you have tqdm 4.41.0 which is
incompatible.

fastai 2.7.7 requires torch<1.13,>=1.7, but you have torch 1.4.0 which is
incompatible.

en-core-web-sm 3.4.0 requires spacy<3.5.0,>=3.4.0, but you have spacy 2.2.3
which is incompatible.

cmdstanpy 1.0.4 requires numpy>=1.21, but you have numpy 1.19.2 which is
incompatible.
Successfully installed blis-0.4.1 catalogue-1.0.0 gensim-3.8.1 jsonlines-1.2.0
nltk-3.4.5 numpy-1.19.2 plac-1.1.3 regex-2019.12.20 sacremoses-0.0.53
sentencepiece-0.1.85 spacy-2.2.3 srsly-1.0.5 tensorboardX-2.0 thinc-7.3.1
tokenizers-0.8.1rc2 torch-1.4.0 tqdm-4.41.0 transformers-3.1.0
/content/ditto

[nltk_data] Downloading package stopwords to /root/nltk_data...
[nltk_data]   Unzipping corpora/stopwords.zip.

[ ]: True

```

## 0.2 Install fp16 optimization

```
[ ]: !git clone https://github.com/NVIDIA/apex
      %cd apex
      !pip install -v --no-cache-dir ./
      %cd ..
```

```
Cloning into 'apex'...
remote: Enumerating objects: 10091, done.
remote: Counting objects: 100% (207/207), done.
remote: Compressing objects: 100% (145/145), done.
remote: Total 10091 (delta 100), reused 130 (delta 58), pack-reused 9884
Receiving objects: 100% (10091/10091), 14.95 MiB | 6.86 MiB/s, done.
Resolving deltas: 100% (6905/6905), done.
/content/ditto/apex
Using pip 21.1.3 from /usr/local/lib/python3.7/dist-packages/pip (python 3.7)
Value for scheme.platlib does not match. Please report this to
<https://github.com/pypa/pip/issues/9617>
distutils: /usr/local/lib/python3.7/dist-packages
sysconfig: /usr/lib/python3.7/site-packages
Value for scheme.purelib does not match. Please report this to
<https://github.com/pypa/pip/issues/9617>
distutils: /usr/local/lib/python3.7/dist-packages
sysconfig: /usr/lib/python3.7/site-packages
Value for scheme.headers does not match. Please report this to
<https://github.com/pypa/pip/issues/9617>
distutils: /usr/local/include/python3.7/UNKNOWN
sysconfig: /usr/include/python3.7m/UNKNOWN
Value for scheme.scripts does not match. Please report this to
<https://github.com/pypa/pip/issues/9617>
distutils: /usr/local/bin
sysconfig: /usr/bin
Value for scheme.data does not match. Please report this to
<https://github.com/pypa/pip/issues/9617>
distutils: /usr/local
sysconfig: /usr
Additional context:
user = False
home = None
root = None
prefix = None
Non-user install because site-packages writable
Created temporary directory: /tmp/pip-ephem-wheel-cache-uaq3hjj9
Created temporary directory: /tmp/pip-req-tracker-vxk5jpbx
Initialized build tracking at /tmp/pip-req-tracker-vxk5jpbx
Created build tracker: /tmp/pip-req-tracker-vxk5jpbx
Entered build tracker: /tmp/pip-req-tracker-vxk5jpbx
```

```

Created temporary directory: /tmp/pip-install-9r_0cibc
Looking in indexes: https://pypi.org/simple, https://us-python.pkg.dev/colab-
wheels/public/simple/
Processing /content/ditto/apex
  Created temporary directory: /tmp/pip-req-build-m44xc_r_
  DEPRECATION: A future pip version will change local packages to be built
in-place without first copying to a temporary directory. We recommend you use
--use-feature=in-tree-build to test your packages with this new behavior before
it becomes the default.

  pip 21.3 will remove support for this functionality. You can find discussion
regarding this at https://github.com/pypa/pip/issues/7555.
  Added file:///content/ditto/apex to build tracker '/tmp/pip-req-tracker-
vxk5jpbx'
  Running setup.py (path:/tmp/pip-req-build-m44xc_r_/setup.py) egg_info for
package from file:///content/ditto/apex
  Created temporary directory: /tmp/pip-pip-egg-info-ik2umxmt
  Running command python setup.py egg_info

  torch.__version__ = 1.4.0

  running egg_info
  creating /tmp/pip-pip-egg-info-ik2umxmt/apex.egg-info
  writing /tmp/pip-pip-egg-info-ik2umxmt/apex.egg-info/PKG-INFO
  writing dependency_links to /tmp/pip-pip-egg-info-ik2umxmt/apex.egg-
info/dependency_links.txt
  writing top-level names to /tmp/pip-pip-egg-info-ik2umxmt/apex.egg-
info/top_level.txt
  writing manifest file '/tmp/pip-pip-egg-info-ik2umxmt/apex.egg-
info/SOURCES.txt'
  adding license file 'LICENSE'
  writing manifest file '/tmp/pip-pip-egg-info-ik2umxmt/apex.egg-
info/SOURCES.txt'
  Source in /tmp/pip-req-build-m44xc_r_ has version 0.1, which satisfies
requirement apex==0.1 from file:///content/ditto/apex
  Removed apex==0.1 from file:///content/ditto/apex from build tracker
'/tmp/pip-req-tracker-vxk5jpbx'
Created temporary directory: /tmp/pip-unpack-p7l_z5qh
Building wheels for collected packages: apex
  Created temporary directory: /tmp/pip-wheel-6o3njxx1
  Building wheel for apex (setup.py) ... Destination directory:
/tmp/pip-wheel-6o3njxx1
  Running command /usr/bin/python3 -u -c 'import io, os, sys, setuptools,
tokenize; sys.argv[0] = ""'/tmp/pip-req-build-m44xc_r_/setup.py'''';

```



```
__file__='''/tmp/pip-req-build-m44xc_r/setup.py''';f = getattr(tokenize,
''''open''', open)(__file__) if os.path.exists(__file__) else
io.StringIO('''from setuptools import setup; setup()''');code =
f.read().replace('''\r\n''', '''\n''');f.close();exec(compile(code,
__file__, '''exec'''))' bdist_wheel -d /tmp/pip-wheel-6o3njxx1
```

```
torch.__version__ = 1.4.0
```

```
running bdist_wheel
running build
running build_py
creating build
creating build/lib
creating build/lib/apex
copying apex/__init__.py -> build/lib/apex
copying apex/_autocast_utils.py -> build/lib/apex
creating build/lib/apex/amp
copying apex/amp/scaler.py -> build/lib/apex/amp
copying apex/amp/_initialize.py -> build/lib/apex/amp
copying apex/amp/handle.py -> build/lib/apex/amp
copying apex/amp/opt.py -> build/lib/apex/amp
copying apex/amp/_amp_state.py -> build/lib/apex/amp
copying apex/amp/__init__.py -> build/lib/apex/amp
copying apex/amp/wrap.py -> build/lib/apex/amp
copying apex/amp/rnn_compat.py -> build/lib/apex/amp
copying apex/amp/_process_optimizer.py -> build/lib/apex/amp
copying apex/amp/utils.py -> build/lib/apex/amp
copying apex/amp/amp.py -> build/lib/apex/amp
copying apex/amp/frontend.py -> build/lib/apex/amp
copying apex/amp/compat.py -> build/lib/apex/amp
copying apex/amp/_version__.py -> build/lib/apex/amp
creating build/lib/apex/optimizers
copying apex/optimizers/fused_lamb.py -> build/lib/apex/optimizers
copying apex/optimizers/fused_adagrad.py -> build/lib/apex/optimizers
copying apex/optimizers/__init__.py -> build/lib/apex/optimizers
copying apex/optimizers/fused_mixed_precision_lamb.py ->
build/lib/apex/optimizers
copying apex/optimizers/fused_novograd.py -> build/lib/apex/optimizers
copying apex/optimizers/fused_sgd.py -> build/lib/apex/optimizers
copying apex/optimizers/fused_adam.py -> build/lib/apex/optimizers
creating build/lib/apex/contrib
copying apex/contrib/__init__.py -> build/lib/apex/contrib
creating build/lib/apex/fused_dense
copying apex/fused_dense/__init__.py -> build/lib/apex/fused_dense
copying apex/fused_dense/fused_dense.py -> build/lib/apex/fused_dense
creating build/lib/apex/normalization
```

```

copying apex/normalization/__init__.py -> build/lib/apex/normalization
copying apex/normalization/fused_layer_norm.py -> build/lib/apex/normalization
creating build/lib/apex/mlp
copying apex/mlp/mlp.py -> build/lib/apex/mlp
copying apex/mlp/__init__.py -> build/lib/apex/mlp
creating build/lib/apex/RNN
copying apex/RNN/cells.py -> build/lib/apex/RNN
copying apex/RNN/__init__.py -> build/lib/apex/RNN
copying apex/RNN/RNNBackend.py -> build/lib/apex/RNN
copying apex/RNN/models.py -> build/lib/apex/RNN
creating build/lib/apex/parallel
copying apex/parallel/LARC.py -> build/lib/apex/parallel
copying apex/parallel/multiproc.py -> build/lib/apex/parallel
copying apex/parallel/__init__.py -> build/lib/apex/parallel
copying apex/parallel/distributed.py -> build/lib/apex/parallel
copying apex/parallel/optimized_sync_batchnorm.py -> build/lib/apex/parallel
copying apex/parallel/sync_batchnorm_kernel.py -> build/lib/apex/parallel
copying apex/parallel/optimized_sync_batchnorm_kernel.py ->
build/lib/apex/parallel
copying apex/parallel/sync_batchnorm.py -> build/lib/apex/parallel
creating build/lib/apex/transformer
copying apex/transformer/parallel_state.py -> build/lib/apex/transformer
copying apex/transformer/enums.py -> build/lib/apex/transformer
copying apex/transformer/__init__.py -> build/lib/apex/transformer
copying apex/transformer/microbatches.py -> build/lib/apex/transformer
copying apex/transformer/utils.py -> build/lib/apex/transformer
copying apex/transformer/log_util.py -> build/lib/apex/transformer
creating build/lib/apex/fp16_utils
copying apex/fp16_utils/fp16_optimizer.py -> build/lib/apex/fp16_utils
copying apex/fp16_utils/__init__.py -> build/lib/apex/fp16_utils
copying apex/fp16_utils/loss_scaler.py -> build/lib/apex/fp16_utils
copying apex/fp16_utils/fp16util.py -> build/lib/apex/fp16_utils
creating build/lib/apex/multi_tensor_apply
copying apex/multi_tensor_apply/__init__.py ->
build/lib/apex/multi_tensor_apply
copying apex/multi_tensor_apply/multi_tensor_apply.py ->
build/lib/apex/multi_tensor_apply
creating build/lib/apex/amp/lists
copying apex/amp/lists/functional_overrides.py -> build/lib/apex/amp/lists
copying apex/amp/lists/tensor_overrides.py -> build/lib/apex/amp/lists
copying apex/amp/lists/__init__.py -> build/lib/apex/amp/lists
copying apex/amp/lists/torch_overrides.py -> build/lib/apex/amp/lists
creating build/lib/apex/contrib/xentropy
copying apex/contrib/xentropy/__init__.py -> build/lib/apex/contrib/xentropy
copying apex/contrib/xentropy/softmax_xentropy.py ->
build/lib/apex/contrib/xentropy
creating build/lib/apex/contrib/multihead_attn
copying apex/contrib/multihead_attn/fast_self_multihead_attn_func.py ->

```

```

build/lib/apex/contrib/multihead_attn
  copying apex/contrib/multihead_attn/fast_self_multihead_attn_norm_add_func.py
-> build/lib/apex/contrib/multihead_attn
  copying apex/contrib/multihead_attn/self_multihead_attn_func.py ->
build/lib/apex/contrib/multihead_attn
  copying apex/contrib/multihead_attn/self_multihead_attn.py ->
build/lib/apex/contrib/multihead_attn
  copying apex/contrib/multihead_attn/__init__.py ->
build/lib/apex/contrib/multihead_attn
  copying
apex/contrib/multihead_attn/fast_encdec_multihead_attn_norm_add_func.py ->
build/lib/apex/contrib/multihead_attn
  copying apex/contrib/multihead_attn/encdec_multihead_attn_func.py ->
build/lib/apex/contrib/multihead_attn
  copying apex/contrib/multihead_attn/mask_softmax_dropout_func.py ->
build/lib/apex/contrib/multihead_attn
  copying apex/contrib/multihead_attn/encdec_multihead_attn.py ->
build/lib/apex/contrib/multihead_attn
  copying apex/contrib/multihead_attn/fast_encdec_multihead_attn_func.py ->
build/lib/apex/contrib/multihead_attn
  creating build/lib/apex/contrib/optimizers
  copying apex/contrib/optimizers/fused_lamb.py ->
build/lib/apex/contrib/optimizers
  copying apex/contrib/optimizers/fp16_optimizer.py ->
build/lib/apex/contrib/optimizers
  copying apex/contrib/optimizers/__init__.py ->
build/lib/apex/contrib/optimizers
  copying apex/contrib/optimizers/distributed_fused_lamb.py ->
build/lib/apex/contrib/optimizers
  copying apex/contrib/optimizers/distributed_fused_adam.py ->
build/lib/apex/contrib/optimizers
  copying apex/contrib/optimizers/fused_sgd.py ->
build/lib/apex/contrib/optimizers
  copying apex/contrib/optimizers/fused_adam.py ->
build/lib/apex/contrib/optimizers
  creating build/lib/apex/contrib/clip_grad
  copying apex/contrib/clip_grad/__init__.py -> build/lib/apex/contrib/clip_grad
  copying apex/contrib/clip_grad/clip_grad.py ->
build/lib/apex/contrib/clip_grad
  creating build/lib/apex/contrib/groupbn
  copying apex/contrib/groupbn/batch_norm.py -> build/lib/apex/contrib/groupbn
  copying apex/contrib/groupbn/__init__.py -> build/lib/apex/contrib/groupbn
  creating build/lib/apex/contrib/transducer
  copying apex/contrib/transducer/transducer.py ->
build/lib/apex/contrib/transducer
  copying apex/contrib/transducer/__init__.py ->
build/lib/apex/contrib/transducer
  creating build/lib/apex/contrib/peer_memory

```

```

    copying apex/contrib/peer_memory/peer_halo_exchanger_1d.py ->
build/lib/apex/contrib/peer_memory
    copying apex/contrib/peer_memory/peer_halo_exchange_module_tests.py ->
build/lib/apex/contrib/peer_memory
    copying apex/contrib/peer_memory/peer_memory.py ->
build/lib/apex/contrib/peer_memory
    copying apex/contrib/peer_memory/__init__.py ->
build/lib/apex/contrib/peer_memory
    creating build/lib/apex/contrib/sparsity
    copying apex/contrib/sparsity/permutation_lib.py ->
build/lib/apex/contrib/sparsity
    copying apex/contrib/sparsity/__init__.py -> build/lib/apex/contrib/sparsity
    copying apex/contrib/sparsity/sparse_masklib.py ->
build/lib/apex/contrib/sparsity
    copying apex/contrib/sparsity/asp.py -> build/lib/apex/contrib/sparsity
    creating build/lib/apex/contrib/conv_bias_relu
    copying apex/contrib/conv_bias_relu/__init__.py ->
build/lib/apex/contrib/conv_bias_relu
    copying apex/contrib/conv_bias_relu/conv_bias_relu.py ->
build/lib/apex/contrib/conv_bias_relu
    creating build/lib/apex/contrib/bottleneck
    copying apex/contrib/bottleneck/bottleneck.py ->
build/lib/apex/contrib/bottleneck
    copying apex/contrib/bottleneck/bottleneck_module_test.py ->
build/lib/apex/contrib/bottleneck
    copying apex/contrib/bottleneck/__init__.py ->
build/lib/apex/contrib/bottleneck
    copying apex/contrib/bottleneck/test.py -> build/lib/apex/contrib/bottleneck
    copying apex/contrib/bottleneck/halo_exchangers.py ->
build/lib/apex/contrib/bottleneck
    creating build/lib/apex/contrib/fmha
    copying apex/contrib/fmha/__init__.py -> build/lib/apex/contrib/fmha
    copying apex/contrib/fmha/fmha.py -> build/lib/apex/contrib/fmha
    creating build/lib/apex/contrib/focal_loss
    copying apex/contrib/focal_loss/__init__.py ->
build/lib/apex/contrib/focal_loss
    copying apex/contrib/focal_loss/focal_loss.py ->
build/lib/apex/contrib/focal_loss
    creating build/lib/apex/contrib/layer_norm
    copying apex/contrib/layer_norm/__init__.py ->
build/lib/apex/contrib/layer_norm
    copying apex/contrib/layer_norm/layer_norm.py ->
build/lib/apex/contrib/layer_norm
    creating build/lib/apex/contrib/index_mul_2d
    copying apex/contrib/index_mul_2d/index_mul_2d.py ->
build/lib/apex/contrib/index_mul_2d
    copying apex/contrib/index_mul_2d/__init__.py ->
build/lib/apex/contrib/index_mul_2d

```

```

    creating build/lib/apex/contrib/sparsity/permutation_search_kernels
    copying apex/contrib/sparsity/permutation_search_kernels/exhaustive_search.py
-> build/lib/apex/contrib/sparsity/permutation_search_kernels
    copying apex/contrib/sparsity/permutation_search_kernels/__init__.py ->
build/lib/apex/contrib/sparsity/permutation_search_kernels
    copying
apex/contrib/sparsity/permutation_search_kernels/permutation_utilities.py ->
build/lib/apex/contrib/sparsity/permutation_search_kernels
    copying apex/contrib/sparsity/permutation_search_kernels/call_permutation_sear
ch_kernels.py -> build/lib/apex/contrib/sparsity/permutation_search_kernels
    creating build/lib/apex/transformer/amp
    copying apex/transformer/amp/__init__.py -> build/lib/apex/transformer/amp
    copying apex/transformer/amp/grad_scaler.py -> build/lib/apex/transformer/amp
    creating build/lib/apex/transformer/layers
    copying apex/transformer/layers/__init__.py ->
build/lib/apex/transformer/layers
    copying apex/transformer/layers/layer_norm.py ->
build/lib/apex/transformer/layers
    creating build/lib/apex/transformer/functional
    copying apex/transformer/functional/fused_softmax.py ->
build/lib/apex/transformer/functional
    copying apex/transformer/functional/__init__.py ->
build/lib/apex/transformer/functional
    creating build/lib/apex/transformer/pipeline_parallel
    copying apex/transformer/pipeline_parallel/p2p_communication.py ->
build/lib/apex/transformer/pipeline_parallel
    copying apex/transformer/pipeline_parallel/__init__.py ->
build/lib/apex/transformer/pipeline_parallel
    copying apex/transformer/pipeline_parallel/_timers.py ->
build/lib/apex/transformer/pipeline_parallel
    copying apex/transformer/pipeline_parallel/utils.py ->
build/lib/apex/transformer/pipeline_parallel
    creating build/lib/apex/transformer/tensor_parallel
    copying apex/transformer/tensor_parallel/cross_entropy.py ->
build/lib/apex/transformer/tensor_parallel
    copying apex/transformer/tensor_parallel/__init__.py ->
build/lib/apex/transformer/tensor_parallel
    copying apex/transformer/tensor_parallel/data.py ->
build/lib/apex/transformer/tensor_parallel
    copying apex/transformer/tensor_parallel/memory.py ->
build/lib/apex/transformer/tensor_parallel
    copying apex/transformer/tensor_parallel/random.py ->
build/lib/apex/transformer/tensor_parallel
    copying apex/transformer/tensor_parallel/utils.py ->
build/lib/apex/transformer/tensor_parallel
    copying apex/transformer/tensor_parallel/layers.py ->
build/lib/apex/transformer/tensor_parallel
    copying apex/transformer/tensor_parallel/mappings.py ->

```

```

build/lib/apex/transformer/tensor_parallel
  creating build/lib/apex/transformer/testing
  copying apex/transformer/testing/distributed_test_base.py ->
build/lib/apex/transformer/testing
  copying apex/transformer/testing/__init__.py ->
build/lib/apex/transformer/testing
  copying apex/transformer/testing/standalone_bert.py ->
build/lib/apex/transformer/testing
  copying apex/transformer/testing/global_vars.py ->
build/lib/apex/transformer/testing
  copying apex/transformer/testing/commons.py ->
build/lib/apex/transformer/testing
  copying apex/transformer/testing/standalone_gpt.py ->
build/lib/apex/transformer/testing
  copying apex/transformer/testing/arguments.py ->
build/lib/apex/transformer/testing
  copying apex/transformer/testing/standalone_transformer_lm.py ->
build/lib/apex/transformer/testing
  creating build/lib/apex/transformer/_data
  copying apex/transformer/_data/__init__.py -> build/lib/apex/transformer/_data
  copying apex/transformer/_data/_batch_sampler.py ->
build/lib/apex/transformer/_data
  creating build/lib/apex/transformer/pipeline_parallel/schedules
  copying apex/transformer/pipeline_parallel/schedules/fwd_bwd_pipeline_parallel_with_interleaving.py -> build/lib/apex/transformer/pipeline_parallel/schedules
  copying apex/transformer/pipeline_parallel/schedules/common.py ->
build/lib/apex/transformer/pipeline_parallel/schedules
  copying apex/transformer/pipeline_parallel/schedules/fwd_bwd_pipeline_parallel_without_interleaving.py -> build/lib/apex/transformer/pipeline_parallel/schedules
  copying apex/transformer/pipeline_parallel/schedules/fwd_bwd_no_pipeline_parallel.py -> build/lib/apex/transformer/pipeline_parallel/schedules
  copying apex/transformer/pipeline_parallel/schedules/__init__.py ->
build/lib/apex/transformer/pipeline_parallel/schedules
  installing to build/bdist.linux-x86_64/wheel
  running install
  running install_lib
  creating build/bdist.linux-x86_64
  creating build/bdist.linux-x86_64/wheel
  creating build/bdist.linux-x86_64/wheel/apex
  creating build/bdist.linux-x86_64/wheel/apex/amp
  copying build/lib/apex/amp/scaler.py ->
build/bdist.linux-x86_64/wheel/apex/amp
  copying build/lib/apex/amp/_initialize.py ->
build/bdist.linux-x86_64/wheel/apex/amp
  copying build/lib/apex/amp/handle.py ->
build/bdist.linux-x86_64/wheel/apex/amp
  copying build/lib/apex/amp/opt.py -> build/bdist.linux-x86_64/wheel/apex/amp
  copying build/lib/apex/amp/_amp_state.py ->

```

```

build/bdist.linux-x86_64/wheel/apex/amp
  copying build/lib/apex/amp/__init__.py ->
build/bdist.linux-x86_64/wheel/apex/amp
  copying build/lib/apex/amp/wrap.py -> build/bdist.linux-x86_64/wheel/apex/amp
  copying build/lib/apex/amp/rnn_compat.py ->
build/bdist.linux-x86_64/wheel/apex/amp
  copying build/lib/apex/amp/_process_optimizer.py ->
build/bdist.linux-x86_64/wheel/apex/amp
  creating build/bdist.linux-x86_64/wheel/apex/amp/lists
  copying build/lib/apex/amp/lists/functional_overrides.py ->
build/bdist.linux-x86_64/wheel/apex/amp/lists
  copying build/lib/apex/amp/lists/tensor_overrides.py ->
build/bdist.linux-x86_64/wheel/apex/amp/lists
  copying build/lib/apex/amp/lists/__init__.py ->
build/bdist.linux-x86_64/wheel/apex/amp/lists
  copying build/lib/apex/amp/lists/torch_overrides.py ->
build/bdist.linux-x86_64/wheel/apex/amp/lists
  copying build/lib/apex/amp/utils.py -> build/bdist.linux-x86_64/wheel/apex/amp
  copying build/lib/apex/amp/amp.py -> build/bdist.linux-x86_64/wheel/apex/amp
  copying build/lib/apex/amp/frontend.py ->
build/bdist.linux-x86_64/wheel/apex/amp
  copying build/lib/apex/amp/compat.py ->
build/bdist.linux-x86_64/wheel/apex/amp
  copying build/lib/apex/amp/__version__.py ->
build/bdist.linux-x86_64/wheel/apex/amp
  creating build/bdist.linux-x86_64/wheel/apex/optimizers
  copying build/lib/apex/optimizers/fused_lamb.py ->
build/bdist.linux-x86_64/wheel/apex/optimizers
  copying build/lib/apex/optimizers/fused_adagrad.py ->
build/bdist.linux-x86_64/wheel/apex/optimizers
  copying build/lib/apex/optimizers/__init__.py ->
build/bdist.linux-x86_64/wheel/apex/optimizers
  copying build/lib/apex/optimizers/fused_mixed_precision_lamb.py ->
build/bdist.linux-x86_64/wheel/apex/optimizers
  copying build/lib/apex/optimizers/fused_novograd.py ->
build/bdist.linux-x86_64/wheel/apex/optimizers
  copying build/lib/apex/optimizers/fused_sgd.py ->
build/bdist.linux-x86_64/wheel/apex/optimizers
  copying build/lib/apex/optimizers/fused_adam.py ->
build/bdist.linux-x86_64/wheel/apex/optimizers
  creating build/bdist.linux-x86_64/wheel/apex/contrib
  creating build/bdist.linux-x86_64/wheel/apex/contrib/xentropy
  copying build/lib/apex/contrib/xentropy/__init__.py ->
build/bdist.linux-x86_64/wheel/apex/contrib/xentropy
  copying build/lib/apex/contrib/xentropy/softmax_xentropy.py ->
build/bdist.linux-x86_64/wheel/apex/contrib/xentropy
  creating build/bdist.linux-x86_64/wheel/apex/contrib/multihead_attn
  copying build/lib/apex/contrib/multihead_attn/fast_self_multihead_attn_func.py

```

```

-> build/bdist.linux-x86_64/wheel/apex/contrib/multihead_attn
    copying
build/lib/apex/contrib/multihead_attn/fast_self_multihead_attn_norm_add_func.py
-> build/bdist.linux-x86_64/wheel/apex/contrib/multihead_attn
    copying build/lib/apex/contrib/multihead_attn/self_multihead_attn_func.py ->
build/bdist.linux-x86_64/wheel/apex/contrib/multihead_attn
    copying build/lib/apex/contrib/multihead_attn/self_multihead_attn.py ->
build/bdist.linux-x86_64/wheel/apex/contrib/multihead_attn
    copying build/lib/apex/contrib/multihead_attn/__init__.py ->
build/bdist.linux-x86_64/wheel/apex/contrib/multihead_attn
    copying build/lib/apex/contrib/multihead_attn/fast_encdec_multihead_attn_norm_
add_func.py -> build/bdist.linux-x86_64/wheel/apex/contrib/multihead_attn
    copying build/lib/apex/contrib/multihead_attn/encdec_multihead_attn_func.py ->
build/bdist.linux-x86_64/wheel/apex/contrib/multihead_attn
    copying build/lib/apex/contrib/multihead_attn/mask_softmax_dropout_func.py ->
build/bdist.linux-x86_64/wheel/apex/contrib/multihead_attn
    copying build/lib/apex/contrib/multihead_attn/encdec_multihead_attn.py ->
build/bdist.linux-x86_64/wheel/apex/contrib/multihead_attn
    copying
build/lib/apex/contrib/multihead_attn/fast_encdec_multihead_attn_func.py ->
build/bdist.linux-x86_64/wheel/apex/contrib/multihead_attn
    creating build/bdist.linux-x86_64/wheel/apex/contrib/optimizers
    copying build/lib/apex/contrib/optimizers/fused_lamb.py ->
build/bdist.linux-x86_64/wheel/apex/contrib/optimizers
    copying build/lib/apex/contrib/optimizers/fp16_optimizer.py ->
build/bdist.linux-x86_64/wheel/apex/contrib/optimizers
    copying build/lib/apex/contrib/optimizers/__init__.py ->
build/bdist.linux-x86_64/wheel/apex/contrib/optimizers
    copying build/lib/apex/contrib/optimizers/distributed_fused_lamb.py ->
build/bdist.linux-x86_64/wheel/apex/contrib/optimizers
    copying build/lib/apex/contrib/optimizers/distributed_fused_adam.py ->
build/bdist.linux-x86_64/wheel/apex/contrib/optimizers
    copying build/lib/apex/contrib/optimizers/fused_sgd.py ->
build/bdist.linux-x86_64/wheel/apex/contrib/optimizers
    copying build/lib/apex/contrib/optimizers/fused_adam.py ->
build/bdist.linux-x86_64/wheel/apex/contrib/optimizers
    creating build/bdist.linux-x86_64/wheel/apex/contrib/clip_grad
    copying build/lib/apex/contrib/clip_grad/__init__.py ->
build/bdist.linux-x86_64/wheel/apex/contrib/clip_grad
    copying build/lib/apex/contrib/clip_grad/clip_grad.py ->
build/bdist.linux-x86_64/wheel/apex/contrib/clip_grad
    creating build/bdist.linux-x86_64/wheel/apex/contrib/groupbn
    copying build/lib/apex/contrib/groupbn/batch_norm.py ->
build/bdist.linux-x86_64/wheel/apex/contrib/groupbn
    copying build/lib/apex/contrib/groupbn/__init__.py ->
build/bdist.linux-x86_64/wheel/apex/contrib/groupbn
    copying build/lib/apex/contrib/__init__.py ->
build/bdist.linux-x86_64/wheel/apex/contrib

```



```

    creating build/bdist.linux-x86_64/wheel/apex/contrib/transducer
    copying build/lib/apex/contrib/transducer/transducer.py ->
build/bdist.linux-x86_64/wheel/apex/contrib/transducer
    copying build/lib/apex/contrib/transducer/__init__.py ->
build/bdist.linux-x86_64/wheel/apex/contrib/transducer
    creating build/bdist.linux-x86_64/wheel/apex/contrib/peer_memory
    copying build/lib/apex/contrib/peer_memory/peer_halo_exchanger_1d.py ->
build/bdist.linux-x86_64/wheel/apex/contrib/peer_memory
    copying build/lib/apex/contrib/peer_memory/peer_halo_exchange_module_tests.py
-> build/bdist.linux-x86_64/wheel/apex/contrib/peer_memory
    copying build/lib/apex/contrib/peer_memory/peer_memory.py ->
build/bdist.linux-x86_64/wheel/apex/contrib/peer_memory
    copying build/lib/apex/contrib/peer_memory/__init__.py ->
build/bdist.linux-x86_64/wheel/apex/contrib/peer_memory
    creating build/bdist.linux-x86_64/wheel/apex/contrib/sparsity
    copying build/lib/apex/contrib/sparsity/permutation_lib.py ->
build/bdist.linux-x86_64/wheel/apex/contrib/sparsity
    creating
build/bdist.linux-x86_64/wheel/apex/contrib/sparsity/permutation_search_kernels
    copying
build/lib/apex/contrib/sparsity/permutation_search_kernels/exhaustive_search.py
->
build/bdist.linux-x86_64/wheel/apex/contrib/sparsity/permutation_search_kernels
    copying build/lib/apex/contrib/sparsity/permutation_search_kernels/__init__.py
->
build/bdist.linux-x86_64/wheel/apex/contrib/sparsity/permutation_search_kernels
    copying build/lib/apex/contrib/sparsity/permutation_search_kernels/permutation
_utilities.py ->
build/bdist.linux-x86_64/wheel/apex/contrib/sparsity/permutation_search_kernels
    copying build/lib/apex/contrib/sparsity/permutation_search_kernels/call_permut
ation_search_kernels.py ->
build/bdist.linux-x86_64/wheel/apex/contrib/sparsity/permutation_search_kernels
    copying build/lib/apex/contrib/sparsity/__init__.py ->
build/bdist.linux-x86_64/wheel/apex/contrib/sparsity
    copying build/lib/apex/contrib/sparsity/sparse_masklib.py ->
build/bdist.linux-x86_64/wheel/apex/contrib/sparsity
    copying build/lib/apex/contrib/sparsity/asp.py ->
build/bdist.linux-x86_64/wheel/apex/contrib/sparsity
    creating build/bdist.linux-x86_64/wheel/apex/contrib/conv_bias_relu
    copying build/lib/apex/contrib/conv_bias_relu/__init__.py ->
build/bdist.linux-x86_64/wheel/apex/contrib/conv_bias_relu
    copying build/lib/apex/contrib/conv_bias_relu/conv_bias_relu.py ->
build/bdist.linux-x86_64/wheel/apex/contrib/conv_bias_relu
    creating build/bdist.linux-x86_64/wheel/apex/contrib/bottleneck
    copying build/lib/apex/contrib/bottleneck/bottleneck.py ->
build/bdist.linux-x86_64/wheel/apex/contrib/bottleneck
    copying build/lib/apex/contrib/bottleneck/bottleneck_module_test.py ->
build/bdist.linux-x86_64/wheel/apex/contrib/bottleneck

```

```

    copying build/lib/apex/contrib/bottleneck/__init__.py ->
build/bdist.linux-x86_64/wheel/apex/contrib/bottleneck
    copying build/lib/apex/contrib/bottleneck/test.py ->
build/bdist.linux-x86_64/wheel/apex/contrib/bottleneck
    copying build/lib/apex/contrib/bottleneck/halo_exchangers.py ->
build/bdist.linux-x86_64/wheel/apex/contrib/bottleneck
    creating build/bdist.linux-x86_64/wheel/apex/contrib/fmha
    copying build/lib/apex/contrib/fmha/__init__.py ->
build/bdist.linux-x86_64/wheel/apex/contrib/fmha
    copying build/lib/apex/contrib/fmha/fmha.py ->
build/bdist.linux-x86_64/wheel/apex/contrib/fmha
    creating build/bdist.linux-x86_64/wheel/apex/contrib/focal_loss
    copying build/lib/apex/contrib/focal_loss/__init__.py ->
build/bdist.linux-x86_64/wheel/apex/contrib/focal_loss
    copying build/lib/apex/contrib/focal_loss/focal_loss.py ->
build/bdist.linux-x86_64/wheel/apex/contrib/focal_loss
    creating build/bdist.linux-x86_64/wheel/apex/contrib/layer_norm
    copying build/lib/apex/contrib/layer_norm/__init__.py ->
build/bdist.linux-x86_64/wheel/apex/contrib/layer_norm
    copying build/lib/apex/contrib/layer_norm/layer_norm.py ->
build/bdist.linux-x86_64/wheel/apex/contrib/layer_norm
    creating build/bdist.linux-x86_64/wheel/apex/contrib/index_mul_2d
    copying build/lib/apex/contrib/index_mul_2d/index_mul_2d.py ->
build/bdist.linux-x86_64/wheel/apex/contrib/index_mul_2d
    copying build/lib/apex/contrib/index_mul_2d/__init__.py ->
build/bdist.linux-x86_64/wheel/apex/contrib/index_mul_2d
    creating build/bdist.linux-x86_64/wheel/apex/fused_dense
    copying build/lib/apex/fused_dense/__init__.py ->
build/bdist.linux-x86_64/wheel/apex/fused_dense
    copying build/lib/apex/fused_dense/fused_dense.py ->
build/bdist.linux-x86_64/wheel/apex/fused_dense
    creating build/bdist.linux-x86_64/wheel/apex/normalization
    copying build/lib/apex/normalization/__init__.py ->
build/bdist.linux-x86_64/wheel/apex/normalization
    copying build/lib/apex/normalization/fused_layer_norm.py ->
build/bdist.linux-x86_64/wheel/apex/normalization
    copying build/lib/apex/__init__.py -> build/bdist.linux-x86_64/wheel/apex
    creating build/bdist.linux-x86_64/wheel/apex/mlp
    copying build/lib/apex/mlp/mlp.py -> build/bdist.linux-x86_64/wheel/apex/mlp
    copying build/lib/apex/mlp/__init__.py ->
build/bdist.linux-x86_64/wheel/apex/mlp
    creating build/bdist.linux-x86_64/wheel/apex/RNN
    copying build/lib/apex/RNN/cells.py -> build/bdist.linux-x86_64/wheel/apex/RNN
    copying build/lib/apex/RNN/__init__.py ->
build/bdist.linux-x86_64/wheel/apex/RNN
    copying build/lib/apex/RNN/RNNBackend.py ->
build/bdist.linux-x86_64/wheel/apex/RNN
    copying build/lib/apex/RNN/models.py ->

```

```

build/bdist.linux-x86_64/wheel/apex/RNN
    copying build/lib/apex/_autocast_utils.py ->
build/bdist.linux-x86_64/wheel/apex
    creating build/bdist.linux-x86_64/wheel/apex/parallel
    copying build/lib/apex/parallel/LARC.py ->
build/bdist.linux-x86_64/wheel/apex/parallel
    copying build/lib/apex/parallel/multiproc.py ->
build/bdist.linux-x86_64/wheel/apex/parallel
    copying build/lib/apex/parallel/__init__.py ->
build/bdist.linux-x86_64/wheel/apex/parallel
    copying build/lib/apex/parallel/distributed.py ->
build/bdist.linux-x86_64/wheel/apex/parallel
    copying build/lib/apex/parallel/optimized_sync_batchnorm.py ->
build/bdist.linux-x86_64/wheel/apex/parallel
    copying build/lib/apex/parallel/sync_batchnorm_kernel.py ->
build/bdist.linux-x86_64/wheel/apex/parallel
    copying build/lib/apex/parallel/optimized_sync_batchnorm_kernel.py ->
build/bdist.linux-x86_64/wheel/apex/parallel
    copying build/lib/apex/parallel/sync_batchnorm.py ->
build/bdist.linux-x86_64/wheel/apex/parallel
    creating build/bdist.linux-x86_64/wheel/apex/transformer
    creating build/bdist.linux-x86_64/wheel/apex/transformer/amp
    copying build/lib/apex/transformer/amp/__init__.py ->
build/bdist.linux-x86_64/wheel/apex/transformer/amp
    copying build/lib/apex/transformer/amp/grad_scaler.py ->
build/bdist.linux-x86_64/wheel/apex/transformer/amp
    creating build/bdist.linux-x86_64/wheel/apex/transformer/layers
    copying build/lib/apex/transformer/layers/__init__.py ->
build/bdist.linux-x86_64/wheel/apex/transformer/layers
    copying build/lib/apex/transformer/layers/layer_norm.py ->
build/bdist.linux-x86_64/wheel/apex/transformer/layers
    copying build/lib/apex/transformer/parallel_state.py ->
build/bdist.linux-x86_64/wheel/apex/transformer
    creating build/bdist.linux-x86_64/wheel/apex/transformer/functional
    copying build/lib/apex/transformer/functional/fused_softmax.py ->
build/bdist.linux-x86_64/wheel/apex/transformer/functional
    copying build/lib/apex/transformer/functional/__init__.py ->
build/bdist.linux-x86_64/wheel/apex/transformer/functional
    copying build/lib/apex/transformer/enums.py ->
build/bdist.linux-x86_64/wheel/apex/transformer
    copying build/lib/apex/transformer/__init__.py ->
build/bdist.linux-x86_64/wheel/apex/transformer
    creating build/bdist.linux-x86_64/wheel/apex/transformer/pipeline_parallel
    copying build/lib/apex/transformer/pipeline_parallel/p2p_communication.py ->
build/bdist.linux-x86_64/wheel/apex/transformer/pipeline_parallel
    copying build/lib/apex/transformer/pipeline_parallel/__init__.py ->
build/bdist.linux-x86_64/wheel/apex/transformer/pipeline_parallel
    creating

```

```

build/bdist.linux-x86_64/wheel/apex/transformer/pipeline_parallel/schedules
  copying build/lib/apex/transformer/pipeline_parallel/schedules/fwd_bwd_pipeline_parallel_with_interleaving.py ->
build/bdist.linux-x86_64/wheel/apex/transformer/pipeline_parallel/schedules
  copying build/lib/apex/transformer/pipeline_parallel/schedules/common.py ->
build/bdist.linux-x86_64/wheel/apex/transformer/pipeline_parallel/schedules
  copying build/lib/apex/transformer/pipeline_parallel/schedules/fwd_bwd_pipeline_parallel_without_interleaving.py ->
build/bdist.linux-x86_64/wheel/apex/transformer/pipeline_parallel/schedules
  copying
build/lib/apex/transformer/pipeline_parallel/schedules/fwd_bwd_no_pipelining.py
-> build/bdist.linux-x86_64/wheel/apex/transformer/pipeline_parallel/schedules
  copying build/lib/apex/transformer/pipeline_parallel/schedules/__init__.py ->
build/bdist.linux-x86_64/wheel/apex/transformer/pipeline_parallel/schedules
  copying build/lib/apex/transformer/pipeline_parallel/_timers.py ->
build/bdist.linux-x86_64/wheel/apex/transformer/pipeline_parallel
  copying build/lib/apex/transformer/pipeline_parallel/utils.py ->
build/bdist.linux-x86_64/wheel/apex/transformer/pipeline_parallel
  copying build/lib/apex/transformer/microbatches.py ->
build/bdist.linux-x86_64/wheel/apex/transformer
  creating build/bdist.linux-x86_64/wheel/apex/transformer/tensor_parallel
  copying build/lib/apex/transformer/tensor_parallel/cross_entropy.py ->
build/bdist.linux-x86_64/wheel/apex/transformer/tensor_parallel
  copying build/lib/apex/transformer/tensor_parallel/__init__.py ->
build/bdist.linux-x86_64/wheel/apex/transformer/tensor_parallel
  copying build/lib/apex/transformer/tensor_parallel/data.py ->
build/bdist.linux-x86_64/wheel/apex/transformer/tensor_parallel
  copying build/lib/apex/transformer/tensor_parallel/memory.py ->
build/bdist.linux-x86_64/wheel/apex/transformer/tensor_parallel
  copying build/lib/apex/transformer/tensor_parallel/random.py ->
build/bdist.linux-x86_64/wheel/apex/transformer/tensor_parallel
  copying build/lib/apex/transformer/tensor_parallel/utils.py ->
build/bdist.linux-x86_64/wheel/apex/transformer/tensor_parallel
  copying build/lib/apex/transformer/tensor_parallel/layers.py ->
build/bdist.linux-x86_64/wheel/apex/transformer/tensor_parallel
  copying build/lib/apex/transformer/tensor_parallel/mappings.py ->
build/bdist.linux-x86_64/wheel/apex/transformer/tensor_parallel
  creating build/bdist.linux-x86_64/wheel/apex/transformer/testing
  copying build/lib/apex/transformer/testing/distributed_test_base.py ->
build/bdist.linux-x86_64/wheel/apex/transformer/testing
  copying build/lib/apex/transformer/testing/__init__.py ->
build/bdist.linux-x86_64/wheel/apex/transformer/testing
  copying build/lib/apex/transformer/testing/standalone_bert.py ->
build/bdist.linux-x86_64/wheel/apex/transformer/testing
  copying build/lib/apex/transformer/testing/global_vars.py ->
build/bdist.linux-x86_64/wheel/apex/transformer/testing
  copying build/lib/apex/transformer/testing/commons.py ->
build/bdist.linux-x86_64/wheel/apex/transformer/testing

```

```

    copying build/lib/apex/transformer/testing/standalone_gpt.py ->
build/bdist.linux-x86_64/wheel/apex/transformer/testing
    copying build/lib/apex/transformer/testing/arguments.py ->
build/bdist.linux-x86_64/wheel/apex/transformer/testing
    copying build/lib/apex/transformer/testing/standalone_transformer_lm.py ->
build/bdist.linux-x86_64/wheel/apex/transformer/testing
    copying build/lib/apex/transformer/utils.py ->
build/bdist.linux-x86_64/wheel/apex/transformer
    copying build/lib/apex/transformer/log_util.py ->
build/bdist.linux-x86_64/wheel/apex/transformer
    creating build/bdist.linux-x86_64/wheel/apex/transformer/_data
    copying build/lib/apex/transformer/_data/__init__.py ->
build/bdist.linux-x86_64/wheel/apex/transformer/_data
    copying build/lib/apex/transformer/_data/_batch_sampler.py ->
build/bdist.linux-x86_64/wheel/apex/transformer/_data
    creating build/bdist.linux-x86_64/wheel/apex/fp16_utils
    copying build/lib/apex/fp16_utils/fp16_optimizer.py ->
build/bdist.linux-x86_64/wheel/apex/fp16_utils
    copying build/lib/apex/fp16_utils/__init__.py ->
build/bdist.linux-x86_64/wheel/apex/fp16_utils
    copying build/lib/apex/fp16_utils/loss_scaler.py ->
build/bdist.linux-x86_64/wheel/apex/fp16_utils
    copying build/lib/apex/fp16_utils/fp16util.py ->
build/bdist.linux-x86_64/wheel/apex/fp16_utils
    creating build/bdist.linux-x86_64/wheel/apex/multi_tensor_apply
    copying build/lib/apex/multi_tensor_apply/__init__.py ->
build/bdist.linux-x86_64/wheel/apex/multi_tensor_apply
    copying build/lib/apex/multi_tensor_apply/multi_tensor_apply.py ->
build/bdist.linux-x86_64/wheel/apex/multi_tensor_apply
    running install_egg_info
    running egg_info
    creating apex.egg-info
    writing apex.egg-info/PKG-INFO
    writing dependency_links to apex.egg-info/dependency_links.txt
    writing top-level names to apex.egg-info/top_level.txt
    writing manifest file 'apex.egg-info/SOURCES.txt'
    adding license file 'LICENSE'
    writing manifest file 'apex.egg-info/SOURCES.txt'
    Copying apex.egg-info to build/bdist.linux-x86_64/wheel/apex-0.1-py3.7.egg-
info
    running install_scripts
    adding license file "LICENSE" (matched pattern "LICEN[CS]E*")
    creating build/bdist.linux-x86_64/wheel/apex-0.1.dist-info/WHEEL
    creating '/tmp/pip-wheel-6o3njxx1/apex-0.1-py3-none-any.whl' and adding
'build/bdist.linux-x86_64/wheel' to it
    adding 'apex/__init__.py'
    adding 'apex/_autocast_utils.py'
    adding 'apex/RNN/RNNBackend.py'

```

```

adding 'apex/RNN/__init__.py'
adding 'apex/RNN/cells.py'
adding 'apex/RNN/models.py'
adding 'apex/amp/__init__.py'
adding 'apex/amp/__version__.py'
adding 'apex/amp/_amp_state.py'
adding 'apex/amp/_initialize.py'
adding 'apex/amp/_process_optimizer.py'
adding 'apex/amp/amp.py'
adding 'apex/amp/compat.py'
adding 'apex/amp/frontend.py'
adding 'apex/amp/handle.py'
adding 'apex/amp/opt.py'
adding 'apex/amp/rnn_compat.py'
adding 'apex/amp/scaler.py'
adding 'apex/amp/utils.py'
adding 'apex/amp/wrap.py'
adding 'apex/amp/lists/__init__.py'
adding 'apex/amp/lists/functional_overrides.py'
adding 'apex/amp/lists/tensor_overrides.py'
adding 'apex/amp/lists/torch_overrides.py'
adding 'apex/contrib/__init__.py'
adding 'apex/contrib/bottleneck/__init__.py'
adding 'apex/contrib/bottleneck/bottleneck.py'
adding 'apex/contrib/bottleneck/bottleneck_module_test.py'
adding 'apex/contrib/bottleneck/halo_exchangers.py'
adding 'apex/contrib/bottleneck/test.py'
adding 'apex/contrib/clip_grad/__init__.py'
adding 'apex/contrib/clip_grad/clip_grad.py'
adding 'apex/contrib/conv_bias_relu/__init__.py'
adding 'apex/contrib/conv_bias_relu/conv_bias_relu.py'
adding 'apex/contrib/fmha/__init__.py'
adding 'apex/contrib/fmha/fmha.py'
adding 'apex/contrib/focal_loss/__init__.py'
adding 'apex/contrib/focal_loss/focal_loss.py'
adding 'apex/contrib/groupbn/__init__.py'
adding 'apex/contrib/groupbn/batch_norm.py'
adding 'apex/contrib/index_mul_2d/__init__.py'
adding 'apex/contrib/index_mul_2d/index_mul_2d.py'
adding 'apex/contrib/layer_norm/__init__.py'
adding 'apex/contrib/layer_norm/layer_norm.py'
adding 'apex/contrib/multihead_attn/__init__.py'
adding 'apex/contrib/multihead_attn/encdec_multihead_attn.py'
adding 'apex/contrib/multihead_attn/encdec_multihead_attn_func.py'
adding 'apex/contrib/multihead_attn/fast_encdec_multihead_attn_func.py'
adding
'apex/contrib/multihead_attn/fast_encdec_multihead_attn_norm_add_func.py'
    adding 'apex/contrib/multihead_attn/fast_self_multihead_attn_func.py'

```

```

adding 'apex/contrib/multihead_attn/fast_self_multihead_attn_norm_add_func.py'
adding 'apex/contrib/multihead_attn/mask_softmax_dropout_func.py'
adding 'apex/contrib/multihead_attn/self_multihead_attn.py'
adding 'apex/contrib/multihead_attn/self_multihead_attn_func.py'
adding 'apex/contrib/optimizers/__init__.py'
adding 'apex/contrib/optimizers/distributed_fused_adam.py'
adding 'apex/contrib/optimizers/distributed_fused_lamb.py'
adding 'apex/contrib/optimizers/fp16_optimizer.py'
adding 'apex/contrib/optimizers/fused_adam.py'
adding 'apex/contrib/optimizers/fused_lamb.py'
adding 'apex/contrib/optimizers/fused_sgd.py'
adding 'apex/contrib/peer_memory/__init__.py'
adding 'apex/contrib/peer_memory/peer_halo_exchange_module_tests.py'
adding 'apex/contrib/peer_memory/peer_halo_exchanger_1d.py'
adding 'apex/contrib/peer_memory/peer_memory.py'
adding 'apex/contrib/sparsity/__init__.py'
adding 'apex/contrib/sparsity/asp.py'
adding 'apex/contrib/sparsity/permutation_lib.py'
adding 'apex/contrib/sparsity/sparse_masklib.py'
adding 'apex/contrib/sparsity/permutation_search_kernels/__init__.py'
adding 'apex/contrib/sparsity/permutation_search_kernels/call_permutation_sear
ch_kernels.py'
    adding 'apex/contrib/sparsity/permutation_search_kernels/exhaustive_search.py'
    adding
'apex/contrib/sparsity/permutation_search_kernels/permutation_utilities.py'
    adding 'apex/contrib/transducer/__init__.py'
    adding 'apex/contrib/transducer/transducer.py'
    adding 'apex/contrib/xentropy/__init__.py'
    adding 'apex/contrib/xentropy/softmax_xentropy.py'
    adding 'apex/fp16_utils/__init__.py'
    adding 'apex/fp16_utils/fp16_optimizer.py'
    adding 'apex/fp16_utils/fp16util.py'
    adding 'apex/fp16_utils/loss_scaler.py'
    adding 'apex/fused_dense/__init__.py'
    adding 'apex/fused_dense/fused_dense.py'
    adding 'apex/mlp/__init__.py'
    adding 'apex/mlp/mlp.py'
    adding 'apex/multi_tensor_apply/__init__.py'
    adding 'apex/multi_tensor_apply/multi_tensor_apply.py'
    adding 'apex/normalization/__init__.py'
    adding 'apex/normalization/fused_layer_norm.py'
    adding 'apex/optimizers/__init__.py'
    adding 'apex/optimizers/fused_adagrad.py'
    adding 'apex/optimizers/fused_adam.py'
    adding 'apex/optimizers/fused_lamb.py'
    adding 'apex/optimizers/fused_mixed_precision_lamb.py'
    adding 'apex/optimizers/fused_novograd.py'
    adding 'apex/optimizers/fused_sgd.py'

```

```

adding 'apex/parallel/LARC.py'
adding 'apex/parallel/__init__.py'
adding 'apex/parallel/distributed.py'
adding 'apex/parallel/multiproc.py'
adding 'apex/parallel/optimized_sync_batchnorm.py'
adding 'apex/parallel/optimized_sync_batchnorm_kernel.py'
adding 'apex/parallel/sync_batchnorm.py'
adding 'apex/parallel/sync_batchnorm_kernel.py'
adding 'apex/transformer/__init__.py'
adding 'apex/transformer/enums.py'
adding 'apex/transformer/log_util.py'
adding 'apex/transformer/microbatches.py'
adding 'apex/transformer/parallel_state.py'
adding 'apex/transformer/utils.py'
adding 'apex/transformer/_data/__init__.py'
adding 'apex/transformer/_data/_batch_sampler.py'
adding 'apex/transformer/amp/__init__.py'
adding 'apex/transformer/amp/grad_scaler.py'
adding 'apex/transformer/functional/__init__.py'
adding 'apex/transformer/functional/fused_softmax.py'
adding 'apex/transformer/layers/__init__.py'
adding 'apex/transformer/layers/layer_norm.py'
adding 'apex/transformer/pipeline_parallel/__init__.py'
adding 'apex/transformer/pipeline_parallel/_timers.py'
adding 'apex/transformer/pipeline_parallel/p2p_communication.py'
adding 'apex/transformer/pipeline_parallel/utils.py'
adding 'apex/transformer/pipeline_parallel/schedules/__init__.py'
adding 'apex/transformer/pipeline_parallel/schedules/common.py'
adding 'apex/transformer/pipeline_parallel/schedules/fwd_bwd_no_pipelining.py'
adding 'apex/transformer/pipeline_parallel/schedules/fwd_bwd_pipelining_with_i
nterleaving.py'
adding 'apex/transformer/pipeline_parallel/schedules/fwd_bwd_pipelining_withou
t_interleaving.py'
adding 'apex/transformer/tensor_parallel/__init__.py'
adding 'apex/transformer/tensor_parallel/cross_entropy.py'
adding 'apex/transformer/tensor_parallel/data.py'
adding 'apex/transformer/tensor_parallel/layers.py'
adding 'apex/transformer/tensor_parallel/mappings.py'
adding 'apex/transformer/tensor_parallel/memory.py'
adding 'apex/transformer/tensor_parallel/random.py'
adding 'apex/transformer/tensor_parallel/utils.py'
adding 'apex/transformer/testing/__init__.py'
adding 'apex/transformer/testing/arguments.py'
adding 'apex/transformer/testing/commons.py'
adding 'apex/transformer/testing/distributed_test_base.py'
adding 'apex/transformer/testing/global_vars.py'
adding 'apex/transformer/testing/standalone_bert.py'
adding 'apex/transformer/testing/standalone_gpt.py'

```



```

adding 'apex/transformer/testing/standalone_transformer_lm.py'
adding 'apex-0.1.dist-info/LICENSE'
adding 'apex-0.1.dist-info/METADATA'
adding 'apex-0.1.dist-info/WHEEL'
adding 'apex-0.1.dist-info/top_level.txt'
adding 'apex-0.1.dist-info/RECORD'
removing build/bdist.linux-x86_64/wheel
done
Created wheel for apex: filename=apex-0.1-py3-none-any.whl size=298114
sha256=980195a05358dcb3f4e31131ae3398300a36a640d5742beb81b3121d33a11321
Stored in directory: /tmp/pip-ephem-wheel-cache-
uaq3hjj9/wheels/6b/22/a2/1665526ee3c3061243260fd522525eca31398e04bfa5ad7e1f
Successfully built apex
Installing collected packages: apex
Value for scheme.platlib does not match. Please report this to
<https://github.com/pypa/pip/issues/9617>
distutils: /usr/local/lib/python3.7/dist-packages
sysconfig: /usr/lib/python3.7/site-packages
Value for scheme.purelib does not match. Please report this to
<https://github.com/pypa/pip/issues/9617>
distutils: /usr/local/lib/python3.7/dist-packages
sysconfig: /usr/lib/python3.7/site-packages
Value for scheme.headers does not match. Please report this to
<https://github.com/pypa/pip/issues/9617>
distutils: /usr/local/include/python3.7/apex
sysconfig: /usr/include/python3.7m/apex
Value for scheme.scripts does not match. Please report this to
<https://github.com/pypa/pip/issues/9617>
distutils: /usr/local/bin
sysconfig: /usr/bin
Value for scheme.data does not match. Please report this to
<https://github.com/pypa/pip/issues/9617>
distutils: /usr/local
sysconfig: /usr
Additional context:
user = False
home = None
root = None
prefix = None

Value for scheme.platlib does not match. Please report this to
<https://github.com/pypa/pip/issues/9617>
distutils: /usr/local/lib/python3.7/dist-packages
sysconfig: /usr/lib/python3.7/site-packages
Value for scheme.purelib does not match. Please report this to
<https://github.com/pypa/pip/issues/9617>
distutils: /usr/local/lib/python3.7/dist-packages
sysconfig: /usr/lib/python3.7/site-packages

```

```

Value for scheme.headers does not match. Please report this to
<https://github.com/pypa/pip/issues/9617>
distutils: /usr/local/include/python3.7/UNKNOWN
sysconfig: /usr/include/python3.7m/UNKNOWN
Value for scheme.scripts does not match. Please report this to
<https://github.com/pypa/pip/issues/9617>
distutils: /usr/local/bin
sysconfig: /usr/bin
Value for scheme.data does not match. Please report this to
<https://github.com/pypa/pip/issues/9617>
distutils: /usr/local
sysconfig: /usr
Additional context:
user = False
home = None
root = None
prefix = None
Successfully installed apex-0.1
Removed build tracker: '/tmp/pip-req-tracker-vxk5jpbx'
/content/ditto

```

```

[ ]: # some issue with colab
!pip install --upgrade "urllib3==1.25.4" awscli

```

```

Looking in indexes: https://pypi.org/simple, https://us-python.pkg.dev/colab-
wheels/public/simple/
Collecting urllib3==1.25.4
  Downloading urllib3-1.25.4-py2.py3-none-any.whl (125 kB)
    |                               | 125 kB 32.3 MB/s
Collecting awscli
  Downloading awscli-1.25.48-py3-none-any.whl (3.9 MB)
    |                               | 3.9 MB 56.1 MB/s
Collecting colorama<0.4.5,>=0.2.5
  Downloading colorama-0.4.4-py2.py3-none-any.whl (16 kB)
Collecting docutils<0.17,>=0.10
  Downloading docutils-0.16-py2.py3-none-any.whl (548 kB)
    |                               | 548 kB 69.0 MB/s
Collecting rsa<4.8,>=3.1.2
  Downloading rsa-4.7.2-py3-none-any.whl (34 kB)
Collecting boto3==1.27.48
  Downloading boto3-1.27.48-py3-none-any.whl (9.0 MB)
    |                               | 9.0 MB 53.1 MB/s
Requirement already satisfied: PyYAML<5.5,>=3.10 in
/usr/local/lib/python3.7/dist-packages (from awscli) (3.13)
Collecting s3transfer<0.7.0,>=0.6.0
  Downloading s3transfer-0.6.0-py3-none-any.whl (79 kB)
    |                               | 79 kB 9.3 MB/s
Requirement already satisfied: python-dateutil<3.0.0,>=2.1 in

```

```

/usr/local/lib/python3.7/dist-packages (from botocore==1.27.48->awscli) (2.8.2)
Collecting jmespath<2.0.0,>=0.7.1
  Downloading jmespath-1.0.1-py3-none-any.whl (20 kB)
Requirement already satisfied: six>=1.5 in /usr/local/lib/python3.7/dist-
packages (from python-dateutil<3.0.0,>=2.1->botocore==1.27.48->awscli) (1.15.0)
Requirement already satisfied: pyasn1>=0.1.3 in /usr/local/lib/python3.7/dist-
packages (from rsa<4.8,>=3.1.2->awscli) (0.4.8)
Installing collected packages: urllib3, jmespath, botocore, s3transfer, rsa,
docutils, colorama, awscli
  Attempting uninstall: urllib3
    Found existing installation: urllib3 1.24.3
    Uninstalling urllib3-1.24.3:
      Successfully uninstalled urllib3-1.24.3
  Attempting uninstall: rsa
    Found existing installation: rsa 4.9
    Uninstalling rsa-4.9:
      Successfully uninstalled rsa-4.9
  Attempting uninstall: docutils
    Found existing installation: docutils 0.17.1
    Uninstalling docutils-0.17.1:
      Successfully uninstalled docutils-0.17.1
ERROR: pip's dependency resolver does not currently take into account all
the packages that are installed. This behaviour is the source of the following
dependency conflicts.
panel 0.12.1 requires tqdm>=4.48.0, but you have tqdm 4.41.0 which is
incompatible.
fastai 2.7.7 requires torch<1.13,>=1.7, but you have torch 1.4.0 which is
incompatible.
Successfully installed awscli-1.25.48 botocore-1.27.48 colorama-0.4.4
docutils-0.16 jmespath-1.0.1 rsa-4.7.2 s3transfer-0.6.0 urllib3-1.25.4

```

```
[ ]: !pip install torch==1.7.1
```

```

Looking in indexes: https://pypi.org/simple, https://us-python.pkg.dev/colab-
wheels/public/simple/
Collecting torch==1.7.1
  Downloading torch-1.7.1-cp37-cp37m-manylinux1_x86_64.whl (776.8 MB)
    |                                     | 776.8 MB 17 kB/s
Requirement already satisfied: typing-extensions in
/usr/local/lib/python3.7/dist-packages (from torch==1.7.1) (4.1.1)
Requirement already satisfied: numpy in /usr/local/lib/python3.7/dist-packages
(from torch==1.7.1) (1.19.2)
Installing collected packages: torch
  Attempting uninstall: torch
    Found existing installation: torch 1.4.0

```

```

Uninstalling torch-1.4.0:
  Successfully uninstalled torch-1.4.0
ERROR: pip's dependency resolver does not currently take into account all
the packages that are installed. This behaviour is the source of the following
dependency conflicts.

torchvision 0.13.0+cu113 requires torch==1.12.0, but you have torch 1.7.1 which
is incompatible.

torchtext 0.13.0 requires torch==1.12.0, but you have torch 1.7.1 which is
incompatible.

torchaudio 0.12.0+cu113 requires torch==1.12.0, but you have torch 1.7.1 which
is incompatible.
Successfully installed torch-1.7.1

```

### 0.3 importa config.json

```
[ ]: !git clone https://github.com/pauloh48/IC.git
```

```

Cloning into 'IC'...
remote: Enumerating objects: 985, done.
remote: Counting objects: 100% (55/55), done.
remote: Compressing objects: 100% (51/51), done.
remote: Total 985 (delta 16), reused 7 (delta 0), pack-reused 930
Receiving objects: 100% (985/985), 252.10 MiB | 17.29 MiB/s, done.
Resolving deltas: 100% (600/600), done.
Checking out files: 100% (1274/1274), done.

```

### 0.4 remove e move

```
[ ]: !ls
```

```

apex          ditto.jpg    LICENSE     requirements.txt  run_all_wdc.py
blocking      ditto_light matcher.py  results_ditto    Snippext_public
configs.json  IC          output      run_all_er_magellan.py train_ditto.py
data          input       README.md   run_all_vary_size.py

```

```
[ ]: !rm configs.json
!mv -n /content/ditto/IC/datasesErros/AADatasets/configs.json /content/ditto/
```

# 1 DA

## 1.1 DA STRUCTURED

### 1.1.1 Beer - Running the matcher

```
[ ]: # Step 1:
# train the wdc model with the save_model flag on
# it will produce two files *_dev.pt and *_test.pt which are the
# best checkpoints on the dev set and the test set.
!CUDA_VISIBLE_DEVICES=0 python train_ditto.py \
  --task Structured/Beer \
  --batch_size 32 \
  --max_len 256 \
  --lr 3e-5 \
  --n_epochs 15 \
  --finetuning \
  --lm roberta \
  --fp16 \
  --save_model \
  --da entry_swap \
  --da attr_del \
  --summarize

# Step 2: pick a checkpoint, rename it to wdc_all_small.pt
!mv *_dev.pt checkpoints/Beer.pt

# Step 3: run the matcher
!CUDA_VISIBLE_DEVICES=0 python matcher.py \
  --task Structured/Beer \
  --input_path input/input_small.jsonl \
  --output_path output/output_small.jsonl \
  --lm roberta \
  --use_gpu \
  --fp16 \
  --checkpoint_path checkpoints/
```

Selected optimization level 02: FP16 training with FP32 batchnorm and FP32 master weights.

Defaults for this optimization level are:

enabled	: True
opt_level	: 02
cast_model_type	: torch.float16
patch_torch_functions	: False
keep_batchnorm_fp32	: True

```

master_weights      : True
loss_scale          : dynamic
Processing user overrides (additional kwargs that are not None)...
After processing overrides, optimization options are:
enabled              : True
opt_level            : 02
cast_model_type      : torch.float16
patch_torch_functions : False
keep_batchnorm_fp32   : True
master_weights       : True
loss_scale           : dynamic
Warning: multi_tensor_applier fused unscale kernel is unavailable, possibly
because apex was installed without --cuda_ext --cpp_ext. Using Python fallback.
Original ImportError was: ModuleNotFoundError("No module named 'amp_C'")
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 32768.0
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:136:
UserWarning: Detected call of `lr_scheduler.step()` before `optimizer.step()`.
In PyTorch 1.1.0 and later, you should call them in the opposite order:
`optimizer.step()` before `lr_scheduler.step()`. Failure to do this will result
in PyTorch skipping the first value of the learning rate schedule. See more
details at https://pytorch.org/docs/stable/optim.html#how-to-adjust-learning-
rate
    "https://pytorch.org/docs/stable/optim.html#how-to-adjust-learning-rate",
UserWarning)
step: 0, loss: 0.4154115319252014
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
    warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 1: dev_f1=0.288659793814433, f1=0.27083333333333337,
best_f1=0.27083333333333337
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.4659673273563385
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.

```

```

    warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 2: dev_f1=0.4057971014492754, f1=0.36666666666666664,
best_f1=0.36666666666666664
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.41500335931777954
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
    warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 3: dev_f1=0.48888888888888893, f1=0.4782608695652174,
best_f1=0.4782608695652174
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.2803691625595093
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
    warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 4: dev_f1=0.7857142857142857, f1=0.689655172413793,
best_f1=0.689655172413793
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.1263187676668167
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
    warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 5: dev_f1=0.88, f1=0.8275862068965518, best_f1=0.8275862068965518
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.13079461455345154
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")

```

```

epoch 6: dev_f1=0.88, f1=0.8666666666666666, best_f1=0.8275862068965518
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.14868633449077606
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 7: dev_f1=0.8750000000000001, f1=0.8484848484848484,
best_f1=0.8275862068965518
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.02690334990620613
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
    warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 8: dev_f1=0.9032258064516129, f1=0.8484848484848484,
best_f1=0.8484848484848484
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.02735351398587227
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 9: dev_f1=0.88, f1=0.8571428571428571, best_f1=0.8484848484848484
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.007574683520942926
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 10: dev_f1=0.8275862068965518, f1=0.9032258064516129,
best_f1=0.8484848484848484
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.005371555220335722
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 11: dev_f1=0.9032258064516129, f1=0.8484848484848484,

```



```

best_f1=0.8484848484848484
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.004237750545144081
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 12: dev_f1=0.8750000000000001, f1=0.8484848484848484,
best_f1=0.8484848484848484
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0014295863220468163
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 13: dev_f1=0.8666666666666666, f1=0.9333333333333333,
best_f1=0.8484848484848484
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0009547629742883146
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 14: dev_f1=0.8666666666666666, f1=0.9333333333333333,
best_f1=0.8484848484848484
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.008380559273064137
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 15: dev_f1=0.8666666666666666, f1=0.9333333333333333,
best_f1=0.8484848484848484
mv: cannot stat '*_dev.pt': No such file or directory
Selected optimization level 02: FP16 training with FP32 batchnorm and FP32
master weights.

```

Defaults for this optimization level are:

```

enabled           : True
opt_level         : 02
cast_model_type   : torch.float16
patch_torch_functions : False
keep_batchnorm_fp32 : True
master_weights    : True

```

```

loss_scale          : dynamic
Processing user overrides (additional kwargs that are not None)...
After processing overrides, optimization options are:
enabled             : True
opt_level           : 02
cast_model_type     : torch.float16
patch_torch_functions : False
keep_batchnorm_fp32  : True
master_weights      : True
loss_scale          : dynamic
Warning: multi_tensor_applier fused unscale kernel is unavailable, possibly
because apex was installed without --cuda_ext --cpp_ext. Using Python fallback.
Original ImportError was: ModuleNotFoundError("No module named 'amp_C'")
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
91it [00:00, 121477.30it/s]
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
load_f1 = 0.8484848484848484
real_f1 = 0.8484848484848484
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
4398it [00:22, 197.13it/s]

```

### 1.1.2 DBLP-ACM - Running the matcher

```

[ ]: # Step 1:
# train the wdc model with the save_model flag on
# it will produce two files *_dev.pt and *_test.pt which are the
# best checkpoints on the dev set and the test set.
!CUDA_VISIBLE_DEVICES=0 python train_ditto.py \
--task Structured/DBLP-ACM \
--batch_size 32 \
--max_len 256 \
--lr 3e-5 \
--n_epochs 15 \
--finetuning \
--lm roberta \
--fp16 \
--save_model \
--da entry_swap \
--da span_shuffle \
--summarize

```

```

# Step 2: pick a checkpoint, rename it to wdc_all_small.pt
!mv *_dev.pt checkpoints/DBLP-ACM.pt

# Step 3: run the matcher
!CUDA_VISIBLE_DEVICES=0 python matcher.py \
  --task Structured/DBLP-ACM \
  --input_path input/input_small.jsonl \
  --output_path output/output_small.jsonl \
  --lm roberta \
  --use_gpu \
  --fp16 \
  --checkpoint_path checkpoints/

```

Selected optimization level 02: FP16 training with FP32 batchnorm and FP32 master weights.

Defaults for this optimization level are:

```

enabled           : True
opt_level         : 02
cast_model_type   : torch.float16
patch_torch_functions : False
keep_batchnorm_fp32 : True
master_weights    : True
loss_scale        : dynamic

```

Processing user overrides (additional kwargs that are not None)...

After processing overrides, optimization options are:

```

enabled           : True
opt_level         : 02
cast_model_type   : torch.float16
patch_torch_functions : False
keep_batchnorm_fp32 : True
master_weights    : True
loss_scale        : dynamic

```

Warning: multi\_tensor\_applier fused unscale kernel is unavailable, possibly because apex was installed without --cuda\_ext --cpp\_ext. Using Python fallback. Original ImportError was: ModuleNotFoundError("No module named 'amp\_C'")  
/usr/local/lib/python3.7/dist-packages/apex/amp/\_initialize.py:25: UserWarning: An input tensor was not cuda.

```
warnings.warn("An input tensor was not cuda.")
```

```
step: 0, loss: 0.5671074390411377
```

Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 32768.0

```
step: 10, loss: 0.47985145449638367
```

Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0

```
step: 20, loss: 0.4191245138645172
```

```
step: 30, loss: 0.08247772604227066
```

```
step: 40, loss: 0.08271973580121994
```

```

step: 50, loss: 0.032454587519168854
step: 60, loss: 0.03776514530181885
step: 70, loss: 0.13104094564914703
step: 80, loss: 0.021380912512540817
step: 90, loss: 0.09809982776641846
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 8192.0
step: 100, loss: 0.13809894025325775
step: 110, loss: 0.04692845791578293
step: 120, loss: 0.009137021377682686
step: 130, loss: 0.011951196938753128
step: 140, loss: 0.012425188906490803
step: 150, loss: 0.12833093106746674
step: 160, loss: 0.003193509764969349
step: 170, loss: 0.035150788724422455
step: 180, loss: 0.03719522804021835
step: 190, loss: 0.057785823941230774
step: 200, loss: 0.08778048306703568
step: 210, loss: 0.02781258523464203
step: 220, loss: 0.0035418367478996515
step: 230, loss: 0.0012541580945253372
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
    warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 1: dev_f1=0.9799554565701558, f1=0.9842696629213483,
best_f1=0.9842696629213483
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0014790502609685063
step: 10, loss: 0.07919950038194656
step: 20, loss: 0.004347498528659344
step: 30, loss: 0.004712313879281282
step: 40, loss: 0.041969072073698044
step: 50, loss: 0.004023781977593899
step: 60, loss: 0.0034560745116323233
step: 70, loss: 0.004301468376070261
step: 80, loss: 0.05595320463180542
step: 90, loss: 0.013379971496760845
step: 100, loss: 0.005545927211642265
step: 110, loss: 0.0021327603608369827
step: 120, loss: 0.0015976213617250323
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 4096.0
step: 130, loss: 0.054356418550014496
step: 140, loss: 0.003046093275770545

```

```

step: 150, loss: 0.20496129989624023
step: 160, loss: 0.005105345044285059
step: 170, loss: 0.01978541538119316
step: 180, loss: 0.007009084802120924
step: 190, loss: 0.003440568456426263
step: 200, loss: 0.0018805558793246746
step: 210, loss: 0.003201015293598175
step: 220, loss: 0.00044273477396927774
step: 230, loss: 0.0011874429183080792
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
  warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 2: dev_f1=0.9898762654668166, f1=0.9910112359550561,
best_f1=0.9910112359550561
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0025209845043718815
step: 10, loss: 0.004391310270875692
step: 20, loss: 0.012599860318005085
step: 30, loss: 0.001802419195882976
step: 40, loss: 0.009442877024412155
step: 50, loss: 0.005308907013386488
step: 60, loss: 0.0017452433239668608
step: 70, loss: 0.0005996973486617208
step: 80, loss: 0.00025133107556030154
step: 90, loss: 0.001401627785526216
step: 100, loss: 0.0010733773233368993
step: 110, loss: 0.0009086185600608587
step: 120, loss: 0.0003887132916133851
step: 130, loss: 0.001376024098135531
step: 140, loss: 0.0005628651124425232
step: 150, loss: 0.0018320259405300021
step: 160, loss: 0.0006338105886243284
step: 170, loss: 0.0015782007249072194
step: 180, loss: 0.0011708528036251664
step: 190, loss: 0.02744399756193161
step: 200, loss: 0.008189702406525612
step: 210, loss: 0.0053874533623456955
step: 220, loss: 0.008714491501450539
step: 230, loss: 0.0026153556536883116
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")

```

```

/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
  warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 3: dev_f1=0.9921787709497207, f1=0.983277591973244,
best_f1=0.983277591973244
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0014536030357703567
step: 10, loss: 0.0034974899608641863
step: 20, loss: 0.0014745895750820637
step: 30, loss: 0.0008955167722888291
step: 40, loss: 0.030131585896015167
step: 50, loss: 0.0009111224790103734
step: 60, loss: 0.0009702247334644198
step: 70, loss: 0.0006724658887833357
step: 80, loss: 0.0003223354578949511
step: 90, loss: 0.0009691737941466272
step: 100, loss: 0.0008902979898266494
step: 110, loss: 0.0007182049448601902
step: 120, loss: 0.008873582817614079
step: 130, loss: 0.00810712855309248
step: 140, loss: 0.0028454032726585865
step: 150, loss: 0.00034542885259725153
step: 160, loss: 0.007604964543133974
step: 170, loss: 0.00207811058498919
step: 180, loss: 0.18995894491672516
step: 190, loss: 0.005128169432282448
step: 200, loss: 0.005755744408816099
step: 210, loss: 0.0015046655898913741
step: 220, loss: 0.005950813181698322
step: 230, loss: 0.024714544415473938
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
epoch 4: dev_f1=0.9733333333333333, f1=0.9752252252252253,
best_f1=0.983277591973244
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.007973679341375828
step: 10, loss: 0.011620083823800087
step: 20, loss: 0.1403733789920807
step: 30, loss: 0.0044068447314202785
step: 40, loss: 0.0573316365480423
step: 50, loss: 0.0010852683335542679
step: 60, loss: 0.02409256622195244

```

```

step: 70, loss: 0.0007049487321637571
step: 80, loss: 0.015515705570578575
step: 90, loss: 0.067782923579216
step: 100, loss: 0.0008227106300182641
step: 110, loss: 0.0016641629626974463
step: 120, loss: 0.0019089656416326761
step: 130, loss: 0.0007520374492742121
step: 140, loss: 0.0011392179876565933
step: 150, loss: 0.07296255975961685
step: 160, loss: 0.00036219682078808546
step: 170, loss: 0.002609746064990759
step: 180, loss: 0.0017790003912523389
step: 190, loss: 0.010373826138675213
step: 200, loss: 0.001144345267675817
step: 210, loss: 0.07855494320392609
step: 220, loss: 0.0022523216903209686
step: 230, loss: 0.015591231174767017
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
epoch 5: dev_f1=0.9844097995545658, f1=0.9865470852017937,
best_f1=0.983277591973244
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0010677569080144167
step: 10, loss: 0.0012822055723518133
step: 20, loss: 0.005616090726107359
step: 30, loss: 0.00241192989051342
step: 40, loss: 0.0003750717733055353
step: 50, loss: 0.0007891462882980704
step: 60, loss: 0.0012055960251018405
step: 70, loss: 0.01203882321715355
step: 80, loss: 0.000633243762422353
step: 90, loss: 0.03170650452375412
step: 100, loss: 0.0066549996845424175
step: 110, loss: 0.01729119010269642
step: 120, loss: 0.0008226163918152452
step: 130, loss: 0.00028373164241202176
step: 140, loss: 0.00035930160083808005
step: 150, loss: 0.00011312525748508051
step: 160, loss: 0.006749371998012066
step: 170, loss: 0.0001925750111695379
step: 180, loss: 0.0004879772604908794
step: 190, loss: 0.00021996712894178927
step: 200, loss: 0.0029523137491196394
step: 210, loss: 0.0005344763631001115
step: 220, loss: 0.0028859060257673264

```

```

step: 230, loss: 0.0008123806910589337
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 6: dev_f1=0.9887640449438202, f1=0.9865168539325843,
best_f1=0.983277591973244
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0009740086970850825
step: 10, loss: 0.0006638234481215477
step: 20, loss: 0.00485842814669013
step: 30, loss: 0.0004963363171555102
step: 40, loss: 0.001121259992942214
step: 50, loss: 0.0008728244574740529
step: 60, loss: 0.0012674611061811447
step: 70, loss: 0.0006428274209611118
step: 80, loss: 0.00041143319685943425
step: 90, loss: 0.0007134013576433063
step: 100, loss: 0.0004321990709286183
step: 110, loss: 0.0005229822709225118
step: 120, loss: 0.0005170622607693076
step: 130, loss: 0.0012349099852144718
step: 140, loss: 0.0001409387623425573
step: 150, loss: 0.0031023980118334293
step: 160, loss: 0.0005596535629592836
step: 170, loss: 0.0003452827804721892
step: 180, loss: 0.0003638321068137884
step: 190, loss: 0.00024450640194118023
step: 200, loss: 0.0077522010542452335
step: 210, loss: 0.0002597422571852803
step: 220, loss: 0.0003254019538871944
step: 230, loss: 0.0002873638004530221
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 7: dev_f1=0.9898762654668166, f1=0.9909706546275394,
best_f1=0.983277591973244
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0001851027482189238
step: 10, loss: 0.00044253742089495063
step: 20, loss: 0.0017107290914282203
step: 30, loss: 0.0003689018776640296
step: 40, loss: 0.0010575944324955344
step: 50, loss: 0.004365863278508186
step: 60, loss: 0.00042304868111386895

```



```

step: 70, loss: 6.150262197479606e-05
step: 80, loss: 0.05243987590074539
step: 90, loss: 0.0003143465146422386
step: 100, loss: 0.00021519487199839205
step: 110, loss: 0.0003807856992352754
step: 120, loss: 0.002145972801372409
step: 130, loss: 0.0038101417012512684
step: 140, loss: 0.0019400327000766993
step: 150, loss: 0.07538451999425888
step: 160, loss: 0.0007874083239585161
step: 170, loss: 0.012259595096111298
step: 180, loss: 0.0010326518677175045
step: 190, loss: 0.00015990679094102234
step: 200, loss: 0.0021591791883111
step: 210, loss: 0.0009997774614021182
step: 220, loss: 0.0012460681609809399
step: 230, loss: 9.957564907381311e-05
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
epoch 8: dev_f1=0.9899216125419933, f1=0.9865470852017937,
best_f1=0.983277591973244
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 8.466464350931346e-05
step: 10, loss: 0.0007894054288044572
step: 20, loss: 0.00010936307808151469
step: 30, loss: 0.0013765027979388833
step: 40, loss: 7.109118450898677e-05
step: 50, loss: 0.0003239623038098216
step: 60, loss: 3.618448317865841e-05
step: 70, loss: 0.02064158022403717
step: 80, loss: 7.257295510498807e-05
step: 90, loss: 0.03587751463055611
step: 100, loss: 0.0002597889688331634
step: 110, loss: 9.328306623501703e-05
step: 120, loss: 0.021451864391565323
step: 130, loss: 0.0025987867265939713
step: 140, loss: 0.0001598864037077874
step: 150, loss: 0.00033894338412210345
step: 160, loss: 0.0005521757411770523
step: 170, loss: 0.00021950327209196985
step: 180, loss: 4.990581874153577e-05
step: 190, loss: 4.6888813812984154e-05
step: 200, loss: 7.793511758791283e-05
step: 210, loss: 0.00039794945041649044
step: 220, loss: 7.680597627768293e-05

```

```

step: 230, loss: 3.2631469366606325e-05
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 9: dev_f1=0.9910514541387023, f1=0.988814317673378,
best_f1=0.983277591973244
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 7.071581057971343e-05
step: 10, loss: 0.00016266433522105217
step: 20, loss: 0.0003131897246930748
step: 30, loss: 8.547033212380484e-05
step: 40, loss: 0.0006272830069065094
step: 50, loss: 0.0003457460552453995
step: 60, loss: 5.764341403846629e-05
step: 70, loss: 0.0002346957044210285
step: 80, loss: 3.597064642235637e-05
step: 90, loss: 5.765317473560572e-05
step: 100, loss: 4.9458525609225035e-05
step: 110, loss: 0.00020577746909111738
step: 120, loss: 7.010247645666823e-05
step: 130, loss: 5.0114584155380726e-05
step: 140, loss: 5.3114457841729745e-05
step: 150, loss: 0.0003092144033871591
step: 160, loss: 3.3970569347729906e-05
step: 170, loss: 3.562862548278645e-05
step: 180, loss: 0.00016986358969006687
step: 190, loss: 4.3653781176544726e-05
step: 200, loss: 9.455094550503418e-05
step: 210, loss: 0.0002745024103205651
step: 220, loss: 7.895485032349825e-05
step: 230, loss: 3.0366452847374603e-05
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 10: dev_f1=0.9910514541387023, f1=0.988814317673378,
best_f1=0.983277591973244
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 5.19760069437325e-05
step: 10, loss: 4.684944360633381e-05
step: 20, loss: 2.9421215003822e-05
step: 30, loss: 2.9592458304250613e-05
step: 40, loss: 1.671854806772899e-05
step: 50, loss: 0.00018230854766443372
step: 60, loss: 0.017095211893320084

```

```

step: 70, loss: 0.0004995561903342605
step: 80, loss: 0.0029265331104397774
step: 90, loss: 0.03620157763361931
step: 100, loss: 0.00020328420214354992
step: 110, loss: 0.0025265098083764315
step: 120, loss: 0.00011799686762969941
step: 130, loss: 0.00013340491568669677
step: 140, loss: 0.005080714356154203
step: 150, loss: 0.00010540850780671462
step: 160, loss: 0.04229341447353363
step: 170, loss: 0.0005135167739354074
step: 180, loss: 4.790304956259206e-05
step: 190, loss: 0.00022509184782393277
step: 200, loss: 0.0016430513933300972
step: 210, loss: 5.978989429422654e-05
step: 220, loss: 9.068589861271903e-05
step: 230, loss: 6.407829641830176e-05
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
  warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 11: dev_f1=0.9932735426008968, f1=0.9887892376681614,
best_f1=0.9887892376681614
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00013676678645424545
step: 10, loss: 4.516767876339145e-05
step: 20, loss: 0.023036373779177666
step: 30, loss: 0.043101295828819275
step: 40, loss: 4.544468174572103e-05
step: 50, loss: 7.369848026428372e-05
step: 60, loss: 3.929965168936178e-05
step: 70, loss: 5.0044636736856773e-05
step: 80, loss: 4.338146391091868e-05
step: 90, loss: 0.0005507634487003088
step: 100, loss: 1.8897902918979526e-05
step: 110, loss: 3.020956319232937e-05
step: 120, loss: 2.1960058802505955e-05
step: 130, loss: 2.3301005057874136e-05
step: 140, loss: 6.195006426423788e-05
step: 150, loss: 4.420828918227926e-05
step: 160, loss: 0.00014409341383725405
step: 170, loss: 0.0004754573747050017
step: 180, loss: 4.1579187382012606e-05

```

```

step: 190, loss: 0.007906164973974228
step: 200, loss: 2.367302295169793e-05
step: 210, loss: 0.0008293139399029315
step: 220, loss: 0.010203195735812187
step: 230, loss: 0.000113119043817278
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
epoch 12: dev_f1=0.9932735426008968, f1=0.9876543209876544,
best_f1=0.9887892376681614
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 3.9723290683468804e-05
step: 10, loss: 0.003101753070950508
step: 20, loss: 3.225559339625761e-05
step: 30, loss: 0.0001386810326948762
step: 40, loss: 8.805269317235798e-05
step: 50, loss: 0.0008769120322540402
step: 60, loss: 6.071915777283721e-05
step: 70, loss: 3.495404598652385e-05
step: 80, loss: 3.568766260286793e-05
step: 90, loss: 2.2321022697724402e-05
step: 100, loss: 4.8209938540821895e-05
step: 110, loss: 4.004605216323398e-05
step: 120, loss: 2.4839297111611813e-05
step: 130, loss: 3.260925950598903e-05
step: 140, loss: 2.2689850084134378e-05
step: 150, loss: 1.579482341185212e-05
step: 160, loss: 0.00816418882459402
step: 170, loss: 1.8383943825028837e-05
step: 180, loss: 0.030757732689380646
step: 190, loss: 2.8001706596114673e-05
step: 200, loss: 1.1637513125606347e-05
step: 210, loss: 7.131808524718508e-05
step: 220, loss: 2.0782750652870163e-05
step: 230, loss: 2.349052192585077e-05
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
epoch 13: dev_f1=0.9932735426008968, f1=0.9865168539325843,
best_f1=0.9887892376681614
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 2.9310887839528732e-05
step: 10, loss: 2.9258681024657562e-05
step: 20, loss: 2.261164627270773e-05

```

```

step: 30, loss: 3.6181274481350556e-05
step: 40, loss: 6.11483192187734e-05
step: 50, loss: 1.598463495611213e-05
step: 60, loss: 2.199315895268228e-05
step: 70, loss: 2.0760466213687323e-05
step: 80, loss: 2.770204264379572e-05
step: 90, loss: 0.0003355088410899043
step: 100, loss: 6.892907549627125e-05
step: 110, loss: 2.5610455850255676e-05
step: 120, loss: 1.1644980986602604e-05
step: 130, loss: 4.663681465899572e-05
step: 140, loss: 0.00011534363875398412
step: 150, loss: 2.9476001145667396e-05
step: 160, loss: 2.509615478629712e-05
step: 170, loss: 2.6555953809292987e-05
step: 180, loss: 1.8156648366129957e-05
step: 190, loss: 1.758282996888738e-05
step: 200, loss: 2.9218628696980886e-05
step: 210, loss: 2.8168271455797367e-05
step: 220, loss: 2.6042545869131573e-05
step: 230, loss: 0.0002735708258114755
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 14: dev_f1=0.9921612541993281, f1=0.9876543209876544,
best_f1=0.9887892376681614
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.028113827109336853
step: 10, loss: 2.90669959213119e-05
step: 20, loss: 9.2992493591737e-05
step: 30, loss: 1.672618236625567e-05
step: 40, loss: 1.348899877484655e-05
step: 50, loss: 1.714332574920263e-05
step: 60, loss: 0.03815336152911186
step: 70, loss: 2.6522751795710064e-05
step: 80, loss: 1.678565786278341e-05
step: 90, loss: 1.6562242308282293e-05
step: 100, loss: 1.4498406017082743e-05
step: 110, loss: 0.0003836562391370535
step: 120, loss: 0.03429282456636429
step: 130, loss: 9.973470150725916e-05
step: 140, loss: 0.017204012721776962
step: 150, loss: 3.735154677997343e-05
step: 160, loss: 0.009615052491426468
step: 170, loss: 1.3284081433084793e-05
step: 180, loss: 2.436232716718223e-05

```

```

step: 190, loss: 0.00017512566410005093
step: 200, loss: 0.00036662063212133944
step: 210, loss: 0.03224378079175949
step: 220, loss: 1.8279042706126347e-05
step: 230, loss: 2.370304355281405e-05
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 15: dev_f1=0.9921612541993281, f1=0.9876543209876544,
best_f1=0.9887892376681614
mv: cannot stat '*_dev.pt': No such file or directory
Selected optimization level 02:  FP16 training with FP32 batchnorm and FP32
master weights.

Defaults for this optimization level are:
enabled                : True
opt_level              : 02
cast_model_type        : torch.float16
patch_torch_functions  : False
keep_batchnorm_fp32    : True
master_weights         : True
loss_scale             : dynamic
Processing user overrides (additional kwargs that are not None)...
After processing overrides, optimization options are:
enabled                : True
opt_level              : 02
cast_model_type        : torch.float16
patch_torch_functions  : False
keep_batchnorm_fp32    : True
master_weights         : True
loss_scale             : dynamic
Warning:  multi_tensor_applier fused unscale kernel is unavailable, possibly
because apex was installed without --cuda_ext --cpp_ext. Using Python fallback.
Original ImportError was: ModuleNotFoundError("No module named 'amp_C'")
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
2473it [00:12, 193.02it/s]
load_f1 = 0.9910112359550561
real_f1 = 0.9898989898989898
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
4398it [00:24, 175.95it/s]

```

### 1.1.3 DBLP-GoogleScholar - Running the matcher

```
[ ]: # Step 1:
# train the wdc model with the save_model flag on
# it will produce two files *_dev.pt and *_test.pt which are the
# best checkpoints on the dev set and the test set.
!CUDA_VISIBLE_DEVICES=0 python train_ditto.py \
  --task Structured/DBLP-GoogleScholar \
  --batch_size 32 \
  --max_len 256 \
  --lr 3e-5 \
  --n_epochs 15 \
  --finetuning \
  --lm roberta \
  --fp16 \
  --save_model \
  --da entry_swap \
  --da span_shuffle \
  --summarize

# Step 2: pick a checkpoint, rename it to wdc_all_small.pt
!mv *_dev.pt checkpoints/DBLP-GoogleScholar.pt

# Step 3: run the matcher
!CUDA_VISIBLE_DEVICES=0 python matcher.py \
  --task Structured/DBLP-GoogleScholar \
  --input_path input/input_small.jsonl \
  --output_path output/output_small.jsonl \
  --lm roberta \
  --use_gpu \
  --fp16 \
  --checkpoint_path checkpoints/
```

Selected optimization level 02: FP16 training with FP32 batchnorm and FP32 master weights.

Defaults for this optimization level are:

```
enabled           : True
opt_level         : 02
cast_model_type   : torch.float16
patch_torch_functions : False
keep_batchnorm_fp32 : True
master_weights    : True
loss_scale        : dynamic
```

Processing user overrides (additional kwargs that are not None)...

After processing overrides, optimization options are:

```
enabled           : True
```

```

opt_level                : 02
cast_model_type          : torch.float16
patch_torch_functions    : False
keep_batchnorm_fp32      : True
master_weights           : True
loss_scale               : dynamic
Warning: multi_tensor_applier fused unscale kernel is unavailable, possibly
because apex was installed without --cuda_ext --cpp_ext. Using Python fallback.
Original ImportError was: ModuleNotFoundError("No module named 'amp_C'")
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.6191593408584595
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 32768.0
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 10, loss: 0.4077439606189728
step: 20, loss: 0.2822803556919098
step: 30, loss: 0.2271384596824646
step: 40, loss: 0.17222130298614502
step: 50, loss: 0.05823099613189697
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 8192.0
step: 60, loss: 0.11382749676704407
step: 70, loss: 0.21582360565662384
step: 80, loss: 0.07141651213169098
step: 90, loss: 0.216706782579422
step: 100, loss: 0.09077068418264389
step: 110, loss: 0.050569187849760056
step: 120, loss: 0.1847688853740692
step: 130, loss: 0.16977937519550323
step: 140, loss: 0.2260921597480774
step: 150, loss: 0.12325126677751541
step: 160, loss: 0.07444331049919128
step: 170, loss: 0.013755100779235363
step: 180, loss: 0.03048725798726082
step: 190, loss: 0.07170520722866058
step: 200, loss: 0.05887242406606674
step: 210, loss: 0.031485188752412796
step: 220, loss: 0.07918784767389297
step: 230, loss: 0.21517588198184967
step: 240, loss: 0.026539362967014313
step: 250, loss: 0.02275099977850914
step: 260, loss: 0.07803288847208023
step: 270, loss: 0.4112647771835327
step: 280, loss: 0.028572341427206993
step: 290, loss: 0.06973178684711456
step: 300, loss: 0.023601120337843895
step: 310, loss: 0.04392337426543236
step: 320, loss: 0.026726024225354195

```



```

step: 330, loss: 0.10781388729810715
step: 340, loss: 0.2762228846549988
step: 350, loss: 0.029508111998438835
step: 360, loss: 0.1253817081451416
step: 370, loss: 0.06336256861686707
step: 380, loss: 0.25487884879112244
step: 390, loss: 0.04456920549273491
step: 400, loss: 0.08092498034238815
step: 410, loss: 0.2543220818042755
step: 420, loss: 0.028635751456022263
step: 430, loss: 0.020116152241826057
step: 440, loss: 0.026196612045168877
step: 450, loss: 0.016138678416609764
step: 460, loss: 0.003949678037315607
step: 470, loss: 0.010208545252680779
step: 480, loss: 0.13772085309028625
step: 490, loss: 0.08686598390340805
step: 500, loss: 0.008031654171645641
step: 510, loss: 0.035160139203071594
step: 520, loss: 0.03854973241686821
step: 530, loss: 0.020906923338770866
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
    warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 1: dev_f1=0.945707656612529, f1=0.9472710453283996,
best_f1=0.9472710453283996
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0622544027864933
step: 10, loss: 0.03738543018698692
step: 20, loss: 0.030048472806811333
step: 30, loss: 0.16834095120429993
step: 40, loss: 0.03423043340444565
step: 50, loss: 0.04560071974992752
step: 60, loss: 0.011592110618948936
step: 70, loss: 0.032029084861278534
step: 80, loss: 0.021775778383016586
step: 90, loss: 0.010118266567587852
step: 100, loss: 0.10779895633459091
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 4096.0
step: 110, loss: 0.0036188499070703983
step: 120, loss: 0.01990727335214615
step: 130, loss: 0.006092436145991087

```

```

step: 140, loss: 0.03408123552799225
step: 150, loss: 0.021761784330010414
step: 160, loss: 0.07503492385149002
step: 170, loss: 0.021459560841321945
step: 180, loss: 0.023388901725411415
step: 190, loss: 0.01119261048734188
step: 200, loss: 0.05087420716881752
step: 210, loss: 0.03548319265246391
step: 220, loss: 0.0012431828072294593
step: 230, loss: 0.11710131168365479
step: 240, loss: 0.033468473702669144
step: 250, loss: 0.010192517191171646
step: 260, loss: 0.009157455526292324
step: 270, loss: 0.0028391415253281593
step: 280, loss: 0.046937234699726105
step: 290, loss: 0.024005575105547905
step: 300, loss: 0.048732321709394455
step: 310, loss: 0.012182306498289108
step: 320, loss: 0.06412695348262787
step: 330, loss: 0.012193466536700726
step: 340, loss: 0.09608493000268936
step: 350, loss: 0.0023252503015100956
step: 360, loss: 0.10053453594446182
step: 370, loss: 0.003242039354518056
step: 380, loss: 0.2015904039144516
step: 390, loss: 0.005677428562194109
step: 400, loss: 0.0335179939866066
step: 410, loss: 0.002850604709237814
step: 420, loss: 0.04292679205536842
step: 430, loss: 0.2991647720336914
step: 440, loss: 0.015391753986477852
step: 450, loss: 0.013987390324473381
step: 460, loss: 0.02886856719851494
step: 470, loss: 0.015376484952867031
step: 480, loss: 0.003431477816775441
step: 490, loss: 0.019990963861346245
step: 500, loss: 0.000640292011667043
step: 510, loss: 0.024805238470435143
step: 520, loss: 0.23318715393543243
step: 530, loss: 0.12358393520116806
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
  warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 2: dev_f1=0.9585253456221198, f1=0.9524697110904008,

```

```
best_f1=0.9524697110904008
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.13245941698551178
step: 10, loss: 0.039170727133750916
step: 20, loss: 0.005329549312591553
step: 30, loss: 0.028677038848400116
step: 40, loss: 0.03605968877673149
step: 50, loss: 0.007922032848000526
step: 60, loss: 0.004578847438097
step: 70, loss: 0.0077791898511350155
step: 80, loss: 0.12006475031375885
step: 90, loss: 0.04562300071120262
step: 100, loss: 0.02579054795205593
step: 110, loss: 0.04356274753808975
step: 120, loss: 0.12804241478443146
step: 130, loss: 0.030684716999530792
step: 140, loss: 0.0015957978321239352
step: 150, loss: 0.025450704619288445
step: 160, loss: 0.017757659777998924
step: 170, loss: 0.01097193919122219
step: 180, loss: 0.009060442447662354
step: 190, loss: 0.0008119305712170899
step: 200, loss: 0.01052272878587246
step: 210, loss: 0.03671496734023094
step: 220, loss: 0.0966002494096756
step: 230, loss: 0.03767688572406769
step: 240, loss: 0.010131198912858963
step: 250, loss: 0.034599144011735916
step: 260, loss: 0.08069316297769547
step: 270, loss: 0.003056152258068323
step: 280, loss: 0.000703713740222156
step: 290, loss: 0.004434501752257347
step: 300, loss: 0.16193906962871552
step: 310, loss: 0.039028339087963104
step: 320, loss: 0.1453939974308014
step: 330, loss: 0.0014936462976038456
step: 340, loss: 0.00471483962610364
step: 350, loss: 0.16425460577011108
step: 360, loss: 0.009729267098009586
step: 370, loss: 0.023503076285123825
step: 380, loss: 0.01478164829313755
step: 390, loss: 0.002803854178637266
step: 400, loss: 0.14544939994812012
step: 410, loss: 0.19501975178718567
step: 420, loss: 0.015247012488543987
step: 430, loss: 0.014390232041478157
```

```

step: 440, loss: 0.21302072703838348
step: 450, loss: 0.05321158096194267
step: 460, loss: 0.06302308291196823
step: 470, loss: 0.015605553984642029
step: 480, loss: 0.02001015469431877
step: 490, loss: 0.06801258772611618
step: 500, loss: 0.0338791199028492
step: 510, loss: 0.029169591143727303
step: 520, loss: 0.0032172598876059055
step: 530, loss: 0.011478235013782978
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
    warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 3: dev_f1=0.96040987424313, f1=0.9537166900420756,
best_f1=0.9537166900420756
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.011437208391726017
step: 10, loss: 0.002331034978851676
step: 20, loss: 0.006784633733332157
step: 30, loss: 0.15619319677352905
step: 40, loss: 0.006288479547947645
step: 50, loss: 0.10746195912361145
step: 60, loss: 0.024632902815937996
step: 70, loss: 0.03681943565607071
step: 80, loss: 0.08358553051948547
step: 90, loss: 0.007943171076476574
step: 100, loss: 0.0023600482381880283
step: 110, loss: 0.028566481545567513
step: 120, loss: 0.002778715454041958
step: 130, loss: 0.01700524240732193
step: 140, loss: 0.015404419973492622
step: 150, loss: 0.0006439434364438057
step: 160, loss: 0.027032658457756042
step: 170, loss: 0.05819905921816826
step: 180, loss: 0.09496992081403732
step: 190, loss: 0.02028956450521946
step: 200, loss: 0.03173322603106499
step: 210, loss: 0.0008565220632590353
step: 220, loss: 0.10952343046665192
step: 230, loss: 0.009347244165837765
step: 240, loss: 0.011397207155823708
step: 250, loss: 0.11032754927873611

```

```

step: 260, loss: 0.0036099832504987717
step: 270, loss: 0.05051577091217041
step: 280, loss: 0.010908281430602074
step: 290, loss: 0.02542947232723236
step: 300, loss: 0.0021855521481484175
step: 310, loss: 0.0015882233856245875
step: 320, loss: 0.022616883739829063
step: 330, loss: 0.045918866991996765
step: 340, loss: 0.008651601150631905
step: 350, loss: 0.06068083643913269
step: 360, loss: 0.03541715443134308
step: 370, loss: 0.0010002526687458158
step: 380, loss: 0.03246420621871948
step: 390, loss: 0.0006435333052650094
step: 400, loss: 0.0014079087413847446
step: 410, loss: 0.000581505533773452
step: 420, loss: 0.006551777943968773
step: 430, loss: 0.0018284249817952514
step: 440, loss: 0.009291908703744411
step: 450, loss: 0.010525413788855076
step: 460, loss: 0.0012847322504967451
step: 470, loss: 0.0012897135457023978
step: 480, loss: 0.003740502055734396
step: 490, loss: 0.0016089846612885594
step: 500, loss: 0.0787815973162651
step: 510, loss: 0.046659354120492935
step: 520, loss: 0.008407559245824814
step: 530, loss: 0.10627290606498718
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
    warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 4: dev_f1=0.9635514018691589, f1=0.9598880597014925,
best_f1=0.9598880597014925
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0010223883436992764
step: 10, loss: 0.012324441224336624
step: 20, loss: 0.001257707248441875
step: 30, loss: 0.004582146182656288
step: 40, loss: 0.0006315391510725021
step: 50, loss: 0.13918772339820862
step: 60, loss: 0.043390627950429916
step: 70, loss: 0.00132843351457268

```

step: 80, loss: 0.00035004367236979306  
step: 90, loss: 0.038693856447935104  
step: 100, loss: 0.002114871982485056  
step: 110, loss: 0.03437766060233116  
step: 120, loss: 0.14639519155025482  
step: 130, loss: 0.0169671643525362  
step: 140, loss: 0.0015737989451736212  
step: 150, loss: 0.023864492774009705  
step: 160, loss: 0.006262865848839283  
step: 170, loss: 0.024350155144929886  
step: 180, loss: 0.0006836349493823946  
step: 190, loss: 0.003105561248958111  
step: 200, loss: 0.003658756148070097  
step: 210, loss: 0.0005043565179221332  
step: 220, loss: 0.005528157111257315  
step: 230, loss: 0.004709419794380665  
step: 240, loss: 0.0013977562775835395  
step: 250, loss: 0.1932898312807083  
step: 260, loss: 0.004308817442506552  
step: 270, loss: 0.025770002976059914  
step: 280, loss: 0.0021519246511161327  
step: 290, loss: 0.06113787740468979  
step: 300, loss: 0.03648833930492401  
step: 310, loss: 0.014231943525373936  
step: 320, loss: 0.005543784238398075  
step: 330, loss: 0.0009773862548172474  
step: 340, loss: 0.0030040668789297342  
step: 350, loss: 0.021148785948753357  
step: 360, loss: 0.0001904723176266998  
step: 370, loss: 0.004359767772257328  
step: 380, loss: 0.000909928698092699  
step: 390, loss: 0.011286560446023941  
step: 400, loss: 0.0004221255367156118  
step: 410, loss: 0.020207226276397705  
step: 420, loss: 0.157780721783638  
step: 430, loss: 0.03237021714448929  
step: 440, loss: 0.00018952581740450114  
step: 450, loss: 0.007381519768387079  
step: 460, loss: 0.011201058514416218  
step: 470, loss: 0.027140872552990913  
step: 480, loss: 0.019635507836937904  
step: 490, loss: 0.00274753849953413  
step: 500, loss: 0.05012281984090805  
step: 510, loss: 0.0035918389912694693  
step: 520, loss: 0.18301405012607574  
step: 530, loss: 0.0079826470464468  
/usr/local/lib/python3.7/dist-packages/apex/amp/\_initialize.py:25: UserWarning:  
An input tensor was not cuda.

```
warnings.warn("An input tensor was not cuda.")
epoch 5: dev_f1=0.9606373008434864, f1=0.9529137529137528,
best_f1=0.9598880597014925
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.019215868785977364
step: 10, loss: 0.0005271111149340868
step: 20, loss: 0.0002752312575466931
step: 30, loss: 0.0006090879905968904
step: 40, loss: 0.0006556178559549153
step: 50, loss: 0.0002725361555349082
step: 60, loss: 0.002879456151276827
step: 70, loss: 0.0003714909835252911
step: 80, loss: 0.0004850046243518591
step: 90, loss: 0.0015332824550569057
step: 100, loss: 0.004028681665658951
step: 110, loss: 0.0011646529892459512
step: 120, loss: 0.00349681219086051
step: 130, loss: 0.00031037861481308937
step: 140, loss: 0.0004711270739790052
step: 150, loss: 0.0005813258467242122
step: 160, loss: 0.014890192076563835
step: 170, loss: 0.00030825682915747166
step: 180, loss: 0.0005041901022195816
step: 190, loss: 0.003786448622122407
step: 200, loss: 0.0006652909214608371
step: 210, loss: 0.0015269223367795348
step: 220, loss: 0.007198486942797899
step: 230, loss: 0.00636103842407465
step: 240, loss: 0.007149077486246824
step: 250, loss: 0.017274145036935806
step: 260, loss: 0.000757101341150701
step: 270, loss: 0.0030151340179145336
step: 280, loss: 0.0023263394832611084
step: 290, loss: 0.00282531906850636
step: 300, loss: 0.004456549417227507
step: 310, loss: 0.021306248381733894
step: 320, loss: 3.2717543945182115e-05
step: 330, loss: 0.0001317142741754651
step: 340, loss: 0.0007494565797969699
step: 350, loss: 0.0014136886456981301
step: 360, loss: 0.001807152177207172
step: 370, loss: 0.0003362861170899123
step: 380, loss: 0.0010576414642855525
step: 390, loss: 0.0172821544110775
step: 400, loss: 0.0015549804084002972
step: 410, loss: 0.0002982633304782212
```

```

step: 420, loss: 0.0029753181152045727
step: 430, loss: 0.00019607541617006063
step: 440, loss: 0.00011348522821208462
step: 450, loss: 0.1375592201948166
step: 460, loss: 0.008931919001042843
step: 470, loss: 0.002682164777070284
step: 480, loss: 0.006253768224269152
step: 490, loss: 0.005622431170195341
step: 500, loss: 0.007048876490443945
step: 510, loss: 0.09934942424297333
step: 520, loss: 0.0063791461288928986
step: 530, loss: 0.005193713121116161
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
epoch 6: dev_f1=0.9595912679981421, f1=0.9525581395348838,
best_f1=0.9598880597014925
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.03323918953537941
step: 10, loss: 0.0009675824549049139
step: 20, loss: 0.0006614711601287127
step: 30, loss: 0.010823414660990238
step: 40, loss: 0.010050254873931408
step: 50, loss: 0.0017017771024256945
step: 60, loss: 0.0011431341990828514
step: 70, loss: 0.00032781073241494596
step: 80, loss: 0.00031673209741711617
step: 90, loss: 7.799198647262529e-05
step: 100, loss: 0.00011826406262116507
step: 110, loss: 4.195240035187453e-05
step: 120, loss: 0.0001782144099706784
step: 130, loss: 8.225647616200149e-05
step: 140, loss: 0.0001808380038710311
step: 150, loss: 0.0013603012776002288
step: 160, loss: 7.056831964291632e-05
step: 170, loss: 0.06472835689783096
step: 180, loss: 0.0003264379338361323
step: 190, loss: 0.0650622546672821
step: 200, loss: 7.235106750158593e-05
step: 210, loss: 0.0009849043563008308
step: 220, loss: 4.621763582690619e-05
step: 230, loss: 3.298169031040743e-05
step: 240, loss: 0.0003843351441901177
step: 250, loss: 0.002384065417572856
step: 260, loss: 0.002731796819716692
step: 270, loss: 9.081375173991546e-05

```



```

step: 280, loss: 0.014379402622580528
step: 290, loss: 0.0008694474236108363
step: 300, loss: 0.00010272972576785833
step: 310, loss: 0.0014049311866983771
step: 320, loss: 0.035822544246912
step: 330, loss: 0.00255437265150249
step: 340, loss: 6.62117381580174e-05
step: 350, loss: 0.0004407570813782513
step: 360, loss: 0.07556220889091492
step: 370, loss: 0.017817554995417595
step: 380, loss: 0.00327162048779428
step: 390, loss: 0.0017082785489037633
step: 400, loss: 0.007195282727479935
step: 410, loss: 0.0007894678274169564
step: 420, loss: 0.07037287205457687
step: 430, loss: 0.00013432223931886256
step: 440, loss: 0.004557319451123476
step: 450, loss: 0.014094260521233082
step: 460, loss: 0.0006656795158050954
step: 470, loss: 0.10038387030363083
step: 480, loss: 0.0055579678155481815
step: 490, loss: 0.0082769226282835
step: 500, loss: 0.00022255268413573503
step: 510, loss: 8.023736882023513e-05
step: 520, loss: 5.0733931857394055e-05
step: 530, loss: 0.0011438550427556038
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
epoch 7: dev_f1=0.9626517273576097, f1=0.9552656104380243,
best_f1=0.9598880597014925
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.003391494508832693
step: 10, loss: 0.002417145064100623
step: 20, loss: 7.400538015644997e-05
step: 30, loss: 4.9692833272274584e-05
step: 40, loss: 2.9141783670638688e-05
step: 50, loss: 0.00013006599328946322
step: 60, loss: 5.099218469695188e-05
step: 70, loss: 0.0004260367713868618
step: 80, loss: 0.006066943984478712
step: 90, loss: 0.0001302035088883713
step: 100, loss: 5.276675074128434e-05
step: 110, loss: 0.0012824288569390774
step: 120, loss: 0.00084951106691733
step: 130, loss: 0.00034704094287008047

```

```

step: 140, loss: 6.155134178698063e-05
step: 150, loss: 0.0004873658181168139
step: 160, loss: 0.006670610513538122
step: 170, loss: 0.01324212085455656
step: 180, loss: 0.0011302211787551641
step: 190, loss: 0.0025089161936193705
step: 200, loss: 0.004553694278001785
step: 210, loss: 0.08381740003824234
step: 220, loss: 0.00010396636935183778
step: 230, loss: 0.06909656524658203
step: 240, loss: 0.0034617960918694735
step: 250, loss: 0.001650583348236978
step: 260, loss: 0.0002983365557156503
step: 270, loss: 0.022579409182071686
step: 280, loss: 0.00043945939978584647
step: 290, loss: 0.00010096792539115995
step: 300, loss: 0.008981972932815552
step: 310, loss: 0.0006345713045448065
step: 320, loss: 0.013219457119703293
step: 330, loss: 0.0005758090992458165
step: 340, loss: 0.02241913229227066
step: 350, loss: 0.00015499917208217084
step: 360, loss: 0.0007931822910904884
step: 370, loss: 0.03604559600353241
step: 380, loss: 3.591665154090151e-05
step: 390, loss: 0.009377666749060154
step: 400, loss: 0.005987972021102905
step: 410, loss: 0.07254276424646378
step: 420, loss: 0.0001514885516371578
step: 430, loss: 0.002629674505442381
step: 440, loss: 0.004306665156036615
step: 450, loss: 9.123738709604368e-05
step: 460, loss: 0.09247393906116486
step: 470, loss: 0.10044418275356293
step: 480, loss: 0.0019602146930992603
step: 490, loss: 0.05231654644012451
step: 500, loss: 0.001615294604562223
step: 510, loss: 0.006273510865867138
step: 520, loss: 6.78756259731017e-05
step: 530, loss: 0.00011326384264975786
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
epoch 8: dev_f1=0.957169459962756, f1=0.9574468085106383,
best_f1=0.9598880597014925
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")

```

step: 0, loss: 8.194973634090275e-05  
step: 10, loss: 0.006800548173487186  
step: 20, loss: 0.00034635732299648225  
step: 30, loss: 0.0631413385272026  
step: 40, loss: 7.311847730306908e-05  
step: 50, loss: 0.0031437973957508802  
step: 60, loss: 0.00075108977034688  
step: 70, loss: 0.0014141822466626763  
step: 80, loss: 0.02177572250366211  
step: 90, loss: 0.06467873603105545  
step: 100, loss: 0.0004375424177851528  
step: 110, loss: 0.0007266577449627221  
step: 120, loss: 0.00047592518967576325  
step: 130, loss: 8.222572796512395e-05  
step: 140, loss: 4.910404823021963e-05  
step: 150, loss: 0.00044609943870455027  
step: 160, loss: 0.0002590710064396262  
step: 170, loss: 0.0045014070346951485  
step: 180, loss: 0.005618065595626831  
step: 190, loss: 9.036753908731043e-05  
step: 200, loss: 0.0008250775863416493  
step: 210, loss: 0.018854480236768723  
step: 220, loss: 0.0020374897867441177  
step: 230, loss: 0.000444055040134117  
step: 240, loss: 0.0017668860964477062  
step: 250, loss: 8.55286925798282e-05  
step: 260, loss: 0.0037090093828737736  
step: 270, loss: 0.00040302364504896104  
step: 280, loss: 0.023401236161589622  
step: 290, loss: 0.0003007986815646291  
step: 300, loss: 0.00020870623120572418  
step: 310, loss: 0.004512799438089132  
step: 320, loss: 0.0010202855337411165  
step: 330, loss: 0.0001368932134937495  
step: 340, loss: 0.007350650615990162  
step: 350, loss: 0.011162502691149712  
step: 360, loss: 0.006780368275940418  
step: 370, loss: 0.0001420416811015457  
step: 380, loss: 0.02047775499522686  
step: 390, loss: 0.0010228788014501333  
step: 400, loss: 0.015018254518508911  
step: 410, loss: 0.00012136146688135341  
step: 420, loss: 7.530189759563655e-05  
step: 430, loss: 0.013425460085272789  
step: 440, loss: 4.0602735680295154e-05  
step: 450, loss: 6.886901974212378e-05  
step: 460, loss: 2.5927196475095116e-05  
step: 470, loss: 0.0004708735505118966

```

step: 480, loss: 0.0002219370799139142
step: 490, loss: 7.407293014694005e-05
step: 500, loss: 0.0003647859557531774
step: 510, loss: 0.00013470729754772037
step: 520, loss: 0.0015770515892654657
step: 530, loss: 4.618244565790519e-05
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 9: dev_f1=0.9526963103122044, f1=0.947417840375587,
best_f1=0.9598880597014925
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00023575240629725158
step: 10, loss: 0.0014187884517014027
step: 20, loss: 0.0011460725218057632
step: 30, loss: 0.001131319673731923
step: 40, loss: 0.03713192790746689
step: 50, loss: 2.9595885280286893e-05
step: 60, loss: 0.001962811453267932
step: 70, loss: 6.944450433366e-05
step: 80, loss: 7.449885015375912e-05
step: 90, loss: 2.0078599845874123e-05
step: 100, loss: 0.0004402605118229985
step: 110, loss: 0.039727963507175446
step: 120, loss: 2.1434516384033486e-05
step: 130, loss: 2.4704875613679178e-05
step: 140, loss: 5.311744462233037e-05
step: 150, loss: 0.0006248124991543591
step: 160, loss: 0.04058718681335449
step: 170, loss: 4.598492159857415e-05
step: 180, loss: 7.973409083206207e-05
step: 190, loss: 1.2490721928770654e-05
step: 200, loss: 4.19801363022998e-05
step: 210, loss: 0.0025693143252283335
step: 220, loss: 0.0009735928615555167
step: 230, loss: 3.0451687052845955e-05
step: 240, loss: 6.290714372880757e-05
step: 250, loss: 0.0011779136257246137
step: 260, loss: 0.002959741046652198
step: 270, loss: 2.1080797523609363e-05
step: 280, loss: 0.015714775770902634
step: 290, loss: 0.00023326512018684298
step: 300, loss: 0.002357285004109144
step: 310, loss: 0.06749503314495087
step: 320, loss: 0.010625123977661133
step: 330, loss: 0.00027407691231928766

```

```

step: 340, loss: 5.986438554828055e-05
step: 350, loss: 0.0034122623037546873
step: 360, loss: 0.00014166266191750765
step: 370, loss: 0.00015032049850560725
step: 380, loss: 0.0018999746534973383
step: 390, loss: 6.32200826657936e-05
step: 400, loss: 2.568411400716286e-05
step: 410, loss: 0.00011530292977113277
step: 420, loss: 2.182553907914553e-05
step: 430, loss: 2.023469642153941e-05
step: 440, loss: 2.9769133107038215e-05
step: 450, loss: 0.01925039105117321
step: 460, loss: 4.151804751018062e-05
step: 470, loss: 0.0036734258756041527
step: 480, loss: 2.9201823053881526e-05
step: 490, loss: 5.2745090215466917e-05
step: 500, loss: 0.004743392113596201
step: 510, loss: 5.717982639907859e-05
step: 520, loss: 1.861059536167886e-05
step: 530, loss: 0.015376479364931583
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
epoch 10: dev_f1=0.9621672115833722, f1=0.9572490706319703,
best_f1=0.9598880597014925
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 2.4436898456769995e-05
step: 10, loss: 3.8809081161161885e-05
step: 20, loss: 0.002818276872858405
step: 30, loss: 2.006012982747052e-05
step: 40, loss: 2.3769633116899058e-05
step: 50, loss: 1.8186356101068668e-05
step: 60, loss: 5.7596535043558106e-05
step: 70, loss: 2.545401184761431e-05
step: 80, loss: 5.404738840297796e-05
step: 90, loss: 0.0008121664286591113
step: 100, loss: 0.000253313104622066
step: 110, loss: 2.5368228307343088e-05
step: 120, loss: 5.188398790778592e-05
step: 130, loss: 0.030257567763328552
step: 140, loss: 1.442403117835056e-05
step: 150, loss: 0.011192210018634796
step: 160, loss: 0.001441397238522768
step: 170, loss: 0.00036961573641747236
step: 180, loss: 1.2885556316177826e-05
step: 190, loss: 0.00037486664950847626

```

```

step: 200, loss: 7.234450094983913e-06
step: 210, loss: 1.4569316590495873e-05
step: 220, loss: 0.00010204187856288627
step: 230, loss: 1.1227773029531818e-05
step: 240, loss: 0.0011837505735456944
step: 250, loss: 2.1433996153064072e-05
step: 260, loss: 0.0041143763810396194
step: 270, loss: 0.00309391925111413
step: 280, loss: 0.002737829927355051
step: 290, loss: 0.0005113457445986569
step: 300, loss: 3.133973586955108e-05
step: 310, loss: 2.0624576791306026e-05
step: 320, loss: 0.0004365153145045042
step: 330, loss: 6.273362487263512e-06
step: 340, loss: 0.00040533446008339524
step: 350, loss: 2.6940380848827772e-05
step: 360, loss: 0.000530688907019794
step: 370, loss: 0.0019341274164617062
step: 380, loss: 0.007201549131423235
step: 390, loss: 0.002998006995767355
step: 400, loss: 0.00019904582586605102
step: 410, loss: 0.003777158912271261
step: 420, loss: 0.00895161833614111
step: 430, loss: 0.0018791741458699107
step: 440, loss: 4.354642442194745e-05
step: 450, loss: 0.000741496798582375
step: 460, loss: 0.0006305438000708818
step: 470, loss: 0.0009179407497867942
step: 480, loss: 0.0027918629348278046
step: 490, loss: 0.004798323847353458
step: 500, loss: 0.0007376823341473937
step: 510, loss: 1.234918090631254e-05
step: 520, loss: 2.8975491659366526e-05
step: 530, loss: 0.002852268749848008
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
epoch 11: dev_f1=0.9562413634269922, f1=0.9478021978021979,
best_f1=0.9598880597014925
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 3.982735506724566e-05
step: 10, loss: 0.00022991352307144552
step: 20, loss: 0.0053024897351861
step: 30, loss: 8.236536814365536e-06
step: 40, loss: 0.0005673238192684948
step: 50, loss: 8.102452738967258e-06

```

step: 60, loss: 0.00022842899488750845  
step: 70, loss: 2.8271442715777084e-05  
step: 80, loss: 0.00014244488556869328  
step: 90, loss: 0.0012580545153468847  
step: 100, loss: 0.02548016607761383  
step: 110, loss: 1.2941262866661418e-05  
step: 120, loss: 3.7913210690021515e-05  
step: 130, loss: 0.0005901426775380969  
step: 140, loss: 0.0008412291645072401  
step: 150, loss: 6.020659566274844e-05  
step: 160, loss: 0.0012673146557062864  
step: 170, loss: 1.6718729966669343e-05  
step: 180, loss: 1.4833794011792634e-05  
step: 190, loss: 0.0018436352256685495  
step: 200, loss: 0.0024799592792987823  
step: 210, loss: 0.00013161389506421983  
step: 220, loss: 0.001190868322737515  
step: 230, loss: 0.00023795576998963952  
step: 240, loss: 0.0007967556593939662  
step: 250, loss: 0.0003468292998149991  
step: 260, loss: 3.303530320408754e-05  
step: 270, loss: 0.003244180465117097  
step: 280, loss: 0.003090172540396452  
step: 290, loss: 0.0032428144477307796  
step: 300, loss: 0.002337096957489848  
step: 310, loss: 0.0004199244431219995  
step: 320, loss: 0.00017147288599517196  
step: 330, loss: 0.011728182435035706  
step: 340, loss: 5.2056344429729506e-05  
step: 350, loss: 1.3239188774605282e-05  
step: 360, loss: 9.461390436626971e-05  
step: 370, loss: 0.00463811494410038  
step: 380, loss: 0.00021381318219937384  
step: 390, loss: 0.01721576601266861  
step: 400, loss: 8.225381861848291e-06  
step: 410, loss: 0.0017809721175581217  
step: 420, loss: 0.00021156849106773734  
step: 430, loss: 9.346690058009699e-05  
step: 440, loss: 0.0001718334387987852  
step: 450, loss: 0.0013651045737788081  
step: 460, loss: 0.0010772062232717872  
step: 470, loss: 0.009127081371843815  
step: 480, loss: 0.03543628752231598  
step: 490, loss: 0.00013904384104534984  
step: 500, loss: 0.00066026346758008  
step: 510, loss: 0.017244383692741394  
step: 520, loss: 0.00762605806812644  
step: 530, loss: 0.00021797404042445123

```
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 12: dev_f1=0.9618039576622182, f1=0.9515096065873743,
best_f1=0.9598880597014925
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 2.6804382287082262e-05
step: 10, loss: 0.00011088293103966862
step: 20, loss: 0.002189642284065485
step: 30, loss: 7.193480087153148e-06
step: 40, loss: 7.981899398146197e-05
step: 50, loss: 0.0030648468527942896
step: 60, loss: 0.00018145782814826816
step: 70, loss: 4.5472279452951625e-05
step: 80, loss: 0.0003727650619111955
step: 90, loss: 4.030102354590781e-05
step: 100, loss: 0.00010229033068753779
step: 110, loss: 0.0011613621609285474
step: 120, loss: 9.886794941849075e-06
step: 130, loss: 1.581268588779494e-05
step: 140, loss: 4.202879063086584e-05
step: 150, loss: 0.0028098453767597675
step: 160, loss: 0.005150959827005863
step: 170, loss: 0.0013265963643789291
step: 180, loss: 1.8953371181851253e-05
step: 190, loss: 0.0020400278735905886
step: 200, loss: 2.471291008987464e-05
step: 210, loss: 1.057223926181905e-05
step: 220, loss: 4.226907549309544e-05
step: 230, loss: 0.0007889391854405403
step: 240, loss: 0.0009843179723247886
step: 250, loss: 0.00040576912579126656
step: 260, loss: 7.404421194223687e-05
step: 270, loss: 0.006205705460160971
step: 280, loss: 1.0430698239360936e-05
step: 290, loss: 7.498967988794902e-06
step: 300, loss: 6.781835691072047e-05
step: 310, loss: 3.845888204523362e-05
step: 320, loss: 1.1499721040308941e-05
step: 330, loss: 1.2125661669415422e-05
step: 340, loss: 3.3007127058226615e-05
step: 350, loss: 5.75554395254585e-06
step: 360, loss: 0.1085880845785141
step: 370, loss: 3.951898906962015e-05
step: 380, loss: 7.130151516321348e-06
step: 390, loss: 0.0003189159033354372
```



```

step: 400, loss: 1.4926492440281436e-05
step: 410, loss: 1.2066035196767189e-05
step: 420, loss: 7.2940274549182504e-06
step: 430, loss: 8.596711268182844e-05
step: 440, loss: 6.329230927804019e-06
step: 450, loss: 0.005160611588507891
step: 460, loss: 4.7087029088288546e-05
step: 470, loss: 0.004474576562643051
step: 480, loss: 5.885929112992017e-06
step: 490, loss: 9.712566679809242e-05
step: 500, loss: 9.145461262960453e-06
step: 510, loss: 9.909042091749143e-06
step: 520, loss: 6.984867923165439e-06
step: 530, loss: 1.970413541130256e-05
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 13: dev_f1=0.9619312906220985, f1=0.9518072289156626,
best_f1=0.9598880597014925
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0005460300599224865
step: 10, loss: 0.001759365084581077
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 20, loss: 7.212100172182545e-06
step: 30, loss: 0.0001508352142991498
step: 40, loss: 0.0007691893843002617
step: 50, loss: 0.00014349758566822857
step: 60, loss: 8.426558451901656e-06
step: 70, loss: 7.949703103804495e-06
step: 80, loss: 1.520252408226952e-05
step: 90, loss: 6.925267825863557e-06
step: 100, loss: 1.4319351976155303e-05
step: 110, loss: 0.005218227859586477
step: 120, loss: 8.899636668502353e-06
step: 130, loss: 9.409933227289002e-06
step: 140, loss: 9.362931450596079e-05
step: 150, loss: 0.000529133016243577
step: 160, loss: 8.445171260973439e-06
step: 170, loss: 1.4412621567316819e-05
step: 180, loss: 9.607430911273696e-06
step: 190, loss: 0.00503189954906702
step: 200, loss: 8.55319194670301e-06
step: 210, loss: 1.3630469766212627e-05
step: 220, loss: 7.547385848738486e-06
step: 230, loss: 2.1494832253665663e-05
step: 240, loss: 0.0011380633804947138

```

```

step: 250, loss: 0.006772337015718222
step: 260, loss: 1.1246427675359882e-05
step: 270, loss: 0.0003582470235414803
step: 280, loss: 9.484507245360874e-06
step: 290, loss: 5.483600034494884e-06
step: 300, loss: 1.7396190742147155e-05
step: 310, loss: 6.686853339488152e-06
step: 320, loss: 8.944341061578598e-06
step: 330, loss: 0.0008559752604924142
step: 340, loss: 0.0007186290458776057
step: 350, loss: 7.182319677667692e-06
step: 360, loss: 1.5845958841964602e-05
step: 370, loss: 4.034532685182057e-05
step: 380, loss: 0.004922507796436548
step: 390, loss: 0.00045343866804614663
step: 400, loss: 0.0016327040502801538
step: 410, loss: 6.973693416512106e-06
step: 420, loss: 1.2922310816065874e-05
step: 430, loss: 8.27004714665236e-06
step: 440, loss: 1.4327099052025005e-05
step: 450, loss: 4.136566349188797e-05
step: 460, loss: 0.04305169731378555
step: 470, loss: 8.288710887427442e-06
step: 480, loss: 5.591632088908227e-06
step: 490, loss: 7.18602223059861e-06
step: 500, loss: 0.003059163223952055
step: 510, loss: 0.022353630512952805
step: 520, loss: 0.00010215257498202845
step: 530, loss: 0.00024859930272214115
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 14: dev_f1=0.9633070134695774, f1=0.9499536607970344,
best_f1=0.9598880597014925
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 9.897907148115337e-06
step: 10, loss: 0.0022308328188955784
step: 20, loss: 0.0015891940565779805
step: 30, loss: 0.0010975869372487068
step: 40, loss: 7.72618022892857e-06
step: 50, loss: 0.0012615085579454899
step: 60, loss: 1.2263457392691635e-05
step: 70, loss: 4.152791734668426e-05
step: 80, loss: 5.293613412504783e-06
step: 90, loss: 3.193138400092721e-05
step: 100, loss: 4.179757524980232e-06

```

```
step: 110, loss: 8.355740646948107e-06
step: 120, loss: 6.992293947405415e-06
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 8192.0
step: 130, loss: 7.375962923106272e-06
step: 140, loss: 0.00024893187219277024
step: 150, loss: 9.819650585995987e-06
step: 160, loss: 6.139239758340409e-06
step: 170, loss: 5.7331867537868675e-06
step: 180, loss: 4.184384670224972e-05
step: 190, loss: 0.0004207938036415726
step: 200, loss: 0.0015129060484468937
step: 210, loss: 4.462881406652741e-06
step: 220, loss: 1.3868279893358704e-05
step: 230, loss: 1.4211152119969483e-05
step: 240, loss: 7.014661150606116e-06
step: 250, loss: 4.585815986501984e-06
step: 260, loss: 5.204211447562557e-06
step: 270, loss: 9.443474482395686e-06
step: 280, loss: 7.744773029116914e-06
step: 290, loss: 2.235779720649589e-05
step: 300, loss: 1.3577765457739588e-05
step: 310, loss: 0.027184465900063515
step: 320, loss: 8.415338015765883e-06
step: 330, loss: 8.527109457645565e-06
step: 340, loss: 1.2494067959778477e-05
step: 350, loss: 0.0018867531325668097
step: 360, loss: 6.7539058363763615e-06
step: 370, loss: 0.00011114456719951704
step: 380, loss: 6.150421995698707e-06
step: 390, loss: 0.0007044583326205611
step: 400, loss: 0.006966532673686743
step: 410, loss: 9.782387678569648e-06
step: 420, loss: 9.487946954322979e-06
step: 430, loss: 4.723641268356005e-06
step: 440, loss: 5.796514415123966e-06
step: 450, loss: 0.013276465237140656
step: 460, loss: 0.0012963010231032968
step: 470, loss: 5.200141094974242e-05
step: 480, loss: 0.0039204019121825695
step: 490, loss: 9.256636258214712e-05
step: 500, loss: 0.0017830593278631568
step: 510, loss: 7.212357013486326e-05
step: 520, loss: 6.720358214806765e-06
step: 530, loss: 4.4032767618773505e-06
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
epoch 15: dev_f1=0.9614849187935035, f1=0.950485886163813,
```

```

best_f1=0.9598880597014925
mv: cannot stat '*_dev.pt': No such file or directory
Selected optimization level O2:  FP16 training with FP32 batchnorm and FP32
master weights.

Defaults for this optimization level are:
enabled                : True
opt_level              : O2
cast_model_type        : torch.float16
patch_torch_functions  : False
keep_batchnorm_fp32    : True
master_weights         : True
loss_scale             : dynamic
Processing user overrides (additional kwargs that are not None)...
After processing overrides, optimization options are:
enabled                : True
opt_level              : O2
cast_model_type        : torch.float16
patch_torch_functions  : False
keep_batchnorm_fp32    : True
master_weights         : True
loss_scale             : dynamic
Warning:  multi_tensor_applier fused unscale kernel is unavailable, possibly
because apex was installed without --cuda_ext --cpp_ext. Using Python fallback.
Original ImportError was: ModuleNotFoundError("No module named 'amp_C'")
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
5742it [00:24, 233.59it/s]
load_f1 = 0.9630322882545623
real_f1 = 0.9608208955223881
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
4398it [00:24, 178.84it/s]

```

#### 1.1.4 Amazon-Google - Running the matcher

```

[ ]: # Step 1:
      # train the wdc model with the save_model flag on
      # it will produce two files *_dev.pt and *_test.pt which are the
      # best checkpoints on the dev set and the test set.
      !CUDA_VISIBLE_DEVICES=0 python train_ditto.py \

```

```

--task Structured/Amazon-Google \
--batch_size 32 \
--max_len 256 \
--lr 3e-5 \
--n_epochs 15 \
--finetuning \
--lm roberta \
--fp16 \
--save_model \
--da entry_swap \
--da span_del \
--summarize

# Step 2: pick a checkpoint, rename it to wdc_all_small.pt
!mv *_dev.pt checkpoints/Amazon-Google.pt

# Step 3: run the matcher
!CUDA_VISIBLE_DEVICES=0 python matcher.py \
  --task Structured/Amazon-Google \
  --input_path input/input_small.jsonl \
  --output_path output/output_small.jsonl \
  --lm roberta \
  --use_gpu \
  --fp16 \
  --checkpoint_path checkpoints/

```

Selected optimization level 02: FP16 training with FP32 batchnorm and FP32 master weights.

Defaults for this optimization level are:

```

enabled           : True
opt_level         : 02
cast_model_type   : torch.float16
patch_torch_functions : False
keep_batchnorm_fp32 : True
master_weights    : True
loss_scale        : dynamic

```

Processing user overrides (additional kwargs that are not None)...

After processing overrides, optimization options are:

```

enabled           : True
opt_level         : 02
cast_model_type   : torch.float16
patch_torch_functions : False
keep_batchnorm_fp32 : True
master_weights    : True
loss_scale        : dynamic

```

Warning: multi\_tensor\_applier fused unscale kernel is unavailable, possibly

```

because apex was installed without --cuda_ext --cpp_ext. Using Python fallback.
Original ImportError was: ModuleNotFoundError("No module named 'amp_C'")
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 32768.0
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:136:
UserWarning: Detected call of `lr_scheduler.step()` before `optimizer.step()`.
In PyTorch 1.1.0 and later, you should call them in the opposite order:
`optimizer.step()` before `lr_scheduler.step()`. Failure to do this will result
in PyTorch skipping the first value of the learning rate schedule. See more
details at https://pytorch.org/docs/stable/optim.html#how-to-adjust-learning-
rate
    "https://pytorch.org/docs/stable/optim.html#how-to-adjust-learning-rate",
UserWarning)
step: 0, loss: 0.5087079405784607
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 10, loss: 0.4774567484855652
step: 20, loss: 0.5314369797706604
step: 30, loss: 0.34352314472198486
step: 40, loss: 0.3624773323535919
step: 50, loss: 0.43399307131767273
step: 60, loss: 0.5160776972770691
step: 70, loss: 0.2836116850376129
step: 80, loss: 0.273493230342865
step: 90, loss: 0.22214940190315247
step: 100, loss: 0.2258540242910385
step: 110, loss: 0.2733578383922577
step: 120, loss: 0.38331371545791626
step: 130, loss: 0.17471939325332642
step: 140, loss: 0.4007999897003174
step: 150, loss: 0.19831892848014832
step: 160, loss: 0.3363288640975952
step: 170, loss: 0.053871069103479385
step: 180, loss: 0.2101452648639679
step: 190, loss: 0.4744056165218353
step: 200, loss: 0.18008288741111755
step: 210, loss: 0.24412089586257935
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
    warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 1: dev_f1=0.6601941747572816, f1=0.6212914485165792,
best_f1=0.6212914485165792
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:

```

```

An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.17235885560512543
step: 10, loss: 0.04857509583234787
step: 20, loss: 0.274850457906723
step: 30, loss: 0.14680765569210052
step: 40, loss: 0.14490428566932678
step: 50, loss: 0.268877774477005
step: 60, loss: 0.2703332304954529
step: 70, loss: 0.1657954901456833
step: 80, loss: 0.16897204518318176
step: 90, loss: 0.17017248272895813
step: 100, loss: 0.28907471895217896
step: 110, loss: 0.2651827335357666
step: 120, loss: 0.11166428029537201
step: 130, loss: 0.09072253853082657
step: 140, loss: 0.08403093367815018
step: 150, loss: 0.22418063879013062
step: 160, loss: 0.07182472199201584
step: 170, loss: 0.2738984525203705
step: 180, loss: 0.1059408038854599
step: 190, loss: 0.3072206676006317
step: 200, loss: 0.12812598049640656
step: 210, loss: 0.11166214197874069
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
  warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 2: dev_f1=0.7353535353535353, f1=0.734020618556701,
best_f1=0.734020618556701
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.047881416976451874
step: 10, loss: 0.0631445050239563
step: 20, loss: 0.14832216501235962
step: 30, loss: 0.04478134587407112
step: 40, loss: 0.2480638027191162
step: 50, loss: 0.18051975965499878
step: 60, loss: 0.15596126019954681
step: 70, loss: 0.05033668875694275
step: 80, loss: 0.08327453583478928
step: 90, loss: 0.02028059959411621
step: 100, loss: 0.1525912582874298
step: 110, loss: 0.15173988044261932

```

```

step: 120, loss: 0.2182266265153885
step: 130, loss: 0.13411776721477509
step: 140, loss: 0.07230184227228165
step: 150, loss: 0.09915217757225037
step: 160, loss: 0.1704520881175995
step: 170, loss: 0.18390725553035736
step: 180, loss: 0.09811750054359436
step: 190, loss: 0.013913603499531746
step: 200, loss: 0.19485491514205933
step: 210, loss: 0.16868257522583008
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
epoch 3: dev_f1=0.7302752293577981, f1=0.7352380952380952,
best_f1=0.734020618556701
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.17390593886375427
step: 10, loss: 0.0980515107512474
step: 20, loss: 0.10149959474802017
step: 30, loss: 0.07974205166101456
step: 40, loss: 0.025829549878835678
step: 50, loss: 0.16178125143051147
step: 60, loss: 0.23674358427524567
step: 70, loss: 0.0538601279258728
step: 80, loss: 0.2175426334142685
step: 90, loss: 0.11089641600847244
step: 100, loss: 0.3320313096046448
step: 110, loss: 0.5259600877761841
step: 120, loss: 0.15093070268630981
step: 130, loss: 0.17224204540252686
step: 140, loss: 0.09232865273952484
step: 150, loss: 0.07625947892665863
step: 160, loss: 0.06788399070501328
step: 170, loss: 0.03504882752895355
step: 180, loss: 0.06330589950084686
step: 190, loss: 0.1618059128522873
step: 200, loss: 0.02344256266951561
step: 210, loss: 0.2082320898771286
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
epoch 4: dev_f1=0.7155963302752294, f1=0.7335907335907336,
best_f1=0.734020618556701
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")

```



```

step: 0, loss: 0.13117799162864685
step: 10, loss: 0.019646478816866875
step: 20, loss: 0.138966903090477
step: 30, loss: 0.023593630641698837
step: 40, loss: 0.06965843588113785
step: 50, loss: 0.2167157232761383
step: 60, loss: 0.14395178854465485
step: 70, loss: 0.14851947128772736
step: 80, loss: 0.029616571962833405
step: 90, loss: 0.07266194373369217
step: 100, loss: 0.012227359227836132
step: 110, loss: 0.05946933105587959
step: 120, loss: 0.09587131440639496
step: 130, loss: 0.08849955350160599
step: 140, loss: 0.12020353972911835
step: 150, loss: 0.036730341613292694
step: 160, loss: 0.06012054532766342
step: 170, loss: 0.02032618224620819
step: 180, loss: 0.06881731003522873
step: 190, loss: 0.11058749258518219
step: 200, loss: 0.04155959188938141
step: 210, loss: 0.06229764223098755
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
    warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 5: dev_f1=0.7545787545787546, f1=0.7442748091603054,
best_f1=0.7442748091603054
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.018341228365898132
step: 10, loss: 0.21932390332221985
step: 20, loss: 0.07150208950042725
step: 30, loss: 0.022010663524270058
step: 40, loss: 0.1291370540857315
step: 50, loss: 0.06096072494983673
step: 60, loss: 0.07389744371175766
step: 70, loss: 0.03771643713116646
step: 80, loss: 0.0727577954530716
step: 90, loss: 0.04176122322678566
step: 100, loss: 0.13910746574401855
step: 110, loss: 0.06439133733510971
step: 120, loss: 0.008904874324798584
step: 130, loss: 0.04256527125835419

```

```

step: 140, loss: 0.09460028260946274
step: 150, loss: 0.05588623136281967
step: 160, loss: 0.044187650084495544
step: 170, loss: 0.06474630534648895
step: 180, loss: 0.11887712776660919
step: 190, loss: 0.06133975833654404
step: 200, loss: 0.09975995123386383
step: 210, loss: 0.1535065621137619
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 6: dev_f1=0.753731343283582, f1=0.7346153846153847,
best_f1=0.7442748091603054
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.17992155253887177
step: 10, loss: 0.021933138370513916
step: 20, loss: 0.007417724933475256
step: 30, loss: 0.0632174015045166
step: 40, loss: 0.04401933029294014
step: 50, loss: 0.16056878864765167
step: 60, loss: 0.09273963421583176
step: 70, loss: 0.15757517516613007
step: 80, loss: 0.10874888300895691
step: 90, loss: 0.08651656657457352
step: 100, loss: 0.05357329547405243
step: 110, loss: 0.046917859464883804
step: 120, loss: 0.023951096460223198
step: 130, loss: 0.09532587230205536
step: 140, loss: 0.10808809101581573
step: 150, loss: 0.05954562872648239
step: 160, loss: 0.2006005495786667
step: 170, loss: 0.1352982223033905
step: 180, loss: 0.0605921670794487
step: 190, loss: 0.14003126323223114
step: 200, loss: 0.02286289632320404
step: 210, loss: 0.19851289689540863
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
    warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 7: dev_f1=0.7561436672967863, f1=0.7481203007518797,
best_f1=0.7481203007518797
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:

```

```

An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.17550000548362732
step: 10, loss: 0.060317207127809525
step: 20, loss: 0.05155697464942932
step: 30, loss: 0.06879827380180359
step: 40, loss: 0.05269152298569679
step: 50, loss: 0.05053946003317833
step: 60, loss: 0.09521905332803726
step: 70, loss: 0.10879514366388321
step: 80, loss: 0.07093507796525955
step: 90, loss: 0.09083785861730576
step: 100, loss: 0.11095267534255981
step: 110, loss: 0.02053961157798767
step: 120, loss: 0.09912631660699844
step: 130, loss: 0.014253777451813221
step: 140, loss: 0.03874773159623146
step: 150, loss: 0.19634157419204712
step: 160, loss: 0.04883129894733429
step: 170, loss: 0.06653182208538055
step: 180, loss: 0.04739196598529816
step: 190, loss: 0.01899772696197033
step: 200, loss: 0.020500192418694496
step: 210, loss: 0.12366221845149994
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
epoch 8: dev_f1=0.7444444444444445, f1=0.7454545454545455,
best_f1=0.7481203007518797
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.02493422105908394
step: 10, loss: 0.03330384939908981
step: 20, loss: 0.08665484935045242
step: 30, loss: 0.001517444266937673
step: 40, loss: 0.0494159460067749
step: 50, loss: 0.09631858021020889
step: 60, loss: 0.04387855902314186
step: 70, loss: 0.04395195469260216
step: 80, loss: 0.03428317978978157
step: 90, loss: 0.24068759381771088
step: 100, loss: 0.09087762236595154
step: 110, loss: 0.10998496413230896
step: 120, loss: 0.041853006929159164
step: 130, loss: 0.09179778397083282
step: 140, loss: 0.13190937042236328
step: 150, loss: 0.008006545715034008

```

```

step: 160, loss: 0.0892661064863205
step: 170, loss: 0.06212396174669266
step: 180, loss: 0.017450012266635895
step: 190, loss: 0.04361650347709656
step: 200, loss: 0.025583084672689438
step: 210, loss: 0.06839414685964584
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
    warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 9: dev_f1=0.7692307692307692, f1=0.7539062500000001,
best_f1=0.7539062500000001
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00870334543287754
step: 10, loss: 0.02332278899848461
step: 20, loss: 0.013957204297184944
step: 30, loss: 0.00461026793345809
step: 40, loss: 0.0062928879633545876
step: 50, loss: 0.010929900221526623
step: 60, loss: 0.004411147441715002
step: 70, loss: 0.10902358591556549
step: 80, loss: 0.030264651402831078
step: 90, loss: 0.09500733762979507
step: 100, loss: 0.03147079795598984
step: 110, loss: 0.006726596504449844
step: 120, loss: 0.06106991693377495
step: 130, loss: 0.037609122693538666
step: 140, loss: 0.1656302809715271
step: 150, loss: 0.035282306373119354
step: 160, loss: 0.03401955962181091
step: 170, loss: 0.08728695660829544
step: 180, loss: 0.03776802122592926
step: 190, loss: 0.03408362343907356
step: 200, loss: 0.17871885001659393
step: 210, loss: 0.023403478786349297
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 10: dev_f1=0.7539062500000001, f1=0.7475538160469667,
best_f1=0.7539062500000001
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")

```

```

step: 0, loss: 0.05853407457470894
step: 10, loss: 0.009416641667485237
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 20, loss: 0.027973229065537453
step: 30, loss: 0.02140512503683567
step: 40, loss: 0.026064656674861908
step: 50, loss: 0.005871306173503399
step: 60, loss: 0.010612179525196552
step: 70, loss: 0.018985217437148094
step: 80, loss: 0.08129885047674179
step: 90, loss: 0.10830448567867279
step: 100, loss: 0.023817218840122223
step: 110, loss: 0.02684369683265686
step: 120, loss: 0.06833091378211975
step: 130, loss: 0.014624331146478653
step: 140, loss: 0.1299831122159958
step: 150, loss: 0.1159881204366684
step: 160, loss: 0.006180590484291315
step: 170, loss: 0.030194677412509918
step: 180, loss: 0.006650436203926802
step: 190, loss: 0.041639454662799835
step: 200, loss: 0.06983467936515808
step: 210, loss: 0.0861789807677269
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 11: dev_f1=0.7611940298507464, f1=0.7509578544061304,
best_f1=0.7539062500000001
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.033226244151592255
step: 10, loss: 0.09625004231929779
step: 20, loss: 0.07434773445129395
step: 30, loss: 0.05764491483569145
step: 40, loss: 0.005784506443887949
step: 50, loss: 0.029525067657232285
step: 60, loss: 0.0040895878337323666
step: 70, loss: 0.09420320391654968
step: 80, loss: 0.03473209962248802
step: 90, loss: 0.03375258296728134
step: 100, loss: 0.009792177006602287
step: 110, loss: 0.019094476476311684
step: 120, loss: 0.011441687121987343
step: 130, loss: 0.10475098341703415
step: 140, loss: 0.02386561408638954
step: 150, loss: 0.0006162550416775048
step: 160, loss: 0.006815367378294468

```

```

step: 170, loss: 0.05144413933157921
step: 180, loss: 0.0803360641002655
step: 190, loss: 0.006196016911417246
step: 200, loss: 0.043872784823179245
step: 210, loss: 0.004261911381036043
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
epoch 12: dev_f1=0.7642585551330798, f1=0.7475538160469667,
best_f1=0.7539062500000001
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.010209405794739723
step: 10, loss: 0.014589368365705013
step: 20, loss: 0.005433621816337109
step: 30, loss: 0.001533053582534194
step: 40, loss: 0.015538062900304794
step: 50, loss: 0.15046533942222595
step: 60, loss: 0.020188281312584877
step: 70, loss: 0.06151655316352844
step: 80, loss: 0.04550769552588463
step: 90, loss: 0.025085125118494034
step: 100, loss: 0.007737987674772739
step: 110, loss: 0.08754206448793411
step: 120, loss: 0.07958382368087769
step: 130, loss: 0.006351995747536421
step: 140, loss: 0.019842594861984253
step: 150, loss: 0.0070068505592644215
step: 160, loss: 0.1215878576040268
step: 170, loss: 0.0379587784409523
step: 180, loss: 0.007078151218593121
step: 190, loss: 0.012770205736160278
step: 200, loss: 0.04004000127315521
step: 210, loss: 0.027581488713622093
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
epoch 13: dev_f1=0.7596899224806202, f1=0.7490039840637451,
best_f1=0.7539062500000001
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.013516712002456188
step: 10, loss: 0.061887308955192566
step: 20, loss: 0.005929530598223209
step: 30, loss: 0.0014599598944187164
step: 40, loss: 0.038114383816719055

```

```

step: 50, loss: 0.00788319855928421
step: 60, loss: 0.17072123289108276
step: 70, loss: 0.012025313451886177
step: 80, loss: 0.018666138872504234
step: 90, loss: 0.0062801651656627655
step: 100, loss: 0.024689283221960068
step: 110, loss: 0.0013274055672809482
step: 120, loss: 0.0006820561829954386
step: 130, loss: 0.016950882971286774
step: 140, loss: 0.016356853768229485
step: 150, loss: 0.019209561869502068
step: 160, loss: 0.06445909291505814
step: 170, loss: 0.03330611065030098
step: 180, loss: 0.0028447892982512712
step: 190, loss: 0.05474916473031044
step: 200, loss: 0.01968732289969921
step: 210, loss: 0.0034930279944092035
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
  warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 14: dev_f1=0.7745664739884391, f1=0.7475149105367793,
best_f1=0.7475149105367793
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.01587391272187233
step: 10, loss: 0.006885927636176348
step: 20, loss: 0.041580647230148315
step: 30, loss: 0.0351688526570797
step: 40, loss: 0.014934325590729713
step: 50, loss: 0.21391254663467407
step: 60, loss: 0.16692164540290833
step: 70, loss: 0.014432111755013466
step: 80, loss: 0.06850115209817886
step: 90, loss: 0.02262389287352562
step: 100, loss: 0.004603591747581959
step: 110, loss: 0.054974064230918884
step: 120, loss: 0.04679933190345764
step: 130, loss: 0.014069756492972374
step: 140, loss: 0.020300157368183136
step: 150, loss: 0.014648540876805782
step: 160, loss: 0.010125060565769672
step: 170, loss: 0.01941419206559658
step: 180, loss: 0.0038038170896470547

```

```

step: 190, loss: 0.0038823564536869526
step: 200, loss: 0.002046853769570589
step: 210, loss: 0.07849474251270294
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 15: dev_f1=0.7692307692307692, f1=0.7480314960629921,
best_f1=0.7475149105367793
mv: cannot stat '*_dev.pt': No such file or directory
Selected optimization level 02: FP16 training with FP32 batchnorm and FP32
master weights.

Defaults for this optimization level are:
enabled                : True
opt_level              : 02
cast_model_type        : torch.float16
patch_torch_functions  : False
keep_batchnorm_fp32    : True
master_weights         : True
loss_scale             : dynamic
Processing user overrides (additional kwargs that are not None)...
After processing overrides, optimization options are:
enabled                : True
opt_level              : 02
cast_model_type        : torch.float16
patch_torch_functions  : False
keep_batchnorm_fp32    : True
master_weights         : True
loss_scale             : dynamic
Warning: multi_tensor_applier fused unscale kernel is unavailable, possibly
because apex was installed without --cuda_ext --cpp_ext. Using Python fallback.
Original ImportError was: ModuleNotFoundError("No module named 'amp_C'")
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
2293it [00:05, 402.03it/s]
load_f1 = 0.769825918762089
real_f1 = 0.7582205029013539
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
4398it [00:24, 176.79it/s]

```



### 1.1.5 Walmart-Amazon - Running the matcher

```
[ ]: # Step 1:
# train the wdc model with the save_model flag on
# it will produce two files *_dev.pt and *_test.pt which are the
# best checkpoints on the dev set and the test set.
!CUDA_VISIBLE_DEVICES=0 python train_ditto.py \
  --task Structured/Walmart-Amazon \
  --batch_size 32 \
  --max_len 256 \
  --lr 3e-5 \
  --n_epochs 15 \
  --finetuning \
  --lm roberta \
  --fp16 \
  --save_model \
  --da entry_swap \
  --da span_del \
  --summarize

# Step 2: pick a checkpoint, rename it to wdc_all_small.pt
!mv *_dev.pt checkpoints/Walmart-Amazon.pt

# Step 3: run the matcher
!CUDA_VISIBLE_DEVICES=0 python matcher.py \
  --task Structured/Walmart-Amazon \
  --input_path input/input_small.jsonl \
  --output_path output/output_small.jsonl \
  --lm roberta \
  --use_gpu \
  --fp16 \
  --checkpoint_path checkpoints/
```

Selected optimization level 02: FP16 training with FP32 batchnorm and FP32 master weights.

Defaults for this optimization level are:

```
enabled           : True
opt_level         : 02
cast_model_type   : torch.float16
patch_torch_functions : False
keep_batchnorm_fp32 : True
master_weights    : True
loss_scale        : dynamic
```

Processing user overrides (additional kwargs that are not None)...

After processing overrides, optimization options are:

```
enabled           : True
```

```

opt_level          : 02
cast_model_type    : torch.float16
patch_torch_functions : False
keep_batchnorm_fp32 : True
master_weights     : True
loss_scale         : dynamic
Warning: multi_tensor_applier fused unscale kernel is unavailable, possibly
because apex was installed without --cuda_ext --cpp_ext. Using Python fallback.
Original ImportError was: ModuleNotFoundError("No module named 'amp_C'")
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 32768.0
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:136:
UserWarning: Detected call of `lr_scheduler.step()` before `optimizer.step()`.
In PyTorch 1.1.0 and later, you should call them in the opposite order:
`optimizer.step()` before `lr_scheduler.step()`. Failure to do this will result
in PyTorch skipping the first value of the learning rate schedule. See more
details at https://pytorch.org/docs/stable/optim.html#how-to-adjust-learning-
rate
    "https://pytorch.org/docs/stable/optim.html#how-to-adjust-learning-rate",
UserWarning)
step: 0, loss: 0.42519837617874146
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 10, loss: 0.49262097477912903
step: 20, loss: 0.25443094968795776
step: 30, loss: 0.35470888018608093
step: 40, loss: 0.23254264891147614
step: 50, loss: 0.3135702610015869
step: 60, loss: 0.47000622749328613
step: 70, loss: 0.4512122571468353
step: 80, loss: 0.15822654962539673
step: 90, loss: 0.30605047941207886
step: 100, loss: 0.3746773600578308
step: 110, loss: 0.21293774247169495
step: 120, loss: 0.3143717050552368
step: 130, loss: 0.3134027421474457
step: 140, loss: 0.1749909520149231
step: 150, loss: 0.20466552674770355
step: 160, loss: 0.161247119307518
step: 170, loss: 0.4348885715007782
step: 180, loss: 0.11996207386255264
step: 190, loss: 0.12998434901237488
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or

```

```

loading the scheduler.
  warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 1: dev_f1=0.6335078534031413, f1=0.6974358974358974,
best_f1=0.6974358974358974
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.3117220401763916
step: 10, loss: 0.15564008057117462
step: 20, loss: 0.6883885860443115
step: 30, loss: 0.09601804614067078
step: 40, loss: 0.29113614559173584
step: 50, loss: 0.390726774930954
step: 60, loss: 0.2592054605484009
step: 70, loss: 0.19059480726718903
step: 80, loss: 0.11630687117576599
step: 90, loss: 0.11022332310676575
step: 100, loss: 0.054571617394685745
step: 110, loss: 0.07332774251699448
step: 120, loss: 0.018257586285471916
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 8192.0
step: 130, loss: 0.1255171149969101
step: 140, loss: 0.1283843219280243
step: 150, loss: 0.29226386547088623
step: 160, loss: 0.17170673608779907
step: 170, loss: 0.04388231784105301
step: 180, loss: 0.031633298844099045
step: 190, loss: 0.07944706827402115
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
  warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 2: dev_f1=0.8277634961439588, f1=0.8103896103896104,
best_f1=0.8103896103896104
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.06007594242691994
step: 10, loss: 0.07782287150621414
step: 20, loss: 0.04516815394163132
step: 30, loss: 0.005431006662547588
step: 40, loss: 0.12596675753593445
step: 50, loss: 0.049019087105989456
step: 60, loss: 0.07057425379753113
step: 70, loss: 0.19508954882621765

```

```

step: 80, loss: 0.14163418114185333
step: 90, loss: 0.04321567341685295
step: 100, loss: 0.14238201081752777
step: 110, loss: 0.15025688707828522
step: 120, loss: 0.16015765070915222
step: 130, loss: 0.09779257327318192
step: 140, loss: 0.048834506422281265
step: 150, loss: 0.22237978875637054
step: 160, loss: 0.06715165823698044
step: 170, loss: 0.2104005515575409
step: 180, loss: 0.08797887712717056
step: 190, loss: 0.07109083980321884
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
    warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 3: dev_f1=0.8585858585858585, f1=0.8511749347258485,
best_f1=0.8511749347258485
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0067596216686069965
step: 10, loss: 0.10111147165298462
step: 20, loss: 0.060712918639183044
step: 30, loss: 0.013467442244291306
step: 40, loss: 0.10049852728843689
step: 50, loss: 0.032763637602329254
step: 60, loss: 0.32647788524627686
step: 70, loss: 0.11899971216917038
step: 80, loss: 0.005772092379629612
step: 90, loss: 0.09193910658359528
step: 100, loss: 0.03884662315249443
step: 110, loss: 0.10797078907489777
step: 120, loss: 0.061302028596401215
step: 130, loss: 0.015185272321105003
step: 140, loss: 0.07999838143587112
step: 150, loss: 0.07620568573474884
step: 160, loss: 0.015615209937095642
step: 170, loss: 0.07687258720397949
step: 180, loss: 0.07385557144880295
step: 190, loss: 0.007852639071643353
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 4: dev_f1=0.845, f1=0.8542199488491048, best_f1=0.8511749347258485

```

```

/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.07160668820142746
step: 10, loss: 0.033777590841054916
step: 20, loss: 0.002901835134252906
step: 30, loss: 0.010362586937844753
step: 40, loss: 0.07222159206867218
step: 50, loss: 0.01700681820511818
step: 60, loss: 0.004111722111701965
step: 70, loss: 0.041244667023420334
step: 80, loss: 0.004511515609920025
step: 90, loss: 0.2277565747499466
step: 100, loss: 0.08510793000459671
step: 110, loss: 0.1536320149898529
step: 120, loss: 0.05266404524445534
step: 130, loss: 0.2727840542793274
step: 140, loss: 0.05607549101114273
step: 150, loss: 0.02351256087422371
step: 160, loss: 0.026408705860376358
step: 170, loss: 0.008089411072432995
step: 180, loss: 0.03060591220855713
step: 190, loss: 0.01141273882240057
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 5: dev_f1=0.843989769820972, f1=0.8533333333333334,
best_f1=0.8511749347258485
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.036892369389534
step: 10, loss: 0.012796327471733093
step: 20, loss: 0.014998981729149818
step: 30, loss: 0.03864656388759613
step: 40, loss: 0.04421468451619148
step: 50, loss: 0.11525248736143112
step: 60, loss: 0.05704374983906746
step: 70, loss: 0.09103813767433167
step: 80, loss: 0.010376355610787868
step: 90, loss: 0.058944202959537506
step: 100, loss: 0.030459774658083916
step: 110, loss: 0.004324568435549736
step: 120, loss: 0.010579921305179596
step: 130, loss: 0.034200940281152725
step: 140, loss: 0.01511384267359972
step: 150, loss: 0.039062369614839554
step: 160, loss: 0.022421376779675484

```

```

step: 170, loss: 0.19301870465278625
step: 180, loss: 0.036764875054359436
step: 190, loss: 0.05673690885305405
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 6: dev_f1=0.8549618320610687, f1=0.8496042216358839,
best_f1=0.8511749347258485
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.013282150961458683
step: 10, loss: 0.0037319250404834747
step: 20, loss: 0.0026404550299048424
step: 30, loss: 0.005860566161572933
step: 40, loss: 0.008418888784945011
step: 50, loss: 0.021838266402482986
step: 60, loss: 0.061184752732515335
step: 70, loss: 0.001427473733201623
step: 80, loss: 0.0006745949503965676
step: 90, loss: 0.15113532543182373
step: 100, loss: 0.217110276222229
step: 110, loss: 0.00964153092354536
step: 120, loss: 0.0044830176047980785
step: 130, loss: 0.009273331612348557
step: 140, loss: 0.025018353015184402
step: 150, loss: 0.017888866364955902
step: 160, loss: 0.04201151803135872
step: 170, loss: 0.2137060910463333
step: 180, loss: 0.028231462463736534
step: 190, loss: 0.04175820201635361
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
    warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 7: dev_f1=0.8652849740932642, f1=0.8471849865951743,
best_f1=0.8471849865951743
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.12990424036979675
step: 10, loss: 0.018130140379071236
step: 20, loss: 0.002364660846069455
step: 30, loss: 0.06436213850975037
step: 40, loss: 0.030637478455901146

```

```

step: 50, loss: 0.014377090148627758
step: 60, loss: 0.007250247523188591
step: 70, loss: 0.0015129635576158762
step: 80, loss: 0.04889574646949768
step: 90, loss: 0.013196267187595367
step: 100, loss: 0.002416604431346059
step: 110, loss: 0.02667548879981041
step: 120, loss: 0.01755247637629509
step: 130, loss: 0.0015636419411748648
step: 140, loss: 0.00523914210498333
step: 150, loss: 0.025827279314398766
step: 160, loss: 0.0007773885736241937
step: 170, loss: 0.006839154288172722
step: 180, loss: 0.035871099680662155
step: 190, loss: 0.0037388091441243887
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 8: dev_f1=0.8640000000000001, f1=0.8478260869565217,
best_f1=0.8471849865951743
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0013065151870250702
step: 10, loss: 0.0006938920123502612
step: 20, loss: 0.007177694700658321
step: 30, loss: 0.004698943346738815
step: 40, loss: 0.15668156743049622
step: 50, loss: 0.002358654048293829
step: 60, loss: 0.037269964814186096
step: 70, loss: 0.0004931216826662421
step: 80, loss: 0.06509992480278015
step: 90, loss: 0.015145082958042622
step: 100, loss: 0.0019509291741997004
step: 110, loss: 0.017015693709254265
step: 120, loss: 0.0019651209004223347
step: 130, loss: 0.006934191100299358
step: 140, loss: 0.06822686642408371
step: 150, loss: 0.09446941316127777
step: 160, loss: 0.001521510537713766
step: 170, loss: 0.004519714042544365
step: 180, loss: 0.05386125668883324
step: 190, loss: 0.00045160437002778053
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or

```

```

loading the scheduler.
  warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 9: dev_f1=0.8724489795918368, f1=0.8404255319148937,
best_f1=0.8404255319148937
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0016947268741205335
step: 10, loss: 0.0008610695367679
step: 20, loss: 0.012007921002805233
step: 30, loss: 0.008507183752954006
step: 40, loss: 0.0015261992812156677
step: 50, loss: 0.021080687642097473
step: 60, loss: 0.0006287593860179186
step: 70, loss: 0.05535305663943291
step: 80, loss: 0.0017644894542172551
step: 90, loss: 0.006325101479887962
step: 100, loss: 0.0007785455673001707
step: 110, loss: 0.05426338315010071
step: 120, loss: 0.0013003733474761248
step: 130, loss: 0.0033453283831477165
step: 140, loss: 0.000835855258628726
step: 150, loss: 0.00546653289347887
step: 160, loss: 0.0027834177017211914
step: 170, loss: 0.0036018940154463053
step: 180, loss: 0.06765148788690567
step: 190, loss: 0.002827668096870184
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
epoch 10: dev_f1=0.8652849740932642, f1=0.851063829787234,
best_f1=0.8404255319148937
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.005538342520594597
step: 10, loss: 0.00035924825351685286
step: 20, loss: 0.0016066035022959113
step: 30, loss: 0.006704061292111874
step: 40, loss: 0.002590559422969818
step: 50, loss: 0.0031281241681426764
step: 60, loss: 0.0167473703622818
step: 70, loss: 0.0004084960091859102
step: 80, loss: 0.03362641483545303
step: 90, loss: 0.00565130403265357
step: 100, loss: 0.06713923811912537
step: 110, loss: 0.13974228501319885
step: 120, loss: 0.0019455014262348413

```



```

step: 130, loss: 0.10186154395341873
step: 140, loss: 0.001984872855246067
step: 150, loss: 0.0008799316128715873
step: 160, loss: 0.0002553017111495137
step: 170, loss: 0.0010939621133729815
step: 180, loss: 0.0007367627113126218
step: 190, loss: 0.001046471414156258
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 11: dev_f1=0.8556430446194225, f1=0.8548812664907651,
best_f1=0.8404255319148937
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0007310706423595548
step: 10, loss: 0.03289418667554855
step: 20, loss: 0.01109650731086731
step: 30, loss: 0.008922642096877098
step: 40, loss: 0.0030263070948421955
step: 50, loss: 0.010974058881402016
step: 60, loss: 0.04019380360841751
step: 70, loss: 0.0011112340725958347
step: 80, loss: 0.030838575214147568
step: 90, loss: 0.0006059205625206232
step: 100, loss: 0.002848140662536025
step: 110, loss: 0.010696874000132084
step: 120, loss: 0.05967095494270325
step: 130, loss: 0.049113232642412186
step: 140, loss: 0.004320444539189339
step: 150, loss: 0.003065314143896103
step: 160, loss: 0.014795280061662197
step: 170, loss: 0.004672123119235039
step: 180, loss: 0.0010401488980278373
step: 190, loss: 0.0008802851079963148
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 12: dev_f1=0.8578811369509043, f1=0.8511749347258485,
best_f1=0.8404255319148937
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.002385365776717663
step: 10, loss: 0.0037960419431328773
step: 20, loss: 0.0003662369854282588
step: 30, loss: 0.02788768894970417
step: 40, loss: 0.0020300426986068487

```

```

step: 50, loss: 0.04751814529299736
step: 60, loss: 0.12245797365903854
step: 70, loss: 0.05182773247361183
step: 80, loss: 0.0018528848886489868
step: 90, loss: 0.0008572076912969351
step: 100, loss: 0.007530053146183491
step: 110, loss: 0.0037066787481307983
step: 120, loss: 0.013093327172100544
step: 130, loss: 0.0045518772676587105
step: 140, loss: 0.02006971649825573
step: 150, loss: 0.0010975359473377466
step: 160, loss: 0.006123843137174845
step: 170, loss: 0.005985279101878405
step: 180, loss: 0.0009102766634896398
step: 190, loss: 0.11635497212409973
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 13: dev_f1=0.8514851485148514, f1=0.8520408163265306,
best_f1=0.8404255319148937
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.006880058906972408
step: 10, loss: 0.008143438957631588
step: 20, loss: 0.0010991543531417847
step: 30, loss: 0.0011429026490077376
step: 40, loss: 0.0006689713918603957
step: 50, loss: 0.0016123296227306128
step: 60, loss: 0.0003251114394515753
step: 70, loss: 0.0018957179272547364
step: 80, loss: 0.001614138251170516
step: 90, loss: 0.0019061864586547017
step: 100, loss: 0.0007429263205267489
step: 110, loss: 0.004351452458649874
step: 120, loss: 0.013156868517398834
step: 130, loss: 0.000547088508028537
step: 140, loss: 0.004239530302584171
step: 150, loss: 0.0019324974855408072
step: 160, loss: 0.0017330427654087543
step: 170, loss: 0.0069195725955069065
step: 180, loss: 0.005995747167617083
step: 190, loss: 0.005686990916728973
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 14: dev_f1=0.8542713567839197, f1=0.8505154639175257,
best_f1=0.8404255319148937

```

```
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
```

```
    warnings.warn("An input tensor was not cuda.")
```

```
step: 0, loss: 0.00027367344591766596
step: 10, loss: 0.009529859758913517
step: 20, loss: 0.009077994152903557
step: 30, loss: 0.011064193211495876
step: 40, loss: 0.12691128253936768
step: 50, loss: 0.0003504744090605527
step: 60, loss: 0.05592336878180504
step: 70, loss: 0.004443465266376734
step: 80, loss: 0.0006697827484458685
step: 90, loss: 0.0055524930357933044
step: 100, loss: 0.0009196246974170208
step: 110, loss: 0.0714801624417305
step: 120, loss: 0.0005165040493011475
step: 130, loss: 0.0003490750677883625
step: 140, loss: 0.03161097317934036
step: 150, loss: 0.000605968467425555
step: 160, loss: 0.0009678759961389005
step: 170, loss: 0.0437847264111042
step: 180, loss: 0.0010986332781612873
step: 190, loss: 0.014306805096566677
```

```
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
```

```
    warnings.warn("An input tensor was not cuda.")
```

```
epoch 15: dev_f1=0.8571428571428572, f1=0.8483290488431877,
best_f1=0.8404255319148937
```

```
mv: cannot stat '*_dev.pt': No such file or directory
```

```
Selected optimization level 02:  FP16 training with FP32 batchnorm and FP32
master weights.
```

```
Defaults for this optimization level are:
```

```
enabled                : True
opt_level              : 02
cast_model_type       : torch.float16
patch_torch_functions : False
keep_batchnorm_fp32   : True
master_weights        : True
loss_scale             : dynamic
```

```
Processing user overrides (additional kwargs that are not None)...
```

```
After processing overrides, optimization options are:
```

```
enabled                : True
opt_level              : 02
cast_model_type       : torch.float16
patch_torch_functions : False
keep_batchnorm_fp32   : True
master_weights        : True
```

```

loss_scale          : dynamic
Warning: multi_tensor_applier fused unscale kernel is unavailable, possibly
because apex was installed without --cuda_ext --cpp_ext. Using Python fallback.
Original ImportError was: ModuleNotFoundError("No module named 'amp_C'")
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
2049it [00:09, 223.92it/s]
load_f1 = 0.837696335078534
real_f1 = 0.8091603053435115
733it [00:00, 3292.82it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
4398it [00:25, 175.85it/s]

```

## 1.2 DA TEXTUAL

### 1.2.1 Abt-Buy - Running the matcher

```

[ ]: # Step 1:
# train the wdc model with the save_model flag on
# it will produce two files *_dev.pt and *_test.pt which are the
# best checkpoints on the dev set and the test set.
!CUDA_VISIBLE_DEVICES=0 python train_ditto.py \
    --task Textual/Abt-Buy \
    --batch_size 32 \
    --max_len 256 \
    --lr 3e-5 \
    --n_epochs 15 \
    --finetuning \
    --lm roberta \
    --fp16 \
    --save_model \
    --da entry_swap \
    --da span_shuffle \
    --summarize

# Step 2: pick a checkpoint, rename it to wdc_all_small.pt
!mv *_dev.pt checkpoints/Abt-Buy.pt

# Step 3: run the matcher
!CUDA_VISIBLE_DEVICES=0 python matcher.py \
    --task Textual/Abt-Buy \

```

```

--input_path input/input_small.jsonl \
--output_path output/output_small.jsonl \
--lm roberta \
--use_gpu \
--fp16 \
--checkpoint_path checkpoints/

```

```

Downloading: 100% 481/481 [00:00<00:00, 441kB/s]
Downloading: 100% 899k/899k [00:01<00:00, 808kB/s]
Downloading: 100% 456k/456k [00:00<00:00, 516kB/s]
Downloading: 100% 501M/501M [00:43<00:00, 11.5MB/s]
Selected optimization level O2:  FP16 training with FP32 batchnorm and FP32
master weights.

```

Defaults for this optimization level are:

```

enabled          : True
opt_level        : O2
cast_model_type  : torch.float16
patch_torch_functions : False
keep_batchnorm_fp32 : True
master_weights   : True
loss_scale       : dynamic

```

Processing user overrides (additional kwargs that are not None)...

After processing overrides, optimization options are:

```

enabled          : True
opt_level        : O2
cast_model_type  : torch.float16
patch_torch_functions : False
keep_batchnorm_fp32 : True
master_weights   : True
loss_scale       : dynamic

```

Warning: multi\_tensor\_applier fused unscale kernel is unavailable, possibly because apex was installed without --cuda\_ext --cpp\_ext. Using Python fallback.

Original ImportError was: ModuleNotFoundError("No module named 'amp\_C'")

/usr/local/lib/python3.7/dist-packages/apex/amp/\_initialize.py:25: UserWarning: An input tensor was not cuda.

```
warnings.warn("An input tensor was not cuda.")
```

Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 32768.0

/usr/local/lib/python3.7/dist-packages/torch/optim/lr\_scheduler.py:136:

UserWarning: Detected call of `lr\_scheduler.step()` before `optimizer.step()`.

In PyTorch 1.1.0 and later, you should call them in the opposite order:

`optimizer.step()` before `lr\_scheduler.step()`. Failure to do this will result in PyTorch skipping the first value of the learning rate schedule. See more details at <https://pytorch.org/docs/stable/optim.html#how-to-adjust-learning-rate>

"<https://pytorch.org/docs/stable/optim.html#how-to-adjust-learning-rate>", UserWarning)

```

step: 0, loss: 0.4992743730545044
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 10, loss: 0.5149590969085693
step: 20, loss: 0.3152698874473572
step: 30, loss: 0.4059995710849762
step: 40, loss: 0.4724136292934418
step: 50, loss: 0.3166466951370239
step: 60, loss: 0.5581367015838623
step: 70, loss: 0.31684958934783936
step: 80, loss: 0.25878360867500305
step: 90, loss: 0.23105081915855408
step: 100, loss: 0.17188826203346252
step: 110, loss: 0.4106925129890442
step: 120, loss: 0.2905150353908539
step: 130, loss: 0.3468267321586609
step: 140, loss: 0.39579468965530396
step: 150, loss: 0.3048207461833954
step: 160, loss: 0.38648077845573425
step: 170, loss: 0.30176830291748047
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
  warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 1: dev_f1=0.2690909090909091, f1=0.24159402241594022,
best_f1=0.24159402241594022
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.31771156191825867
step: 10, loss: 0.44500741362571716
step: 20, loss: 0.3005259335041046
step: 30, loss: 0.2827572524547577
step: 40, loss: 0.058875929564237595
step: 50, loss: 0.4049058258533478
step: 60, loss: 0.17933858931064606
step: 70, loss: 0.4184848666191101
step: 80, loss: 0.15001937747001648
step: 90, loss: 0.027196712791919708
step: 100, loss: 0.2361665666103363
step: 110, loss: 0.05457082763314247
step: 120, loss: 0.07672691345214844
step: 130, loss: 0.15259158611297607
step: 140, loss: 0.2428784966468811
step: 150, loss: 0.03312280401587486
step: 160, loss: 0.2257433384656906

```

```

step: 170, loss: 0.07424275577068329
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
    warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 2: dev_f1=0.8090452261306532, f1=0.8160377358490567,
best_f1=0.8160377358490567
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.18540534377098083
step: 10, loss: 0.10335139185190201
step: 20, loss: 0.03769667446613312
step: 30, loss: 0.038398582488298416
step: 40, loss: 0.15483224391937256
step: 50, loss: 0.1079142764210701
step: 60, loss: 0.023575855419039726
step: 70, loss: 0.4232199788093567
step: 80, loss: 0.12755757570266724
step: 90, loss: 0.2021597921848297
step: 100, loss: 0.048302628099918365
step: 110, loss: 0.12009163200855255
step: 120, loss: 0.14426591992378235
step: 130, loss: 0.06552333384752274
step: 140, loss: 0.13861198723316193
step: 150, loss: 0.020858360454440117
step: 160, loss: 0.013455689884722233
step: 170, loss: 0.01685524731874466
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
    warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 3: dev_f1=0.8395061728395062, f1=0.883054892601432,
best_f1=0.883054892601432
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.09072988480329514
step: 10, loss: 0.12357615679502487
step: 20, loss: 0.009019266813993454
step: 30, loss: 0.11004280298948288
step: 40, loss: 0.06740178167819977

```

```

step: 50, loss: 0.03228345513343811
step: 60, loss: 0.01889943517744541
step: 70, loss: 0.0033503971062600613
step: 80, loss: 0.1976657509803772
step: 90, loss: 0.15162025392055511
step: 100, loss: 0.12475141882896423
step: 110, loss: 0.20357951521873474
step: 120, loss: 0.22669556736946106
step: 130, loss: 0.04008098319172859
step: 140, loss: 0.11616142094135284
step: 150, loss: 0.15982255339622498
step: 160, loss: 0.0007584920967929065
step: 170, loss: 0.08358477056026459
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
    warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 4: dev_f1=0.8472906403940887, f1=0.8685446009389671,
best_f1=0.8685446009389671
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.013238953426480293
step: 10, loss: 0.027621671557426453
step: 20, loss: 0.024501919746398926
step: 30, loss: 0.01735861413180828
step: 40, loss: 0.012120098806917667
step: 50, loss: 0.018734214827418327
step: 60, loss: 0.03299267217516899
step: 70, loss: 0.04384471848607063
step: 80, loss: 0.0004148368025198579
step: 90, loss: 0.00741899199783802
step: 100, loss: 0.011683289892971516
step: 110, loss: 0.038062967360019684
step: 120, loss: 0.12452545017004013
step: 130, loss: 0.009247769601643085
step: 140, loss: 0.04579365253448486
step: 150, loss: 0.03301149606704712
step: 160, loss: 0.004520802292972803
step: 170, loss: 0.00270856567658484
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or

```



```

loading the scheduler.
  warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 5: dev_f1=0.8686868686868687, f1=0.9193154034229829,
best_f1=0.9193154034229829
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.11552701890468597
step: 10, loss: 0.010939611122012138
step: 20, loss: 0.009804259985685349
step: 30, loss: 0.005590333137661219
step: 40, loss: 0.0017325502121821046
step: 50, loss: 0.0018142288317903876
step: 60, loss: 0.017565211281180382
step: 70, loss: 0.007646622601896524
step: 80, loss: 0.002137969946488738
step: 90, loss: 0.0951584056019783
step: 100, loss: 0.0036865556612610817
step: 110, loss: 0.012702311389148235
step: 120, loss: 0.033324405550956726
step: 130, loss: 0.008724745362997055
step: 140, loss: 0.01867571845650673
step: 150, loss: 0.021128153428435326
step: 160, loss: 0.011519634164869785
step: 170, loss: 0.01890159212052822
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
epoch 6: dev_f1=0.8433734939759036, f1=0.8715596330275228,
best_f1=0.9193154034229829
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.07492846250534058
step: 10, loss: 0.0661037340760231
step: 20, loss: 0.005714171100407839
step: 30, loss: 0.0009402483701705933
step: 40, loss: 0.00035918355570174754
step: 50, loss: 0.0006913523538969457
step: 60, loss: 0.006171655375510454
step: 70, loss: 0.04245752468705177
step: 80, loss: 0.07142733037471771
step: 90, loss: 0.03102007508277893
step: 100, loss: 0.0006345793954096735
step: 110, loss: 0.04522620886564255
step: 120, loss: 0.0023016850464046
step: 130, loss: 0.00829820055514574
step: 140, loss: 0.0018982832552865148

```

```

step: 150, loss: 0.1413213461637497
step: 160, loss: 0.010161090642213821
step: 170, loss: 0.03297847509384155
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
    warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 7: dev_f1=0.8708860759493672, f1=0.9073170731707318,
best_f1=0.9073170731707318
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.02676568180322647
step: 10, loss: 0.0028889549430459738
step: 20, loss: 0.003795414697378874
step: 30, loss: 0.00031484285136684775
step: 40, loss: 0.0002729994885157794
step: 50, loss: 0.0017581016290932894
step: 60, loss: 0.001786147360689938
step: 70, loss: 0.025841012597084045
step: 80, loss: 0.0026862232480198145
step: 90, loss: 0.026184413582086563
step: 100, loss: 0.0003350032202433795
step: 110, loss: 0.08146178722381592
step: 120, loss: 0.12816327810287476
step: 130, loss: 0.0021995038259774446
step: 140, loss: 0.024868851527571678
step: 150, loss: 0.004695902578532696
step: 160, loss: 0.00030052365036681294
step: 170, loss: 0.07061399519443512
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 8: dev_f1=0.8658227848101266, f1=0.9130434782608696,
best_f1=0.9073170731707318
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.03450389578938484
step: 10, loss: 0.014578831382095814
step: 20, loss: 0.0006422706646844745
step: 30, loss: 0.0013891967246308923
step: 40, loss: 0.00528375431895256
step: 50, loss: 0.00549403065815568
step: 60, loss: 0.03586910292506218

```

```

step: 70, loss: 0.0007000962505117059
step: 80, loss: 0.0009294756455346942
step: 90, loss: 0.002750708721578121
step: 100, loss: 0.01018882542848587
step: 110, loss: 0.24199452996253967
step: 120, loss: 0.0025151788722723722
step: 130, loss: 0.0007648275932297111
step: 140, loss: 0.0011554239317774773
step: 150, loss: 0.09898634254932404
step: 160, loss: 0.0004574506892822683
step: 170, loss: 0.000581660948228091
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
    warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 9: dev_f1=0.8723897911832947, f1=0.8634361233480176,
best_f1=0.8634361233480176
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0002869022428058088
step: 10, loss: 0.011478711850941181
step: 20, loss: 0.03593425825238228
step: 30, loss: 0.03992736339569092
step: 40, loss: 0.0002441757242195308
step: 50, loss: 0.10066021233797073
step: 60, loss: 0.0014660884626209736
step: 70, loss: 0.00019017823797184974
step: 80, loss: 0.02587348408997059
step: 90, loss: 0.0001628061436349526
step: 100, loss: 0.000634491560049355
step: 110, loss: 0.007854991592466831
step: 120, loss: 0.0007010870613157749
step: 130, loss: 9.943930490408093e-05
step: 140, loss: 0.00271398201584816
step: 150, loss: 0.004061439540237188
step: 160, loss: 7.49807441025041e-05
step: 170, loss: 0.03618945553898811
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
    warnings.warn(SAVE_STATE_WARNING, UserWarning)

```

```

epoch 10: dev_f1=0.8894472361809046, f1=0.9086538461538461,
best_f1=0.9086538461538461
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.024214129894971848
step: 10, loss: 0.0011905630817636847
step: 20, loss: 0.0002457008813507855
step: 30, loss: 0.01038284320384264
step: 40, loss: 0.0006219960050657392
step: 50, loss: 0.0057398551143705845
step: 60, loss: 0.026068346574902534
step: 70, loss: 0.00244216900318861
step: 80, loss: 0.021778060123324394
step: 90, loss: 0.0009025782928802073
step: 100, loss: 0.015761112794280052
step: 110, loss: 0.0011340078199282289
step: 120, loss: 0.00017929237219505012
step: 130, loss: 0.017420096322894096
step: 140, loss: 0.060299284756183624
step: 150, loss: 0.0009635772439651191
step: 160, loss: 0.02360038459300995
step: 170, loss: 0.021316183730959892
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 11: dev_f1=0.883248730964467, f1=0.9156626506024096,
best_f1=0.9086538461538461
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0010736151598393917
step: 10, loss: 0.001981608336791396
step: 20, loss: 0.000459972390672192
step: 30, loss: 0.11940279603004456
step: 40, loss: 0.00029453024035319686
step: 50, loss: 0.008374221622943878
step: 60, loss: 0.023578345775604248
step: 70, loss: 0.0001613030763110146
step: 80, loss: 0.0011326613603159785
step: 90, loss: 0.009001855738461018
step: 100, loss: 0.00013824671623297036
step: 110, loss: 0.18212245404720306
step: 120, loss: 0.01699843257665634
step: 130, loss: 0.0049617537297308445
step: 140, loss: 0.0007304151076823473
step: 150, loss: 0.0036975766997784376
step: 160, loss: 0.0007714692037552595

```

```

step: 170, loss: 0.00023816082102712244
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 12: dev_f1=0.8849104859335037, f1=0.892601431980907,
best_f1=0.9086538461538461
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0009022425510920584
step: 10, loss: 8.763685764279217e-05
step: 20, loss: 0.0002447802689857781
step: 30, loss: 0.056387413293123245
step: 40, loss: 0.0001884981174953282
step: 50, loss: 0.017384931445121765
step: 60, loss: 0.000338102865498513
step: 70, loss: 0.06306470185518265
step: 80, loss: 0.00017541732813697308
step: 90, loss: 0.0012171759735792875
step: 100, loss: 0.011648095212876797
step: 110, loss: 0.0009721452952362597
step: 120, loss: 0.021209660917520523
step: 130, loss: 9.917630086420104e-05
step: 140, loss: 0.06752771884202957
step: 150, loss: 0.00035561935510486364
step: 160, loss: 0.0007119930814951658
step: 170, loss: 0.00017820240464061499
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 13: dev_f1=0.8872180451127819, f1=0.8995215311004785,
best_f1=0.9086538461538461
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0030524758622050285
step: 10, loss: 0.0006017381674610078
step: 20, loss: 0.02271653339266777
step: 30, loss: 0.0005846977001056075
step: 40, loss: 0.00017974637739825994
step: 50, loss: 7.171341712819412e-05
step: 60, loss: 0.00429697846993804
step: 70, loss: 0.01721441000699997
step: 80, loss: 0.00033187007647939026
step: 90, loss: 0.0001133692276198417
step: 100, loss: 9.284700354328379e-05
step: 110, loss: 0.012213476002216339
step: 120, loss: 0.00015811689081601799

```

```

step: 130, loss: 0.10458385199308395
step: 140, loss: 0.02400333806872368
step: 150, loss: 0.042396482080221176
step: 160, loss: 0.0007078899652697146
step: 170, loss: 0.030309146270155907
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
epoch 14: dev_f1=0.8780487804878048, f1=0.8909512761020881,
best_f1=0.9086538461538461
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 3.627972910180688e-05
step: 10, loss: 0.0007242856663651764
step: 20, loss: 7.275497046066448e-05
step: 30, loss: 0.000126652026665397
step: 40, loss: 0.00011885187996085733
step: 50, loss: 4.224668373353779e-05
step: 60, loss: 0.009714147076010704
step: 70, loss: 0.04420870915055275
step: 80, loss: 0.00553145119920373
step: 90, loss: 0.004372117109596729
step: 100, loss: 0.030186716467142105
step: 110, loss: 0.00012285371485631913
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 120, loss: 9.117510489886627e-05
step: 130, loss: 0.00016359762230422348
step: 140, loss: 0.015889663249254227
step: 150, loss: 0.0003933851548936218
step: 160, loss: 0.00011606898624449968
step: 170, loss: 0.003147586714476347
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
epoch 15: dev_f1=0.8753056234718827, f1=0.8878504672897196,
best_f1=0.9086538461538461
mv: cannot stat '*_dev.pt': No such file or directory
Selected optimization level 02: FP16 training with FP32 batchnorm and FP32
master weights.

```

Defaults for this optimization level are:

```

enabled           : True
opt_level         : 02
cast_model_type   : torch.float16
patch_torch_functions : False
keep_batchnorm_fp32 : True
master_weights    : True

```

```

loss_scale          : dynamic
Processing user overrides (additional kwargs that are not None)...
After processing overrides, optimization options are:
enabled             : True
opt_level           : 02
cast_model_type     : torch.float16
patch_torch_functions : False
keep_batchnorm_fp32  : True
master_weights      : True
loss_scale          : dynamic
Warning: multi_tensor_applier fused unscale kernel is unavailable, possibly
because apex was installed without --cuda_ext --cpp_ext. Using Python fallback.
Original ImportError was: ModuleNotFoundError("No module named 'amp_C'")
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
1916it [00:08, 224.60it/s]
load_f1 = 0.31471282454760025
real_f1 = 0.2923299565846599
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
4398it [00:32, 134.19it/s]

```

## 1.3 DA DIRTY

### 1.3.1 DBLP-ACM - Running the matcher

```

[ ]: # Step 1:
# train the wdc model with the save_model flag on
# it will produce two files *_dev.pt and *_test.pt which are the
# best checkpoints on the dev set and the test set.
!CUDA_VISIBLE_DEVICES=0 python train_ditto.py \
--task Dirty/DBLP-ACM \
--batch_size 32 \
--max_len 256 \
--lr 3e-5 \
--n_epochs 15 \
--finetuning \
--lm roberta \
--fp16 \
--save_model \
--da entry_swap \

```

```

--da span_shuffle \
--summarize

# Step 2: pick a checkpoint, rename it to wdc_all_small.pt
!mv *_dev.pt checkpoints/DBLP-ACM.pt

# Step 3: run the matcher
!CUDA_VISIBLE_DEVICES=0 python matcher.py \
--task Dirty/DBLP-ACM \
--input_path input/input_small.jsonl \
--output_path output/output_small.jsonl \
--lm roberta \
--use_gpu \
--fp16 \
--checkpoint_path checkpoints/

```

```

Downloading: 100% 481/481 [00:00<00:00, 442kB/s]
Downloading: 100% 899k/899k [00:01<00:00, 806kB/s]
Downloading: 100% 456k/456k [00:00<00:00, 503kB/s]
Downloading: 100% 501M/501M [00:14<00:00, 34.2MB/s]
Selected optimization level O2: FP16 training with FP32 batchnorm and FP32
master weights.

```

Defaults for this optimization level are:

```

enabled           : True
opt_level         : 02
cast_model_type   : torch.float16
patch_torch_functions : False
keep_batchnorm_fp32 : True
master_weights    : True
loss_scale        : dynamic

```

Processing user overrides (additional kwargs that are not None)...

After processing overrides, optimization options are:

```

enabled           : True
opt_level         : 02
cast_model_type   : torch.float16
patch_torch_functions : False
keep_batchnorm_fp32 : True
master_weights    : True
loss_scale        : dynamic

```

```

Warning: multi_tensor_applier fused unscale kernel is unavailable, possibly
because apex was installed without --cuda_ext --cpp_ext. Using Python fallback.
Original ImportError was: ModuleNotFoundError("No module named 'amp_C'")
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.

```

```

    warnings.warn("An input tensor was not cuda.")

```

```

Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 32768.0

```



```

/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:136:
UserWarning: Detected call of `lr_scheduler.step()` before `optimizer.step()`.
In PyTorch 1.1.0 and later, you should call them in the opposite order:
`optimizer.step()` before `lr_scheduler.step()`. Failure to do this will result
in PyTorch skipping the first value of the learning rate schedule. See more
details at https://pytorch.org/docs/stable/optim.html#how-to-adjust-learning-
rate
  "https://pytorch.org/docs/stable/optim.html#how-to-adjust-learning-rate",
UserWarning)
step: 0, loss: 0.5653526186943054
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 10, loss: 0.4643578827381134
step: 20, loss: 0.49901312589645386
step: 30, loss: 0.2882493734359741
step: 40, loss: 0.1141531690955162
step: 50, loss: 0.10771293938159943
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 8192.0
step: 60, loss: 0.06836507469415665
step: 70, loss: 0.157709538936615
step: 80, loss: 0.04314645379781723
step: 90, loss: 0.08492382615804672
step: 100, loss: 0.19916673004627228
step: 110, loss: 0.020319579169154167
step: 120, loss: 0.03591519966721535
step: 130, loss: 0.0678887739777565
step: 140, loss: 0.020156804472208023
step: 150, loss: 0.124221570789814
step: 160, loss: 0.004119172226637602
step: 170, loss: 0.015417952090501785
step: 180, loss: 0.07873085141181946
step: 190, loss: 0.004793692845851183
step: 200, loss: 0.006475328933447599
step: 210, loss: 0.10388005524873734
step: 220, loss: 0.1202043667435646
step: 230, loss: 0.010245664045214653
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
  warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 1: dev_f1=0.9655937846836848, f1=0.9719416386083053,
best_f1=0.9719416386083053
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.004158055409789085

```

```

step: 10, loss: 0.1534455269575119
step: 20, loss: 0.03199005872011185
step: 30, loss: 0.011666901409626007
step: 40, loss: 0.010388257913291454
step: 50, loss: 0.0019265325972810388
step: 60, loss: 0.007800444960594177
step: 70, loss: 0.00407247431576252
step: 80, loss: 0.0027694765012711287
step: 90, loss: 0.154832661151886
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 4096.0
step: 100, loss: 0.0022030312102288008
step: 110, loss: 0.045256536453962326
step: 120, loss: 0.007113532163202763
step: 130, loss: 0.008157373405992985
step: 140, loss: 0.0013281687861308455
step: 150, loss: 0.21103639900684357
step: 160, loss: 0.0065613300539553165
step: 170, loss: 0.0029702167958021164
step: 180, loss: 0.0047409385442733765
step: 190, loss: 0.0128141138702631
step: 200, loss: 0.09492477774620056
step: 210, loss: 0.023019704967737198
step: 220, loss: 0.009796560741961002
step: 230, loss: 0.007362592965364456
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
    warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 2: dev_f1=0.9670828603859251, f1=0.9576174112256587,
best_f1=0.9576174112256587
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00742878345772624
step: 10, loss: 0.0028591412119567394
step: 20, loss: 0.01805797964334488
step: 30, loss: 0.0033430629409849644
step: 40, loss: 0.027229689061641693
step: 50, loss: 0.013560730032622814
step: 60, loss: 0.006986552383750677
step: 70, loss: 0.019978182390332222
step: 80, loss: 0.11568071693181992
step: 90, loss: 0.006257195957005024
step: 100, loss: 0.010269597172737122
step: 110, loss: 0.01838352344930172

```

```

step: 120, loss: 0.0008026221767067909
step: 130, loss: 0.013825259171426296
step: 140, loss: 0.0014024681877344847
step: 150, loss: 0.07322575151920319
step: 160, loss: 0.012549437582492828
step: 170, loss: 0.0008474257192574441
step: 180, loss: 0.00936993770301342
step: 190, loss: 0.016006771475076675
step: 200, loss: 0.003666842123493552
step: 210, loss: 0.004146941937506199
step: 220, loss: 0.12329482287168503
step: 230, loss: 0.023037608712911606
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
    warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 3: dev_f1=0.9733333333333333, f1=0.9719416386083053,
best_f1=0.9719416386083053
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.013476537540555
step: 10, loss: 0.0028736700769513845
step: 20, loss: 0.0013018157333135605
step: 30, loss: 0.00215071439743042
step: 40, loss: 0.024962030351161957
step: 50, loss: 0.006122929975390434
step: 60, loss: 0.0023567313328385353
step: 70, loss: 0.03359336033463478
step: 80, loss: 0.057410918176174164
step: 90, loss: 0.13180236518383026
step: 100, loss: 0.009863422252237797
step: 110, loss: 0.0008840752416290343
step: 120, loss: 0.003703658701851964
step: 130, loss: 0.013613749295473099
step: 140, loss: 0.00906631164252758
step: 150, loss: 0.003431426826864481
step: 160, loss: 0.0036234466824680567
step: 170, loss: 0.001898562302812934
step: 180, loss: 0.14273816347122192
step: 190, loss: 0.002901380881667137
step: 200, loss: 0.1424737125635147
step: 210, loss: 0.10544876754283905
step: 220, loss: 0.00505143404006958
step: 230, loss: 0.005176607519388199

```

```

/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
    warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 4: dev_f1=0.9776286353467561, f1=0.9730941704035874,
best_f1=0.9730941704035874
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0013980745570734143
step: 10, loss: 0.0019892845302820206
step: 20, loss: 0.18822714686393738
step: 30, loss: 0.0021047142799943686
step: 40, loss: 0.003348980564624071
step: 50, loss: 0.0014935892540961504
step: 60, loss: 0.03899079188704491
step: 70, loss: 0.003470516763627529
step: 80, loss: 0.01121610403060913
step: 90, loss: 0.018371477723121643
step: 100, loss: 0.045755110681056976
step: 110, loss: 0.0049308668822050095
step: 120, loss: 0.00042575394036248326
step: 130, loss: 0.00030639031319878995
step: 140, loss: 0.0008790859719738364
step: 150, loss: 0.01858150213956833
step: 160, loss: 0.0104781873524189
step: 170, loss: 0.028934728354215622
step: 180, loss: 0.00495340907946229
step: 190, loss: 0.06208401173353195
step: 200, loss: 0.0038224037270992994
step: 210, loss: 0.02238515391945839
step: 220, loss: 0.003022256772965193
step: 230, loss: 0.013158362358808517
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
    warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 5: dev_f1=0.9853107344632768, f1=0.9684684684684683,
best_f1=0.9684684684684683
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")

```

```

step: 0, loss: 0.001610240200534463
step: 10, loss: 0.0008460261742584407
step: 20, loss: 0.004271899815648794
step: 30, loss: 0.0009202811634168029
step: 40, loss: 0.0004485475074034184
step: 50, loss: 0.0002883897686842829
step: 60, loss: 0.008057836443185806
step: 70, loss: 0.11616355925798416
step: 80, loss: 0.000694811053108424
step: 90, loss: 0.0013041668571531773
step: 100, loss: 0.0003693896869663149
step: 110, loss: 0.02369525283575058
step: 120, loss: 0.00525098480284214
step: 130, loss: 0.0018184977816417813
step: 140, loss: 0.00047194829676300287
step: 150, loss: 6.936363934073597e-05
step: 160, loss: 0.025868454948067665
step: 170, loss: 0.0012307530269026756
step: 180, loss: 0.0027830987237393856
step: 190, loss: 0.0003934645210392773
step: 200, loss: 0.004498567432165146
step: 210, loss: 0.0021000169217586517
step: 220, loss: 0.08456979691982269
step: 230, loss: 0.003507626708596945
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
epoch 6: dev_f1=0.9831271091113611, f1=0.9752252252252253,
best_f1=0.9684684684684683
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.004333187825977802
step: 10, loss: 0.0007021133205853403
step: 20, loss: 0.03083408996462822
step: 30, loss: 0.0006429990171454847
step: 40, loss: 0.0013718446716666222
step: 50, loss: 0.0010254248045384884
step: 60, loss: 0.00041364054777659476
step: 70, loss: 0.004508493468165398
step: 80, loss: 0.0006098246667534113
step: 90, loss: 0.00026060055824927986
step: 100, loss: 0.000632200506515801
step: 110, loss: 0.00040024309419095516
step: 120, loss: 0.0012033309321850538
step: 130, loss: 0.0038422029465436935
step: 140, loss: 0.0006442643352784216
step: 150, loss: 0.02689295820891857

```

```

step: 160, loss: 0.0003431091899983585
step: 170, loss: 0.0017261962639167905
step: 180, loss: 0.0003462855238467455
step: 190, loss: 0.001320339972153306
step: 200, loss: 0.017210375517606735
step: 210, loss: 0.0066986531019210815
step: 220, loss: 0.0024624832440167665
step: 230, loss: 0.003733329940587282
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
    warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 7: dev_f1=0.9864253393665158, f1=0.9761634506242906,
best_f1=0.9761634506242906
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0041058920323848724
step: 10, loss: 0.012205389328300953
step: 20, loss: 0.009488441050052643
step: 30, loss: 0.0027455806266516447
step: 40, loss: 0.0004516522749327123
step: 50, loss: 0.0003950442769564688
step: 60, loss: 0.0013167011784389615
step: 70, loss: 0.00016252833302132785
step: 80, loss: 0.04268491268157959
step: 90, loss: 0.0015831211348995566
step: 100, loss: 0.00030289552523754537
step: 110, loss: 0.0026659283321350813
step: 120, loss: 0.0003392602375242859
step: 130, loss: 0.0029551431071013212
step: 140, loss: 0.00014061795081943274
step: 150, loss: 0.1393090933561325
step: 160, loss: 0.008104812353849411
step: 170, loss: 0.0014796907780691981
step: 180, loss: 0.0002192731772083789
step: 190, loss: 0.000928745255805552
step: 200, loss: 0.002592414151877165
step: 210, loss: 0.00053770950762555
step: 220, loss: 0.0010343537433072925
step: 230, loss: 0.000499240995850414
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 8: dev_f1=0.9830890642615557, f1=0.9753363228699552,

```

```

best_f1=0.9761634506242906
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0012579535832628608
step: 10, loss: 0.0004367890360299498
step: 20, loss: 0.0006778659881092608
step: 30, loss: 0.0004483624070417136
step: 40, loss: 0.0001484711974626407
step: 50, loss: 9.501935710432008e-05
step: 60, loss: 7.185026333900169e-05
step: 70, loss: 0.032542020082473755
step: 80, loss: 0.0008709742687642574
step: 90, loss: 0.09094784408807755
step: 100, loss: 0.007792097050696611
step: 110, loss: 0.001751365838572383
step: 120, loss: 0.03858497738838196
step: 130, loss: 0.005653258878737688
step: 140, loss: 0.00010873347491724417
step: 150, loss: 8.498874376527965e-05
step: 160, loss: 0.00018453397206030786
step: 170, loss: 0.0003335446526762098
step: 180, loss: 0.0027600424364209175
step: 190, loss: 0.0003139347827527672
step: 200, loss: 0.00012024874740745872
step: 210, loss: 0.0026327897794544697
step: 220, loss: 0.001161748543381691
step: 230, loss: 0.0014144128654152155
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 9: dev_f1=0.9831271091113611, f1=0.9786276715410572,
best_f1=0.9761634506242906
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00035051049781031907
step: 10, loss: 0.0005257455050013959
step: 20, loss: 0.0003668394929263741
step: 30, loss: 0.00028221349930390716
step: 40, loss: 0.00019524095114320517
step: 50, loss: 9.039072028826922e-05
step: 60, loss: 0.00013045889500062913
step: 70, loss: 0.0006673886091448367
step: 80, loss: 0.002314646728336811
step: 90, loss: 4.715402246802114e-05
step: 100, loss: 3.908686267095618e-05
step: 110, loss: 7.431587437167764e-05

```

```

step: 120, loss: 0.0022174245677888393
step: 130, loss: 0.0060041905380785465
step: 140, loss: 0.0007108622812665999
step: 150, loss: 0.00043935145367868245
step: 160, loss: 3.035546797036659e-05
step: 170, loss: 0.00010765019396785647
step: 180, loss: 5.383405368775129e-05
step: 190, loss: 5.137881817063317e-05
step: 200, loss: 9.644669626140967e-05
step: 210, loss: 0.00010861812916118652
step: 220, loss: 0.0002521072456147522
step: 230, loss: 0.00010389746603323147
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 10: dev_f1=0.983050847457627, f1=0.9774266365688488,
best_f1=0.9761634506242906
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 3.5819026379613206e-05
step: 10, loss: 0.00013643948477692902
step: 20, loss: 5.2229108405299485e-05
step: 30, loss: 0.00013450872211251408
step: 40, loss: 2.041019433818292e-05
step: 50, loss: 7.675762753933668e-05
step: 60, loss: 0.015238584019243717
step: 70, loss: 0.0007568348664790392
step: 80, loss: 0.0015206231037154794
step: 90, loss: 0.02200561948120594
step: 100, loss: 0.0001221069396706298
step: 110, loss: 6.743847188772634e-05
step: 120, loss: 0.0004469526174943894
step: 130, loss: 6.611938442802057e-05
step: 140, loss: 0.0012633174192160368
step: 150, loss: 9.410416532773525e-05
step: 160, loss: 0.0392119325697422
step: 170, loss: 0.00017153109365608543
step: 180, loss: 0.00036813801852986217
step: 190, loss: 4.699225246440619e-05
step: 200, loss: 9.71916742855683e-05
step: 210, loss: 5.9675221564248204e-05
step: 220, loss: 9.950601088348776e-05
step: 230, loss: 0.00011614786490099505
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 11: dev_f1=0.9820224719101124, f1=0.9774774774774775,

```



```

best_f1=0.9761634506242906
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 8.80830775713548e-05
step: 10, loss: 4.166170765529387e-05
step: 20, loss: 0.00847566220909357
step: 30, loss: 0.034904513508081436
step: 40, loss: 4.6060977183515206e-05
step: 50, loss: 0.0016406916547566652
step: 60, loss: 0.0005571219953708351
step: 70, loss: 0.00014287656813394278
step: 80, loss: 7.375591667369008e-05
step: 90, loss: 0.0012291582534089684
step: 100, loss: 3.706867209984921e-05
step: 110, loss: 2.9633774829562753e-05
step: 120, loss: 4.283886300981976e-05
step: 130, loss: 0.00019098962366115302
step: 140, loss: 3.721033863257617e-05
step: 150, loss: 4.084591273567639e-05
step: 160, loss: 4.8360609071096405e-05
step: 170, loss: 6.451712397392839e-05
step: 180, loss: 3.3276457543252036e-05
step: 190, loss: 0.00353290350176394
step: 200, loss: 3.071011451538652e-05
step: 210, loss: 0.0009342096745967865
step: 220, loss: 0.01150739286094904
step: 230, loss: 3.662884410005063e-05
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 12: dev_f1=0.9832026875699889, f1=0.9765363128491621,
best_f1=0.9761634506242906
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00010500167991267517
step: 10, loss: 5.6467986723873764e-05
step: 20, loss: 4.315531623433344e-05
step: 30, loss: 3.713902333402075e-05
step: 40, loss: 0.00019760621944442391
step: 50, loss: 4.855987572227605e-05
step: 60, loss: 3.2304495107382536e-05
step: 70, loss: 3.009540341736283e-05
step: 80, loss: 2.3710859750281088e-05
step: 90, loss: 2.812135426211171e-05
step: 100, loss: 4.379855818115175e-05
step: 110, loss: 4.220545815769583e-05

```

```

step: 120, loss: 3.24424727295991e-05
step: 130, loss: 3.147392635582946e-05
step: 140, loss: 0.0003677031199913472
step: 150, loss: 3.1896463042357937e-05
step: 160, loss: 5.921244519413449e-05
step: 170, loss: 3.326189471408725e-05
step: 180, loss: 0.03282272815704346
step: 190, loss: 3.304973506601527e-05
step: 200, loss: 2.9391017960733734e-05
step: 210, loss: 3.4204251278424636e-05
step: 220, loss: 6.146913074189797e-05
step: 230, loss: 0.0002617955906316638
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 13: dev_f1=0.9810055865921787, f1=0.9776785714285714,
best_f1=0.9761634506242906
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 7.157656364142895e-05
step: 10, loss: 6.942651816643775e-05
step: 20, loss: 9.579179459251463e-05
step: 30, loss: 0.00011092681234003976
step: 40, loss: 5.6694534578127787e-05
step: 50, loss: 6.304492126218975e-05
step: 60, loss: 9.032605885295197e-05
step: 70, loss: 0.00010172421025345102
step: 80, loss: 0.00034548112307675183
step: 90, loss: 0.00021564462804235518
step: 100, loss: 5.167575727682561e-05
step: 110, loss: 6.0687125369440764e-05
step: 120, loss: 2.1456893591675907e-05
step: 130, loss: 9.113644773606211e-05
step: 140, loss: 0.0009024674072861671
step: 150, loss: 0.00025391957024112344
step: 160, loss: 4.566429925034754e-05
step: 170, loss: 7.322344754356891e-05
step: 180, loss: 4.916971010970883e-05
step: 190, loss: 3.30524671880994e-05
step: 200, loss: 6.231795123312622e-05
step: 210, loss: 3.938815643778071e-05
step: 220, loss: 4.329231887822971e-05
step: 230, loss: 7.392240513581783e-05
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 14: dev_f1=0.9832026875699889, f1=0.9765363128491621,

```

```

best_f1=0.9761634506242906
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.020531291142106056
step: 10, loss: 4.2815740016521886e-05
step: 20, loss: 4.50916777481325e-05
step: 30, loss: 0.00010912853758782148
step: 40, loss: 2.7383342967368662e-05
step: 50, loss: 3.463602115516551e-05
step: 60, loss: 0.0321800671517849
step: 70, loss: 3.9801783714210615e-05
step: 80, loss: 0.00040387571789324284
step: 90, loss: 0.00043759268010035157
step: 100, loss: 2.6656964109861292e-05
step: 110, loss: 4.207872916595079e-05
step: 120, loss: 0.04123203456401825
step: 130, loss: 3.64463085134048e-05
step: 140, loss: 0.007559966295957565
step: 150, loss: 0.00011226793139940128
step: 160, loss: 0.01275935210287571
step: 170, loss: 2.6999119654647075e-05
step: 180, loss: 3.82195103156846e-05
step: 190, loss: 0.0006591253913938999
step: 200, loss: 4.972760143573396e-05
step: 210, loss: 0.011676562018692493
step: 220, loss: 4.132174944970757e-05
step: 230, loss: 5.4359104979084805e-05
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
epoch 15: dev_f1=0.9832026875699889, f1=0.9765886287625419,
best_f1=0.9761634506242906
mv: cannot stat '*_dev.pt': No such file or directory
Selected optimization level 02:  FP16 training with FP32 batchnorm and FP32
master weights.

Defaults for this optimization level are:
enabled                : True
opt_level              : 02
cast_model_type       : torch.float16
patch_torch_functions : False
keep_batchnorm_fp32   : True
master_weights        : True
loss_scale             : dynamic
Processing user overrides (additional kwargs that are not None)...
After processing overrides, optimization options are:
enabled                : True

```

```

opt_level                : 02
cast_model_type          : torch.float16
patch_torch_functions    : False
keep_batchnorm_fp32      : True
master_weights           : True
loss_scale                : dynamic
Warning: multi_tensor_applier fused unscale kernel is unavailable, possibly
because apex was installed without --cuda_ext --cpp_ext. Using Python fallback.
Original ImportError was: ModuleNotFoundError("No module named 'amp_C'")
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
Oit [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
2473it [00:16, 146.08it/s]
load_f1 = 0.9831649831649831
real_f1 = 0.9843400447427293
Oit [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
4398it [00:31, 141.31it/s]

```

### 1.3.2 DBLP-GoogleScholar - Running the matcher

```

[ ]: # Step 1:
# train the wdc model with the save_model flag on
# it will produce two files *_dev.pt and *_test.pt which are the
# best checkpoints on the dev set and the test set.
!CUDA_VISIBLE_DEVICES=0 python train_ditto.py \
  --task Dirty/DBLP-GoogleScholar \
  --batch_size 32 \
  --max_len 256 \
  --lr 3e-5 \
  --n_epochs 15 \
  --finetuning \
  --lm roberta \
  --fp16 \
  --save_model \
  --da span_shuffle \
  --da attr_del \
  --summarize

# Step 2: pick a checkpoint, rename it to wdc_all_small.pt
!mv *_dev.pt checkpoints/DBLP-GoogleScholar/.pt

```

```
# Step 3: run the matcher
!CUDA_VISIBLE_DEVICES=0 python matcher.py \
  --task Dirty/DBLP-GoogleScholar \
  --input_path input/input_small.jsonl \
  --output_path output/output_small.jsonl \
  --lm roberta \
  --use_gpu \
  --fp16 \
  --checkpoint_path checkpoints/
```

Selected optimization level O2: FP16 training with FP32 batchnorm and FP32 master weights.

Defaults for this optimization level are:

```
enabled           : True
opt_level         : O2
cast_model_type   : torch.float16
patch_torch_functions : False
keep_batchnorm_fp32 : True
master_weights    : True
loss_scale        : dynamic
```

Processing user overrides (additional kwargs that are not None)...

After processing overrides, optimization options are:

```
enabled           : True
opt_level         : O2
cast_model_type   : torch.float16
patch_torch_functions : False
keep_batchnorm_fp32 : True
master_weights    : True
loss_scale        : dynamic
```

Warning: multi\_tensor\_applier fused unscale kernel is unavailable, possibly because apex was installed without --cuda\_ext --cpp\_ext. Using Python fallback. Original ImportError was: ModuleNotFoundError("No module named 'amp\_C'")  
/usr/local/lib/python3.7/dist-packages/apex/amp/\_initialize.py:25: UserWarning: An input tensor was not cuda.

```
warnings.warn("An input tensor was not cuda.")
```

```
step: 0, loss: 0.6220995783805847
```

Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 32768.0

Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0

```
step: 10, loss: 0.46018269658088684
```

```
step: 20, loss: 0.25874605774879456
```

```
step: 30, loss: 0.35186150670051575
```

```
step: 40, loss: 0.36562830209732056
```

```
step: 50, loss: 0.28664061427116394
```

```
step: 60, loss: 0.03824127838015556
```

```
step: 70, loss: 0.2442900389432907
```

```
step: 80, loss: 0.09753329306840897
```

Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 8192.0

step: 90, loss: 0.15068809688091278  
step: 100, loss: 0.3232463598251343  
step: 110, loss: 0.09268205612897873  
step: 120, loss: 0.23379811644554138  
step: 130, loss: 0.0900815799832344  
step: 140, loss: 0.36847588419914246  
step: 150, loss: 0.09800883382558823  
step: 160, loss: 0.05749029666185379  
step: 170, loss: 0.027111612260341644  
step: 180, loss: 0.04822789505124092  
step: 190, loss: 0.020641565322875977  
step: 200, loss: 0.05839652195572853  
step: 210, loss: 0.0330008789896965  
step: 220, loss: 0.08478628844022751  
step: 230, loss: 0.187966451048851  
step: 240, loss: 0.045401155948638916  
step: 250, loss: 0.0192114245146513  
step: 260, loss: 0.07591702789068222  
step: 270, loss: 0.3299505412578583  
step: 280, loss: 0.0339970625936985  
step: 290, loss: 0.06084625422954559  
step: 300, loss: 0.01541550736874342  
step: 310, loss: 0.02007734589278698  
step: 320, loss: 0.07142022997140884  
step: 330, loss: 0.1337699443101883  
step: 340, loss: 0.20856989920139313  
step: 350, loss: 0.06713208556175232  
step: 360, loss: 0.10663660615682602  
step: 370, loss: 0.1711544394493103  
step: 380, loss: 0.2961120307445526  
step: 390, loss: 0.014148079790174961  
step: 400, loss: 0.0653475746512413  
step: 410, loss: 0.26778629422187805  
step: 420, loss: 0.02768022008240223  
step: 430, loss: 0.04237481579184532  
step: 440, loss: 0.07951059937477112  
step: 450, loss: 0.06959761679172516  
step: 460, loss: 0.02261577732861042  
step: 470, loss: 0.015024599619209766  
step: 480, loss: 0.16755476593971252  
step: 490, loss: 0.09413103014230728  
step: 500, loss: 0.030003776773810387  
step: 510, loss: 0.0589117631316185  
step: 520, loss: 0.1581350713968277  
step: 530, loss: 0.004681815858930349

/usr/local/lib/python3.7/dist-packages/apex/amp/\_initialize.py:25: UserWarning:  
An input tensor was not cuda.

```

warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 1: dev_f1=0.9522484932777004, f1=0.9502093997208004,
best_f1=0.9502093997208004
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.040576476603746414
step: 10, loss: 0.06894225627183914
step: 20, loss: 0.07753065228462219
step: 30, loss: 0.0443691611289978
step: 40, loss: 0.053189102560281754
step: 50, loss: 0.047224175184965134
step: 60, loss: 0.019656790420413017
step: 70, loss: 0.011071533896028996
step: 80, loss: 0.008603844791650772
step: 90, loss: 0.006526530254632235
step: 100, loss: 0.0857081338763237
step: 110, loss: 0.005497694946825504
step: 120, loss: 0.04459374025464058
step: 130, loss: 0.01935845986008644
step: 140, loss: 0.05514761805534363
step: 150, loss: 0.06562155485153198
step: 160, loss: 0.010747167281806469
step: 170, loss: 0.022017408162355423
step: 180, loss: 0.031148554757237434
step: 190, loss: 0.015696141868829727
step: 200, loss: 0.25273391604423523
step: 210, loss: 0.03371093049645424
step: 220, loss: 0.002991314046084881
step: 230, loss: 0.09547992795705795
step: 240, loss: 0.08724504709243774
step: 250, loss: 0.007897819392383099
step: 260, loss: 0.053174637258052826
step: 270, loss: 0.005285949446260929
step: 280, loss: 0.07107830047607422
step: 290, loss: 0.06077795848250389
step: 300, loss: 0.027615470811724663
step: 310, loss: 0.031127719208598137
step: 320, loss: 0.07659658789634705
step: 330, loss: 0.07902789115905762
step: 340, loss: 0.043753527104854584
step: 350, loss: 0.002718232572078705
step: 360, loss: 0.046432893723249435
step: 370, loss: 0.012129968963563442

```

```
step: 380, loss: 0.07862125337123871
step: 390, loss: 0.0046530477702617645
step: 400, loss: 0.06518727540969849
step: 410, loss: 0.010180456563830376
step: 420, loss: 0.05368325859308243
step: 430, loss: 0.19712592661380768
step: 440, loss: 0.014726459980010986
step: 450, loss: 0.028228171169757843
step: 460, loss: 0.08414071798324585
step: 470, loss: 0.06451766937971115
step: 480, loss: 0.09529086947441101
step: 490, loss: 0.027209360152482986
step: 500, loss: 0.00574244000017643
step: 510, loss: 0.024654997512698174
step: 520, loss: 0.2913465201854706
step: 530, loss: 0.1314954310655594
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
epoch 2: dev_f1=0.9474174034434621, f1=0.949041608228144,
best_f1=0.9502093997208004
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.13315869867801666
step: 10, loss: 0.22156988084316254
step: 20, loss: 0.06189050152897835
step: 30, loss: 0.049689751118421555
step: 40, loss: 0.10523219406604767
step: 50, loss: 0.003618407528847456
step: 60, loss: 0.003657776629552245
step: 70, loss: 0.014558925293385983
step: 80, loss: 0.013875165022909641
step: 90, loss: 0.036888010799884796
step: 100, loss: 0.06241995468735695
step: 110, loss: 0.0969686508178711
step: 120, loss: 0.11877737939357758
step: 130, loss: 0.13370461761951447
step: 140, loss: 0.01298846397548914
step: 150, loss: 0.03753621131181717
step: 160, loss: 0.017298003658652306
step: 170, loss: 0.02086917869746685
step: 180, loss: 0.016072619706392288
step: 190, loss: 0.013509715907275677
step: 200, loss: 0.05185863748192787
step: 210, loss: 0.02484813518822193
step: 220, loss: 0.057166703045368195
step: 230, loss: 0.0061321319080889225
```



```
step: 240, loss: 0.017784319818019867
step: 250, loss: 0.0644318014383316
step: 260, loss: 0.15536494553089142
step: 270, loss: 0.02167317643761635
step: 280, loss: 0.0021576625294983387
step: 290, loss: 0.01791970618069172
step: 300, loss: 0.0922400951385498
step: 310, loss: 0.10344720631837845
step: 320, loss: 0.01092175580561161
step: 330, loss: 0.0014581003924831748
step: 340, loss: 0.0045140087604522705
step: 350, loss: 0.08345610648393631
step: 360, loss: 0.016220349818468094
step: 370, loss: 0.06355808675289154
step: 380, loss: 0.01561252772808075
step: 390, loss: 0.0028379641007632017
step: 400, loss: 0.12343228608369827
step: 410, loss: 0.02977079711854458
step: 420, loss: 0.022933494299650192
step: 430, loss: 0.008352674543857574
step: 440, loss: 0.18774336576461792
step: 450, loss: 0.011813710443675518
step: 460, loss: 0.046382710337638855
step: 470, loss: 0.04035310074687004
step: 480, loss: 0.18896448612213135
step: 490, loss: 0.017633827403187752
step: 500, loss: 0.022514402866363525
step: 510, loss: 0.005707920528948307
step: 520, loss: 0.001307376311160624
step: 530, loss: 0.005519031547009945
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 3: dev_f1=0.9522024367385192, f1=0.947219604147031,
best_f1=0.9502093997208004
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.04353241249918938
step: 10, loss: 0.00905388779938221
step: 20, loss: 0.11163633316755295
step: 30, loss: 0.0966157615184784
step: 40, loss: 0.006096699740737677
step: 50, loss: 0.09195584803819656
step: 60, loss: 0.01950271613895893
step: 70, loss: 0.007777578663080931
step: 80, loss: 0.07563381642103195
step: 90, loss: 0.12589994072914124
```

```
step: 100, loss: 0.0020456251222640276
step: 110, loss: 0.07431525737047195
step: 120, loss: 0.0077696568332612514
step: 130, loss: 0.029963750392198563
step: 140, loss: 0.014040226116776466
step: 150, loss: 0.00476118316873908
step: 160, loss: 0.04038294404745102
step: 170, loss: 0.02947939559817314
step: 180, loss: 0.15323378145694733
step: 190, loss: 0.013295622542500496
step: 200, loss: 0.17998452484607697
step: 210, loss: 0.0026890847366303205
step: 220, loss: 0.002279002917930484
step: 230, loss: 0.013276160694658756
step: 240, loss: 0.01901201717555523
step: 250, loss: 0.21865372359752655
step: 260, loss: 0.005761159583926201
step: 270, loss: 0.032934922724962234
step: 280, loss: 0.005360045004636049
step: 290, loss: 0.057967886328697205
step: 300, loss: 0.005537816323339939
step: 310, loss: 0.00458274781703949
step: 320, loss: 0.05277222767472267
step: 330, loss: 0.0966847762465477
step: 340, loss: 0.00474744476377964
step: 350, loss: 0.11277569830417633
step: 360, loss: 0.18099118769168854
step: 370, loss: 0.009770327247679234
step: 380, loss: 0.012434961274266243
step: 390, loss: 0.004044263623654842
step: 400, loss: 0.010403613559901714
step: 410, loss: 0.0005171055090613663
step: 420, loss: 0.012419437989592552
step: 430, loss: 0.006235524546355009
step: 440, loss: 0.0013011008268222213
step: 450, loss: 0.004336845595389605
step: 460, loss: 0.10511554032564163
step: 470, loss: 0.003929750993847847
step: 480, loss: 0.04776535555720329
step: 490, loss: 0.002322489395737648
step: 500, loss: 0.02544153295457363
step: 510, loss: 0.021702954545617104
step: 520, loss: 0.01384657807648182
step: 530, loss: 0.1792636215686798
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
epoch 4: dev_f1=0.952073732718894, f1=0.9486823855755894,
```

```
best_f1=0.9502093997208004
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.002079945057630539
step: 10, loss: 0.030505908653140068
step: 20, loss: 0.015127471648156643
step: 30, loss: 0.023599939420819283
step: 40, loss: 0.046567775309085846
step: 50, loss: 0.09429557621479034
step: 60, loss: 0.06236672028899193
step: 70, loss: 0.0036504403688013554
step: 80, loss: 0.021242761984467506
step: 90, loss: 0.1736477166414261
step: 100, loss: 0.11258286237716675
step: 110, loss: 0.08598726242780685
step: 120, loss: 0.08995289355516434
step: 130, loss: 0.04419514536857605
step: 140, loss: 0.010553301312029362
step: 150, loss: 0.09590096026659012
step: 160, loss: 0.009518500417470932
step: 170, loss: 0.04138444364070892
step: 180, loss: 0.020797571167349815
step: 190, loss: 0.0016423510387539864
step: 200, loss: 0.007392189931124449
step: 210, loss: 0.003846706822514534
step: 220, loss: 0.011505356058478355
step: 230, loss: 0.006631764583289623
step: 240, loss: 0.008939115330576897
step: 250, loss: 0.19143545627593994
step: 260, loss: 0.009812464006245136
step: 270, loss: 0.0033495284151285887
step: 280, loss: 0.001245236722752452
step: 290, loss: 0.007851913571357727
step: 300, loss: 0.11160670220851898
step: 310, loss: 0.04899410158395767
step: 320, loss: 0.008085322566330433
step: 330, loss: 0.001729214796796441
step: 340, loss: 0.052524399012327194
step: 350, loss: 0.0010377588914707303
step: 360, loss: 0.0002560149587225169
step: 370, loss: 0.0004293549573048949
step: 380, loss: 0.0006252863095141947
step: 390, loss: 0.06494998931884766
step: 400, loss: 0.050549909472465515
step: 410, loss: 0.06788583844900131
step: 420, loss: 0.2693500518798828
step: 430, loss: 0.012816214933991432
```

```

step: 440, loss: 0.004061530344188213
step: 450, loss: 0.007940136827528477
step: 460, loss: 0.06012168899178505
step: 470, loss: 0.0512610524892807
step: 480, loss: 0.07223386317491531
step: 490, loss: 0.048663489520549774
step: 500, loss: 0.0070884195156395435
step: 510, loss: 0.032938264310359955
step: 520, loss: 0.038586195558309555
step: 530, loss: 0.039740342646837234
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
epoch 5: dev_f1=0.9519774011299436, f1=0.945575011831519,
best_f1=0.9502093997208004
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.07841164618730545
step: 10, loss: 0.0007668411126360297
step: 20, loss: 0.003331171814352274
step: 30, loss: 0.0007364160846918821
step: 40, loss: 0.10373173654079437
step: 50, loss: 0.0018935787957161665
step: 60, loss: 0.010987627319991589
step: 70, loss: 0.0009452589438296854
step: 80, loss: 0.030199330300092697
step: 90, loss: 0.07229900360107422
step: 100, loss: 0.008940701372921467
step: 110, loss: 0.0008207168430089951
step: 120, loss: 0.053853750228881836
step: 130, loss: 0.005264386534690857
step: 140, loss: 0.00032568935421295464
step: 150, loss: 0.006772220600396395
step: 160, loss: 0.048691198229789734
step: 170, loss: 0.003907994367182255
step: 180, loss: 0.015257110819220543
step: 190, loss: 0.02723156474530697
step: 200, loss: 0.02766750194132328
step: 210, loss: 0.07634546607732773
step: 220, loss: 0.0021133925765752792
step: 230, loss: 0.0019214104395359755
step: 240, loss: 0.0005607249913737178
step: 250, loss: 0.01742023415863514
step: 260, loss: 0.001944981748238206
step: 270, loss: 0.017780225723981857
step: 280, loss: 0.02554284781217575
step: 290, loss: 0.01160943042486906

```

```

step: 300, loss: 0.012547031976282597
step: 310, loss: 0.06837674230337143
step: 320, loss: 0.0003337835951242596
step: 330, loss: 0.0016014830907806754
step: 340, loss: 0.007465733215212822
step: 350, loss: 0.013357097283005714
step: 360, loss: 0.0767502635717392
step: 370, loss: 0.04577222838997841
step: 380, loss: 0.0009716362692415714
step: 390, loss: 0.004556936677545309
step: 400, loss: 0.023634962737560272
step: 410, loss: 0.0004982881946489215
step: 420, loss: 0.0061807711608707905
step: 430, loss: 0.005353153683245182
step: 440, loss: 0.0003799407568294555
step: 450, loss: 0.29692354798316956
step: 460, loss: 0.0032828927505761385
step: 470, loss: 0.002069916343316436
step: 480, loss: 0.0029779395554214716
step: 490, loss: 0.005214491859078407
step: 500, loss: 0.044892650097608566
step: 510, loss: 0.09540855884552002
step: 520, loss: 0.0076578413136303425
step: 530, loss: 0.02947019226849079
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
  warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 6: dev_f1=0.9592592592592593, f1=0.9539594843462248,
best_f1=0.9539594843462248
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.004865831695497036
step: 10, loss: 0.0005167645285837352
step: 20, loss: 0.0027787310536950827
step: 30, loss: 0.03388155996799469
step: 40, loss: 0.0016274109948426485
step: 50, loss: 0.002122757025063038
step: 60, loss: 0.03493393957614899
step: 70, loss: 0.008183763362467289
step: 80, loss: 0.00209212233312428
step: 90, loss: 0.0017597883706912398
step: 100, loss: 0.03357226774096489
step: 110, loss: 0.0005024773417972028

```

```
step: 120, loss: 0.002457507885992527
step: 130, loss: 0.0003274019982200116
step: 140, loss: 0.02076239138841629
step: 150, loss: 0.0022055329754948616
step: 160, loss: 0.00014057620137464255
step: 170, loss: 0.0028056823648512363
step: 180, loss: 0.2732882499694824
step: 190, loss: 0.01347102876752615
step: 200, loss: 0.00263238325715065
step: 210, loss: 0.0011405821423977613
step: 220, loss: 0.0002050383045570925
step: 230, loss: 0.00011400476068956777
step: 240, loss: 0.018451685085892677
step: 250, loss: 0.02336827479302883
step: 260, loss: 0.005718936212360859
step: 270, loss: 0.002328599337488413
step: 280, loss: 0.03371158614754677
step: 290, loss: 0.0067588756792247295
step: 300, loss: 0.0005730839329771698
step: 310, loss: 0.0009155845618806779
step: 320, loss: 0.11490516364574432
step: 330, loss: 0.007901564240455627
step: 340, loss: 0.0028183001559227705
step: 350, loss: 0.0010569700971245766
step: 360, loss: 0.03445464000105858
step: 370, loss: 0.13122156262397766
step: 380, loss: 0.010344508104026318
step: 390, loss: 0.016688162460923195
step: 400, loss: 0.05002420023083687
step: 410, loss: 0.0009734589257277548
step: 420, loss: 0.008982553146779537
step: 430, loss: 0.00033725451794452965
step: 440, loss: 0.03149170055985451
step: 450, loss: 0.0013217985397204757
step: 460, loss: 0.0014570787316188216
step: 470, loss: 0.14313358068466187
step: 480, loss: 0.002441124524921179
step: 490, loss: 0.008696368895471096
step: 500, loss: 0.0029782853089272976
step: 510, loss: 0.001434776932001114
step: 520, loss: 0.013386916369199753
step: 530, loss: 0.0008901808760128915
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
epoch 7: dev_f1=0.955868544600939, f1=0.9557522123893806,
best_f1=0.9539594843462248
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
```

An input tensor was not cuda.

```
warnings.warn("An input tensor was not cuda.")
```

```
step: 0, loss: 0.0020342995412647724
step: 10, loss: 0.001812645117752254
step: 20, loss: 0.013897578231990337
step: 30, loss: 0.0029892087914049625
step: 40, loss: 0.0016081752255558968
step: 50, loss: 0.0006079890299588442
step: 60, loss: 0.0013634221395477653
step: 70, loss: 0.003463258733972907
step: 80, loss: 0.003927947953343391
step: 90, loss: 0.00014064155402593315
step: 100, loss: 0.000560895714443177
step: 110, loss: 0.0005220547318458557
step: 120, loss: 0.0011322714854031801
step: 130, loss: 0.008665384724736214
step: 140, loss: 0.004501112271100283
step: 150, loss: 0.0013001016341149807
step: 160, loss: 0.0052083986811339855
step: 170, loss: 0.11670540273189545
step: 180, loss: 0.06429964303970337
step: 190, loss: 0.03156013786792755
step: 200, loss: 0.007297051139175892
step: 210, loss: 0.06024012714624405
step: 220, loss: 0.005550413858145475
step: 230, loss: 0.08199623227119446
step: 240, loss: 0.00040549194090999663
step: 250, loss: 0.0010902980575338006
step: 260, loss: 0.0002817099157255143
step: 270, loss: 0.001599396811798215
step: 280, loss: 0.00022559650824405253
step: 290, loss: 0.0043789977207779884
step: 300, loss: 0.0001772077230270952
step: 310, loss: 0.002530162688344717
step: 320, loss: 0.0014532561181113124
step: 330, loss: 0.0013841389445587993
step: 340, loss: 0.0733814463019371
step: 350, loss: 0.0005485026631504297
step: 360, loss: 0.0025376868434250355
step: 370, loss: 0.05952279642224312
step: 380, loss: 0.0030205962248146534
step: 390, loss: 0.10238918662071228
step: 400, loss: 0.005253826733678579
step: 410, loss: 0.052986711263656616
step: 420, loss: 0.004019595682621002
step: 430, loss: 0.009568557143211365
step: 440, loss: 0.013012484647333622
step: 450, loss: 0.0015586597146466374
```

```

step: 460, loss: 0.005668224301189184
step: 470, loss: 0.07186820358037949
step: 480, loss: 0.05559270456433296
step: 490, loss: 0.0065951780416071415
step: 500, loss: 0.0017378490883857012
step: 510, loss: 0.01995081640779972
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 520, loss: 0.0007183534908108413
step: 530, loss: 0.005498358979821205
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
epoch 8: dev_f1=0.9560036663611365, f1=0.9540282203004097,
best_f1=0.9539594843462248
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.008239658549427986
step: 10, loss: 0.0012470752699300647
step: 20, loss: 0.00025235445355065167
step: 30, loss: 0.08577308058738708
step: 40, loss: 0.0033307583071291447
step: 50, loss: 0.00147683871909976
step: 60, loss: 0.0038855518214404583
step: 70, loss: 0.008324488997459412
step: 80, loss: 0.005849581677466631
step: 90, loss: 0.07447028160095215
step: 100, loss: 0.0025585477706044912
step: 110, loss: 0.011258771643042564
step: 120, loss: 0.0025761828292161226
step: 130, loss: 0.0012394149089232087
step: 140, loss: 0.0003596724709495902
step: 150, loss: 0.008177393116056919
step: 160, loss: 0.0058524697087705135
step: 170, loss: 0.0010043912334367633
step: 180, loss: 0.05635413900017738
step: 190, loss: 0.0015153108397498727
step: 200, loss: 0.0011215524282306433
step: 210, loss: 0.0031575867906212807
step: 220, loss: 0.0019500323105603456
step: 230, loss: 0.0016169605078175664
step: 240, loss: 0.0005732394056394696
step: 250, loss: 0.015053575858473778
step: 260, loss: 0.0015010717324912548
step: 270, loss: 0.0021461909636855125
step: 280, loss: 0.0036093133967369795
step: 290, loss: 0.0005573968519456685
step: 300, loss: 0.0008531014900654554

```



```

step: 310, loss: 0.06473781168460846
step: 320, loss: 0.0021443390287458897
step: 330, loss: 0.000726130441762507
step: 340, loss: 0.018763257190585136
step: 350, loss: 0.028458047658205032
step: 360, loss: 0.0020725687500089407
step: 370, loss: 0.0012813783250749111
step: 380, loss: 0.001970808021724224
step: 390, loss: 0.00031597967608831823
step: 400, loss: 0.13493435084819794
step: 410, loss: 0.008240445517003536
step: 420, loss: 0.0007146669668145478
step: 430, loss: 0.007373179774731398
step: 440, loss: 0.06212560832500458
step: 450, loss: 0.0017504239222034812
step: 460, loss: 0.00023983039136510342
step: 470, loss: 0.0006173981237225235
step: 480, loss: 0.0004911806900054216
step: 490, loss: 0.05328383296728134
step: 500, loss: 0.005797483492642641
step: 510, loss: 0.05856496840715408
step: 520, loss: 0.007212300784885883
step: 530, loss: 0.014489993453025818
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
    warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 9: dev_f1=0.9593721144967682, f1=0.9522497704315885,
best_f1=0.9522497704315885
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0016440245090052485
step: 10, loss: 0.0018900214927271008
step: 20, loss: 0.0003483092878013849
step: 30, loss: 0.0022123288363218307
step: 40, loss: 0.00030864326981827617
step: 50, loss: 0.0006785705918446183
step: 60, loss: 0.001634475076571107
step: 70, loss: 0.0002755924651864916
step: 80, loss: 0.0011286615626886487
step: 90, loss: 0.00015335330681409687
step: 100, loss: 0.0004717374686151743
step: 110, loss: 0.004087789915502071
step: 120, loss: 0.00017008223221637309

```

```

step: 130, loss: 0.00030676269670948386
step: 140, loss: 0.04448797181248665
step: 150, loss: 0.0002669829409569502
step: 160, loss: 0.1290857493877411
step: 170, loss: 0.0076614790596067905
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 8192.0
step: 180, loss: 0.01700947806239128
step: 190, loss: 0.0003561284684110433
step: 200, loss: 0.0006476484704762697
step: 210, loss: 0.005727415904402733
step: 220, loss: 0.0007362990872934461
step: 230, loss: 0.00042601002496667206
step: 240, loss: 0.00034206898999400437
step: 250, loss: 0.006025257054716349
step: 260, loss: 0.006573019549250603
step: 270, loss: 0.00011393261956982315
step: 280, loss: 0.00696506816893816
step: 290, loss: 0.0016315720276907086
step: 300, loss: 0.022633638232946396
step: 310, loss: 0.033074717968702316
step: 320, loss: 0.002355947159230709
step: 330, loss: 0.010685759596526623
step: 340, loss: 0.004485756158828735
step: 350, loss: 0.06691712886095047
step: 360, loss: 0.0001955145999090746
step: 370, loss: 0.0012187432730570436
step: 380, loss: 0.014316132292151451
step: 390, loss: 0.0006361127598211169
step: 400, loss: 0.0056075528264045715
step: 410, loss: 0.0046405247412621975
step: 420, loss: 0.0027243434451520443
step: 430, loss: 0.0003280726377852261
step: 440, loss: 0.00012503369362093508
step: 450, loss: 0.011926049366593361
step: 460, loss: 0.0008867542492225766
step: 470, loss: 0.007220722734928131
step: 480, loss: 0.026759065687656403
step: 490, loss: 0.0013186876894906163
step: 500, loss: 0.00037479374441318214
step: 510, loss: 0.008994472213089466
step: 520, loss: 0.10444801300764084
step: 530, loss: 0.010934092104434967
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
epoch 10: dev_f1=0.958256029684601, f1=0.954398894518655,
best_f1=0.9522497704315885
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:

```

An input tensor was not cuda.

warnings.warn("An input tensor was not cuda.")

step: 0, loss: 0.00044661349966190755  
step: 10, loss: 0.04984966292977333  
step: 20, loss: 0.010245734825730324  
step: 30, loss: 0.0004480077186599374  
step: 40, loss: 0.001831148169003427  
step: 50, loss: 0.005999900866299868  
step: 60, loss: 0.00290857395157218  
step: 70, loss: 0.0006035584374330938  
step: 80, loss: 0.0005897387745790184  
step: 90, loss: 0.002768036676570773  
step: 100, loss: 0.0008621340384706855  
step: 110, loss: 0.00019038061145693064  
step: 120, loss: 0.00032418829505331814  
step: 130, loss: 0.00032646849285811186  
step: 140, loss: 0.0004675605450756848  
step: 150, loss: 0.00016722897998988628  
step: 160, loss: 0.00043768290197476745  
step: 170, loss: 0.002397853648290038  
step: 180, loss: 0.00023248398792929947  
step: 190, loss: 0.00023847840202506632  
step: 200, loss: 0.00018098248983733356  
step: 210, loss: 0.00046371592907235026  
step: 220, loss: 0.00846363976597786  
step: 230, loss: 0.0001775924174580723  
step: 240, loss: 0.0019275384256616235  
step: 250, loss: 0.00014339199697133154  
step: 260, loss: 0.005799429025501013  
step: 270, loss: 0.0034391111694276333  
step: 280, loss: 0.0002900071849580854  
step: 290, loss: 0.00017194633255712688  
step: 300, loss: 0.00011467342847026885  
step: 310, loss: 0.0010030351113528013  
step: 320, loss: 0.003075862070545554  
step: 330, loss: 7.188013842096552e-05  
step: 340, loss: 0.04282508045434952  
step: 350, loss: 0.002609025686979294  
step: 360, loss: 0.005535015370696783  
step: 370, loss: 0.0008921112748794258  
step: 380, loss: 0.004241599701344967  
step: 390, loss: 0.0006784526631236076  
step: 400, loss: 0.005021393299102783  
step: 410, loss: 0.006404706742614508  
step: 420, loss: 0.010707276873290539  
step: 430, loss: 0.001255268114618957  
step: 440, loss: 0.036555830389261246  
step: 450, loss: 0.02864532172679901

```

step: 460, loss: 0.0002669476962182671
step: 470, loss: 0.0006914978730492294
step: 480, loss: 0.0019161371747031808
step: 490, loss: 0.000261978100752458
step: 500, loss: 0.0008798626950010657
step: 510, loss: 0.0006867655902169645
step: 520, loss: 0.00016837404109537601
step: 530, loss: 0.02657889574766159
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 11: dev_f1=0.9549632352941175, f1=0.9520328917313842,
best_f1=0.9522497704315885
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.011017844080924988
step: 10, loss: 0.0006289714947342873
step: 20, loss: 0.04125593602657318
step: 30, loss: 0.0002827949938364327
step: 40, loss: 0.005035103298723698
step: 50, loss: 0.002173352986574173
step: 60, loss: 0.013719767332077026
step: 70, loss: 0.003008890198543668
step: 80, loss: 0.004799809772521257
step: 90, loss: 0.003937371075153351
step: 100, loss: 0.11494971811771393
step: 110, loss: 0.03665545955300331
step: 120, loss: 0.0010109302820637822
step: 130, loss: 0.0029335960280150175
step: 140, loss: 0.008331472054123878
step: 150, loss: 0.00037514796713367105
step: 160, loss: 0.02575875259935856
step: 170, loss: 0.0026034582406282425
step: 180, loss: 0.00044122632243670523
step: 190, loss: 0.0014490587636828423
step: 200, loss: 0.0007217274978756905
step: 210, loss: 0.00011963244469370693
step: 220, loss: 0.00013921760546509176
step: 230, loss: 0.0023996413219720125
step: 240, loss: 0.001063894247636199
step: 250, loss: 0.028086131438612938
step: 260, loss: 0.0004932645242661238
step: 270, loss: 0.0005680341855622828
step: 280, loss: 0.0012092320248484612
step: 290, loss: 0.000738428148906678
step: 300, loss: 0.0013240536209195852
step: 310, loss: 0.0004868164542131126

```

```

step: 320, loss: 0.043001603335142136
step: 330, loss: 0.001334042870439589
step: 340, loss: 0.00048431314644403756
step: 350, loss: 0.0002325347304577008
step: 360, loss: 0.0002410233428236097
step: 370, loss: 0.0015515625709667802
step: 380, loss: 0.00046319200191646814
step: 390, loss: 0.004326803144067526
step: 400, loss: 0.001134891645051539
step: 410, loss: 0.0010992061579599977
step: 420, loss: 0.0003414530656300485
step: 430, loss: 0.006588819436728954
step: 440, loss: 0.03785236179828644
step: 450, loss: 0.025659654289484024
step: 460, loss: 0.00010628376912791282
step: 470, loss: 0.0020228633657097816
step: 480, loss: 0.0006256368942558765
step: 490, loss: 0.0066866022534668446
step: 500, loss: 0.0003001584846060723
step: 510, loss: 0.00020021709497086704
step: 520, loss: 0.005749882198870182
step: 530, loss: 0.0005122580332681537
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 12: dev_f1=0.9545667447306792, f1=0.9499298081422555,
best_f1=0.9522497704315885
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00011366670514689758
step: 10, loss: 0.00011939931573579088
step: 20, loss: 0.005529301706701517
step: 30, loss: 0.00011243348853895441
step: 40, loss: 0.013600563630461693
step: 50, loss: 0.0013273033546283841
step: 60, loss: 6.345040310407057e-05
step: 70, loss: 0.0021800154354423285
step: 80, loss: 0.0007888527470640838
step: 90, loss: 0.0001272865483770147
step: 100, loss: 0.00010088522685691714
step: 110, loss: 0.04884488508105278
step: 120, loss: 0.0008878840017132461
step: 130, loss: 0.0007281553116627038
step: 140, loss: 0.00034316154778935015
step: 150, loss: 0.001218644785694778
step: 160, loss: 0.00042515649693086743
step: 170, loss: 0.0034660357050597668

```

```

step: 180, loss: 0.00036349723814055324
step: 190, loss: 0.003572560613974929
step: 200, loss: 0.035524141043424606
step: 210, loss: 6.007847696309909e-05
step: 220, loss: 0.22416876256465912
step: 230, loss: 0.0008710871916264296
step: 240, loss: 0.002978821750730276
step: 250, loss: 0.0005371790612116456
step: 260, loss: 0.00011473797349026427
step: 270, loss: 0.038563989102840424
step: 280, loss: 0.0009046572959050536
step: 290, loss: 0.0005397353088483214
step: 300, loss: 0.00010214740905212238
step: 310, loss: 0.00022763226297684014
step: 320, loss: 0.00010660308180376887
step: 330, loss: 0.0008482438861392438
step: 340, loss: 0.003772516967728734
step: 350, loss: 0.0006723494152538478
step: 360, loss: 0.10870381444692612
step: 370, loss: 0.059214770793914795
step: 380, loss: 0.00035447406116873026
step: 390, loss: 0.0021403844002634287
step: 400, loss: 0.0002473237400408834
step: 410, loss: 0.00011544572043931112
step: 420, loss: 0.0031245406717061996
step: 430, loss: 0.0002445233112666756
step: 440, loss: 0.0011609173379838467
step: 450, loss: 0.0068633463233709335
step: 460, loss: 0.0028629221487790346
step: 470, loss: 0.008312264457345009
step: 480, loss: 7.274074596352875e-05
step: 490, loss: 0.007613101042807102
step: 500, loss: 0.0002043042768491432
step: 510, loss: 0.0002257469022879377
step: 520, loss: 0.00012039607099723071
step: 530, loss: 8.230777166318148e-05
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
epoch 13: dev_f1=0.957089552238806, f1=0.949767441860465,
best_f1=0.9522497704315885
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0005924998549744487
step: 10, loss: 0.0004305032780393958
step: 20, loss: 0.0014842270174995065
step: 30, loss: 0.0011697359150275588

```

step: 40, loss: 0.000628769223112613  
step: 50, loss: 0.00014237470168154687  
step: 60, loss: 0.0006040733424015343  
step: 70, loss: 0.00031986497924663126  
step: 80, loss: 0.0006707264692522585  
step: 90, loss: 0.00012362583947833627  
step: 100, loss: 0.00042085646418854594  
step: 110, loss: 0.0012498649302870035  
step: 120, loss: 4.4072610762668774e-05  
step: 130, loss: 0.000679337652400136  
step: 140, loss: 0.0007678211550228298  
step: 150, loss: 0.0016785112675279379  
step: 160, loss: 7.246969471452758e-05  
step: 170, loss: 0.007391317281872034  
step: 180, loss: 0.0005795789766125381  
step: 190, loss: 0.0012828508624807  
step: 200, loss: 5.764073284808546e-05  
step: 210, loss: 0.0013859412865713239  
step: 220, loss: 6.050572483218275e-05  
step: 230, loss: 0.00011807979171862826  
step: 240, loss: 0.0008198938448913395  
step: 250, loss: 0.0007626402657479048  
step: 260, loss: 0.0002346329129068181  
step: 270, loss: 0.0002986726467497647  
step: 280, loss: 0.00038479649811051786  
step: 290, loss: 0.01197914220392704  
step: 300, loss: 5.6875967857195064e-05  
step: 310, loss: 9.127744124270976e-05  
step: 320, loss: 4.351737879915163e-05  
step: 330, loss: 0.001615534769371152  
step: 340, loss: 0.0008611671510152519  
step: 350, loss: 5.4878128139534965e-05  
step: 360, loss: 0.0036951997317373753  
step: 370, loss: 0.00012066881754435599  
step: 380, loss: 0.0016077994368970394  
step: 390, loss: 7.155180355766788e-05  
step: 400, loss: 0.0013608563458546996  
step: 410, loss: 5.1511680794646963e-05  
step: 420, loss: 0.0002508389006834477  
step: 430, loss: 0.0001266900944756344  
step: 440, loss: 0.020914658904075623  
step: 450, loss: 9.67768719419837e-05  
step: 460, loss: 0.04740050807595253  
step: 470, loss: 0.11036046594381332  
step: 480, loss: 0.00024507686612196267  
step: 490, loss: 0.01538396067917347  
step: 500, loss: 0.00627095764502883  
step: 510, loss: 0.0359807163476944

```

step: 520, loss: 0.0013025426305830479
step: 530, loss: 0.0019181956304237247
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 14: dev_f1=0.9559573481687529, f1=0.9525126786537575,
best_f1=0.9522497704315885
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00010069518611999229
step: 10, loss: 0.0033658251632004976
step: 20, loss: 0.0008678161539137363
step: 30, loss: 0.00951469037681818
step: 40, loss: 0.005810786038637161
step: 50, loss: 0.0013346108607947826
step: 60, loss: 0.00015581419575028121
step: 70, loss: 9.522197797195986e-05
step: 80, loss: 0.0005335198366083205
step: 90, loss: 0.00010988918074872345
step: 100, loss: 0.0002188269718317315
step: 110, loss: 0.004680713173002005
step: 120, loss: 4.858368993154727e-05
step: 130, loss: 5.2322200644994155e-05
step: 140, loss: 0.0007352797547355294
step: 150, loss: 0.00012451267684809864
step: 160, loss: 0.000153423345182091
step: 170, loss: 7.035303860902786e-05
step: 180, loss: 0.01542546134442091
step: 190, loss: 0.0012278426438570023
step: 200, loss: 0.0014776376774534583
step: 210, loss: 0.0002994289679918438
step: 220, loss: 7.911283319117501e-05
step: 230, loss: 0.0005751786520704627
step: 240, loss: 0.0002054483920801431
step: 250, loss: 0.00484504597261548
step: 260, loss: 0.0005119965062476695
step: 270, loss: 0.00014684940106235445
step: 280, loss: 7.741591252852231e-05
step: 290, loss: 0.0016419898020103574
step: 300, loss: 4.225103111821227e-05
step: 310, loss: 0.061236560344696045
step: 320, loss: 0.0003194574383087456
step: 330, loss: 0.0008947330061346292
step: 340, loss: 4.721519144368358e-05
step: 350, loss: 0.0008297233725897968
step: 360, loss: 0.0007379455491900444
step: 370, loss: 0.11833175271749496

```



```

step: 380, loss: 0.00021077592100482434
step: 390, loss: 0.0001852939312811941
step: 400, loss: 0.0014831670559942722
step: 410, loss: 0.00014698770246468484
step: 420, loss: 0.002397325122728944
step: 430, loss: 4.269065539119765e-05
step: 440, loss: 0.006299130618572235
step: 450, loss: 0.003258979879319668
step: 460, loss: 0.006587498355656862
step: 470, loss: 4.904902016278356e-05
step: 480, loss: 0.006236694287508726
step: 490, loss: 0.00011217351129744202
step: 500, loss: 0.0018366625299677253
step: 510, loss: 5.608003993984312e-05
step: 520, loss: 4.050742063554935e-05
step: 530, loss: 6.383559957612306e-05
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
epoch 15: dev_f1=0.9560336763330215, f1=0.9510945505356311,
best_f1=0.9522497704315885
mv: cannot stat '*_dev.pt': No such file or directory
Selected optimization level 02:  FP16 training with FP32 batchnorm and FP32
master weights.

Defaults for this optimization level are:
enabled                : True
opt_level              : 02
cast_model_type        : torch.float16
patch_torch_functions  : False
keep_batchnorm_fp32    : True
master_weights         : True
loss_scale             : dynamic
Processing user overrides (additional kwargs that are not None)...
After processing overrides, optimization options are:
enabled                : True
opt_level              : 02
cast_model_type        : torch.float16
patch_torch_functions  : False
keep_batchnorm_fp32    : True
master_weights         : True
loss_scale             : dynamic
Warning:  multi_tensor_applier fused unscale kernel is unavailable, possibly
because apex was installed without --cuda_ext --cpp_ext. Using Python fallback.
Original ImportError was: ModuleNotFoundError("No module named 'amp_C'")
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")

```

```

0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
5742it [00:33, 172.48it/s]
load_f1 = 0.9587772116720705
real_f1 = 0.9562790697674419
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
4398it [00:30, 144.71it/s]

```

## 2 BASELINE

### 2.1 BASELINE STRUCTURED

#### 2.1.1 Beer - Running the matcher

```

[ ]: # Step 1:
# train the wdc model with the save_model flag on
# it will produce two files *_dev.pt and *_test.pt which are the
# best checkpoints on the dev set and the test set.
!CUDA_VISIBLE_DEVICES=0 python train_ditto.py \
  --task Structured/Beer \
  --batch_size 32 \
  --max_len 256 \
  --lr 3e-5 \
  --n_epochs 15 \
  --finetuning \
  --lm roberta \
  --fp16 \
  --save_model

# Step 2: pick a checkpoint, rename it to wdc_all_small.pt
!mv *_dev.pt checkpoints/Beer.pt

# Step 3: run the matcher
!CUDA_VISIBLE_DEVICES=0 python matcher.py \
  --task Structured/Beer \
  --input_path input/input_small.jsonl \
  --output_path output/output_small.jsonl \
  --lm roberta \
  --use_gpu \
  --fp16 \
  --checkpoint_path checkpoints/

```

Selected optimization level 02: FP16 training with FP32 batchnorm and FP32 master weights.

Defaults for this optimization level are:

```
enabled          : True
opt_level        : 02
cast_model_type  : torch.float16
patch_torch_functions : False
keep_batchnorm_fp32 : True
master_weights   : True
loss_scale       : dynamic
```

Processing user overrides (additional kwargs that are not None)...

After processing overrides, optimization options are:

```
enabled          : True
opt_level        : 02
cast_model_type  : torch.float16
patch_torch_functions : False
keep_batchnorm_fp32 : True
master_weights   : True
loss_scale       : dynamic
```

Warning: multi\_tensor\_applier fused unscale kernel is unavailable, possibly because apex was installed without --cuda\_ext --cpp\_ext. Using Python fallback. Original ImportError was: ModuleNotFoundError("No module named 'amp\_C'")  
/usr/local/lib/python3.7/dist-packages/apex/amp/\_initialize.py:25: UserWarning: An input tensor was not cuda.

```
warnings.warn("An input tensor was not cuda.")
```

Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 32768.0  
/usr/local/lib/python3.7/dist-packages/torch/optim/lr\_scheduler.py:136:

UserWarning: Detected call of `lr\_scheduler.step()` before `optimizer.step()`. In PyTorch 1.1.0 and later, you should call them in the opposite order: `optimizer.step()` before `lr\_scheduler.step()`. Failure to do this will result in PyTorch skipping the first value of the learning rate schedule. See more details at <https://pytorch.org/docs/stable/optim.html#how-to-adjust-learning-rate>

```
"https://pytorch.org/docs/stable/optim.html#how-to-adjust-learning-rate",
UserWarning)
```

```
step: 0, loss: 0.45558738708496094
```

Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0  
/usr/local/lib/python3.7/dist-packages/apex/amp/\_initialize.py:25: UserWarning: An input tensor was not cuda.

```
warnings.warn("An input tensor was not cuda.")
```

```
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
```

UserWarning: Please also save or load the state of the optimizer when saving or loading the scheduler.

```
warnings.warn(SAVE_STATE_WARNING, UserWarning)
```

```
epoch 1: dev_f1=0.29508196721311475, f1=0.3492063492063492,
best_f1=0.3492063492063492
```

/usr/local/lib/python3.7/dist-packages/apex/amp/\_initialize.py:25: UserWarning:

```

An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.46229082345962524
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
  warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 2: dev_f1=0.3529411764705882, f1=0.36, best_f1=0.36
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.41859403252601624
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
epoch 3: dev_f1=0.3384615384615385, f1=0.36923076923076925, best_f1=0.36
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.19331644475460052
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
  warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 4: dev_f1=0.3835616438356164, f1=0.41791044776119407,
best_f1=0.41791044776119407
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.2679484784603119
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 8192.0
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
  warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 5: dev_f1=0.5405405405405405, f1=0.45, best_f1=0.45
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")

```

```

step: 0, loss: 0.3793691396713257
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
    warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 6: dev_f1=0.8, f1=0.7878787878787878, best_f1=0.7878787878787878
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.24580125510692596
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
    warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 7: dev_f1=0.896551724137931, f1=0.7777777777777778,
best_f1=0.7777777777777778
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.09316761791706085
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
    warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 8: dev_f1=0.9285714285714286, f1=0.8484848484848484,
best_f1=0.8484848484848484
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.1037096306681633
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 9: dev_f1=0.9285714285714286, f1=0.9333333333333333,
best_f1=0.8484848484848484
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.09037686139345169

```

```

/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 10: dev_f1=0.9285714285714286, f1=0.9032258064516129,
best_f1=0.8484848484848484
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.1205442026257515
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 11: dev_f1=0.8666666666666666, f1=0.9333333333333333,
best_f1=0.8484848484848484
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.007705621886998415
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 12: dev_f1=0.896551724137931, f1=0.9333333333333333,
best_f1=0.8484848484848484
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0043306718580424786
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 13: dev_f1=0.896551724137931, f1=0.9333333333333333,
best_f1=0.8484848484848484
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.006871100515127182
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 14: dev_f1=0.896551724137931, f1=0.9333333333333333,
best_f1=0.8484848484848484
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00536206504330039
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")

```

```
epoch 15: dev_f1=0.896551724137931, f1=0.9333333333333333,
best_f1=0.8484848484848484
mv: cannot stat '*_dev.pt': No such file or directory
Selected optimization level O2:  FP16 training with FP32 batchnorm and FP32
master weights.
```

Defaults for this optimization level are:

```
enabled          : True
opt_level        : 02
cast_model_type  : torch.float16
patch_torch_functions : False
keep_batchnorm_fp32 : True
master_weights   : True
loss_scale       : dynamic
```

Processing user overrides (additional kwargs that are not None)...

After processing overrides, optimization options are:

```
enabled          : True
opt_level        : 02
cast_model_type  : torch.float16
patch_torch_functions : False
keep_batchnorm_fp32 : True
master_weights   : True
loss_scale       : dynamic
```

Warning: multi\_tensor\_applier fused unscale kernel is unavailable, possibly because apex was installed without --cuda\_ext --cpp\_ext. Using Python fallback.

Original ImportError was: ModuleNotFoundError("No module named 'amp\_C'")

/usr/local/lib/python3.7/dist-packages/apex/amp/\_initialize.py:25: UserWarning:

An input tensor was not cuda.

```
warnings.warn("An input tensor was not cuda.")
```

91it [00:00, 86411.97it/s]

/usr/local/lib/python3.7/dist-packages/apex/amp/\_initialize.py:25: UserWarning:

An input tensor was not cuda.

```
warnings.warn("An input tensor was not cuda.")
```

load\_f1 = 0.9285714285714286

real\_f1 = 0.9285714285714286

0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-

packages/apex/amp/\_initialize.py:25: UserWarning: An input tensor was not cuda.

```
warnings.warn("An input tensor was not cuda.")
```

4398it [00:31, 137.55it/s]

## 2.1.2 DBLP-ACM - Running the matcher

```
[ ]: # Step 1:
      # train the wdc model with the save_model flag on
      # it will produce two files *_dev.pt and *_test.pt which are the
      # best checkpoints on the dev set and the test set.
```

```

!CUDA_VISIBLE_DEVICES=0 python train_ditto.py \
--task Structured/DBLP-ACM \
--batch_size 32 \
--max_len 256 \
--lr 3e-5 \
--n_epochs 15 \
--finetuning \
--lm roberta \
--fp16 \
--save_model

# Step 2: pick a checkpoint, rename it to wdc_all_small.pt
!mv *_dev.pt checkpoints/DBLP-ACM.pt

# Step 3: run the matcher
!CUDA_VISIBLE_DEVICES=0 python matcher.py \
--task Structured/DBLP-ACM \
--input_path input/input_small.jsonl \
--output_path output/output_small.jsonl \
--lm roberta \
--use_gpu \
--fp16 \
--checkpoint_path checkpoints/

```

Selected optimization level 02: FP16 training with FP32 batchnorm and FP32 master weights.

Defaults for this optimization level are:

```

enabled           : True
opt_level         : 02
cast_model_type   : torch.float16
patch_torch_functions : False
keep_batchnorm_fp32 : True
master_weights    : True
loss_scale        : dynamic

```

Processing user overrides (additional kwargs that are not None)...

After processing overrides, optimization options are:

```

enabled           : True
opt_level         : 02
cast_model_type   : torch.float16
patch_torch_functions : False
keep_batchnorm_fp32 : True
master_weights    : True
loss_scale        : dynamic

```

Warning: multi\_tensor\_applier fused unscale kernel is unavailable, possibly because apex was installed without --cuda\_ext --cpp\_ext. Using Python fallback. Original ImportError was: ModuleNotFoundError("No module named 'amp\_C'")



```

/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.5989105105400085
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 32768.0
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 10, loss: 0.5612415671348572
step: 20, loss: 0.5401861667633057
step: 30, loss: 0.27648764848709106
step: 40, loss: 0.3303438127040863
step: 50, loss: 0.6211916208267212
step: 60, loss: 0.4691833257675171
step: 70, loss: 0.40822964906692505
step: 80, loss: 0.5833286643028259
step: 90, loss: 0.20114406943321228
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 8192.0
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 4096.0
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 2048.0
step: 100, loss: 0.7094266414642334
step: 110, loss: 0.5356987118721008
step: 120, loss: 0.40811777114868164
step: 130, loss: 0.06902533769607544
step: 140, loss: 0.17354393005371094
step: 150, loss: 0.16554084420204163
step: 160, loss: 0.010831266641616821
step: 170, loss: 0.07784163951873779
step: 180, loss: 0.1833755522966385
step: 190, loss: 0.021450506523251534
step: 200, loss: 0.047976285219192505
step: 210, loss: 0.027953559532761574
step: 220, loss: 0.13470974564552307
step: 230, loss: 0.03073183447122574
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
    warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 1: dev_f1=0.958659217877095, f1=0.9684684684684683,
best_f1=0.9684684684684683
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0031366157345473766
step: 10, loss: 0.1480739265680313
step: 20, loss: 0.03846191242337227
step: 30, loss: 0.02299189753830433

```

```

step: 40, loss: 0.015359929762780666
step: 50, loss: 0.0032919629011303186
step: 60, loss: 0.0028804743196815252
step: 70, loss: 0.00826624222099781
step: 80, loss: 0.0020373633597046137
step: 90, loss: 0.0032388619147241116
step: 100, loss: 0.006173631642013788
step: 110, loss: 0.002884246874600649
step: 120, loss: 0.00749822985380888
step: 130, loss: 0.019778529182076454
step: 140, loss: 0.0011629884829744697
step: 150, loss: 0.03012288361787796
step: 160, loss: 0.006452504545450211
step: 170, loss: 0.0017934325151145458
step: 180, loss: 0.002940745558589697
step: 190, loss: 0.001279144431464374
step: 200, loss: 0.02146335318684578
step: 210, loss: 0.008200593292713165
step: 220, loss: 0.004838415887206793
step: 230, loss: 0.0020427058916538954
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
    warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 2: dev_f1=0.9732142857142857, f1=0.9743016759776536,
best_f1=0.9743016759776536
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.009824033826589584
step: 10, loss: 0.003004397265613079
step: 20, loss: 0.001242601196281612
step: 30, loss: 0.0021336995996534824
step: 40, loss: 0.007519421633332968
step: 50, loss: 0.01295954454690218
step: 60, loss: 0.005699780769646168
step: 70, loss: 0.004441419616341591
step: 80, loss: 0.16532468795776367
step: 90, loss: 0.0029526259750127792
step: 100, loss: 0.001993587240576744
step: 110, loss: 0.01029951311647892
step: 120, loss: 0.004613158758729696
step: 130, loss: 0.038323625922203064
step: 140, loss: 0.0008731122361496091
step: 150, loss: 0.009325896389782429

```

```

step: 160, loss: 0.0030512413941323757
step: 170, loss: 0.0008945261361077428
step: 180, loss: 0.008663469925522804
step: 190, loss: 0.02109738066792488
step: 200, loss: 0.003128986805677414
step: 210, loss: 0.0020465096458792686
step: 220, loss: 0.003710263641551137
step: 230, loss: 0.0275584626942873
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
    warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 3: dev_f1=0.9757709251101322, f1=0.9639344262295082,
best_f1=0.9639344262295082
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.002738357987254858
step: 10, loss: 0.001017172122374177
step: 20, loss: 0.002525061136111617
step: 30, loss: 0.002065344713628292
step: 40, loss: 0.033749137073755264
step: 50, loss: 0.0026460099034011364
step: 60, loss: 0.0032017268240451813
step: 70, loss: 0.004259426612406969
step: 80, loss: 0.0012629893608391285
step: 90, loss: 0.06716633588075638
step: 100, loss: 0.0009052269160747528
step: 110, loss: 0.00057419971562922
step: 120, loss: 0.04819945991039276
step: 130, loss: 0.0326153002679348
step: 140, loss: 0.0024596904404461384
step: 150, loss: 0.005140560679137707
step: 160, loss: 0.0016310716746374965
step: 170, loss: 0.003108698409050703
step: 180, loss: 0.10970981419086456
step: 190, loss: 0.00192246213555336
step: 200, loss: 0.16041986644268036
step: 210, loss: 0.002132433233782649
step: 220, loss: 0.0011998737463727593
step: 230, loss: 0.04939480870962143
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:

```

UserWarning: Please also save or load the state of the optimizer when saving or loading the scheduler.

```
warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 4: dev_f1=0.9910313901345291, f1=0.9843749999999999,
best_f1=0.9843749999999999
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
```

```
warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0011631299275904894
step: 10, loss: 0.0011640320299193263
step: 20, loss: 0.01608184352517128
step: 30, loss: 0.0013618721859529614
step: 40, loss: 0.0028459844179451466
step: 50, loss: 0.017906032502651215
step: 60, loss: 0.0015554871642962098
step: 70, loss: 0.0014077466912567616
step: 80, loss: 0.07249671965837479
step: 90, loss: 0.04435456916689873
step: 100, loss: 0.001258311909623444
step: 110, loss: 0.005365455057471991
step: 120, loss: 0.00033639383036643267
step: 130, loss: 0.0029716994613409042
step: 140, loss: 0.008396903984248638
step: 150, loss: 0.06038181483745575
step: 160, loss: 0.00032360057230107486
step: 170, loss: 0.0005670661921612918
step: 180, loss: 0.0005817526835016906
step: 190, loss: 0.06160658970475197
step: 200, loss: 0.1531653106212616
step: 210, loss: 0.005923604127019644
step: 220, loss: 0.003058329224586487
step: 230, loss: 0.01673269271850586
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
```

```
warnings.warn("An input tensor was not cuda.")
epoch 5: dev_f1=0.9854096520763187, f1=0.9854096520763187,
best_f1=0.9843749999999999
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
```

```
warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0007240509148687124
step: 10, loss: 0.0015528338262811303
step: 20, loss: 0.002868335461243987
step: 30, loss: 0.0012362999841570854
step: 40, loss: 0.00045130131184123456
step: 50, loss: 0.0006443949532695115
step: 60, loss: 0.0009274479234591126
step: 70, loss: 0.0015263656387105584
```

```

step: 80, loss: 0.0006588458782061934
step: 90, loss: 0.0027202023193240166
step: 100, loss: 0.034248366951942444
step: 110, loss: 0.057503100484609604
step: 120, loss: 0.001424951246008277
step: 130, loss: 0.002185423392802477
step: 140, loss: 0.0013258948456496
step: 150, loss: 0.0002902027918025851
step: 160, loss: 0.005548511166125536
step: 170, loss: 0.0009133715066127479
step: 180, loss: 0.0003237794735468924
step: 190, loss: 0.0009166253730654716
step: 200, loss: 0.0036450885236263275
step: 210, loss: 0.0017357670003548265
step: 220, loss: 0.010004700161516666
step: 230, loss: 0.0013111160369589925
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 6: dev_f1=0.9831649831649831, f1=0.9820627802690582,
best_f1=0.9843749999999999
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0012875701067969203
step: 10, loss: 0.0004689678899012506
step: 20, loss: 0.00020441495871637017
step: 30, loss: 0.0002090372727252543
step: 40, loss: 0.0013242049608379602
step: 50, loss: 0.000665652914904058
step: 60, loss: 0.00044301999150775373
step: 70, loss: 0.0010397378355264664
step: 80, loss: 0.0003990131663158536
step: 90, loss: 0.0004015109152533114
step: 100, loss: 0.0002542596193961799
step: 110, loss: 0.0002673449052963406
step: 120, loss: 0.00033143151085823774
step: 130, loss: 0.0020597833208739758
step: 140, loss: 0.0005280699115246534
step: 150, loss: 0.02271730825304985
step: 160, loss: 0.0004667823086492717
step: 170, loss: 0.0003695274062920362
step: 180, loss: 0.0001902406947920099
step: 190, loss: 0.00026363329379819334
step: 200, loss: 0.002698127645999193
step: 210, loss: 0.007625384256243706
step: 220, loss: 0.0019426788203418255
step: 230, loss: 0.00577044440433383

```

```

/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 7: dev_f1=0.9887892376681614, f1=0.9832402234636871,
best_f1=0.9843749999999999
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.004417187999933958
step: 10, loss: 0.0014964895090088248
step: 20, loss: 0.00025685253785923123
step: 30, loss: 0.00021261375513859093
step: 40, loss: 0.00036853726487606764
step: 50, loss: 0.000188512698514387
step: 60, loss: 0.0002023606066359207
step: 70, loss: 0.0001248459448106587
step: 80, loss: 0.032404012978076935
step: 90, loss: 0.00029604037990793586
step: 100, loss: 0.00034224087721668184
step: 110, loss: 0.007348833605647087
step: 120, loss: 0.002462110249325633
step: 130, loss: 0.0010706016328185797
step: 140, loss: 0.00037429857184179127
step: 150, loss: 0.05393144115805626
step: 160, loss: 0.00042556560947559774
step: 170, loss: 0.0003039255680050701
step: 180, loss: 0.0003877721610479057
step: 190, loss: 0.0005272007547318935
step: 200, loss: 0.0002259925240650773
step: 210, loss: 0.0013750084908679128
step: 220, loss: 0.0003582006029319018
step: 230, loss: 0.0007855491130612791
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 8: dev_f1=0.9865168539325843, f1=0.9767441860465117,
best_f1=0.9843749999999999
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0005201890016905963
step: 10, loss: 0.0007307572523131967
step: 20, loss: 0.0005034747300669551
step: 30, loss: 0.0006478161667473614
step: 40, loss: 0.007347366772592068
step: 50, loss: 0.003224710701033473
step: 60, loss: 0.0003940419410355389
step: 70, loss: 0.04528970271348953

```

```

step: 80, loss: 0.0013060008641332388
step: 90, loss: 0.05093928426504135
step: 100, loss: 0.00022722153516951948
step: 110, loss: 0.00014500340330414474
step: 120, loss: 0.02434440702199936
step: 130, loss: 0.0002763473894447088
step: 140, loss: 0.00027460281853564084
step: 150, loss: 0.0005864472477696836
step: 160, loss: 0.0005406261188909411
step: 170, loss: 0.0001621503324713558
step: 180, loss: 0.0002810795558616519
step: 190, loss: 0.0001368898811051622
step: 200, loss: 0.00024641057825647295
step: 210, loss: 0.0013570651644840837
step: 220, loss: 0.0004825804498977959
step: 230, loss: 0.00018812935741152614
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
    warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 9: dev_f1=0.9910514541387023, f1=0.9866666666666666,
best_f1=0.9866666666666666
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00026611899374984205
step: 10, loss: 0.00038096148637123406
step: 20, loss: 0.0002196576097048819
step: 30, loss: 0.0001817287557059899
step: 40, loss: 0.0010186841245740652
step: 50, loss: 0.00027282597147859633
step: 60, loss: 0.0006467801867984235
step: 70, loss: 0.0009546633809804916
step: 80, loss: 0.00018504101899452507
step: 90, loss: 0.0004927964764647186
step: 100, loss: 0.0004917503683827817
step: 110, loss: 0.013095632195472717
step: 120, loss: 0.00021303810353856534
step: 130, loss: 0.0004574639315251261
step: 140, loss: 0.0005635945126414299
step: 150, loss: 0.0035353058483451605
step: 160, loss: 0.00010016648593591526
step: 170, loss: 0.0002372002782067284
step: 180, loss: 0.0009802111890166998
step: 190, loss: 0.0005290290573611856

```

```

step: 200, loss: 0.0011180980363860726
step: 210, loss: 0.0005799478967674077
step: 220, loss: 0.030736098065972328
step: 230, loss: 0.00016385811613872647
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 10: dev_f1=0.9910514541387023, f1=0.9833147942157954,
best_f1=0.9866666666666666
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00011255406570853665
step: 10, loss: 0.0002606013440527022
step: 20, loss: 0.00024725537514314055
step: 30, loss: 0.0003244333202019334
step: 40, loss: 5.556818359764293e-05
step: 50, loss: 0.00030562555184587836
step: 60, loss: 0.008643551729619503
step: 70, loss: 0.00036098837153986096
step: 80, loss: 0.0070640225894749165
step: 90, loss: 0.013012617826461792
step: 100, loss: 0.0008664233610033989
step: 110, loss: 0.0005961524439044297
step: 120, loss: 0.0005412950413301587
step: 130, loss: 0.000562345958314836
step: 140, loss: 0.006927298381924629
step: 150, loss: 0.0005915404763072729
step: 160, loss: 0.04558851942420006
step: 170, loss: 0.0011530567426234484
step: 180, loss: 0.0006092570838518441
step: 190, loss: 0.00030510526266880333
step: 200, loss: 0.0011025579879060388
step: 210, loss: 0.00026960481773130596
step: 220, loss: 0.00030451943166553974
step: 230, loss: 0.0002756322792265564
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 11: dev_f1=0.9910514541387023, f1=0.9855715871254161,
best_f1=0.9866666666666666
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00029331358382478356
step: 10, loss: 0.0001741125452099368
step: 20, loss: 0.011883888393640518
step: 30, loss: 0.034485846757888794

```



```

step: 40, loss: 0.0003745924914255738
step: 50, loss: 0.000590460782404989
step: 60, loss: 0.00032472473685629666
step: 70, loss: 0.00014764499792363495
step: 80, loss: 8.329077536473051e-05
step: 90, loss: 0.00019837342551909387
step: 100, loss: 8.997447730507702e-05
step: 110, loss: 7.083793025230989e-05
step: 120, loss: 0.00013602752005681396
step: 130, loss: 9.177423635264859e-05
step: 140, loss: 0.00010404812201159075
step: 150, loss: 0.00012134180724387988
step: 160, loss: 0.001184165826998651
step: 170, loss: 0.00012575836444739252
step: 180, loss: 0.00010427006054669619
step: 190, loss: 0.0002743953955359757
step: 200, loss: 0.00010852736886590719
step: 210, loss: 0.00013949902495369315
step: 220, loss: 0.014963945373892784
step: 230, loss: 0.00014452211325988173
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
epoch 12: dev_f1=0.9910514541387023, f1=0.9821826280623607,
best_f1=0.9866666666666666
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.06675239652395248
step: 10, loss: 0.00039761324296705425
step: 20, loss: 0.00024127110373228788
step: 30, loss: 0.002408253261819482
step: 40, loss: 0.000242241847445257
step: 50, loss: 0.0003409437194932252
step: 60, loss: 0.00014636678679380566
step: 70, loss: 7.552713213954121e-05
step: 80, loss: 6.766025762772188e-05
step: 90, loss: 8.344749949174002e-05
step: 100, loss: 0.0002203705080319196
step: 110, loss: 9.827035682974383e-05
step: 120, loss: 0.00010834734712261707
step: 130, loss: 0.00010858116002054885
step: 140, loss: 0.0002821226662490517
step: 150, loss: 3.292943802080117e-05
step: 160, loss: 0.005335872061550617
step: 170, loss: 7.957371417433023e-05
step: 180, loss: 0.039241231977939606
step: 190, loss: 0.00012826899182982743

```

```

step: 200, loss: 3.0489689379464835e-05
step: 210, loss: 0.00014843793178442866
step: 220, loss: 8.839017391437665e-05
step: 230, loss: 8.255915599875152e-05
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 13: dev_f1=0.9910112359550561, f1=0.987598647125141,
best_f1=0.9866666666666666
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 7.546253618784249e-05
step: 10, loss: 5.1884046115446836e-05
step: 20, loss: 6.200120697030798e-05
step: 30, loss: 0.00011072041525039822
step: 40, loss: 5.375055116019212e-05
step: 50, loss: 7.237934914883226e-05
step: 60, loss: 5.2692779718199745e-05
step: 70, loss: 9.610201232135296e-05
step: 80, loss: 4.8918711399892345e-05
step: 90, loss: 0.0007046370301395655
step: 100, loss: 0.0001063542440533638
step: 110, loss: 7.956640183692798e-05
step: 120, loss: 2.449253406666685e-05
step: 130, loss: 0.00017193981329910457
step: 140, loss: 6.24028907623142e-05
step: 150, loss: 0.0002927746681962162
step: 160, loss: 0.00010442516213515773
step: 170, loss: 8.812749001663178e-05
step: 180, loss: 6.560831388924271e-05
step: 190, loss: 6.089786984375678e-05
step: 200, loss: 0.00021872618526685983
step: 210, loss: 6.411103822756559e-05
step: 220, loss: 6.60723599139601e-05
step: 230, loss: 3.674708568723872e-05
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
    warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 14: dev_f1=0.9921259842519685, f1=0.9876265466816648,
best_f1=0.9876265466816648
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")

```

```

step: 0, loss: 0.02201012708246708
step: 10, loss: 6.744319398421794e-05
step: 20, loss: 0.00016940583009272814
step: 30, loss: 8.82823733263649e-05
step: 40, loss: 4.647558671422303e-05
step: 50, loss: 7.529809226980433e-05
step: 60, loss: 0.0249770600348711
step: 70, loss: 6.280160596361384e-05
step: 80, loss: 6.89948647050187e-05
step: 90, loss: 5.7539680710760877e-05
step: 100, loss: 3.932060644729063e-05
step: 110, loss: 7.899983756942675e-05
step: 120, loss: 0.03916870802640915
step: 130, loss: 6.333456985885277e-05
step: 140, loss: 0.012955511920154095
step: 150, loss: 0.00013220180699136108
step: 160, loss: 0.013721791096031666
step: 170, loss: 4.922922744299285e-05
step: 180, loss: 7.380324677797034e-05
step: 190, loss: 7.7407545177266e-05
step: 200, loss: 0.0014064114075154066
step: 210, loss: 0.026796484366059303
step: 220, loss: 7.27566730347462e-05
step: 230, loss: 8.00534980953671e-05
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
epoch 15: dev_f1=0.9921259842519685, f1=0.9865168539325843,
best_f1=0.9876265466816648
mv: cannot stat '*_dev.pt': No such file or directory
Selected optimization level 02: FP16 training with FP32 batchnorm and FP32
master weights.

Defaults for this optimization level are:
enabled                : True
opt_level              : 02
cast_model_type        : torch.float16
patch_torch_functions  : False
keep_batchnorm_fp32    : True
master_weights         : True
loss_scale             : dynamic
Processing user overrides (additional kwargs that are not None)...
After processing overrides, optimization options are:
enabled                : True
opt_level              : 02
cast_model_type        : torch.float16
patch_torch_functions  : False
keep_batchnorm_fp32    : True

```

```

master_weights      : True
loss_scale          : dynamic
Warning: multi_tensor_applier fused unscale kernel is unavailable, possibly
because apex was installed without --cuda_ext --cpp_ext. Using Python fallback.
Original ImportError was: ModuleNotFoundError("No module named 'amp_C'")
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
2473it [00:15, 155.25it/s]
load_f1 = 0.988814317673378
real_f1 = 0.9899441340782122
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
4398it [00:31, 140.20it/s]

```

### 2.1.3 DBLP-GoogleScholar - Running the matcher

```

[ ]: # Step 1:
# train the wdc model with the save_model flag on
# it will produce two files *_dev.pt and *_test.pt which are the
# best checkpoints on the dev set and the test set.
!CUDA_VISIBLE_DEVICES=0 python train_ditto.py \
--task Structured/DBLP-GoogleScholar \
--batch_size 32 \
--max_len 256 \
--lr 3e-5 \
--n_epochs 15 \
--finetuning \
--lm roberta \
--fp16 \
--save_model

# Step 2: pick a checkpoint, rename it to wdc_all_small.pt
!mv *_dev.pt checkpoints/DBLP-GoogleScholar.pt

# Step 3: run the matcher
!CUDA_VISIBLE_DEVICES=0 python matcher.py \
--task Structured/DBLP-GoogleScholar \
--input_path input/input_small.jsonl \
--output_path output/output_small.jsonl \
--lm roberta \

```

```
--use_gpu \  
--fp16 \  
--checkpoint_path checkpoints/
```

Selected optimization level 02: FP16 training with FP32 batchnorm and FP32 master weights.

Defaults for this optimization level are:

```
enabled          : True  
opt_level        : 02  
cast_model_type  : torch.float16  
patch_torch_functions : False  
keep_batchnorm_fp32 : True  
master_weights   : True  
loss_scale       : dynamic
```

Processing user overrides (additional kwargs that are not None)...

After processing overrides, optimization options are:

```
enabled          : True  
opt_level        : 02  
cast_model_type  : torch.float16  
patch_torch_functions : False  
keep_batchnorm_fp32 : True  
master_weights   : True  
loss_scale       : dynamic
```

Warning: multi\_tensor\_applier fused unscale kernel is unavailable, possibly because apex was installed without --cuda\_ext --cpp\_ext. Using Python fallback. Original ImportError was: ModuleNotFoundError("No module named 'amp\_C'")  
/usr/local/lib/python3.7/dist-packages/apex/amp/\_initialize.py:25: UserWarning: An input tensor was not cuda.

```
warnings.warn("An input tensor was not cuda.")
```

```
step: 0, loss: 0.6261507272720337
```

Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 32768.0

```
step: 10, loss: 0.4832388162612915
```

```
step: 20, loss: 0.28477954864501953
```

```
step: 30, loss: 0.3453848958015442
```

Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0

```
step: 40, loss: 0.3678410053253174
```

```
step: 50, loss: 0.5453017354011536
```

```
step: 60, loss: 0.2595318853855133
```

Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 8192.0

```
step: 70, loss: 0.2761323153972626
```

```
step: 80, loss: 0.24748417735099792
```

```
step: 90, loss: 0.2238815277814865
```

```
step: 100, loss: 0.040835846215486526
```

```
step: 110, loss: 0.09313009679317474
```

```
step: 120, loss: 0.12996555864810944
```

```
step: 130, loss: 0.1199643462896347
```

```

step: 140, loss: 0.32608306407928467
step: 150, loss: 0.13164567947387695
step: 160, loss: 0.1082160696387291
step: 170, loss: 0.03109813667833805
step: 180, loss: 0.08766434341669083
step: 190, loss: 0.075374536216259
step: 200, loss: 0.043275222182273865
step: 210, loss: 0.037510521709918976
step: 220, loss: 0.021877698600292206
step: 230, loss: 0.20856665074825287
step: 240, loss: 0.14690914750099182
step: 250, loss: 0.013146796263754368
step: 260, loss: 0.057932961732149124
step: 270, loss: 0.28157803416252136
step: 280, loss: 0.03242737054824829
step: 290, loss: 0.08025506883859634
step: 300, loss: 0.022276293486356735
step: 310, loss: 0.052786026149988174
step: 320, loss: 0.03820597007870674
step: 330, loss: 0.06262264400720596
step: 340, loss: 0.4210493862628937
step: 350, loss: 0.0477135069668293
step: 360, loss: 0.09189213067293167
step: 370, loss: 0.011759153567254543
step: 380, loss: 0.09175212681293488
step: 390, loss: 0.03661119192838669
step: 400, loss: 0.09365256130695343
step: 410, loss: 0.23014546930789948
step: 420, loss: 0.028256110846996307
step: 430, loss: 0.025045646354556084
step: 440, loss: 0.07877602428197861
step: 450, loss: 0.036947645246982574
step: 460, loss: 0.00954330526292324
step: 470, loss: 0.012602795846760273
step: 480, loss: 0.18472425639629364
step: 490, loss: 0.031469181180000305
step: 500, loss: 0.01373907458037138
step: 510, loss: 0.02576138637959957
step: 520, loss: 0.05511888489127159
step: 530, loss: 0.01056948583573103
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
  warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 1: dev_f1=0.9518072289156626, f1=0.9498607242339833,

```

```
best_f1=0.9498607242339833
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.09896969050168991
step: 10, loss: 0.02218763157725334
step: 20, loss: 0.024833237752318382
step: 30, loss: 0.08609700202941895
step: 40, loss: 0.03897666558623314
step: 50, loss: 0.02828032709658146
step: 60, loss: 0.01831570640206337
step: 70, loss: 0.02351672761142254
step: 80, loss: 0.005414848681539297
step: 90, loss: 0.001365917967632413
step: 100, loss: 0.10557740926742554
step: 110, loss: 0.011632286012172699
step: 120, loss: 0.017988009378314018
step: 130, loss: 0.003163954010233283
step: 140, loss: 0.048460256308317184
step: 150, loss: 0.017804473638534546
step: 160, loss: 0.0261080265045166
step: 170, loss: 0.007136539556086063
step: 180, loss: 0.007903761230409145
step: 190, loss: 0.016535062342882156
step: 200, loss: 0.10088392347097397
step: 210, loss: 0.028131192550063133
step: 220, loss: 0.0017143873265013099
step: 230, loss: 0.04166759178042412
step: 240, loss: 0.024883421137928963
step: 250, loss: 0.023029131814837456
step: 260, loss: 0.005623762030154467
step: 270, loss: 0.009269479662179947
step: 280, loss: 0.28183624148368835
step: 290, loss: 0.030838852748274803
step: 300, loss: 0.029810769483447075
step: 310, loss: 0.025622332468628883
step: 320, loss: 0.08167332410812378
step: 330, loss: 0.04586514085531235
step: 340, loss: 0.11796154081821442
step: 350, loss: 0.002843202091753483
step: 360, loss: 0.09461421519517899
step: 370, loss: 0.014198094606399536
step: 380, loss: 0.19894710183143616
step: 390, loss: 0.0029805232770740986
step: 400, loss: 0.04305163025856018
step: 410, loss: 0.011391455307602882
step: 420, loss: 0.05678892135620117
step: 430, loss: 0.11830360442399979
```

```

step: 440, loss: 0.01531317550688982
step: 450, loss: 0.014768188819289207
step: 460, loss: 0.0080588199198246
step: 470, loss: 0.010934348218142986
step: 480, loss: 0.006303567439317703
step: 490, loss: 0.060496941208839417
step: 500, loss: 0.0029679317958652973
step: 510, loss: 0.002409321255981922
step: 520, loss: 0.2565983831882477
step: 530, loss: 0.1297360211610794
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
  warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 2: dev_f1=0.9528214616096208, f1=0.9467345993515517,
best_f1=0.9467345993515517
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.1689760386943817
step: 10, loss: 0.03253495320677757
step: 20, loss: 0.006319291889667511
step: 30, loss: 0.055213119834661484
step: 40, loss: 0.02543659880757332
step: 50, loss: 0.005755640100687742
step: 60, loss: 0.007968193851411343
step: 70, loss: 0.0043789902701973915
step: 80, loss: 0.05354796722531319
step: 90, loss: 0.03607821464538574
step: 100, loss: 0.0641474574804306
step: 110, loss: 0.028763655573129654
step: 120, loss: 0.11236711591482162
step: 130, loss: 0.0040620616637170315
step: 140, loss: 0.00929277203977108
step: 150, loss: 0.02047841064631939
step: 160, loss: 0.021816862747073174
step: 170, loss: 0.0008507719030603766
step: 180, loss: 0.003233718452975154
step: 190, loss: 0.005613232031464577
step: 200, loss: 0.037458810955286026
step: 210, loss: 0.01277080550789833
step: 220, loss: 0.035006627440452576
step: 230, loss: 0.034162070602178574
step: 240, loss: 0.03433449938893318
step: 250, loss: 0.07639051228761673

```



```

step: 260, loss: 0.09849750995635986
step: 270, loss: 0.002721794182434678
step: 280, loss: 0.0006136539159342647
step: 290, loss: 0.006477898918092251
step: 300, loss: 0.0518612340092659
step: 310, loss: 0.02556593157351017
step: 320, loss: 0.06888636201620102
step: 330, loss: 0.001788213150575757
step: 340, loss: 0.013339156284928322
step: 350, loss: 0.13411091268062592
step: 360, loss: 0.013482272624969482
step: 370, loss: 0.015721462666988373
step: 380, loss: 0.010470548644661903
step: 390, loss: 0.006172674708068371
step: 400, loss: 0.05958117917180061
step: 410, loss: 0.029532751068472862
step: 420, loss: 0.023969730362296104
step: 430, loss: 0.022138496860861778
step: 440, loss: 0.26504990458488464
step: 450, loss: 0.01401457004249096
step: 460, loss: 0.09374116361141205
step: 470, loss: 0.005063348915427923
step: 480, loss: 0.13520267605781555
step: 490, loss: 0.024017732590436935
step: 500, loss: 0.05524349957704544
step: 510, loss: 0.03618059679865837
step: 520, loss: 0.0012637936742976308
step: 530, loss: 0.0009828669717535377
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
    warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 3: dev_f1=0.9587772116720705, f1=0.9583333333333334,
best_f1=0.9583333333333334
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.007118174806237221
step: 10, loss: 0.004212765023112297
step: 20, loss: 0.0020309241954237223
step: 30, loss: 0.12840181589126587
step: 40, loss: 0.055610910058021545
step: 50, loss: 0.07166479527950287
step: 60, loss: 0.03671345114707947
step: 70, loss: 0.03822116553783417

```

step: 80, loss: 0.01191917434334755  
step: 90, loss: 0.1607871651649475  
step: 100, loss: 0.0022081362549215555  
step: 110, loss: 0.016192354261875153  
step: 120, loss: 0.00019495320157147944  
step: 130, loss: 0.0016793526010587811  
step: 140, loss: 0.005265803541988134  
step: 150, loss: 0.001600744784809649  
step: 160, loss: 0.0045031229965388775  
step: 170, loss: 0.01794564723968506  
step: 180, loss: 0.06474735587835312  
step: 190, loss: 0.022864649072289467  
step: 200, loss: 0.040690407156944275  
step: 210, loss: 0.0026288542430847883  
step: 220, loss: 0.0034646480344235897  
step: 230, loss: 0.027504855766892433  
step: 240, loss: 0.05992363393306732  
step: 250, loss: 0.1748620867729187  
step: 260, loss: 0.0028346520848572254  
step: 270, loss: 0.019794650375843048  
step: 280, loss: 0.02639423869550228  
step: 290, loss: 0.04155852273106575  
step: 300, loss: 0.0031267590820789337  
step: 310, loss: 0.0015452175866812468  
step: 320, loss: 0.09511062502861023  
step: 330, loss: 0.009759322740137577  
step: 340, loss: 0.003692511236295104  
step: 350, loss: 0.016877425834536552  
step: 360, loss: 0.011521786451339722  
step: 370, loss: 0.0031037514563649893  
step: 380, loss: 0.013693446293473244  
step: 390, loss: 0.0014684053603559732  
step: 400, loss: 0.0013909811386838555  
step: 410, loss: 0.0003633899614214897  
step: 420, loss: 0.0018901372095569968  
step: 430, loss: 0.0011897962540388107  
step: 440, loss: 0.0023356156889349222  
step: 450, loss: 0.019448261708021164  
step: 460, loss: 0.11652522534132004  
step: 470, loss: 0.010724560357630253  
step: 480, loss: 0.00983720924705267  
step: 490, loss: 0.002022241475060582  
step: 500, loss: 0.05458232760429382  
step: 510, loss: 0.08570367097854614  
step: 520, loss: 0.003794898744672537  
step: 530, loss: 0.05107255280017853  
/usr/local/lib/python3.7/dist-packages/apex/amp/\_initialize.py:25: UserWarning:  
An input tensor was not cuda.

```

    warnings.warn("An input tensor was not cuda.")
epoch 4: dev_f1=0.9585253456221198, f1=0.9576816927322908,
best_f1=0.9583333333333334
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0017119760159403086
step: 10, loss: 0.02507363073527813
step: 20, loss: 0.004226445220410824
step: 30, loss: 0.0023241431917995214
step: 40, loss: 0.0003951990802306682
step: 50, loss: 0.11129932105541229
step: 60, loss: 0.006775146350264549
step: 70, loss: 0.010263471864163876
step: 80, loss: 0.00018021650612354279
step: 90, loss: 0.0034555247984826565
step: 100, loss: 0.11992792040109634
step: 110, loss: 0.009549487382173538
step: 120, loss: 0.18747037649154663
step: 130, loss: 0.01010195817798376
step: 140, loss: 0.001263139769434929
step: 150, loss: 0.02165260910987854
step: 160, loss: 0.007246533874422312
step: 170, loss: 0.03055974654853344
step: 180, loss: 0.0004489026323426515
step: 190, loss: 0.02732185460627079
step: 200, loss: 0.007901416160166264
step: 210, loss: 0.0007467004470527172
step: 220, loss: 0.0033380426466464996
step: 230, loss: 0.021869130432605743
step: 240, loss: 0.012098864652216434
step: 250, loss: 0.12182274460792542
step: 260, loss: 0.01219202857464552
step: 270, loss: 0.0021595479920506477
step: 280, loss: 0.000664988998323679
step: 290, loss: 0.0005505797453224659
step: 300, loss: 0.002575464779511094
step: 310, loss: 0.10452187806367874
step: 320, loss: 0.13334746658802032
step: 330, loss: 0.004289676900953054
step: 340, loss: 0.016988597810268402
step: 350, loss: 0.0008124572341330349
step: 360, loss: 0.002282365458086133
step: 370, loss: 0.0002037363883573562
step: 380, loss: 0.00021560906316153705
step: 390, loss: 0.08985967189073563
step: 400, loss: 0.005854750983417034
step: 410, loss: 0.06456971913576126

```

```

step: 420, loss: 0.14053522050380707
step: 430, loss: 0.0027675852179527283
step: 440, loss: 0.0007912172586657107
step: 450, loss: 0.012761357240378857
step: 460, loss: 0.0013819807209074497
step: 470, loss: 0.02873128093779087
step: 480, loss: 0.006299444939941168
step: 490, loss: 0.05625488981604576
step: 500, loss: 0.0043204729445278645
step: 510, loss: 0.003440834814682603
step: 520, loss: 0.12328189611434937
step: 530, loss: 0.05518217757344246
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 5: dev_f1=0.9574266792809839, f1=0.9508041627246925,
best_f1=0.9583333333333334
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.016772698611021042
step: 10, loss: 0.0018388086464256048
step: 20, loss: 0.001701298519037664
step: 30, loss: 0.0007863035425543785
step: 40, loss: 0.0002329180424567312
step: 50, loss: 0.0004716092371381819
step: 60, loss: 0.000376610696548596
step: 70, loss: 0.0006038041319698095
step: 80, loss: 0.0006107730441726744
step: 90, loss: 0.0008064981084316969
step: 100, loss: 0.0015291572781279683
step: 110, loss: 0.0017874842742457986
step: 120, loss: 0.005353391170501709
step: 130, loss: 0.0005470300675369799
step: 140, loss: 0.005627516191452742
step: 150, loss: 0.002141143660992384
step: 160, loss: 0.017795879393815994
step: 170, loss: 0.003027799306437373
step: 180, loss: 0.009483406320214272
step: 190, loss: 0.0318799652159214
step: 200, loss: 0.002156200585886836
step: 210, loss: 0.0005623915349133313
step: 220, loss: 0.005111464764922857
step: 230, loss: 0.003958373796194792
step: 240, loss: 0.00011591627117013559
step: 250, loss: 0.05543944984674454
step: 260, loss: 0.016284240409731865
step: 270, loss: 0.006564405746757984

```

```

step: 280, loss: 0.013057233765721321
step: 290, loss: 0.0028666716534644365
step: 300, loss: 0.009427901357412338
step: 310, loss: 0.01674296334385872
step: 320, loss: 0.0005397186614573002
step: 330, loss: 0.0042747375555336475
step: 340, loss: 0.0017502617556601763
step: 350, loss: 0.042802754789590836
step: 360, loss: 0.12925505638122559
step: 370, loss: 0.00926640722900629
step: 380, loss: 0.0013722273288294673
step: 390, loss: 0.010994416661560535
step: 400, loss: 0.05304832383990288
step: 410, loss: 0.000930170645006001
step: 420, loss: 0.05871826037764549
step: 430, loss: 0.0009979207534343004
step: 440, loss: 0.0004888674593530595
step: 450, loss: 0.2811400890350342
step: 460, loss: 0.003136364044621587
step: 470, loss: 0.003608326194807887
step: 480, loss: 0.003034242894500494
step: 490, loss: 0.005066507495939732
step: 500, loss: 0.0010231932392343879
step: 510, loss: 0.11002951115369797
step: 520, loss: 0.001532659400254488
step: 530, loss: 0.002623808803036809
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
epoch 6: dev_f1=0.9534883720930233, f1=0.9503745318352059,
best_f1=0.9583333333333334
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.01111783366650343
step: 10, loss: 0.0014535477384924889
step: 20, loss: 0.008188081905245781
step: 30, loss: 0.00569525919854641
step: 40, loss: 0.0029334628488868475
step: 50, loss: 0.0011924038408324122
step: 60, loss: 0.000902579864487052
step: 70, loss: 0.0003409300115890801
step: 80, loss: 0.00048190276720561087
step: 90, loss: 0.015470761805772781
step: 100, loss: 7.417058805003762e-05
step: 110, loss: 4.582144902087748e-05
step: 120, loss: 0.0016954264137893915
step: 130, loss: 4.465907113626599e-05

```

```

step: 140, loss: 0.0018751706229522824
step: 150, loss: 0.0007503863307647407
step: 160, loss: 6.145953375380486e-05
step: 170, loss: 0.00289184064604342
step: 180, loss: 0.009505007416009903
step: 190, loss: 0.006679447367787361
step: 200, loss: 0.0017666758503764868
step: 210, loss: 0.0014707864029332995
step: 220, loss: 0.005018474068492651
step: 230, loss: 4.077299308846705e-05
step: 240, loss: 0.01414349302649498
step: 250, loss: 0.0005745358066633344
step: 260, loss: 0.009162865579128265
step: 270, loss: 0.002660488011315465
step: 280, loss: 0.03082103282213211
step: 290, loss: 0.0005116563988849521
step: 300, loss: 0.0003857312840409577
step: 310, loss: 0.0017443703254684806
step: 320, loss: 0.0057974923402071
step: 330, loss: 0.0003187003021594137
step: 340, loss: 0.0018745746929198503
step: 350, loss: 0.0004974674084223807
step: 360, loss: 0.0004634768993128091
step: 370, loss: 0.006669232156127691
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 8192.0
step: 380, loss: 0.0029035797342658043
step: 390, loss: 0.002628875430673361
step: 400, loss: 0.003111779224127531
step: 410, loss: 0.0006939959130249918
step: 420, loss: 0.03988223895430565
step: 430, loss: 0.0001912718144012615
step: 440, loss: 0.009312516078352928
step: 450, loss: 0.00426238588988781
step: 460, loss: 0.013042639009654522
step: 470, loss: 0.16183853149414062
step: 480, loss: 0.003532549599185586
step: 490, loss: 0.09352166205644608
step: 500, loss: 0.0033931348007172346
step: 510, loss: 0.0021147739607840776
step: 520, loss: 0.0006494582630693913
step: 530, loss: 0.0007042685174383223
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
epoch 7: dev_f1=0.9570494864612512, f1=0.9538461538461538,
best_f1=0.9583333333333334
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.

```

```
warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.000349023612216115
step: 10, loss: 0.0012588623212650418
step: 20, loss: 0.0008901949622668326
step: 30, loss: 0.0004864431684836745
step: 40, loss: 0.0002355135657126084
step: 50, loss: 0.00014522313722409308
step: 60, loss: 0.00024278719502035528
step: 70, loss: 0.05390032008290291
step: 80, loss: 0.0018259760690853
step: 90, loss: 0.00017717824084684253
step: 100, loss: 0.0004056718316860497
step: 110, loss: 0.00027615739963948727
step: 120, loss: 0.0004917666083201766
step: 130, loss: 0.0019039430189877748
step: 140, loss: 5.4844891565153375e-05
step: 150, loss: 0.008071832358837128
step: 160, loss: 6.328485324047506e-05
step: 170, loss: 0.11987938731908798
step: 180, loss: 0.0005494561046361923
step: 190, loss: 0.012334225699305534
step: 200, loss: 0.0008886667201295495
step: 210, loss: 0.03921395540237427
step: 220, loss: 7.14315174263902e-05
step: 230, loss: 0.1676296591758728
step: 240, loss: 0.0024012047797441483
step: 250, loss: 0.0044902171939611435
step: 260, loss: 0.001030148472636938
step: 270, loss: 0.04712044820189476
step: 280, loss: 0.0010438986355438828
step: 290, loss: 0.0008507276652380824
step: 300, loss: 0.0004146149440202862
step: 310, loss: 0.002134164795279503
step: 320, loss: 0.0690557211637497
step: 330, loss: 0.000899866980034858
step: 340, loss: 0.027507733553647995
step: 350, loss: 0.0006149989203549922
step: 360, loss: 0.0007790903327986598
step: 370, loss: 0.2043483704328537
step: 380, loss: 0.00031514529837295413
step: 390, loss: 0.005949445068836212
step: 400, loss: 0.0008542902069166303
step: 410, loss: 0.0013664050493389368
step: 420, loss: 0.0005612573586404324
step: 430, loss: 0.005764591507613659
step: 440, loss: 0.0025254746433347464
step: 450, loss: 0.0006498463917523623
step: 460, loss: 0.0037556132301688194
```

```

step: 470, loss: 0.08362945169210434
step: 480, loss: 0.03350213170051575
step: 490, loss: 0.0019693549256771803
step: 500, loss: 0.000586183276027441
step: 510, loss: 0.07406290620565414
step: 520, loss: 0.0005529209738597274
step: 530, loss: 7.277474651345983e-05
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 8: dev_f1=0.9570041608876559, f1=0.955637707948244,
best_f1=0.9583333333333334
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 3.5511791793396696e-05
step: 10, loss: 0.0020101163536310196
step: 20, loss: 0.0011578527046367526
step: 30, loss: 0.11392512917518616
step: 40, loss: 0.0009838104015216231
step: 50, loss: 0.0003872126981150359
step: 60, loss: 0.0013905104715377092
step: 70, loss: 0.004286007955670357
step: 80, loss: 0.005597591400146484
step: 90, loss: 0.08584943413734436
step: 100, loss: 0.0037328905891627073
step: 110, loss: 0.005685309413820505
step: 120, loss: 0.006937746424227953
step: 130, loss: 0.00029120242106728256
step: 140, loss: 0.0004968232242390513
step: 150, loss: 0.0016323860036209226
step: 160, loss: 0.004887556191533804
step: 170, loss: 0.004247824661433697
step: 180, loss: 0.00010292748629581183
step: 190, loss: 2.4377208319492638e-05
step: 200, loss: 2.699462311284151e-05
step: 210, loss: 0.00025907153030857444
step: 220, loss: 0.003770157927647233
step: 230, loss: 0.0031493392307311296
step: 240, loss: 0.000934614974539727
step: 250, loss: 0.003951606806367636
step: 260, loss: 0.020551396533846855
step: 270, loss: 0.00476103276014328
step: 280, loss: 0.009690470062196255
step: 290, loss: 0.0037014209665358067
step: 300, loss: 0.0013685603626072407
step: 310, loss: 0.10709870606660843
step: 320, loss: 0.0011162100126966834

```



```

step: 330, loss: 0.0007006803061813116
step: 340, loss: 0.018330633640289307
step: 350, loss: 0.06612804532051086
step: 360, loss: 9.125468204729259e-05
step: 370, loss: 0.0004941271035932004
step: 380, loss: 0.0008950269548222423
step: 390, loss: 0.04928738623857498
step: 400, loss: 0.0898575410246849
step: 410, loss: 0.07749626785516739
step: 420, loss: 0.00048246042570099235
step: 430, loss: 0.00205324892885983
step: 440, loss: 8.943633292801678e-05
step: 450, loss: 0.019686656072735786
step: 460, loss: 8.934510697145015e-05
step: 470, loss: 0.00018307712161913514
step: 480, loss: 4.364674896351062e-05
step: 490, loss: 4.6528002712875605e-05
step: 500, loss: 3.330618710606359e-05
step: 510, loss: 0.003232477232813835
step: 520, loss: 0.0018517260905355215
step: 530, loss: 0.0029341564513742924
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 9: dev_f1=0.9579288025889967, f1=0.9542302357836338,
best_f1=0.9583333333333334
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.02969362400472164
step: 10, loss: 0.0010913523146882653
step: 20, loss: 0.000327634799759835
step: 30, loss: 0.001717464765533805
step: 40, loss: 0.00014867795107420534
step: 50, loss: 0.0003241621016059071
step: 60, loss: 0.002116971183568239
step: 70, loss: 4.986844214727171e-05
step: 80, loss: 6.691004091408104e-05
step: 90, loss: 4.342686224845238e-05
step: 100, loss: 0.0005899665993638337
step: 110, loss: 0.003298116847872734
step: 120, loss: 0.004940463230013847
step: 130, loss: 0.010750704444944859
step: 140, loss: 3.627130718086846e-05
step: 150, loss: 0.0018965811468660831
step: 160, loss: 0.047878023236989975
step: 170, loss: 0.00020156281243544072
step: 180, loss: 0.0011705218348652124

```

```

step: 190, loss: 0.00012181799684185535
step: 200, loss: 0.0012687875423580408
step: 210, loss: 0.01795872300863266
step: 220, loss: 0.027586666867136955
step: 230, loss: 0.00325515097938478
step: 240, loss: 0.0002614277764223516
step: 250, loss: 0.00019211390463169664
step: 260, loss: 0.0002143003512173891
step: 270, loss: 0.00018585917132440954
step: 280, loss: 0.029612209647893906
step: 290, loss: 0.0066001079976558685
step: 300, loss: 0.007338155526667833
step: 310, loss: 0.04610951989889145
step: 320, loss: 0.0004808977246284485
step: 330, loss: 0.00025064306100830436
step: 340, loss: 3.589984407881275e-05
step: 350, loss: 0.00015311772585846484
step: 360, loss: 6.03376938670408e-05
step: 370, loss: 0.00019763552700169384
step: 380, loss: 0.0022881038021296263
step: 390, loss: 0.00016465944645460695
step: 400, loss: 0.0002946701133623719
step: 410, loss: 0.0024052190128713846
step: 420, loss: 0.00016382080502808094
step: 430, loss: 0.0037420084699988365
step: 440, loss: 0.00030626595253124833
step: 450, loss: 0.0070612444542348385
step: 460, loss: 0.0001105665578506887
step: 470, loss: 0.0013277101097628474
step: 480, loss: 7.549075235147029e-05
step: 490, loss: 0.0030148851219564676
step: 500, loss: 0.003565294900909066
step: 510, loss: 5.76675301999785e-05
step: 520, loss: 0.009381919167935848
step: 530, loss: 0.0018062974559143186
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
  warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 10: dev_f1=0.9620608899297424, f1=0.9585468095016302,
best_f1=0.9585468095016302
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00010630970791680738

```

step: 10, loss: 0.001056175446137786  
step: 20, loss: 0.00011784930393332615  
step: 30, loss: 0.00012221421638969332  
step: 40, loss: 0.00195872038602829  
step: 50, loss: 4.792453546542674e-05  
step: 60, loss: 8.449466258753091e-05  
step: 70, loss: 7.4321091233287e-05  
step: 80, loss: 0.0008559729321859777  
step: 90, loss: 0.0016228987369686365  
step: 100, loss: 0.0002804899704642594  
step: 110, loss: 5.761873399023898e-05  
step: 120, loss: 0.00014361334615387022  
step: 130, loss: 2.5714422008604743e-05  
step: 140, loss: 0.00014488209853880107  
step: 150, loss: 0.00866605807095766  
step: 160, loss: 0.0001337838330073282  
step: 170, loss: 4.467836697585881e-05  
step: 180, loss: 8.532328502042219e-05  
step: 190, loss: 0.0009567135712131858  
step: 200, loss: 0.0002548767370171845  
step: 210, loss: 0.00010085547546623275  
step: 220, loss: 0.042180150747299194  
step: 230, loss: 1.5052427443151828e-05  
step: 240, loss: 0.0019380629528313875  
step: 250, loss: 0.00013777069398202002  
step: 260, loss: 0.0026727207005023956  
step: 270, loss: 0.001674878061749041  
step: 280, loss: 0.00011118927068309858  
step: 290, loss: 4.066325709572993e-05  
step: 300, loss: 3.1472958653466776e-05  
step: 310, loss: 3.6773566534975544e-05  
step: 320, loss: 0.0002623393083922565  
step: 330, loss: 2.2335854737320915e-05  
step: 340, loss: 0.00032744373311288655  
step: 350, loss: 3.495700366329402e-05  
step: 360, loss: 0.00019272971258033067  
step: 370, loss: 0.0021370570175349712  
step: 380, loss: 0.003937097731977701  
step: 390, loss: 0.0011491063050925732  
step: 400, loss: 9.257868805434555e-05  
step: 410, loss: 0.0007850569090805948  
step: 420, loss: 0.0002894085773732513  
step: 430, loss: 0.0009200386120937765  
step: 440, loss: 3.4892986150225624e-05  
step: 450, loss: 0.0007756390259601176  
step: 460, loss: 0.0005656020366586745  
step: 470, loss: 0.00048747393884696066  
step: 480, loss: 0.0012503924081102014

```
step: 490, loss: 0.011318088509142399
step: 500, loss: 0.04887128993868828
step: 510, loss: 0.00012383033754304051
step: 520, loss: 7.476071186829358e-05
step: 530, loss: 0.0029225226026028395
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
epoch 11: dev_f1=0.9563152896486229, f1=0.9520225776105363,
best_f1=0.9585468095016302
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 3.9212409319588915e-05
step: 10, loss: 7.442606147378683e-05
step: 20, loss: 0.02942023053765297
step: 30, loss: 0.0009456989355385303
step: 40, loss: 0.00097364786779508
step: 50, loss: 9.548027446726337e-05
step: 60, loss: 4.293314123060554e-05
step: 70, loss: 0.00014857211499474943
step: 80, loss: 0.00017401455261278898
step: 90, loss: 0.001112910802476108
step: 100, loss: 0.2106889933347702
step: 110, loss: 0.00240869028493762
step: 120, loss: 0.0006014917162247002
step: 130, loss: 0.001863558660261333
step: 140, loss: 0.0012364281574264169
step: 150, loss: 0.0028908746317029
step: 160, loss: 0.0007267193286679685
step: 170, loss: 0.00012983284250367433
step: 180, loss: 4.713851376436651e-05
step: 190, loss: 7.546322740381584e-05
step: 200, loss: 0.0009213596349582076
step: 210, loss: 0.0002861736575141549
step: 220, loss: 0.006174338515847921
step: 230, loss: 7.996701606316492e-05
step: 240, loss: 0.0017325361259281635
step: 250, loss: 2.4049335479503497e-05
step: 260, loss: 1.4755478332517669e-05
step: 270, loss: 0.0006175221060402691
step: 280, loss: 3.563423524610698e-05
step: 290, loss: 9.944174234988168e-05
step: 300, loss: 0.0009384374716319144
step: 310, loss: 0.0005424321279861033
step: 320, loss: 8.93840697244741e-05
step: 330, loss: 0.004430897068232298
step: 340, loss: 5.054696521256119e-05
```

```

step: 350, loss: 0.009826885536313057
step: 360, loss: 7.050814747344702e-05
step: 370, loss: 0.015232102945446968
step: 380, loss: 0.0013459392357617617
step: 390, loss: 0.0014701135223731399
step: 400, loss: 7.73992360336706e-05
step: 410, loss: 8.149215864250436e-05
step: 420, loss: 0.0012619573390111327
step: 430, loss: 7.475606980733573e-05
step: 440, loss: 0.002372003858909011
step: 450, loss: 0.022451123222708702
step: 460, loss: 0.0007114650798030198
step: 470, loss: 0.008758545853197575
step: 480, loss: 0.005176210775971413
step: 490, loss: 3.3338226785417646e-05
step: 500, loss: 4.13515699619893e-05
step: 510, loss: 0.00343899754807353
step: 520, loss: 0.0013268142938613892
step: 530, loss: 4.8585647164145485e-05
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
epoch 12: dev_f1=0.9597754911131898, f1=0.9570093457943926,
best_f1=0.9585468095016302
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 1.4695888239657506e-05
step: 10, loss: 0.00018298652139492333
step: 20, loss: 2.955880518129561e-05
step: 30, loss: 9.552596020512283e-05
step: 40, loss: 0.003775854827836156
step: 50, loss: 0.0036116288974881172
step: 60, loss: 1.222610899276333e-05
step: 70, loss: 3.644269600044936e-05
step: 80, loss: 0.00017865766130853444
step: 90, loss: 3.6906836612615734e-05
step: 100, loss: 2.4640929041197523e-05
step: 110, loss: 0.0003021266602445394
step: 120, loss: 1.916969267767854e-05
step: 130, loss: 1.5842888387851417e-05
step: 140, loss: 8.074467041296884e-05
step: 150, loss: 0.002041837200522423
step: 160, loss: 0.0010296466061845422
step: 170, loss: 0.0013982906239107251
step: 180, loss: 3.547434971551411e-05
step: 190, loss: 0.002325656358152628
step: 200, loss: 0.00017452091560699046

```

```

step: 210, loss: 4.679974517785013e-05
step: 220, loss: 0.000473153282655403
step: 230, loss: 0.0006290411110967398
step: 240, loss: 0.0005652750260196626
step: 250, loss: 0.00035667934571392834
step: 260, loss: 0.00010631029726937413
step: 270, loss: 8.464162237942219e-05
step: 280, loss: 3.485543129500002e-05
step: 290, loss: 8.068874194577802e-06
step: 300, loss: 3.9084930904209614e-05
step: 310, loss: 2.7498335839482024e-05
step: 320, loss: 1.5902707673376426e-05
step: 330, loss: 1.5303236068575643e-05
step: 340, loss: 0.000374706054572016
step: 350, loss: 9.49185141507769e-06
step: 360, loss: 0.089700847864151
step: 370, loss: 0.00029959637322463095
step: 380, loss: 8.313115540659055e-05
step: 390, loss: 0.0002513393119443208
step: 400, loss: 2.9889459256082773e-05
step: 410, loss: 0.00012906150368507951
step: 420, loss: 1.5954832633724436e-05
step: 430, loss: 0.00010597238724585623
step: 440, loss: 9.886731277219951e-06
step: 450, loss: 6.194377783685923e-05
step: 460, loss: 0.002514162566512823
step: 470, loss: 0.0010867916280403733
step: 480, loss: 7.074233508319594e-06
step: 490, loss: 2.3520020477008075e-05
step: 500, loss: 9.108142876357306e-06
step: 510, loss: 5.763875742559321e-05
step: 520, loss: 0.0002499518741387874
step: 530, loss: 1.7023718100972474e-05
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 13: dev_f1=0.9614299153339605, f1=0.9535864978902954,
best_f1=0.9585468095016302
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0001967531570699066
step: 10, loss: 0.0011428197612985969
step: 20, loss: 1.270283610210754e-05
step: 30, loss: 0.0029313673730939627
step: 40, loss: 0.00032762312912382185
step: 50, loss: 0.00021944423497188836
step: 60, loss: 1.2356575098237954e-05

```

step: 70, loss: 7.01135941199027e-05  
step: 80, loss: 3.646684854174964e-05  
step: 90, loss: 1.095585867005866e-05  
step: 100, loss: 1.7124439182225615e-05  
step: 110, loss: 7.540828664787114e-05  
step: 120, loss: 1.2989860806555953e-05  
step: 130, loss: 5.448700540000573e-05  
step: 140, loss: 0.00014497715164907277  
step: 150, loss: 0.0013168994337320328  
step: 160, loss: 1.3209630196797661e-05  
step: 170, loss: 0.000448630191385746  
step: 180, loss: 0.00021035433746874332  
step: 190, loss: 9.212490112986416e-06  
step: 200, loss: 4.229098340147175e-05  
step: 210, loss: 2.523216971894726e-05  
step: 220, loss: 1.1861026905535255e-05  
step: 230, loss: 1.0199697499047033e-05  
step: 240, loss: 0.000525592768099159  
step: 250, loss: 0.0006291005993261933  
step: 260, loss: 0.0001567913859616965  
step: 270, loss: 0.0008184657781384885  
step: 280, loss: 8.96662641025614e-06  
step: 290, loss: 9.685516488389112e-06  
step: 300, loss: 3.798305260716006e-05  
step: 310, loss: 7.707519216637593e-06  
step: 320, loss: 1.551172317704186e-05  
step: 330, loss: 0.0008190104854293168  
step: 340, loss: 0.0008624345064163208  
step: 350, loss: 1.0415745236969087e-05  
step: 360, loss: 3.513701813062653e-05  
step: 370, loss: 0.0023181417491286993  
step: 380, loss: 0.0012700591469183564  
step: 390, loss: 0.0004379998426884413  
step: 400, loss: 0.0027772276662290096  
step: 410, loss: 1.441954736947082e-05  
step: 420, loss: 3.774671131395735e-05  
step: 430, loss: 1.4952920537325554e-05  
step: 440, loss: 1.7813588783610612e-05  
step: 450, loss: 1.035614786815131e-05  
step: 460, loss: 0.0175832100212574  
step: 470, loss: 1.1507260751386639e-05  
step: 480, loss: 0.00044728166540153325  
step: 490, loss: 1.0751043191703502e-05  
step: 500, loss: 0.0007765316404402256  
step: 510, loss: 0.0071149179711937904  
step: 520, loss: 0.0009189354022964835  
step: 530, loss: 0.001534131821244955  
/usr/local/lib/python3.7/dist-packages/apex/amp/\_initialize.py:25: UserWarning:

```

An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
epoch 14: dev_f1=0.961646398503274, f1=0.957089552238806,
best_f1=0.9585468095016302
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 9.078417861019261e-06
step: 10, loss: 0.002086926717311144
step: 20, loss: 0.001133267069235444
step: 30, loss: 0.002745558274909854
step: 40, loss: 3.1572730222251266e-05
step: 50, loss: 0.0009566740482114255
step: 60, loss: 1.4893372281221673e-05
step: 70, loss: 2.031857002293691e-05
step: 80, loss: 9.4434817583533e-06
step: 90, loss: 1.8029246348305605e-05
step: 100, loss: 4.7348171392513905e-06
step: 110, loss: 0.0008642185130156577
step: 120, loss: 8.214165063691325e-06
step: 130, loss: 2.1086927517899312e-05
step: 140, loss: 1.0881327398237772e-05
step: 150, loss: 1.0728695087891538e-05
step: 160, loss: 9.107380174100399e-05
step: 170, loss: 7.975689186423551e-06
step: 180, loss: 1.2617372703971341e-05
step: 190, loss: 0.0009087413200177252
step: 200, loss: 0.0004228758334647864
step: 210, loss: 5.900222095078789e-05
step: 220, loss: 1.0561038834566716e-05
step: 230, loss: 1.351885475742165e-05
step: 240, loss: 8.657480066176504e-06
step: 250, loss: 2.920942279160954e-05
step: 260, loss: 8.195544978661928e-06
step: 270, loss: 1.37049928525812e-05
step: 280, loss: 1.139542473538313e-05
step: 290, loss: 0.00011836644262075424
step: 300, loss: 0.00018793037452269346
step: 310, loss: 0.04115761071443558
step: 320, loss: 1.1462553629826289e-05
step: 330, loss: 1.2635998245968949e-05
step: 340, loss: 1.1525892659847159e-05
step: 350, loss: 0.0036157805006951094
step: 360, loss: 3.1120620405999944e-05
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 370, loss: 1.2989856259082444e-05
step: 380, loss: 9.227426744473632e-06
step: 390, loss: 1.2784884347638581e-05

```



```

step: 400, loss: 0.012805861420929432
step: 410, loss: 3.3060754503821954e-05
step: 420, loss: 8.35945684229955e-06
step: 430, loss: 5.308504569256911e-06
step: 440, loss: 7.76713750383351e-06
step: 450, loss: 0.002318869112059474
step: 460, loss: 0.0011575366370379925
step: 470, loss: 8.33711055747699e-06
step: 480, loss: 0.002209307625889778
step: 490, loss: 1.4759070836589672e-05
step: 500, loss: 0.0008755812887102365
step: 510, loss: 4.006074959761463e-05
step: 520, loss: 0.00025085380184464157
step: 530, loss: 8.549410267733037e-06
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 15: dev_f1=0.961989676208353, f1=0.9565217391304348,
best_f1=0.9585468095016302
mv: cannot stat '*_dev.pt': No such file or directory
Selected optimization level 02: FP16 training with FP32 batchnorm and FP32
master weights.

Defaults for this optimization level are:
enabled                : True
opt_level              : 02
cast_model_type        : torch.float16
patch_torch_functions  : False
keep_batchnorm_fp32    : True
master_weights         : True
loss_scale             : dynamic
Processing user overrides (additional kwargs that are not None)...
After processing overrides, optimization options are:
enabled                : True
opt_level              : 02
cast_model_type        : torch.float16
patch_torch_functions  : False
keep_batchnorm_fp32    : True
master_weights         : True
loss_scale             : dynamic
Warning: multi_tensor_applier fused unscale kernel is unavailable, possibly
because apex was installed without --cuda_ext --cpp_ext. Using Python fallback.
Original ImportError was: ModuleNotFoundError("No module named 'amp_C'")
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
Oit [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.

```

```

    warnings.warn("An input tensor was not cuda.")
5742it [00:32, 175.19it/s]
load_f1 = 0.9622377622377621
real_f1 = 0.9622377622377621
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
4398it [00:31, 140.35it/s]

```

## 2.1.4 Amazon-Google - Running the matcher

```

[ ]: # Step 1:
# train the wdc model with the save_model flag on
# it will produce two files *_dev.pt and *_test.pt which are the
# best checkpoints on the dev set and the test set.
!CUDA_VISIBLE_DEVICES=0 python train_ditto.py \
    --task Structured/Amazon-Google \
    --batch_size 32 \
    --max_len 256 \
    --lr 3e-5 \
    --n_epochs 15 \
    --finetuning \
    --lm roberta \
    --fp16 \
    --save_model

# Step 2: pick a checkpoint, rename it to wdc_all_small.pt
!mv *_dev.pt checkpoints/Amazon-Google.pt

# Step 3: run the matcher
!CUDA_VISIBLE_DEVICES=0 python matcher.py \
    --task Structured/Amazon-Google \
    --input_path input/input_small.jsonl \
    --output_path output/output_small.jsonl \
    --lm roberta \
    --use_gpu \
    --fp16 \
    --checkpoint_path checkpoints/

```

Selected optimization level 02: FP16 training with FP32 batchnorm and FP32 master weights.

Defaults for this optimization level are:

```

enabled          : True
opt_level        : 02
cast_model_type  : torch.float16

```

```

patch_torch_functions : False
keep_batchnorm_fp32   : True
master_weights        : True
loss_scale             : dynamic
Processing user overrides (additional kwargs that are not None)...
After processing overrides, optimization options are:
enabled               : True
opt_level             : 02
cast_model_type       : torch.float16
patch_torch_functions : False
keep_batchnorm_fp32   : True
master_weights        : True
loss_scale            : dynamic
Warning: multi_tensor_applier fused unscale kernel is unavailable, possibly
because apex was installed without --cuda_ext --cpp_ext. Using Python fallback.
Original ImportError was: ModuleNotFoundError("No module named 'amp_C'")
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 32768.0
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:136:
UserWarning: Detected call of `lr_scheduler.step()` before `optimizer.step()`.
In PyTorch 1.1.0 and later, you should call them in the opposite order:
`optimizer.step()` before `lr_scheduler.step()`. Failure to do this will result
in PyTorch skipping the first value of the learning rate schedule. See more
details at https://pytorch.org/docs/stable/optim.html#how-to-adjust-learning-
rate
    "https://pytorch.org/docs/stable/optim.html#how-to-adjust-learning-rate",
UserWarning)
step: 0, loss: 0.5243884325027466
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 10, loss: 0.40292757749557495
step: 20, loss: 0.45489832758903503
step: 30, loss: 0.32912489771842957
step: 40, loss: 0.3305967450141907
step: 50, loss: 0.41397231817245483
step: 60, loss: 0.563334584236145
step: 70, loss: 0.35427913069725037
step: 80, loss: 0.3320746123790741
step: 90, loss: 0.26558637619018555
step: 100, loss: 0.24686889350414276
step: 110, loss: 0.2400158941745758
step: 120, loss: 0.3557991683483124
step: 130, loss: 0.24991892278194427
step: 140, loss: 0.432404100894928
step: 150, loss: 0.3535897433757782
step: 160, loss: 0.4554344117641449
step: 170, loss: 0.22042366862297058

```

```

step: 180, loss: 0.2846716046333313
step: 190, loss: 0.46392813324928284
step: 200, loss: 0.1774752140045166
step: 210, loss: 0.30242452025413513
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
    warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 1: dev_f1=0.6137566137566137, f1=0.6177024482109228,
best_f1=0.6177024482109228
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.28983110189437866
step: 10, loss: 0.15553100407123566
step: 20, loss: 0.3878655731678009
step: 30, loss: 0.37307456135749817
step: 40, loss: 0.4655243456363678
step: 50, loss: 0.1848212331533432
step: 60, loss: 0.3679233491420746
step: 70, loss: 0.1731424480676651
step: 80, loss: 0.18657654523849487
step: 90, loss: 0.17646023631095886
step: 100, loss: 0.3038082420825958
step: 110, loss: 0.2419905662536621
step: 120, loss: 0.07651379704475403
step: 130, loss: 0.0818910226225853
step: 140, loss: 0.1680273860692978
step: 150, loss: 0.27829787135124207
step: 160, loss: 0.05323468893766403
step: 170, loss: 0.3951944410800934
step: 180, loss: 0.14127019047737122
step: 190, loss: 0.22444120049476624
step: 200, loss: 0.11896607279777527
step: 210, loss: 0.21001720428466797
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
    warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 2: dev_f1=0.7028862478777589, f1=0.677304964539007,
best_f1=0.677304964539007
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:

```

```

An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.03839067369699478
step: 10, loss: 0.03881911188364029
step: 20, loss: 0.1311763972043991
step: 30, loss: 0.05334322899580002
step: 40, loss: 0.2804754972457886
step: 50, loss: 0.12832516431808472
step: 60, loss: 0.2006114423274994
step: 70, loss: 0.1155007854104042
step: 80, loss: 0.15381814539432526
step: 90, loss: 0.048220664262771606
step: 100, loss: 0.193673774600029
step: 110, loss: 0.06680047512054443
step: 120, loss: 0.1660921722650528
step: 130, loss: 0.14301562309265137
step: 140, loss: 0.06822014600038528
step: 150, loss: 0.16251640021800995
step: 160, loss: 0.09120673686265945
step: 170, loss: 0.2920011878013611
step: 180, loss: 0.10128739476203918
step: 190, loss: 0.02746955305337906
step: 200, loss: 0.09582357108592987
step: 210, loss: 0.12953418493270874
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
  warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 3: dev_f1=0.7245841035120147, f1=0.7121212121212123,
best_f1=0.7121212121212123
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.028844503685832024
step: 10, loss: 0.1518724113702774
step: 20, loss: 0.07890728861093521
step: 30, loss: 0.06031574308872223
step: 40, loss: 0.03397657349705696
step: 50, loss: 0.10660656541585922
step: 60, loss: 0.163803368806839
step: 70, loss: 0.07099054753780365
step: 80, loss: 0.04691605642437935
step: 90, loss: 0.05678679421544075
step: 100, loss: 0.16820596158504486
step: 110, loss: 0.34135177731513977

```

```

step: 120, loss: 0.1394708752632141
step: 130, loss: 0.23323103785514832
step: 140, loss: 0.13578803837299347
step: 150, loss: 0.042862746864557266
step: 160, loss: 0.14409895241260529
step: 170, loss: 0.05276333913207054
step: 180, loss: 0.025682374835014343
step: 190, loss: 0.09417524933815002
step: 200, loss: 0.0382833406329155
step: 210, loss: 0.2244662046432495
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
    warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 4: dev_f1=0.7319778188539741, f1=0.7669172932330827,
best_f1=0.7669172932330827
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.12612725794315338
step: 10, loss: 0.009873013943433762
step: 20, loss: 0.07317094504833221
step: 30, loss: 0.026698613539338112
step: 40, loss: 0.05306970700621605
step: 50, loss: 0.07079976052045822
step: 60, loss: 0.03632355108857155
step: 70, loss: 0.18207651376724243
step: 80, loss: 0.07307180762290955
step: 90, loss: 0.049605678766965866
step: 100, loss: 0.012928936630487442
step: 110, loss: 0.09690920263528824
step: 120, loss: 0.11220980435609818
step: 130, loss: 0.02819094806909561
step: 140, loss: 0.06758684664964676
step: 150, loss: 0.0515696257352829
step: 160, loss: 0.02291046269237995
step: 170, loss: 0.029317351058125496
step: 180, loss: 0.09642213582992554
step: 190, loss: 0.1335943192243576
step: 200, loss: 0.06553708016872406
step: 210, loss: 0.08544692397117615
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:

```

UserWarning: Please also save or load the state of the optimizer when saving or loading the scheduler.

```
warnings.warn(SAVE_STATE_WARNING, UserWarning)
```

epoch 5: dev\_f1=0.7411347517730497, f1=0.7396768402154399,

best\_f1=0.7396768402154399

/usr/local/lib/python3.7/dist-packages/apex/amp/\_initialize.py:25: UserWarning: An input tensor was not cuda.

```
warnings.warn("An input tensor was not cuda.")
```

step: 0, loss: 0.023940548300743103

step: 10, loss: 0.030120832845568657

step: 20, loss: 0.01811383105814457

step: 30, loss: 0.035458456724882126

step: 40, loss: 0.033249519765377045

step: 50, loss: 0.04121655225753784

step: 60, loss: 0.09490775316953659

step: 70, loss: 0.05148450285196304

step: 80, loss: 0.09592711925506592

step: 90, loss: 0.05965791270136833

step: 100, loss: 0.13238954544067383

step: 110, loss: 0.0921064242720604

step: 120, loss: 0.006813114043325186

step: 130, loss: 0.009552429430186749

step: 140, loss: 0.08102936297655106

step: 150, loss: 0.052426159381866455

step: 160, loss: 0.05680863559246063

step: 170, loss: 0.027854852378368378

step: 180, loss: 0.04705891013145447

step: 190, loss: 0.03297683969140053

step: 200, loss: 0.0539361797273159

step: 210, loss: 0.048446714878082275

/usr/local/lib/python3.7/dist-packages/apex/amp/\_initialize.py:25: UserWarning: An input tensor was not cuda.

```
warnings.warn("An input tensor was not cuda.")
```

/usr/local/lib/python3.7/dist-packages/torch/optim/lr\_scheduler.py:216:

UserWarning: Please also save or load the state of the optimizer when saving or loading the scheduler.

```
warnings.warn(SAVE_STATE_WARNING, UserWarning)
```

epoch 6: dev\_f1=0.7657992565055762, f1=0.7609942638623327,

best\_f1=0.7609942638623327

/usr/local/lib/python3.7/dist-packages/apex/amp/\_initialize.py:25: UserWarning: An input tensor was not cuda.

```
warnings.warn("An input tensor was not cuda.")
```

step: 0, loss: 0.0765802413225174

step: 10, loss: 0.0009854190284386277

step: 20, loss: 0.0029265237972140312

step: 30, loss: 0.2947807013988495

step: 40, loss: 0.06937815248966217

step: 50, loss: 0.030564798042178154

```

step: 60, loss: 0.09624069184064865
step: 70, loss: 0.07769086211919785
step: 80, loss: 0.04156409204006195
step: 90, loss: 0.04758178070187569
step: 100, loss: 0.026235688477754593
step: 110, loss: 0.12011558562517166
step: 120, loss: 0.04919973015785217
step: 130, loss: 0.05483458191156387
step: 140, loss: 0.026817750185728073
step: 150, loss: 0.01700894720852375
step: 160, loss: 0.12092697620391846
step: 170, loss: 0.037091925740242004
step: 180, loss: 0.006425225175917149
step: 190, loss: 0.0383683517575264
step: 200, loss: 0.027000272646546364
step: 210, loss: 0.0904494971036911
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
epoch 7: dev_f1=0.7408829174664108, f1=0.7555555555555556,
best_f1=0.7609942638623327
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.10188741981983185
step: 10, loss: 0.0038962634280323982
step: 20, loss: 0.0945989191532135
step: 30, loss: 0.0712553933262825
step: 40, loss: 0.0592183843255043
step: 50, loss: 0.010885336436331272
step: 60, loss: 0.03193378075957298
step: 70, loss: 0.03949188441038132
step: 80, loss: 0.06498377025127411
step: 90, loss: 0.05132715776562691
step: 100, loss: 0.19724030792713165
step: 110, loss: 0.015028207562863827
step: 120, loss: 0.07914671301841736
step: 130, loss: 0.10249156504869461
step: 140, loss: 0.07497628033161163
step: 150, loss: 0.048713259398937225
step: 160, loss: 0.0745980367064476
step: 170, loss: 0.13148871064186096
step: 180, loss: 0.01216201763600111
step: 190, loss: 0.12133575230836868
step: 200, loss: 0.03143659979104996
step: 210, loss: 0.11945132911205292
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.

```



```

    warnings.warn("An input tensor was not cuda.")
epoch 8: dev_f1=0.7283464566929133, f1=0.763265306122449,
best_f1=0.7609942638623327
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.04193602874875069
step: 10, loss: 0.0933217778801918
step: 20, loss: 0.11987560987472534
step: 30, loss: 0.0015712647000327706
step: 40, loss: 0.04738759249448776
step: 50, loss: 0.03468037396669388
step: 60, loss: 0.0007012215792201459
step: 70, loss: 0.10793599486351013
step: 80, loss: 0.0033379103988409042
step: 90, loss: 0.007580465637147427
step: 100, loss: 0.059815168380737305
step: 110, loss: 0.09215406328439713
step: 120, loss: 0.09674745798110962
step: 130, loss: 0.04631650447845459
step: 140, loss: 0.09008044004440308
step: 150, loss: 0.11698154360055923
step: 160, loss: 0.07434658706188202
step: 170, loss: 0.04037019610404968
step: 180, loss: 0.07395915687084198
step: 190, loss: 0.015377569012343884
step: 200, loss: 0.0020821955986320972
step: 210, loss: 0.012771873734891415
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 9: dev_f1=0.7378277153558052, f1=0.7642585551330798,
best_f1=0.7609942638623327
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0023417570628225803
step: 10, loss: 0.033475108444690704
step: 20, loss: 0.003992828540503979
step: 30, loss: 0.0024888208135962486
step: 40, loss: 0.0008284489740617573
step: 50, loss: 0.012039068154990673
step: 60, loss: 0.002520275302231312
step: 70, loss: 0.004989134147763252
step: 80, loss: 0.02694028988480568
step: 90, loss: 0.0516836978495121
step: 100, loss: 0.032060958445072174
step: 110, loss: 0.0033176583237946033

```

```

step: 120, loss: 0.043172791600227356
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 130, loss: 0.011040942743420601
step: 140, loss: 0.11778946220874786
step: 150, loss: 0.0397784523665905
step: 160, loss: 0.01103151310235262
step: 170, loss: 0.03945482522249222
step: 180, loss: 0.030130626633763313
step: 190, loss: 0.017811425030231476
step: 200, loss: 0.03731345385313034
step: 210, loss: 0.009011439047753811
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
epoch 10: dev_f1=0.7500000000000001, f1=0.771266540642722,
best_f1=0.7609942638623327
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.061355456709861755
step: 10, loss: 0.0027701915241777897
step: 20, loss: 0.011184193193912506
step: 30, loss: 0.06554067134857178
step: 40, loss: 0.009844967164099216
step: 50, loss: 0.005602544639259577
step: 60, loss: 0.0015660569770261645
step: 70, loss: 0.008942599408328533
step: 80, loss: 0.029388444498181343
step: 90, loss: 0.18829256296157837
step: 100, loss: 0.0018384400755167007
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 8192.0
step: 110, loss: 0.0473119355738163
step: 120, loss: 0.00463560875505209
step: 130, loss: 0.0014326302334666252
step: 140, loss: 0.07529426366090775
step: 150, loss: 0.020087260752916336
step: 160, loss: 0.0009661871590651572
step: 170, loss: 0.037069279700517654
step: 180, loss: 0.013769460842013359
step: 190, loss: 0.06048693507909775
step: 200, loss: 0.10112755000591278
step: 210, loss: 0.004710113164037466
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
epoch 11: dev_f1=0.7415730337078651, f1=0.7588785046728973,
best_f1=0.7609942638623327
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:

```

```

An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.02421334758400917
step: 10, loss: 0.00851280614733696
step: 20, loss: 0.020401788875460625
step: 30, loss: 0.003053627908229828
step: 40, loss: 0.0005337647162377834
step: 50, loss: 0.014833934605121613
step: 60, loss: 0.0001338611327810213
step: 70, loss: 0.006422910373657942
step: 80, loss: 0.2596467435359955
step: 90, loss: 0.015255727805197239
step: 100, loss: 0.00036078214179724455
step: 110, loss: 0.0014761155471205711
step: 120, loss: 0.0005810815491713583
step: 130, loss: 0.06885822862386703
step: 140, loss: 0.00221213954500854
step: 150, loss: 0.0002655235002748668
step: 160, loss: 0.00353041454218328
step: 170, loss: 0.07127515226602554
step: 180, loss: 0.018916329368948936
step: 190, loss: 0.010448228567838669
step: 200, loss: 0.00023000565124675632
step: 210, loss: 0.00029045145493000746
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
epoch 12: dev_f1=0.7395626242544732, f1=0.7670682730923696,
best_f1=0.7609942638623327
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.019276762381196022
step: 10, loss: 0.022088490426540375
step: 20, loss: 0.00024410267360508442
step: 30, loss: 0.00031457695877179503
step: 40, loss: 0.015321695245802402
step: 50, loss: 0.10443748533725739
step: 60, loss: 0.0028999452479183674
step: 70, loss: 0.000750401639379561
step: 80, loss: 0.0378311350941658
step: 90, loss: 0.03437839075922966
step: 100, loss: 0.02370702661573887
step: 110, loss: 0.03513679280877113
step: 120, loss: 0.0029362947680056095
step: 130, loss: 0.008087826892733574
step: 140, loss: 0.024889735504984856
step: 150, loss: 0.004012423101812601

```

```

step: 160, loss: 0.0022092973813414574
step: 170, loss: 0.005232335068285465
step: 180, loss: 0.00018102613103110343
step: 190, loss: 0.01782781444489956
step: 200, loss: 0.033859122544527054
step: 210, loss: 0.01429899875074625
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 13: dev_f1=0.7418738049713194, f1=0.7624521072796935,
best_f1=0.7609942638623327
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.000334554206347093
step: 10, loss: 0.08361169695854187
step: 20, loss: 0.003376903012394905
step: 30, loss: 0.0003655703039839864
step: 40, loss: 0.01757080852985382
step: 50, loss: 0.0019879986066371202
step: 60, loss: 0.015291525050997734
step: 70, loss: 0.013414213433861732
step: 80, loss: 0.021929292008280754
step: 90, loss: 0.00018056205590255558
step: 100, loss: 0.00048633693950250745
step: 110, loss: 0.0005986107280477881
step: 120, loss: 0.0001250401692232117
step: 130, loss: 0.007153257727622986
step: 140, loss: 0.0008118004770949483
step: 150, loss: 0.003835089271888137
step: 160, loss: 0.00014245160855352879
step: 170, loss: 0.030508460476994514
step: 180, loss: 0.0003679692163132131
step: 190, loss: 0.051826588809490204
step: 200, loss: 0.021271273493766785
step: 210, loss: 0.0009851650102064013
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 14: dev_f1=0.7283702213279678, f1=0.7710843373493976,
best_f1=0.7609942638623327
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.04246653988957405
step: 10, loss: 0.0002953496004920453
step: 20, loss: 0.018598996102809906
step: 30, loss: 0.0038139137905091047

```

```

step: 40, loss: 0.007795573212206364
step: 50, loss: 0.00023318050079979002
step: 60, loss: 0.010447831824421883
step: 70, loss: 0.007631276268512011
step: 80, loss: 0.07339036464691162
step: 90, loss: 0.00028938849573023617
step: 100, loss: 0.0015640300698578358
step: 110, loss: 0.011780875734984875
step: 120, loss: 0.0009698583744466305
step: 130, loss: 0.00014571168867405504
step: 140, loss: 0.00432116212323308
step: 150, loss: 0.00509232934564352
step: 160, loss: 0.0009356563678011298
step: 170, loss: 0.02086532488465309
step: 180, loss: 0.0001719063293421641
step: 190, loss: 0.004382380750030279
step: 200, loss: 0.0004230072954669595
step: 210, loss: 0.05403356999158859
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
epoch 15: dev_f1=0.7227926078028748, f1=0.769857433808554,
best_f1=0.7609942638623327
mv: cannot stat '*_dev.pt': No such file or directory
Selected optimization level 02: FP16 training with FP32 batchnorm and FP32
master weights.

Defaults for this optimization level are:
enabled                : True
opt_level              : 02
cast_model_type        : torch.float16
patch_torch_functions  : False
keep_batchnorm_fp32    : True
master_weights         : True
loss_scale             : dynamic
Processing user overrides (additional kwargs that are not None)...
After processing overrides, optimization options are:
enabled                : True
opt_level              : 02
cast_model_type        : torch.float16
patch_torch_functions  : False
keep_batchnorm_fp32    : True
master_weights         : True
loss_scale             : dynamic
Warning: multi_tensor_applier fused unscale kernel is unavailable, possibly
because apex was installed without --cuda_ext --cpp_ext. Using Python fallback.
Original ImportError was: ModuleNotFoundError("No module named 'amp_C'")
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:

```

An input tensor was not cuda.

```
warnings.warn("An input tensor was not cuda.")
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
2293it [00:09, 243.84it/s]
load_f1 = 0.764378478664193
real_f1 = 0.7629629629629631
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
4398it [00:31, 141.27it/s]
```

### 2.1.5 Walmart-Amazon - Running the matcher

```
[ ]: # Step 1:
# train the wdc model with the save_model flag on
# it will produce two files *_dev.pt and *_test.pt which are the
# best checkpoints on the dev set and the test set.
!CUDA_VISIBLE_DEVICES=0 python train_ditto.py \
  --task Structured/Walmart-Amazon \
  --batch_size 32 \
  --max_len 256 \
  --lr 3e-5 \
  --n_epochs 15 \
  --finetuning \
  --lm roberta \
  --fp16 \
  --save_model

# Step 2: pick a checkpoint, rename it to wdc_all_small.pt
!mv *_dev.pt checkpoints/Walmart-Amazon.pt

# Step 3: run the matcher
!CUDA_VISIBLE_DEVICES=0 python matcher.py \
  --task Structured/Walmart-Amazon \
  --input_path input/input_small.jsonl \
  --output_path output/output_small.jsonl \
  --lm roberta \
  --use_gpu \
  --fp16 \
  --checkpoint_path checkpoints/
```

```
Downloading: 100% 481/481 [00:00<00:00, 415kB/s]
Downloading: 100% 899k/899k [00:01<00:00, 828kB/s]
Downloading: 100% 456k/456k [00:01<00:00, 427kB/s]
```

Downloading: 100% 501M/501M [00:10<00:00, 45.9MB/s]  
Selected optimization level 02: FP16 training with FP32 batchnorm and FP32 master weights.

Defaults for this optimization level are:

enabled : True  
opt\_level : 02  
cast\_model\_type : torch.float16  
patch\_torch\_functions : False  
keep\_batchnorm\_fp32 : True  
master\_weights : True  
loss\_scale : dynamic

Processing user overrides (additional kwargs that are not None)...

After processing overrides, optimization options are:

enabled : True  
opt\_level : 02  
cast\_model\_type : torch.float16  
patch\_torch\_functions : False  
keep\_batchnorm\_fp32 : True  
master\_weights : True  
loss\_scale : dynamic

Warning: multi\_tensor\_applier fused unscale kernel is unavailable, possibly because apex was installed without --cuda\_ext --cpp\_ext. Using Python fallback. Original ImportError was: ModuleNotFoundError("No module named 'amp\_C'")  
/usr/local/lib/python3.7/dist-packages/apex/amp/\_initialize.py:25: UserWarning: An input tensor was not cuda.

warnings.warn("An input tensor was not cuda.")

Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 32768.0  
/usr/local/lib/python3.7/dist-packages/torch/optim/lr\_scheduler.py:136:

UserWarning: Detected call of `lr\_scheduler.step()` before `optimizer.step()`. In PyTorch 1.1.0 and later, you should call them in the opposite order: `optimizer.step()` before `lr\_scheduler.step()`. Failure to do this will result in PyTorch skipping the first value of the learning rate schedule. See more details at <https://pytorch.org/docs/stable/optim.html#how-to-adjust-learning-rate>

"<https://pytorch.org/docs/stable/optim.html#how-to-adjust-learning-rate>",  
UserWarning)

step: 0, loss: 0.5017979741096497

Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0

step: 10, loss: 0.40285807847976685

step: 20, loss: 0.27302366495132446

step: 30, loss: 0.4278043210506439

step: 40, loss: 0.2448272705078125

step: 50, loss: 0.3206791281700134

step: 60, loss: 0.5824809074401855

step: 70, loss: 0.4220569133758545

step: 80, loss: 0.14262695610523224

step: 90, loss: 0.30581510066986084

```

step: 100, loss: 0.5216330885887146
step: 110, loss: 0.23688894510269165
step: 120, loss: 0.3006764352321625
step: 130, loss: 0.3317714035511017
step: 140, loss: 0.17916011810302734
step: 150, loss: 0.2981002926826477
step: 160, loss: 0.22049608826637268
step: 170, loss: 0.3744588792324066
step: 180, loss: 0.16198481619358063
step: 190, loss: 0.17646296322345734
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
    warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 1: dev_f1=0.17216770740410348, f1=0.17216770740410348,
best_f1=0.17216770740410348
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.3968813419342041
step: 10, loss: 0.32711702585220337
step: 20, loss: 0.614321231842041
step: 30, loss: 0.24806298315525055
step: 40, loss: 0.5501478910446167
step: 50, loss: 0.30991318821907043
step: 60, loss: 0.47705158591270447
step: 70, loss: 0.3180234432220459
step: 80, loss: 0.15000316500663757
step: 90, loss: 0.31508761644363403
step: 100, loss: 0.2594681978225708
step: 110, loss: 0.3866274058818817
step: 120, loss: 0.2329753339290619
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 8192.0
step: 130, loss: 0.5201711654663086
step: 140, loss: 0.32154330611228943
step: 150, loss: 0.3197554349899292
step: 160, loss: 0.31268438696861267
step: 170, loss: 0.24141918122768402
step: 180, loss: 0.17675065994262695
step: 190, loss: 0.2543953061103821
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 2: dev_f1=0.17216770740410348, f1=0.17216770740410348,
best_f1=0.17216770740410348

```



```
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:  
An input tensor was not cuda.
```

```
    warnings.warn("An input tensor was not cuda.")
```

```
step: 0, loss: 0.3834710419178009  
step: 10, loss: 0.3753359913825989  
step: 20, loss: 0.4597625434398651  
step: 30, loss: 0.3245965540409088  
step: 40, loss: 0.08409526199102402  
step: 50, loss: 0.3902059495449066  
step: 60, loss: 0.15876856446266174  
step: 70, loss: 0.3913951516151428  
step: 80, loss: 0.3078456223011017  
step: 90, loss: 0.36614561080932617  
step: 100, loss: 0.53153395652771  
step: 110, loss: 0.6550169587135315  
step: 120, loss: 0.367667019367218  
step: 130, loss: 0.15240758657455444  
step: 140, loss: 0.3598162829875946  
step: 150, loss: 0.3458768427371979  
step: 160, loss: 0.5673648715019226  
step: 170, loss: 0.4608994424343109  
step: 180, loss: 0.38291794061660767  
step: 190, loss: 0.16840149462223053
```

```
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:  
An input tensor was not cuda.
```

```
    warnings.warn("An input tensor was not cuda.")
```

```
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
```

```
UserWarning: Please also save or load the state of the optimizer when saving or  
loading the scheduler.
```

```
    warnings.warn(SAVE_STATE_WARNING, UserWarning)
```

```
epoch 3: dev_f1=0.23747680890538034, f1=0.22047244094488191,  
best_f1=0.22047244094488191
```

```
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:  
An input tensor was not cuda.
```

```
    warnings.warn("An input tensor was not cuda.")
```

```
step: 0, loss: 0.24378782510757446  
step: 10, loss: 0.20499679446220398  
step: 20, loss: 0.2615089416503906  
step: 30, loss: 0.228849396109581  
step: 40, loss: 0.6322762370109558  
step: 50, loss: 0.2634703516960144  
step: 60, loss: 0.35484346747398376  
step: 70, loss: 0.2650216519832611  
step: 80, loss: 0.29524126648902893  
step: 90, loss: 0.16274045407772064  
step: 100, loss: 0.3528725504875183  
step: 110, loss: 0.3694107234477997  
step: 120, loss: 0.24575737118721008
```

```

step: 130, loss: 0.43232494592666626
step: 140, loss: 0.3185557425022125
step: 150, loss: 0.2518003582954407
step: 160, loss: 0.3377344608306885
step: 170, loss: 0.43147242069244385
step: 180, loss: 0.38398969173431396
step: 190, loss: 0.149248868227005
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 4: dev_f1=0.17216770740410348, f1=0.17216770740410348,
best_f1=0.22047244094488191
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.41427287459373474
step: 10, loss: 0.3880874812602997
step: 20, loss: 0.1724477857351303
step: 30, loss: 0.11580988019704819
step: 40, loss: 0.3113260567188263
step: 50, loss: 0.4951784312725067
step: 60, loss: 0.23299026489257812
step: 70, loss: 0.37513914704322815
step: 80, loss: 0.3635166883468628
step: 90, loss: 0.292759507894516
step: 100, loss: 0.46079716086387634
step: 110, loss: 0.3490128219127655
step: 120, loss: 0.21151922643184662
step: 130, loss: 0.5685973763465881
step: 140, loss: 0.3793469965457916
step: 150, loss: 0.31308749318122864
step: 160, loss: 0.17396119236946106
step: 170, loss: 0.3715536594390869
step: 180, loss: 0.2455863058567047
step: 190, loss: 0.31009456515312195
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 5: dev_f1=0.17216770740410348, f1=0.17216770740410348,
best_f1=0.22047244094488191
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.31379854679107666
step: 10, loss: 0.2475428283214569
step: 20, loss: 0.31684648990631104
step: 30, loss: 0.4761769771575928
step: 40, loss: 0.25231775641441345

```

```

step: 50, loss: 0.308819055557251
step: 60, loss: 0.44306978583335876
step: 70, loss: 0.3146838843822479
step: 80, loss: 0.3252356946468353
step: 90, loss: 0.23737996816635132
step: 100, loss: 0.4742829501628876
step: 110, loss: 0.2495454102754593
step: 120, loss: 0.4567742943763733
step: 130, loss: 0.5346904993057251
step: 140, loss: 0.20419888198375702
step: 150, loss: 0.39198121428489685
step: 160, loss: 0.3787420690059662
step: 170, loss: 0.38625526428222656
step: 180, loss: 0.18110550940036774
step: 190, loss: 0.31147530674934387
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 6: dev_f1=0.17216770740410348, f1=0.17216770740410348,
best_f1=0.22047244094488191
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.3085685968399048
step: 10, loss: 0.3129045367240906
step: 20, loss: 0.30666548013687134
step: 30, loss: 0.1961202174425125
step: 40, loss: 0.31056466698646545
step: 50, loss: 0.08343295753002167
step: 60, loss: 0.1604185253381729
step: 70, loss: 0.1560460925102234
step: 80, loss: 0.23984239995479584
step: 90, loss: 0.23844219744205475
step: 100, loss: 0.5679486989974976
step: 110, loss: 0.42836135625839233
step: 120, loss: 0.38256412744522095
step: 130, loss: 0.314273864030838
step: 140, loss: 0.24950289726257324
step: 150, loss: 0.30525678396224976
step: 160, loss: 0.38754788041114807
step: 170, loss: 0.3063633441925049
step: 180, loss: 0.24315683543682098
step: 190, loss: 0.3079119920730591
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 7: dev_f1=0.17216770740410348, f1=0.17216770740410348,
best_f1=0.22047244094488191

```

```

/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.24438558518886566
step: 10, loss: 0.39531823992729187
step: 20, loss: 0.23273324966430664
step: 30, loss: 0.37999606132507324
step: 40, loss: 0.1614466905593872
step: 50, loss: 0.29455599188804626
step: 60, loss: 0.4401976764202118
step: 70, loss: 0.2403377741575241
step: 80, loss: 0.3834962546825409
step: 90, loss: 0.2331942915916443
step: 100, loss: 0.29231810569763184
step: 110, loss: 0.376844584941864
step: 120, loss: 0.3324770927429199
step: 130, loss: 0.2888263165950775
step: 140, loss: 0.36028462648391724
step: 150, loss: 0.29349711537361145
step: 160, loss: 0.1635567545890808
step: 170, loss: 0.2242169976234436
step: 180, loss: 0.27776429057121277
step: 190, loss: 0.2686429023742676
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 8: dev_f1=0.2054113749309774, f1=0.20907079646017698,
best_f1=0.22047244094488191
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.23398298025131226
step: 10, loss: 0.0954328328371048
step: 20, loss: 0.30480799078941345
step: 30, loss: 0.1766965091228485
step: 40, loss: 0.28146645426750183
step: 50, loss: 0.3504149317741394
step: 60, loss: 0.3713768720626831
step: 70, loss: 0.15471315383911133
step: 80, loss: 0.35215821862220764
step: 90, loss: 0.751875638961792
step: 100, loss: 0.3423362970352173
step: 110, loss: 0.33901891112327576
step: 120, loss: 0.5568284392356873
step: 130, loss: 0.29157257080078125
step: 140, loss: 0.2923683226108551
step: 150, loss: 0.24052301049232483
step: 160, loss: 0.20920532941818237

```

```

step: 170, loss: 0.525087833404541
step: 180, loss: 0.3777618706226349
step: 190, loss: 0.23756098747253418
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 9: dev_f1=0.21052631578947367, f1=0.21733668341708542,
best_f1=0.22047244094488191
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.2881380319595337
step: 10, loss: 0.15165960788726807
step: 20, loss: 0.30104219913482666
step: 30, loss: 0.35307547450065613
step: 40, loss: 0.3286910057067871
step: 50, loss: 0.11486823856830597
step: 60, loss: 0.2811933159828186
step: 70, loss: 0.3207237720489502
step: 80, loss: 0.41485396027565
step: 90, loss: 0.13272646069526672
step: 100, loss: 0.20261342823505402
step: 110, loss: 0.19733569025993347
step: 120, loss: 0.34048938751220703
step: 130, loss: 0.42876139283180237
step: 140, loss: 0.25090643763542175
step: 150, loss: 0.22118723392486572
step: 160, loss: 0.11191124469041824
step: 170, loss: 0.1575315147638321
step: 180, loss: 0.22961872816085815
step: 190, loss: 0.13138067722320557
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
    warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 10: dev_f1=0.5925925925925926, f1=0.6217391304347826,
best_f1=0.6217391304347826
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.3351566791534424
step: 10, loss: 0.04755463823676109
step: 20, loss: 0.1120537668466568
step: 30, loss: 0.1787048578262329
step: 40, loss: 0.010723310522735119

```

```

step: 50, loss: 0.13892249763011932
step: 60, loss: 0.10930854082107544
step: 70, loss: 0.26314830780029297
step: 80, loss: 0.3830452561378479
step: 90, loss: 0.35612550377845764
step: 100, loss: 0.059961073100566864
step: 110, loss: 0.1792360097169876
step: 120, loss: 0.20771917700767517
step: 130, loss: 0.4489981234073639
step: 140, loss: 0.33146458864212036
step: 150, loss: 0.052619773894548416
step: 160, loss: 0.09713487327098846
step: 170, loss: 0.3830624222755432
step: 180, loss: 0.10303975641727448
step: 190, loss: 0.047737568616867065
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
    warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 11: dev_f1=0.7153652392947103, f1=0.7362924281984334,
best_f1=0.7362924281984334
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.22973304986953735
step: 10, loss: 0.12338857352733612
step: 20, loss: 0.14567290246486664
step: 30, loss: 0.19848376512527466
step: 40, loss: 0.03279941901564598
step: 50, loss: 0.10755103826522827
step: 60, loss: 0.16185003519058228
step: 70, loss: 0.15456917881965637
step: 80, loss: 0.2029474824666977
step: 90, loss: 0.2265867292881012
step: 100, loss: 0.25500771403312683
step: 110, loss: 0.4177381098270416
step: 120, loss: 0.029450489208102226
step: 130, loss: 0.0372810959815979
step: 140, loss: 0.19836364686489105
step: 150, loss: 0.1407596319913864
step: 160, loss: 0.05603223294019699
step: 170, loss: 0.2862500846385956
step: 180, loss: 0.10007712990045547
step: 190, loss: 0.11228465288877487
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:

```

```

An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
  warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 12: dev_f1=0.7379134860050889, f1=0.7647058823529412,
best_f1=0.7647058823529412
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.1629633903503418
step: 10, loss: 0.06041986495256424
step: 20, loss: 0.03008899837732315
step: 30, loss: 0.10824183374643326
step: 40, loss: 0.11937393248081207
step: 50, loss: 0.06900127232074738
step: 60, loss: 0.07977078855037689
step: 70, loss: 0.09366366267204285
step: 80, loss: 0.03588911518454552
step: 90, loss: 0.05961694195866585
step: 100, loss: 0.03802022337913513
step: 110, loss: 0.006725197192281485
step: 120, loss: 0.0588214173913002
step: 130, loss: 0.12664255499839783
step: 140, loss: 0.19482989609241486
step: 150, loss: 0.07219048589468002
step: 160, loss: 0.20487229526042938
step: 170, loss: 0.13235284388065338
step: 180, loss: 0.04117277264595032
step: 190, loss: 0.30997639894485474
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
  warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 13: dev_f1=0.7546174142480211, f1=0.777142857142857,
best_f1=0.777142857142857
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.11389471590518951
step: 10, loss: 0.05307191237807274
step: 20, loss: 0.07719743251800537
step: 30, loss: 0.018949100747704506
step: 40, loss: 0.0116178710013628

```

```

step: 50, loss: 0.021350862458348274
step: 60, loss: 0.012054765596985817
step: 70, loss: 0.029689613729715347
step: 80, loss: 0.04757576808333397
step: 90, loss: 0.10080884397029877
step: 100, loss: 0.007213676813989878
step: 110, loss: 0.01148483157157898
step: 120, loss: 0.029204463586211205
step: 130, loss: 0.01872873678803444
step: 140, loss: 0.09808875620365143
step: 150, loss: 0.05223127827048302
step: 160, loss: 0.056166451424360275
step: 170, loss: 0.014814157038927078
step: 180, loss: 0.05743377283215523
step: 190, loss: 0.25966036319732666
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
    warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 14: dev_f1=0.761904761904762, f1=0.7758620689655173,
best_f1=0.7758620689655173
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.16292016208171844
step: 10, loss: 0.10124858468770981
step: 20, loss: 0.12962716817855835
step: 30, loss: 0.01110763382166624
step: 40, loss: 0.022503728047013283
step: 50, loss: 0.01872311159968376
step: 60, loss: 0.013269441202282906
step: 70, loss: 0.07874743640422821
step: 80, loss: 0.06329129636287689
step: 90, loss: 0.24327898025512695
step: 100, loss: 0.14799253642559052
step: 110, loss: 0.04501137137413025
step: 120, loss: 0.05953183397650719
step: 130, loss: 0.04558004438877106
step: 140, loss: 0.19392187893390656
step: 150, loss: 0.006697145290672779
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 8192.0
step: 160, loss: 0.0501188300549984
step: 170, loss: 0.19538787007331848
step: 180, loss: 0.01957465149462223
step: 190, loss: 0.055690113455057144

```



```

/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
    warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 15: dev_f1=0.765625, f1=0.7796610169491526, best_f1=0.7796610169491526
mv: cannot stat '*_dev.pt': No such file or directory
Selected optimization level 02: FP16 training with FP32 batchnorm and FP32
master weights.

Defaults for this optimization level are:
enabled                : True
opt_level              : 02
cast_model_type        : torch.float16
patch_torch_functions  : False
keep_batchnorm_fp32    : True
master_weights         : True
loss_scale             : dynamic
Processing user overrides (additional kwargs that are not None)...
After processing overrides, optimization options are:
enabled                : True
opt_level              : 02
cast_model_type        : torch.float16
patch_torch_functions  : False
keep_batchnorm_fp32    : True
master_weights         : True
loss_scale             : dynamic
Warning: multi_tensor_applier fused unscale kernel is unavailable, possibly
because apex was installed without --cuda_ext --cpp_ext. Using Python fallback.
Original ImportError was: ModuleNotFoundError("No module named 'amp_C'")
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
2049it [00:12, 162.98it/s]
load_f1 = 0.7584415584415585
real_f1 = 0.7556675062972291
733it [00:00, 3555.35it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
4398it [00:30, 143.94it/s]

```

## 2.2 BASELINE TEXTUAL

### 2.2.1 Abt-Buy - Running the matcher

```
[ ]: # Step 1:
# train the wdc model with the save_model flag on
# it will produce two files *_dev.pt and *_test.pt which are the
# best checkpoints on the dev set and the test set.
!CUDA_VISIBLE_DEVICES=0 python train_ditto.py \
  --task Textual/Abt-Buy \
  --batch_size 32 \
  --max_len 256 \
  --lr 3e-5 \
  --n_epochs 15 \
  --finetuning \
  --lm roberta \
  --fp16 \
  --save_model

# Step 2: pick a checkpoint, rename it to wdc_all_small.pt
!mv *_dev.pt checkpoints/Abt-Buy.pt

# Step 3: run the matcher
!CUDA_VISIBLE_DEVICES=0 python matcher.py \
  --task Textual/Abt-Buy \
  --input_path input/input_small.jsonl \
  --output_path output/output_small.jsonl \
  --lm roberta \
  --use_gpu \
  --fp16 \
  --checkpoint_path checkpoints/
```

Downloading: 100% 481/481 [00:00<00:00, 327kB/s]

Downloading: 100% 899k/899k [00:00<00:00, 5.92MB/s]

Downloading: 100% 456k/456k [00:00<00:00, 3.99MB/s]

Downloading: 100% 501M/501M [00:07<00:00, 65.6MB/s]

Selected optimization level 02: FP16 training with FP32 batchnorm and FP32 master weights.

Defaults for this optimization level are:

enabled	: True
opt_level	: 02
cast_model_type	: torch.float16
patch_torch_functions	: False
keep_batchnorm_fp32	: True
master_weights	: True
loss_scale	: dynamic

```

Processing user overrides (additional kwargs that are not None)...
After processing overrides, optimization options are:
enabled                : True
opt_level              : 02
cast_model_type        : torch.float16
patch_torch_functions  : False
keep_batchnorm_fp32    : True
master_weights         : True
loss_scale              : dynamic
Warning: multi_tensor_applier fused unscale kernel is unavailable, possibly
because apex was installed without --cuda_ext --cpp_ext. Using Python fallback.
Original ImportError was: ModuleNotFoundError("No module named 'amp_C'")
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 32768.0
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:136:
UserWarning: Detected call of `lr_scheduler.step()` before `optimizer.step()`.
In PyTorch 1.1.0 and later, you should call them in the opposite order:
`optimizer.step()` before `lr_scheduler.step()`. Failure to do this will result
in PyTorch skipping the first value of the learning rate schedule. See more
details at https://pytorch.org/docs/stable/optim.html#how-to-adjust-learning-
rate
    "https://pytorch.org/docs/stable/optim.html#how-to-adjust-learning-rate",
UserWarning)
step: 0, loss: 0.49295246601104736
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 10, loss: 0.4649890065193176
step: 20, loss: 0.3968137204647064
step: 30, loss: 0.4020492732524872
step: 40, loss: 0.5705543756484985
step: 50, loss: 0.3436630666255951
step: 60, loss: 0.6153914928436279
step: 70, loss: 0.36175376176834106
step: 80, loss: 0.24484983086585999
step: 90, loss: 0.25098779797554016
step: 100, loss: 0.12849262356758118
step: 110, loss: 0.4095313847064972
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 8192.0
step: 120, loss: 0.3121628165245056
step: 130, loss: 0.3164382576942444
step: 140, loss: 0.37694495916366577
step: 150, loss: 0.31976577639579773
step: 160, loss: 0.4134734869003296
step: 170, loss: 0.3271818459033966
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")

```

```

/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
    warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 1: dev_f1=0.24489795918367352, f1=0.23633879781420766,
best_f1=0.23633879781420766
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.30893266201019287
step: 10, loss: 0.4193643033504486
step: 20, loss: 0.25482049584388733
step: 30, loss: 0.22085751593112946
step: 40, loss: 0.08630675822496414
step: 50, loss: 0.35223257541656494
step: 60, loss: 0.17274771630764008
step: 70, loss: 0.2419525682926178
step: 80, loss: 0.16215671598911285
step: 90, loss: 0.08547740429639816
step: 100, loss: 0.2788526117801666
step: 110, loss: 0.08936242014169693
step: 120, loss: 0.10113023221492767
step: 130, loss: 0.22300158441066742
step: 140, loss: 0.22306327521800995
step: 150, loss: 0.16276614367961884
step: 160, loss: 0.1426355540752411
step: 170, loss: 0.04304930940270424
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
    warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 2: dev_f1=0.826923076923077, f1=0.7941176470588236,
best_f1=0.7941176470588236
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.15682852268218994
step: 10, loss: 0.0736805722117424
step: 20, loss: 0.016053732484579086
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 4096.0
step: 30, loss: 0.05139761418104172
step: 40, loss: 0.03212784603238106
step: 50, loss: 0.09880807995796204
step: 60, loss: 0.06310883164405823
step: 70, loss: 0.19281744956970215

```

```

step: 80, loss: 0.07826120406389236
step: 90, loss: 0.08654481917619705
step: 100, loss: 0.05900871008634567
step: 110, loss: 0.025954782962799072
step: 120, loss: 0.023260265588760376
step: 130, loss: 0.06731081753969193
step: 140, loss: 0.1573471575975418
step: 150, loss: 0.012186951003968716
step: 160, loss: 0.14556744694709778
step: 170, loss: 0.02025310881435871
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
    warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 3: dev_f1=0.8523002421307506, f1=0.8738317757009346,
best_f1=0.8738317757009346
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.05553210899233818
step: 10, loss: 0.09351874887943268
step: 20, loss: 0.005422241520136595
step: 30, loss: 0.07221023738384247
step: 40, loss: 0.09631325304508209
step: 50, loss: 0.03367120772600174
step: 60, loss: 0.033317212015390396
step: 70, loss: 0.005301726050674915
step: 80, loss: 0.1706358939409256
step: 90, loss: 0.20311330258846283
step: 100, loss: 0.07110562920570374
step: 110, loss: 0.03849765658378601
step: 120, loss: 0.20465362071990967
step: 130, loss: 0.10723356157541275
step: 140, loss: 0.10617174953222275
step: 150, loss: 0.046412818133831024
step: 160, loss: 0.0007809597882442176
step: 170, loss: 0.0583357959985733
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 4: dev_f1=0.8416289592760181, f1=0.8546637744034706,
best_f1=0.8738317757009346
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")

```

```

step: 0, loss: 0.016569357365369797
step: 10, loss: 0.03046363778412342
step: 20, loss: 0.017840204760432243
step: 30, loss: 0.003280090866610408
step: 40, loss: 0.06662017852067947
step: 50, loss: 0.04174260050058365
step: 60, loss: 0.03108222782611847
step: 70, loss: 0.014453683979809284
step: 80, loss: 0.0007560029625892639
step: 90, loss: 0.05755985155701637
step: 100, loss: 0.037164103239774704
step: 110, loss: 0.16041499376296997
step: 120, loss: 0.033952172845602036
step: 130, loss: 0.002370049711316824
step: 140, loss: 0.026518110185861588
step: 150, loss: 0.13457772135734558
step: 160, loss: 0.0024513357784599066
step: 170, loss: 0.004858340602368116
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 5: dev_f1=0.8495145631067961, f1=0.8720379146919431,
best_f1=0.8738317757009346
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.08535362035036087
step: 10, loss: 0.014635118655860424
step: 20, loss: 0.06816995143890381
step: 30, loss: 0.0004309157666284591
step: 40, loss: 0.0028782987501472235
step: 50, loss: 0.001812261063605547
step: 60, loss: 0.023898007348179817
step: 70, loss: 0.03625553846359253
step: 80, loss: 0.0008473354391753674
step: 90, loss: 0.04742775857448578
step: 100, loss: 0.00261751189827919
step: 110, loss: 0.021124953404068947
step: 120, loss: 0.012842721305787563
step: 130, loss: 0.011076517403125763
step: 140, loss: 0.031174909323453903
step: 150, loss: 0.0015946110943332314
step: 160, loss: 0.051993176341056824
step: 170, loss: 0.013313358649611473
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:

```

UserWarning: Please also save or load the state of the optimizer when saving or loading the scheduler.

```
warnings.warn(SAVE_STATE_WARNING, UserWarning)
```

epoch 6: dev\_f1=0.8728179551122195, f1=0.8921568627450982,

best\_f1=0.8921568627450982

/usr/local/lib/python3.7/dist-packages/apex/amp/\_initialize.py:25: UserWarning: An input tensor was not cuda.

```
warnings.warn("An input tensor was not cuda.")
```

step: 0, loss: 0.0919945165514946

step: 10, loss: 0.04762427136301994

step: 20, loss: 0.0009330146131105721

step: 30, loss: 0.12738123536109924

step: 40, loss: 0.00012871180661022663

step: 50, loss: 0.002185692312195897

step: 60, loss: 0.00859749410301447

step: 70, loss: 0.007382649928331375

step: 80, loss: 0.0032559477258473635

step: 90, loss: 0.06422232836484909

step: 100, loss: 0.0005004914128221571

step: 110, loss: 0.2282676100730896

step: 120, loss: 0.019786231219768524

step: 130, loss: 0.04413633421063423

step: 140, loss: 0.0006158050382509828

step: 150, loss: 0.12296055257320404

step: 160, loss: 0.0010908639524132013

step: 170, loss: 0.03362209349870682

/usr/local/lib/python3.7/dist-packages/apex/amp/\_initialize.py:25: UserWarning: An input tensor was not cuda.

```
warnings.warn("An input tensor was not cuda.")
```

epoch 7: dev\_f1=0.8702290076335877, f1=0.8811881188118812,

best\_f1=0.8921568627450982

/usr/local/lib/python3.7/dist-packages/apex/amp/\_initialize.py:25: UserWarning: An input tensor was not cuda.

```
warnings.warn("An input tensor was not cuda.")
```

step: 0, loss: 0.050743672996759415

step: 10, loss: 0.0005756033933721483

step: 20, loss: 0.014954871498048306

step: 30, loss: 0.0018485200125724077

step: 40, loss: 0.0002971018257085234

step: 50, loss: 0.002165545942261815

step: 60, loss: 0.0017335864249616861

step: 70, loss: 0.03969765454530716

step: 80, loss: 0.0005983088631182909

step: 90, loss: 0.09312620013952255

step: 100, loss: 0.000501199800055474

step: 110, loss: 0.07453451305627823

step: 120, loss: 0.06887151300907135

step: 130, loss: 0.049807108938694

```

step: 140, loss: 0.03210831806063652
step: 150, loss: 0.0012967647053301334
step: 160, loss: 0.0001802234910428524
step: 170, loss: 0.030307911336421967
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
    warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 8: dev_f1=0.8746803069053709, f1=0.8793969849246231,
best_f1=0.8793969849246231
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.007035400252789259
step: 10, loss: 0.004304367583245039
step: 20, loss: 0.0007080160430632532
step: 30, loss: 0.0003362132702022791
step: 40, loss: 0.010903763584792614
step: 50, loss: 0.0009328075684607029
step: 60, loss: 0.011553088203072548
step: 70, loss: 0.0003776691446546465
step: 80, loss: 0.0002678402524907142
step: 90, loss: 0.0010567625286057591
step: 100, loss: 0.011032181791961193
step: 110, loss: 0.09243927150964737
step: 120, loss: 0.004296565894037485
step: 130, loss: 0.02182496152818203
step: 140, loss: 0.0006049330113455653
step: 150, loss: 0.14374370872974396
step: 160, loss: 0.005187489092350006
step: 170, loss: 0.0036957997363060713
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 9: dev_f1=0.8578680203045685, f1=0.8697788697788698,
best_f1=0.8793969849246231
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.004121360369026661
step: 10, loss: 0.002728319726884365
step: 20, loss: 0.03940843045711517
step: 30, loss: 0.06238892674446106
step: 40, loss: 0.0007673298823647201
step: 50, loss: 0.0011597221018746495

```



```

step: 60, loss: 0.002184252953156829
step: 70, loss: 0.001716134836897254
step: 80, loss: 0.019616398960351944
step: 90, loss: 0.0003804256266448647
step: 100, loss: 0.003638056106865406
step: 110, loss: 0.009455070830881596
step: 120, loss: 0.0006204373785294592
step: 130, loss: 6.218723865458742e-05
step: 140, loss: 0.0011928018648177385
step: 150, loss: 0.00024704582756385207
step: 160, loss: 7.655336958123371e-05
step: 170, loss: 0.038799453526735306
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 10: dev_f1=0.8691099476439791, f1=0.8793969849246231,
best_f1=0.8793969849246231
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0657721534371376
step: 10, loss: 9.293990297010168e-05
step: 20, loss: 5.132916339789517e-05
step: 30, loss: 0.01136160921305418
step: 40, loss: 0.0005098354304209352
step: 50, loss: 0.0007965848781168461
step: 60, loss: 0.0002177968417527154
step: 70, loss: 0.0006850123172625899
step: 80, loss: 0.029587652534246445
step: 90, loss: 0.0005754366866312921
step: 100, loss: 0.00953845027834177
step: 110, loss: 0.00018786561849992722
step: 120, loss: 0.028009459376335144
step: 130, loss: 0.008446482941508293
step: 140, loss: 0.0003018675197381526
step: 150, loss: 0.00013786378258373588
step: 160, loss: 0.0001842210185714066
step: 170, loss: 0.0017329794354736805
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 11: dev_f1=0.8697788697788698, f1=0.8530805687203792,
best_f1=0.8793969849246231
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.011636652983725071
step: 10, loss: 0.0031627006828784943

```

```

step: 20, loss: 0.00036214786814525723
step: 30, loss: 0.04127237945795059
step: 40, loss: 0.00012877136759925634
step: 50, loss: 0.00012035838881274685
step: 60, loss: 0.0057648480869829655
step: 70, loss: 6.072482938179746e-05
step: 80, loss: 7.739387365290895e-05
step: 90, loss: 0.0019077147589996457
step: 100, loss: 0.00011030705354642123
step: 110, loss: 0.002265969291329384
step: 120, loss: 0.019506996497511864
step: 130, loss: 0.0030539329163730145
step: 140, loss: 0.00018959819863084704
step: 150, loss: 0.019766731187701225
step: 160, loss: 0.0011247028596699238
step: 170, loss: 0.00035351881524547935
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 12: dev_f1=0.8682170542635659, f1=0.8712871287128712,
best_f1=0.8793969849246231
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0001631212217034772
step: 10, loss: 8.771504508331418e-05
step: 20, loss: 0.0017275204882025719
step: 30, loss: 0.055646274238824844
step: 40, loss: 0.00014816998736932874
step: 50, loss: 0.0031082609202712774
step: 60, loss: 6.931878306204453e-05
step: 70, loss: 0.049939364194869995
step: 80, loss: 0.00010217855015071109
step: 90, loss: 0.0003239402431063354
step: 100, loss: 0.003920883871614933
step: 110, loss: 0.00018711989105213434
step: 120, loss: 0.009025434963405132
step: 130, loss: 8.414990588789806e-05
step: 140, loss: 0.0008127372711896896
step: 150, loss: 0.00021571392426267266
step: 160, loss: 8.310064731631428e-05
step: 170, loss: 0.00011838624050142244
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.

```

```

warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 13: dev_f1=0.8753180661577608, f1=0.8712871287128712,
best_f1=0.8712871287128712
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00486680818721652
step: 10, loss: 0.00011472195910755545
step: 20, loss: 0.004678803030401468
step: 30, loss: 0.0004215176741126925
step: 40, loss: 8.604295726399869e-05
step: 50, loss: 0.0008196118869818747
step: 60, loss: 0.0036548234056681395
step: 70, loss: 0.001083275070413947
step: 80, loss: 0.0006865970208309591
step: 90, loss: 0.00011215020640520379
step: 100, loss: 7.646088488399982e-05
step: 110, loss: 0.016130220144987106
step: 120, loss: 8.30555145512335e-05
step: 130, loss: 0.00029457840719260275
step: 140, loss: 0.045302197337150574
step: 150, loss: 0.012220476754009724
step: 160, loss: 0.00022515661839861423
step: 170, loss: 0.0374695360660553
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 14: dev_f1=0.8686868686868687, f1=0.8743961352657006,
best_f1=0.8712871287128712
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 2.3204533135867678e-05
step: 10, loss: 8.935861842473969e-05
step: 20, loss: 8.62750384840183e-05
step: 30, loss: 0.00016462066560052335
step: 40, loss: 5.658472582581453e-05
step: 50, loss: 6.106172077124938e-05
step: 60, loss: 0.005123381037265062
step: 70, loss: 0.06005507707595825
step: 80, loss: 6.791601481381804e-05
step: 90, loss: 0.0027973903343081474
step: 100, loss: 0.030124828219413757
step: 110, loss: 0.00011175779218319803
step: 120, loss: 6.604586087632924e-05
step: 130, loss: 0.0014990413328632712
step: 140, loss: 0.005014572758227587
step: 150, loss: 7.274348172359169e-05

```

```

step: 160, loss: 6.618408224312589e-05
step: 170, loss: 0.0002455657522659749
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 15: dev_f1=0.8708860759493672, f1=0.8807785888077859,
best_f1=0.8712871287128712
mv: cannot stat '*_dev.pt': No such file or directory
Selected optimization level 02: FP16 training with FP32 batchnorm and FP32
master weights.

Defaults for this optimization level are:
enabled                : True
opt_level              : 02
cast_model_type        : torch.float16
patch_torch_functions  : False
keep_batchnorm_fp32    : True
master_weights         : True
loss_scale             : dynamic
Processing user overrides (additional kwargs that are not None)...
After processing overrides, optimization options are:
enabled                : True
opt_level              : 02
cast_model_type        : torch.float16
patch_torch_functions  : False
keep_batchnorm_fp32    : True
master_weights         : True
loss_scale             : dynamic
Warning: multi_tensor_applier fused unscale kernel is unavailable, possibly
because apex was installed without --cuda_ext --cpp_ext. Using Python fallback.
Original ImportError was: ModuleNotFoundError("No module named 'amp_C'")
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
1916it [00:06, 305.33it/s]
load_f1 = 0.8711340206185566
real_f1 = 0.8753180661577608
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
4398it [00:22, 199.46it/s]

```

## 2.3 BASELINE DIRTY

### 2.3.1 DBLP-ACM - Running the matcher

```
[ ]: # Step 1:
# train the wdc model with the save_model flag on
# it will produce two files *_dev.pt and *_test.pt which are the
# best checkpoints on the dev set and the test set.
!CUDA_VISIBLE_DEVICES=0 python train_ditto.py \
--task Dirty/DBLP-ACM \
--batch_size 32 \
--max_len 256 \
--lr 3e-5 \
--n_epochs 15 \
--finetuning \
--lm roberta \
--fp16 \
--save_model

# Step 2: pick a checkpoint, rename it to wdc_all_small.pt
!mv *_dev.pt checkpoints/DBLP-ACM.pt

# Step 3: run the matcher
!CUDA_VISIBLE_DEVICES=0 python matcher.py \
--task Dirty/DBLP-ACM \
--input_path input/input_small.jsonl \
--output_path output/output_small.jsonl \
--lm roberta \
--use_gpu \
--fp16 \
--checkpoint_path checkpoints/
```

Selected optimization level 02: FP16 training with FP32 batchnorm and FP32 master weights.

Defaults for this optimization level are:

```
enabled           : True
opt_level         : 02
cast_model_type   : torch.float16
patch_torch_functions : False
keep_batchnorm_fp32 : True
master_weights    : True
loss_scale        : dynamic
```

Processing user overrides (additional kwargs that are not None)...

After processing overrides, optimization options are:

```
enabled           : True
opt_level         : 02
```

```

cast_model_type      : torch.float16
patch_torch_functions : False
keep_batchnorm_fp32  : True
master_weights       : True
loss_scale            : dynamic
Warning: multi_tensor_applier fused unscale kernel is unavailable, possibly
because apex was installed without --cuda_ext --cpp_ext. Using Python fallback.
Original ImportError was: ModuleNotFoundError("No module named 'amp_C'")
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.5872743725776672
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 32768.0
step: 10, loss: 0.5099325776100159
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 20, loss: 0.5698271989822388
step: 30, loss: 0.3333955407142639
step: 40, loss: 0.3278418481349945
step: 50, loss: 0.5487825870513916
step: 60, loss: 0.47559335827827454
step: 70, loss: 0.358742892742157
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 8192.0
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 4096.0
step: 80, loss: 0.3821769952774048
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 2048.0
step: 90, loss: 0.15119969844818115
step: 100, loss: 0.18866941332817078
step: 110, loss: 0.17000100016593933
step: 120, loss: 0.11574134230613708
step: 130, loss: 0.04324913024902344
step: 140, loss: 0.008990056812763214
step: 150, loss: 0.18256883323192596
step: 160, loss: 0.017279144376516342
step: 170, loss: 0.058843228965997696
step: 180, loss: 0.16927450895309448
step: 190, loss: 0.09010049700737
step: 200, loss: 0.009412254206836224
step: 210, loss: 0.04900772497057915
step: 220, loss: 0.08434779942035675
step: 230, loss: 0.012266837060451508
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
    warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 1: dev_f1=0.9669603524229076, f1=0.9711111111111111,

```

```

best_f1=0.9711111111111111
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.08190497010946274
step: 10, loss: 0.09623055905103683
step: 20, loss: 0.011278809979557991
step: 30, loss: 0.02263423427939415
step: 40, loss: 0.02527354657649994
step: 50, loss: 0.00700604310259223
step: 60, loss: 0.01644701324403286
step: 70, loss: 0.026431629434227943
step: 80, loss: 0.009111886844038963
step: 90, loss: 0.01356492843478918
step: 100, loss: 0.002717742696404457
step: 110, loss: 0.11990631371736526
step: 120, loss: 0.018122291192412376
step: 130, loss: 0.018551291897892952
step: 140, loss: 0.0029663576278835535
step: 150, loss: 0.05861876904964447
step: 160, loss: 0.013117908500134945
step: 170, loss: 0.0013656324008479714
step: 180, loss: 0.005801367107778788
step: 190, loss: 0.02068961225450039
step: 200, loss: 0.015887469053268433
step: 210, loss: 0.010985451750457287
step: 220, loss: 0.007253946736454964
step: 230, loss: 0.005255059339106083
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
    warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 2: dev_f1=0.9733924611973392, f1=0.9665924276169264,
best_f1=0.9665924276169264
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.004456210415810347
step: 10, loss: 0.011726051568984985
step: 20, loss: 0.0031221231911331415
step: 30, loss: 0.004595710895955563
step: 40, loss: 0.01486132387071848
step: 50, loss: 0.04629812389612198
step: 60, loss: 0.10483291745185852
step: 70, loss: 0.00686053279787302

```

```

step: 80, loss: 0.16760970652103424
step: 90, loss: 0.009201752953231335
step: 100, loss: 0.06735212355852127
step: 110, loss: 0.023744702339172363
step: 120, loss: 0.0033186948858201504
step: 130, loss: 0.00632140226662159
step: 140, loss: 0.0019473977154120803
step: 150, loss: 0.02843341790139675
step: 160, loss: 0.0038914289325475693
step: 170, loss: 0.0010325273033231497
step: 180, loss: 0.008029312826693058
step: 190, loss: 0.013186142779886723
step: 200, loss: 0.006613117642700672
step: 210, loss: 0.02090434916317463
step: 220, loss: 0.1794889271259308
step: 230, loss: 0.01477225311100483
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 3: dev_f1=0.9692982456140351, f1=0.9670329670329669,
best_f1=0.9665924276169264
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.012749374844133854
step: 10, loss: 0.0012830038322135806
step: 20, loss: 0.001513621537014842
step: 30, loss: 0.002902439795434475
step: 40, loss: 0.06166795268654823
step: 50, loss: 0.10093651711940765
step: 60, loss: 0.038992613554000854
step: 70, loss: 0.025580406188964844
step: 80, loss: 0.11511841416358948
step: 90, loss: 0.017714980989694595
step: 100, loss: 0.12234054505825043
step: 110, loss: 0.001223209546878934
step: 120, loss: 0.007742349989712238
step: 130, loss: 0.007899426855146885
step: 140, loss: 0.005041440483182669
step: 150, loss: 0.02750539407134056
step: 160, loss: 0.009625103324651718
step: 170, loss: 0.0011996098328381777
step: 180, loss: 0.10772860050201416
step: 190, loss: 0.0021057825069874525
step: 200, loss: 0.1364026665687561
step: 210, loss: 0.027275823056697845
step: 220, loss: 0.004555503372102976
step: 230, loss: 0.07931166142225266

```



```

/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 4: dev_f1=0.9713024282560706, f1=0.9766407119021134,
best_f1=0.9665924276169264
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0005549120833165944
step: 10, loss: 0.02246227115392685
step: 20, loss: 0.15465983748435974
step: 30, loss: 0.021373676136136055
step: 40, loss: 0.0031964057125151157
step: 50, loss: 0.006231122184544802
step: 60, loss: 0.013395335525274277
step: 70, loss: 0.013094251044094563
step: 80, loss: 0.026316635310649872
step: 90, loss: 0.10045596212148666
step: 100, loss: 0.003657999448478222
step: 110, loss: 0.018498437479138374
step: 120, loss: 0.2541341185569763
step: 130, loss: 0.015413300134241581
step: 140, loss: 0.006099367514252663
step: 150, loss: 0.04638609662652016
step: 160, loss: 0.0022530648857355118
step: 170, loss: 0.014248036779463291
step: 180, loss: 0.0061170002445578575
step: 190, loss: 0.012248476035892963
step: 200, loss: 0.11826805025339127
step: 210, loss: 0.0034074706491082907
step: 220, loss: 0.009114273823797703
step: 230, loss: 0.037757180631160736
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 5: dev_f1=0.9680968096809681, f1=0.9680968096809681,
best_f1=0.9665924276169264
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0016478396719321609
step: 10, loss: 0.0013436840381473303
step: 20, loss: 0.006768169347196817
step: 30, loss: 0.0009865175234153867
step: 40, loss: 0.000335660413838923
step: 50, loss: 0.0003493119729682803
step: 60, loss: 0.005831732414662838
step: 70, loss: 0.19680801033973694

```

```

step: 80, loss: 0.008668278343975544
step: 90, loss: 0.00612218352034688
step: 100, loss: 0.0010105941910296679
step: 110, loss: 0.010572398081421852
step: 120, loss: 0.0014363054651767015
step: 130, loss: 0.0042663016356527805
step: 140, loss: 0.000750904728192836
step: 150, loss: 0.0001482705702073872
step: 160, loss: 0.0010386533103883266
step: 170, loss: 0.0006328575545921922
step: 180, loss: 0.0004135183698963374
step: 190, loss: 0.002088485285639763
step: 200, loss: 0.005350726190954447
step: 210, loss: 0.0041990638710558414
step: 220, loss: 0.009506221860647202
step: 230, loss: 0.0025797963608056307
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
    warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 6: dev_f1=0.9767441860465117, f1=0.9713024282560706,
best_f1=0.9713024282560706
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.003012506291270256
step: 10, loss: 0.0028817574493587017
step: 20, loss: 0.0004903033259324729
step: 30, loss: 0.00018720701336860657
step: 40, loss: 0.0024762817192822695
step: 50, loss: 0.0026849531568586826
step: 60, loss: 0.0005445978604257107
step: 70, loss: 0.0003895171685144305
step: 80, loss: 0.0031676036305725574
step: 90, loss: 0.00040810287464410067
step: 100, loss: 0.00022340107534546405
step: 110, loss: 0.0002571202057879418
step: 120, loss: 0.022956877946853638
step: 130, loss: 0.0009366738959215581
step: 140, loss: 0.00046342163113877177
step: 150, loss: 0.031061820685863495
step: 160, loss: 0.0022063502110540867
step: 170, loss: 0.00739000178873539
step: 180, loss: 0.0028030958492308855
step: 190, loss: 0.00036674647708423436

```

```

step: 200, loss: 0.008594865910708904
step: 210, loss: 0.028604650869965553
step: 220, loss: 0.00022666914446745068
step: 230, loss: 0.0020050692837685347
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
    warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 7: dev_f1=0.9799107142857142, f1=0.9744160177975528,
best_f1=0.9744160177975528
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0023532414343208075
step: 10, loss: 0.005414221901446581
step: 20, loss: 0.002429486019536853
step: 30, loss: 0.006354004144668579
step: 40, loss: 0.01149299181997776
step: 50, loss: 0.0006800394621677697
step: 60, loss: 0.0010718264384195209
step: 70, loss: 0.0006248150020837784
step: 80, loss: 0.046233881264925
step: 90, loss: 0.0003906053025275469
step: 100, loss: 0.0003735637874342501
step: 110, loss: 0.0008772140718065202
step: 120, loss: 0.001149091636762023
step: 130, loss: 0.0006879912689328194
step: 140, loss: 0.00014558889961335808
step: 150, loss: 0.0659605860710144
step: 160, loss: 0.00031799040152691305
step: 170, loss: 0.0030157985165715218
step: 180, loss: 0.00021875901438761503
step: 190, loss: 0.0010014231083914638
step: 200, loss: 0.012520035728812218
step: 210, loss: 0.0012683868408203125
step: 220, loss: 0.0022459409665316343
step: 230, loss: 0.005206719972193241
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
    warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 8: dev_f1=0.9809203142536477, f1=0.9754464285714286,

```

```

best_f1=0.9754464285714286
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0030692776199430227
step: 10, loss: 0.0007177707739174366
step: 20, loss: 0.0032199870329350233
step: 30, loss: 0.00033886160235852003
step: 40, loss: 0.0037637881468981504
step: 50, loss: 0.000860509870108217
step: 60, loss: 0.0002857657673303038
step: 70, loss: 0.03458059951663017
step: 80, loss: 0.003535535419359803
step: 90, loss: 0.052968963980674744
step: 100, loss: 0.0009989205282181501
step: 110, loss: 0.00011491012264741585
step: 120, loss: 0.046407587826251984
step: 130, loss: 0.00015081708261277527
step: 140, loss: 0.0004426319501362741
step: 150, loss: 0.00018281287339050323
step: 160, loss: 0.008480140939354897
step: 170, loss: 7.302960875676945e-05
step: 180, loss: 0.0017837664345279336
step: 190, loss: 7.148984877858311e-05
step: 200, loss: 0.0006979302852414548
step: 210, loss: 0.01713605783879757
step: 220, loss: 0.0008771762368269265
step: 230, loss: 0.0001020865238388069
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
    warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 9: dev_f1=0.9821428571428571, f1=0.9733333333333333,
best_f1=0.9733333333333333
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00041924204560928047
step: 10, loss: 0.00014999715494923294
step: 20, loss: 0.00010493014997337013
step: 30, loss: 5.385587428463623e-05
step: 40, loss: 0.0002587079943623394
step: 50, loss: 7.849165558582172e-05
step: 60, loss: 0.00021914715762250125
step: 70, loss: 0.002315536607056856

```

```

step: 80, loss: 0.0001244825398316607
step: 90, loss: 0.00018189576803706586
step: 100, loss: 0.000102177684311755
step: 110, loss: 0.00012261066876817495
step: 120, loss: 0.0012053338577970862
step: 130, loss: 0.000149999643326737
step: 140, loss: 0.0001378638989990577
step: 150, loss: 0.0005065057775937021
step: 160, loss: 3.971529076807201e-05
step: 170, loss: 7.158138760132715e-05
step: 180, loss: 0.0005119054112583399
step: 190, loss: 0.00021853583166375756
step: 200, loss: 0.00013973725435789675
step: 210, loss: 0.008972009643912315
step: 220, loss: 0.00017964158905670047
step: 230, loss: 0.0001275061658816412
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
    warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 10: dev_f1=0.9832402234636871, f1=0.975609756097561,
best_f1=0.975609756097561
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 6.984319043112919e-05
step: 10, loss: 0.00014987138274591416
step: 20, loss: 0.00010475190356373787
step: 30, loss: 0.00011262301995884627
step: 40, loss: 6.00942294113338e-05
step: 50, loss: 6.962355109862983e-05
step: 60, loss: 0.013090379536151886
step: 70, loss: 0.00035235032555647194
step: 80, loss: 0.0016907422104850411
step: 90, loss: 0.09345663338899612
step: 100, loss: 0.00011274939606664702
step: 110, loss: 0.0006699131336063147
step: 120, loss: 7.918224582681432e-05
step: 130, loss: 8.186361083062366e-05
step: 140, loss: 0.0025599945802241564
step: 150, loss: 9.669272549217567e-05
step: 160, loss: 0.03528353571891785
step: 170, loss: 6.808734178775921e-05
step: 180, loss: 4.3736454244935885e-05
step: 190, loss: 4.1464274545433e-05

```

```

step: 200, loss: 0.0016232857014983892
step: 210, loss: 9.38895609579049e-05
step: 220, loss: 0.0005041266558691859
step: 230, loss: 0.00025230730534531176
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 11: dev_f1=0.9787709497206705, f1=0.9735099337748344,
best_f1=0.975609756097561
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00575395580381155
step: 10, loss: 6.441452569561079e-05
step: 20, loss: 0.015012871474027634
step: 30, loss: 0.25620555877685547
step: 40, loss: 0.00010428769019199535
step: 50, loss: 0.0009304180275648832
step: 60, loss: 0.0005264138453640044
step: 70, loss: 6.0881739045726135e-05
step: 80, loss: 4.226497185300104e-05
step: 90, loss: 0.001745456364005804
step: 100, loss: 5.000248711439781e-05
step: 110, loss: 5.1946393796242774e-05
step: 120, loss: 0.00010009602556237951
step: 130, loss: 9.11293609533459e-05
step: 140, loss: 0.00023626096663065255
step: 150, loss: 5.942175630480051e-05
step: 160, loss: 0.00028694604407064617
step: 170, loss: 8.473070920445025e-05
step: 180, loss: 0.00011118405382148921
step: 190, loss: 0.0002865852147806436
step: 200, loss: 4.998469739803113e-05
step: 210, loss: 8.215790148824453e-05
step: 220, loss: 0.009935164824128151
step: 230, loss: 4.7211520723067224e-05
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 12: dev_f1=0.9788182831661093, f1=0.9756637168141594,
best_f1=0.975609756097561
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 9.922556637320668e-05
step: 10, loss: 7.13850895408541e-05
step: 20, loss: 5.8106677897740155e-05
step: 30, loss: 4.08399646403268e-05

```

```

step: 40, loss: 0.03771453723311424
step: 50, loss: 0.0003565285005606711
step: 60, loss: 6.058235157979652e-05
step: 70, loss: 0.00012416655954439193
step: 80, loss: 4.858163447352126e-05
step: 90, loss: 6.239418871700764e-05
step: 100, loss: 7.208463648566976e-05
step: 110, loss: 8.145451283780858e-05
step: 120, loss: 6.306348950602114e-05
step: 130, loss: 5.616445923806168e-05
step: 140, loss: 5.8394547522766516e-05
step: 150, loss: 5.902126576984301e-05
step: 160, loss: 0.009890657849609852
step: 170, loss: 4.3051350075984374e-05
step: 180, loss: 0.025973524898290634
step: 190, loss: 7.821179315214977e-05
step: 200, loss: 2.5547362383804284e-05
step: 210, loss: 9.857045370154083e-05
step: 220, loss: 3.432367884670384e-05
step: 230, loss: 4.964060281054117e-05
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 13: dev_f1=0.9776286353467561, f1=0.9743016759776536,
best_f1=0.975609756097561
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 3.1540970667265356e-05
step: 10, loss: 3.220404687453993e-05
step: 20, loss: 9.770039469003677e-05
step: 30, loss: 9.339842654298991e-05
step: 40, loss: 3.9503054722445086e-05
step: 50, loss: 4.4920168875250965e-05
step: 60, loss: 3.14030512527097e-05
step: 70, loss: 3.508004010654986e-05
step: 80, loss: 2.569624484749511e-05
step: 90, loss: 6.513250264106318e-05
step: 100, loss: 3.7928984966129065e-05
step: 110, loss: 7.036125316517428e-05
step: 120, loss: 1.9099303244729526e-05
step: 130, loss: 5.226644498179667e-05
step: 140, loss: 3.733626726898365e-05
step: 150, loss: 6.339232641039416e-05
step: 160, loss: 4.666273162001744e-05
step: 170, loss: 6.082730760681443e-05
step: 180, loss: 3.5846827813657e-05
step: 190, loss: 3.651350561995059e-05

```

```

step: 200, loss: 6.0359048802638426e-05
step: 210, loss: 5.273452552501112e-05
step: 220, loss: 4.398195233079605e-05
step: 230, loss: 4.425595761858858e-05
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 14: dev_f1=0.9799107142857142, f1=0.9755011135857461,
best_f1=0.975609756097561
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.003384074429050088
step: 10, loss: 4.196301961201243e-05
step: 20, loss: 0.00019480941409710795
step: 30, loss: 4.58814793091733e-05
step: 40, loss: 2.3394313757307827e-05
step: 50, loss: 2.8132442821515724e-05
step: 60, loss: 0.0340624563395977
step: 70, loss: 4.0484199416823685e-05
step: 80, loss: 0.00011495107173686847
step: 90, loss: 5.627320570056327e-05
step: 100, loss: 2.380411751801148e-05
step: 110, loss: 4.283887756173499e-05
step: 120, loss: 0.03533351421356201
step: 130, loss: 3.757154627237469e-05
step: 140, loss: 0.0036675776354968548
step: 150, loss: 4.1896328184520826e-05
step: 160, loss: 0.00836010742932558
step: 170, loss: 2.1296942577464506e-05
step: 180, loss: 4.53410430054646e-05
step: 190, loss: 0.0002432349865557626
step: 200, loss: 0.00041384450742043555
step: 210, loss: 0.021361498162150383
step: 220, loss: 3.152216231683269e-05
step: 230, loss: 4.5091841457178816e-05
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 15: dev_f1=0.9799107142857142, f1=0.9755011135857461,
best_f1=0.975609756097561
mv: cannot stat '*_dev.pt': No such file or directory
Selected optimization level 02:  FP16 training with FP32 batchnorm and FP32
master weights.

```

Defaults for this optimization level are:

```

enabled          : True
opt_level        : 02

```



```

cast_model_type      : torch.float16
patch_torch_functions : False
keep_batchnorm_fp32  : True
master_weights       : True
loss_scale            : dynamic
Processing user overrides (additional kwargs that are not None)...
After processing overrides, optimization options are:
enabled              : True
opt_level            : 02
cast_model_type      : torch.float16
patch_torch_functions : False
keep_batchnorm_fp32  : True
master_weights       : True
loss_scale            : dynamic
Warning: multi_tensor_applier fused unscale kernel is unavailable, possibly
because apex was installed without --cuda_ext --cpp_ext. Using Python fallback.
Original ImportError was: ModuleNotFoundError("No module named 'amp_C'")
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
2473it [00:12, 195.00it/s]
load_f1 = 0.9810055865921787
real_f1 = 0.9832402234636871
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
4398it [00:22, 197.74it/s]

```

### 2.3.2 DBLP-GoogleScholar - Running the matcher

```

[ ]: # Step 1:
# train the wdc model with the save_model flag on
# it will produce two files *_dev.pt and *_test.pt which are the
# best checkpoints on the dev set and the test set.
!CUDA_VISIBLE_DEVICES=0 python train_ditto.py \
--task Dirty/DBLP-GoogleScholar \
--batch_size 32 \
--max_len 256 \
--lr 3e-5 \
--n_epochs 15 \
--finetuning \
--lm roberta \
--fp16 \

```

```

--save_model

# Step 2: pick a checkpoint, rename it to wdc_all_small.pt
!mv *_dev.pt checkpoints/DBLP-GoogleScholar/.pt

# Step 3: run the matcher
!CUDA_VISIBLE_DEVICES=0 python matcher.py \
  --task Dirty/DBLP-GoogleScholar \
  --input_path input/input_small.jsonl \
  --output_path output/output_small.jsonl \
  --lm roberta \
  --use_gpu \
  --fp16 \
  --checkpoint_path checkpoints/

```

Selected optimization level O2: FP16 training with FP32 batchnorm and FP32 master weights.

Defaults for this optimization level are:

```

enabled          : True
opt_level        : O2
cast_model_type  : torch.float16
patch_torch_functions : False
keep_batchnorm_fp32 : True
master_weights   : True
loss_scale       : dynamic

```

Processing user overrides (additional kwargs that are not None)...

After processing overrides, optimization options are:

```

enabled          : True
opt_level        : O2
cast_model_type  : torch.float16
patch_torch_functions : False
keep_batchnorm_fp32 : True
master_weights   : True
loss_scale       : dynamic

```

Warning: multi\_tensor\_applier fused unscale kernel is unavailable, possibly because apex was installed without --cuda\_ext --cpp\_ext. Using Python fallback. Original ImportError was: ModuleNotFoundError("No module named 'amp\_C'")  
/usr/local/lib/python3.7/dist-packages/apex/amp/\_initialize.py:25: UserWarning: An input tensor was not cuda.

```
warnings.warn("An input tensor was not cuda.")
```

```
step: 0, loss: 0.632512092590332
```

Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 32768.0

Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0

```
step: 10, loss: 0.47821590304374695
```

```
step: 20, loss: 0.304843008518219
```

```
step: 30, loss: 0.3609592914581299
```

Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 8192.0

step: 40, loss: 0.36418187618255615  
step: 50, loss: 0.33218616247177124  
step: 60, loss: 0.2623429000377655  
step: 70, loss: 0.14370469748973846  
step: 80, loss: 0.12101243436336517  
step: 90, loss: 0.15666015446186066  
step: 100, loss: 0.29152047634124756  
step: 110, loss: 0.09635920822620392  
step: 120, loss: 0.26928386092185974  
step: 130, loss: 0.0640420913696289  
step: 140, loss: 0.20983819663524628  
step: 150, loss: 0.06331848353147507  
step: 160, loss: 0.056574009358882904  
step: 170, loss: 0.06629613041877747  
step: 180, loss: 0.13640345633029938  
step: 190, loss: 0.11747559905052185  
step: 200, loss: 0.03139985725283623  
step: 210, loss: 0.017232418060302734  
step: 220, loss: 0.016418328508734703  
step: 230, loss: 0.09447745233774185  
step: 240, loss: 0.02553003653883934  
step: 250, loss: 0.050870418548583984  
step: 260, loss: 0.06201979145407677  
step: 270, loss: 0.34433719515800476  
step: 280, loss: 0.016718856990337372  
step: 290, loss: 0.06906534731388092  
step: 300, loss: 0.008922220207750797  
step: 310, loss: 0.04603397846221924  
step: 320, loss: 0.03167346864938736  
step: 330, loss: 0.03282228857278824  
step: 340, loss: 0.3836583197116852  
step: 350, loss: 0.09381546080112457  
step: 360, loss: 0.09082112461328506  
step: 370, loss: 0.11535491049289703  
step: 380, loss: 0.08492658287286758  
step: 390, loss: 0.010826357640326023  
step: 400, loss: 0.027312278747558594  
step: 410, loss: 0.25676026940345764  
step: 420, loss: 0.04013616219162941  
step: 430, loss: 0.007868318818509579  
step: 440, loss: 0.011532526463270187  
step: 450, loss: 0.010206212289631367  
step: 460, loss: 0.010381487198174  
step: 470, loss: 0.02118496038019657  
step: 480, loss: 0.15577441453933716  
step: 490, loss: 0.04967203363776207  
step: 500, loss: 0.004924983251839876

```

step: 510, loss: 0.021423889324069023
step: 520, loss: 0.04278968647122383
step: 530, loss: 0.021924154832959175
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
    warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 1: dev_f1=0.9439472444653791, f1=0.9429373246024322,
best_f1=0.9429373246024322
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.06591565907001495
step: 10, loss: 0.07094181329011917
step: 20, loss: 0.015856966376304626
step: 30, loss: 0.020467352122068405
step: 40, loss: 0.06688369065523148
step: 50, loss: 0.04875821992754936
step: 60, loss: 0.013731512241065502
step: 70, loss: 0.011911310255527496
step: 80, loss: 0.02485044114291668
step: 90, loss: 0.007314532063901424
step: 100, loss: 0.2067243754863739
step: 110, loss: 0.015742730349302292
step: 120, loss: 0.03729936107993126
step: 130, loss: 0.0025472829584032297
step: 140, loss: 0.09603863954544067
step: 150, loss: 0.027865095064044
step: 160, loss: 0.027191918343305588
step: 170, loss: 0.03636207804083824
step: 180, loss: 0.06649821251630783
step: 190, loss: 0.006924291141331196
step: 200, loss: 0.030588584020733833
step: 210, loss: 0.038898661732673645
step: 220, loss: 0.00054267788073048
step: 230, loss: 0.09482748806476593
step: 240, loss: 0.09100601822137833
step: 250, loss: 0.019755665212869644
step: 260, loss: 0.04143468663096428
step: 270, loss: 0.0013514079619199038
step: 280, loss: 0.023606695234775543
step: 290, loss: 0.05975749343633652
step: 300, loss: 0.028055042028427124
step: 310, loss: 0.05180191993713379
step: 320, loss: 0.10643380880355835

```

```

step: 330, loss: 0.042721111327409744
step: 340, loss: 0.2292410284280777
step: 350, loss: 0.006965332664549351
step: 360, loss: 0.20815065503120422
step: 370, loss: 0.006794458720833063
step: 380, loss: 0.10850764811038971
step: 390, loss: 0.010973509401082993
step: 400, loss: 0.05493314936757088
step: 410, loss: 0.026054592803120613
step: 420, loss: 0.029837869107723236
step: 430, loss: 0.2880135476589203
step: 440, loss: 0.021501822397112846
step: 450, loss: 0.02841144986450672
step: 460, loss: 0.02655569463968277
step: 470, loss: 0.014688894152641296
step: 480, loss: 0.008208667859435081
step: 490, loss: 0.03872936591506004
step: 500, loss: 0.001687282812781632
step: 510, loss: 0.0189681239426136
step: 520, loss: 0.16332027316093445
step: 530, loss: 0.062347233295440674
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
    warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 2: dev_f1=0.9552376557452701, f1=0.9478098788443615,
best_f1=0.9478098788443615
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.10485652089118958
step: 10, loss: 0.017095983028411865
step: 20, loss: 0.004983112681657076
step: 30, loss: 0.03194184973835945
step: 40, loss: 0.02299218624830246
step: 50, loss: 0.004527238663285971
step: 60, loss: 0.005435890052467585
step: 70, loss: 0.0015555620193481445
step: 80, loss: 0.1487845778465271
step: 90, loss: 0.0015093495603650808
step: 100, loss: 0.01719539240002632
step: 110, loss: 0.0025418070144951344
step: 120, loss: 0.06764630228281021
step: 130, loss: 0.027822114527225494
step: 140, loss: 0.003369836835190654

```

```

step: 150, loss: 0.005017888266593218
step: 160, loss: 0.028152313083410263
step: 170, loss: 0.0011873142793774605
step: 180, loss: 0.004753880202770233
step: 190, loss: 0.001451240386813879
step: 200, loss: 0.018305392935872078
step: 210, loss: 0.012409931980073452
step: 220, loss: 0.047566067427396774
step: 230, loss: 0.05098574608564377
step: 240, loss: 0.03397158905863762
step: 250, loss: 0.023309782147407532
step: 260, loss: 0.08176667988300323
step: 270, loss: 0.0064054690301418304
step: 280, loss: 0.01678161509335041
step: 290, loss: 0.04181935265660286
step: 300, loss: 0.07643892616033554
step: 310, loss: 0.04522094503045082
step: 320, loss: 0.02928355149924755
step: 330, loss: 0.0013938812771812081
step: 340, loss: 0.0027245201636105776
step: 350, loss: 0.2163325995206833
step: 360, loss: 0.02385341376066208
step: 370, loss: 0.02569674700498581
step: 380, loss: 0.003001714823767543
step: 390, loss: 0.0037700561806559563
step: 400, loss: 0.19481879472732544
step: 410, loss: 0.05540171638131142
step: 420, loss: 0.035472750663757324
step: 430, loss: 0.004011963959783316
step: 440, loss: 0.19151701033115387
step: 450, loss: 0.05466032773256302
step: 460, loss: 0.08969353884458542
step: 470, loss: 0.20771528780460358
step: 480, loss: 0.18597561120986938
step: 490, loss: 0.035098060965538025
step: 500, loss: 0.0246339850127697
step: 510, loss: 0.018758652731776237
step: 520, loss: 0.0009074449772015214
step: 530, loss: 0.006330975331366062
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
epoch 3: dev_f1=0.9487058823529413, f1=0.9423984891406987,
best_f1=0.9478098788443615
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.008456002920866013

```

step: 10, loss: 0.011423856019973755  
step: 20, loss: 0.0014469909947365522  
step: 30, loss: 0.10133658349514008  
step: 40, loss: 0.01081536803394556  
step: 50, loss: 0.01656797155737877  
step: 60, loss: 0.000687698251567781  
step: 70, loss: 0.08652425557374954  
step: 80, loss: 0.06972683221101761  
step: 90, loss: 0.010743942111730576  
step: 100, loss: 0.00633673882111907  
step: 110, loss: 0.08038035780191422  
step: 120, loss: 0.0015805438160896301  
step: 130, loss: 0.004723512567579746  
step: 140, loss: 0.030254468321800232  
step: 150, loss: 0.002057112054899335  
step: 160, loss: 0.002711030188947916  
step: 170, loss: 0.009193957783281803  
step: 180, loss: 0.03950412571430206  
step: 190, loss: 0.16851449012756348  
step: 200, loss: 0.037299420684576035  
step: 210, loss: 0.00034835320548154414  
step: 220, loss: 0.003434260142967105  
step: 230, loss: 0.004431311972439289  
step: 240, loss: 0.040508344769477844  
step: 250, loss: 0.039444636553525925  
step: 260, loss: 0.000605270906817168  
step: 270, loss: 0.29814982414245605  
step: 280, loss: 0.0066528585739433765  
step: 290, loss: 0.028124578297138214  
step: 300, loss: 0.0014104912988841534  
step: 310, loss: 0.010355031117796898  
step: 320, loss: 0.02028394676744938  
step: 330, loss: 0.003247372806072235  
step: 340, loss: 0.00154962670058012  
step: 350, loss: 0.12649837136268616  
step: 360, loss: 0.0358465202152729  
step: 370, loss: 0.0018853778019547462  
step: 380, loss: 0.004588227719068527  
step: 390, loss: 0.00038268754724413157  
step: 400, loss: 0.04441848397254944  
step: 410, loss: 0.002364774001762271  
step: 420, loss: 0.03793562203645706  
step: 430, loss: 0.0002459692186675966  
step: 440, loss: 0.002529296325519681  
step: 450, loss: 0.011083213612437248  
step: 460, loss: 0.0026391709689050913  
step: 470, loss: 0.019930509850382805  
step: 480, loss: 0.005900876596570015

```

step: 490, loss: 0.0003232211456634104
step: 500, loss: 0.031227953732013702
step: 510, loss: 0.005619008559733629
step: 520, loss: 0.013773844577372074
step: 530, loss: 0.16811878979206085
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
  warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 4: dev_f1=0.958256029684601, f1=0.9564007421150279,
best_f1=0.9564007421150279
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0006190515123307705
step: 10, loss: 0.0023009334690868855
step: 20, loss: 0.0031030632089823484
step: 30, loss: 0.0042786612175405025
step: 40, loss: 0.0008334664162248373
step: 50, loss: 0.07506918907165527
step: 60, loss: 0.021941229701042175
step: 70, loss: 0.0009464834001846611
step: 80, loss: 0.0008461130782961845
step: 90, loss: 0.004664721433073282
step: 100, loss: 0.14591951668262482
step: 110, loss: 0.0008372287265956402
step: 120, loss: 0.11373180150985718
step: 130, loss: 0.006288468837738037
step: 140, loss: 0.0013856911100447178
step: 150, loss: 0.011184370145201683
step: 160, loss: 0.0036252494901418686
step: 170, loss: 0.02672404795885086
step: 180, loss: 0.0016580534866079688
step: 190, loss: 0.049954090267419815
step: 200, loss: 0.019401300698518753
step: 210, loss: 0.0011427103308960795
step: 220, loss: 0.02552901580929756
step: 230, loss: 0.002571725519374013
step: 240, loss: 0.008642599917948246
step: 250, loss: 0.20869769155979156
step: 260, loss: 0.0001978947111638263
step: 270, loss: 0.05656758323311806
step: 280, loss: 0.004467048682272434
step: 290, loss: 0.0018445326713845134
step: 300, loss: 0.09038788825273514

```



```

step: 310, loss: 0.023155517876148224
step: 320, loss: 0.007976425811648369
step: 330, loss: 0.0004886194365099072
step: 340, loss: 0.014545843936502934
step: 350, loss: 0.0008337053586728871
step: 360, loss: 0.00032043003011494875
step: 370, loss: 0.0002420971868559718
step: 380, loss: 0.00038379186298698187
step: 390, loss: 0.00229543331079185
step: 400, loss: 0.000618137011770159
step: 410, loss: 0.1999557614326477
step: 420, loss: 0.19274446368217468
step: 430, loss: 0.0028213171754032373
step: 440, loss: 0.0002865991264116019
step: 450, loss: 0.008143395185470581
step: 460, loss: 0.0039070285856723785
step: 470, loss: 0.02054218202829361
step: 480, loss: 0.032902467995882034
step: 490, loss: 0.009657539427280426
step: 500, loss: 0.039112478494644165
step: 510, loss: 0.008642693050205708
step: 520, loss: 0.09253557026386261
step: 530, loss: 0.09567809849977493
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 5: dev_f1=0.9564814814814815, f1=0.9458583988894032,
best_f1=0.9564007421150279
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.010741626843810081
step: 10, loss: 0.004043180029839277
step: 20, loss: 0.0034272705670446157
step: 30, loss: 0.0007838328601792455
step: 40, loss: 0.00010454530274728313
step: 50, loss: 0.00032526100403629243
step: 60, loss: 0.00011009679292328656
step: 70, loss: 0.016033874824643135
step: 80, loss: 0.00012742636317852885
step: 90, loss: 0.0004270304343663156
step: 100, loss: 0.0013713828520849347
step: 110, loss: 0.0008492504130117595
step: 120, loss: 0.004451966378837824
step: 130, loss: 0.0015223643276840448
step: 140, loss: 0.00046393671073019505
step: 150, loss: 6.747720908606425e-05
step: 160, loss: 0.0007292858208529651

```

```

step: 170, loss: 0.00014019667287357152
step: 180, loss: 7.098536298144609e-05
step: 190, loss: 0.024850767105817795
step: 200, loss: 0.003728751791641116
step: 210, loss: 0.0022248493041843176
step: 220, loss: 0.006846742704510689
step: 230, loss: 0.0036058537662029266
step: 240, loss: 0.00017126384773291647
step: 250, loss: 0.0012060317676514387
step: 260, loss: 0.0006011008517816663
step: 270, loss: 9.044289618032053e-05
step: 280, loss: 0.0001284545287489891
step: 290, loss: 0.2534312605857849
step: 300, loss: 0.004618779290467501
step: 310, loss: 0.02861066907644272
step: 320, loss: 0.0002441690012346953
step: 330, loss: 0.11914689838886261
step: 340, loss: 0.0013545163674280047
step: 350, loss: 0.021685998886823654
step: 360, loss: 0.08975614607334137
step: 370, loss: 0.009727894328534603
step: 380, loss: 0.0006245304248295724
step: 390, loss: 0.0011725805234164
step: 400, loss: 0.013044312596321106
step: 410, loss: 0.00010143185500055552
step: 420, loss: 0.015490359626710415
step: 430, loss: 0.00017545332957524806
step: 440, loss: 0.00031488039530813694
step: 450, loss: 0.2291971892118454
step: 460, loss: 0.0007767652859911323
step: 470, loss: 0.0036457334645092487
step: 480, loss: 0.004410532768815756
step: 490, loss: 0.005588975735008717
step: 500, loss: 0.023951327428221703
step: 510, loss: 0.1313994973897934
step: 520, loss: 0.0006693225004710257
step: 530, loss: 0.0003755016077775508
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 6: dev_f1=0.958256029684601, f1=0.9503940658321743,
best_f1=0.9564007421150279
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.002006492344662547
step: 10, loss: 0.003388255601748824
step: 20, loss: 0.012780279852449894

```

step: 30, loss: 0.047378119081258774  
step: 40, loss: 0.0008677271544001997  
step: 50, loss: 0.011445249430835247  
step: 60, loss: 0.0028368602506816387  
step: 70, loss: 0.0015715609770268202  
step: 80, loss: 0.0021742552053183317  
step: 90, loss: 0.00015309602895285934  
step: 100, loss: 0.0018180697225034237  
step: 110, loss: 3.9787846617400646e-05  
step: 120, loss: 0.0001406555384164676  
step: 130, loss: 3.0381837859749794e-05  
step: 140, loss: 0.014050453901290894  
step: 150, loss: 0.00020854281319770962  
step: 160, loss: 6.0815978940809146e-05  
step: 170, loss: 0.007417713291943073  
step: 180, loss: 0.00018127942166756839  
step: 190, loss: 0.00019401581084821373  
step: 200, loss: 0.0002955982054118067  
step: 210, loss: 0.010456673800945282  
step: 220, loss: 0.006952312309294939  
step: 230, loss: 5.9846068324986845e-05  
step: 240, loss: 0.04870786517858505  
step: 250, loss: 0.041424959897994995  
step: 260, loss: 0.016238437965512276  
step: 270, loss: 0.0002946761669591069  
step: 280, loss: 0.0051855649799108505  
step: 290, loss: 0.0004714527167379856  
step: 300, loss: 0.0004646709421649575  
step: 310, loss: 0.0009846030734479427  
step: 320, loss: 0.00010544768156250939  
step: 330, loss: 0.00045304125524125993  
step: 340, loss: 0.011408772319555283  
step: 350, loss: 5.517418685485609e-05  
step: 360, loss: 0.061657682061195374  
step: 370, loss: 0.005986471194773912  
step: 380, loss: 0.0074192616157233715  
step: 390, loss: 0.010111789219081402  
step: 400, loss: 0.010859573259949684  
step: 410, loss: 0.006148964166641235  
step: 420, loss: 0.18727244436740875  
step: 430, loss: 0.0007879493059590459  
step: 440, loss: 0.002568078227341175  
step: 450, loss: 0.0011938181705772877  
step: 460, loss: 0.03558376431465149  
step: 470, loss: 0.1564113348722458  
step: 480, loss: 0.008335615508258343  
step: 490, loss: 0.005148235242813826  
step: 500, loss: 0.006472214125096798

```

step: 510, loss: 0.0008604726172052324
step: 520, loss: 0.003224159823730588
step: 530, loss: 0.001472368952818215
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
    warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 7: dev_f1=0.9592969472710452, f1=0.9564007421150279,
best_f1=0.9564007421150279
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0005802921368740499
step: 10, loss: 0.0018955037230625749
step: 20, loss: 0.00016504304949194193
step: 30, loss: 3.816025491687469e-05
step: 40, loss: 2.1852047211723402e-05
step: 50, loss: 0.014569539576768875
step: 60, loss: 3.7407022318802774e-05
step: 70, loss: 0.00010688340989872813
step: 80, loss: 0.00043652390013448894
step: 90, loss: 3.191695213899948e-05
step: 100, loss: 6.544209463754669e-05
step: 110, loss: 0.00040021390304900706
step: 120, loss: 0.002832359168678522
step: 130, loss: 0.028078023344278336
step: 140, loss: 4.2857660446316004e-05
step: 150, loss: 0.00010001069313148037
step: 160, loss: 2.1792360712424852e-05
step: 170, loss: 0.0026685744524002075
step: 180, loss: 6.465931073762476e-05
step: 190, loss: 0.0010698882397264242
step: 200, loss: 0.0003481924068182707
step: 210, loss: 0.06567975133657455
step: 220, loss: 6.509088416351005e-05
step: 230, loss: 0.1194232627749443
step: 240, loss: 0.002652379684150219
step: 250, loss: 0.00020940686226822436
step: 260, loss: 0.00024390565522480756
step: 270, loss: 0.008904526941478252
step: 280, loss: 0.0004722516459878534
step: 290, loss: 0.00039385267882607877
step: 300, loss: 9.298406075686216e-05
step: 310, loss: 0.0013255361700430512
step: 320, loss: 0.0011140021961182356

```

```

step: 330, loss: 0.04377530887722969
step: 340, loss: 0.02039158158004284
step: 350, loss: 0.0004740389413200319
step: 360, loss: 0.00022737777908332646
step: 370, loss: 0.051261063665151596
step: 380, loss: 7.980526424944401e-05
step: 390, loss: 0.009278997778892517
step: 400, loss: 0.014209326356649399
step: 410, loss: 0.00021435710368677974
step: 420, loss: 0.00010228764585917816
step: 430, loss: 0.00067438546102494
step: 440, loss: 9.341332770418376e-05
step: 450, loss: 0.00020313615095801651
step: 460, loss: 0.006553042680025101
step: 470, loss: 0.1516941338777542
step: 480, loss: 0.011757586151361465
step: 490, loss: 0.004068458918482065
step: 500, loss: 0.0004955580225214362
step: 510, loss: 0.00935134757310152
step: 520, loss: 0.00019777526904363185
step: 530, loss: 0.00039197251317091286
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 8: dev_f1=0.954119850187266, f1=0.9510945505356311,
best_f1=0.9564007421150279
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00026247819187119603
step: 10, loss: 0.003125803079456091
step: 20, loss: 0.0002133020752808079
step: 30, loss: 0.027335958555340767
step: 40, loss: 8.72898890520446e-05
step: 50, loss: 0.0002953330986201763
step: 60, loss: 0.0008249810198321939
step: 70, loss: 0.0061074839904904366
step: 80, loss: 0.0012640525819733739
step: 90, loss: 0.08080190420150757
step: 100, loss: 0.0004583239206112921
step: 110, loss: 0.00028824852779507637
step: 120, loss: 0.00049494270933792
step: 130, loss: 0.0005595213151536882
step: 140, loss: 0.00017027021385729313
step: 150, loss: 0.006626083981245756
step: 160, loss: 0.0012327342992648482
step: 170, loss: 0.0005558373522944748
step: 180, loss: 0.00039286460378207266

```

```

step: 190, loss: 3.723421468748711e-05
step: 200, loss: 3.262402969994582e-05
step: 210, loss: 0.00034270971082150936
step: 220, loss: 4.5626627979800105e-05
step: 230, loss: 5.0797927542589605e-05
step: 240, loss: 0.0017670546658337116
step: 250, loss: 0.0005383932148106396
step: 260, loss: 2.889626739488449e-05
step: 270, loss: 0.00013516480976250023
step: 280, loss: 0.00028806261252611876
step: 290, loss: 2.6050269298139028e-05
step: 300, loss: 0.0009578876197338104
step: 310, loss: 0.0006979306926950812
step: 320, loss: 4.845180592383258e-05
step: 330, loss: 2.621417115733493e-05
step: 340, loss: 0.00011173867824254557
step: 350, loss: 0.022925766184926033
step: 360, loss: 3.252277383580804e-05
step: 370, loss: 2.2995420295046642e-05
step: 380, loss: 2.587517701613251e-05
step: 390, loss: 0.0001883386867120862
step: 400, loss: 0.07437648624181747
step: 410, loss: 5.498707832884975e-05
step: 420, loss: 2.351295916014351e-05
step: 430, loss: 0.00018494906544219702
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 440, loss: 0.0004862667992711067
step: 450, loss: 0.00041225849417969584
step: 460, loss: 0.00039058676338754594
step: 470, loss: 2.53644375334261e-05
step: 480, loss: 0.0016046379460021853
step: 490, loss: 0.00029450582223944366
step: 500, loss: 0.00015396709204651415
step: 510, loss: 7.708664634265006e-05
step: 520, loss: 0.0015867836773395538
step: 530, loss: 0.08462828397750854
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
epoch 9: dev_f1=0.958139534883721, f1=0.9562383612662941,
best_f1=0.9564007421150279
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0005638575530610979
step: 10, loss: 0.000719812698662281
step: 20, loss: 0.010658925399184227
step: 30, loss: 0.0006875750841572881

```

step: 40, loss: 7.003804785199463e-05  
step: 50, loss: 0.001452406169846654  
step: 60, loss: 0.0012369458563625813  
step: 70, loss: 2.7618330932455137e-05  
step: 80, loss: 0.0002753459557425231  
step: 90, loss: 6.744888378307223e-05  
step: 100, loss: 0.0007252544164657593  
step: 110, loss: 0.0032704323530197144  
step: 120, loss: 0.00010350175580242649  
step: 130, loss: 0.00013180643145460635  
step: 140, loss: 1.85740391316358e-05  
step: 150, loss: 0.002286733128130436  
step: 160, loss: 0.06843852996826172  
step: 170, loss: 1.8141730834031478e-05  
step: 180, loss: 0.0005590346991084516  
step: 190, loss: 1.0907564501394518e-05  
step: 200, loss: 2.5750501663424075e-05  
step: 210, loss: 0.00027589613455347717  
step: 220, loss: 0.0010220097610726953  
step: 230, loss: 1.54783483594656e-05  
step: 240, loss: 6.249726720852777e-05  
step: 250, loss: 0.00018598543829284608  
step: 260, loss: 0.0003969097451772541  
step: 270, loss: 3.0671257263747975e-05  
step: 280, loss: 0.0009282889077439904  
step: 290, loss: 0.000622378196567297  
step: 300, loss: 0.002705527935177088  
step: 310, loss: 0.09354385733604431  
step: 320, loss: 0.0013440410839393735  
step: 330, loss: 0.0014048476004973054  
step: 340, loss: 0.00010763741011032835  
step: 350, loss: 0.0005200892919674516  
step: 360, loss: 6.164974183775485e-05  
step: 370, loss: 7.683722651563585e-05  
step: 380, loss: 0.002194994827732444  
step: 390, loss: 0.00012231638538651168  
step: 400, loss: 0.0009445395553484559  
step: 410, loss: 0.0003248065186198801  
step: 420, loss: 0.0004969395231455564  
step: 430, loss: 0.00022577612253371626  
step: 440, loss: 3.8871374272275716e-05  
step: 450, loss: 0.002285382244735956  
step: 460, loss: 0.00035326386569067836  
step: 470, loss: 0.0022747989278286695  
step: 480, loss: 0.001328315120190382  
step: 490, loss: 0.0010829950915649533  
step: 500, loss: 0.004010385368019342  
step: 510, loss: 0.00018742332758847624

```

step: 520, loss: 0.019941767677664757
step: 530, loss: 0.0004062123771291226
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
epoch 10: dev_f1=0.9576547231270358, f1=0.9529576152771309,
best_f1=0.9564007421150279
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00024353018670808524
step: 10, loss: 0.00041960019734688103
step: 20, loss: 0.00013002997729927301
step: 30, loss: 0.00015241431538015604
step: 40, loss: 6.310607568593696e-05
step: 50, loss: 0.00018166138033848256
step: 60, loss: 0.00012349546886980534
step: 70, loss: 0.00015841316781006753
step: 80, loss: 0.0006300644599832594
step: 90, loss: 0.0008543393341824412
step: 100, loss: 0.0006661798688583076
step: 110, loss: 8.15362436696887e-05
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 8192.0
step: 120, loss: 0.0135755380615592
step: 130, loss: 0.0002999774587806314
step: 140, loss: 7.722269219812006e-05
step: 150, loss: 0.00016187936125788838
step: 160, loss: 0.0014196262927725911
step: 170, loss: 0.0007955781766213477
step: 180, loss: 8.711456030141562e-05
step: 190, loss: 0.002240904839709401
step: 200, loss: 0.0012083349283784628
step: 210, loss: 0.0006281092646531761
step: 220, loss: 0.000676055671647191
step: 230, loss: 0.004796546883881092
step: 240, loss: 0.0019139244686812162
step: 250, loss: 0.0012501670280471444
step: 260, loss: 0.0036472962237894535
step: 270, loss: 0.002793998923152685
step: 280, loss: 0.002931365044787526
step: 290, loss: 0.010042168200016022
step: 300, loss: 9.083282202482224e-05
step: 310, loss: 0.0011894929921254516
step: 320, loss: 0.007489004638046026
step: 330, loss: 7.396248111035675e-05
step: 340, loss: 0.0071934363804757595
step: 350, loss: 0.00023796198365744203
step: 360, loss: 0.0003161860804539174

```



```

step: 370, loss: 0.002557523548603058
step: 380, loss: 0.0008602682501077652
step: 390, loss: 0.004988003056496382
step: 400, loss: 0.002292242366820574
step: 410, loss: 0.0015902942977845669
step: 420, loss: 4.7388271923409775e-05
step: 430, loss: 0.0020446653943508863
step: 440, loss: 4.4723001337843016e-05
step: 450, loss: 0.0018387424061074853
step: 460, loss: 0.0007592116016894579
step: 470, loss: 9.31549584493041e-05
step: 480, loss: 0.0002572982048150152
step: 490, loss: 2.0775380107806996e-05
step: 500, loss: 5.639456503558904e-05
step: 510, loss: 8.424619591096416e-05
step: 520, loss: 9.406034223502502e-05
step: 530, loss: 0.0025366772897541523
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
epoch 11: dev_f1=0.9583917718560077, f1=0.951356407857811,
best_f1=0.9564007421150279
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.004611690063029528
step: 10, loss: 0.00030614668503403664
step: 20, loss: 0.0001025549863697961
step: 30, loss: 2.9859076676075347e-05
step: 40, loss: 0.00030154723208397627
step: 50, loss: 3.434344398556277e-05
step: 60, loss: 0.001453861128538847
step: 70, loss: 0.0008334047743119299
step: 80, loss: 2.777119334496092e-05
step: 90, loss: 0.0001529255969217047
step: 100, loss: 0.0008013342157937586
step: 110, loss: 3.193176235072315e-05
step: 120, loss: 0.00016197608783841133
step: 130, loss: 6.923050386831164e-05
step: 140, loss: 4.75661399832461e-05
step: 150, loss: 2.852271791198291e-05
step: 160, loss: 6.0982376453466713e-05
step: 170, loss: 3.632313018897548e-05
step: 180, loss: 0.0017586068715900183
step: 190, loss: 0.0042219581082463264
step: 200, loss: 0.0022038801107555628
step: 210, loss: 3.2479285437148064e-05
step: 220, loss: 0.00014740774349775165

```

```

step: 230, loss: 2.3043990950100124e-05
step: 240, loss: 3.2543692213948816e-05
step: 250, loss: 0.0006568097742274404
step: 260, loss: 3.433303936617449e-05
step: 270, loss: 0.0023629292845726013
step: 280, loss: 3.2334392017219216e-05
step: 290, loss: 0.0013751956867054105
step: 300, loss: 0.002780070062726736
step: 310, loss: 0.013146420009434223
step: 320, loss: 0.0013050464913249016
step: 330, loss: 0.0010164420818910003
step: 340, loss: 0.0014761477941647172
step: 350, loss: 0.003274243324995041
step: 360, loss: 4.541029193205759e-05
step: 370, loss: 0.0012361998669803143
step: 380, loss: 0.0001413322752341628
step: 390, loss: 0.00026302612968720496
step: 400, loss: 7.901060598669574e-05
step: 410, loss: 2.9667016860912554e-05
step: 420, loss: 9.374826186103746e-05
step: 430, loss: 7.969048601808026e-05
step: 440, loss: 3.724876296473667e-05
step: 450, loss: 0.0002437161310808733
step: 460, loss: 2.448523991915863e-05
step: 470, loss: 0.0018856209935620427
step: 480, loss: 0.001953471451997757
step: 490, loss: 2.5133820599876344e-05
step: 500, loss: 2.7058276828029193e-05
step: 510, loss: 0.00020866947306785733
step: 520, loss: 0.0003823688020929694
step: 530, loss: 3.2874206226551905e-05
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
    warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 12: dev_f1=0.9598506069094305, f1=0.9520260829063809,
best_f1=0.9520260829063809
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 2.2600330339628272e-05
step: 10, loss: 8.372668526135385e-05
step: 20, loss: 2.5413242838112637e-05
step: 30, loss: 1.472205804020632e-05
step: 40, loss: 0.00042337103514000773

```

step: 50, loss: 0.002499842317774892  
step: 60, loss: 3.368251782376319e-05  
step: 70, loss: 2.1010089767514728e-05  
step: 80, loss: 0.0003474997356534004  
step: 90, loss: 2.727523497014772e-05  
step: 100, loss: 1.8070904843625613e-05  
step: 110, loss: 1.4509765605907887e-05  
step: 120, loss: 5.412329119280912e-05  
step: 130, loss: 0.00020088048768229783  
step: 140, loss: 0.00028112524887546897  
step: 150, loss: 0.001178201986476779  
step: 160, loss: 0.00010362293687649071  
step: 170, loss: 0.001345670665614307  
step: 180, loss: 2.3531854822067544e-05  
step: 190, loss: 2.862683504645247e-05  
step: 200, loss: 2.0752780983457342e-05  
step: 210, loss: 2.2101110516814515e-05  
step: 220, loss: 1.744879955367651e-05  
step: 230, loss: 0.0007206139853224158  
step: 240, loss: 0.0009158212342299521  
step: 250, loss: 2.405725172138773e-05  
step: 260, loss: 4.127688953303732e-05  
step: 270, loss: 0.0035942026879638433  
step: 280, loss: 2.7807254809886217e-05  
step: 290, loss: 1.5857973266975023e-05  
step: 300, loss: 2.2738271582056768e-05  
step: 310, loss: 2.0495703211054206e-05  
step: 320, loss: 0.002106370171532035  
step: 330, loss: 2.382994171057362e-05  
step: 340, loss: 0.0007698661065660417  
step: 350, loss: 1.1112358151876833e-05  
step: 360, loss: 0.08324563503265381  
step: 370, loss: 2.2505999368149787e-05  
step: 380, loss: 2.027581285801716e-05  
step: 390, loss: 3.859448042931035e-05  
step: 400, loss: 2.708502324821893e-05  
step: 410, loss: 2.4392391424044035e-05  
step: 420, loss: 2.2837912183604203e-05  
step: 430, loss: 3.5756569559453055e-05  
step: 440, loss: 1.2740174497594126e-05  
step: 450, loss: 1.8156595615437254e-05  
step: 460, loss: 4.918028207612224e-05  
step: 470, loss: 0.0012038907734677196  
step: 480, loss: 1.0520050636841916e-05  
step: 490, loss: 4.050131610711105e-05  
step: 500, loss: 2.535427120164968e-05  
step: 510, loss: 0.005986299831420183  
step: 520, loss: 1.4029210433363914e-05

```
step: 530, loss: 1.3749815479968674e-05
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 13: dev_f1=0.9582355701548568, f1=0.9525598872710193,
best_f1=0.9520260829063809
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0004605374706443399
step: 10, loss: 0.0005443136324174702
step: 20, loss: 1.3332609341887292e-05
step: 30, loss: 5.881129618501291e-05
step: 40, loss: 1.2945185517310165e-05
step: 50, loss: 5.158610292710364e-05
step: 60, loss: 3.440114596742205e-05
step: 70, loss: 3.857609408441931e-05
step: 80, loss: 0.0008094048243947327
step: 90, loss: 1.704280475678388e-05
step: 100, loss: 0.00020995766681153327
step: 110, loss: 2.4314138499903493e-05
step: 120, loss: 1.820140823838301e-05
step: 130, loss: 1.9374814655748196e-05
step: 140, loss: 0.0001720556610962376
step: 150, loss: 0.0021259181667119265
step: 160, loss: 3.000185824930668e-05
step: 170, loss: 0.003448472823947668
step: 180, loss: 4.229298065183684e-05
step: 190, loss: 0.000593487813603133
step: 200, loss: 8.1846650573425e-05
step: 210, loss: 0.0035685000475496054
step: 220, loss: 0.0001236482639797032
step: 230, loss: 0.002658864250406623
step: 240, loss: 0.002412019995972514
step: 250, loss: 0.005792991258203983
step: 260, loss: 5.190488445805386e-05
step: 270, loss: 0.001496081822551787
step: 280, loss: 5.257780139800161e-05
step: 290, loss: 5.1776631153188646e-05
step: 300, loss: 7.300058496184647e-05
step: 310, loss: 5.559607961913571e-05
step: 320, loss: 6.017841587890871e-05
step: 330, loss: 0.001637404435314238
step: 340, loss: 8.65768306539394e-05
step: 350, loss: 1.8477137928130105e-05
step: 360, loss: 0.0004068944544997066
step: 370, loss: 0.00027534892433322966
step: 380, loss: 0.0003102893242612481
```

```

step: 390, loss: 0.0008831100421957672
step: 400, loss: 0.002396673196926713
step: 410, loss: 2.8314356313785538e-05
step: 420, loss: 2.264544127683621e-05
step: 430, loss: 8.474278001813218e-05
step: 440, loss: 3.6080815334571525e-05
step: 450, loss: 2.8407877834979445e-05
step: 460, loss: 0.017696693539619446
step: 470, loss: 2.008265073527582e-05
step: 480, loss: 0.00017049651069100946
step: 490, loss: 0.0006229100399650633
step: 500, loss: 0.0017115904483944178
step: 510, loss: 0.010732829570770264
step: 520, loss: 2.700655568332877e-05
step: 530, loss: 0.002510816091671586
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 14: dev_f1=0.957889865802869, f1=0.951061865189289,
best_f1=0.9520260829063809
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 4.137407449888997e-05
step: 10, loss: 0.002046150155365467
step: 20, loss: 0.0019225124269723892
step: 30, loss: 0.0023788383696228266
step: 40, loss: 3.816638854914345e-05
step: 50, loss: 0.0011817218037322164
step: 60, loss: 5.131219222676009e-05
step: 70, loss: 2.430258427921217e-05
step: 80, loss: 2.046631016128231e-05
step: 90, loss: 0.0009640518110245466
step: 100, loss: 1.596247238921933e-05
step: 110, loss: 0.000742452044505626
step: 120, loss: 3.5719644074561074e-05
step: 130, loss: 3.858367199427448e-05
step: 140, loss: 1.6532521840417758e-05
step: 150, loss: 2.2973292288952507e-05
step: 160, loss: 2.0972813217667863e-05
step: 170, loss: 1.787737346603535e-05
step: 180, loss: 3.09148381347768e-05
step: 190, loss: 0.001504075014963746
step: 200, loss: 0.0012845526216551661
step: 210, loss: 0.00011048493615817279
step: 220, loss: 2.391940142842941e-05
step: 230, loss: 0.00020500940445344895
step: 240, loss: 2.2894910216564313e-05

```

```

step: 250, loss: 2.71334847639082e-05
step: 260, loss: 1.3578537618741393e-05
step: 270, loss: 1.915129178087227e-05
step: 280, loss: 1.4427881978917867e-05
step: 290, loss: 0.00035380732151679695
step: 300, loss: 0.0004088127752766013
step: 310, loss: 0.037189334630966187
step: 320, loss: 2.2619527953793295e-05
step: 330, loss: 3.01588952424936e-05
step: 340, loss: 0.0004932628362439573
step: 350, loss: 0.0022387478966265917
step: 360, loss: 4.8390480515081435e-05
step: 370, loss: 0.00040824885945767164
step: 380, loss: 0.00043279011151753366
step: 390, loss: 0.0012755633797496557
step: 400, loss: 0.005371847189962864
step: 410, loss: 2.231392136309296e-05
step: 420, loss: 0.0001287075283471495
step: 430, loss: 1.4632687452831306e-05
step: 440, loss: 0.002102309837937355
step: 450, loss: 0.0018004585290327668
step: 460, loss: 0.008596030995249748
step: 470, loss: 1.5407573300763033e-05
step: 480, loss: 0.0021782820113003254
step: 490, loss: 0.0004548184806481004
step: 500, loss: 0.0015908076893538237
step: 510, loss: 1.7694848793325946e-05
step: 520, loss: 2.6407169571029954e-05
step: 530, loss: 1.4435246157518122e-05
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
epoch 15: dev_f1=0.9592215013901761, f1=0.9542725173210163,
best_f1=0.9520260829063809
mv: cannot stat '*_dev.pt': No such file or directory
Selected optimization level 02: FP16 training with FP32 batchnorm and FP32
master weights.

Defaults for this optimization level are:
enabled                : True
opt_level              : 02
cast_model_type        : torch.float16
patch_torch_functions  : False
keep_batchnorm_fp32    : True
master_weights         : True
loss_scale             : dynamic
Processing user overrides (additional kwargs that are not None)...
After processing overrides, optimization options are:

```

```

enabled                : True
opt_level              : 02
cast_model_type        : torch.float16
patch_torch_functions  : False
keep_batchnorm_fp32    : True
master_weights         : True
loss_scale             : dynamic
Warning: multi_tensor_applier fused unscale kernel is unavailable, possibly
because apex was installed without --cuda_ext --cpp_ext. Using Python fallback.
Original ImportError was: ModuleNotFoundError("No module named 'amp_C'")
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
Oit [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
5742it [00:21, 261.46it/s]
load_f1 = 0.960710944808232
real_f1 = 0.957825679475164
Oit [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
4398it [00:21, 200.88it/s]

```

### 3 DK

```

[ ]: # Para ejecutar dk necesita
      #!pip install -r requirements.txt
      !python -m spacy download en_core_web_lg

```

```

Looking in indexes: https://pypi.org/simple, https://us-python.pkg.dev/colab-
wheels/public/simple/
Collecting en_core_web_lg==2.2.5
  Downloading https://github.com/explosion/spacy-
models/releases/download/en_core_web_lg-2.2.5/en_core_web_lg-2.2.5.tar.gz (827.9
MB)
    |                               | 827.9 MB 508 kB/s
Requirement already satisfied: spacy>=2.2.2 in
/usr/local/lib/python3.7/dist-packages (from en_core_web_lg==2.2.5) (2.2.3)
Requirement already satisfied: numpy>=1.15.0 in /usr/local/lib/python3.7/dist-
packages (from spacy>=2.2.2->en_core_web_lg==2.2.5) (1.19.2)
Requirement already satisfied: preshed<3.1.0,>=3.0.2 in
/usr/local/lib/python3.7/dist-packages (from
spacy>=2.2.2->en_core_web_lg==2.2.5) (3.0.6)
Requirement already satisfied: blis<0.5.0,>=0.4.0 in
/usr/local/lib/python3.7/dist-packages (from

```

```

spacy>=2.2.2->en_core_web_lg==2.2.5) (0.4.1)
Requirement already satisfied: plac<1.2.0,>=0.9.6 in
/usr/local/lib/python3.7/dist-packages (from
spacy>=2.2.2->en_core_web_lg==2.2.5) (1.1.3)
Requirement already satisfied: srsly<1.1.0,>=0.1.0 in
/usr/local/lib/python3.7/dist-packages (from
spacy>=2.2.2->en_core_web_lg==2.2.5) (1.0.5)
Requirement already satisfied: wasabi<1.1.0,>=0.4.0 in
/usr/local/lib/python3.7/dist-packages (from
spacy>=2.2.2->en_core_web_lg==2.2.5) (0.10.1)
Requirement already satisfied: thinc<7.4.0,>=7.3.0 in
/usr/local/lib/python3.7/dist-packages (from
spacy>=2.2.2->en_core_web_lg==2.2.5) (7.3.1)
Requirement already satisfied: requests<3.0.0,>=2.13.0 in
/usr/local/lib/python3.7/dist-packages (from
spacy>=2.2.2->en_core_web_lg==2.2.5) (2.23.0)
Requirement already satisfied: catalogue<1.1.0,>=0.0.7 in
/usr/local/lib/python3.7/dist-packages (from
spacy>=2.2.2->en_core_web_lg==2.2.5) (1.0.0)
Requirement already satisfied: cymem<2.1.0,>=2.0.2 in
/usr/local/lib/python3.7/dist-packages (from
spacy>=2.2.2->en_core_web_lg==2.2.5) (2.0.6)
Requirement already satisfied: murmurhash<1.1.0,>=0.28.0 in
/usr/local/lib/python3.7/dist-packages (from
spacy>=2.2.2->en_core_web_lg==2.2.5) (1.0.7)
Requirement already satisfied: setuptools in /usr/local/lib/python3.7/dist-
packages (from spacy>=2.2.2->en_core_web_lg==2.2.5) (57.4.0)
Requirement already satisfied: importlib-metadata>=0.20 in
/usr/local/lib/python3.7/dist-packages (from
catalogue<1.1.0,>=0.0.7->spacy>=2.2.2->en_core_web_lg==2.2.5) (4.12.0)
Requirement already satisfied: zipp>=0.5 in /usr/local/lib/python3.7/dist-
packages (from importlib-
metadata>=0.20->catalogue<1.1.0,>=0.0.7->spacy>=2.2.2->en_core_web_lg==2.2.5)
(3.8.1)
Requirement already satisfied: typing-extensions>=3.6.4 in
/usr/local/lib/python3.7/dist-packages (from importlib-
metadata>=0.20->catalogue<1.1.0,>=0.0.7->spacy>=2.2.2->en_core_web_lg==2.2.5)
(4.1.1)
Requirement already satisfied: urllib3!=1.25.0,!1.25.1,<1.26,>=1.21.1 in
/usr/local/lib/python3.7/dist-packages (from
requests<3.0.0,>=2.13.0->spacy>=2.2.2->en_core_web_lg==2.2.5) (1.25.4)
Requirement already satisfied: certifi>=2017.4.17 in
/usr/local/lib/python3.7/dist-packages (from
requests<3.0.0,>=2.13.0->spacy>=2.2.2->en_core_web_lg==2.2.5) (2022.6.15)
Requirement already satisfied: chardet<4,>=3.0.2 in
/usr/local/lib/python3.7/dist-packages (from
requests<3.0.0,>=2.13.0->spacy>=2.2.2->en_core_web_lg==2.2.5) (3.0.4)
Requirement already satisfied: idna<3,>=2.5 in /usr/local/lib/python3.7/dist-

```



```

packages (from requests<3.0.0,>=2.13.0->spacy>=2.2.2->en_core_web_lg==2.2.5)
(2.10)
Requirement already satisfied: tqdm<5.0.0,>=4.10.0 in
/usr/local/lib/python3.7/dist-packages (from
thinc<7.4.0,>=7.3.0->spacy>=2.2.2->en_core_web_lg==2.2.5) (4.41.0)
Building wheels for collected packages: en-core-web-lg
  Building wheel for en-core-web-lg (setup.py) ... done
  Created wheel for en-core-web-lg: filename=en_core_web_lg-2.2.5-py3-none-
any.whl size=829180942
sha256=5b6bfa33c3e216d7810e580e25742abd48c44e7ec6a2bfc8c64bb27c279b9001
  Stored in directory: /tmp/pip-ephem-wheel-cache-
um5isvoo/wheels/11/95/ba/2c36cc368c0bd339b44a791c2c1881a1fb714b78c29a4cb8f5
Successfully built en-core-web-lg
Installing collected packages: en-core-web-lg
Successfully installed en-core-web-lg-2.2.5
  Download and installation successful
You can now load the model via spacy.load('en_core_web_lg')

```

## 3.1 DK STRUCTURED

### 3.1.1 Beer - Running the matcher

```

[ ]: # Step 1:
# train the wdc model with the save_model flag on
# it will produce two files *_dev.pt and *_test.pt which are the
# best checkpoints on the dev set and the test set.
!CUDA_VISIBLE_DEVICES=0 python train_ditto.py \
  --task Structured/Beer \
  --batch_size 32 \
  --max_len 256 \
  --lr 3e-5 \
  --n_epochs 15 \
  --finetuning \
  --lm roberta \
  --fp16 \
  --save_model \
  --dk general \
  --dk product \
  --summarize

# Step 2: pick a checkpoint, rename it to wdc_all_small.pt
!mv *_dev.pt checkpoints/Beer.pt

# Step 3: run the matcher
!CUDA_VISIBLE_DEVICES=0 python matcher.py \
  --task Structured/Beer \

```

```

--input_path input/input_small.jsonl \
--output_path output/output_small.jsonl \
--lm roberta \
--use_gpu \
--fp16 \
--checkpoint_path checkpoints/

```

```

Downloading: 100% 481/481 [00:00<00:00, 472kB/s]
Downloading: 100% 899k/899k [00:00<00:00, 24.2MB/s]
Downloading: 100% 456k/456k [00:00<00:00, 7.64MB/s]
Downloading: 100% 501M/501M [00:07<00:00, 69.4MB/s]
Selected optimization level O2:  FP16 training with FP32 batchnorm and FP32
master weights.

```

Defaults for this optimization level are:

```

enabled          : True
opt_level        : O2
cast_model_type  : torch.float16
patch_torch_functions : False
keep_batchnorm_fp32 : True
master_weights   : True
loss_scale       : dynamic

```

Processing user overrides (additional kwargs that are not None)...

After processing overrides, optimization options are:

```

enabled          : True
opt_level        : O2
cast_model_type  : torch.float16
patch_torch_functions : False
keep_batchnorm_fp32 : True
master_weights   : True
loss_scale       : dynamic

```

Warning: multi\_tensor\_applier fused unscale kernel is unavailable, possibly because apex was installed without --cuda\_ext --cpp\_ext. Using Python fallback.

Original ImportError was: ModuleNotFoundError("No module named 'amp\_C'")

/usr/local/lib/python3.7/dist-packages/apex/amp/\_initialize.py:25: UserWarning: An input tensor was not cuda.

```
warnings.warn("An input tensor was not cuda.")
```

Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 32768.0

/usr/local/lib/python3.7/dist-packages/torch/optim/lr\_scheduler.py:136:

UserWarning: Detected call of `lr\_scheduler.step()` before `optimizer.step()`. In PyTorch 1.1.0 and later, you should call them in the opposite order:

`optimizer.step()` before `lr\_scheduler.step()`. Failure to do this will result in PyTorch skipping the first value of the learning rate schedule. See more details at <https://pytorch.org/docs/stable/optim.html#how-to-adjust-learning-rate>

"<https://pytorch.org/docs/stable/optim.html#how-to-adjust-learning-rate>", UserWarning)

```

step: 0, loss: 0.4385996162891388
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
    warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 1: dev_f1=0.3076923076923077, f1=0.3218390804597701,
best_f1=0.3218390804597701
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.45578837394714355
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 2: dev_f1=0.2666666666666667, f1=0.2666666666666667,
best_f1=0.3218390804597701
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.4194447994232178
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
    warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 3: dev_f1=0.3181818181818182, f1=0.3333333333333333,
best_f1=0.3333333333333333
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.3055943548679352
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
    warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 4: dev_f1=0.45901639344262296, f1=0.41379310344827586,
best_f1=0.41379310344827586
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.

```

```

    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.32068753242492676
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
    warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 5: dev_f1=0.7058823529411764, f1=0.6842105263157894,
best_f1=0.6842105263157894
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.21823152899742126
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
    warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 6: dev_f1=0.8461538461538461, f1=0.7586206896551724,
best_f1=0.7586206896551724
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.4100833535194397
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
    warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 7: dev_f1=0.8666666666666666, f1=0.8235294117647058,
best_f1=0.8235294117647058
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.2587135136127472
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
    warnings.warn(SAVE_STATE_WARNING, UserWarning)

```

```

epoch 8: dev_f1=0.8750000000000001, f1=0.8, best_f1=0.8
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.1449996829032898
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 9: dev_f1=0.8750000000000001, f1=0.8235294117647058, best_f1=0.8
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.11484058946371078
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
    warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 10: dev_f1=0.896551724137931, f1=0.9333333333333333,
best_f1=0.9333333333333333
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.06272298842668533
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 11: dev_f1=0.896551724137931, f1=0.9333333333333333,
best_f1=0.9333333333333333
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00845396053045988
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 12: dev_f1=0.8666666666666666, f1=0.9333333333333333,
best_f1=0.9333333333333333
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.005486245732754469
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 13: dev_f1=0.8750000000000001, f1=0.9333333333333333,

```

```

best_f1=0.9333333333333333
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0031760192941874266
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 14: dev_f1=0.8750000000000001, f1=0.9333333333333333,
best_f1=0.9333333333333333
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.004166550002992153
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 15: dev_f1=0.8750000000000001, f1=0.9333333333333333,
best_f1=0.9333333333333333
mv: cannot stat '*_dev.pt': No such file or directory
Selected optimization level 02:  FP16 training with FP32 batchnorm and FP32
master weights.

```

Defaults for this optimization level are:

```

enabled          : True
opt_level        : 02
cast_model_type  : torch.float16
patch_torch_functions : False
keep_batchnorm_fp32 : True
master_weights   : True
loss_scale       : dynamic

```

Processing user overrides (additional kwargs that are not None)...

After processing overrides, optimization options are:

```

enabled          : True
opt_level        : 02
cast_model_type  : torch.float16
patch_torch_functions : False
keep_batchnorm_fp32 : True
master_weights   : True
loss_scale       : dynamic

```

Warning: multi\_tensor\_applier fused unscale kernel is unavailable, possibly because apex was installed without --cuda\_ext --cpp\_ext. Using Python fallback. Original ImportError was: ModuleNotFoundError("No module named 'amp\_C'")

```

/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")

```

```

91it [00:00, 80557.55it/s]

```

```

/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:

```

```

An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
load_f1 = 0.9333333333333333
real_f1 = 0.9032258064516129
Oit [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
4398it [00:18, 231.76it/s]

```

### 3.1.2 DBLP-ACM - Running the matcher

```

[ ]: # Step 1:
# train the wdc model with the save_model flag on
# it will produce two files *_dev.pt and *_test.pt which are the
# best checkpoints on the dev set and the test set.
!CUDA_VISIBLE_DEVICES=0 python train_ditto.py \
  --task Structured/DBLP-ACM \
  --batch_size 32 \
  --max_len 256 \
  --lr 3e-5 \
  --n_epochs 15 \
  --finetuning \
  --lm roberta \
  --fp16 \
  --save_model \
  --dk general \
  --dk product \
  --summarize

# Step 2: pick a checkpoint, rename it to wdc_all_small.pt
!mv *_dev.pt checkpoints/DBLP-ACM.pt

# Step 3: run the matcher
!CUDA_VISIBLE_DEVICES=0 python matcher.py \
  --task Structured/DBLP-ACM \
  --input_path input/input_small.jsonl \
  --output_path output/output_small.jsonl \
  --lm roberta \
  --use_gpu \
  --fp16 \
  --checkpoint_path checkpoints/

```

```

Downloading: 100% 481/481 [00:00<00:00, 439kB/s]
Downloading: 100% 899k/899k [00:01<00:00, 805kB/s]
Downloading: 100% 456k/456k [00:04<00:00, 102kB/s]

```

Downloading: 100% 501M/501M [00:07<00:00, 69.8MB/s]  
Selected optimization level 02: FP16 training with FP32 batchnorm and FP32 master weights.

Defaults for this optimization level are:

enabled : True  
opt\_level : 02  
cast\_model\_type : torch.float16  
patch\_torch\_functions : False  
keep\_batchnorm\_fp32 : True  
master\_weights : True  
loss\_scale : dynamic

Processing user overrides (additional kwargs that are not None)...

After processing overrides, optimization options are:

enabled : True  
opt\_level : 02  
cast\_model\_type : torch.float16  
patch\_torch\_functions : False  
keep\_batchnorm\_fp32 : True  
master\_weights : True  
loss\_scale : dynamic

Warning: multi\_tensor\_applier fused unscale kernel is unavailable, possibly because apex was installed without --cuda\_ext --cpp\_ext. Using Python fallback. Original ImportError was: ModuleNotFoundError("No module named 'amp\_C'")  
/usr/local/lib/python3.7/dist-packages/apex/amp/\_initialize.py:25: UserWarning: An input tensor was not cuda.

warnings.warn("An input tensor was not cuda.")

step: 0, loss: 0.5662547945976257

Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 32768.0

Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0

step: 10, loss: 0.4485122561454773

step: 20, loss: 0.5422874689102173

step: 30, loss: 0.275655061006546

step: 40, loss: 0.19556918740272522

step: 50, loss: 0.17422613501548767

step: 60, loss: 0.06566894054412842

step: 70, loss: 0.09551999717950821

step: 80, loss: 0.021579276770353317

step: 90, loss: 0.017425719648599625

step: 100, loss: 0.12330068647861481

step: 110, loss: 0.050625357776880264

step: 120, loss: 0.012410424649715424

Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 8192.0

step: 130, loss: 0.0219658724963665

step: 140, loss: 0.09316825866699219

step: 150, loss: 0.10747630894184113

step: 160, loss: 0.08586094528436661

step: 170, loss: 0.024715784937143326



```

step: 180, loss: 0.060394901782274246
step: 190, loss: 0.19945284724235535
step: 200, loss: 0.07492364197969437
step: 210, loss: 0.03489534929394722
step: 220, loss: 0.07761987298727036
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 4096.0
step: 230, loss: 0.005162075161933899
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
    warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 1: dev_f1=0.9627959413754228, f1=0.9652076318742986,
best_f1=0.9652076318742986
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.010092166252434254
step: 10, loss: 0.062332525849342346
step: 20, loss: 0.0016465213848277926
step: 30, loss: 0.005375273991376162
step: 40, loss: 0.07132387161254883
step: 50, loss: 0.009768444113433361
step: 60, loss: 0.0018433054210618138
step: 70, loss: 0.039729416370391846
step: 80, loss: 0.0018060039728879929
step: 90, loss: 0.004182989709079266
step: 100, loss: 0.005208002869039774
step: 110, loss: 0.006959820166230202
step: 120, loss: 0.018855925649404526
step: 130, loss: 0.0026598291005939245
step: 140, loss: 0.003418551990762353
step: 150, loss: 0.12432592362165451
step: 160, loss: 0.0032180803827941418
step: 170, loss: 0.0009840911952778697
step: 180, loss: 0.007669174112379551
step: 190, loss: 0.0173801202327013
step: 200, loss: 0.017375946044921875
step: 210, loss: 0.02105928212404251
step: 220, loss: 0.06432797759771347
step: 230, loss: 0.002245519310235977
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or

```

```

loading the scheduler.
  warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 2: dev_f1=0.9866369710467707, f1=0.9788182831661093,
best_f1=0.9788182831661093
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.005556079093366861
step: 10, loss: 0.0029447078704833984
step: 20, loss: 0.0016880730399861932
step: 30, loss: 0.0004531051963567734
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 2048.0
step: 40, loss: 0.0015563045162707567
step: 50, loss: 0.08364392817020416
step: 60, loss: 0.06664874404668808
step: 70, loss: 0.002828830387443304
step: 80, loss: 0.02709854207932949
step: 90, loss: 0.0016016154550015926
step: 100, loss: 0.003588153049349785
step: 110, loss: 0.012281649746000767
step: 120, loss: 0.003794738557189703
step: 130, loss: 0.0028927307575941086
step: 140, loss: 0.003426210256293416
step: 150, loss: 0.14763645827770233
step: 160, loss: 0.0025488128885626793
step: 170, loss: 0.006283605471253395
step: 180, loss: 0.007031416520476341
step: 190, loss: 0.0878966823220253
step: 200, loss: 0.0273915845900774
step: 210, loss: 0.0035661717411130667
step: 220, loss: 0.014681362546980381
step: 230, loss: 0.00892640184611082
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
  warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 3: dev_f1=0.987598647125141, f1=0.983050847457627,
best_f1=0.983050847457627
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00542760593816638
step: 10, loss: 0.0060324654914438725
step: 20, loss: 0.007099132053554058
step: 30, loss: 0.007700997870415449

```

```

step: 40, loss: 0.04394876956939697
step: 50, loss: 0.01239671092480421
step: 60, loss: 0.004397708456963301
step: 70, loss: 0.01689024455845356
step: 80, loss: 0.0008891685283742845
step: 90, loss: 0.0037283929996192455
step: 100, loss: 0.0019218393135815859
step: 110, loss: 0.0023465969134122133
step: 120, loss: 0.08615165203809738
step: 130, loss: 0.0032018476631492376
step: 140, loss: 0.0008260990725830197
step: 150, loss: 0.0002863790141418576
step: 160, loss: 0.0008511998457834125
step: 170, loss: 0.0005725004011765122
step: 180, loss: 0.07159902155399323
step: 190, loss: 0.00175578857306391
step: 200, loss: 0.005456387996673584
step: 210, loss: 0.005669794511049986
step: 220, loss: 0.0007406021468341351
step: 230, loss: 0.056724511086940765
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
  warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 4: dev_f1=0.9909706546275394, f1=0.9864864864864865,
best_f1=0.9864864864864865
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0005394132458604872
step: 10, loss: 0.0014569518389180303
step: 20, loss: 0.004238917957991362
step: 30, loss: 0.0026614987291395664
step: 40, loss: 0.001682139583863318
step: 50, loss: 0.004574143793433905
step: 60, loss: 0.26823386549949646
step: 70, loss: 0.014182514511048794
step: 80, loss: 0.012076355516910553
step: 90, loss: 0.018159491941332817
step: 100, loss: 0.00035794952418655157
step: 110, loss: 0.002150734420865774
step: 120, loss: 0.000755027518607676
step: 130, loss: 0.0002645210188347846
step: 140, loss: 0.014039614237844944
step: 150, loss: 0.04650953784584999

```

```

step: 160, loss: 0.000787738710641861
step: 170, loss: 0.008442816324532032
step: 180, loss: 0.001333677559159696
step: 190, loss: 0.02733791060745716
step: 200, loss: 0.0014036481734365225
step: 210, loss: 0.018229922279715538
step: 220, loss: 0.0032983236014842987
step: 230, loss: 0.00799618475139141
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 5: dev_f1=0.9865771812080537, f1=0.9865168539325843,
best_f1=0.9864864864864865
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00037755173980258405
step: 10, loss: 0.0023023737594485283
step: 20, loss: 0.0014156909892335534
step: 30, loss: 0.006448559463024139
step: 40, loss: 0.0731714516878128
step: 50, loss: 0.0033392233308404684
step: 60, loss: 0.0033198597375303507
step: 70, loss: 0.015959208831191063
step: 80, loss: 0.02949732355773449
step: 90, loss: 0.03451518714427948
step: 100, loss: 0.0012438042322173715
step: 110, loss: 0.04237718507647514
step: 120, loss: 0.0012730053858831525
step: 130, loss: 0.0008378314669243991
step: 140, loss: 0.0006036839913576841
step: 150, loss: 0.0005137632833793759
step: 160, loss: 0.0427587665617466
step: 170, loss: 0.0016689153853803873
step: 180, loss: 0.004386911168694496
step: 190, loss: 0.0012572597479447722
step: 200, loss: 0.0013562310487031937
step: 210, loss: 0.0014481873949989676
step: 220, loss: 0.009323344565927982
step: 230, loss: 0.00572619354352355
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 6: dev_f1=0.9831649831649831, f1=0.9740698985343857,
best_f1=0.9864864864864865
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")

```

```

step: 0, loss: 0.0032383922953158617
step: 10, loss: 0.0011737225577235222
step: 20, loss: 0.0011502623092383146
step: 30, loss: 0.0005516939563676715
step: 40, loss: 0.0016520947683602571
step: 50, loss: 0.001028487691655755
step: 60, loss: 0.0005908644525334239
step: 70, loss: 0.0007301056757569313
step: 80, loss: 0.0013972497545182705
step: 90, loss: 0.0017784875817596912
step: 100, loss: 0.019040660932660103
step: 110, loss: 0.000936942407861352
step: 120, loss: 0.0020388381090015173
step: 130, loss: 0.0010572924511507154
step: 140, loss: 0.0004511881561484188
step: 150, loss: 0.0030865732114762068
step: 160, loss: 0.0006083078333176672
step: 170, loss: 0.000602036016061902
step: 180, loss: 0.0004907264956273139
step: 190, loss: 0.000809038057923317
step: 200, loss: 0.05170955881476402
step: 210, loss: 0.00031886136275716126
step: 220, loss: 0.0006224018870852888
step: 230, loss: 0.006817021407186985
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
epoch 7: dev_f1=0.9877641824249165, f1=0.9777777777777777,
best_f1=0.9864864864864865
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.001301645883359015
step: 10, loss: 0.0035023875534534454
step: 20, loss: 0.0015414035879075527
step: 30, loss: 0.0012236121110618114
step: 40, loss: 0.004149250686168671
step: 50, loss: 0.0015896798577159643
step: 60, loss: 0.0004994368064217269
step: 70, loss: 0.0003914523986168206
step: 80, loss: 0.02862504869699478
step: 90, loss: 0.0006726954015903175
step: 100, loss: 0.0007280520512722433
step: 110, loss: 0.0015709740109741688
step: 120, loss: 0.002593046287074685
step: 130, loss: 0.0014148103073239326
step: 140, loss: 0.00034026725916191936
step: 150, loss: 0.040609270334243774

```

```

step: 160, loss: 0.010640614666044712
step: 170, loss: 0.001657323562540114
step: 180, loss: 0.00040220413939096034
step: 190, loss: 0.0016956597100943327
step: 200, loss: 0.0040147071704268456
step: 210, loss: 0.002503158524632454
step: 220, loss: 0.0018699637148529291
step: 230, loss: 0.0007159435772337019
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 8: dev_f1=0.9887892376681614, f1=0.9799107142857142,
best_f1=0.9864864864864865
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.000680374912917614
step: 10, loss: 0.001023370074108243
step: 20, loss: 0.0011761263012886047
step: 30, loss: 0.0009621634962968528
step: 40, loss: 0.0008422463433817029
step: 50, loss: 0.0005194359691813588
step: 60, loss: 0.000539666332770139
step: 70, loss: 0.020230289548635483
step: 80, loss: 0.00013445359945762902
step: 90, loss: 0.02080669440329075
step: 100, loss: 0.00034667018917389214
step: 110, loss: 0.00029661471489816904
step: 120, loss: 0.049271319061517715
step: 130, loss: 0.0005823067622259259
step: 140, loss: 0.0006074621342122555
step: 150, loss: 0.0005377255147323012
step: 160, loss: 0.0006870491779409349
step: 170, loss: 0.00022643145348411053
step: 180, loss: 0.00031251070322468877
step: 190, loss: 0.00021645912784151733
step: 200, loss: 0.000294159515760839
step: 210, loss: 0.0009024245082400739
step: 220, loss: 0.0004512495070230216
step: 230, loss: 0.0002619293227326125
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 9: dev_f1=0.9887640449438202, f1=0.987598647125141,
best_f1=0.9864864864864865
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")

```

```

step: 0, loss: 0.0004939327482134104
step: 10, loss: 0.0007191720069386065
step: 20, loss: 0.0006390886846929789
step: 30, loss: 0.0003945661010220647
step: 40, loss: 0.002313872566446662
step: 50, loss: 0.0013107710983604193
step: 60, loss: 0.0007109853904694319
step: 70, loss: 0.007469384931027889
step: 80, loss: 0.0003715681959874928
step: 90, loss: 0.0025387024506926537
step: 100, loss: 0.00035657850094139576
step: 110, loss: 0.002026601927354932
step: 120, loss: 0.00025448130327276886
step: 130, loss: 0.00025130168069154024
step: 140, loss: 0.00035632491926662624
step: 150, loss: 0.0012094430858269334
step: 160, loss: 0.00015479032299481332
step: 170, loss: 0.00018166658992413431
step: 180, loss: 0.001060948707163334
step: 190, loss: 0.0004525747208390385
step: 200, loss: 0.0008627531351521611
step: 210, loss: 0.01614469662308693
step: 220, loss: 0.10343147814273834
step: 230, loss: 0.0006436426192522049
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
epoch 10: dev_f1=0.9888641425389755, f1=0.9724366041896362,
best_f1=0.9864864864864865
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0005545876338146627
step: 10, loss: 0.0008922313572838902
step: 20, loss: 0.015548458322882652
step: 30, loss: 0.0003159425687044859
step: 40, loss: 0.00019484107906464487
step: 50, loss: 0.0015679559437558055
step: 60, loss: 0.02051178552210331
step: 70, loss: 0.008949113078415394
step: 80, loss: 0.0022387930657714605
step: 90, loss: 0.1555340439081192
step: 100, loss: 0.0004863126087002456
step: 110, loss: 0.0006234395550563931
step: 120, loss: 0.0006117299199104309
step: 130, loss: 0.00043046247446909547
step: 140, loss: 0.018112413585186005
step: 150, loss: 0.001061095972545445

```

```

step: 160, loss: 0.02290591411292553
step: 170, loss: 0.004224915523082018
step: 180, loss: 0.0007926127873361111
step: 190, loss: 0.0005112451035529375
step: 200, loss: 0.01355063822120428
step: 210, loss: 0.00025782122975215316
step: 220, loss: 0.015372851863503456
step: 230, loss: 0.00025691092014312744
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 11: dev_f1=0.987736900780379, f1=0.9810479375696767,
best_f1=0.9864864864864865
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00038331825635395944
step: 10, loss: 0.0002626294444780797
step: 20, loss: 0.018021192401647568
step: 30, loss: 0.0519116185605526
step: 40, loss: 0.01188963558524847
step: 50, loss: 0.0016869157552719116
step: 60, loss: 0.0012213176814839244
step: 70, loss: 0.0004427306412253529
step: 80, loss: 0.0001413205754943192
step: 90, loss: 0.0018985919887199998
step: 100, loss: 0.00023821511422283947
step: 110, loss: 0.0001933345920406282
step: 120, loss: 0.00036675765295512974
step: 130, loss: 0.0002783126838039607
step: 140, loss: 0.0007157988729886711
step: 150, loss: 0.0003477303544059396
step: 160, loss: 0.00032888533314689994
step: 170, loss: 0.0004886134993284941
step: 180, loss: 0.00032310065580531955
step: 190, loss: 0.014188751578330994
step: 200, loss: 0.0002230829995824024
step: 210, loss: 0.0040044705383479595
step: 220, loss: 0.01654999330639839
step: 230, loss: 0.0004574921913444996
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 12: dev_f1=0.9876543209876544, f1=0.9887387387387387,
best_f1=0.9864864864864865
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")

```



```

step: 0, loss: 0.0014438460348173976
step: 10, loss: 0.0003135956358164549
step: 20, loss: 0.0008405527332797647
step: 30, loss: 0.0014576561516150832
step: 40, loss: 0.0015699687646701932
step: 50, loss: 0.0042252084240317345
step: 60, loss: 0.0005432268371805549
step: 70, loss: 0.0012712812749668956
step: 80, loss: 0.0006415712414309382
step: 90, loss: 0.0005177650600671768
step: 100, loss: 0.001475082361139357
step: 110, loss: 0.0008020942914299667
step: 120, loss: 0.00040865884511731565
step: 130, loss: 0.00035577788366936147
step: 140, loss: 0.00037492942647077143
step: 150, loss: 0.00014294673746917397
step: 160, loss: 0.024740882217884064
step: 170, loss: 0.00019844585040118545
step: 180, loss: 0.07588525861501694
step: 190, loss: 0.0005595365073531866
step: 200, loss: 9.998896712204441e-05
step: 210, loss: 0.0004630813782569021
step: 220, loss: 0.0004182550183031708
step: 230, loss: 0.0003206853580195457
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
epoch 13: dev_f1=0.9899441340782122, f1=0.9843400447427293,
best_f1=0.9864864864864865
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00023358987527899444
step: 10, loss: 0.00019644170242827386
step: 20, loss: 0.00021714599279221147
step: 30, loss: 0.0005016314098611474
step: 40, loss: 0.0002065735898213461
step: 50, loss: 0.00014705942885484546
step: 60, loss: 0.0002153892710339278
step: 70, loss: 0.00022806394554208964
step: 80, loss: 0.00029816656024195254
step: 90, loss: 0.0007311016088351607
step: 100, loss: 0.0006830907659605145
step: 110, loss: 0.0003219421487301588
step: 120, loss: 7.459561311407015e-05
step: 130, loss: 0.0008099743281491101
step: 140, loss: 0.0004843317437916994
step: 150, loss: 0.0011621437733992934

```

```

step: 160, loss: 0.0006769965984858572
step: 170, loss: 0.00031035239226184785
step: 180, loss: 0.0001932236336870119
step: 190, loss: 0.0004070275754202157
step: 200, loss: 0.0005370063590817153
step: 210, loss: 0.00025211251340806484
step: 220, loss: 0.0002568960189819336
step: 230, loss: 0.0004553057951852679
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 14: dev_f1=0.9899441340782122, f1=0.9843400447427293,
best_f1=0.9864864864864865
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.03374123573303223
step: 10, loss: 0.00021034496603533626
step: 20, loss: 0.0003968041273765266
step: 30, loss: 0.0011040880344808102
step: 40, loss: 0.00012794544454663992
step: 50, loss: 0.00014281843323260546
step: 60, loss: 0.022369341924786568
step: 70, loss: 0.0196545347571373
step: 80, loss: 0.000329464121023193
step: 90, loss: 0.031638436019420624
step: 100, loss: 9.11760944291018e-05
step: 110, loss: 0.0027386611327528954
step: 120, loss: 0.03284180536866188
step: 130, loss: 0.00021415117953438312
step: 140, loss: 0.01177544891834259
step: 150, loss: 0.00031761228456161916
step: 160, loss: 0.0143959429115057
step: 170, loss: 0.0003313203633297235
step: 180, loss: 0.00018820134573616087
step: 190, loss: 0.002060250611975789
step: 200, loss: 0.0019731619395315647
step: 210, loss: 0.03206922486424446
step: 220, loss: 0.00016830419190227985
step: 230, loss: 0.00039425131399184465
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 15: dev_f1=0.9887892376681614, f1=0.9887387387387387,
best_f1=0.9864864864864865
mv: cannot stat '*_dev.pt': No such file or directory
Selected optimization level O2:  FP16 training with FP32 batchnorm and FP32
master weights.

```

```

Defaults for this optimization level are:
enabled                : True
opt_level              : 02
cast_model_type        : torch.float16
patch_torch_functions  : False
keep_batchnorm_fp32    : True
master_weights         : True
loss_scale             : dynamic
Processing user overrides (additional kwargs that are not None)...
After processing overrides, optimization options are:
enabled                : True
opt_level              : 02
cast_model_type        : torch.float16
patch_torch_functions  : False
keep_batchnorm_fp32    : True
master_weights         : True
loss_scale             : dynamic
Warning: multi_tensor_applier fused unscale kernel is unavailable, possibly
because apex was installed without --cuda_ext --cpp_ext. Using Python fallback.
Original ImportError was: ModuleNotFoundError("No module named 'amp_C'")
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
2473it [00:15, 155.31it/s]
load_f1 = 0.9864253393665158
real_f1 = 0.9842342342342343
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
4398it [00:31, 140.10it/s]

```

### 3.1.3 DBLP-GoogleScholar - Running the matcher

```

[ ]: # Step 1:
# train the wdc model with the save_model flag on
# it will produce two files *_dev.pt and *_test.pt which are the
# best checkpoints on the dev set and the test set.
!CUDA_VISIBLE_DEVICES=0 python train_ditto.py \
  --task Structured/DBLP-GoogleScholar \
  --batch_size 32 \
  --max_len 256 \
  --lr 3e-5 \

```

```

--n_epochs 15 \
--finetuning \
--lm roberta \
--fp16 \
--save_model \
--dk general \
--dk product \
--summarize

# Step 2: pick a checkpoint, rename it to wdc_all_small.pt
!mv *_dev.pt checkpoints/DBLP-GoogleScholar.pt

# Step 3: run the matcher
!CUDA_VISIBLE_DEVICES=0 python matcher.py \
  --task Structured/DBLP-GoogleScholar \
  --input_path input/input_small.jsonl \
  --output_path output/output_small.jsonl \
  --lm roberta \
  --use_gpu \
  --fp16 \
  --checkpoint_path checkpoints/

```

Selected optimization level 02: FP16 training with FP32 batchnorm and FP32 master weights.

Defaults for this optimization level are:

```

enabled          : True
opt_level        : 02
cast_model_type  : torch.float16
patch_torch_functions : False
keep_batchnorm_fp32 : True
master_weights   : True
loss_scale       : dynamic

```

Processing user overrides (additional kwargs that are not None)...

After processing overrides, optimization options are:

```

enabled          : True
opt_level        : 02
cast_model_type  : torch.float16
patch_torch_functions : False
keep_batchnorm_fp32 : True
master_weights   : True
loss_scale       : dynamic

```

Warning: multi\_tensor\_applier fused unscale kernel is unavailable, possibly because apex was installed without --cuda\_ext --cpp\_ext. Using Python fallback.

Original ImportError was: ModuleNotFoundError("No module named 'amp\_C'")

/usr/local/lib/python3.7/dist-packages/apex/amp/\_initialize.py:25: UserWarning:

An input tensor was not cuda.

```
warnings.warn("An input tensor was not cuda.")
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 32768.0
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:136:
UserWarning: Detected call of `lr_scheduler.step()` before `optimizer.step()`.
In PyTorch 1.1.0 and later, you should call them in the opposite order:
`optimizer.step()` before `lr_scheduler.step()`. Failure to do this will result
in PyTorch skipping the first value of the learning rate schedule. See more
details at https://pytorch.org/docs/stable/optim.html#how-to-adjust-learning-
rate
```

```
"https://pytorch.org/docs/stable/optim.html#how-to-adjust-learning-rate",
UserWarning)
step: 0, loss: 0.6775714755058289
step: 10, loss: 0.449269562959671
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 20, loss: 0.27703484892845154
step: 30, loss: 0.36706626415252686
step: 40, loss: 0.26585251092910767
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 8192.0
step: 50, loss: 0.21886998414993286
step: 60, loss: 0.1553020030260086
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 4096.0
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 2048.0
step: 70, loss: 0.3159073293209076
step: 80, loss: 0.14651434123516083
step: 90, loss: 0.13455165922641754
step: 100, loss: 0.3153856098651886
step: 110, loss: 0.08482921868562698
step: 120, loss: 0.21351060271263123
step: 130, loss: 0.08873189985752106
step: 140, loss: 0.37729334831237793
step: 150, loss: 0.10476398468017578
step: 160, loss: 0.17170748114585876
step: 170, loss: 0.09322548657655716
step: 180, loss: 0.030849670991301537
step: 190, loss: 0.03647378087043762
step: 200, loss: 0.030584778636693954
step: 210, loss: 0.03567684814333916
step: 220, loss: 0.08798112720251083
step: 230, loss: 0.2040260285139084
step: 240, loss: 0.055502235889434814
step: 250, loss: 0.03352527320384979
step: 260, loss: 0.05450577661395073
step: 270, loss: 0.31967148184776306
step: 280, loss: 0.03456927090883255
step: 290, loss: 0.058662187308073044
step: 300, loss: 0.009540271013975143
step: 310, loss: 0.050187986344099045
```

```

step: 320, loss: 0.09555114805698395
step: 330, loss: 0.1968139111995697
step: 340, loss: 0.4342917203903198
step: 350, loss: 0.05598962679505348
step: 360, loss: 0.10461142659187317
step: 370, loss: 0.0293413158506155
step: 380, loss: 0.1381348967552185
step: 390, loss: 0.05644365772604942
step: 400, loss: 0.0755830854177475
step: 410, loss: 0.24522894620895386
step: 420, loss: 0.019730372354388237
step: 430, loss: 0.018375366926193237
step: 440, loss: 0.010278520174324512
step: 450, loss: 0.04656166210770607
step: 460, loss: 0.008283083327114582
step: 470, loss: 0.016593072563409805
step: 480, loss: 0.17563143372535706
step: 490, loss: 0.07435767352581024
step: 500, loss: 0.05492879077792168
step: 510, loss: 0.0327971912920475
step: 520, loss: 0.055218182504177094
step: 530, loss: 0.030617229640483856
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
    warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 1: dev_f1=0.9456721915285451, f1=0.9450346420323326,
best_f1=0.9450346420323326
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.09810837358236313
step: 10, loss: 0.0630357638001442
step: 20, loss: 0.011606965214014053
step: 30, loss: 0.08286548405885696
step: 40, loss: 0.056020718067884445
step: 50, loss: 0.028846316039562225
step: 60, loss: 0.04725749418139458
step: 70, loss: 0.025641046464443207
step: 80, loss: 0.02496451511979103
step: 90, loss: 0.004877262283116579
step: 100, loss: 0.20900489389896393
step: 110, loss: 0.013948453590273857
step: 120, loss: 0.03552059829235077
step: 130, loss: 0.004002220928668976

```

```

step: 140, loss: 0.07604458183050156
step: 150, loss: 0.007977311499416828
step: 160, loss: 0.025884905830025673
step: 170, loss: 0.011888443492352962
step: 180, loss: 0.042288243770599365
step: 190, loss: 0.008607610128819942
step: 200, loss: 0.2558286488056183
step: 210, loss: 0.01560529787093401
step: 220, loss: 0.001144877402111888
step: 230, loss: 0.10459515452384949
step: 240, loss: 0.04194289445877075
step: 250, loss: 0.04571183770895004
step: 260, loss: 0.07986241579055786
step: 270, loss: 0.004272874910384417
step: 280, loss: 0.026753772050142288
step: 290, loss: 0.04554508626461029
step: 300, loss: 0.05947870388627052
step: 310, loss: 0.07534275203943253
step: 320, loss: 0.10419599711894989
step: 330, loss: 0.02218678593635559
step: 340, loss: 0.06233440339565277
step: 350, loss: 0.001307959551922977
step: 360, loss: 0.11654495447874069
step: 370, loss: 0.0060955556109547615
step: 380, loss: 0.16446875035762787
step: 390, loss: 0.017481116577982903
step: 400, loss: 0.009268097579479218
step: 410, loss: 0.006511148065328598
step: 420, loss: 0.15487410128116608
step: 430, loss: 0.2338123470544815
step: 440, loss: 0.01522996835410595
step: 450, loss: 0.008293935097754002
step: 460, loss: 0.0199239794164896
step: 470, loss: 0.05079210177063942
step: 480, loss: 0.021326513960957527
step: 490, loss: 0.07257412374019623
step: 500, loss: 0.0017918695230036974
step: 510, loss: 0.013522895984351635
step: 520, loss: 0.15545888245105743
step: 530, loss: 0.0796181857585907
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
epoch 2: dev_f1=0.9443938012762079, f1=0.9524245196706312,
best_f1=0.9450346420323326
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")

```

step: 0, loss: 0.06921159476041794  
step: 10, loss: 0.05637933313846588  
step: 20, loss: 0.006347017828375101  
step: 30, loss: 0.03039713390171528  
step: 40, loss: 0.06238972768187523  
step: 50, loss: 0.0060856458730995655  
step: 60, loss: 0.009730132296681404  
step: 70, loss: 0.005740143358707428  
step: 80, loss: 0.07586688548326492  
step: 90, loss: 0.03230586275458336  
step: 100, loss: 0.03500644490122795  
step: 110, loss: 0.03007364831864834  
step: 120, loss: 0.08902011066675186  
step: 130, loss: 0.021245259791612625  
step: 140, loss: 0.012843374162912369  
step: 150, loss: 0.010255493223667145  
step: 160, loss: 0.025782285258173943  
step: 170, loss: 0.0006896283593960106  
step: 180, loss: 0.010289070196449757  
step: 190, loss: 0.0013341447338461876  
step: 200, loss: 0.008803170174360275  
step: 210, loss: 0.0541258342564106  
step: 220, loss: 0.09181491285562515  
step: 230, loss: 0.021182281896471977  
step: 240, loss: 0.07648244500160217  
step: 250, loss: 0.042386844754219055  
step: 260, loss: 0.13881194591522217  
step: 270, loss: 0.0019812840037047863  
step: 280, loss: 0.005960775539278984  
step: 290, loss: 0.0009406442986801267  
step: 300, loss: 0.1318594515323639  
step: 310, loss: 0.049931708723306656  
step: 320, loss: 0.0235524233430624  
step: 330, loss: 0.0020574985537678003  
step: 340, loss: 0.008481154218316078  
step: 350, loss: 0.09126932173967361  
step: 360, loss: 0.004183666314929724  
step: 370, loss: 0.006780591793358326  
step: 380, loss: 0.006729988381266594  
step: 390, loss: 0.004826351534575224  
step: 400, loss: 0.1299617737531662  
step: 410, loss: 0.05564441159367561  
step: 420, loss: 0.006151057314127684  
step: 430, loss: 0.020650990307331085  
step: 440, loss: 0.10636912286281586  
step: 450, loss: 0.010157179087400436  
step: 460, loss: 0.030637376010417938  
step: 470, loss: 0.021519428119063377



```

step: 480, loss: 0.0808047279715538
step: 490, loss: 0.0480131171643734
step: 500, loss: 0.08395352959632874
step: 510, loss: 0.013792045414447784
step: 520, loss: 0.013183106668293476
step: 530, loss: 0.0047925496473908424
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
    warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 3: dev_f1=0.9524688509460083, f1=0.9483870967741935,
best_f1=0.9483870967741935
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00911211408674717
step: 10, loss: 0.0032302651088684797
step: 20, loss: 0.14758971333503723
step: 30, loss: 0.11898619681596756
step: 40, loss: 0.016124015673995018
step: 50, loss: 0.09340111911296844
step: 60, loss: 0.0008223954355344176
step: 70, loss: 0.05783393979072571
step: 80, loss: 0.12821193039417267
step: 90, loss: 0.010941939428448677
step: 100, loss: 0.001999747008085251
step: 110, loss: 0.012150303460657597
step: 120, loss: 0.0026922973338514566
step: 130, loss: 0.0372530035674572
step: 140, loss: 0.123703233897686
step: 150, loss: 0.0015817716484889388
step: 160, loss: 0.002415415132418275
step: 170, loss: 0.020819000899791718
step: 180, loss: 0.08033502846956253
step: 190, loss: 0.06241115182638168
step: 200, loss: 0.057877637445926666
step: 210, loss: 0.0012928301002830267
step: 220, loss: 0.03639798238873482
step: 230, loss: 0.05142475292086601
step: 240, loss: 0.04494553431868553
step: 250, loss: 0.10084488242864609
step: 260, loss: 0.006895888596773148
step: 270, loss: 0.06869551539421082
step: 280, loss: 0.021118097007274628
step: 290, loss: 0.018794333562254906

```

```

step: 300, loss: 0.0015766246942803264
step: 310, loss: 0.0006563930073752999
step: 320, loss: 0.2075570821762085
step: 330, loss: 0.03798667713999748
step: 340, loss: 0.012055127881467342
step: 350, loss: 0.22798942029476166
step: 360, loss: 0.022287102416157722
step: 370, loss: 0.0017284760251641273
step: 380, loss: 0.0031715910881757736
step: 390, loss: 0.0013897718163207173
step: 400, loss: 0.004282000940293074
step: 410, loss: 0.004679826553910971
step: 420, loss: 0.005241666920483112
step: 430, loss: 0.015618884935975075
step: 440, loss: 0.004203180316835642
step: 450, loss: 0.017689174041152
step: 460, loss: 0.009082941338419914
step: 470, loss: 0.002286497037857771
step: 480, loss: 0.002164932433515787
step: 490, loss: 0.0010369938099756837
step: 500, loss: 0.0182415209710598
step: 510, loss: 0.012988460250198841
step: 520, loss: 0.006775032263249159
step: 530, loss: 0.11703415215015411
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
    warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 4: dev_f1=0.9556282111163008, f1=0.9496944052656323,
best_f1=0.9496944052656323
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0030435046646744013
step: 10, loss: 0.026454158127307892
step: 20, loss: 0.008385528810322285
step: 30, loss: 0.0046784826554358006
step: 40, loss: 0.0012625291710719466
step: 50, loss: 0.010594228282570839
step: 60, loss: 0.033589720726013184
step: 70, loss: 0.001913595013320446
step: 80, loss: 0.00036962691228836775
step: 90, loss: 0.05616912618279457
step: 100, loss: 0.0051515959203243256
step: 110, loss: 0.005329690873622894

```

```
step: 120, loss: 0.012434477917850018
step: 130, loss: 0.007941946387290955
step: 140, loss: 0.00568993529304862
step: 150, loss: 0.018932167440652847
step: 160, loss: 0.007605873979628086
step: 170, loss: 0.07965158671140671
step: 180, loss: 0.0027565513737499714
step: 190, loss: 0.010531620122492313
step: 200, loss: 0.0024961894378066063
step: 210, loss: 0.0016474478179588914
step: 220, loss: 0.004637460689991713
step: 230, loss: 0.01850535161793232
step: 240, loss: 0.0007972351158969104
step: 250, loss: 0.16596311330795288
step: 260, loss: 0.05274005979299545
step: 270, loss: 0.014358457177877426
step: 280, loss: 0.012029866687953472
step: 290, loss: 0.07045725733041763
step: 300, loss: 0.0006836290122009814
step: 310, loss: 0.004920268431305885
step: 320, loss: 0.011062243022024632
step: 330, loss: 0.0008053545607253909
step: 340, loss: 0.0005136673571541905
step: 350, loss: 0.00014043044939171523
step: 360, loss: 0.00031777526601217687
step: 370, loss: 0.00011370173160685226
step: 380, loss: 9.227003465639427e-05
step: 390, loss: 0.00013936107279732823
step: 400, loss: 0.009073423221707344
step: 410, loss: 0.008130544796586037
step: 420, loss: 0.18449766933918
step: 430, loss: 0.050303030759096146
step: 440, loss: 0.0005343296215869486
step: 450, loss: 0.0066834804601967335
step: 460, loss: 0.013622349128127098
step: 470, loss: 0.014500060118734837
step: 480, loss: 0.028180541470646858
step: 490, loss: 0.03058738447725773
step: 500, loss: 0.009491863660514355
step: 510, loss: 0.008703119121491909
step: 520, loss: 0.0625351220369339
step: 530, loss: 0.07660859078168869
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
epoch 5: dev_f1=0.9555867227676483, f1=0.9555451567618156,
best_f1=0.9496944052656323
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
```

An input tensor was not cuda.

```
warnings.warn("An input tensor was not cuda.")
```

```
step: 0, loss: 0.022222543135285378
step: 10, loss: 0.0012354404898360372
step: 20, loss: 0.0003991265839431435
step: 30, loss: 0.0030671569984406233
step: 40, loss: 0.0027047311887145042
step: 50, loss: 0.000626762630417943
step: 60, loss: 0.012814801186323166
step: 70, loss: 0.0014727362431585789
step: 80, loss: 0.0009069137158803642
step: 90, loss: 0.0014700442552566528
step: 100, loss: 0.001081937924027443
step: 110, loss: 0.0008940811385400593
step: 120, loss: 0.01124727725982666
step: 130, loss: 0.0027386480942368507
step: 140, loss: 0.0035494097974151373
step: 150, loss: 0.0010510531719774008
step: 160, loss: 0.02805536799132824
step: 170, loss: 0.0010087387636303902
step: 180, loss: 0.0005655468557961285
step: 190, loss: 0.0010170461609959602
step: 200, loss: 0.006322277709841728
step: 210, loss: 0.0015347234439104795
step: 220, loss: 0.000666979409288615
step: 230, loss: 0.0026670778170228004
step: 240, loss: 0.02507805824279785
step: 250, loss: 0.08292016386985779
step: 260, loss: 0.0028178521897643805
step: 270, loss: 0.0012181438505649567
step: 280, loss: 0.0006983507191762328
step: 290, loss: 0.02437603659927845
step: 300, loss: 0.01262535247951746
step: 310, loss: 0.05484820157289505
step: 320, loss: 5.637907815980725e-05
step: 330, loss: 0.02969811111688614
step: 340, loss: 0.00015257037011906505
step: 350, loss: 0.005213508382439613
step: 360, loss: 0.003021489828824997
step: 370, loss: 0.002798601519316435
step: 380, loss: 0.0005216128774918616
step: 390, loss: 0.002148787723854184
step: 400, loss: 0.01596560701727867
step: 410, loss: 0.00011333443399053067
step: 420, loss: 0.034542981535196304
step: 430, loss: 0.014540894888341427
step: 440, loss: 0.008177137933671474
step: 450, loss: 0.2844940721988678
```

```
step: 460, loss: 0.010237907990813255
step: 470, loss: 0.007119993679225445
step: 480, loss: 0.0038456497713923454
step: 490, loss: 0.021241992712020874
step: 500, loss: 0.0006762293050996959
step: 510, loss: 0.037120237946510315
step: 520, loss: 0.012513192370533943
step: 530, loss: 0.0038367293309420347
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 6: dev_f1=0.9543336439888165, f1=0.9541454377026402,
best_f1=0.9496944052656323
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0023759030736982822
step: 10, loss: 0.000920497695915401
step: 20, loss: 0.004723466467112303
step: 30, loss: 0.0033888397738337517
step: 40, loss: 0.011521675623953342
step: 50, loss: 0.0012933460529893637
step: 60, loss: 0.001765513326972723
step: 70, loss: 0.0002049884933512658
step: 80, loss: 0.0007434629951603711
step: 90, loss: 0.00021086388733237982
step: 100, loss: 0.04392366111278534
step: 110, loss: 9.398609836352989e-05
step: 120, loss: 0.00019212921324651688
step: 130, loss: 0.0001136041828431189
step: 140, loss: 0.00025462982011958957
step: 150, loss: 0.00041470746509730816
step: 160, loss: 5.9392568800831214e-05
step: 170, loss: 0.0027638126630336046
step: 180, loss: 0.0011979510309174657
step: 190, loss: 0.1325436681509018
step: 200, loss: 0.00036114052636548877
step: 210, loss: 0.008458994328975677
step: 220, loss: 0.0005288011743687093
step: 230, loss: 0.00016169979062397033
step: 240, loss: 0.0008665717905387282
step: 250, loss: 0.04422271251678467
step: 260, loss: 0.009584946557879448
step: 270, loss: 9.437735570827499e-05
step: 280, loss: 0.015546075068414211
step: 290, loss: 0.002071881666779518
step: 300, loss: 0.00010468198888702318
step: 310, loss: 0.0008435624768026173
```

```

step: 320, loss: 0.0007637631497345865
step: 330, loss: 0.00034601788502186537
step: 340, loss: 0.0008113629883155227
step: 350, loss: 0.0026489754673093557
step: 360, loss: 0.008283625356853008
step: 370, loss: 0.02939814329147339
step: 380, loss: 0.00346759264357388
step: 390, loss: 0.020030638203024864
step: 400, loss: 0.06117437034845352
step: 410, loss: 0.003672547172755003
step: 420, loss: 0.16525740921497345
step: 430, loss: 0.003939103335142136
step: 440, loss: 0.0024365403223782778
step: 450, loss: 0.0004921308718621731
step: 460, loss: 0.029830552637577057
step: 470, loss: 0.14912042021751404
step: 480, loss: 0.00394086679443717
step: 490, loss: 0.0029408454429358244
step: 500, loss: 0.0015594459837302566
step: 510, loss: 0.0006674730102531612
step: 520, loss: 0.0006544640054926276
step: 530, loss: 0.0019758727867156267
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
    warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 7: dev_f1=0.9570041608876559, f1=0.9536178107606679,
best_f1=0.9536178107606679
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0007092467276379466
step: 10, loss: 0.0019427456427365541
step: 20, loss: 0.001370061538182199
step: 30, loss: 0.0010444812942296267
step: 40, loss: 0.000710940919816494
step: 50, loss: 0.0010459369514137506
step: 60, loss: 0.0006638591294176877
step: 70, loss: 0.00579700106754899
step: 80, loss: 0.007509931921958923
step: 90, loss: 0.0007865577936172485
step: 100, loss: 0.00016152004536706954
step: 110, loss: 0.00021053693490102887
step: 120, loss: 0.00029531429754570127
step: 130, loss: 0.0002626087225507945

```

```

step: 140, loss: 0.0001360632450086996
step: 150, loss: 0.0002726692473515868
step: 160, loss: 0.0009099793969653547
step: 170, loss: 0.00020154182857368141
step: 180, loss: 0.00018070806981995702
step: 190, loss: 0.005932454951107502
step: 200, loss: 0.00034353043884038925
step: 210, loss: 0.07584492117166519
step: 220, loss: 8.861599053489044e-05
step: 230, loss: 0.10975994914770126
step: 240, loss: 0.002395962830632925
step: 250, loss: 0.0015325959539040923
step: 260, loss: 0.004606103990226984
step: 270, loss: 0.004748022183775902
step: 280, loss: 0.0007713073282502592
step: 290, loss: 0.005806860513985157
step: 300, loss: 0.00014537187234964222
step: 310, loss: 0.012187885120511055
step: 320, loss: 0.008370415307581425
step: 330, loss: 0.00025279284454882145
step: 340, loss: 0.1341787725687027
step: 350, loss: 0.0005468878080137074
step: 360, loss: 0.00017460979870520532
step: 370, loss: 0.03113444149494171
step: 380, loss: 0.00018037502013612539
step: 390, loss: 0.006833519786596298
step: 400, loss: 0.0012647216208279133
step: 410, loss: 0.0002697092422749847
step: 420, loss: 0.00014718080637976527
step: 430, loss: 0.0020679777953773737
step: 440, loss: 0.004155788104981184
step: 450, loss: 0.0006863567978143692
step: 460, loss: 0.005142324138432741
step: 470, loss: 0.12730807065963745
step: 480, loss: 0.03659738600254059
step: 490, loss: 0.025076141580939293
step: 500, loss: 0.001447100192308426
step: 510, loss: 0.018460214138031006
step: 520, loss: 0.0004063313826918602
step: 530, loss: 0.00015972372784744948
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
  warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 8: dev_f1=0.9574567554932211, f1=0.9579831932773109,

```

```
best_f1=0.9579831932773109
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.000527264317497611
step: 10, loss: 0.006772484164685011
step: 20, loss: 0.0035807581152766943
step: 30, loss: 0.13800528645515442
step: 40, loss: 0.0013169944286346436
step: 50, loss: 0.0008254349231719971
step: 60, loss: 0.0009848104091361165
step: 70, loss: 0.031070010736584663
step: 80, loss: 0.004031067248433828
step: 90, loss: 0.06832918524742126
step: 100, loss: 0.0012820269912481308
step: 110, loss: 0.001632659463211894
step: 120, loss: 0.000665434286929667
step: 130, loss: 0.0012531171087175608
step: 140, loss: 0.00016312238585669547
step: 150, loss: 0.0009711540187709033
step: 160, loss: 0.00014095680671744049
step: 170, loss: 0.0013650296023115516
step: 180, loss: 0.00014486920554190874
step: 190, loss: 6.510750972665846e-05
step: 200, loss: 5.6508364650653675e-05
step: 210, loss: 0.0001447330432711169
step: 220, loss: 0.035621531307697296
step: 230, loss: 0.0010565727716311812
step: 240, loss: 0.00026824692031368613
step: 250, loss: 0.00037415348924696445
step: 260, loss: 0.0006888009957037866
step: 270, loss: 0.00016311787476297468
step: 280, loss: 0.019379019737243652
step: 290, loss: 0.00032413104781880975
step: 300, loss: 0.000985861523076892
step: 310, loss: 0.06670594215393066
step: 320, loss: 0.0020162484142929316
step: 330, loss: 0.00018970068776980042
step: 340, loss: 0.0009389750775881112
step: 350, loss: 0.013419106602668762
step: 360, loss: 0.0028044863138347864
step: 370, loss: 4.877408719039522e-05
step: 380, loss: 0.0008875321364030242
step: 390, loss: 6.744213897036389e-05
step: 400, loss: 0.18487192690372467
step: 410, loss: 0.0018797909142449498
step: 420, loss: 0.00027231330750510097
step: 430, loss: 0.0056131272576749325
```



```
step: 440, loss: 9.534729178994894e-05
step: 450, loss: 0.0017026993446052074
step: 460, loss: 6.166035745991394e-05
step: 470, loss: 0.0004930214490741491
step: 480, loss: 4.053381780977361e-05
step: 490, loss: 0.009498215280473232
step: 500, loss: 0.0005605333717539907
step: 510, loss: 0.040494613349437714
step: 520, loss: 0.00398498447611928
step: 530, loss: 0.09619572758674622
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 9: dev_f1=0.9557933922754769, f1=0.9504672897196261,
best_f1=0.9579831932773109
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.01048320159316063
step: 10, loss: 0.0012065389892086387
step: 20, loss: 6.640858191531152e-05
step: 30, loss: 0.0010574753396213055
step: 40, loss: 5.0505179387982935e-05
step: 50, loss: 0.00017689255764707923
step: 60, loss: 0.01922173611819744
step: 70, loss: 4.855017323279753e-05
step: 80, loss: 0.0030990741215646267
step: 90, loss: 8.584026363678277e-05
step: 100, loss: 0.0023262896575033665
step: 110, loss: 0.004465123172849417
step: 120, loss: 0.0009413861553184688
step: 130, loss: 0.0527384914457798
step: 140, loss: 0.0002962627331726253
step: 150, loss: 0.001052394392900169
step: 160, loss: 0.06568920612335205
step: 170, loss: 0.00010413491690997034
step: 180, loss: 0.0006116244476288557
step: 190, loss: 7.996409112820402e-05
step: 200, loss: 0.0031383312307298183
step: 210, loss: 0.012952612712979317
step: 220, loss: 0.002074482850730419
step: 230, loss: 0.0002139749121852219
step: 240, loss: 0.0006849697092548013
step: 250, loss: 0.0026544046122580767
step: 260, loss: 0.0002878204395528883
step: 270, loss: 0.00020251887326594442
step: 280, loss: 0.027232080698013306
step: 290, loss: 0.00012819255061913282
```

```

step: 300, loss: 0.000269292009761557
step: 310, loss: 0.05909416079521179
step: 320, loss: 0.010164095088839531
step: 330, loss: 0.0013343057362362742
step: 340, loss: 0.0008323509246110916
step: 350, loss: 0.000857024744618684
step: 360, loss: 8.622064342489466e-05
step: 370, loss: 0.00105995312333107
step: 380, loss: 0.00211132294498384
step: 390, loss: 0.00011216674465686083
step: 400, loss: 0.010883253067731857
step: 410, loss: 9.148505341727287e-05
step: 420, loss: 3.8613277865806594e-05
step: 430, loss: 0.00547782564535737
step: 440, loss: 5.966362732579e-05
step: 450, loss: 0.0008468531304970384
step: 460, loss: 9.901041630655527e-05
step: 470, loss: 0.001471252995543182
step: 480, loss: 0.0003581551427487284
step: 490, loss: 0.0008438496734015644
step: 500, loss: 0.0006142658530734479
step: 510, loss: 0.0001440163905499503
step: 520, loss: 0.0045981756411492825
step: 530, loss: 0.0031095861922949553
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
    warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 10: dev_f1=0.9596662030598052, f1=0.9578508568781844,
best_f1=0.9578508568781844
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 4.197437738184817e-05
step: 10, loss: 0.00012731280003208667
step: 20, loss: 0.00031342398142442107
step: 30, loss: 0.0004190107574686408
step: 40, loss: 3.776845915126614e-05
step: 50, loss: 4.905207606498152e-05
step: 60, loss: 0.0003100900794379413
step: 70, loss: 4.778135917149484e-05
step: 80, loss: 0.0008289095712825656
step: 90, loss: 0.0008228371734730899
step: 100, loss: 0.00016361792222596705
step: 110, loss: 5.749591946369037e-05

```

```
step: 120, loss: 0.00013347815547604114
step: 130, loss: 0.00017287924129050225
step: 140, loss: 0.007050833199173212
step: 150, loss: 0.00038372029666788876
step: 160, loss: 0.00023582001449540257
step: 170, loss: 0.00015294563490897417
step: 180, loss: 0.00019486644305288792
step: 190, loss: 0.0014922372065484524
step: 200, loss: 0.00023077664081938565
step: 210, loss: 0.0003280541568528861
step: 220, loss: 0.029425600543618202
step: 230, loss: 2.6004328901763074e-05
step: 240, loss: 0.002721017925068736
step: 250, loss: 0.0003076442517340183
step: 260, loss: 0.0057882764376699924
step: 270, loss: 0.00200277310796082
step: 280, loss: 0.00034477884764783084
step: 290, loss: 0.00017083215061575174
step: 300, loss: 9.903869067784399e-05
step: 310, loss: 0.00021129593369551003
step: 320, loss: 0.0007696214015595615
step: 330, loss: 6.463967292802408e-05
step: 340, loss: 0.000170059414813295
step: 350, loss: 0.0006906230119057
step: 360, loss: 9.253079770132899e-05
step: 370, loss: 0.002609682036563754
step: 380, loss: 0.0030597366858273745
step: 390, loss: 0.0031820847652852535
step: 400, loss: 0.06467800587415695
step: 410, loss: 0.00163236353546381
step: 420, loss: 0.0003419088898226619
step: 430, loss: 0.0017772025894373655
step: 440, loss: 0.0018707925919443369
step: 450, loss: 0.002329602837562561
step: 460, loss: 0.003964666742831469
step: 470, loss: 0.0003899637085851282
step: 480, loss: 0.0005564295570366085
step: 490, loss: 5.640463132294826e-05
step: 500, loss: 0.0011150692589581013
step: 510, loss: 0.005213841795921326
step: 520, loss: 5.048726961831562e-05
step: 530, loss: 0.0018010472413152456
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
epoch 11: dev_f1=0.9592887225081891, f1=0.9514837494112105,
best_f1=0.9578508568781844
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
```

An input tensor was not cuda.

```
warnings.warn("An input tensor was not cuda.")
```

```
step: 0, loss: 0.002714832779020071
step: 10, loss: 4.1652179788798094e-05
step: 20, loss: 0.006096032913774252
step: 30, loss: 0.000504227529745549
step: 40, loss: 0.0005110796773806214
step: 50, loss: 4.592726691043936e-05
step: 60, loss: 4.459418778424151e-05
step: 70, loss: 0.0004697665572166443
step: 80, loss: 3.6557896237354726e-05
step: 90, loss: 0.00018210598500445485
step: 100, loss: 0.2145005613565445
step: 110, loss: 0.0001972866739379242
step: 120, loss: 0.0002435580245219171
step: 130, loss: 0.00104712531901896
step: 140, loss: 0.00022346334299072623
step: 150, loss: 0.00014968292089179158
step: 160, loss: 0.00024197516904678196
step: 170, loss: 9.283867257181555e-05
step: 180, loss: 9.97856623143889e-05
step: 190, loss: 0.00023001516819931567
step: 200, loss: 0.0023264519404619932
step: 210, loss: 9.313719783676788e-05
step: 220, loss: 0.0001403099304297939
step: 230, loss: 0.010196336545050144
step: 240, loss: 0.0012418113183230162
step: 250, loss: 4.104797335457988e-05
step: 260, loss: 3.594209920265712e-05
step: 270, loss: 0.0012628734111785889
step: 280, loss: 0.00010199782991549
step: 290, loss: 0.0005887232255190611
step: 300, loss: 0.000835371611174196
step: 310, loss: 7.584747072542086e-05
step: 320, loss: 9.274847980123013e-05
step: 330, loss: 0.0017583513399586082
step: 340, loss: 0.0001457003818359226
step: 350, loss: 0.0001784678897820413
step: 360, loss: 0.0018296937923878431
step: 370, loss: 0.001998729771003127
step: 380, loss: 0.002197608584538102
step: 390, loss: 0.0008968422771431506
step: 400, loss: 7.538402860518545e-05
step: 410, loss: 0.023077422752976418
step: 420, loss: 0.00641119247302413
step: 430, loss: 0.00023610907373949885
step: 440, loss: 0.0021567281801253557
step: 450, loss: 0.0025044262874871492
```

```

step: 460, loss: 0.0007191378972493112
step: 470, loss: 0.016299765557050705
step: 480, loss: 0.006657349411398172
step: 490, loss: 0.0001228178443852812
step: 500, loss: 3.87938525818754e-05
step: 510, loss: 0.015922220423817635
step: 520, loss: 0.004447873216122389
step: 530, loss: 0.00028301787097007036
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
    warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 12: dev_f1=0.9610511496949787, f1=0.9539040451552211,
best_f1=0.9539040451552211
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 8.609184442320839e-05
step: 10, loss: 6.663807289442047e-05
step: 20, loss: 9.8741365945898e-05
step: 30, loss: 2.6592897484079003e-05
step: 40, loss: 8.376500045415014e-05
step: 50, loss: 0.009884467348456383
step: 60, loss: 0.00020542048150673509
step: 70, loss: 0.00022700736008118838
step: 80, loss: 0.00023502440308220685
step: 90, loss: 6.170971028041095e-05
step: 100, loss: 8.781409997027367e-05
step: 110, loss: 0.00027807336300611496
step: 120, loss: 4.578692096401937e-05
step: 130, loss: 2.8774082238669507e-05
step: 140, loss: 0.10939031094312668
step: 150, loss: 0.0016480167396366596
step: 160, loss: 0.0013101794756948948
step: 170, loss: 0.0025379923172295094
step: 180, loss: 0.00010587689757812768
step: 190, loss: 0.003038342110812664
step: 200, loss: 8.343692752532661e-05
step: 210, loss: 7.513087621191517e-05
step: 220, loss: 5.484323264681734e-05
step: 230, loss: 0.0006265864358283579
step: 240, loss: 0.0012645367532968521
step: 250, loss: 0.0012799759861081839
step: 260, loss: 0.00012179769692011178
step: 270, loss: 0.00017574269440956414

```

```

step: 280, loss: 4.2979721911251545e-05
step: 290, loss: 4.8673507990315557e-05
step: 300, loss: 3.5258035495644435e-05
step: 310, loss: 0.00012184569641249254
step: 320, loss: 0.00015239484491758049
step: 330, loss: 0.0005651636747643352
step: 340, loss: 0.001619564602151513
step: 350, loss: 3.388871482457034e-05
step: 360, loss: 0.09892618656158447
step: 370, loss: 0.00019313966913614422
step: 380, loss: 9.494255937170237e-05
step: 390, loss: 0.00010777372517623007
step: 400, loss: 0.05785692110657692
step: 410, loss: 0.0002140777967397124
step: 420, loss: 8.194266411010176e-05
step: 430, loss: 0.00038510188460350037
step: 440, loss: 3.5910306905861944e-05
step: 450, loss: 0.0005403912509791553
step: 460, loss: 0.0004387784574646503
step: 470, loss: 0.0014689265517517924
step: 480, loss: 1.8051996448775753e-05
step: 490, loss: 0.0004062563239131123
step: 500, loss: 0.022722458466887474
step: 510, loss: 0.0004107132845092565
step: 520, loss: 9.437373228138313e-05
step: 530, loss: 8.329110278282315e-05
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 13: dev_f1=0.9602220166512488, f1=0.9564007421150279,
best_f1=0.9539040451552211
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0017521639820188284
step: 10, loss: 0.0004190272302366793
step: 20, loss: 3.4008236980298534e-05
step: 30, loss: 0.0006543969502672553
step: 40, loss: 0.0008408491848967969
step: 50, loss: 0.0005259978934191167
step: 60, loss: 5.221438186708838e-05
step: 70, loss: 7.89727782830596e-05
step: 80, loss: 0.00010778757859952748
step: 90, loss: 2.1270327124511823e-05
step: 100, loss: 3.398327316972427e-05
step: 110, loss: 0.0009423049632459879
step: 120, loss: 4.408427412272431e-05
step: 130, loss: 0.00013579321966972202

```

```

step: 140, loss: 0.0003920981544069946
step: 150, loss: 0.001127749215811491
step: 160, loss: 5.235938442638144e-05
step: 170, loss: 0.021934228017926216
step: 180, loss: 4.9915215640794486e-05
step: 190, loss: 8.519520633853972e-05
step: 200, loss: 5.117345426697284e-05
step: 210, loss: 0.00011357049515936524
step: 220, loss: 3.141954221064225e-05
step: 230, loss: 2.198213041992858e-05
step: 240, loss: 0.000955189869273454
step: 250, loss: 0.0009140796028077602
step: 260, loss: 0.0001230015914188698
step: 270, loss: 0.0003018575662281364
step: 280, loss: 6.129528628662229e-05
step: 290, loss: 2.3221511582960375e-05
step: 300, loss: 0.000660555437207222
step: 310, loss: 5.5078384320950136e-05
step: 320, loss: 0.00011710264516295865
step: 330, loss: 0.0010435954900458455
step: 340, loss: 0.0011896335054188967
step: 350, loss: 3.478049984551035e-05
step: 360, loss: 0.00013210636097937822
step: 370, loss: 0.00034204419353045523
step: 380, loss: 0.002248843666166067
step: 390, loss: 4.439937765710056e-05
step: 400, loss: 0.0029274208936840296
step: 410, loss: 2.8090300475014374e-05
step: 420, loss: 0.0005753906443715096
step: 430, loss: 0.0002601279702503234
step: 440, loss: 4.8582307499600574e-05
step: 450, loss: 3.765594374272041e-05
step: 460, loss: 0.04399310052394867
step: 470, loss: 3.353305874043144e-05
step: 480, loss: 6.977710290811956e-05
step: 490, loss: 2.8467276933952235e-05
step: 500, loss: 0.0016359809087589383
step: 510, loss: 0.004456216003745794
step: 520, loss: 0.0012086920905858278
step: 530, loss: 0.0015695176552981138
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
epoch 14: dev_f1=0.9601482854494902, f1=0.956884561891516,
best_f1=0.9539040451552211
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")

```

step: 0, loss: 0.00021629147522617131  
step: 10, loss: 0.002404229948297143  
step: 20, loss: 0.0008023494738154113  
step: 30, loss: 0.005891375709325075  
step: 40, loss: 3.973021011915989e-05  
step: 50, loss: 0.0011208361247554421  
step: 60, loss: 0.0010951166041195393  
step: 70, loss: 2.4957937057479285e-05  
step: 80, loss: 3.0009136025910266e-05  
step: 90, loss: 2.340125138289295e-05  
step: 100, loss: 1.609195169294253e-05  
step: 110, loss: 0.004820634610950947  
step: 120, loss: 2.546071664255578e-05  
step: 130, loss: 3.07790505758021e-05  
step: 140, loss: 1.7739108443493024e-05  
step: 150, loss: 3.615107198129408e-05  
step: 160, loss: 4.4688626076094806e-05  
step: 170, loss: 7.669891783734784e-05  
step: 180, loss: 4.426475425134413e-05  
step: 190, loss: 0.0012066575000062585  
step: 200, loss: 0.000782245013397187  
step: 210, loss: 0.00011339785123709589  
step: 220, loss: 3.111825208179653e-05  
step: 230, loss: 3.390893107280135e-05  
step: 240, loss: 3.1010400562081486e-05  
step: 250, loss: 2.484117248968687e-05  
step: 260, loss: 1.56197802425595e-05  
step: 270, loss: 2.9371118216658942e-05  
step: 280, loss: 1.4997686776041519e-05  
step: 290, loss: 0.00010296133405063301  
step: 300, loss: 5.300325938151218e-05  
step: 310, loss: 0.019543061032891273  
step: 320, loss: 3.14732278638985e-05  
step: 330, loss: 2.8150814614491537e-05  
step: 340, loss: 4.035453457618132e-05  
step: 350, loss: 0.0009154824074357748  
step: 360, loss: 6.791755731683224e-05  
step: 370, loss: 0.0001612349005881697  
step: 380, loss: 5.191522359382361e-05  
step: 390, loss: 3.0046303436392918e-05  
step: 400, loss: 0.014381522312760353  
step: 410, loss: 4.0286093280883506e-05  
step: 420, loss: 2.658629273355473e-05  
step: 430, loss: 1.0181051948165987e-05  
step: 440, loss: 0.00016837415751069784  
step: 450, loss: 0.005360594484955072  
step: 460, loss: 0.00279070227406919  
step: 470, loss: 1.9389453882467933e-05



```

step: 480, loss: 0.0021967622451484203
step: 490, loss: 4.320513835409656e-05
step: 500, loss: 0.0018083479953929782
step: 510, loss: 0.00011201125016668811
step: 520, loss: 0.0004633413045667112
step: 530, loss: 2.1903266315348446e-05
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 15: dev_f1=0.9606003752345216, f1=0.9553361542078043,
best_f1=0.9539040451552211
mv: cannot stat '*_dev.pt': No such file or directory
Selected optimization level 02:  FP16 training with FP32 batchnorm and FP32
master weights.

Defaults for this optimization level are:
enabled                : True
opt_level              : 02
cast_model_type        : torch.float16
patch_torch_functions  : False
keep_batchnorm_fp32    : True
master_weights         : True
loss_scale             : dynamic
Processing user overrides (additional kwargs that are not None)...
After processing overrides, optimization options are:
enabled                : True
opt_level              : 02
cast_model_type        : torch.float16
patch_torch_functions  : False
keep_batchnorm_fp32    : True
master_weights         : True
loss_scale             : dynamic
Warning:  multi_tensor_applier fused unscale kernel is unavailable, possibly
because apex was installed without --cuda_ext --cpp_ext. Using Python fallback.
Original ImportError was: ModuleNotFoundError("No module named 'amp_C'")
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
5742it [00:32, 174.31it/s]
load_f1 = 0.9598506069094305
real_f1 = 0.9592887225081891
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
4398it [00:31, 139.55it/s]

```

### 3.1.4 Amazon-Google - Running the matcher

```
[ ]: # Step 1:
# train the wdc model with the save_model flag on
# it will produce two files *_dev.pt and *_test.pt which are the
# best checkpoints on the dev set and the test set.
!CUDA_VISIBLE_DEVICES=0 python train_ditto.py \
  --task Structured/Amazon-Google \
  --batch_size 32 \
  --max_len 256 \
  --lr 3e-5 \
  --n_epochs 15 \
  --finetuning \
  --lm roberta \
  --fp16 \
  --save_model \
  --dk general \
  --dk product \
  --summarize

# Step 2: pick a checkpoint, rename it to wdc_all_small.pt
!mv *_dev.pt checkpoints/Amazon-Google.pt

# Step 3: run the matcher
!CUDA_VISIBLE_DEVICES=0 python matcher.py \
  --task Structured/Amazon-Google \
  --input_path input/input_small.jsonl \
  --output_path output/output_small.jsonl \
  --lm roberta \
  --use_gpu \
  --fp16 \
  --checkpoint_path checkpoints/
```

Selected optimization level 02: FP16 training with FP32 batchnorm and FP32 master weights.

Defaults for this optimization level are:

```
enabled           : True
opt_level         : 02
cast_model_type   : torch.float16
patch_torch_functions : False
keep_batchnorm_fp32 : True
master_weights    : True
loss_scale        : dynamic
```

Processing user overrides (additional kwargs that are not None)...

After processing overrides, optimization options are:

```

enabled                : True
opt_level              : 02
cast_model_type        : torch.float16
patch_torch_functions  : False
keep_batchnorm_fp32    : True
master_weights         : True
loss_scale             : dynamic
Warning: multi_tensor_applier fused unscale kernel is unavailable, possibly
because apex was installed without --cuda_ext --cpp_ext. Using Python fallback.
Original ImportError was: ModuleNotFoundError("No module named 'amp_C'")
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 32768.0
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:136:
UserWarning: Detected call of `lr_scheduler.step()` before `optimizer.step()`.
In PyTorch 1.1.0 and later, you should call them in the opposite order:
`optimizer.step()` before `lr_scheduler.step()`. Failure to do this will result
in PyTorch skipping the first value of the learning rate schedule. See more
details at https://pytorch.org/docs/stable/optim.html#how-to-adjust-learning-
rate
    "https://pytorch.org/docs/stable/optim.html#how-to-adjust-learning-rate",
UserWarning)
step: 0, loss: 0.5039608478546143
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 10, loss: 0.5276949405670166
step: 20, loss: 0.4924947917461395
step: 30, loss: 0.30206239223480225
step: 40, loss: 0.37483763694763184
step: 50, loss: 0.47839710116386414
step: 60, loss: 0.47494223713874817
step: 70, loss: 0.3067350387573242
step: 80, loss: 0.38532495498657227
step: 90, loss: 0.2607560455799103
step: 100, loss: 0.24982643127441406
step: 110, loss: 0.2543222904205322
step: 120, loss: 0.36980435252189636
step: 130, loss: 0.25893667340278625
step: 140, loss: 0.45089781284332275
step: 150, loss: 0.3883163332939148
step: 160, loss: 0.5193416476249695
step: 170, loss: 0.23885388672351837
step: 180, loss: 0.36944684386253357
step: 190, loss: 0.6341403722763062
step: 200, loss: 0.3719358444213867
step: 210, loss: 0.5620338916778564
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.

```

```

    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
    warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 1: dev_f1=0.18519984170953702, f1=0.18519984170953702,
best_f1=0.18519984170953702
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.36942920088768005
step: 10, loss: 0.19082127511501312
step: 20, loss: 0.49043548107147217
step: 30, loss: 0.511960506439209
step: 40, loss: 0.4471459686756134
step: 50, loss: 0.24471111595630646
step: 60, loss: 0.31107261776924133
step: 70, loss: 0.4342854619026184
step: 80, loss: 0.3275313973426819
step: 90, loss: 0.3887665271759033
step: 100, loss: 0.48290666937828064
step: 110, loss: 0.3960782587528229
step: 120, loss: 0.24185961484909058
step: 130, loss: 0.17213299870491028
step: 140, loss: 0.22724105417728424
step: 150, loss: 0.4399692714214325
step: 160, loss: 0.1330382525920868
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 8192.0
step: 170, loss: 0.4540374279022217
step: 180, loss: 0.388158917427063
step: 190, loss: 0.32955116033554077
step: 200, loss: 0.15546458959579468
step: 210, loss: 0.32465606927871704
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
    warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 2: dev_f1=0.3708838821490467, f1=0.2888888888888889,
best_f1=0.2888888888888889
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.2189028114080429
step: 10, loss: 0.1718169003725052
step: 20, loss: 0.42861467599868774

```

```

step: 30, loss: 0.2499190717935562
step: 40, loss: 0.44566109776496887
step: 50, loss: 0.48342418670654297
step: 60, loss: 0.49839067459106445
step: 70, loss: 0.19425469636917114
step: 80, loss: 0.40974530577659607
step: 90, loss: 0.20536665618419647
step: 100, loss: 0.38178253173828125
step: 110, loss: 0.12515826523303986
step: 120, loss: 0.22183477878570557
step: 130, loss: 0.19781365990638733
step: 140, loss: 0.37190401554107666
step: 150, loss: 0.3281078636646271
step: 160, loss: 0.18382462859153748
step: 170, loss: 0.3946031332015991
step: 180, loss: 0.242489755153656
step: 190, loss: 0.15228745341300964
step: 200, loss: 0.22370561957359314
step: 210, loss: 0.25762036442756653
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
epoch 3: dev_f1=0.21325361235675142, f1=0.21335992023928216,
best_f1=0.2888888888888889
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.3012012243270874
step: 10, loss: 0.32308539748191833
step: 20, loss: 0.32878926396369934
step: 30, loss: 0.23701229691505432
step: 40, loss: 0.2284218668937683
step: 50, loss: 0.23491793870925903
step: 60, loss: 0.5267866849899292
step: 70, loss: 0.2350674420595169
step: 80, loss: 0.23107442259788513
step: 90, loss: 0.4394078850746155
step: 100, loss: 0.37758544087409973
step: 110, loss: 0.5683178901672363
step: 120, loss: 0.3649739921092987
step: 130, loss: 0.6444141864776611
step: 140, loss: 0.5143036246299744
step: 150, loss: 0.4045545756816864
step: 160, loss: 0.30380603671073914
step: 170, loss: 0.21232573688030243
step: 180, loss: 0.06905702501535416
step: 190, loss: 0.1594865322113037
step: 200, loss: 0.24073612689971924

```

```

step: 210, loss: 0.4850398898124695
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
    warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 4: dev_f1=0.41620111731843573, f1=0.4083916083916084,
best_f1=0.4083916083916084
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.3790893256664276
step: 10, loss: 0.2692098915576935
step: 20, loss: 0.3009776771068573
step: 30, loss: 0.1904250830411911
step: 40, loss: 0.3185644745826721
step: 50, loss: 0.4141671657562256
step: 60, loss: 0.3728751242160797
step: 70, loss: 0.23656442761421204
step: 80, loss: 0.37866005301475525
step: 90, loss: 0.3489248752593994
step: 100, loss: 0.12032734602689743
step: 110, loss: 0.18096032738685608
step: 120, loss: 0.1930607408285141
step: 130, loss: 0.269550085067749
step: 140, loss: 0.5492989420890808
step: 150, loss: 0.21221399307250977
step: 160, loss: 0.18664133548736572
step: 170, loss: 0.17969539761543274
step: 180, loss: 0.2249513417482376
step: 190, loss: 0.35818013548851013
step: 200, loss: 0.34344226121902466
step: 210, loss: 0.21785472333431244
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
    warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 5: dev_f1=0.5693430656934306, f1=0.570342205323194,
best_f1=0.570342205323194
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.09683047235012054

```

```

step: 10, loss: 0.3295392692089081
step: 20, loss: 0.3778570592403412
step: 30, loss: 0.18593180179595947
step: 40, loss: 0.2663487195968628
step: 50, loss: 0.36091241240501404
step: 60, loss: 0.15134555101394653
step: 70, loss: 0.233612060546875
step: 80, loss: 0.2816230356693268
step: 90, loss: 0.2454068809747696
step: 100, loss: 0.2673706114292145
step: 110, loss: 0.0838029682636261
step: 120, loss: 0.21187514066696167
step: 130, loss: 0.133993998169899
step: 140, loss: 0.21491383016109467
step: 150, loss: 0.060145460069179535
step: 160, loss: 0.13814102113246918
step: 170, loss: 0.1902678906917572
step: 180, loss: 0.23111622035503387
step: 190, loss: 0.15282560884952545
step: 200, loss: 0.2872544229030609
step: 210, loss: 0.18070101737976074
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please save or load the state of the optimizer when saving or
loading the scheduler.
    warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 6: dev_f1=0.618042226487524, f1=0.6247544204322201,
best_f1=0.6247544204322201
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.18134227395057678
step: 10, loss: 0.052857208997011185
step: 20, loss: 0.14578814804553986
step: 30, loss: 0.16068969666957855
step: 40, loss: 0.14311206340789795
step: 50, loss: 0.19785048067569733
step: 60, loss: 0.18178364634513855
step: 70, loss: 0.10578012466430664
step: 80, loss: 0.3078190088272095
step: 90, loss: 0.3880179226398468
step: 100, loss: 0.14689967036247253
step: 110, loss: 0.1715793013572693
step: 120, loss: 0.2269052118062973
step: 130, loss: 0.2311730682849884
step: 140, loss: 0.1881096065044403

```

```

step: 150, loss: 0.23596109449863434
step: 160, loss: 0.23069782555103302
step: 170, loss: 0.21646466851234436
step: 180, loss: 0.10768461972475052
step: 190, loss: 0.21387387812137604
step: 200, loss: 0.1299157738685608
step: 210, loss: 0.21350513398647308
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
    warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 7: dev_f1=0.6715867158671587, f1=0.6740331491712707,
best_f1=0.6740331491712707
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.28282952308654785
step: 10, loss: 0.21632497012615204
step: 20, loss: 0.09920565038919449
step: 30, loss: 0.1178782656788826
step: 40, loss: 0.12304522842168808
step: 50, loss: 0.0961448922753334
step: 60, loss: 0.05291176214814186
step: 70, loss: 0.2284115105867386
step: 80, loss: 0.235401451587677
step: 90, loss: 0.13060304522514343
step: 100, loss: 0.2532830834388733
step: 110, loss: 0.10599178820848465
step: 120, loss: 0.22775615751743317
step: 130, loss: 0.03262428939342499
step: 140, loss: 0.08497640490531921
step: 150, loss: 0.366544634103775
step: 160, loss: 0.44757217168807983
step: 170, loss: 0.12767106294631958
step: 180, loss: 0.1805449277162552
step: 190, loss: 0.09207423776388168
step: 200, loss: 0.09537352621555328
step: 210, loss: 0.297841876745224
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
    warnings.warn(SAVE_STATE_WARNING, UserWarning)

```



```

epoch 8: dev_f1=0.6731707317073172, f1=0.6613162118780096,
best_f1=0.6613162118780096
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.13168495893478394
step: 10, loss: 0.17411600053310394
step: 20, loss: 0.15976178646087646
step: 30, loss: 0.15301229059696198
step: 40, loss: 0.11647573113441467
step: 50, loss: 0.15266071259975433
step: 60, loss: 0.09277140349149704
step: 70, loss: 0.10861051827669144
step: 80, loss: 0.06724216043949127
step: 90, loss: 0.17927497625350952
step: 100, loss: 0.10054528713226318
step: 110, loss: 0.1692476123571396
step: 120, loss: 0.15796644985675812
step: 130, loss: 0.11012309789657593
step: 140, loss: 0.10493528842926025
step: 150, loss: 0.06314685195684433
step: 160, loss: 0.40531638264656067
step: 170, loss: 0.14213243126869202
step: 180, loss: 0.15441715717315674
step: 190, loss: 0.14973758161067963
step: 200, loss: 0.0688885822892189
step: 210, loss: 0.07491251081228256
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
    warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 9: dev_f1=0.6963249516441006, f1=0.6990291262135924,
best_f1=0.6990291262135924
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.04776874557137489
step: 10, loss: 0.05979682877659798
step: 20, loss: 0.0874796062707901
step: 30, loss: 0.04461584612727165
step: 40, loss: 0.08196377009153366
step: 50, loss: 0.08599451929330826
step: 60, loss: 0.04427826777100563
step: 70, loss: 0.08203944563865662
step: 80, loss: 0.05040135979652405

```

```

step: 90, loss: 0.1366734504699707
step: 100, loss: 0.2018503099679947
step: 110, loss: 0.01917940191924572
step: 120, loss: 0.20533104240894318
step: 130, loss: 0.09772206842899323
step: 140, loss: 0.06519373506307602
step: 150, loss: 0.13051484525203705
step: 160, loss: 0.10249945521354675
step: 170, loss: 0.07466105371713638
step: 180, loss: 0.12361498922109604
step: 190, loss: 0.04415680468082428
step: 200, loss: 0.1486314982175827
step: 210, loss: 0.03932402282953262
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
    warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 10: dev_f1=0.7171314741035857, f1=0.7203219315895373,
best_f1=0.7203219315895373
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.08019912987947464
step: 10, loss: 0.07717067003250122
step: 20, loss: 0.035877905786037445
step: 30, loss: 0.12856987118721008
step: 40, loss: 0.0889354944229126
step: 50, loss: 0.15451227128505707
step: 60, loss: 0.1330929696559906
step: 70, loss: 0.05815843865275383
step: 80, loss: 0.08376666903495789
step: 90, loss: 0.17005492746829987
step: 100, loss: 0.08212636411190033
step: 110, loss: 0.05609112232923508
step: 120, loss: 0.07274188846349716
step: 130, loss: 0.061511795967817307
step: 140, loss: 0.10189841687679291
step: 150, loss: 0.07109363377094269
step: 160, loss: 0.0175953172147274
step: 170, loss: 0.034575898200273514
step: 180, loss: 0.034723564982414246
step: 190, loss: 0.32831454277038574
step: 200, loss: 0.1396457701921463
step: 210, loss: 0.03462647646665573
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:

```

```

An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
epoch 11: dev_f1=0.7159090909090908, f1=0.7093023255813954,
best_f1=0.7203219315895373
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.04660595953464508
step: 10, loss: 0.2292862832546234
step: 20, loss: 0.14135919511318207
step: 30, loss: 0.08211042732000351
step: 40, loss: 0.008501829579472542
step: 50, loss: 0.09287387877702713
step: 60, loss: 0.011567989364266396
step: 70, loss: 0.20489667356014252
step: 80, loss: 0.13518336415290833
step: 90, loss: 0.05438108742237091
step: 100, loss: 0.007518633268773556
step: 110, loss: 0.09693491458892822
step: 120, loss: 0.009481823071837425
step: 130, loss: 0.05340252071619034
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 8192.0
step: 140, loss: 0.1017698347568512
step: 150, loss: 0.0003988801909144968
step: 160, loss: 0.02587440423667431
step: 170, loss: 0.09154849499464035
step: 180, loss: 0.17008322477340698
step: 190, loss: 0.06168736517429352
step: 200, loss: 0.024430468678474426
step: 210, loss: 0.0752914696931839
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
  warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 12: dev_f1=0.7356321839080459, f1=0.7137254901960784,
best_f1=0.7137254901960784
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.06490655988454819
step: 10, loss: 0.04539094865322113
step: 20, loss: 0.014670870266854763
step: 30, loss: 0.029402391985058784
step: 40, loss: 0.045229222625494
step: 50, loss: 0.25133422017097473

```

```

step: 60, loss: 0.029743578284978867
step: 70, loss: 0.04402705654501915
step: 80, loss: 0.14268770813941956
step: 90, loss: 0.0609925240278244
step: 100, loss: 0.04411311447620392
step: 110, loss: 0.08944496512413025
step: 120, loss: 0.04875824600458145
step: 130, loss: 0.019185377284884453
step: 140, loss: 0.037420645356178284
step: 150, loss: 0.012204120866954327
step: 160, loss: 0.02401277795433998
step: 170, loss: 0.05636696144938469
step: 180, loss: 0.03788504749536514
step: 190, loss: 0.058496471494436264
step: 200, loss: 0.08619816601276398
step: 210, loss: 0.055673521012067795
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 13: dev_f1=0.7251908396946565, f1=0.720616570327553,
best_f1=0.7137254901960784
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.019669990986585617
step: 10, loss: 0.11741850525140762
step: 20, loss: 0.05198117718100548
step: 30, loss: 0.016497334465384483
step: 40, loss: 0.07193869352340698
step: 50, loss: 0.024510055780410767
step: 60, loss: 0.053262028843164444
step: 70, loss: 0.02466363087296486
step: 80, loss: 0.010913201607763767
step: 90, loss: 0.006949007045477629
step: 100, loss: 0.08113381266593933
step: 110, loss: 0.03141424059867859
step: 120, loss: 0.0014867429854348302
step: 130, loss: 0.12823155522346497
step: 140, loss: 0.011472761631011963
step: 150, loss: 0.053451795130968094
step: 160, loss: 0.002009207848459482
step: 170, loss: 0.11046876758337021
step: 180, loss: 0.021054832264780998
step: 190, loss: 0.07963293790817261
step: 200, loss: 0.04862348362803459
step: 210, loss: 0.04510489106178284
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.

```

```

    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
    warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 14: dev_f1=0.7356746765249538, f1=0.7296296296296296,
best_f1=0.7296296296296296
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.012025415897369385
step: 10, loss: 0.00973700825124979
step: 20, loss: 0.10904834419488907
step: 30, loss: 0.05340706557035446
step: 40, loss: 0.058779846876859665
step: 50, loss: 0.02581978589296341
step: 60, loss: 0.15765638649463654
step: 70, loss: 0.032390475273132324
step: 80, loss: 0.05740189179778099
step: 90, loss: 0.020019054412841797
step: 100, loss: 0.05882083252072334
step: 110, loss: 0.1128852590918541
step: 120, loss: 0.024470867589116096
step: 130, loss: 0.012729880400002003
step: 140, loss: 0.0312504768371582
step: 150, loss: 0.0637526884675026
step: 160, loss: 0.1632453054189682
step: 170, loss: 0.034307826310396194
step: 180, loss: 0.026624176651239395
step: 190, loss: 0.009797211736440659
step: 200, loss: 0.04661265388131142
step: 210, loss: 0.048128459602594376
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 15: dev_f1=0.7333333333333334, f1=0.7214953271028036,
best_f1=0.7296296296296296
mv: cannot stat '*_dev.pt': No such file or directory
Selected optimization level 02: FP16 training with FP32 batchnorm and FP32
master weights.

```

Defaults for this optimization level are:

```

enabled           : True
opt_level         : 02
cast_model_type   : torch.float16
patch_torch_functions : False
keep_batchnorm_fp32 : True
master_weights    : True

```

```

loss_scale          : dynamic
Processing user overrides (additional kwargs that are not None)...
After processing overrides, optimization options are:
enabled             : True
opt_level           : 02
cast_model_type     : torch.float16
patch_torch_functions : False
keep_batchnorm_fp32  : True
master_weights      : True
loss_scale          : dynamic
Warning: multi_tensor_applier fused unscale kernel is unavailable, possibly
because apex was installed without --cuda_ext --cpp_ext. Using Python fallback.
Original ImportError was: ModuleNotFoundError("No module named 'amp_C'")
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
2293it [00:09, 241.63it/s]
load_f1 = 0.7323943661971831
real_f1 = 0.7305389221556887
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
4398it [00:30, 142.21it/s]

```

### 3.1.5 Walmart-Amazon - Running the matcher

```

[ ]: # Step 1:
# train the wdc model with the save_model flag on
# it will produce two files *_dev.pt and *_test.pt which are the
# best checkpoints on the dev set and the test set.
!CUDA_VISIBLE_DEVICES=0 python train_ditto.py \
--task Structured/Walmart-Amazon \
--batch_size 32 \
--max_len 256 \
--lr 3e-5 \
--n_epochs 15 \
--finetuning \
--lm roberta \
--fp16 \
--save_model \
--dk general \
--dk product \
--summarize

```

```

# Step 2: pick a checkpoint, rename it to wdc_all_small.pt
!mv *_dev.pt checkpoints/Walmart-Amazon.pt

# Step 3: run the matcher
!CUDA_VISIBLE_DEVICES=0 python matcher.py \
  --task Structured/Walmart-Amazon \
  --input_path input/input_small.jsonl \
  --output_path output/output_small.jsonl \
  --lm roberta \
  --use_gpu \
  --fp16 \
  --checkpoint_path checkpoints/

```

Selected optimization level 02: FP16 training with FP32 batchnorm and FP32 master weights.

Defaults for this optimization level are:

```

enabled          : True
opt_level        : 02
cast_model_type  : torch.float16
patch_torch_functions : False
keep_batchnorm_fp32 : True
master_weights   : True
loss_scale       : dynamic

```

Processing user overrides (additional kwargs that are not None)...

After processing overrides, optimization options are:

```

enabled          : True
opt_level        : 02
cast_model_type  : torch.float16
patch_torch_functions : False
keep_batchnorm_fp32 : True
master_weights   : True
loss_scale       : dynamic

```

Warning: multi\_tensor\_applier fused unscale kernel is unavailable, possibly because apex was installed without --cuda\_ext --cpp\_ext. Using Python fallback. Original ImportError was: ModuleNotFoundError("No module named 'amp\_C'")  
/usr/local/lib/python3.7/dist-packages/apex/amp/\_initialize.py:25: UserWarning: An input tensor was not cuda.

```
warnings.warn("An input tensor was not cuda.")
```

Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 32768.0  
/usr/local/lib/python3.7/dist-packages/torch/optim/lr\_scheduler.py:136:

UserWarning: Detected call of `lr\_scheduler.step()` before `optimizer.step()`.

In PyTorch 1.1.0 and later, you should call them in the opposite order:

`optimizer.step()` before `lr\_scheduler.step()`. Failure to do this will result in PyTorch skipping the first value of the learning rate schedule. See more

details at <https://pytorch.org/docs/stable/optim.html#how-to-adjust-learning-rate>

```

"https://pytorch.org/docs/stable/optim.html#how-to-adjust-learning-rate",
UserWarning)
step: 0, loss: 0.4691135883331299
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 10, loss: 0.3853090703487396
step: 20, loss: 0.3259918987751007
step: 30, loss: 0.391821026802063
step: 40, loss: 0.26484400033950806
step: 50, loss: 0.31452804803848267
step: 60, loss: 0.551618218421936
step: 70, loss: 0.41518208384513855
step: 80, loss: 0.15303455293178558
step: 90, loss: 0.29905202984809875
step: 100, loss: 0.4344649314880371
step: 110, loss: 0.23063236474990845
step: 120, loss: 0.3313947021961212
step: 130, loss: 0.31731167435646057
step: 140, loss: 0.18756505846977234
step: 150, loss: 0.3008309304714203
step: 160, loss: 0.23009970784187317
step: 170, loss: 0.3559926152229309
step: 180, loss: 0.18785125017166138
step: 190, loss: 0.12423822283744812
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
  warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 1: dev_f1=0.6572104018912529, f1=0.6634615384615385,
best_f1=0.6634615384615385
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.3219646215438843
step: 10, loss: 0.2380756437778473
step: 20, loss: 0.7734664678573608
step: 30, loss: 0.18772700428962708
step: 40, loss: 0.2507811188697815
step: 50, loss: 0.3686348795890808
step: 60, loss: 0.10854455083608627
step: 70, loss: 0.27908268570899963
step: 80, loss: 0.05784813314676285
step: 90, loss: 0.18878144025802612
step: 100, loss: 0.13887354731559753

```



```

step: 110, loss: 0.12715759873390198
step: 120, loss: 0.15888828039169312
step: 130, loss: 0.08185741305351257
step: 140, loss: 0.23267890512943268
step: 150, loss: 0.2107401043176651
step: 160, loss: 0.08929447829723358
step: 170, loss: 0.019340721890330315
step: 180, loss: 0.014845963567495346
step: 190, loss: 0.17985165119171143
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
  warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 2: dev_f1=0.7921348314606742, f1=0.8222222222222222,
best_f1=0.8222222222222222
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.06887119263410568
step: 10, loss: 0.06144954264163971
step: 20, loss: 0.01966044493019581
step: 30, loss: 0.023156072944402695
step: 40, loss: 0.12175646424293518
step: 50, loss: 0.04947442188858986
step: 60, loss: 0.08457416296005249
step: 70, loss: 0.14763911068439484
step: 80, loss: 0.1266554445028305
step: 90, loss: 0.1444232165813446
step: 100, loss: 0.12906821072101593
step: 110, loss: 0.13294290006160736
step: 120, loss: 0.22298382222652435
step: 130, loss: 0.025597216561436653
step: 140, loss: 0.08835399150848389
step: 150, loss: 0.08610740303993225
step: 160, loss: 0.04398145526647568
step: 170, loss: 0.31902265548706055
step: 180, loss: 0.116116464138031
step: 190, loss: 0.02466554194688797
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
  warnings.warn(SAVE_STATE_WARNING, UserWarning)

```

```

epoch 3: dev_f1=0.8196721311475409, f1=0.8228882833787465,
best_f1=0.8228882833787465
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.005400074180215597
step: 10, loss: 0.030946874991059303
step: 20, loss: 0.05231070891022682
step: 30, loss: 0.008046082220971584
step: 40, loss: 0.054088905453681946
step: 50, loss: 0.004994620103389025
step: 60, loss: 0.11659962683916092
step: 70, loss: 0.022318001836538315
step: 80, loss: 0.0020341486670076847
step: 90, loss: 0.11909382790327072
step: 100, loss: 0.13046161830425262
step: 110, loss: 0.08242791146039963
step: 120, loss: 0.02783004753291607
step: 130, loss: 0.006658082362264395
step: 140, loss: 0.1210264191031456
step: 150, loss: 0.013531466946005821
step: 160, loss: 0.006244523450732231
step: 170, loss: 0.07714327424764633
step: 180, loss: 0.023060046136379242
step: 190, loss: 0.013946822844445705
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
    warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 4: dev_f1=0.8253164556962026, f1=0.8244274809160306,
best_f1=0.8244274809160306
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.008290833793580532
step: 10, loss: 0.031941529363393784
step: 20, loss: 0.003433695062994957
step: 30, loss: 0.008301248773932457
step: 40, loss: 0.10144662111997604
step: 50, loss: 0.009761710651218891
step: 60, loss: 0.04555553197860718
step: 70, loss: 0.016152111813426018
step: 80, loss: 0.002840715227648616
step: 90, loss: 0.16934865713119507
step: 100, loss: 0.12374921143054962

```

```

step: 110, loss: 0.11817629635334015
step: 120, loss: 0.07868447154760361
step: 130, loss: 0.32001015543937683
step: 140, loss: 0.010259955190122128
step: 150, loss: 0.08368067443370819
step: 160, loss: 0.012345942668616772
step: 170, loss: 0.06355711072683334
step: 180, loss: 0.014005100354552269
step: 190, loss: 0.003914212808012962
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
  warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 5: dev_f1=0.8302872062663187, f1=0.83289124668435,
best_f1=0.83289124668435
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.013303179293870926
step: 10, loss: 0.005385270342230797
step: 20, loss: 0.0015271519077941775
step: 30, loss: 0.011601099744439125
step: 40, loss: 0.0016828810330480337
step: 50, loss: 0.002513917861506343
step: 60, loss: 0.0028052288107573986
step: 70, loss: 0.025799138471484184
step: 80, loss: 0.02770214155316353
step: 90, loss: 0.001387166092172265
step: 100, loss: 0.0025417094584554434
step: 110, loss: 0.006499547511339188
step: 120, loss: 0.040921617299318314
step: 130, loss: 0.08243847638368607
step: 140, loss: 0.0032643761951476336
step: 150, loss: 0.022526005282998085
step: 160, loss: 0.04804699495434761
step: 170, loss: 0.12393823266029358
step: 180, loss: 0.04555885121226311
step: 190, loss: 0.005170760676264763
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
  warnings.warn(SAVE_STATE_WARNING, UserWarning)

```

```

epoch 6: dev_f1=0.8383838383838383, f1=0.8244274809160306,
best_f1=0.8244274809160306
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.004298600368201733
step: 10, loss: 0.0010013948194682598
step: 20, loss: 0.0010283946758136153
step: 30, loss: 0.008594553917646408
step: 40, loss: 0.001394793507643044
step: 50, loss: 0.00028368373750708997
step: 60, loss: 0.01629166677594185
step: 70, loss: 0.0006284145638346672
step: 80, loss: 0.004402901045978069
step: 90, loss: 0.0030013788491487503
step: 100, loss: 0.2607640326023102
step: 110, loss: 0.022359251976013184
step: 120, loss: 0.02792074717581272
step: 130, loss: 0.00174996186979115
step: 140, loss: 0.001885953824967146
step: 150, loss: 0.0007632009801454842
step: 160, loss: 0.009183363988995552
step: 170, loss: 0.01948663778603077
step: 180, loss: 0.009612328372895718
step: 190, loss: 0.0021833637729287148
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 7: dev_f1=0.8129675810473815, f1=0.8261964735516373,
best_f1=0.8244274809160306
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.11469709128141403
step: 10, loss: 0.004858647007495165
step: 20, loss: 0.0020844130776822567
step: 30, loss: 0.07449706643819809
step: 40, loss: 0.021871108561754227
step: 50, loss: 0.0033038482069969177
step: 60, loss: 0.0024246270768344402
step: 70, loss: 0.00034663095721043646
step: 80, loss: 0.16435883939266205
step: 90, loss: 0.0440070815384388
step: 100, loss: 0.0009348164894618094
step: 110, loss: 0.007955481298267841
step: 120, loss: 0.03656250610947609
step: 130, loss: 0.003434270853176713
step: 140, loss: 0.006856782361865044

```

```

step: 150, loss: 0.001942233880981803
step: 160, loss: 0.0006730316090397537
step: 170, loss: 0.0003903588803950697
step: 180, loss: 0.03389723226428032
step: 190, loss: 0.0007253731018863618
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
    warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 8: dev_f1=0.8578947368421053, f1=0.8461538461538463,
best_f1=0.8461538461538463
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.000997734721750021
step: 10, loss: 0.0004711602523457259
step: 20, loss: 0.015841711312532425
step: 30, loss: 0.006139430217444897
step: 40, loss: 0.03358506038784981
step: 50, loss: 0.03824548423290253
step: 60, loss: 0.002052373019978404
step: 70, loss: 0.0007272843504324555
step: 80, loss: 0.012042196467518806
step: 90, loss: 0.18618208169937134
step: 100, loss: 0.0010420364560559392
step: 110, loss: 0.00043454914703033864
step: 120, loss: 0.0446719191968441
step: 130, loss: 0.00028333874070085585
step: 140, loss: 0.029240664094686508
step: 150, loss: 0.04847419261932373
step: 160, loss: 0.020827634260058403
step: 170, loss: 0.0014630573568865657
step: 180, loss: 0.005679488647729158
step: 190, loss: 0.00020502146799117327
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 9: dev_f1=0.8525469168900803, f1=0.8489583333333334,
best_f1=0.8461538461538463
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00020571793720591813
step: 10, loss: 0.0003001403820235282
step: 20, loss: 0.01216123253107071

```

```

step: 30, loss: 0.0006546913064084947
step: 40, loss: 0.0006890833610668778
step: 50, loss: 0.00023929036979097873
step: 60, loss: 0.0009888762142509222
step: 70, loss: 0.00044900906505063176
step: 80, loss: 0.000751251180190593
step: 90, loss: 0.0002489601611159742
step: 100, loss: 0.0002425996062811464
step: 110, loss: 0.00015280152729246765
step: 120, loss: 0.00030346144922077656
step: 130, loss: 0.00031027739169076085
step: 140, loss: 0.0002641918836161494
step: 150, loss: 0.00020412217418197542
step: 160, loss: 0.0009741149842739105
step: 170, loss: 0.0009949618251994252
step: 180, loss: 0.0036811737809330225
step: 190, loss: 0.0009738988010212779
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
epoch 10: dev_f1=0.8407310704960834, f1=0.8477157360406091,
best_f1=0.8461538461538463
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0026155903469771147
step: 10, loss: 0.0006433665985241532
step: 20, loss: 0.0022949629928916693
step: 30, loss: 0.00031360890716314316
step: 40, loss: 0.0002278113242937252
step: 50, loss: 0.0006329861935228109
step: 60, loss: 0.001048272242769599
step: 70, loss: 0.0015886655310168862
step: 80, loss: 0.0012914120452478528
step: 90, loss: 0.00021548046788666397
step: 100, loss: 0.0003617672191467136
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 110, loss: 0.19970542192459106
step: 120, loss: 0.0010597596410661936
step: 130, loss: 0.001158013823442161
step: 140, loss: 0.0006364589789882302
step: 150, loss: 0.001090492820367217
step: 160, loss: 0.0012156281154602766
step: 170, loss: 0.0011336236493662
step: 180, loss: 0.006168735679239035
step: 190, loss: 0.006064615212380886
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.

```

```

    warnings.warn("An input tensor was not cuda.")
epoch 11: dev_f1=0.8379052369077307, f1=0.8246913580246913,
best_f1=0.8461538461538463
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0009430220234207809
step: 10, loss: 0.002501661656424403
step: 20, loss: 0.0002776039473246783
step: 30, loss: 0.0007929270504973829
step: 40, loss: 0.0001833472342696041
step: 50, loss: 0.001080608693882823
step: 60, loss: 0.0003192989097442478
step: 70, loss: 0.00021110233501531184
step: 80, loss: 0.001737608341500163
step: 90, loss: 0.0003254182229284197
step: 100, loss: 0.0014607952907681465
step: 110, loss: 0.05586221069097519
step: 120, loss: 0.006378286983817816
step: 130, loss: 0.0006763627170585096
step: 140, loss: 0.00021046116307843477
step: 150, loss: 0.0013294096570461988
step: 160, loss: 0.00011248742521274835
step: 170, loss: 0.0008784684468992054
step: 180, loss: 0.00019672435882966965
step: 190, loss: 0.005808783695101738
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 12: dev_f1=0.851063829787234, f1=0.8421052631578947,
best_f1=0.8461538461538463
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0003081108152400702
step: 10, loss: 0.00013020781625527889
step: 20, loss: 0.00026279326993972063
step: 30, loss: 0.05336915701627731
step: 40, loss: 0.0009550407994538546
step: 50, loss: 0.000199893896933645
step: 60, loss: 0.006039745640009642
step: 70, loss: 0.0003864331229124218
step: 80, loss: 0.0005751855787821114
step: 90, loss: 0.0005580530851148069
step: 100, loss: 0.0008134047384373844
step: 110, loss: 0.0008036293438635767
step: 120, loss: 0.0001279988937312737
step: 130, loss: 0.00010297414701199159

```

```

step: 140, loss: 0.0003973609418608248
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 8192.0
step: 150, loss: 0.0003315091016702354
step: 160, loss: 0.0004224676813464612
step: 170, loss: 7.613775233039632e-05
step: 180, loss: 8.221697498811409e-05
step: 190, loss: 0.13417769968509674
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 13: dev_f1=0.8431876606683805, f1=0.843989769820972,
best_f1=0.8461538461538463
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0005207151989452541
step: 10, loss: 0.00033281269134022295
step: 20, loss: 0.0018447081092745066
step: 30, loss: 0.00018011665088124573
step: 40, loss: 0.00018478816491551697
step: 50, loss: 0.00013986980775371194
step: 60, loss: 0.0003384422161616385
step: 70, loss: 0.00030393939232453704
step: 80, loss: 0.0001727686612866819
step: 90, loss: 0.00011315088340779766
step: 100, loss: 0.00012319926463533193
step: 110, loss: 0.00016420149768237025
step: 120, loss: 0.0003229044668842107
step: 130, loss: 0.00018349842866882682
step: 140, loss: 0.00013936519098933786
step: 150, loss: 0.00017165657482109964
step: 160, loss: 0.0003789953771047294
step: 170, loss: 6.080358798499219e-05
step: 180, loss: 0.00021228687546681613
step: 190, loss: 0.0005119417910464108
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 14: dev_f1=0.8541666666666666, f1=0.8445595854922279,
best_f1=0.8461538461538463
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00018468826601747423
step: 10, loss: 0.000993959023617208
step: 20, loss: 0.006056775338947773
step: 30, loss: 0.00015894036914687604
step: 40, loss: 0.0005692841368727386

```



```

step: 50, loss: 8.81839805515483e-05
step: 60, loss: 9.045602200785652e-05
step: 70, loss: 0.00016841785691212863
step: 80, loss: 5.520862032426521e-05
step: 90, loss: 0.0004645286244340241
step: 100, loss: 0.00017007268615998328
step: 110, loss: 0.00012739773956127465
step: 120, loss: 0.000172575528267771
step: 130, loss: 0.0002583693712949753
step: 140, loss: 0.0001368590019410476
step: 150, loss: 6.787427992094308e-05
step: 160, loss: 0.0001012529683066532
step: 170, loss: 0.000294017925625667
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 4096.0
step: 180, loss: 0.00016961140499915928
step: 190, loss: 0.00010891784768318757
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
    warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 15: dev_f1=0.8601036269430051, f1=0.8453608247422681,
best_f1=0.8453608247422681
mv: cannot stat '*_dev.pt': No such file or directory
Selected optimization level 02:  FP16 training with FP32 batchnorm and FP32
master weights.

Defaults for this optimization level are:
enabled                : True
opt_level              : 02
cast_model_type        : torch.float16
patch_torch_functions  : False
keep_batchnorm_fp32    : True
master_weights         : True
loss_scale             : dynamic
Processing user overrides (additional kwargs that are not None)...
After processing overrides, optimization options are:
enabled                : True
opt_level              : 02
cast_model_type        : torch.float16
patch_torch_functions  : False
keep_batchnorm_fp32    : True
master_weights         : True
loss_scale             : dynamic
Warning: multi_tensor_applier fused unscale kernel is unavailable, possibly
because apex was installed without --cuda_ext --cpp_ext. Using Python fallback.

```

```

Original ImportError was: ModuleNotFoundError("No module named 'amp_C'")
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
2049it [00:12, 162.07it/s]
load_f1 = 0.7688564476885644
real_f1 = 0.7378190255220417
733it [00:00, 3539.91it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
4398it [00:31, 139.27it/s]

```

## 3.2 DK TEXTUAL

### 3.2.1 Abt-Buy - Running the matcher

```

[ ]: # Step 1:
# train the wdc model with the save_model flag on
# it will produce two files *_dev.pt and *_test.pt which are the
# best checkpoints on the dev set and the test set.
!CUDA_VISIBLE_DEVICES=0 python train_ditto.py \
    --task Textual/Abt-Buy \
    --batch_size 32 \
    --max_len 256 \
    --lr 3e-5 \
    --n_epochs 15 \
    --finetuning \
    --lm roberta \
    --fp16 \
    --save_model \
    --dk general \
    --dk product \
    --summarize

# Step 2: pick a checkpoint, rename it to wdc_all_small.pt
!mv *_dev.pt checkpoints/Abt-Buy.pt

# Step 3: run the matcher
!CUDA_VISIBLE_DEVICES=0 python matcher.py \
    --task Textual/Abt-Buy \
    --input_path input/input_small.jsonl \
    --output_path output/output_small.jsonl \

```

```
--lm roberta \  
--use_gpu \  
--fp16 \  
--checkpoint_path checkpoints/
```

Selected optimization level 02: FP16 training with FP32 batchnorm and FP32 master weights.

Defaults for this optimization level are:

```
enabled          : True  
opt_level        : 02  
cast_model_type  : torch.float16  
patch_torch_functions : False  
keep_batchnorm_fp32 : True  
master_weights   : True  
loss_scale       : dynamic
```

Processing user overrides (additional kwargs that are not None)...

After processing overrides, optimization options are:

```
enabled          : True  
opt_level        : 02  
cast_model_type  : torch.float16  
patch_torch_functions : False  
keep_batchnorm_fp32 : True  
master_weights   : True  
loss_scale       : dynamic
```

Warning: multi\_tensor\_applier fused unscale kernel is unavailable, possibly because apex was installed without --cuda\_ext --cpp\_ext. Using Python fallback. Original ImportError was: ModuleNotFoundError("No module named 'amp\_C'")  
/usr/local/lib/python3.7/dist-packages/apex/amp/\_initialize.py:25: UserWarning: An input tensor was not cuda.

```
warnings.warn("An input tensor was not cuda.")
```

Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 32768.0  
/usr/local/lib/python3.7/dist-packages/torch/optim/lr\_scheduler.py:136:

UserWarning: Detected call of `lr\_scheduler.step()` before `optimizer.step()`.

In PyTorch 1.1.0 and later, you should call them in the opposite order:

`optimizer.step()` before `lr\_scheduler.step()`. Failure to do this will result in PyTorch skipping the first value of the learning rate schedule. See more details at <https://pytorch.org/docs/stable/optim.html#how-to-adjust-learning-rate>

```
"https://pytorch.org/docs/stable/optim.html#how-to-adjust-learning-rate",  
UserWarning)
```

```
step: 0, loss: 0.526259183883667
```

Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0

```
step: 10, loss: 0.4676077961921692
```

```
step: 20, loss: 0.2999970614910126
```

```
step: 30, loss: 0.40644142031669617
```

```
step: 40, loss: 0.6064562797546387
```

```

step: 50, loss: 0.3317360281944275
step: 60, loss: 0.5633382797241211
step: 70, loss: 0.34134018421173096
step: 80, loss: 0.24885126948356628
step: 90, loss: 0.2036239057779312
step: 100, loss: 0.17885735630989075
step: 110, loss: 0.4324948191642761
step: 120, loss: 0.29962629079818726
step: 130, loss: 0.3155773878097534
step: 140, loss: 0.38777464628219604
step: 150, loss: 0.33260294795036316
step: 160, loss: 0.40174585580825806
step: 170, loss: 0.3263002038002014
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
    warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 1: dev_f1=0.24169184290030213, f1=0.24096385542168672,
best_f1=0.24096385542168672
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.32118627429008484
step: 10, loss: 0.48827245831489563
step: 20, loss: 0.31402069330215454
step: 30, loss: 0.3077103793621063
step: 40, loss: 0.0764283835887909
step: 50, loss: 0.4213221073150635
step: 60, loss: 0.1713540405035019
step: 70, loss: 0.4601700007915497
step: 80, loss: 0.2180217057466507
step: 90, loss: 0.25475290417671204
step: 100, loss: 0.48652708530426025
step: 110, loss: 0.2708509564399719
step: 120, loss: 0.22727084159851074
step: 130, loss: 0.4445948004722595
step: 140, loss: 0.40001705288887024
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 8192.0
step: 150, loss: 0.2061198651790619
step: 160, loss: 0.4448425769805908
step: 170, loss: 0.11599242687225342
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:

```

UserWarning: Please also save or load the state of the optimizer when saving or loading the scheduler.

```
warnings.warn(SAVE_STATE_WARNING, UserWarning)
```

epoch 2: dev\_f1=0.7375886524822693, f1=0.7448275862068966,

best\_f1=0.7448275862068966

/usr/local/lib/python3.7/dist-packages/apex/amp/\_initialize.py:25: UserWarning: An input tensor was not cuda.

```
warnings.warn("An input tensor was not cuda.")
```

step: 0, loss: 0.25822874903678894

step: 10, loss: 0.030980896204710007

step: 20, loss: 0.04135896638035774

step: 30, loss: 0.12145643681287766

step: 40, loss: 0.16157716512680054

step: 50, loss: 0.415581077337265

step: 60, loss: 0.048399969935417175

step: 70, loss: 0.2423153668642044

step: 80, loss: 0.08154647797346115

step: 90, loss: 0.38736000657081604

step: 100, loss: 0.03548164665699005

step: 110, loss: 0.044926807284355164

step: 120, loss: 0.15315575897693634

step: 130, loss: 0.13177728652954102

step: 140, loss: 0.1778516322374344

step: 150, loss: 0.1054215207695961

step: 160, loss: 0.07556106150150299

step: 170, loss: 0.04704862833023071

/usr/local/lib/python3.7/dist-packages/apex/amp/\_initialize.py:25: UserWarning: An input tensor was not cuda.

```
warnings.warn("An input tensor was not cuda.")
```

/usr/local/lib/python3.7/dist-packages/torch/optim/lr\_scheduler.py:216:

UserWarning: Please also save or load the state of the optimizer when saving or loading the scheduler.

```
warnings.warn(SAVE_STATE_WARNING, UserWarning)
```

epoch 3: dev\_f1=0.8066825775656326, f1=0.834862385321101,

best\_f1=0.834862385321101

/usr/local/lib/python3.7/dist-packages/apex/amp/\_initialize.py:25: UserWarning: An input tensor was not cuda.

```
warnings.warn("An input tensor was not cuda.")
```

step: 0, loss: 0.05357023701071739

step: 10, loss: 0.031765785068273544

step: 20, loss: 0.03207718953490257

step: 30, loss: 0.324153870344162

step: 40, loss: 0.10962729156017303

step: 50, loss: 0.08191410452127457

step: 60, loss: 0.1705000251531601

step: 70, loss: 0.002432645298540592

step: 80, loss: 0.3131241798400879

step: 90, loss: 0.16147680580615997

```

step: 100, loss: 0.10523732751607895
step: 110, loss: 0.06471998989582062
step: 120, loss: 0.23448362946510315
step: 130, loss: 0.06289661675691605
step: 140, loss: 0.0999491736292839
step: 150, loss: 0.3037404716014862
step: 160, loss: 0.0525614395737648
step: 170, loss: 0.024413615465164185
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
    warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 4: dev_f1=0.8578680203045685, f1=0.8904761904761905,
best_f1=0.8904761904761905
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.012146693654358387
step: 10, loss: 0.08410604298114777
step: 20, loss: 0.09747122973203659
step: 30, loss: 0.04537736997008324
step: 40, loss: 0.05388794466853142
step: 50, loss: 0.06349106878042221
step: 60, loss: 0.11657988280057907
step: 70, loss: 0.012362470850348473
step: 80, loss: 0.001200590981170535
step: 90, loss: 0.03095550462603569
step: 100, loss: 0.016954490914940834
step: 110, loss: 0.09029677510261536
step: 120, loss: 0.055496711283922195
step: 130, loss: 0.005254071671515703
step: 140, loss: 0.11180946975946426
step: 150, loss: 0.026895839720964432
step: 160, loss: 0.036203544586896896
step: 170, loss: 0.0033216264564543962
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
    warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 5: dev_f1=0.8634146341463416, f1=0.8941176470588235,
best_f1=0.8941176470588235
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:

```

```

An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.13292568922042847
step: 10, loss: 0.04566536471247673
step: 20, loss: 0.1040763109922409
step: 30, loss: 0.005742212291806936
step: 40, loss: 0.016689477488398552
step: 50, loss: 0.0027289066929370165
step: 60, loss: 0.016595881432294846
step: 70, loss: 0.006915337406098843
step: 80, loss: 0.002927230903878808
step: 90, loss: 0.073667012155056
step: 100, loss: 0.0014246230712160468
step: 110, loss: 0.044049859046936035
step: 120, loss: 0.01566057652235031
step: 130, loss: 0.0016709534684196115
step: 140, loss: 0.018967757001519203
step: 150, loss: 0.006091041024774313
step: 160, loss: 0.07850073277950287
step: 170, loss: 0.02733246050775051
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
  warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 6: dev_f1=0.8661800486618004, f1=0.8951048951048951,
best_f1=0.8951048951048951
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.06047496944665909
step: 10, loss: 0.05335426330566406
step: 20, loss: 0.0016577932983636856
step: 30, loss: 0.002052642870694399
step: 40, loss: 0.0004562009416986257
step: 50, loss: 0.001662872964516282
step: 60, loss: 0.0012060514418408275
step: 70, loss: 0.006085401866585016
step: 80, loss: 0.04039088636636734
step: 90, loss: 0.051758892834186554
step: 100, loss: 0.007262060418725014
step: 110, loss: 0.09805256128311157
step: 120, loss: 0.009725759737193584
step: 130, loss: 0.0020403622183948755
step: 140, loss: 0.024709617719054222
step: 150, loss: 0.13964036107063293

```

```

step: 160, loss: 0.0072141047567129135
step: 170, loss: 0.04722701758146286
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
    warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 7: dev_f1=0.8682170542635659, f1=0.8985507246376813,
best_f1=0.8985507246376813
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.03920692950487137
step: 10, loss: 0.0010775765404105186
step: 20, loss: 0.02160722017288208
step: 30, loss: 0.002042656298726797
step: 40, loss: 0.0006599866319447756
step: 50, loss: 0.05079014226794243
step: 60, loss: 0.04663019999861717
step: 70, loss: 0.013566660694777966
step: 80, loss: 0.0010014285799115896
step: 90, loss: 0.06363712251186371
step: 100, loss: 0.003376785898581147
step: 110, loss: 0.11837334930896759
step: 120, loss: 0.036786917597055435
step: 130, loss: 0.0010177004151046276
step: 140, loss: 0.17298513650894165
step: 150, loss: 0.00512123154476285
step: 160, loss: 0.0014676552964374423
step: 170, loss: 0.2542550265789032
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
    warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 8: dev_f1=0.8762886597938145, f1=0.8968824940047962,
best_f1=0.8968824940047962
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.01747852936387062
step: 10, loss: 0.013178122229874134
step: 20, loss: 0.013725925236940384
step: 30, loss: 0.062211111187934875

```



```

step: 40, loss: 0.002737360307946801
step: 50, loss: 0.0014659189619123936
step: 60, loss: 0.05598519742488861
step: 70, loss: 0.0031253641936928034
step: 80, loss: 0.0016911689890548587
step: 90, loss: 0.007049905136227608
step: 100, loss: 0.05234752967953682
step: 110, loss: 0.014826495200395584
step: 120, loss: 0.011356919072568417
step: 130, loss: 0.007300740573555231
step: 140, loss: 0.0011832289164885879
step: 150, loss: 0.23643267154693604
step: 160, loss: 0.00810317974537611
step: 170, loss: 0.010020800866186619
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 9: dev_f1=0.859375, f1=0.9037037037037038, best_f1=0.8968824940047962
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0009039605502039194
step: 10, loss: 0.004769334569573402
step: 20, loss: 0.10659591853618622
step: 30, loss: 0.06239695847034454
step: 40, loss: 0.010614097118377686
step: 50, loss: 0.04456174746155739
step: 60, loss: 0.02188006602227688
step: 70, loss: 0.004275296814739704
step: 80, loss: 0.021339450031518936
step: 90, loss: 0.003093011910095811
step: 100, loss: 0.002198953414335847
step: 110, loss: 0.0034566475078463554
step: 120, loss: 0.0002469085156917572
step: 130, loss: 0.0025919086765497923
step: 140, loss: 0.0019981006626039743
step: 150, loss: 0.0004306854971218854
step: 160, loss: 0.0001035911263898015
step: 170, loss: 0.009770944714546204
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 10: dev_f1=0.8720626631853786, f1=0.9041769041769042,
best_f1=0.8968824940047962
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.07711940258741379

```

```

step: 10, loss: 0.001643862808123231
step: 20, loss: 0.00010491551802260801
step: 30, loss: 0.00826895609498024
step: 40, loss: 0.005507994908839464
step: 50, loss: 0.01568532921373844
step: 60, loss: 0.05576768517494202
step: 70, loss: 0.0007173542398959398
step: 80, loss: 0.09003850072622299
step: 90, loss: 0.006670298986136913
step: 100, loss: 0.007406516000628471
step: 110, loss: 0.002026598434895277
step: 120, loss: 0.0005652210675179958
step: 130, loss: 0.00856452900916338
step: 140, loss: 0.0007779439329169691
step: 150, loss: 0.00045758349006064236
step: 160, loss: 0.09005970507860184
step: 170, loss: 0.011063113808631897
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
  warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 11: dev_f1=0.8826530612244897, f1=0.9113300492610837,
best_f1=0.9113300492610837
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.000936072610784322
step: 10, loss: 0.0020540417172014713
step: 20, loss: 0.004089764319360256
step: 30, loss: 0.06641535460948944
step: 40, loss: 0.0004935658653266728
step: 50, loss: 0.008677913807332516
step: 60, loss: 0.24889437854290009
step: 70, loss: 0.00013857119483873248
step: 80, loss: 0.00029874945175834
step: 90, loss: 0.00787544809281826
step: 100, loss: 0.0010103180538862944
step: 110, loss: 0.0677090585231781
step: 120, loss: 0.020842695608735085
step: 130, loss: 0.02650636062026024
step: 140, loss: 0.00033881881972774863
step: 150, loss: 0.0058447751216590405
step: 160, loss: 0.0009881352307274938
step: 170, loss: 0.0003448155475780368
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:

```

```

An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
epoch 12: dev_f1=0.8717948717948718, f1=0.9090909090909091,
best_f1=0.9113300492610837
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.000383829785278067
step: 10, loss: 0.00017135175585281104
step: 20, loss: 0.0021314462646842003
step: 30, loss: 0.037486132234334946
step: 40, loss: 0.0003198837803211063
step: 50, loss: 0.020174704492092133
step: 60, loss: 0.0005627700593322515
step: 70, loss: 0.09005226194858551
step: 80, loss: 0.0006397562101483345
step: 90, loss: 0.00554729625582695
step: 100, loss: 0.03680283948779106
step: 110, loss: 0.0006063829641789198
step: 120, loss: 0.0026001823134720325
step: 130, loss: 0.0001439416955690831
step: 140, loss: 0.03226327523589134
step: 150, loss: 0.00013003074855078012
step: 160, loss: 0.0001574331836309284
step: 170, loss: 0.0011659832671284676
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
epoch 13: dev_f1=0.8616187989556136, f1=0.9046454767726161,
best_f1=0.9113300492610837
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.005687010008841753
step: 10, loss: 0.00012188372056698427
step: 20, loss: 0.010376008227467537
step: 30, loss: 0.00227175559848547
step: 40, loss: 0.005613543093204498
step: 50, loss: 0.000521500245667994
step: 60, loss: 0.0066056204959750175
step: 70, loss: 0.002844626549631357
step: 80, loss: 0.0002577303966972977
step: 90, loss: 0.018820641562342644
step: 100, loss: 0.00013576952915173024
step: 110, loss: 0.01157868281006813
step: 120, loss: 0.001362321898341179
step: 130, loss: 0.0012880951398983598
step: 140, loss: 0.025422263890504837

```

```

step: 150, loss: 0.02890613116323948
step: 160, loss: 0.00023619869898539037
step: 170, loss: 0.03669944405555725
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 14: dev_f1=0.8645833333333333, f1=0.9019607843137256,
best_f1=0.9113300492610837
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 9.4326060207095e-05
step: 10, loss: 0.0002421199023956433
step: 20, loss: 0.00015600059123244137
step: 30, loss: 0.00105092185549438
step: 40, loss: 0.00016430144023615867
step: 50, loss: 0.00039814264164306223
step: 60, loss: 0.0007018899777904153
step: 70, loss: 0.060548458248376846
step: 80, loss: 0.0030192111153155565
step: 90, loss: 0.002151380991563201
step: 100, loss: 0.035825494676828384
step: 110, loss: 0.00016984381363727152
step: 120, loss: 0.000499385001603514
step: 130, loss: 0.00022497806639876217
step: 140, loss: 0.00835784524679184
step: 150, loss: 0.00036758824717253447
step: 160, loss: 0.002541592111811042
step: 170, loss: 0.0007185909198597074
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 15: dev_f1=0.8601036269430052, f1=0.8992628992628993,
best_f1=0.9113300492610837
mv: cannot stat '*_dev.pt': No such file or directory
Selected optimization level O2:  FP16 training with FP32 batchnorm and FP32
master weights.

```

Defaults for this optimization level are:

```

enabled          : True
opt_level        : O2
cast_model_type  : torch.float16
patch_torch_functions : False
keep_batchnorm_fp32 : True
master_weights   : True
loss_scale       : dynamic

```

Processing user overrides (additional kwargs that are not None)...

After processing overrides, optimization options are:

```

enabled                : True
opt_level              : 02
cast_model_type        : torch.float16
patch_torch_functions  : False
keep_batchnorm_fp32    : True
master_weights         : True
loss_scale             : dynamic
Warning: multi_tensor_applier fused unscale kernel is unavailable, possibly
because apex was installed without --cuda_ext --cpp_ext. Using Python fallback.
Original ImportError was: ModuleNotFoundError("No module named 'amp_C'")
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
Oit [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
1916it [00:08, 229.86it/s]
load_f1 = 0.5763293310463122
real_f1 = 0.5672131147540983
Oit [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
4398it [00:32, 134.84it/s]

```

### 3.3 DK DIRTY

#### 3.3.1 DBLP-ACM - Running the matcher

```

[ ]: # Step 1:
# train the wdc model with the save_model flag on
# it will produce two files *_dev.pt and *_test.pt which are the
# best checkpoints on the dev set and the test set.
!CUDA_VISIBLE_DEVICES=0 python train_ditto.py \
--task Dirty/DBLP-ACM \
--batch_size 32 \
--max_len 256 \
--lr 3e-5 \
--n_epochs 15 \
--finetuning \
--lm roberta \
--fp16 \
--save_model \
--dk general \
--dk product \
--summarize

```

```

# Step 2: pick a checkpoint, rename it to wdc_all_small.pt
!mv *_dev.pt checkpoints/DBLP-ACM.pt

# Step 3: run the matcher
!CUDA_VISIBLE_DEVICES=0 python matcher.py \
  --task Dirty/DBLP-ACM \
  --input_path input/input_small.jsonl \
  --output_path output/output_small.jsonl \
  --lm roberta \
  --use_gpu \
  --fp16 \
  --checkpoint_path checkpoints/

```

Selected optimization level 02: FP16 training with FP32 batchnorm and FP32 master weights.

Defaults for this optimization level are:

```

enabled           : True
opt_level         : 02
cast_model_type   : torch.float16
patch_torch_functions : False
keep_batchnorm_fp32 : True
master_weights    : True
loss_scale        : dynamic

```

Processing user overrides (additional kwargs that are not None)...

After processing overrides, optimization options are:

```

enabled           : True
opt_level         : 02
cast_model_type   : torch.float16
patch_torch_functions : False
keep_batchnorm_fp32 : True
master_weights    : True
loss_scale        : dynamic

```

Warning: multi\_tensor\_applier fused unscale kernel is unavailable, possibly because apex was installed without --cuda\_ext --cpp\_ext. Using Python fallback. Original ImportError was: ModuleNotFoundError("No module named 'amp\_C'")  
/usr/local/lib/python3.7/dist-packages/apex/amp/\_initialize.py:25: UserWarning: An input tensor was not cuda.

warnings.warn("An input tensor was not cuda.")

step: 0, loss: 0.5984520316123962

Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 32768.0

step: 10, loss: 0.42718932032585144

Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0

step: 20, loss: 0.5498222708702087

step: 30, loss: 0.3638935685157776

step: 40, loss: 0.34833914041519165

```

step: 50, loss: 0.5551391243934631
step: 60, loss: 0.4008510112762451
step: 70, loss: 0.09775950014591217
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 8192.0
step: 80, loss: 0.27070823311805725
step: 90, loss: 0.4374099671840668
step: 100, loss: 0.14118532836437225
step: 110, loss: 0.1554541289806366
step: 120, loss: 0.024204542860388756
step: 130, loss: 0.019812086597085
step: 140, loss: 0.03315583989024162
step: 150, loss: 0.16039584577083588
step: 160, loss: 0.031033506616950035
step: 170, loss: 0.13657228648662567
step: 180, loss: 0.02576395682990551
step: 190, loss: 0.12410871684551239
step: 200, loss: 0.05258909612894058
step: 210, loss: 0.0173347070813179
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 4096.0
step: 220, loss: 0.0877537727355957
step: 230, loss: 0.005990250967442989
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please save or load the state of the optimizer when saving or
loading the scheduler.
    warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 1: dev_f1=0.9568106312292359, f1=0.9664429530201343,
best_f1=0.9664429530201343
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.006020561326295137
step: 10, loss: 0.13558059930801392
step: 20, loss: 0.019164470955729485
step: 30, loss: 0.011827470734715462
step: 40, loss: 0.06214994564652443
step: 50, loss: 0.006011516787111759
step: 60, loss: 0.007652635220438242
step: 70, loss: 0.0067881024442613125
step: 80, loss: 0.00464149983599782
step: 90, loss: 0.02060634084045887
step: 100, loss: 0.0022585620172321796
step: 110, loss: 0.042050834745168686
step: 120, loss: 0.011355521157383919
step: 130, loss: 0.01747809164226055
step: 140, loss: 0.00259228702634573

```

```

step: 150, loss: 0.31457433104515076
step: 160, loss: 0.00465061329305172
step: 170, loss: 0.0022357008419930935
step: 180, loss: 0.010901732370257378
step: 190, loss: 0.024962766095995903
step: 200, loss: 0.016204029321670532
step: 210, loss: 0.06863158196210861
step: 220, loss: 0.005816054064780474
step: 230, loss: 0.0018054797546938062
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
  warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 2: dev_f1=0.9751693002257337, f1=0.9682539682539683,
best_f1=0.9682539682539683
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.007156009320169687
step: 10, loss: 0.034903742372989655
step: 20, loss: 0.11783359199762344
step: 30, loss: 0.011844633147120476
step: 40, loss: 0.0380195751786232
step: 50, loss: 0.012967475689947605
step: 60, loss: 0.0558200441300869
step: 70, loss: 0.0054636928252875805
step: 80, loss: 0.1476423293352127
step: 90, loss: 0.10238035768270493
step: 100, loss: 0.01059657707810402
step: 110, loss: 0.011778812855482101
step: 120, loss: 0.0066818855702877045
step: 130, loss: 0.008661167696118355
step: 140, loss: 0.004742200952023268
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 2048.0
step: 150, loss: 0.08373145014047623
step: 160, loss: 0.0037326898891478777
step: 170, loss: 0.0012687684502452612
step: 180, loss: 0.014734789729118347
step: 190, loss: 0.007792624179273844
step: 200, loss: 0.007169085089117289
step: 210, loss: 0.006468594539910555
step: 220, loss: 0.17232593894004822
step: 230, loss: 0.012287082150578499
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.

```



```

    warnings.warn("An input tensor was not cuda.")
epoch 3: dev_f1=0.970917225950783, f1=0.9675977653631285,
best_f1=0.9682539682539683
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0077288043685257435
step: 10, loss: 0.0029848553240299225
step: 20, loss: 0.08203281462192535
step: 30, loss: 0.003827344160526991
step: 40, loss: 0.04352016746997833
step: 50, loss: 0.016800791025161743
step: 60, loss: 0.018203679472208023
step: 70, loss: 0.0023665959015488625
step: 80, loss: 0.20315106213092804
step: 90, loss: 0.13556602597236633
step: 100, loss: 0.024290869012475014
step: 110, loss: 0.0027290894649922848
step: 120, loss: 0.009873660281300545
step: 130, loss: 0.06626604497432709
step: 140, loss: 0.007340080104768276
step: 150, loss: 0.0033394074998795986
step: 160, loss: 0.001937682623974979
step: 170, loss: 0.0017892514588311315
step: 180, loss: 0.1484786421060562
step: 190, loss: 0.002166880527511239
step: 200, loss: 0.07611818611621857
step: 210, loss: 0.11050212383270264
step: 220, loss: 0.0021566441282629967
step: 230, loss: 0.008509943261742592
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
    warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 4: dev_f1=0.9797752808988766, f1=0.9797297297297298,
best_f1=0.9797297297297298
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.001224192208610475
step: 10, loss: 0.003430639859288931
step: 20, loss: 0.07083109021186829
step: 30, loss: 0.006844889838248491
step: 40, loss: 0.001847183215431869
step: 50, loss: 0.008182303979992867

```

```

step: 60, loss: 0.057728394865989685
step: 70, loss: 0.00308061670511961
step: 80, loss: 0.0734231173992157
step: 90, loss: 0.05154428258538246
step: 100, loss: 0.0007668087491765618
step: 110, loss: 0.0052461097948253155
step: 120, loss: 0.001042269403114915
step: 130, loss: 0.0006658771308138967
step: 140, loss: 0.0009606596431694925
step: 150, loss: 0.031710464507341385
step: 160, loss: 0.0012809844920411706
step: 170, loss: 0.003281685058027506
step: 180, loss: 0.01692095771431923
step: 190, loss: 0.20092041790485382
step: 200, loss: 0.005792588461190462
step: 210, loss: 0.024459993466734886
step: 220, loss: 0.009705904871225357
step: 230, loss: 0.019521890208125114
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
  warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 5: dev_f1=0.9887892376681614, f1=0.9787234042553192,
best_f1=0.9787234042553192
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.001856919494457543
step: 10, loss: 0.002496912609785795
step: 20, loss: 0.004153925459831953
step: 30, loss: 0.0011371142463758588
step: 40, loss: 0.00040354442899115384
step: 50, loss: 0.0019647893495857716
step: 60, loss: 0.004356268793344498
step: 70, loss: 0.15408381819725037
step: 80, loss: 0.006957049015909433
step: 90, loss: 0.022501006722450256
step: 100, loss: 0.00047880224883556366
step: 110, loss: 0.018222060054540634
step: 120, loss: 0.0010074460878968239
step: 130, loss: 0.0031597367487847805
step: 140, loss: 0.004096813499927521
step: 150, loss: 0.000422885874286294
step: 160, loss: 0.004408087115734816
step: 170, loss: 0.0007118411012925208

```

```

step: 180, loss: 0.02538570761680603
step: 190, loss: 0.0012531386455520988
step: 200, loss: 0.05499431490898132
step: 210, loss: 0.001310888328589499
step: 220, loss: 0.11003171652555466
step: 230, loss: 0.0038769147358834743
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 6: dev_f1=0.9796839729119639, f1=0.9808773903262092,
best_f1=0.9787234042553192
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.005344165489077568
step: 10, loss: 0.009502210654318333
step: 20, loss: 0.010597247630357742
step: 30, loss: 0.0007100831717252731
step: 40, loss: 0.0022086401004344225
step: 50, loss: 0.002013599267229438
step: 60, loss: 0.0020692353136837482
step: 70, loss: 0.0009408267214894295
step: 80, loss: 0.0008143267477862537
step: 90, loss: 0.001745975692756474
step: 100, loss: 0.0006878282292746007
step: 110, loss: 0.001969678793102503
step: 120, loss: 0.007440933957695961
step: 130, loss: 0.0014640484005212784
step: 140, loss: 0.0014250570675358176
step: 150, loss: 0.14179334044456482
step: 160, loss: 0.0008457309450022876
step: 170, loss: 0.009028825908899307
step: 180, loss: 0.0007163768168538809
step: 190, loss: 0.0005429504089988768
step: 200, loss: 0.007314393762499094
step: 210, loss: 0.01134366076439619
step: 220, loss: 0.0009226564434356987
step: 230, loss: 0.0010748824570327997
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 7: dev_f1=0.9798657718120806, f1=0.9764309764309763,
best_f1=0.9787234042553192
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0006236740155145526
step: 10, loss: 0.0025245118886232376

```

```

step: 20, loss: 0.0019261286361142993
step: 30, loss: 0.0008432514732703567
step: 40, loss: 0.0013194632483646274
step: 50, loss: 0.0006240542279556394
step: 60, loss: 0.0007641165284439921
step: 70, loss: 0.00030355615308508277
step: 80, loss: 0.005727794952690601
step: 90, loss: 0.005176333710551262
step: 100, loss: 0.0023358722683042288
step: 110, loss: 0.0016068306285887957
step: 120, loss: 0.0009621850913390517
step: 130, loss: 0.0014229780063033104
step: 140, loss: 0.002437508897855878
step: 150, loss: 0.2890245020389557
step: 160, loss: 0.002241429639980197
step: 170, loss: 0.0026122641284018755
step: 180, loss: 0.0006756905349902809
step: 190, loss: 0.0018797626253217459
step: 200, loss: 0.10150176286697388
step: 210, loss: 0.0022855147253721952
step: 220, loss: 0.0011671800166368484
step: 230, loss: 0.0016348984790965915
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 8: dev_f1=0.9843749999999999, f1=0.974472807991121,
best_f1=0.9787234042553192
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0032875067554414272
step: 10, loss: 0.006346951238811016
step: 20, loss: 0.0029080428648740053
step: 30, loss: 0.004927404690533876
step: 40, loss: 0.026520462706685066
step: 50, loss: 0.0034603625535964966
step: 60, loss: 0.00022220637765713036
step: 70, loss: 0.04965146631002426
step: 80, loss: 0.0002263364876853302
step: 90, loss: 0.04665902629494667
step: 100, loss: 0.0069363596849143505
step: 110, loss: 0.00049687857972458
step: 120, loss: 0.03472033515572548
step: 130, loss: 0.014653257094323635
step: 140, loss: 0.000405135506298393
step: 150, loss: 0.001096007996238768
step: 160, loss: 0.000667120679281652
step: 170, loss: 0.00039408064913004637

```

```

step: 180, loss: 0.0027473787777125835
step: 190, loss: 0.0002025719586526975
step: 200, loss: 0.10513922572135925
step: 210, loss: 0.039572328329086304
step: 220, loss: 0.0010130377486348152
step: 230, loss: 0.00017239096632692963
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 9: dev_f1=0.9854423292273236, f1=0.9764837625979844,
best_f1=0.9787234042553192
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0004228153557050973
step: 10, loss: 0.00043285443098284304
step: 20, loss: 0.0002472025225870311
step: 30, loss: 7.980452937772498e-05
step: 40, loss: 0.0002876388025470078
step: 50, loss: 0.00010029756958829239
step: 60, loss: 0.0001722384913591668
step: 70, loss: 0.001156515907496214
step: 80, loss: 0.00016092907753773034
step: 90, loss: 0.00020249173394404352
step: 100, loss: 0.00039350788574665785
step: 110, loss: 0.00020038132788613439
step: 120, loss: 0.001538734883069992
step: 130, loss: 0.0036600970197468996
step: 140, loss: 0.00038954385672695935
step: 150, loss: 0.0009132542181760073
step: 160, loss: 0.00010533343447605148
step: 170, loss: 0.014876809902489185
step: 180, loss: 0.0008445005514658988
step: 190, loss: 0.0003571596462279558
step: 200, loss: 0.0013234760845080018
step: 210, loss: 0.0005469411262311041
step: 220, loss: 0.0001745390909491107
step: 230, loss: 0.0023034014739096165
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 10: dev_f1=0.9854748603351955, f1=0.9776785714285714,
best_f1=0.9787234042553192
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.000381679943529889
step: 10, loss: 0.0014492508489638567

```

```

step: 20, loss: 0.000559053907636553
step: 30, loss: 0.0003978876047767699
step: 40, loss: 0.00011791673023253679
step: 50, loss: 0.0006526042125187814
step: 60, loss: 0.005382502917200327
step: 70, loss: 0.0003147753595840186
step: 80, loss: 0.00032754690619185567
step: 90, loss: 0.016244016587734222
step: 100, loss: 0.0009447532356716692
step: 110, loss: 0.00335517106577754
step: 120, loss: 0.00024360191309824586
step: 130, loss: 0.0002279071049997583
step: 140, loss: 0.003815488889813423
step: 150, loss: 0.004815369378775358
step: 160, loss: 0.00493348203599453
step: 170, loss: 0.0003492843243293464
step: 180, loss: 0.00015457553672604263
step: 190, loss: 0.00021755915076937526
step: 200, loss: 0.002027699025347829
step: 210, loss: 0.000343534629791975
step: 220, loss: 0.0010689424816519022
step: 230, loss: 0.0005244308267720044
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 11: dev_f1=0.9875706214689265, f1=0.9864253393665158,
best_f1=0.9787234042553192
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.002299182815477252
step: 10, loss: 0.0001738165447022766
step: 20, loss: 0.009769834578037262
step: 30, loss: 0.024843810126185417
step: 40, loss: 0.0004030441341456026
step: 50, loss: 0.0015578794991597533
step: 60, loss: 0.0004227684694342315
step: 70, loss: 0.00011807977716671303
step: 80, loss: 7.520569488406181e-05
step: 90, loss: 0.00037750921910628676
step: 100, loss: 9.43392951739952e-05
step: 110, loss: 0.00017183793534059078
step: 120, loss: 0.0002192926622228697
step: 130, loss: 0.00010143965482711792
step: 140, loss: 0.00048192418762482703
step: 150, loss: 0.00049580232007429
step: 160, loss: 0.00019195780623704195
step: 170, loss: 0.00031413117540068924

```

```

step: 180, loss: 0.00018876430112868547
step: 190, loss: 0.003616440575569868
step: 200, loss: 8.357840124517679e-05
step: 210, loss: 0.08667252957820892
step: 220, loss: 0.0072753834538161755
step: 230, loss: 0.0015311382012441754
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 12: dev_f1=0.9876543209876544, f1=0.9842696629213483,
best_f1=0.9787234042553192
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00014546026068273932
step: 10, loss: 0.004153982736170292
step: 20, loss: 0.00015652984438929707
step: 30, loss: 0.0010686679743230343
step: 40, loss: 0.008905882947146893
step: 50, loss: 0.00033846363658085465
step: 60, loss: 0.00016969481657724828
step: 70, loss: 0.008420511148869991
step: 80, loss: 0.0003580308984965086
step: 90, loss: 0.00012948320363648236
step: 100, loss: 0.0005775026511400938
step: 110, loss: 0.00017047712753992528
step: 120, loss: 0.00018452510994393378
step: 130, loss: 0.00015466794138774276
step: 140, loss: 0.0013479749904945493
step: 150, loss: 6.299790402408689e-05
step: 160, loss: 0.12257194519042969
step: 170, loss: 0.00010900036431849003
step: 180, loss: 0.02305101789534092
step: 190, loss: 0.00024126184871420264
step: 200, loss: 3.5086086427327245e-05
step: 210, loss: 0.0002257694286527112
step: 220, loss: 0.00018706111586652696
step: 230, loss: 0.00013280009443406016
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 13: dev_f1=0.9853107344632768, f1=0.9841986455981941,
best_f1=0.9787234042553192
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00010860967449843884
step: 10, loss: 9.57231386564672e-05

```

```

step: 20, loss: 0.00015718188660684973
step: 30, loss: 0.0008942022686824203
step: 40, loss: 0.00010023942741099745
step: 50, loss: 0.00014212891983333975
step: 60, loss: 0.00021225211094133556
step: 70, loss: 0.0001375207502860576
step: 80, loss: 0.0003052233369089663
step: 90, loss: 0.00011469821038190275
step: 100, loss: 0.00017950090114027262
step: 110, loss: 0.00010985131666529924
step: 120, loss: 2.837790088960901e-05
step: 130, loss: 0.00024418087559752166
step: 140, loss: 0.00011174065002705902
step: 150, loss: 9.943814802682027e-05
step: 160, loss: 0.00013274519005790353
step: 170, loss: 0.00010358324652770534
step: 180, loss: 8.799711940810084e-05
step: 190, loss: 6.553729326697066e-05
step: 200, loss: 0.0020452255848795176
step: 210, loss: 8.597845589974895e-05
step: 220, loss: 0.00010723005834734067
step: 230, loss: 0.0005721079069189727
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 14: dev_f1=0.9853107344632768, f1=0.9853438556933484,
best_f1=0.9787234042553192
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00915380660444498
step: 10, loss: 8.551800419809297e-05
step: 20, loss: 0.00015566794900223613
step: 30, loss: 0.00014363716763909906
step: 40, loss: 4.3223935790592805e-05
step: 50, loss: 8.545375749235973e-05
step: 60, loss: 0.03135685250163078
step: 70, loss: 9.045925980899483e-05
step: 80, loss: 9.179254266200587e-05
step: 90, loss: 8.158941636793315e-05
step: 100, loss: 5.027070801588707e-05
step: 110, loss: 0.0001912541629280895
step: 120, loss: 0.028016023337841034
step: 130, loss: 6.307358125923201e-05
step: 140, loss: 0.006438571494072676
step: 150, loss: 0.00017802843649405986
step: 160, loss: 0.10241960734128952
step: 170, loss: 5.4942349379416555e-05

```



```

step: 180, loss: 9.144929936155677e-05
step: 190, loss: 0.0003152031858917326
step: 200, loss: 0.0001432355202268809
step: 210, loss: 0.020312124863266945
step: 220, loss: 7.466768875019625e-05
step: 230, loss: 0.00012790273467544466
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 15: dev_f1=0.9853438556933484, f1=0.9853438556933484,
best_f1=0.9787234042553192
mv: cannot stat '*_dev.pt': No such file or directory
Selected optimization level 02:  FP16 training with FP32 batchnorm and FP32
master weights.

Defaults for this optimization level are:
enabled                : True
opt_level              : 02
cast_model_type        : torch.float16
patch_torch_functions  : False
keep_batchnorm_fp32    : True
master_weights         : True
loss_scale             : dynamic
Processing user overrides (additional kwargs that are not None)...
After processing overrides, optimization options are:
enabled                : True
opt_level              : 02
cast_model_type        : torch.float16
patch_torch_functions  : False
keep_batchnorm_fp32    : True
master_weights         : True
loss_scale             : dynamic
Warning:  multi_tensor_applier fused unscale kernel is unavailable, possibly
because apex was installed without --cuda_ext --cpp_ext. Using Python fallback.
Original ImportError was: ModuleNotFoundError("No module named 'amp_C'")
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
2473it [00:16, 150.64it/s]
load_f1 = 0.9876265466816648
real_f1 = 0.9843749999999999
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
4398it [00:31, 138.83it/s]

```

### 3.3.2 DBLP-GoogleScholar - Running the matcher

```
[ ]: # Step 1:
# train the wdc model with the save_model flag on
# it will produce two files *_dev.pt and *_test.pt which are the
# best checkpoints on the dev set and the test set.
!CUDA_VISIBLE_DEVICES=0 python train_ditto.py \
--task Dirty/DBLP-GoogleScholar \
--batch_size 32 \
--max_len 256 \
--lr 3e-5 \
--n_epochs 15 \
--finetuning \
--lm roberta \
--fp16 \
--save_model \
--dk general \
--dk product \
--summarize

# Step 2: pick a checkpoint, rename it to wdc_all_small.pt
!mv *_dev.pt checkpoints/DBLP-GoogleScholar/.pt

# Step 3: run the matcher
!CUDA_VISIBLE_DEVICES=0 python matcher.py \
--task Dirty/DBLP-GoogleScholar \
--input_path input/input_small.jsonl \
--output_path output/output_small.jsonl \
--lm roberta \
--use_gpu \
--fp16 \
--checkpoint_path checkpoints/
```

Selected optimization level 02: FP16 training with FP32 batchnorm and FP32 master weights.

Defaults for this optimization level are:

```
enabled           : True
opt_level         : 02
cast_model_type   : torch.float16
patch_torch_functions : False
keep_batchnorm_fp32 : True
master_weights    : True
loss_scale        : dynamic
```

Processing user overrides (additional kwargs that are not None)...

After processing overrides, optimization options are:

```

enabled                : True
opt_level              : 02
cast_model_type        : torch.float16
patch_torch_functions  : False
keep_batchnorm_fp32    : True
master_weights         : True
loss_scale             : dynamic
Warning: multi_tensor_applier fused unscale kernel is unavailable, possibly
because apex was installed without --cuda_ext --cpp_ext. Using Python fallback.
Original ImportError was: ModuleNotFoundError("No module named 'amp_C'")
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.6514492034912109
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 32768.0
step: 10, loss: 0.5018212795257568
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 20, loss: 0.3748336136341095
step: 30, loss: 0.3283054530620575
step: 40, loss: 0.3774401843547821
step: 50, loss: 0.697770893573761
step: 60, loss: 0.37956109642982483
step: 70, loss: 0.4686849117279053
step: 80, loss: 0.5074924230575562
step: 90, loss: 0.5034087300300598
step: 100, loss: 0.5829089879989624
step: 110, loss: 0.4603719711303711
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 8192.0
step: 120, loss: 0.6989585757255554
step: 130, loss: 0.60938560962677
step: 140, loss: 0.2535564601421356
step: 150, loss: 0.1480826735496521
step: 160, loss: 0.32769307494163513
step: 170, loss: 0.17402450740337372
step: 180, loss: 0.09342386573553085
step: 190, loss: 0.10920058935880661
step: 200, loss: 0.06479814648628235
step: 210, loss: 0.04509948566555977
step: 220, loss: 0.011041023768484592
step: 230, loss: 0.281588077545166
step: 240, loss: 0.11108553409576416
step: 250, loss: 0.015403153374791145
step: 260, loss: 0.22885674238204956
step: 270, loss: 0.4121250808238983
step: 280, loss: 0.059919826686382294
step: 290, loss: 0.1471765786409378
step: 300, loss: 0.056142255663871765
step: 310, loss: 0.02375340461730957

```

```

step: 320, loss: 0.0825393944978714
step: 330, loss: 0.16945035755634308
step: 340, loss: 0.30094826221466064
step: 350, loss: 0.12621371448040009
step: 360, loss: 0.08308039605617523
step: 370, loss: 0.2117515653371811
step: 380, loss: 0.20517003536224365
step: 390, loss: 0.02752593904733658
step: 400, loss: 0.10754825174808502
step: 410, loss: 0.25559839606285095
step: 420, loss: 0.012886294163763523
step: 430, loss: 0.012362330220639706
step: 440, loss: 0.030768465250730515
step: 450, loss: 0.10135538130998611
step: 460, loss: 0.009206346236169338
step: 470, loss: 0.02980513870716095
step: 480, loss: 0.06654348969459534
step: 490, loss: 0.04836400970816612
step: 500, loss: 0.002185232937335968
step: 510, loss: 0.028160499408841133
step: 520, loss: 0.028834765776991844
step: 530, loss: 0.008497850969433784
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
    warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 1: dev_f1=0.9430523917995445, f1=0.9421860885275519,
best_f1=0.9421860885275519
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.11382545530796051
step: 10, loss: 0.02413758635520935
step: 20, loss: 0.0031519816257059574
step: 30, loss: 0.035376984626054764
step: 40, loss: 0.05467017740011215
step: 50, loss: 0.019219113513827324
step: 60, loss: 0.011496120132505894
step: 70, loss: 0.010318100452423096
step: 80, loss: 0.005986243486404419
step: 90, loss: 0.0025263838469982147
step: 100, loss: 0.1680520921945572
step: 110, loss: 0.013117462396621704
step: 120, loss: 0.00833122432231903
step: 130, loss: 0.004235878586769104

```

```

step: 140, loss: 0.12899702787399292
step: 150, loss: 0.03092566318809986
step: 160, loss: 0.03446914255619049
step: 170, loss: 0.08170506358146667
step: 180, loss: 0.10075435042381287
step: 190, loss: 0.01793622598052025
step: 200, loss: 0.13647951185703278
step: 210, loss: 0.035381659865379333
step: 220, loss: 0.0014747882960364223
step: 230, loss: 0.05780303105711937
step: 240, loss: 0.03770042955875397
step: 250, loss: 0.031130632385611534
step: 260, loss: 0.023219605907797813
step: 270, loss: 0.03156403824687004
step: 280, loss: 0.04176396131515503
step: 290, loss: 0.055533405393362045
step: 300, loss: 0.04923166707158089
step: 310, loss: 0.02935565449297428
step: 320, loss: 0.03851822018623352
step: 330, loss: 0.06099012494087219
step: 340, loss: 0.022361844778060913
step: 350, loss: 0.0019752124790102243
step: 360, loss: 0.09452594816684723
step: 370, loss: 0.00563944922760129
step: 380, loss: 0.12658368051052094
step: 390, loss: 0.006308695301413536
step: 400, loss: 0.04124089330434799
step: 410, loss: 0.015590551309287548
step: 420, loss: 0.01861557923257351
step: 430, loss: 0.2617884576320648
step: 440, loss: 0.016929542645812035
step: 450, loss: 0.009382077492773533
step: 460, loss: 0.01727825030684471
step: 470, loss: 0.06142552196979523
step: 480, loss: 0.0018306535203009844
step: 490, loss: 0.05941711738705635
step: 500, loss: 0.0021087785717099905
step: 510, loss: 0.020938690751791
step: 520, loss: 0.2956219017505646
step: 530, loss: 0.03488526493310928
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
  warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 2: dev_f1=0.9529520295202951, f1=0.951851851851852,

```

```
best_f1=0.951851851851852
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.07586350291967392
step: 10, loss: 0.015673037618398666
step: 20, loss: 0.010217353701591492
step: 30, loss: 0.04302928224205971
step: 40, loss: 0.030421525239944458
step: 50, loss: 0.007110259961336851
step: 60, loss: 0.010428430512547493
step: 70, loss: 0.010698557831346989
step: 80, loss: 0.03221698850393295
step: 90, loss: 0.02840222604572773
step: 100, loss: 0.010330341756343842
step: 110, loss: 0.03630389645695686
step: 120, loss: 0.18490386009216309
step: 130, loss: 0.0025812475942075253
step: 140, loss: 0.02642999030649662
step: 150, loss: 0.03261502459645271
step: 160, loss: 0.015431663021445274
step: 170, loss: 0.08064527809619904
step: 180, loss: 0.04123909771442413
step: 190, loss: 0.004011665005236864
step: 200, loss: 0.01203322783112526
step: 210, loss: 0.0466834232211113
step: 220, loss: 0.06367429345846176
step: 230, loss: 0.01842690259218216
step: 240, loss: 0.07209501415491104
step: 250, loss: 0.05788882449269295
step: 260, loss: 0.16570229828357697
step: 270, loss: 0.0029410815332084894
step: 280, loss: 0.027508841827511787
step: 290, loss: 0.014876171946525574
step: 300, loss: 0.062249500304460526
step: 310, loss: 0.08166085183620453
step: 320, loss: 0.03946354612708092
step: 330, loss: 0.0025540783535689116
step: 340, loss: 0.0030171149410307407
step: 350, loss: 0.14020203053951263
step: 360, loss: 0.006418651435524225
step: 370, loss: 0.00912783108651638
step: 380, loss: 0.005828164983540773
step: 390, loss: 0.0034174916800111532
step: 400, loss: 0.15089671313762665
step: 410, loss: 0.0654299333691597
step: 420, loss: 0.009035558439791203
step: 430, loss: 0.012724760919809341
```

```

step: 440, loss: 0.22832874953746796
step: 450, loss: 0.019027728587388992
step: 460, loss: 0.054926797747612
step: 470, loss: 0.07107628136873245
step: 480, loss: 0.06641481816768646
step: 490, loss: 0.053254976868629456
step: 500, loss: 0.009410979226231575
step: 510, loss: 0.014588127844035625
step: 520, loss: 0.004474174231290817
step: 530, loss: 0.008749657310545444
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
    warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 3: dev_f1=0.9565217391304348, f1=0.9469234382339127,
best_f1=0.9469234382339127
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.007273627910763025
step: 10, loss: 0.004486429505050182
step: 20, loss: 0.0020670804660767317
step: 30, loss: 0.11036471277475357
step: 40, loss: 0.031175723299384117
step: 50, loss: 0.04552844911813736
step: 60, loss: 0.005629685241729021
step: 70, loss: 0.04541079327464104
step: 80, loss: 0.027166038751602173
step: 90, loss: 0.07094614952802658
step: 100, loss: 0.0013934478629380465
step: 110, loss: 0.07793554663658142
step: 120, loss: 0.0027994192205369473
step: 130, loss: 0.004654242657124996
step: 140, loss: 0.0035764514468610287
step: 150, loss: 0.004017231985926628
step: 160, loss: 0.010565713047981262
step: 170, loss: 0.024984830990433693
step: 180, loss: 0.06700241565704346
step: 190, loss: 0.038829509168863297
step: 200, loss: 0.04764939844608307
step: 210, loss: 0.0010927735129371285
step: 220, loss: 0.0013852508272975683
step: 230, loss: 0.0010727994376793504
step: 240, loss: 0.04455043375492096
step: 250, loss: 0.035370223224163055

```

```

step: 260, loss: 0.0016838323790580034
step: 270, loss: 0.15785565972328186
step: 280, loss: 0.007233244366943836
step: 290, loss: 0.16209843754768372
step: 300, loss: 0.017048748210072517
step: 310, loss: 0.009225661866366863
step: 320, loss: 0.10470637679100037
step: 330, loss: 0.028517546132206917
step: 340, loss: 0.00565674714744091
step: 350, loss: 0.012575945816934109
step: 360, loss: 0.027741076424717903
step: 370, loss: 0.0052744559943675995
step: 380, loss: 0.005656518507748842
step: 390, loss: 0.0017851351294666529
step: 400, loss: 0.006641289684921503
step: 410, loss: 0.0009920260636135936
step: 420, loss: 0.048921458423137665
step: 430, loss: 0.008212201297283173
step: 440, loss: 0.007288556080311537
step: 450, loss: 0.013194775208830833
step: 460, loss: 0.012927507981657982
step: 470, loss: 0.00088180520106107
step: 480, loss: 0.004513909574598074
step: 490, loss: 0.0009227999253198504
step: 500, loss: 0.010668034665286541
step: 510, loss: 0.008774635381996632
step: 520, loss: 0.0064044552855193615
step: 530, loss: 0.11459740251302719
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
    warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 4: dev_f1=0.9614132961413296, f1=0.9505773672055428,
best_f1=0.9505773672055428
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.016741756349802017
step: 10, loss: 0.005453494843095541
step: 20, loss: 0.005332337226718664
step: 30, loss: 0.005420473404228687
step: 40, loss: 0.0016753452364355326
step: 50, loss: 0.06863538175821304
step: 60, loss: 0.005598381161689758
step: 70, loss: 0.0015201537171378732

```



```
step: 80, loss: 0.0019183275289833546
step: 90, loss: 0.07905292510986328
step: 100, loss: 0.05693904310464859
step: 110, loss: 0.014545273967087269
step: 120, loss: 0.07278184592723846
step: 130, loss: 0.008842200972139835
step: 140, loss: 0.0021032774820923805
step: 150, loss: 0.002468785736709833
step: 160, loss: 0.007963579148054123
step: 170, loss: 0.02331322617828846
step: 180, loss: 0.05030115693807602
step: 190, loss: 0.002738191979005933
step: 200, loss: 0.00239853048697114
step: 210, loss: 0.0007321801967918873
step: 220, loss: 0.0006765917059965432
step: 230, loss: 0.003706778399646282
step: 240, loss: 0.009974656626582146
step: 250, loss: 0.07199603319168091
step: 260, loss: 0.0012044006725773215
step: 270, loss: 0.01404645573347807
step: 280, loss: 0.0020143508445471525
step: 290, loss: 0.00181865063495934
step: 300, loss: 0.003963591530919075
step: 310, loss: 0.026806749403476715
step: 320, loss: 0.004487111698836088
step: 330, loss: 0.00011697388981701806
step: 340, loss: 0.004486844874918461
step: 350, loss: 0.0002993004454765469
step: 360, loss: 0.00021803524577990174
step: 370, loss: 0.0003645919496193528
step: 380, loss: 5.281824996927753e-05
step: 390, loss: 0.01474523451179266
step: 400, loss: 0.009776557795703411
step: 410, loss: 0.034315142780542374
step: 420, loss: 0.3069821000099182
step: 430, loss: 0.04071582481265068
step: 440, loss: 0.0019691975321620703
step: 450, loss: 0.005790284834802151
step: 460, loss: 0.004527881275862455
step: 470, loss: 0.026118597015738487
step: 480, loss: 0.0023131680209189653
step: 490, loss: 0.040207237005233765
step: 500, loss: 0.001829130807891488
step: 510, loss: 0.0034737056121230125
step: 520, loss: 0.07845645397901535
step: 530, loss: 0.051782459020614624
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
```

```

    warnings.warn("An input tensor was not cuda.")
epoch 5: dev_f1=0.9562790697674419, f1=0.9449112978524743,
best_f1=0.9505773672055428
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.014607292599976063
step: 10, loss: 0.0006223949603736401
step: 20, loss: 9.018955461215228e-05
step: 30, loss: 0.004631803836673498
step: 40, loss: 0.0001408774551237002
step: 50, loss: 0.046601343899965286
step: 60, loss: 0.04884486272931099
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 8192.0
step: 70, loss: 0.001587378908880055
step: 80, loss: 0.00021150760585442185
step: 90, loss: 0.0481121726334095
step: 100, loss: 0.005644692108035088
step: 110, loss: 0.0066298432648181915
step: 120, loss: 0.03591784089803696
step: 130, loss: 0.005617642775177956
step: 140, loss: 0.0004748800420202315
step: 150, loss: 0.0003717295767273754
step: 160, loss: 0.04080702364444733
step: 170, loss: 0.001791209913790226
step: 180, loss: 0.0024501706939190626
step: 190, loss: 0.04636605456471443
step: 200, loss: 0.043237920850515366
step: 210, loss: 0.013877220451831818
step: 220, loss: 0.0016845628852024674
step: 230, loss: 0.0074521806091070175
step: 240, loss: 0.0018906296463683248
step: 250, loss: 0.008142326958477497
step: 260, loss: 0.0005776575999334455
step: 270, loss: 0.00264237099327147
step: 280, loss: 0.0022651178296655416
step: 290, loss: 0.001337422407232225
step: 300, loss: 0.011400900781154633
step: 310, loss: 0.07606635987758636
step: 320, loss: 4.48975624749437e-05
step: 330, loss: 0.00018576622824184597
step: 340, loss: 0.005205463618040085
step: 350, loss: 0.013933906331658363
step: 360, loss: 0.18112479150295258
step: 370, loss: 0.010616262443363667
step: 380, loss: 0.0048194145783782005
step: 390, loss: 0.004698953125625849
step: 400, loss: 0.022510254755616188

```

```

step: 410, loss: 0.00040239625377580523
step: 420, loss: 0.00046784974983893335
step: 430, loss: 0.00230485200881958
step: 440, loss: 0.0021396519150584936
step: 450, loss: 0.35482361912727356
step: 460, loss: 0.0016741665313020349
step: 470, loss: 0.0013974817702546716
step: 480, loss: 0.0037640356458723545
step: 490, loss: 0.027410492300987244
step: 500, loss: 0.0010226302547380328
step: 510, loss: 0.12328638136386871
step: 520, loss: 0.0003477085556369275
step: 530, loss: 0.0014065688010305166
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
epoch 6: dev_f1=0.9444444444444445, f1=0.9388322520852641,
best_f1=0.9505773672055428
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0022827007342129946
step: 10, loss: 0.004535502754151821
step: 20, loss: 0.001044274540618062
step: 30, loss: 0.09366942942142487
step: 40, loss: 0.027046354487538338
step: 50, loss: 0.050506841391325
step: 60, loss: 0.003457972314208746
step: 70, loss: 0.002765171229839325
step: 80, loss: 7.840709440642968e-05
step: 90, loss: 0.055588752031326294
step: 100, loss: 0.0010046589886769652
step: 110, loss: 5.7952718634624034e-05
step: 120, loss: 9.416989632882178e-05
step: 130, loss: 4.708898995886557e-05
step: 140, loss: 4.387801163829863e-05
step: 150, loss: 0.22637012600898743
step: 160, loss: 0.0002510325866751373
step: 170, loss: 0.021109722554683685
step: 180, loss: 0.0019095191964879632
step: 190, loss: 0.07087036967277527
step: 200, loss: 0.0005693403072655201
step: 210, loss: 0.0020889558363705873
step: 220, loss: 5.5247619457077235e-05
step: 230, loss: 0.008121484890580177
step: 240, loss: 0.021825652569532394
step: 250, loss: 0.0012410989729687572
step: 260, loss: 0.007603295147418976

```

```

step: 270, loss: 0.000725015124771744
step: 280, loss: 0.0005069767357781529
step: 290, loss: 0.0013386700302362442
step: 300, loss: 4.194541907054372e-05
step: 310, loss: 0.0010672343196347356
step: 320, loss: 0.0008061568951234221
step: 330, loss: 0.00016123945533763617
step: 340, loss: 0.018547825515270233
step: 350, loss: 0.0030072559602558613
step: 360, loss: 0.000499499321449548
step: 370, loss: 0.0013722634175792336
step: 380, loss: 0.007280934136360884
step: 390, loss: 0.02165781892836094
step: 400, loss: 0.015676258131861687
step: 410, loss: 0.0023329060059040785
step: 420, loss: 0.10041225701570511
step: 430, loss: 0.0001038494156091474
step: 440, loss: 0.00431694695726037
step: 450, loss: 0.0008768749539740384
step: 460, loss: 7.09716186975129e-05
step: 470, loss: 0.11571270227432251
step: 480, loss: 0.002796467859297991
step: 490, loss: 0.0008213006658479571
step: 500, loss: 0.0006992471171543002
step: 510, loss: 0.00011554529191926122
step: 520, loss: 5.1656570576597005e-05
step: 530, loss: 0.00011037544754799455
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
epoch 7: dev_f1=0.9592887225081891, f1=0.9514925373134329,
best_f1=0.9505773672055428
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0009529628441669047
step: 10, loss: 0.002473542932420969
step: 20, loss: 0.0009955356363207102
step: 30, loss: 7.40067771403119e-05
step: 40, loss: 6.241181108634919e-05
step: 50, loss: 0.0007278419216163456
step: 60, loss: 0.0029835840687155724
step: 70, loss: 0.00044551739119924605
step: 80, loss: 0.004804398864507675
step: 90, loss: 6.194314482854679e-05
step: 100, loss: 6.707354623358697e-05
step: 110, loss: 0.0009324072161689401
step: 120, loss: 0.009232616983354092

```

```

step: 130, loss: 6.657273479504511e-05
step: 140, loss: 3.360856499057263e-05
step: 150, loss: 0.0008470022585242987
step: 160, loss: 3.0248747862060554e-05
step: 170, loss: 0.16657598316669464
step: 180, loss: 0.0004605463473126292
step: 190, loss: 0.13827209174633026
step: 200, loss: 0.013598443940281868
step: 210, loss: 0.03149246796965599
step: 220, loss: 0.0035558249801397324
step: 230, loss: 0.04350147768855095
step: 240, loss: 0.03588453307747841
step: 250, loss: 0.0015345001593232155
step: 260, loss: 0.0005751443095505238
step: 270, loss: 0.002829099539667368
step: 280, loss: 0.000180173316039145
step: 290, loss: 0.007880885154008865
step: 300, loss: 0.011620289646089077
step: 310, loss: 0.10683901607990265
step: 320, loss: 0.015399365685880184
step: 330, loss: 0.00015802847337909043
step: 340, loss: 0.00420196820050478
step: 350, loss: 0.00023088727903086692
step: 360, loss: 0.0012225318932905793
step: 370, loss: 0.016338087618350983
step: 380, loss: 2.9533284759963863e-05
step: 390, loss: 0.0014626202173531055
step: 400, loss: 0.0025537540204823017
step: 410, loss: 0.0001799832098186016
step: 420, loss: 3.966010262956843e-05
step: 430, loss: 0.0016595000633969903
step: 440, loss: 0.001380119938403368
step: 450, loss: 5.788622365798801e-05
step: 460, loss: 0.00034217536449432373
step: 470, loss: 0.09526439756155014
step: 480, loss: 0.006227701902389526
step: 490, loss: 0.011821568943560123
step: 500, loss: 3.0311563023133203e-05
step: 510, loss: 0.0025519696064293385
step: 520, loss: 0.000708109640982002
step: 530, loss: 0.03919866308569908
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
  warnings.warn(SAVE_STATE_WARNING, UserWarning)

```

```
epoch 8: dev_f1=0.9648382559774966, f1=0.9489939167056621,
best_f1=0.9489939167056621
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0005939237307757139
step: 10, loss: 0.01671808399260044
step: 20, loss: 0.00031227045110426843
step: 30, loss: 0.07373615354299545
step: 40, loss: 0.0009949593804776669
step: 50, loss: 0.003282950259745121
step: 60, loss: 0.0010761437006294727
step: 70, loss: 0.009669537656009197
step: 80, loss: 0.005432739853858948
step: 90, loss: 0.07194314897060394
step: 100, loss: 0.001427085604518652
step: 110, loss: 0.0006415662355720997
step: 120, loss: 0.0016898084431886673
step: 130, loss: 0.0001541330711916089
step: 140, loss: 0.005894002038985491
step: 150, loss: 6.199034396559e-05
step: 160, loss: 0.00013073465379420668
step: 170, loss: 0.0015385658480226994
step: 180, loss: 0.0001613395579624921
step: 190, loss: 2.9980657927808352e-05
step: 200, loss: 0.001221961691044271
step: 210, loss: 0.00021240292699076235
step: 220, loss: 0.002789605176076293
step: 230, loss: 0.0016354973195120692
step: 240, loss: 0.0015028647612780333
step: 250, loss: 0.00035225573810748756
step: 260, loss: 8.926431473810226e-05
step: 270, loss: 0.013227446936070919
step: 280, loss: 0.01830524019896984
step: 290, loss: 9.62523918133229e-05
step: 300, loss: 0.01711539551615715
step: 310, loss: 0.07631947845220566
step: 320, loss: 0.0009895280236378312
step: 330, loss: 0.0010959521168842912
step: 340, loss: 0.0033865247387439013
step: 350, loss: 0.014047540724277496
step: 360, loss: 0.00012729964510072023
step: 370, loss: 0.00037288718158379197
step: 380, loss: 0.020838573575019836
step: 390, loss: 0.00033492714283056557
step: 400, loss: 0.012196565046906471
step: 410, loss: 0.0008526762248948216
step: 420, loss: 0.0002042620035354048
```

```

step: 430, loss: 0.027176951989531517
step: 440, loss: 5.40232467756141e-05
step: 450, loss: 0.0027198633179068565
step: 460, loss: 0.0006253889878280461
step: 470, loss: 8.794186578597873e-05
step: 480, loss: 0.0004602974804583937
step: 490, loss: 0.04135938733816147
step: 500, loss: 9.978526213672012e-05
step: 510, loss: 0.0001456728350603953
step: 520, loss: 0.0027909018099308014
step: 530, loss: 0.03614909201860428
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
epoch 9: dev_f1=0.9602246139447823, f1=0.9560336763330215,
best_f1=0.9489939167056621
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0007558684446848929
step: 10, loss: 0.0012416485697031021
step: 20, loss: 0.0002015324280364439
step: 30, loss: 0.001051644328981638
step: 40, loss: 6.717952055623755e-05
step: 50, loss: 5.493815478985198e-05
step: 60, loss: 0.00958230346441269
step: 70, loss: 4.075278411619365e-05
step: 80, loss: 8.248307130998e-05
step: 90, loss: 0.0001840540353441611
step: 100, loss: 0.0002611571690067649
step: 110, loss: 0.0005926615558564663
step: 120, loss: 0.0002335003373445943
step: 130, loss: 0.00015251780860126019
step: 140, loss: 0.00043111739796586335
step: 150, loss: 0.00020249201043043286
step: 160, loss: 0.013908018358051777
step: 170, loss: 6.720233068335801e-05
step: 180, loss: 0.06738343089818954
step: 190, loss: 3.3752476156223565e-05
step: 200, loss: 0.00016326672630384564
step: 210, loss: 0.0421038381755352
step: 220, loss: 0.0007719050045125186
step: 230, loss: 0.00013282620056997985
step: 240, loss: 6.861500150989741e-05
step: 250, loss: 0.0008120521088130772
step: 260, loss: 0.007301235105842352
step: 270, loss: 3.0453307772404514e-05
step: 280, loss: 0.018138684332370758

```

```

step: 290, loss: 2.7651734853861853e-05
step: 300, loss: 0.00016811888781376183
step: 310, loss: 0.027911534532904625
step: 320, loss: 5.5466640333179384e-05
step: 330, loss: 0.00012482519377954304
step: 340, loss: 2.687705637072213e-05
step: 350, loss: 0.0005796945188194513
step: 360, loss: 2.7730104193324223e-05
step: 370, loss: 0.00041802425403147936
step: 380, loss: 0.004128237254917622
step: 390, loss: 0.0020866096019744873
step: 400, loss: 0.0005810183356516063
step: 410, loss: 0.0017513827187940478
step: 420, loss: 0.0001243677834281698
step: 430, loss: 0.0013255870435386896
step: 440, loss: 6.383621803252026e-05
step: 450, loss: 0.0005411768215708435
step: 460, loss: 0.0001019733608700335
step: 470, loss: 0.001330261817201972
step: 480, loss: 0.0006596780149266124
step: 490, loss: 9.834478260017931e-05
step: 500, loss: 0.00022278794494923204
step: 510, loss: 2.4876997485989705e-05
step: 520, loss: 0.021714285016059875
step: 530, loss: 0.0013906399253755808
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
epoch 10: dev_f1=0.9604834960483496, f1=0.9521597770552717,
best_f1=0.9489939167056621
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 9.372473869007081e-05
step: 10, loss: 0.04315844178199768
step: 20, loss: 0.002424001693725586
step: 30, loss: 8.526410965714604e-05
step: 40, loss: 3.3030955819413066e-05
step: 50, loss: 0.0001673151709837839
step: 60, loss: 3.0031849746592343e-05
step: 70, loss: 0.0001292884408030659
step: 80, loss: 0.001156882382929325
step: 90, loss: 0.0009406417375430465
step: 100, loss: 0.0005173757672309875
step: 110, loss: 0.00011994967644568533
step: 120, loss: 3.488957372610457e-05
step: 130, loss: 3.4567827242426574e-05
step: 140, loss: 2.9093083867337555e-05

```



```

step: 150, loss: 0.0005397791974246502
step: 160, loss: 0.0037000258453190327
step: 170, loss: 0.030150754377245903
step: 180, loss: 2.5100000129896216e-05
step: 190, loss: 0.00013907266838941723
step: 200, loss: 5.2108545787632465e-05
step: 210, loss: 0.0003310051979497075
step: 220, loss: 0.0015910703223198652
step: 230, loss: 2.9181901481933892e-05
step: 240, loss: 0.007203155662864447
step: 250, loss: 0.0006310866447165608
step: 260, loss: 0.003193818498402834
step: 270, loss: 0.0023856826592236757
step: 280, loss: 0.0034843862522393465
step: 290, loss: 6.585746450582519e-05
step: 300, loss: 3.2109881431097165e-05
step: 310, loss: 0.008547810837626457
step: 320, loss: 6.150194531073794e-05
step: 330, loss: 0.00019713137589860708
step: 340, loss: 0.00018544652266427875
step: 350, loss: 2.467525700922124e-05
step: 360, loss: 0.0005079206894151866
step: 370, loss: 0.0017985888989642262
step: 380, loss: 0.00029978761449456215
step: 390, loss: 0.0014189800713211298
step: 400, loss: 5.8531048125587404e-05
step: 410, loss: 0.0012913577957078815
step: 420, loss: 0.0003360139380674809
step: 430, loss: 0.000925498956348747
step: 440, loss: 2.517848406569101e-05
step: 450, loss: 0.0011793191079050303
step: 460, loss: 0.0002073423092951998
step: 470, loss: 0.0012905837502330542
step: 480, loss: 0.0009670996805652976
step: 490, loss: 7.482152432203293e-05
step: 500, loss: 0.0013274658704176545
step: 510, loss: 0.1641802191734314
step: 520, loss: 0.0005649367230944335
step: 530, loss: 0.17601093649864197
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
epoch 11: dev_f1=0.9589169000933706, f1=0.947565543071161,
best_f1=0.9489939167056621
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 1.793319097487256e-05

```

step: 10, loss: 0.002217124216258526  
step: 20, loss: 0.00907539576292038  
step: 30, loss: 0.00251760077662766  
step: 40, loss: 0.002372037386521697  
step: 50, loss: 2.9355662263697013e-05  
step: 60, loss: 7.358216680586338e-05  
step: 70, loss: 0.00015405254089273512  
step: 80, loss: 2.0786737877642736e-05  
step: 90, loss: 7.506069960072637e-05  
step: 100, loss: 0.05202285945415497  
step: 110, loss: 7.679929694859311e-05  
step: 120, loss: 0.0015521678142249584  
step: 130, loss: 0.0013153916224837303  
step: 140, loss: 0.00036368530709296465  
step: 150, loss: 8.053462079260498e-05  
step: 160, loss: 6.955632852623239e-05  
step: 170, loss: 9.246846457244828e-05  
step: 180, loss: 2.1501551600522362e-05  
step: 190, loss: 2.0063929696334526e-05  
step: 200, loss: 0.0015266326954588294  
step: 210, loss: 0.005550894886255264  
step: 220, loss: 1.3038385077379644e-05  
step: 230, loss: 2.6149562472710386e-05  
step: 240, loss: 0.0015185467200353742  
step: 250, loss: 2.277072962897364e-05  
step: 260, loss: 1.2621171663340647e-05  
step: 270, loss: 0.005163254216313362  
step: 280, loss: 5.3682215366279706e-05  
step: 290, loss: 0.004031024873256683  
step: 300, loss: 0.001322217402048409  
step: 310, loss: 1.955730840563774e-05  
step: 320, loss: 0.0016922985669225454  
step: 330, loss: 0.000867294380441308  
step: 340, loss: 0.000965824699960649  
step: 350, loss: 4.781128518516198e-05  
step: 360, loss: 0.0020381987560540438  
step: 370, loss: 6.68104985379614e-05  
step: 380, loss: 0.0012699199141934514  
step: 390, loss: 7.861207268433645e-05  
step: 400, loss: 1.161161344498396e-05  
step: 410, loss: 0.0351557694375515  
step: 420, loss: 1.1518502105900552e-05  
step: 430, loss: 1.4908338926034048e-05  
step: 440, loss: 5.74022124055773e-05  
step: 450, loss: 2.1553805709118024e-05  
step: 460, loss: 2.499782931408845e-05  
step: 470, loss: 0.0028032485861331224  
step: 480, loss: 0.0324464850127697

```

step: 490, loss: 1.3846754882251844e-05
step: 500, loss: 4.779902519658208e-05
step: 510, loss: 0.01255557406693697
step: 520, loss: 0.0014699840685352683
step: 530, loss: 0.00010174669296247885
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 12: dev_f1=0.9594027064862343, f1=0.950957496496964,
best_f1=0.9489939167056621
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 1.5608649846399203e-05
step: 10, loss: 6.294997729128227e-05
step: 20, loss: 1.457689722883515e-05
step: 30, loss: 1.1071453627664596e-05
step: 40, loss: 0.0006088813534006476
step: 50, loss: 0.0024092388339340687
step: 60, loss: 1.7068754459614865e-05
step: 70, loss: 1.3962154298496898e-05
step: 80, loss: 1.2520562449935824e-05
step: 90, loss: 0.0002914709330070764
step: 100, loss: 1.157806946139317e-05
step: 110, loss: 1.0248182661598548e-05
step: 120, loss: 1.1924518730666023e-05
step: 130, loss: 0.0014159079873934388
step: 140, loss: 1.3451879567583092e-05
step: 150, loss: 0.0016701295971870422
step: 160, loss: 1.730001349642407e-05
step: 170, loss: 0.0015026129549369216
step: 180, loss: 1.1891002031916287e-05
step: 190, loss: 0.000167959020473063
step: 200, loss: 1.4535823538608383e-05
step: 210, loss: 1.2457202501536813e-05
step: 220, loss: 1.0490337444934994e-05
step: 230, loss: 0.000600476807449013
step: 240, loss: 0.0007444385555572808
step: 250, loss: 0.00037946333759464324
step: 260, loss: 1.3153792679077014e-05
step: 270, loss: 0.00012965381029061973
step: 280, loss: 3.114011633442715e-05
step: 290, loss: 8.821428309602197e-06
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 300, loss: 1.1034211638616398e-05
step: 310, loss: 1.564193189551588e-05
step: 320, loss: 3.1683175620855764e-05
step: 330, loss: 1.3314056559465826e-05

```

```

step: 340, loss: 0.00011372638982720673
step: 350, loss: 1.0527571248530876e-05
step: 360, loss: 0.12538285553455353
step: 370, loss: 1.973053440451622e-05
step: 380, loss: 5.8781639381777495e-05
step: 390, loss: 9.383949873154052e-06
step: 400, loss: 4.399471799843013e-05
step: 410, loss: 1.797749246179592e-05
step: 420, loss: 1.1369467756594531e-05
step: 430, loss: 1.1078912393713836e-05
step: 440, loss: 9.484508154855575e-06
step: 450, loss: 9.798612154554576e-05
step: 460, loss: 1.6614516425761394e-05
step: 470, loss: 0.0019569657742977142
step: 480, loss: 7.1934978222998325e-06
step: 490, loss: 9.279628102376591e-06
step: 500, loss: 7.69267717259936e-06
step: 510, loss: 1.76491685124347e-05
step: 520, loss: 1.2151604096288793e-05
step: 530, loss: 0.00042540807044133544
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 13: dev_f1=0.9584513692162419, f1=0.9438414346389806,
best_f1=0.9489939167056621
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0014860890805721283
step: 10, loss: 0.00104562658816576
step: 20, loss: 9.193950063490774e-06
step: 30, loss: 0.00010416088480269536
step: 40, loss: 0.0010510606225579977
step: 50, loss: 0.0005142392474226654
step: 60, loss: 1.3488981494447216e-05
step: 70, loss: 9.730390047479887e-06
step: 80, loss: 1.5306983186746947e-05
step: 90, loss: 9.05611250345828e-06
step: 100, loss: 1.1097491551481653e-05
step: 110, loss: 0.00019275907834526151
step: 120, loss: 9.715480700833723e-06
step: 130, loss: 4.604926652973518e-05
step: 140, loss: 1.2803640856873244e-05
step: 150, loss: 0.0014969115145504475
step: 160, loss: 1.1369459571142215e-05
step: 170, loss: 9.927820428856649e-06
step: 180, loss: 3.916150308214128e-05
step: 190, loss: 1.1086342055932619e-05

```

```
step: 200, loss: 1.2788726962753572e-05
step: 210, loss: 0.0011845314875245094
step: 220, loss: 1.0274237865814939e-05
step: 230, loss: 1.2285696357139386e-05
step: 240, loss: 0.0011168825440108776
step: 250, loss: 0.001086398377083242
step: 260, loss: 5.646641875500791e-05
step: 270, loss: 0.0001802964834496379
step: 280, loss: 1.761512794473674e-05
step: 290, loss: 7.424451723636594e-06
step: 300, loss: 1.594362038304098e-05
step: 310, loss: 6.434757233364508e-05
step: 320, loss: 1.1060265023843385e-05
step: 330, loss: 0.0011667516082525253
step: 340, loss: 0.0009294469491578639
step: 350, loss: 9.11943061510101e-06
step: 360, loss: 2.5201903554261662e-05
step: 370, loss: 0.00025340195861645043
step: 380, loss: 0.002366946544498205
step: 390, loss: 2.6023710233857855e-05
step: 400, loss: 0.0012685786932706833
step: 410, loss: 7.1264444159169216e-06
step: 420, loss: 9.436074833502062e-06
step: 430, loss: 0.00020295436843298376
step: 440, loss: 0.00010739565914263949
step: 450, loss: 8.467529369227123e-06
step: 460, loss: 0.020650362595915794
step: 470, loss: 1.0393480806669686e-05
step: 480, loss: 1.0777170246001333e-05
step: 490, loss: 1.0296621439920273e-05
step: 500, loss: 0.001447329530492425
step: 510, loss: 0.02193361334502697
step: 520, loss: 0.001434168079867959
step: 530, loss: 0.0012646445538848639
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
epoch 14: dev_f1=0.9614661654135339, f1=0.9507273580478649,
best_f1=0.9489939167056621
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 9.167866664938629e-06
step: 10, loss: 0.0028632236644625664
step: 20, loss: 0.0015357118099927902
step: 30, loss: 0.001696237944997847
step: 40, loss: 1.1149659258080646e-05
step: 50, loss: 0.0010102569358423352
```

step: 60, loss: 2.662189945112914e-05  
step: 70, loss: 1.5831950804567896e-05  
step: 80, loss: 8.36321669339668e-06  
step: 90, loss: 1.0773360372695606e-05  
step: 100, loss: 7.152507805585628e-06  
step: 110, loss: 1.4226437997422181e-05  
step: 120, loss: 7.882666068326216e-06  
step: 130, loss: 8.333418008987792e-06  
step: 140, loss: 1.0043234397016931e-05  
step: 150, loss: 9.264720574719831e-06  
step: 160, loss: 9.238610800821334e-06  
step: 170, loss: 6.761370514141163e-06  
step: 180, loss: 1.8923030438600108e-05  
step: 190, loss: 0.0012309768935665488  
step: 200, loss: 0.0012377689126878977  
step: 210, loss: 6.8135173023620155e-06  
step: 220, loss: 8.977878678706475e-06  
step: 230, loss: 2.323143598914612e-05  
step: 240, loss: 2.3358439648291096e-05  
step: 250, loss: 0.0010330439545214176  
step: 260, loss: 8.843766408972442e-06  
step: 270, loss: 2.017399310716428e-05  
step: 280, loss: 2.1931429728283547e-05  
step: 290, loss: 0.00031238768133334816  
step: 300, loss: 7.469160209438996e-06  
step: 310, loss: 0.021027978509664536  
step: 320, loss: 1.1581731087062508e-05  
step: 330, loss: 1.9992725356132723e-05  
step: 340, loss: 3.114981882390566e-05  
step: 350, loss: 0.0014831286389380693  
step: 360, loss: 8.883268310455605e-05  
step: 370, loss: 5.753185541834682e-05  
step: 380, loss: 1.0613233826006763e-05  
step: 390, loss: 1.9519347915775143e-05  
step: 400, loss: 0.0012663101078942418  
step: 410, loss: 1.0430653674120549e-05  
step: 420, loss: 9.126865734288003e-06  
step: 430, loss: 6.779977866244735e-06  
step: 440, loss: 1.2106765097996686e-05  
step: 450, loss: 0.0029091467149555683  
step: 460, loss: 0.014625897631049156  
step: 470, loss: 8.612813871877734e-06  
step: 480, loss: 0.0023900032974779606  
step: 490, loss: 1.0337533240090124e-05  
step: 500, loss: 0.001371823949739337  
step: 510, loss: 8.977821380540263e-06  
step: 520, loss: 1.7194981410284527e-05  
step: 530, loss: 6.619809028052259e-06

```

/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 15: dev_f1=0.9605263157894737, f1=0.9507273580478649,
best_f1=0.9489939167056621
mv: cannot stat '*_dev.pt': No such file or directory
Selected optimization level 02:  FP16 training with FP32 batchnorm and FP32
master weights.

Defaults for this optimization level are:
enabled                : True
opt_level              : 02
cast_model_type        : torch.float16
patch_torch_functions  : False
keep_batchnorm_fp32    : True
master_weights         : True
loss_scale             : dynamic
Processing user overrides (additional kwargs that are not None)...
After processing overrides, optimization options are:
enabled                : True
opt_level              : 02
cast_model_type        : torch.float16
patch_torch_functions  : False
keep_batchnorm_fp32    : True
master_weights         : True
loss_scale             : dynamic
Warning:  multi_tensor_applier fused unscale kernel is unavailable, possibly
because apex was installed without --cuda_ext --cpp_ext. Using Python fallback.
Original ImportError was: ModuleNotFoundError("No module named 'amp_C'")
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
5742it [00:34, 166.10it/s]
load_f1 = 0.9636194029850746
real_f1 = 0.9631013545072397
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
4398it [00:31, 139.86it/s]

```

## 4 DITTO

### 4.1 DITTO STRUCTURED

#### 4.1.1 Beer - Running the matcher

```
[ ]: # Step 1:
# train the wdc model with the save_model flag on
# it will produce two files *_dev.pt and *_test.pt which are the
# best checkpoints on the dev set and the test set.
!CUDA_VISIBLE_DEVICES=0 python train_ditto.py \
  --task Structured/Beer \
  --batch_size 32 \
  --max_len 256 \
  --lr 3e-5 \
  --n_epochs 15 \
  --finetuning \
  --lm roberta \
  --fp16 \
  --save_model \
  --dk general \
  --dk product \
  --summarize \
  --da entry_swap \
  --da attr_del

# Step 2: pick a checkpoint, rename it to wdc_all_small.pt
!mv *_dev.pt checkpoints/Beer.pt

# Step 3: run the matcher
!CUDA_VISIBLE_DEVICES=0 python matcher.py \
  --task Structured/Beer \
  --input_path input/input_small.jsonl \
  --output_path output/output_small.jsonl \
  --lm roberta \
  --use_gpu \
  --fp16 \
  --checkpoint_path checkpoints/
```

Selected optimization level 02: FP16 training with FP32 batchnorm and FP32 master weights.

Defaults for this optimization level are:

enabled	: True
opt_level	: 02
cast_model_type	: torch.float16



```

patch_torch_functions : False
keep_batchnorm_fp32   : True
master_weights        : True
loss_scale             : dynamic
Processing user overrides (additional kwargs that are not None)...
After processing overrides, optimization options are:
enabled                : True
opt_level              : 02
cast_model_type        : torch.float16
patch_torch_functions : False
keep_batchnorm_fp32    : True
master_weights         : True
loss_scale             : dynamic
Warning: multi_tensor_applier fused unscale kernel is unavailable, possibly
because apex was installed without --cuda_ext --cpp_ext. Using Python fallback.
Original ImportError was: ModuleNotFoundError("No module named 'amp_C'")
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 32768.0
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:136:
UserWarning: Detected call of `lr_scheduler.step()` before `optimizer.step()`.
In PyTorch 1.1.0 and later, you should call them in the opposite order:
`optimizer.step()` before `lr_scheduler.step()`. Failure to do this will result
in PyTorch skipping the first value of the learning rate schedule. See more
details at https://pytorch.org/docs/stable/optim.html#how-to-adjust-learning-
rate
    "https://pytorch.org/docs/stable/optim.html#how-to-adjust-learning-rate",
UserWarning)
step: 0, loss: 0.4366139769554138
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
    warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 1: dev_f1=0.2857142857142857, f1=0.28, best_f1=0.28
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.4049339294433594
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.

```

```

    warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 2: dev_f1=0.5555555555555556, f1=0.4878048780487805,
best_f1=0.4878048780487805
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.3818034827709198
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 3: dev_f1=0.456140350877193, f1=0.39285714285714285,
best_f1=0.4878048780487805
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.24190178513526917
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 4: dev_f1=0.3943661971830986, f1=0.45161290322580644,
best_f1=0.4878048780487805
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.295907199382782
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
    warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 5: dev_f1=0.8148148148148148, f1=0.8125000000000001,
best_f1=0.8125000000000001
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.2090083807706833
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
    warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 6: dev_f1=0.9032258064516129, f1=0.8, best_f1=0.8
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:

```

```

An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.22937597334384918
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
epoch 7: dev_f1=0.9032258064516129, f1=0.8750000000000001, best_f1=0.8
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.09742402285337448
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
epoch 8: dev_f1=0.896551724137931, f1=0.9333333333333333, best_f1=0.8
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.029193125665187836
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
  warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 9: dev_f1=0.923076923076923, f1=0.8571428571428571,
best_f1=0.8571428571428571
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.08958527445793152
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
epoch 10: dev_f1=0.923076923076923, f1=0.8148148148148148,
best_f1=0.8571428571428571
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.01131790317595005
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
epoch 11: dev_f1=0.9032258064516129, f1=0.8750000000000001,
best_f1=0.8571428571428571
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.

```

```

    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0027221518103033304
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 12: dev_f1=0.9032258064516129, f1=0.8750000000000001,
best_f1=0.8571428571428571
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0018444248707965016
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 13: dev_f1=0.9032258064516129, f1=0.8750000000000001,
best_f1=0.8571428571428571
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0055499570444226265
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 14: dev_f1=0.9032258064516129, f1=0.8750000000000001,
best_f1=0.8571428571428571
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0023315513972193003
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 15: dev_f1=0.9032258064516129, f1=0.8750000000000001,
best_f1=0.8571428571428571
mv: cannot stat '*_dev.pt': No such file or directory
Selected optimization level 02:  FP16 training with FP32 batchnorm and FP32
master weights.

```

Defaults for this optimization level are:

```

enabled           : True
opt_level         : 02
cast_model_type   : torch.float16
patch_torch_functions : False
keep_batchnorm_fp32 : True
master_weights    : True
loss_scale        : dynamic

```

Processing user overrides (additional kwargs that are not None)...

After processing overrides, optimization options are:

```

enabled                : True
opt_level              : 02
cast_model_type        : torch.float16
patch_torch_functions  : False
keep_batchnorm_fp32    : True
master_weights         : True
loss_scale             : dynamic
Warning: multi_tensor_applier fused unscale kernel is unavailable, possibly
because apex was installed without --cuda_ext --cpp_ext. Using Python fallback.
Original ImportError was: ModuleNotFoundError("No module named 'amp_C'")
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
91it [00:00, 125059.52it/s]
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
load_f1 = 0.9032258064516129
real_f1 = 0.9032258064516129
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
4398it [00:31, 141.09it/s]

```

#### 4.1.2 DBLP-ACM - Running the matcher

```

[ ]: # Step 1:
# train the wdc model with the save_model flag on
# it will produce two files *_dev.pt and *_test.pt which are the
# best checkpoints on the dev set and the test set.
!CUDA_VISIBLE_DEVICES=0 python train_ditto.py \
--task Structured/DBLP-ACM \
--batch_size 32 \
--max_len 256 \
--lr 3e-5 \
--n_epochs 15 \
--finetuning \
--lm roberta \
--fp16 \
--save_model \
--dk general \
--dk product \
--summarize \
--da entry_swap \
--da attr_del

```

```

# Step 2: pick a checkpoint, rename it to wdc_all_small.pt
!mv *_dev.pt checkpoints/DBLP-ACM.pt

# Step 3: run the matcher
!CUDA_VISIBLE_DEVICES=0 python matcher.py \
  --task Structured/DBLP-ACM \
  --input_path input/input_small.jsonl \
  --output_path output/output_small.jsonl \
  --lm roberta \
  --use_gpu \
  --fp16 \
  --checkpoint_path checkpoints/

```

Selected optimization level 02: FP16 training with FP32 batchnorm and FP32 master weights.

Defaults for this optimization level are:

```

enabled           : True
opt_level         : 02
cast_model_type   : torch.float16
patch_torch_functions : False
keep_batchnorm_fp32 : True
master_weights    : True
loss_scale        : dynamic

```

Processing user overrides (additional kwargs that are not None)...

After processing overrides, optimization options are:

```

enabled           : True
opt_level         : 02
cast_model_type   : torch.float16
patch_torch_functions : False
keep_batchnorm_fp32 : True
master_weights    : True
loss_scale        : dynamic

```

Warning: multi\_tensor\_applier fused unscale kernel is unavailable, possibly because apex was installed without --cuda\_ext --cpp\_ext. Using Python fallback. Original ImportError was: ModuleNotFoundError("No module named 'amp\_C'")  
/usr/local/lib/python3.7/dist-packages/apex/amp/\_initialize.py:25: UserWarning: An input tensor was not cuda.

```
warnings.warn("An input tensor was not cuda.")
```

```
step: 0, loss: 0.5744146704673767
```

Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 32768.0

```
step: 10, loss: 0.45344069600105286
```

Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0

```
step: 20, loss: 0.5010546445846558
```

```
step: 30, loss: 0.33652263879776
```

```
step: 40, loss: 0.20186328887939453
```

```

step: 50, loss: 0.12550565600395203
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 8192.0
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 4096.0
step: 60, loss: 0.3468756377696991
step: 70, loss: 0.06699441373348236
step: 80, loss: 0.05623256042599678
step: 90, loss: 0.13986480236053467
step: 100, loss: 0.09941402822732925
step: 110, loss: 0.02652747370302677
step: 120, loss: 0.008108435198664665
step: 130, loss: 0.008689397014677525
step: 140, loss: 0.02321498841047287
step: 150, loss: 0.13383720815181732
step: 160, loss: 0.004455122165381908
step: 170, loss: 0.13177810609340668
step: 180, loss: 0.11371047049760818
step: 190, loss: 0.0503462590277195
step: 200, loss: 0.010345760732889175
step: 210, loss: 0.026607464998960495
step: 220, loss: 0.06268952041864395
step: 230, loss: 0.021926339715719223
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
    warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 1: dev_f1=0.9745293466223698, f1=0.9733333333333333,
best_f1=0.9733333333333333
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.010169959627091885
step: 10, loss: 0.04196208342909813
step: 20, loss: 0.010709019377827644
step: 30, loss: 0.016405820846557617
step: 40, loss: 0.010388300754129887
step: 50, loss: 0.002956227632239461
step: 60, loss: 0.007123890332877636
step: 70, loss: 0.002709867898374796
step: 80, loss: 0.0018136021681129932
step: 90, loss: 0.003762186737731099
step: 100, loss: 0.00214783544652164
step: 110, loss: 0.06087825074791908
step: 120, loss: 0.028618574142456055
step: 130, loss: 0.0038644291926175356
step: 140, loss: 0.001540781813673675

```

```

step: 150, loss: 0.19969381392002106
step: 160, loss: 0.002071201801300049
step: 170, loss: 0.002409108681604266
step: 180, loss: 0.016680197790265083
step: 190, loss: 0.004880666732788086
step: 200, loss: 0.004762771539390087
step: 210, loss: 0.04719274491071701
step: 220, loss: 0.001344326650723815
step: 230, loss: 0.0014341090572997928
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
  warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 2: dev_f1=0.9855072463768116, f1=0.9854423292273236,
best_f1=0.9854423292273236
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.004364116117358208
step: 10, loss: 0.015947479754686356
step: 20, loss: 0.0008888826705515385
step: 30, loss: 0.013698983937501907
step: 40, loss: 0.026612071320414543
step: 50, loss: 0.055670350790023804
step: 60, loss: 0.002753747161477804
step: 70, loss: 0.0018539937445893884
step: 80, loss: 0.0074896439909935
step: 90, loss: 0.001567987841553986
step: 100, loss: 0.00637240381911397
step: 110, loss: 0.025838566944003105
step: 120, loss: 0.0006334111094474792
step: 130, loss: 0.008005564101040363
step: 140, loss: 0.0013038242468610406
step: 150, loss: 0.15725918114185333
step: 160, loss: 0.002385964384302497
step: 170, loss: 0.002057806821539998
step: 180, loss: 0.0018817492527887225
step: 190, loss: 0.029755916446447372
step: 200, loss: 0.004909756127744913
step: 210, loss: 0.0011140548158437014
step: 220, loss: 0.008977952413260937
step: 230, loss: 0.005919215735048056
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")

```



```

/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
  warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 3: dev_f1=0.9865470852017937, f1=0.9830890642615557,
best_f1=0.9830890642615557
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.004133058246225119
step: 10, loss: 0.01628616638481617
step: 20, loss: 0.01819799281656742
step: 30, loss: 0.0011622612364590168
step: 40, loss: 0.08947204798460007
step: 50, loss: 0.005119292996823788
step: 60, loss: 0.0012767845764756203
step: 70, loss: 0.0018665750976651907
step: 80, loss: 0.001274315407499671
step: 90, loss: 0.005497620441019535
step: 100, loss: 0.0012525058118626475
step: 110, loss: 0.0007195925572887063
step: 120, loss: 0.018077729269862175
step: 130, loss: 0.002791651990264654
step: 140, loss: 0.0009765970753505826
step: 150, loss: 0.0006225879187695682
step: 160, loss: 0.0034576270263642073
step: 170, loss: 0.0476631335914135
step: 180, loss: 0.13902722299098969
step: 190, loss: 0.0029790597036480904
step: 200, loss: 0.002717949915677309
step: 210, loss: 0.002544160932302475
step: 220, loss: 0.000576447113417089
step: 230, loss: 0.0131691200658679
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
epoch 4: dev_f1=0.9865470852017937, f1=0.9887133182844244,
best_f1=0.9830890642615557
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0014292002888396382
step: 10, loss: 0.0017338774632662535
step: 20, loss: 0.09849873930215836
step: 30, loss: 0.0038591965567320585
step: 40, loss: 0.007476997096091509
step: 50, loss: 0.0019672242924571037
step: 60, loss: 0.004237894434481859

```

```

step: 70, loss: 0.0021731099113821983
step: 80, loss: 0.11346733570098877
step: 90, loss: 0.03662688657641411
step: 100, loss: 0.00030278947087936103
step: 110, loss: 0.003158220322802663
step: 120, loss: 0.0021829637698829174
step: 130, loss: 0.0005075312801636755
step: 140, loss: 0.0022674936335533857
step: 150, loss: 0.014204608276486397
step: 160, loss: 0.00037907695514149964
step: 170, loss: 0.01942775957286358
step: 180, loss: 0.02514638751745224
step: 190, loss: 0.15656965970993042
step: 200, loss: 0.004441302269697189
step: 210, loss: 0.002632526447996497
step: 220, loss: 0.00596680399030447
step: 230, loss: 0.03595719113945961
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
  warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 5: dev_f1=0.9887892376681614, f1=0.9830890642615557,
best_f1=0.9830890642615557
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00041014738962985575
step: 10, loss: 0.001749189104884863
step: 20, loss: 0.0015866131288930774
step: 30, loss: 0.002964131301268935
step: 40, loss: 0.0008876680512912571
step: 50, loss: 0.0047487616539001465
step: 60, loss: 0.0060667609795928
step: 70, loss: 0.0004735758120659739
step: 80, loss: 0.003467935835942626
step: 90, loss: 0.02524997852742672
step: 100, loss: 0.0028922120109200478
step: 110, loss: 0.03060409240424633
step: 120, loss: 0.0003576207673177123
step: 130, loss: 0.00028046168154105544
step: 140, loss: 0.0004070204449817538
step: 150, loss: 0.0004663430154323578
step: 160, loss: 0.0018150575924664736
step: 170, loss: 0.001960896188393235
step: 180, loss: 0.0005627111531794071

```

```

step: 190, loss: 0.0005601290613412857
step: 200, loss: 0.03214559331536293
step: 210, loss: 0.0036973287351429462
step: 220, loss: 0.01587549038231373
step: 230, loss: 0.0011393723543733358
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
  warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 6: dev_f1=0.9899441340782122, f1=0.9821029082774049,
best_f1=0.9821029082774049
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0032192848157137632
step: 10, loss: 0.0006561951595358551
step: 20, loss: 0.002004386857151985
step: 30, loss: 0.00034635484917089343
step: 40, loss: 0.004080228507518768
step: 50, loss: 0.000671673275064677
step: 60, loss: 0.0002776330220513046
step: 70, loss: 0.00024192288401536644
step: 80, loss: 0.0008613793761469424
step: 90, loss: 0.003128492273390293
step: 100, loss: 0.00023245066404342651
step: 110, loss: 0.000190523816854693
step: 120, loss: 0.0003104774805251509
step: 130, loss: 0.0026009909342974424
step: 140, loss: 0.0014650702942162752
step: 150, loss: 0.012554198503494263
step: 160, loss: 0.0001542845566291362
step: 170, loss: 0.00032657344127073884
step: 180, loss: 0.00021773669868707657
step: 190, loss: 0.002059723250567913
step: 200, loss: 0.005524806212633848
step: 210, loss: 0.0004083302046637982
step: 220, loss: 0.0012055016122758389
step: 230, loss: 0.0029866062104701996
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
epoch 7: dev_f1=0.9888392857142857, f1=0.9865771812080537,
best_f1=0.9821029082774049
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.

```

```

    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.006495540961623192
step: 10, loss: 0.009920891374349594
step: 20, loss: 0.008920291438698769
step: 30, loss: 0.0018882843432947993
step: 40, loss: 0.0010942703811451793
step: 50, loss: 0.0016670965123921633
step: 60, loss: 0.0008636569255031645
step: 70, loss: 0.0002582334273029119
step: 80, loss: 0.012307208962738514
step: 90, loss: 0.0008008776349015534
step: 100, loss: 0.01615738868713379
step: 110, loss: 0.0004915935569442809
step: 120, loss: 0.0012689844006672502
step: 130, loss: 0.00903334654867649
step: 140, loss: 0.002234546234831214
step: 150, loss: 0.16093386709690094
step: 160, loss: 0.0014113779179751873
step: 170, loss: 0.006393431685864925
step: 180, loss: 0.001226345426402986
step: 190, loss: 0.0031592112500220537
step: 200, loss: 0.004199774470180273
step: 210, loss: 0.01372506469488144
step: 220, loss: 0.0009937530849128962
step: 230, loss: 0.001030503772199154
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
    warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 8: dev_f1=0.9910514541387023, f1=0.9843400447427293,
best_f1=0.9843400447427293
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0010091242147609591
step: 10, loss: 0.0012064785696566105
step: 20, loss: 0.0025514501612633467
step: 30, loss: 0.0006496343994513154
step: 40, loss: 0.0009004768799059093
step: 50, loss: 0.0015115303685888648
step: 60, loss: 0.014787418767809868
step: 70, loss: 0.04777490347623825
step: 80, loss: 0.0004787309153471142
step: 90, loss: 0.03778507560491562
step: 100, loss: 0.0005655853310599923

```

```

step: 110, loss: 0.0005179025465622544
step: 120, loss: 0.011246219277381897
step: 130, loss: 0.0007217422244139016
step: 140, loss: 0.00052103609777987
step: 150, loss: 0.002433663234114647
step: 160, loss: 0.001008257851935923
step: 170, loss: 0.00238283583894372
step: 180, loss: 0.0020048280712217093
step: 190, loss: 0.00013523588131647557
step: 200, loss: 0.0009185908711515367
step: 210, loss: 0.02818208374083042
step: 220, loss: 0.0004314411780796945
step: 230, loss: 0.00014823928358964622
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
epoch 9: dev_f1=0.9910112359550561, f1=0.9854096520763187,
best_f1=0.9843400447427293
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0002441331307636574
step: 10, loss: 0.00032067846041172743
step: 20, loss: 0.00033756185439415276
step: 30, loss: 0.0008925668662413955
step: 40, loss: 0.003360782051458955
step: 50, loss: 0.00020073962514288723
step: 60, loss: 0.0004818038723897189
step: 70, loss: 0.014312963001430035
step: 80, loss: 0.00012415314267855138
step: 90, loss: 0.00012871828221250325
step: 100, loss: 0.00021764582197647542
step: 110, loss: 0.06045817211270332
step: 120, loss: 0.00017856337944976985
step: 130, loss: 0.00021300934895407408
step: 140, loss: 0.002221491653472185
step: 150, loss: 0.001369715784676373
step: 160, loss: 8.11061545391567e-05
step: 170, loss: 0.00019556982442736626
step: 180, loss: 0.025607304647564888
step: 190, loss: 0.0004647917812690139
step: 200, loss: 0.00040412263479083776
step: 210, loss: 0.007775670383125544
step: 220, loss: 0.010953960940241814
step: 230, loss: 0.00026241890736855567
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")

```

```

/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
  warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 10: dev_f1=0.9921787709497207, f1=0.9800443458980044,
best_f1=0.9800443458980044
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00017034111078828573
step: 10, loss: 0.00017514005594421178
step: 20, loss: 0.0002701992343645543
step: 30, loss: 0.001620437833480537
step: 40, loss: 0.00011214771075174212
step: 50, loss: 0.0015391813358291984
step: 60, loss: 0.01290323305875063
step: 70, loss: 0.0002790658618323505
step: 80, loss: 0.00010408302478026599
step: 90, loss: 0.1631278544664383
step: 100, loss: 0.015981921926140785
step: 110, loss: 0.0011329316766932607
step: 120, loss: 0.0009630870772525668
step: 130, loss: 0.0002812727470882237
step: 140, loss: 0.0050029996782541275
step: 150, loss: 0.00022961432114243507
step: 160, loss: 0.0065223220735788345
step: 170, loss: 0.0003461991436779499
step: 180, loss: 0.0008652989636175334
step: 190, loss: 0.00029290103702805936
step: 200, loss: 0.0013690594350919127
step: 210, loss: 0.00024091410159599036
step: 220, loss: 0.0003102388873230666
step: 230, loss: 0.00010019965702667832
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
epoch 11: dev_f1=0.990990990990991, f1=0.9841986455981941,
best_f1=0.9800443458980044
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00023320021864492446
step: 10, loss: 0.0001155879654106684
step: 20, loss: 0.02218201570212841
step: 30, loss: 0.025897055864334106
step: 40, loss: 0.00037983691436238587
step: 50, loss: 0.0016074592713266611
step: 60, loss: 0.0007778447470627725

```

```

step: 70, loss: 0.0002664781641215086
step: 80, loss: 8.8767810666468e-05
step: 90, loss: 0.0003929051454178989
step: 100, loss: 0.00013132703315932304
step: 110, loss: 9.695523476693779e-05
step: 120, loss: 0.0006951487739570439
step: 130, loss: 0.0001905412063933909
step: 140, loss: 0.0022138473577797413
step: 150, loss: 0.0001632424391573295
step: 160, loss: 0.0002021428372245282
step: 170, loss: 0.0007091880543157458
step: 180, loss: 0.0007515514153055847
step: 190, loss: 0.015629403293132782
step: 200, loss: 0.00011514583457028493
step: 210, loss: 0.0005278898170217872
step: 220, loss: 0.007755230646580458
step: 230, loss: 0.0001024099110509269
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
epoch 12: dev_f1=0.9921259842519685, f1=0.9887387387387387,
best_f1=0.9800443458980044
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0006102712359279394
step: 10, loss: 9.354441135656089e-05
step: 20, loss: 0.00011359309428371489
step: 30, loss: 3.770849434658885e-05
step: 40, loss: 0.00014683077461086214
step: 50, loss: 0.0006039651343598962
step: 60, loss: 8.609188807895407e-05
step: 70, loss: 0.0005135739338584244
step: 80, loss: 0.00017283861234318465
step: 90, loss: 5.7657904108054936e-05
step: 100, loss: 0.00013655492512043566
step: 110, loss: 0.0007977366331033409
step: 120, loss: 0.000207144083105959
step: 130, loss: 0.0018843685975298285
step: 140, loss: 0.001231466419994831
step: 150, loss: 0.0002652349940035492
step: 160, loss: 0.018740035593509674
step: 170, loss: 0.0001339399750577286
step: 180, loss: 0.021534327417612076
step: 190, loss: 0.0001318462163908407
step: 200, loss: 3.7685895222239196e-05
step: 210, loss: 0.0002939357073046267
step: 220, loss: 9.13968906388618e-05

```

```

step: 230, loss: 9.369070176035166e-05
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 13: dev_f1=0.990990990990991, f1=0.9864864864864865,
best_f1=0.9800443458980044
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 5.240242899162695e-05
step: 10, loss: 6.87660722178407e-05
step: 20, loss: 0.001007493701763451
step: 30, loss: 0.00010787913925014436
step: 40, loss: 0.00016186620632652193
step: 50, loss: 8.375686593353748e-05
step: 60, loss: 7.14780471753329e-05
step: 70, loss: 0.00010591636237222701
step: 80, loss: 0.00017676925926934928
step: 90, loss: 0.00014188906061463058
step: 100, loss: 0.00023110340407583863
step: 110, loss: 0.00010031989950221032
step: 120, loss: 2.3423748643836007e-05
step: 130, loss: 0.00017570726049598306
step: 140, loss: 0.0042104278691112995
step: 150, loss: 8.130034984787926e-05
step: 160, loss: 6.817550456617028e-05
step: 170, loss: 0.00010364956688135862
step: 180, loss: 8.386967238038778e-05
step: 190, loss: 0.0012535062851384282
step: 200, loss: 0.00020099988614674658
step: 210, loss: 6.378305261023343e-05
step: 220, loss: 0.0024490277282893658
step: 230, loss: 6.730953464284539e-05
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 14: dev_f1=0.9910714285714286, f1=0.9822222222222222,
best_f1=0.9800443458980044
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.03789593279361725
step: 10, loss: 5.664215132128447e-05
step: 20, loss: 0.0003083129122387618
step: 30, loss: 7.369677041424438e-05
step: 40, loss: 3.3541047741891816e-05
step: 50, loss: 5.005256025469862e-05
step: 60, loss: 0.031080465763807297

```



```

step: 70, loss: 0.00013972526357974857
step: 80, loss: 0.002100840676575899
step: 90, loss: 5.976944885333069e-05
step: 100, loss: 3.624114833655767e-05
step: 110, loss: 0.0021860748529434204
step: 120, loss: 0.04147418588399887
step: 130, loss: 7.299063145183027e-05
step: 140, loss: 0.015806546434760094
step: 150, loss: 0.013945146463811398
step: 160, loss: 0.014254129491746426
step: 170, loss: 7.501556683564559e-05
step: 180, loss: 9.562892228132114e-05
step: 190, loss: 5.184079418540932e-05
step: 200, loss: 0.0020273528061807156
step: 210, loss: 0.03098090924322605
step: 220, loss: 5.475361103890464e-05
step: 230, loss: 7.293871749425307e-05
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
epoch 15: dev_f1=0.990990990990991, f1=0.9864864864864865,
best_f1=0.9800443458980044
mv: cannot stat '*_dev.pt': No such file or directory
Selected optimization level 02: FP16 training with FP32 batchnorm and FP32
master weights.

Defaults for this optimization level are:
enabled                : True
opt_level              : 02
cast_model_type        : torch.float16
patch_torch_functions  : False
keep_batchnorm_fp32    : True
master_weights         : True
loss_scale             : dynamic
Processing user overrides (additional kwargs that are not None)...
After processing overrides, optimization options are:
enabled                : True
opt_level              : 02
cast_model_type        : torch.float16
patch_torch_functions  : False
keep_batchnorm_fp32    : True
master_weights         : True
loss_scale             : dynamic
Warning: multi_tensor_applier fused unscale kernel is unavailable, possibly
because apex was installed without --cuda_ext --cpp_ext. Using Python fallback.
Original ImportError was: ModuleNotFoundError("No module named 'amp_C'")
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.

```

```

    warnings.warn("An input tensor was not cuda.")
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
2473it [00:11, 213.42it/s]
load_f1 = 0.9910714285714286
real_f1 = 0.9888392857142857
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
4398it [00:22, 198.96it/s]

```

### 4.1.3 DBLP-GoogleScholar - Running the matcher

```

[ ]: # Step 1:
# train the wdc model with the save_model flag on
# it will produce two files *_dev.pt and *_test.pt which are the
# best checkpoints on the dev set and the test set.
!CUDA_VISIBLE_DEVICES=0 python train_ditto.py \
--task Structured/DBLP-GoogleScholar \
--batch_size 32 \
--max_len 256 \
--lr 3e-5 \
--n_epochs 15 \
--finetuning \
--lm roberta \
--fp16 \
--save_model \
--dk general \
--dk product \
--summarize \
--da entry_swap \
--da attr_del

# Step 2: pick a checkpoint, rename it to wdc_all_small.pt
!mv *_dev.pt checkpoints/DBLP-GoogleScholar.pt

# Step 3: run the matcher
!CUDA_VISIBLE_DEVICES=0 python matcher.py \
--task Structured/DBLP-GoogleScholar \
--input_path input/input_small.jsonl \
--output_path output/output_small.jsonl \
--lm roberta \
--use_gpu \
--fp16 \

```

```
--checkpoint_path checkpoints/
```

```
Downloading: 100% 481/481 [00:00<00:00, 628kB/s]
Downloading: 100% 899k/899k [00:00<00:00, 30.1MB/s]
Downloading: 100% 456k/456k [00:00<00:00, 28.3MB/s]
Downloading: 100% 501M/501M [00:08<00:00, 58.8MB/s]
Selected optimization level 02: FP16 training with FP32 batchnorm and FP32
master weights.
```

```
Defaults for this optimization level are:
```

```
enabled                : True
opt_level              : 02
cast_model_type        : torch.float16
patch_torch_functions  : False
keep_batchnorm_fp32    : True
master_weights         : True
loss_scale             : dynamic
```

```
Processing user overrides (additional kwargs that are not None)...
```

```
After processing overrides, optimization options are:
```

```
enabled                : True
opt_level              : 02
cast_model_type        : torch.float16
patch_torch_functions  : False
keep_batchnorm_fp32    : True
master_weights         : True
loss_scale             : dynamic
```

```
Warning: multi_tensor_applier fused unscale kernel is unavailable, possibly
because apex was installed without --cuda_ext --cpp_ext. Using Python fallback.
Original ImportError was: ModuleNotFoundError("No module named 'amp_C'")
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
```

```
    warnings.warn("An input tensor was not cuda.")
```

```
step: 0, loss: 0.6678732633590698
```

```
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 32768.0
```

```
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
```

```
step: 10, loss: 0.42742785811424255
```

```
step: 20, loss: 0.2619576156139374
```

```
step: 30, loss: 0.28199121356010437
```

```
step: 40, loss: 0.1687605381011963
```

```
step: 50, loss: 0.14239443838596344
```

```
step: 60, loss: 0.3023311197757721
```

```
step: 70, loss: 0.21753738820552826
```

```
step: 80, loss: 0.10192237049341202
```

```
step: 90, loss: 0.10497428476810455
```

```
step: 100, loss: 0.12910997867584229
```

```
step: 110, loss: 0.04584508761763573
```

```
step: 120, loss: 0.08951669931411743
```

```

step: 130, loss: 0.09875155985355377
step: 140, loss: 0.2798832952976227
step: 150, loss: 0.08785443753004074
step: 160, loss: 0.10194317251443863
step: 170, loss: 0.024405211210250854
step: 180, loss: 0.04188816249370575
step: 190, loss: 0.02858145348727703
step: 200, loss: 0.05662674456834793
step: 210, loss: 0.03174491971731186
step: 220, loss: 0.13964787125587463
step: 230, loss: 0.21954241394996643
step: 240, loss: 0.06634300947189331
step: 250, loss: 0.017770037055015564
step: 260, loss: 0.09451739490032196
step: 270, loss: 0.2945121228694916
step: 280, loss: 0.0240713432431221
step: 290, loss: 0.07357558608055115
step: 300, loss: 0.029802091419696808
step: 310, loss: 0.058584485203027725
step: 320, loss: 0.0935981348156929
step: 330, loss: 0.062135662883520126
step: 340, loss: 0.296571284532547
step: 350, loss: 0.05869349092245102
step: 360, loss: 0.08313492685556412
step: 370, loss: 0.02127496898174286
step: 380, loss: 0.18342113494873047
step: 390, loss: 0.08143174648284912
step: 400, loss: 0.07553335279226303
step: 410, loss: 0.2792451083660126
step: 420, loss: 0.028793200850486755
step: 430, loss: 0.05027221888303757
step: 440, loss: 0.01542101614177227
step: 450, loss: 0.04946025460958481
step: 460, loss: 0.004502625670284033
step: 470, loss: 0.019704831764101982
step: 480, loss: 0.16687004268169403
step: 490, loss: 0.1003468707203865
step: 500, loss: 0.04208114370703697
step: 510, loss: 0.02763405442237854
step: 520, loss: 0.06081859767436981
step: 530, loss: 0.008116060867905617
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
  warnings.warn(SAVE_STATE_WARNING, UserWarning)

```

```
epoch 1: dev_f1=0.9527739569005044, f1=0.9504132231404958,  
best_f1=0.9504132231404958  
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:  
An input tensor was not cuda.  
    warnings.warn("An input tensor was not cuda.")  
step: 0, loss: 0.0785687267780304  
step: 10, loss: 0.02448752522468567  
step: 20, loss: 0.006995551288127899  
step: 30, loss: 0.09748803824186325  
step: 40, loss: 0.07696256041526794  
step: 50, loss: 0.03930563107132912  
step: 60, loss: 0.025539124384522438  
step: 70, loss: 0.02368297055363655  
step: 80, loss: 0.020144056528806686  
step: 90, loss: 0.009254774078726768  
step: 100, loss: 0.14501291513442993  
step: 110, loss: 0.010015612468123436  
step: 120, loss: 0.07396154850721359  
step: 130, loss: 0.0067220586352050304  
step: 140, loss: 0.11900641769170761  
step: 150, loss: 0.02083219401538372  
step: 160, loss: 0.027405772358179092  
step: 170, loss: 0.03973178192973137  
step: 180, loss: 0.04061845690011978  
step: 190, loss: 0.011939842253923416  
step: 200, loss: 0.1545330137014389  
step: 210, loss: 0.02998054027557373  
step: 220, loss: 0.018813839182257652  
step: 230, loss: 0.05418140068650246  
step: 240, loss: 0.050973713397979736  
step: 250, loss: 0.00935614574700594  
step: 260, loss: 0.1967242807149887  
step: 270, loss: 0.020475884899497032  
step: 280, loss: 0.058569297194480896  
step: 290, loss: 0.1042582169175148  
step: 300, loss: 0.12476349622011185  
step: 310, loss: 0.03525453805923462  
step: 320, loss: 0.12248539179563522  
step: 330, loss: 0.0985473096370697  
step: 340, loss: 0.06933625042438507  
step: 350, loss: 0.004244054201990366  
step: 360, loss: 0.038160115480422974  
step: 370, loss: 0.004298957996070385  
step: 380, loss: 0.12064820528030396  
step: 390, loss: 0.007578299380838871  
step: 400, loss: 0.1343066245317459  
step: 410, loss: 0.014488071203231812  
step: 420, loss: 0.08015023171901703
```

```

step: 430, loss: 0.1330665498971939
step: 440, loss: 0.007295288145542145
step: 450, loss: 0.03648862987756729
step: 460, loss: 0.06309904158115387
step: 470, loss: 0.056001823395490646
step: 480, loss: 0.002289440715685487
step: 490, loss: 0.04976673051714897
step: 500, loss: 0.005960507318377495
step: 510, loss: 0.008672724477946758
step: 520, loss: 0.1830616444349289
step: 530, loss: 0.11358499526977539
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
  warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 2: dev_f1=0.953759925268566, f1=0.9515294117647058,
best_f1=0.9515294117647058
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.1380801796913147
step: 10, loss: 0.04722166433930397
step: 20, loss: 0.009571759030222893
step: 30, loss: 0.04545724391937256
step: 40, loss: 0.051707006990909576
step: 50, loss: 0.040276285260915756
step: 60, loss: 0.005818306468427181
step: 70, loss: 0.004907283000648022
step: 80, loss: 0.08002738654613495
step: 90, loss: 0.03937723860144615
step: 100, loss: 0.017725730314850807
step: 110, loss: 0.009434565901756287
step: 120, loss: 0.08438288420438766
step: 130, loss: 0.06768260151147842
step: 140, loss: 0.026458486914634705
step: 150, loss: 0.045598551630973816
step: 160, loss: 0.016571596264839172
step: 170, loss: 0.03867824748158455
step: 180, loss: 0.009586643427610397
step: 190, loss: 0.0011078673414885998
step: 200, loss: 0.01511270273476839
step: 210, loss: 0.036347854882478714
step: 220, loss: 0.048083316534757614
step: 230, loss: 0.049133773893117905
step: 240, loss: 0.13184945285320282

```

```

step: 250, loss: 0.08956029266119003
step: 260, loss: 0.16896814107894897
step: 270, loss: 0.03598540276288986
step: 280, loss: 0.009540791623294353
step: 290, loss: 0.0019954522140324116
step: 300, loss: 0.12155781686306
step: 310, loss: 0.028266090899705887
step: 320, loss: 0.029546916484832764
step: 330, loss: 0.005068727303296328
step: 340, loss: 0.013441838324069977
step: 350, loss: 0.17156602442264557
step: 360, loss: 0.021912701427936554
step: 370, loss: 0.0192897729575634
step: 380, loss: 0.0030528921633958817
step: 390, loss: 0.0035038189962506294
step: 400, loss: 0.0817529708147049
step: 410, loss: 0.01597675494849682
step: 420, loss: 0.010714108124375343
step: 430, loss: 0.00959490891546011
step: 440, loss: 0.1962091028690338
step: 450, loss: 0.04946085438132286
step: 460, loss: 0.04972191900014877
step: 470, loss: 0.02125510387122631
step: 480, loss: 0.006516304332762957
step: 490, loss: 0.02010442316532135
step: 500, loss: 0.09068955481052399
step: 510, loss: 0.05087810754776001
step: 520, loss: 0.06362970918416977
step: 530, loss: 0.0019052437273785472
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
    warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 3: dev_f1=0.9569691300280637, f1=0.9508656995788489,
best_f1=0.9508656995788489
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.04293051362037659
step: 10, loss: 0.005273095797747374
step: 20, loss: 0.052186813205480576
step: 30, loss: 0.07252494245767593
step: 40, loss: 0.01503250002861023
step: 50, loss: 0.05589143931865692
step: 60, loss: 0.011473347432911396

```

step: 70, loss: 0.03176479414105415  
step: 80, loss: 0.006038025952875614  
step: 90, loss: 0.08371325582265854  
step: 100, loss: 0.0005527421599254012  
step: 110, loss: 0.004822927061468363  
step: 120, loss: 0.010495101101696491  
step: 130, loss: 0.11502194404602051  
step: 140, loss: 0.03561561554670334  
step: 150, loss: 0.010026547126471996  
step: 160, loss: 0.0016137590864673257  
step: 170, loss: 0.01713201217353344  
step: 180, loss: 0.13389526307582855  
step: 190, loss: 0.051803167909383774  
step: 200, loss: 0.05018250644207001  
step: 210, loss: 0.0006698605720885098  
step: 220, loss: 0.003225987544283271  
step: 230, loss: 0.005878902971744537  
step: 240, loss: 0.0679292231798172  
step: 250, loss: 0.08322737365961075  
step: 260, loss: 0.008204496465623379  
step: 270, loss: 0.04261840507388115  
step: 280, loss: 0.005288249813020229  
step: 290, loss: 0.032494790852069855  
step: 300, loss: 0.0006985210929997265  
step: 310, loss: 0.0015044640749692917  
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 8192.0  
step: 320, loss: 0.05613570660352707  
step: 330, loss: 0.046178560703992844  
step: 340, loss: 0.009633232839405537  
step: 350, loss: 0.12553945183753967  
step: 360, loss: 0.07010402530431747  
step: 370, loss: 0.00146210053935647  
step: 380, loss: 0.001269968575797975  
step: 390, loss: 0.00010484416998224333  
step: 400, loss: 0.0009490179363638163  
step: 410, loss: 0.0020995240192860365  
step: 420, loss: 0.021115625277161598  
step: 430, loss: 0.01685856655240059  
step: 440, loss: 0.010730337351560593  
step: 450, loss: 0.0071905627846717834  
step: 460, loss: 0.01279095932841301  
step: 470, loss: 0.002074559684842825  
step: 480, loss: 0.008632651530206203  
step: 490, loss: 0.003502225037664175  
step: 500, loss: 0.087497279047966  
step: 510, loss: 0.06723099201917648  
step: 520, loss: 0.00384645932354033  
step: 530, loss: 0.10570412874221802



```

/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
    warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 4: dev_f1=0.9622377622377621, f1=0.9572490706319703,
best_f1=0.9572490706319703
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0034687549341470003
step: 10, loss: 0.01580396294593811
step: 20, loss: 0.002253409707918763
step: 30, loss: 0.010011523962020874
step: 40, loss: 0.015421179123222828
step: 50, loss: 0.05951536446809769
step: 60, loss: 0.006489202380180359
step: 70, loss: 0.0029927967116236687
step: 80, loss: 0.0004095137701369822
step: 90, loss: 0.010480280965566635
step: 100, loss: 0.2050837278366089
step: 110, loss: 0.006111598573625088
step: 120, loss: 0.1154947578907013
step: 130, loss: 0.0060766227543354034
step: 140, loss: 0.05467518791556358
step: 150, loss: 0.007861306890845299
step: 160, loss: 0.0023896426428109407
step: 170, loss: 0.16216954588890076
step: 180, loss: 0.004167209379374981
step: 190, loss: 0.006173516623675823
step: 200, loss: 0.0018641396891325712
step: 210, loss: 0.17351005971431732
step: 220, loss: 0.06282095611095428
step: 230, loss: 0.014653552323579788
step: 240, loss: 0.17053382098674774
step: 250, loss: 0.1657439023256302
step: 260, loss: 0.00958640780299902
step: 270, loss: 0.021070022135972977
step: 280, loss: 0.005935984198004007
step: 290, loss: 0.005357255227863789
step: 300, loss: 0.017358746379613876
step: 310, loss: 0.03401844576001167
step: 320, loss: 0.009851985611021519
step: 330, loss: 0.0016988288844004273
step: 340, loss: 0.019293511286377907
step: 350, loss: 0.002482523676007986

```

```
step: 360, loss: 0.001327698351815343
step: 370, loss: 0.000392396526876837
step: 380, loss: 0.0011468228185549378
step: 390, loss: 0.002817795379087329
step: 400, loss: 0.024802634492516518
step: 410, loss: 0.02835521660745144
step: 420, loss: 0.1813429594039917
step: 430, loss: 0.06614868342876434
step: 440, loss: 0.0003667143755592406
step: 450, loss: 0.01616361364722252
step: 460, loss: 0.09578760713338852
step: 470, loss: 0.03394784405827522
step: 480, loss: 0.19386212527751923
step: 490, loss: 0.020561352372169495
step: 500, loss: 0.003644846845418215
step: 510, loss: 0.0006193738081492484
step: 520, loss: 0.09547613561153412
step: 530, loss: 0.02116534113883972
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 5: dev_f1=0.9602989257356376, f1=0.9514018691588785,
best_f1=0.9572490706319703
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.15023919939994812
step: 10, loss: 0.005662649869918823
step: 20, loss: 0.08435112982988358
step: 30, loss: 0.0026066526770591736
step: 40, loss: 0.01148255169391632
step: 50, loss: 0.013224625028669834
step: 60, loss: 0.016216369345784187
step: 70, loss: 0.004463536664843559
step: 80, loss: 0.0013325319159775972
step: 90, loss: 0.0015536196297034621
step: 100, loss: 0.0005887417937628925
step: 110, loss: 0.008851874619722366
step: 120, loss: 0.0014319823822006583
step: 130, loss: 0.0004447655810508877
step: 140, loss: 0.003343722550198436
step: 150, loss: 0.0007038709009066224
step: 160, loss: 0.057842742651700974
step: 170, loss: 0.0019718213006854057
step: 180, loss: 0.0008938505197875202
step: 190, loss: 0.006062686908990145
step: 200, loss: 0.031068019568920135
step: 210, loss: 0.01917371153831482
```

```
step: 220, loss: 0.00494739506393671
step: 230, loss: 0.0020194717217236757
step: 240, loss: 0.01688501611351967
step: 250, loss: 0.022510191425681114
step: 260, loss: 0.0030135998968034983
step: 270, loss: 0.003686292329803109
step: 280, loss: 0.0043168929405510426
step: 290, loss: 0.0015466574113816023
step: 300, loss: 0.0518699511885643
step: 310, loss: 0.012358630076050758
step: 320, loss: 0.00023771209816914052
step: 330, loss: 0.001295381924137473
step: 340, loss: 0.0005725126247853041
step: 350, loss: 0.004677498713135719
step: 360, loss: 0.055760812014341354
step: 370, loss: 0.00741333095356822
step: 380, loss: 0.003242489183321595
step: 390, loss: 0.005166185088455677
step: 400, loss: 0.04101172462105751
step: 410, loss: 0.00011404585529817268
step: 420, loss: 0.01152878813445568
step: 430, loss: 0.003442309098318219
step: 440, loss: 0.0013797628926113248
step: 450, loss: 0.17109444737434387
step: 460, loss: 0.0026748860254883766
step: 470, loss: 0.003942226991057396
step: 480, loss: 0.0019560735672712326
step: 490, loss: 0.004722924437373877
step: 500, loss: 0.00216879416257143
step: 510, loss: 0.07627198845148087
step: 520, loss: 0.006895953323692083
step: 530, loss: 0.006912856362760067
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
epoch 6: dev_f1=0.9579831932773109, f1=0.9498829039812646,
best_f1=0.9572490706319703
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0057591889053583145
step: 10, loss: 0.0005071586929261684
step: 20, loss: 0.11396807432174683
step: 30, loss: 0.013929430395364761
step: 40, loss: 0.037101566791534424
step: 50, loss: 0.003719612956047058
step: 60, loss: 0.0011810851283371449
step: 70, loss: 0.0104710403829813
```

```
step: 80, loss: 0.00027136519202031195
step: 90, loss: 0.0014223585603758693
step: 100, loss: 0.02469244971871376
step: 110, loss: 0.0007955185719765723
step: 120, loss: 0.002686013001948595
step: 130, loss: 0.00018393757636658847
step: 140, loss: 0.001531643676571548
step: 150, loss: 0.014019266702234745
step: 160, loss: 0.0004477996553760022
step: 170, loss: 0.006251508370041847
step: 180, loss: 0.002632019342854619
step: 190, loss: 0.04770032316446304
step: 200, loss: 0.0002337056357646361
step: 210, loss: 0.0010609554592519999
step: 220, loss: 0.0009187511750496924
step: 230, loss: 0.0020790596026927233
step: 240, loss: 0.0054315985180437565
step: 250, loss: 0.003604297060519457
step: 260, loss: 0.0033695485908538103
step: 270, loss: 0.10304024815559387
step: 280, loss: 0.0018872878281399608
step: 290, loss: 0.0054367221891880035
step: 300, loss: 0.00039038274553604424
step: 310, loss: 0.0009393793880008161
step: 320, loss: 0.08377502858638763
step: 330, loss: 0.004725623410195112
step: 340, loss: 0.0018878844566643238
step: 350, loss: 0.002500674454495311
step: 360, loss: 0.04375027120113373
step: 370, loss: 0.08248322457075119
step: 380, loss: 0.013750685378909111
step: 390, loss: 0.009397030808031559
step: 400, loss: 0.018279606476426125
step: 410, loss: 0.002898993669077754
step: 420, loss: 0.07192366570234299
step: 430, loss: 0.00023794315347913653
step: 440, loss: 0.002674964489415288
step: 450, loss: 0.013581543229520321
step: 460, loss: 0.002337624551728368
step: 470, loss: 0.17294111847877502
step: 480, loss: 0.0016375656705349684
step: 490, loss: 0.02794524095952511
step: 500, loss: 0.0021130640525370836
step: 510, loss: 0.0012610402191057801
step: 520, loss: 0.0007217181264422834
step: 530, loss: 0.004500535316765308
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
```

```
warnings.warn("An input tensor was not cuda.")
epoch 7: dev_f1=0.9592215013901761, f1=0.949041608228144,
best_f1=0.9572490706319703
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00035802103229798377
step: 10, loss: 0.004596663173288107
step: 20, loss: 0.004780293442308903
step: 30, loss: 0.007777781691402197
step: 40, loss: 0.0004963743267580867
step: 50, loss: 0.00264219893142581
step: 60, loss: 0.0011776128085330129
step: 70, loss: 0.010830393061041832
step: 80, loss: 0.019900143146514893
step: 90, loss: 0.000777171109803021
step: 100, loss: 0.001461999025195837
step: 110, loss: 0.00028433912666514516
step: 120, loss: 0.00609967764467001
step: 130, loss: 0.00019309567869640887
step: 140, loss: 4.437211828189902e-05
step: 150, loss: 0.001660339767113328
step: 160, loss: 0.014380349777638912
step: 170, loss: 0.07699351757764816
step: 180, loss: 0.0015958102885633707
step: 190, loss: 0.048182208091020584
step: 200, loss: 0.01342703402042389
step: 210, loss: 0.04130052775144577
step: 220, loss: 0.0013375728158280253
step: 230, loss: 0.058660708367824554
step: 240, loss: 0.039963189512491226
step: 250, loss: 0.003491191426292062
step: 260, loss: 0.002245516749098897
step: 270, loss: 0.011569044552743435
step: 280, loss: 0.0011464330600574613
step: 290, loss: 0.0008506861049681902
step: 300, loss: 2.1300918888300657e-05
step: 310, loss: 0.0028909891843795776
step: 320, loss: 0.000715552712790668
step: 330, loss: 0.00013524082896765321
step: 340, loss: 0.06405533850193024
step: 350, loss: 1.580999196448829e-05
step: 360, loss: 0.031120644882321358
step: 370, loss: 0.016985038295388222
step: 380, loss: 7.850342080928385e-05
step: 390, loss: 0.0020567034371197224
step: 400, loss: 0.007878130301833153
step: 410, loss: 0.0038145985454320908
```

```

step: 420, loss: 5.89491828577593e-05
step: 430, loss: 0.0032508675940334797
step: 440, loss: 0.0005343243246898055
step: 450, loss: 0.0012003097217530012
step: 460, loss: 0.014311439357697964
step: 470, loss: 0.11917249113321304
step: 480, loss: 0.014256294816732407
step: 490, loss: 0.01789155974984169
step: 500, loss: 0.00048631831305101514
step: 510, loss: 0.0017420639051124454
step: 520, loss: 3.922198811778799e-05
step: 530, loss: 0.0009360641706734896
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
  warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 8: dev_f1=0.9624826308476147, f1=0.9537037037037037,
best_f1=0.9537037037037037
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00018478845595382154
step: 10, loss: 0.06529224663972855
step: 20, loss: 0.0019125851104035974
step: 30, loss: 0.08939829468727112
step: 40, loss: 0.0008684788481332362
step: 50, loss: 0.00012866487668361515
step: 60, loss: 0.0007137803477235138
step: 70, loss: 0.0023116955999284983
step: 80, loss: 0.04550712928175926
step: 90, loss: 0.057076599448919296
step: 100, loss: 0.0024830601178109646
step: 110, loss: 0.0114351287484169
step: 120, loss: 0.007404487580060959
step: 130, loss: 0.00011931789049413055
step: 140, loss: 0.000638044613879174
step: 150, loss: 0.03345922380685806
step: 160, loss: 0.0003412170917727053
step: 170, loss: 0.025613082572817802
step: 180, loss: 0.004013087134808302
step: 190, loss: 0.044051799923181534
step: 200, loss: 0.0010971238370984793
step: 210, loss: 0.000821715802885592
step: 220, loss: 0.010701416060328484
step: 230, loss: 0.00026751376572065055

```

```

step: 240, loss: 0.0007820535101927817
step: 250, loss: 0.0004288034397177398
step: 260, loss: 0.00010160452802665532
step: 270, loss: 0.008044680580496788
step: 280, loss: 0.023554634302854538
step: 290, loss: 0.00017201392620336264
step: 300, loss: 4.001230263384059e-05
step: 310, loss: 0.14300642907619476
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 8192.0
step: 320, loss: 3.0053854061407037e-05
step: 330, loss: 0.0003704668488353491
step: 340, loss: 0.011880355887115002
step: 350, loss: 0.03057333454489708
step: 360, loss: 0.0057901679538190365
step: 370, loss: 2.2045498553779908e-05
step: 380, loss: 0.0002787526464089751
step: 390, loss: 0.003131596138700843
step: 400, loss: 0.03153271973133087
step: 410, loss: 0.003285862971097231
step: 420, loss: 0.0013946983963251114
step: 430, loss: 0.010125163942575455
step: 440, loss: 0.00012881981092505157
step: 450, loss: 0.005898864474147558
step: 460, loss: 0.003371002385392785
step: 470, loss: 0.0008298125467263162
step: 480, loss: 2.8072758141206577e-05
step: 490, loss: 0.0030415307264775038
step: 500, loss: 0.00013587083958555013
step: 510, loss: 0.008115204982459545
step: 520, loss: 0.005173483397811651
step: 530, loss: 0.0701865702867508
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
    warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 9: dev_f1=0.9632387156817124, f1=0.9505597014925373,
best_f1=0.9505597014925373
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0026594155933707952
step: 10, loss: 0.002403311664238572
step: 20, loss: 0.0006423784652724862
step: 30, loss: 0.01425972394645214
step: 40, loss: 0.002103601349517703

```

step: 50, loss: 0.001064275624230504  
step: 60, loss: 0.0050897798500955105  
step: 70, loss: 0.00010133111209142953  
step: 80, loss: 0.0005901182303205132  
step: 90, loss: 0.0192261915653944  
step: 100, loss: 0.00022335961693897843  
step: 110, loss: 0.007060574367642403  
step: 120, loss: 0.00034733262145891786  
step: 130, loss: 0.006173961795866489  
step: 140, loss: 0.0003842524311039597  
step: 150, loss: 0.07604622095823288  
step: 160, loss: 0.03877782076597214  
step: 170, loss: 0.008887743577361107  
step: 180, loss: 0.0008041900000534952  
step: 190, loss: 0.000149391038576141  
step: 200, loss: 0.018428344279527664  
step: 210, loss: 0.014162681065499783  
step: 220, loss: 0.0005870680324733257  
step: 230, loss: 0.0007401747861877084  
step: 240, loss: 0.0010766731575131416  
step: 250, loss: 0.0030913110822439194  
step: 260, loss: 0.03435574844479561  
step: 270, loss: 0.00039187882794067264  
step: 280, loss: 0.04020993411540985  
step: 290, loss: 0.0004283625748939812  
step: 300, loss: 0.002165599726140499  
step: 310, loss: 0.032928239554166794  
step: 320, loss: 0.1026989221572876  
step: 330, loss: 0.006936361081898212  
step: 340, loss: 0.0021797672379761934  
step: 350, loss: 0.0002705533988773823  
step: 360, loss: 0.00010897684114752337  
step: 370, loss: 0.00602718023583293  
step: 380, loss: 0.0035796407610177994  
step: 390, loss: 0.00014582292351406068  
step: 400, loss: 0.0003477031132206321  
step: 410, loss: 0.004066180437803268  
step: 420, loss: 8.133728988468647e-05  
step: 430, loss: 0.0011177321430295706  
step: 440, loss: 5.439463348011486e-05  
step: 450, loss: 0.0018940864829346538  
step: 460, loss: 0.0001775687123881653  
step: 470, loss: 0.0032219416461884975  
step: 480, loss: 0.00014110740448813885  
step: 490, loss: 0.01402001641690731  
step: 500, loss: 0.00021515130356419832  
step: 510, loss: 3.904594632331282e-05  
step: 520, loss: 0.0015297490172088146



```

step: 530, loss: 0.01886620558798313
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 10: dev_f1=0.9611605053813758, f1=0.9507735583684951,
best_f1=0.9505597014925373
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 7.2344801083090715e-06
step: 10, loss: 0.09671288728713989
step: 20, loss: 0.004745313432067633
step: 30, loss: 9.841352584771812e-05
step: 40, loss: 3.084695345023647e-05
step: 50, loss: 0.00033990497468039393
step: 60, loss: 0.008405117318034172
step: 70, loss: 0.00021458482660818845
step: 80, loss: 0.0006995668518356979
step: 90, loss: 0.026296572759747505
step: 100, loss: 0.008794020861387253
step: 110, loss: 0.0007779246661812067
step: 120, loss: 0.0033213114365935326
step: 130, loss: 0.0005831941380165517
step: 140, loss: 0.00011862950486829504
step: 150, loss: 0.0005670422688126564
step: 160, loss: 0.003522927174344659
step: 170, loss: 0.005426212213933468
step: 180, loss: 0.001112865167669952
step: 190, loss: 0.01471760869026184
step: 200, loss: 0.00019345640612300485
step: 210, loss: 0.006997745484113693
step: 220, loss: 0.016910789534449577
step: 230, loss: 0.0011451991740614176
step: 240, loss: 0.0026529296301305294
step: 250, loss: 5.006076389690861e-05
step: 260, loss: 0.0027991756796836853
step: 270, loss: 0.0013977165799587965
step: 280, loss: 0.00041082993266172707
step: 290, loss: 0.00012445719039533287
step: 300, loss: 0.0001688018091954291
step: 310, loss: 0.02776365913450718
step: 320, loss: 4.931847070110962e-05
step: 330, loss: 6.817250778112793e-06
step: 340, loss: 0.002487519523128867
step: 350, loss: 0.00018757907673716545
step: 360, loss: 0.0012264893157407641
step: 370, loss: 0.003724772250279784
step: 380, loss: 0.0010146163403987885

```

```

step: 390, loss: 0.0018770609749481082
step: 400, loss: 0.00036153028486296535
step: 410, loss: 0.010135024785995483
step: 420, loss: 0.0013925059465691447
step: 430, loss: 0.0022648104932159185
step: 440, loss: 0.00022741426073480397
step: 450, loss: 0.0009835097007453442
step: 460, loss: 0.007886902429163456
step: 470, loss: 0.0008952436619438231
step: 480, loss: 0.0009313944610767066
step: 490, loss: 0.005806954577565193
step: 500, loss: 0.0014946268638595939
step: 510, loss: 0.0015495683765038848
step: 520, loss: 0.0005503483698703349
step: 530, loss: 0.004044108092784882
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 11: dev_f1=0.9594027064862343, f1=0.9495327102803738,
best_f1=0.9505597014925373
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.010764333419501781
step: 10, loss: 0.0007791179232299328
step: 20, loss: 0.011773547157645226
step: 30, loss: 0.052172932773828506
step: 40, loss: 0.003285813145339489
step: 50, loss: 2.0825535102630965e-05
step: 60, loss: 7.599541277159005e-06
step: 70, loss: 0.07336601614952087
step: 80, loss: 0.0005753104924224317
step: 90, loss: 0.008231799118220806
step: 100, loss: 0.010073709301650524
step: 110, loss: 0.0009389764163643122
step: 120, loss: 0.00027518608840182424
step: 130, loss: 0.0009753410122357309
step: 140, loss: 0.0005287298117764294
step: 150, loss: 9.413645602762699e-05
step: 160, loss: 0.002342823427170515
step: 170, loss: 0.0006390862981788814
step: 180, loss: 9.719206900626887e-06
step: 190, loss: 0.0011644366895779967
step: 200, loss: 0.004079232923686504
step: 210, loss: 2.4925760953919962e-05
step: 220, loss: 1.1164521311002318e-05
step: 230, loss: 2.078868965327274e-05
step: 240, loss: 0.006798962131142616

```

```

step: 250, loss: 6.865680916234851e-06
step: 260, loss: 0.0003339808899909258
step: 270, loss: 0.02416927181184292
step: 280, loss: 9.908984793582931e-06
step: 290, loss: 0.0004187375889159739
step: 300, loss: 0.007850799709558487
step: 310, loss: 6.615873280679807e-05
step: 320, loss: 0.0278368778526783
step: 330, loss: 0.049005962908267975
step: 340, loss: 0.0050856443122029305
step: 350, loss: 8.43618981889449e-05
step: 360, loss: 0.0007980524678714573
step: 370, loss: 0.029574241489171982
step: 380, loss: 0.0010785451158881187
step: 390, loss: 0.0039946832694113255
step: 400, loss: 7.752294550300576e-06
step: 410, loss: 0.0005152987432666123
step: 420, loss: 0.004132270347326994
step: 430, loss: 0.000568292336538434
step: 440, loss: 0.012587814591825008
step: 450, loss: 0.0005349352722987533
step: 460, loss: 0.0004914070595987141
step: 470, loss: 0.0004916446632705629
step: 480, loss: 0.004666990600526333
step: 490, loss: 1.0587060387479141e-05
step: 500, loss: 1.1756725143641233e-05
step: 510, loss: 0.02621585875749588
step: 520, loss: 3.478031067061238e-05
step: 530, loss: 3.38695062964689e-05
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 12: dev_f1=0.962686567164179, f1=0.952513966480447,
best_f1=0.9505597014925373
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00024789056624285877
step: 10, loss: 0.00040759294643066823
step: 20, loss: 1.1212886420253199e-05
step: 30, loss: 0.0011600313009694219
step: 40, loss: 0.03842329606413841
step: 50, loss: 0.001923676230944693
step: 60, loss: 0.00013385065540205687
step: 70, loss: 0.001343397656455636
step: 80, loss: 0.004227669909596443
step: 90, loss: 0.00023113255156204104
step: 100, loss: 2.718829273362644e-05

```

```
step: 110, loss: 6.348833267111331e-05
step: 120, loss: 3.453204772085883e-05
step: 130, loss: 0.00016763084568083286
step: 140, loss: 0.023506013676524162
step: 150, loss: 0.003772907657548785
step: 160, loss: 0.004546088166534901
step: 170, loss: 0.20466427505016327
step: 180, loss: 0.00023571375641040504
step: 190, loss: 8.916441584005952e-05
step: 200, loss: 0.00010261764691676944
step: 210, loss: 7.285329047590494e-05
step: 220, loss: 0.0002797955530695617
step: 230, loss: 0.0015082549070939422
step: 240, loss: 0.0013112594606354833
step: 250, loss: 0.00043158960761502385
step: 260, loss: 0.127437025308609
step: 270, loss: 0.00022836599964648485
step: 280, loss: 0.004914438351988792
step: 290, loss: 9.069872612599283e-05
step: 300, loss: 4.843416536459699e-05
step: 310, loss: 0.0008093569194898009
step: 320, loss: 5.982838774798438e-05
step: 330, loss: 0.002055772813037038
step: 340, loss: 0.002902704756706953
step: 350, loss: 0.0001515724725322798
step: 360, loss: 0.09126594662666321
step: 370, loss: 0.03110538423061371
step: 380, loss: 0.00041883974336087704
step: 390, loss: 0.0006972473347559571
step: 400, loss: 0.0008609776268713176
step: 410, loss: 0.0012121694162487984
step: 420, loss: 0.0010691253701224923
step: 430, loss: 8.213154796976596e-05
step: 440, loss: 7.621203258167952e-05
step: 450, loss: 0.0015869243070483208
step: 460, loss: 0.022927580401301384
step: 470, loss: 0.0032776123844087124
step: 480, loss: 2.6691182938520797e-05
step: 490, loss: 0.0001366602664347738
step: 500, loss: 0.0009104557102546096
step: 510, loss: 4.664563311962411e-05
step: 520, loss: 0.0001251984213013202
step: 530, loss: 0.00015430121857207268
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 8192.0
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
epoch 13: dev_f1=0.9577333952624245, f1=0.9512761020881672,
```

```
best_f1=0.9505597014925373
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0023154260125011206
step: 10, loss: 0.010905211791396141
step: 20, loss: 0.00016743458400014788
step: 30, loss: 0.0005620743031613529
step: 40, loss: 0.000801270070951432
step: 50, loss: 0.0023571932688355446
step: 60, loss: 0.0003714955528266728
step: 70, loss: 0.00010365905473008752
step: 80, loss: 0.0008447286090813577
step: 90, loss: 1.326190340478206e-05
step: 100, loss: 4.279095810488798e-05
step: 110, loss: 0.0023221622686833143
step: 120, loss: 3.4835742553696036e-05
step: 130, loss: 0.000428423605626449
step: 140, loss: 0.012036972679197788
step: 150, loss: 0.053274448961019516
step: 160, loss: 0.001881436095573008
step: 170, loss: 0.00035151041811332107
step: 180, loss: 0.00032545390422455966
step: 190, loss: 0.000459502509329468
step: 200, loss: 2.1415591618278995e-05
step: 210, loss: 0.0017738359747454524
step: 220, loss: 2.3569073164253496e-05
step: 230, loss: 2.824008151947055e-05
step: 240, loss: 0.0013764589093625546
step: 250, loss: 0.001534249517135322
step: 260, loss: 0.0010733960662037134
step: 270, loss: 0.000172653395566158
step: 280, loss: 0.00164094811771065
step: 290, loss: 4.79056870972272e-05
step: 300, loss: 3.477983409538865e-05
step: 310, loss: 0.001074000378139317
step: 320, loss: 3.768552778637968e-05
step: 330, loss: 0.0014433579053729773
step: 340, loss: 0.002086825668811798
step: 350, loss: 0.0003341918927617371
step: 360, loss: 0.0007882514037191868
step: 370, loss: 0.00014072423800826073
step: 380, loss: 0.011498860083520412
step: 390, loss: 0.0010110343573614955
step: 400, loss: 0.004091033712029457
step: 410, loss: 0.00021077720157336444
step: 420, loss: 0.0005740420892834663
step: 430, loss: 0.004000141751021147
```

```

step: 440, loss: 0.08000828325748444
step: 450, loss: 0.0004369065281935036
step: 460, loss: 0.03218889236450195
step: 470, loss: 0.00011693067790474743
step: 480, loss: 0.002355035161599517
step: 490, loss: 0.012914017774164677
step: 500, loss: 0.0022173626348376274
step: 510, loss: 0.0156725887209177
step: 520, loss: 0.0013238451210781932
step: 530, loss: 0.0012825870653614402
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
epoch 14: dev_f1=0.9586466165413533, f1=0.9516204790981682,
best_f1=0.9505597014925373
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 7.505342364311218e-05
step: 10, loss: 0.0012665127869695425
step: 20, loss: 0.0027801054529845715
step: 30, loss: 0.004911577329039574
step: 40, loss: 1.14290096462355e-05
step: 50, loss: 0.03994729742407799
step: 60, loss: 0.0009682181407697499
step: 70, loss: 3.021077463927213e-05
step: 80, loss: 4.445790546014905e-05
step: 90, loss: 0.00011711588012985885
step: 100, loss: 1.2934084224980325e-05
step: 110, loss: 0.01479310542345047
step: 120, loss: 0.0006577293388545513
step: 130, loss: 0.005679327994585037
step: 140, loss: 1.577620059833862e-05
step: 150, loss: 4.542768147075549e-05
step: 160, loss: 0.0001082858580048196
step: 170, loss: 1.6119081919896416e-05
step: 180, loss: 0.0033422799315303564
step: 190, loss: 0.007939408533275127
step: 200, loss: 0.006702031474560499
step: 210, loss: 0.00034621654776856303
step: 220, loss: 1.331401745119365e-05
step: 230, loss: 0.0020591006614267826
step: 240, loss: 0.0023071905598044395
step: 250, loss: 0.0011079078540205956
step: 260, loss: 1.556399183755275e-05
step: 270, loss: 9.719212357595097e-06
step: 280, loss: 2.9469407309079543e-05
step: 290, loss: 0.001928593497723341

```

```

step: 300, loss: 0.00027992326067760587
step: 310, loss: 0.040723033249378204
step: 320, loss: 0.0001737466809572652
step: 330, loss: 5.796709956484847e-05
step: 340, loss: 0.00018782936967909336
step: 350, loss: 0.004358815960586071
step: 360, loss: 5.725360824726522e-05
step: 370, loss: 0.04606293886899948
step: 380, loss: 0.0025561617221683264
step: 390, loss: 0.0003810809284914285
step: 400, loss: 0.026341117918491364
step: 410, loss: 0.0006582846399396658
step: 420, loss: 0.015501920133829117
step: 430, loss: 5.031324326409958e-05
step: 440, loss: 4.4540996896103024e-05
step: 450, loss: 0.007045472040772438
step: 460, loss: 0.005483231041580439
step: 470, loss: 0.00019204572890885174
step: 480, loss: 0.0017016020137816668
step: 490, loss: 0.0014794677263125777
step: 500, loss: 0.0006638378836214542
step: 510, loss: 8.789364073891193e-05
step: 520, loss: 1.550410524941981e-05
step: 530, loss: 0.0006877020932734013
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
epoch 15: dev_f1=0.9583528310715957, f1=0.9536299765807962,
best_f1=0.9505597014925373
mv: cannot stat '*_dev.pt': No such file or directory
Selected optimization level 02: FP16 training with FP32 batchnorm and FP32
master weights.

Defaults for this optimization level are:
enabled                : True
opt_level              : 02
cast_model_type        : torch.float16
patch_torch_functions  : False
keep_batchnorm_fp32    : True
master_weights         : True
loss_scale             : dynamic
Processing user overrides (additional kwargs that are not None)...
After processing overrides, optimization options are:
enabled                : True
opt_level              : 02
cast_model_type        : torch.float16
patch_torch_functions  : False
keep_batchnorm_fp32    : True

```

```

master_weights      : True
loss_scale          : dynamic
Warning: multi_tensor_applier fused unscale kernel is unavailable, possibly
because apex was installed without --cuda_ext --cpp_ext. Using Python fallback.
Original ImportError was: ModuleNotFoundError("No module named 'amp_C'")
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
5742it [00:20, 279.32it/s]
load_f1 = 0.9606299212598425
real_f1 = 0.958256029684601
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
4398it [00:22, 199.47it/s]

```

#### 4.1.4 Amazon-Google - Running the matcher

```

[ ]: # Step 1:
# train the wdc model with the save_model flag on
# it will produce two files *_dev.pt and *_test.pt which are the
# best checkpoints on the dev set and the test set.
!CUDA_VISIBLE_DEVICES=0 python train_ditto.py \
    --task Structured/Amazon-Google \
    --batch_size 32 \
    --max_len 256 \
    --lr 3e-5 \
    --n_epochs 15 \
    --finetuning \
    --lm roberta \
    --fp16 \
    --save_model \
    --dk general \
    --dk product \
    --summarize \
    --da entry_swap \
    --da attr_del

# Step 2: pick a checkpoint, rename it to wdc_all_small.pt
!mv *_dev.pt checkpoints/Amazon-Google.pt

# Step 3: run the matcher

```



```
!CUDA_VISIBLE_DEVICES=0 python matcher.py \
--task Structured/Amazon-Google \
--input_path input/input_small.jsonl \
--output_path output/output_small.jsonl \
--lm roberta \
--use_gpu \
--fp16 \
--checkpoint_path checkpoints/
```

Selected optimization level 02: FP16 training with FP32 batchnorm and FP32 master weights.

Defaults for this optimization level are:

```
enabled          : True
opt_level        : 02
cast_model_type  : torch.float16
patch_torch_functions : False
keep_batchnorm_fp32 : True
master_weights   : True
loss_scale       : dynamic
```

Processing user overrides (additional kwargs that are not None)...

After processing overrides, optimization options are:

```
enabled          : True
opt_level        : 02
cast_model_type  : torch.float16
patch_torch_functions : False
keep_batchnorm_fp32 : True
master_weights   : True
loss_scale       : dynamic
```

Warning: multi\_tensor\_applier fused unscale kernel is unavailable, possibly because apex was installed without --cuda\_ext --cpp\_ext. Using Python fallback. Original ImportError was: ModuleNotFoundError("No module named 'amp\_C'")  
/usr/local/lib/python3.7/dist-packages/apex/amp/\_initialize.py:25: UserWarning: An input tensor was not cuda.

```
warnings.warn("An input tensor was not cuda.")
```

Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 32768.0  
/usr/local/lib/python3.7/dist-packages/torch/optim/lr\_scheduler.py:136:

UserWarning: Detected call of `lr\_scheduler.step()` before `optimizer.step()`.

In PyTorch 1.1.0 and later, you should call them in the opposite order: `optimizer.step()` before `lr\_scheduler.step()`. Failure to do this will result in PyTorch skipping the first value of the learning rate schedule. See more details at <https://pytorch.org/docs/stable/optim.html#how-to-adjust-learning-rate>

```
"https://pytorch.org/docs/stable/optim.html#how-to-adjust-learning-rate",
UserWarning)
```

```
step: 0, loss: 0.5204046964645386
```

Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0

```

step: 10, loss: 0.4098491668701172
step: 20, loss: 0.4282356798648834
step: 30, loss: 0.3894853889942169
step: 40, loss: 0.3341982066631317
step: 50, loss: 0.42835357785224915
step: 60, loss: 0.47358977794647217
step: 70, loss: 0.2872677743434906
step: 80, loss: 0.2862933278083801
step: 90, loss: 0.2777577042579651
step: 100, loss: 0.2724088132381439
step: 110, loss: 0.23944641649723053
step: 120, loss: 0.2397891879081726
step: 130, loss: 0.18337029218673706
step: 140, loss: 0.2946924865245819
step: 150, loss: 0.25182849168777466
step: 160, loss: 0.35937580466270447
step: 170, loss: 0.13006819784641266
step: 180, loss: 0.17044341564178467
step: 190, loss: 0.2986518442630768
step: 200, loss: 0.16377143561840057
step: 210, loss: 0.29619669914245605
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please save or load the state of the optimizer when saving or
loading the scheduler.
    warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 1: dev_f1=0.6666666666666667, f1=0.6704545454545455,
best_f1=0.6704545454545455
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.1790124624967575
step: 10, loss: 0.12023264914751053
step: 20, loss: 0.25708332657814026
step: 30, loss: 0.2257596254348755
step: 40, loss: 0.4002583622932434
step: 50, loss: 0.12526799738407135
step: 60, loss: 0.2600589692592621
step: 70, loss: 0.1532895714044571
step: 80, loss: 0.1562805473804474
step: 90, loss: 0.17996913194656372
step: 100, loss: 0.24607102572917938
step: 110, loss: 0.26758286356925964
step: 120, loss: 0.11268910020589828
step: 130, loss: 0.1004338264465332
step: 140, loss: 0.08414091169834137

```

```

step: 150, loss: 0.24495644867420197
step: 160, loss: 0.10901181399822235
step: 170, loss: 0.22295229136943817
step: 180, loss: 0.17631220817565918
step: 190, loss: 0.3294292092323303
step: 200, loss: 0.08788543939590454
step: 210, loss: 0.13550256192684174
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
    warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 2: dev_f1=0.7119266055045872, f1=0.7052238805970149,
best_f1=0.7052238805970149
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.04675284028053284
step: 10, loss: 0.03510666638612747
step: 20, loss: 0.2793932557106018
step: 30, loss: 0.07590377330780029
step: 40, loss: 0.22189883887767792
step: 50, loss: 0.21877732872962952
step: 60, loss: 0.2374219000339508
step: 70, loss: 0.13632868230342865
step: 80, loss: 0.15354453027248383
step: 90, loss: 0.04928334802389145
step: 100, loss: 0.27579665184020996
step: 110, loss: 0.11773397028446198
step: 120, loss: 0.2021217942237854
step: 130, loss: 0.18014009296894073
step: 140, loss: 0.13678131997585297
step: 150, loss: 0.12713778018951416
step: 160, loss: 0.05962232127785683
step: 170, loss: 0.2647070586681366
step: 180, loss: 0.08622056990861893
step: 190, loss: 0.03069785051047802
step: 200, loss: 0.11106795072555542
step: 210, loss: 0.11854992061853409
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 3: dev_f1=0.709278350515464, f1=0.7239263803680981,
best_f1=0.7052238805970149
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.

```

```

    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.022010713815689087
step: 10, loss: 0.11061398684978485
step: 20, loss: 0.05269058048725128
step: 30, loss: 0.08235056698322296
step: 40, loss: 0.10924036800861359
step: 50, loss: 0.07794587314128876
step: 60, loss: 0.12242192029953003
step: 70, loss: 0.12180744111537933
step: 80, loss: 0.08603710681200027
step: 90, loss: 0.04778940975666046
step: 100, loss: 0.18567347526550293
step: 110, loss: 0.47185218334198
step: 120, loss: 0.16523319482803345
step: 130, loss: 0.21356429159641266
step: 140, loss: 0.216279074549675
step: 150, loss: 0.1684330701828003
step: 160, loss: 0.08450277894735336
step: 170, loss: 0.11801943182945251
step: 180, loss: 0.07075013965368271
step: 190, loss: 0.1796753853559494
step: 200, loss: 0.08802621066570282
step: 210, loss: 0.37627142667770386
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
    warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 4: dev_f1=0.7137404580152672, f1=0.7134502923976609,
best_f1=0.7134502923976609
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.10904869437217712
step: 10, loss: 0.03908824920654297
step: 20, loss: 0.1086105927824974
step: 30, loss: 0.020890718325972557
step: 40, loss: 0.054379865527153015
step: 50, loss: 0.08867598325014114
step: 60, loss: 0.02974236197769642
step: 70, loss: 0.19330640137195587
step: 80, loss: 0.08513778448104858
step: 90, loss: 0.10773425549268723
step: 100, loss: 0.03095628134906292
step: 110, loss: 0.06302762776613235
step: 120, loss: 0.18168146908283234

```

```

step: 130, loss: 0.15765085816383362
step: 140, loss: 0.08582073450088501
step: 150, loss: 0.09951235353946686
step: 160, loss: 0.029733771458268166
step: 170, loss: 0.03271640092134476
step: 180, loss: 0.22646912932395935
step: 190, loss: 0.19667287170886993
step: 200, loss: 0.046762414276599884
step: 210, loss: 0.07195688039064407
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
epoch 5: dev_f1=0.7122302158273381, f1=0.7116788321167884,
best_f1=0.7134502923976609
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.011236474849283695
step: 10, loss: 0.03570844978094101
step: 20, loss: 0.07178131490945816
step: 30, loss: 0.007168125826865435
step: 40, loss: 0.08264325559139252
step: 50, loss: 0.022455981001257896
step: 60, loss: 0.1389457881450653
step: 70, loss: 0.029205327853560448
step: 80, loss: 0.07488338649272919
step: 90, loss: 0.043913766741752625
step: 100, loss: 0.1338169425725937
step: 110, loss: 0.1715846210718155
step: 120, loss: 0.013503414578735828
step: 130, loss: 0.010468140244483948
step: 140, loss: 0.17668387293815613
step: 150, loss: 0.0830390676856041
step: 160, loss: 0.06857162714004517
step: 170, loss: 0.05729743465781212
step: 180, loss: 0.049996864050626755
step: 190, loss: 0.02449096366763115
step: 200, loss: 0.10328357666730881
step: 210, loss: 0.13349321484565735
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
  warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 6: dev_f1=0.7210626185958253, f1=0.7372400756143668,
best_f1=0.7372400756143668

```

```

/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.04421515390276909
step: 10, loss: 0.004959010053426027
step: 20, loss: 0.1598052978515625
step: 30, loss: 0.05969657748937607
step: 40, loss: 0.0535624735057354
step: 50, loss: 0.07467935234308243
step: 60, loss: 0.11572115868330002
step: 70, loss: 0.1039196327328682
step: 80, loss: 0.08561795949935913
step: 90, loss: 0.04135006666183472
step: 100, loss: 0.025587525218725204
step: 110, loss: 0.08545288443565369
step: 120, loss: 0.1257496476173401
step: 130, loss: 0.0544908344745636
step: 140, loss: 0.05588633567094803
step: 150, loss: 0.032894205302000046
step: 160, loss: 0.2442503720521927
step: 170, loss: 0.10308369994163513
step: 180, loss: 0.019249632954597473
step: 190, loss: 0.016875281929969788
step: 200, loss: 0.11026990413665771
step: 210, loss: 0.09476551413536072
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
    warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 7: dev_f1=0.7326007326007326, f1=0.7179487179487178,
best_f1=0.7179487179487178
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.07972946763038635
step: 10, loss: 0.07255186885595322
step: 20, loss: 0.2220713347196579
step: 30, loss: 0.09865846484899521
step: 40, loss: 0.006686465349048376
step: 50, loss: 0.0028631568420678377
step: 60, loss: 0.06443697214126587
step: 70, loss: 0.05804392322897911
step: 80, loss: 0.022063661366701126
step: 90, loss: 0.019286517053842545
step: 100, loss: 0.2799510955810547

```

```

step: 110, loss: 0.02044278383255005
step: 120, loss: 0.18131273984909058
step: 130, loss: 0.004655086435377598
step: 140, loss: 0.0689420998096466
step: 150, loss: 0.09974920749664307
step: 160, loss: 0.23271948099136353
step: 170, loss: 0.4114861488342285
step: 180, loss: 0.09556970745325089
step: 190, loss: 0.013293236494064331
step: 200, loss: 0.055995579808950424
step: 210, loss: 0.28492480516433716
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
epoch 8: dev_f1=0.7116104868913857, f1=0.7299813780260707,
best_f1=0.7179487179487178
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.07646532356739044
step: 10, loss: 0.026582272723317146
step: 20, loss: 0.1694105714559555
step: 30, loss: 0.018420936539769173
step: 40, loss: 0.06473279744386673
step: 50, loss: 0.09096425026655197
step: 60, loss: 0.03240804746747017
step: 70, loss: 0.10062642395496368
step: 80, loss: 0.009752719663083553
step: 90, loss: 0.034686923027038574
step: 100, loss: 0.02667817659676075
step: 110, loss: 0.046759232878685
step: 120, loss: 0.062223706394433975
step: 130, loss: 0.1983611136674881
step: 140, loss: 0.11403357982635498
step: 150, loss: 0.10454761236906052
step: 160, loss: 0.12883326411247253
step: 170, loss: 0.03579483553767204
step: 180, loss: 0.012480605393648148
step: 190, loss: 0.007277380209416151
step: 200, loss: 0.028208844363689423
step: 210, loss: 0.039784833788871765
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
epoch 9: dev_f1=0.7169811320754716, f1=0.7131147540983607,
best_f1=0.7179487179487178
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.

```

```

    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.023089220747351646
step: 10, loss: 0.023113111034035683
step: 20, loss: 0.011524368077516556
step: 30, loss: 0.14710554480552673
step: 40, loss: 0.020448768511414528
step: 50, loss: 0.012635812163352966
step: 60, loss: 0.06658707559108734
step: 70, loss: 0.0770399197936058
step: 80, loss: 0.040933478623628616
step: 90, loss: 0.05556020885705948
step: 100, loss: 0.043535053730010986
step: 110, loss: 0.04174056649208069
step: 120, loss: 0.1633533090353012
step: 130, loss: 0.03723622485995293
step: 140, loss: 0.014418663457036018
step: 150, loss: 0.02720007859170437
step: 160, loss: 0.01657729409635067
step: 170, loss: 0.023558611050248146
step: 180, loss: 0.03637181594967842
step: 190, loss: 0.025803688913583755
step: 200, loss: 0.04451366141438484
step: 210, loss: 0.09874773025512695
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
    warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 10: dev_f1=0.7480916030534351, f1=0.7158671586715868,
best_f1=0.7158671586715868
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.06300103664398193
step: 10, loss: 0.037126556038856506
step: 20, loss: 0.04701916128396988
step: 30, loss: 0.0512925423681736
step: 40, loss: 0.06238177791237831
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 50, loss: 0.204104483127594
step: 60, loss: 0.033610228449106216
step: 70, loss: 0.041645459830760956
step: 80, loss: 0.0628957524895668
step: 90, loss: 0.10694035142660141
step: 100, loss: 0.02351948991417885
step: 110, loss: 0.033702317625284195

```



```

step: 120, loss: 0.030656766146421432
step: 130, loss: 0.0016140005318447948
step: 140, loss: 0.05275467410683632
step: 150, loss: 0.028917428106069565
step: 160, loss: 0.007734468672424555
step: 170, loss: 0.03773246333003044
step: 180, loss: 0.0689300149679184
step: 190, loss: 0.0750681459903717
step: 200, loss: 0.02197769656777382
step: 210, loss: 0.004735317546874285
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 11: dev_f1=0.7410358565737052, f1=0.72, best_f1=0.7158671586715868
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.05997230112552643
step: 10, loss: 0.08837976306676865
step: 20, loss: 0.06614800542593002
step: 30, loss: 0.010512243956327438
step: 40, loss: 0.03358918055891991
step: 50, loss: 0.012168730609118938
step: 60, loss: 0.04368404299020767
step: 70, loss: 0.09863171726465225
step: 80, loss: 0.09137008339166641
step: 90, loss: 0.04734048247337341
step: 100, loss: 0.015949612483382225
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 8192.0
step: 110, loss: 0.05234908685088158
step: 120, loss: 0.001547507825307548
step: 130, loss: 0.036625687032938004
step: 140, loss: 0.14955958724021912
step: 150, loss: 0.0014295054133981466
step: 160, loss: 0.04739350453019142
step: 170, loss: 0.08577504754066467
step: 180, loss: 0.03583697974681854
step: 190, loss: 0.06352345645427704
step: 200, loss: 0.03823414817452431
step: 210, loss: 0.005754147656261921
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 12: dev_f1=0.7338403041825093, f1=0.7398843930635838,
best_f1=0.7158671586715868
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")

```

```

step: 0, loss: 0.021057797595858574
step: 10, loss: 0.007029051426798105
step: 20, loss: 0.0014123414875939488
step: 30, loss: 0.0019439365714788437
step: 40, loss: 0.03469059616327286
step: 50, loss: 0.08197388052940369
step: 60, loss: 0.004183086100965738
step: 70, loss: 0.0036033177748322487
step: 80, loss: 0.20866559445858002
step: 90, loss: 0.014241954311728477
step: 100, loss: 0.044689029455184937
step: 110, loss: 0.07940243929624557
step: 120, loss: 0.004653529729694128
step: 130, loss: 0.007755889557301998
step: 140, loss: 0.04727901518344879
step: 150, loss: 0.004796223249286413
step: 160, loss: 0.045341309159994125
step: 170, loss: 0.037410981953144073
step: 180, loss: 0.021597471088171005
step: 190, loss: 0.01778414286673069
step: 200, loss: 0.057582344859838486
step: 210, loss: 0.04113953560590744
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 13: dev_f1=0.7198443579766536, f1=0.7290448343079923,
best_f1=0.7158671586715868
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.011164232157170773
step: 10, loss: 0.07867962121963501
step: 20, loss: 0.010096446610987186
step: 30, loss: 0.1657065451145172
step: 40, loss: 0.026999354362487793
step: 50, loss: 0.03654690086841583
step: 60, loss: 0.02545086480677128
step: 70, loss: 0.0038775885477662086
step: 80, loss: 0.0520847886800766
step: 90, loss: 0.03192787244915962
step: 100, loss: 0.011928868480026722
step: 110, loss: 0.001937358290888369
step: 120, loss: 0.0004757501883432269
step: 130, loss: 0.018711956217885017
step: 140, loss: 0.0036247526295483112
step: 150, loss: 0.05486619099974632
step: 160, loss: 0.009386016987264156
step: 170, loss: 0.03579220175743103

```

```

step: 180, loss: 0.007728705182671547
step: 190, loss: 0.04081472009420395
step: 200, loss: 0.007875371724367142
step: 210, loss: 0.001058233086951077
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 14: dev_f1=0.7239263803680981, f1=0.7211740041928721,
best_f1=0.7158671586715868
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.012762470170855522
step: 10, loss: 0.0038183850701898336
step: 20, loss: 0.06938423216342926
step: 30, loss: 0.0390612222542763
step: 40, loss: 0.012082698754966259
step: 50, loss: 0.11134126782417297
step: 60, loss: 0.047288138419389725
step: 70, loss: 0.031711842864751816
step: 80, loss: 0.05254291370511055
step: 90, loss: 0.030383514240384102
step: 100, loss: 0.01313787791877985
step: 110, loss: 0.07393774390220642
step: 120, loss: 0.06933936476707458
step: 130, loss: 0.018366249278187752
step: 140, loss: 0.013052064925432205
step: 150, loss: 0.06889667361974716
step: 160, loss: 0.006908879615366459
step: 170, loss: 0.026539212092757225
step: 180, loss: 0.00871942937374115
step: 190, loss: 0.07396213710308075
step: 200, loss: 0.0015149995451793075
step: 210, loss: 0.031244728714227676
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 15: dev_f1=0.7221095334685598, f1=0.728395061728395,
best_f1=0.7158671586715868
mv: cannot stat '*_dev.pt': No such file or directory
Selected optimization level 02: FP16 training with FP32 batchnorm and FP32
master weights.

Defaults for this optimization level are:
enabled                : True
opt_level              : 02
cast_model_type        : torch.float16
patch_torch_functions  : False

```

```

keep_batchnorm_fp32      : True
master_weights           : True
loss_scale                : dynamic
Processing user overrides (additional kwargs that are not None)...
After processing overrides, optimization options are:
enabled                   : True
opt_level                 : 02
cast_model_type           : torch.float16
patch_torch_functions     : False
keep_batchnorm_fp32      : True
master_weights           : True
loss_scale                : dynamic
Warning: multi_tensor_applier fused unscale kernel is unavailable, possibly
because apex was installed without --cuda_ext --cpp_ext. Using Python fallback.
Original ImportError was: ModuleNotFoundError("No module named 'amp_C'")
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
2293it [00:04, 534.22it/s]
load_f1 = 0.7453183520599251
real_f1 = 0.738970588235294
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
4398it [00:22, 197.67it/s]

```

#### 4.1.5 Walmart-Amazon - Running the matcher

```

[ ]: # Step 1:
# train the wdc model with the save_model flag on
# it will produce two files *_dev.pt and *_test.pt which are the
# best checkpoints on the dev set and the test set.
!CUDA_VISIBLE_DEVICES=0 python train_ditto.py \
--task Structured/Walmart-Amazon \
--batch_size 32 \
--max_len 256 \
--lr 3e-5 \
--n_epochs 15 \
--finetuning \
--lm roberta \
--fp16 \
--save_model \
--dk general \

```

```

--dk product \
--summarize \
--da entry_swap \
--da attr_del

# Step 2: pick a checkpoint, rename it to wdc_all_small.pt
!mv *_dev.pt checkpoints/Walmart-Amazon.pt

# Step 3: run the matcher
!CUDA_VISIBLE_DEVICES=0 python matcher.py \
  --task Structured/Walmart-Amazon \
  --input_path input/input_small.jsonl \
  --output_path output/output_small.jsonl \
  --lm roberta \
  --use_gpu \
  --fp16 \
  --checkpoint_path checkpoints/

```

Selected optimization level 02: FP16 training with FP32 batchnorm and FP32 master weights.

Defaults for this optimization level are:

```

enabled           : True
opt_level         : 02
cast_model_type   : torch.float16
patch_torch_functions : False
keep_batchnorm_fp32 : True
master_weights    : True
loss_scale        : dynamic

```

Processing user overrides (additional kwargs that are not None)...

After processing overrides, optimization options are:

```

enabled           : True
opt_level         : 02
cast_model_type   : torch.float16
patch_torch_functions : False
keep_batchnorm_fp32 : True
master_weights    : True
loss_scale        : dynamic

```

Warning: multi\_tensor\_applier fused unscale kernel is unavailable, possibly because apex was installed without --cuda\_ext --cpp\_ext. Using Python fallback. Original ImportError was: ModuleNotFoundError("No module named 'amp\_C'")  
 /usr/local/lib/python3.7/dist-packages/apex/amp/\_initialize.py:25: UserWarning: An input tensor was not cuda.

```
warnings.warn("An input tensor was not cuda.")
```

Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 32768.0  
 /usr/local/lib/python3.7/dist-packages/torch/optim/lr\_scheduler.py:136:

UserWarning: Detected call of `lr\_scheduler.step()` before `optimizer.step()`. In PyTorch 1.1.0 and later, you should call them in the opposite order: `optimizer.step()` before `lr\_scheduler.step()`. Failure to do this will result in PyTorch skipping the first value of the learning rate schedule. See more details at <https://pytorch.org/docs/stable/optim.html#how-to-adjust-learning-rate>

"<https://pytorch.org/docs/stable/optim.html#how-to-adjust-learning-rate>", UserWarning)

step: 0, loss: 0.4690689146518707

Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0

step: 10, loss: 0.3849039077758789

step: 20, loss: 0.30028092861175537

step: 30, loss: 0.36388877034187317

step: 40, loss: 0.2581064701080322

step: 50, loss: 0.31211498379707336

step: 60, loss: 0.4626849591732025

step: 70, loss: 0.42527031898498535

step: 80, loss: 0.1780678778886795

step: 90, loss: 0.3053607940673828

step: 100, loss: 0.41784965991973877

step: 110, loss: 0.2549942135810852

step: 120, loss: 0.36262500286102295

step: 130, loss: 0.3139588236808777

step: 140, loss: 0.17823295295238495

step: 150, loss: 0.2953103184700012

step: 160, loss: 0.2247256338596344

step: 170, loss: 0.4037485718727112

step: 180, loss: 0.13802097737789154

step: 190, loss: 0.14659054577350616

/usr/local/lib/python3.7/dist-packages/apex/amp/\_initialize.py:25: UserWarning: An input tensor was not cuda.

warnings.warn("An input tensor was not cuda.")

/usr/local/lib/python3.7/dist-packages/torch/optim/lr\_scheduler.py:216:

UserWarning: Please also save or load the state of the optimizer when saving or loading the scheduler.

warnings.warn(SAVE\_STATE\_WARNING, UserWarning)

epoch 1: dev\_f1=0.5577464788732395, f1=0.5848563968668407,

best\_f1=0.5848563968668407

/usr/local/lib/python3.7/dist-packages/apex/amp/\_initialize.py:25: UserWarning: An input tensor was not cuda.

warnings.warn("An input tensor was not cuda.")

step: 0, loss: 0.292331725358963

step: 10, loss: 0.38111841678619385

Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 8192.0

step: 20, loss: 0.7728230357170105

step: 30, loss: 0.17967988550662994

step: 40, loss: 0.22388514876365662

step: 50, loss: 0.3251321017742157

```

step: 60, loss: 0.26424694061279297
step: 70, loss: 0.1955321729183197
step: 80, loss: 0.08396411687135696
step: 90, loss: 0.21989277005195618
step: 100, loss: 0.1263689547777176
step: 110, loss: 0.14544713497161865
step: 120, loss: 0.14284415543079376
step: 130, loss: 0.1467936635017395
step: 140, loss: 0.3411487638950348
step: 150, loss: 0.1934611201286316
step: 160, loss: 0.1108553558588028
step: 170, loss: 0.02609948255121708
step: 180, loss: 0.06945917755365372
step: 190, loss: 0.03551734238862991
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
    warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 2: dev_f1=0.7969543147208122, f1=0.8099999999999999,
best_f1=0.8099999999999999
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.09314318001270294
step: 10, loss: 0.06861800700426102
step: 20, loss: 0.12869465351104736
step: 30, loss: 0.014325978234410286
step: 40, loss: 0.04844120889902115
step: 50, loss: 0.22288824617862701
step: 60, loss: 0.10518846660852432
step: 70, loss: 0.09669536352157593
step: 80, loss: 0.04767481982707977
step: 90, loss: 0.11924168467521667
step: 100, loss: 0.14687031507492065
step: 110, loss: 0.2367517501115799
step: 120, loss: 0.16729134321212769
step: 130, loss: 0.06298115104436874
step: 140, loss: 0.08419103175401688
step: 150, loss: 0.04974595457315445
step: 160, loss: 0.050276074558496475
step: 170, loss: 0.16056059300899506
step: 180, loss: 0.16514095664024353
step: 190, loss: 0.04119396582245827
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.

```

```

    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
    warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 3: dev_f1=0.8432432432432432, f1=0.8346456692913385,
best_f1=0.8346456692913385
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.021104130893945694
step: 10, loss: 0.033063266426324844
step: 20, loss: 0.05441530421376228
step: 30, loss: 0.03412005305290222
step: 40, loss: 0.06276194751262665
step: 50, loss: 0.06558386981487274
step: 60, loss: 0.149762362241745
step: 70, loss: 0.027639729902148247
step: 80, loss: 0.016188250854611397
step: 90, loss: 0.06019158288836479
step: 100, loss: 0.07128028571605682
step: 110, loss: 0.10552865266799927
step: 120, loss: 0.018638912588357925
step: 130, loss: 0.00822365190833807
step: 140, loss: 0.05825269967317581
step: 150, loss: 0.017081396654248238
step: 160, loss: 0.0050683957524597645
step: 170, loss: 0.02140522375702858
step: 180, loss: 0.03202773258090019
step: 190, loss: 0.015665967017412186
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
    warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 4: dev_f1=0.8556430446194225, f1=0.859375, best_f1=0.859375
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.04221842810511589
step: 10, loss: 0.02071050927042961
step: 20, loss: 0.0029925352428108454
step: 30, loss: 0.02820383757352829
step: 40, loss: 0.017036808654665947
step: 50, loss: 0.013233210891485214
step: 60, loss: 0.000984203303232789

```



```

step: 70, loss: 0.01659468002617359
step: 80, loss: 0.009475253522396088
step: 90, loss: 0.05146421119570732
step: 100, loss: 0.056135524064302444
step: 110, loss: 0.1734054982662201
step: 120, loss: 0.0339876189827919
step: 130, loss: 0.2396274209022522
step: 140, loss: 0.016817139461636543
step: 150, loss: 0.050956517457962036
step: 160, loss: 0.02003243938088417
step: 170, loss: 0.015331226401031017
step: 180, loss: 0.11769809573888779
step: 190, loss: 0.001629914389923215
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 5: dev_f1=0.851063829787234, f1=0.8446866485013624, best_f1=0.859375
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.005358667112886906
step: 10, loss: 0.02273150347173214
step: 20, loss: 0.007924224250018597
step: 30, loss: 0.0218097772449255
step: 40, loss: 0.021826624870300293
step: 50, loss: 0.03813847154378891
step: 60, loss: 0.010386382229626179
step: 70, loss: 0.1727013885974884
step: 80, loss: 0.05578220635652542
step: 90, loss: 0.022107291966676712
step: 100, loss: 0.04135071113705635
step: 110, loss: 0.028455354273319244
step: 120, loss: 0.024370741099119186
step: 130, loss: 0.025815891101956367
step: 140, loss: 0.0043165716342628
step: 150, loss: 0.004980803467333317
step: 160, loss: 0.01862468011677265
step: 170, loss: 0.13321128487586975
step: 180, loss: 0.07322096824645996
step: 190, loss: 0.005274154711514711
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 6: dev_f1=0.8465608465608466, f1=0.8502673796791445, best_f1=0.859375
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.028442542999982834

```

```

step: 10, loss: 0.0017766256351023912
step: 20, loss: 0.003142735455185175
step: 30, loss: 0.01726340688765049
step: 40, loss: 0.0031342406291514635
step: 50, loss: 0.0030741256196051836
step: 60, loss: 0.015061948448419571
step: 70, loss: 0.0012893449747934937
step: 80, loss: 0.0013253566576167941
step: 90, loss: 0.1581718623638153
step: 100, loss: 0.2638171315193176
step: 110, loss: 0.019205953925848007
step: 120, loss: 0.02354400046169758
step: 130, loss: 0.012120923027396202
step: 140, loss: 0.018164684996008873
step: 150, loss: 0.13952438533306122
step: 160, loss: 0.03492804244160652
step: 170, loss: 0.026098035275936127
step: 180, loss: 0.006919483654201031
step: 190, loss: 0.008606594055891037
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
epoch 7: dev_f1=0.8478260869565217, f1=0.8222222222222222, best_f1=0.859375
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.009321805089712143
step: 10, loss: 0.002303696470335126
step: 20, loss: 0.07633596658706665
step: 30, loss: 0.014784086495637894
step: 40, loss: 0.026508856564760208
step: 50, loss: 0.003073060652241111
step: 60, loss: 0.04357469454407692
step: 70, loss: 0.016554515808820724
step: 80, loss: 0.08152362704277039
step: 90, loss: 0.02212093584239483
step: 100, loss: 0.0035046308767050505
step: 110, loss: 0.009643892757594585
step: 120, loss: 0.007601057179272175
step: 130, loss: 0.0028403743635863066
step: 140, loss: 0.025260519236326218
step: 150, loss: 0.03625020012259483
step: 160, loss: 0.01295531541109085
step: 170, loss: 0.0008875473286025226
step: 180, loss: 0.053850289434194565
step: 190, loss: 0.038069140166044235
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.

```

```

    warnings.warn("An input tensor was not cuda.")
epoch 8: dev_f1=0.8390501319261214, f1=0.846153846153846, best_f1=0.859375
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.08153527230024338
step: 10, loss: 0.0006122501799836755
step: 20, loss: 0.04752529785037041
step: 30, loss: 0.011770642362535
step: 40, loss: 0.05870203301310539
step: 50, loss: 0.05445253849029541
step: 60, loss: 0.011233195662498474
step: 70, loss: 0.001430100528523326
step: 80, loss: 0.0013160041999071836
step: 90, loss: 0.00497959041967988
step: 100, loss: 0.0017075386131182313
step: 110, loss: 0.0013884254731237888
step: 120, loss: 0.0035776596050709486
step: 130, loss: 0.0010288155172020197
step: 140, loss: 0.0922045037150383
step: 150, loss: 0.001522966893389821
step: 160, loss: 0.002989289816468954
step: 170, loss: 0.20268063247203827
step: 180, loss: 0.004586648195981979
step: 190, loss: 0.0016671791672706604
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 9: dev_f1=0.8287292817679558, f1=0.8125, best_f1=0.859375
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0003443474997766316
step: 10, loss: 0.0017405642429366708
step: 20, loss: 0.00312843918800354
step: 30, loss: 0.014770988374948502
step: 40, loss: 0.0016419384628534317
step: 50, loss: 0.0007701702998019755
step: 60, loss: 0.11773979663848877
step: 70, loss: 0.01968294382095337
step: 80, loss: 0.003782479092478752
step: 90, loss: 0.0022030319087207317
step: 100, loss: 0.000910909497179091
step: 110, loss: 0.0012847557663917542
step: 120, loss: 0.0011988675687462091
step: 130, loss: 0.12537965178489685
step: 140, loss: 0.0014600034337490797
step: 150, loss: 0.02897210605442524

```

```

step: 160, loss: 0.0007839553873054683
step: 170, loss: 0.0017729285173118114
step: 180, loss: 0.022996144369244576
step: 190, loss: 0.03047446720302105
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 10: dev_f1=0.8378378378378378, f1=0.8101983002832861, best_f1=0.859375
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0016814046539366245
step: 10, loss: 0.002785471500828862
step: 20, loss: 0.0009104141499847174
step: 30, loss: 0.003597357077524066
step: 40, loss: 0.00033274790621362627
step: 50, loss: 0.060556624084711075
step: 60, loss: 0.013330881483852863
step: 70, loss: 0.00048097941908054054
step: 80, loss: 0.0003476650745142251
step: 90, loss: 0.023649433627724648
step: 100, loss: 0.0008085820591077209
step: 110, loss: 0.06896308809518814
step: 120, loss: 0.0009640430798754096
step: 130, loss: 0.0011129515478387475
step: 140, loss: 0.0026682852767407894
step: 150, loss: 0.002154726767912507
step: 160, loss: 0.0006755260983482003
step: 170, loss: 0.0017373483860865235
step: 180, loss: 0.002180117415264249
step: 190, loss: 0.009265189059078693
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 11: dev_f1=0.8194070080862533, f1=0.8044692737430168, best_f1=0.859375
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.021976280957460403
step: 10, loss: 0.016958659514784813
step: 20, loss: 0.021466419100761414
step: 30, loss: 0.007179811131209135
step: 40, loss: 0.015799006447196007
step: 50, loss: 0.0059892903082072735
step: 60, loss: 0.0508558489382267
step: 70, loss: 0.1173737645149231
step: 80, loss: 0.0021476184483617544
step: 90, loss: 0.0014144859742373228

```

```

step: 100, loss: 0.0017228645738214254
step: 110, loss: 0.01436020340770483
step: 120, loss: 0.006534643936902285
step: 130, loss: 0.0015271201264113188
step: 140, loss: 0.026398850604891777
step: 150, loss: 0.0119241988286376
step: 160, loss: 0.00027154237614013255
step: 170, loss: 0.00020137206593062729
step: 180, loss: 0.0002445823047310114
step: 190, loss: 0.002748796483501792
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
epoch 12: dev_f1=0.8241206030150753, f1=0.8195876288659794, best_f1=0.859375
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.001340752700343728
step: 10, loss: 0.010559320449829102
step: 20, loss: 0.0006584238144569099
step: 30, loss: 0.021801885217428207
step: 40, loss: 0.0002521021815482527
step: 50, loss: 0.008318813517689705
step: 60, loss: 0.004449158441275358
step: 70, loss: 0.0004883254878222942
step: 80, loss: 0.00037690423778258264
step: 90, loss: 0.00043206464033573866
step: 100, loss: 0.0006290231831371784
step: 110, loss: 0.00039565464248880744
step: 120, loss: 0.00023994380899239331
step: 130, loss: 0.0018726270645856857
step: 140, loss: 0.0011136865941807628
step: 150, loss: 0.00023756544396746904
step: 160, loss: 0.0002419106604065746
step: 170, loss: 0.000833479396533221
step: 180, loss: 0.0005786455003544688
step: 190, loss: 0.05855401232838631
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
epoch 13: dev_f1=0.827027027027027, f1=0.8171428571428572, best_f1=0.859375
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.003679265035316348
step: 10, loss: 0.03691936284303665
step: 20, loss: 0.0003771228657569736
step: 30, loss: 0.003581759985536337

```

```

step: 40, loss: 0.0030714503955096006
step: 50, loss: 0.0007574474439024925
step: 60, loss: 0.0028355452232062817
step: 70, loss: 0.0010483544319868088
step: 80, loss: 0.0019468714017421007
step: 90, loss: 0.0011381861986592412
step: 100, loss: 0.0004947255365550518
step: 110, loss: 0.00011533858196344227
step: 120, loss: 0.0010089135030284524
step: 130, loss: 0.00022607366554439068
step: 140, loss: 0.001324932905845344
step: 150, loss: 0.028984548524022102
step: 160, loss: 0.03363669663667679
step: 170, loss: 0.0007555675692856312
step: 180, loss: 0.0020354061853140593
step: 190, loss: 0.00020092807244509459
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 14: dev_f1=0.8272251308900523, f1=0.8288770053475937, best_f1=0.859375
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0002557150728534907
step: 10, loss: 0.0017792574362829328
step: 20, loss: 0.0826701894402504
step: 30, loss: 0.001814599963836372
step: 40, loss: 0.0031303318683058023
step: 50, loss: 0.0005379662616178393
step: 60, loss: 0.0011110407067462802
step: 70, loss: 0.011006012558937073
step: 80, loss: 0.0002785956603474915
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 8192.0
step: 90, loss: 0.000905546301510185
step: 100, loss: 0.014423535205423832
step: 110, loss: 0.028934264555573463
step: 120, loss: 0.0007281749276444316
step: 130, loss: 0.005509814713150263
step: 140, loss: 0.002535016043111682
step: 150, loss: 0.0003028371138498187
step: 160, loss: 0.0001980961678782478
step: 170, loss: 0.08437219262123108
step: 180, loss: 0.0004564502742141485
step: 190, loss: 0.0009812954813241959
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 15: dev_f1=0.8301886792452831, f1=0.8324022346368716, best_f1=0.859375

```

```
mv: cannot stat '*_dev.pt': No such file or directory
Selected optimization level O2:  FP16 training with FP32 batchnorm and FP32
master weights.
```

Defaults for this optimization level are:

```
enabled          : True
opt_level        : O2
cast_model_type  : torch.float16
patch_torch_functions : False
keep_batchnorm_fp32 : True
master_weights   : True
loss_scale       : dynamic
```

Processing user overrides (additional kwargs that are not None)...

After processing overrides, optimization options are:

```
enabled          : True
opt_level        : O2
cast_model_type  : torch.float16
patch_torch_functions : False
keep_batchnorm_fp32 : True
master_weights   : True
loss_scale       : dynamic
```

```
Warning: multi_tensor_applier fused unscale kernel is unavailable, possibly
because apex was installed without --cuda_ext --cpp_ext. Using Python fallback.
Original ImportError was: ModuleNotFoundError("No module named 'amp_C'")
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
```

```
    warnings.warn("An input tensor was not cuda.")
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
2049it [00:07, 267.59it/s]
load_f1 = 0.8319559228650137
real_f1 = 0.8260869565217391
733it [00:00, 3441.24it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
4398it [00:22, 194.83it/s]
```

## 4.2 DITTO TEXTUAL

### 4.2.1 Abt-Buy - Running the matcher

```
[ ]: # Step 1:
      # train the wdc model with the save_model flag on
      # it will produce two files *_dev.pt and *_test.pt which are the
      # best checkpoints on the dev set and the test set.
```

```

!CUDA_VISIBLE_DEVICES=0 python train_ditto.py \
  --task Textual/Abt-Buy \
  --batch_size 32 \
  --max_len 256 \
  --lr 3e-5 \
  --n_epochs 15 \
  --finetuning \
  --lm roberta \
  --fp16 \
  --save_model \
  --dk general \
  --dk product \
  --summarize \
  --da entry_swap \
  --da attr_del

# Step 2: pick a checkpoint, rename it to wdc_all_small.pt
!mv *_dev.pt checkpoints/Abt-Buy.pt

# Step 3: run the matcher
!CUDA_VISIBLE_DEVICES=0 python matcher.py \
  --task Textual/Abt-Buy \
  --input_path input/input_small.jsonl \
  --output_path output/output_small.jsonl \
  --lm roberta \
  --use_gpu \
  --fp16 \
  --checkpoint_path checkpoints/

```

Selected optimization level 02: FP16 training with FP32 batchnorm and FP32 master weights.

Defaults for this optimization level are:

```

enabled          : True
opt_level        : 02
cast_model_type  : torch.float16
patch_torch_functions : False
keep_batchnorm_fp32 : True
master_weights   : True
loss_scale       : dynamic

```

Processing user overrides (additional kwargs that are not None)...

After processing overrides, optimization options are:

```

enabled          : True
opt_level        : 02
cast_model_type  : torch.float16
patch_torch_functions : False

```



```

keep_batchnorm_fp32      : True
master_weights           : True
loss_scale                : dynamic
Warning: multi_tensor_applier fused unscale kernel is unavailable, possibly
because apex was installed without --cuda_ext --cpp_ext. Using Python fallback.
Original ImportError was: ModuleNotFoundError("No module named 'amp_C'")
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 32768.0
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:136:
UserWarning: Detected call of `lr_scheduler.step()` before `optimizer.step()`.
In PyTorch 1.1.0 and later, you should call them in the opposite order:
`optimizer.step()` before `lr_scheduler.step()`. Failure to do this will result
in PyTorch skipping the first value of the learning rate schedule. See more
details at https://pytorch.org/docs/stable/optim.html#how-to-adjust-learning-
rate
    "https://pytorch.org/docs/stable/optim.html#how-to-adjust-learning-rate",
UserWarning)
step: 0, loss: 0.48074132204055786
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 10, loss: 0.43326252698898315
step: 20, loss: 0.28317591547966003
step: 30, loss: 0.37013277411460876
step: 40, loss: 0.5071871876716614
step: 50, loss: 0.3003700077533722
step: 60, loss: 0.529095470905304
step: 70, loss: 0.3028623163700104
step: 80, loss: 0.2430952787399292
step: 90, loss: 0.26699134707450867
step: 100, loss: 0.12653286755084991
step: 110, loss: 0.4083147644996643
step: 120, loss: 0.3421889841556549
step: 130, loss: 0.31949758529663086
step: 140, loss: 0.3780047297477722
step: 150, loss: 0.31652912497520447
step: 160, loss: 0.42224112153053284
step: 170, loss: 0.3222968578338623
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
    warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 1: dev_f1=0.1955252918287938, f1=0.2007060010085729,
best_f1=0.2007060010085729
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:

```

```

An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.3212262690067291
step: 10, loss: 0.4593675136566162
step: 20, loss: 0.3130393326282501
step: 30, loss: 0.32043519616127014
step: 40, loss: 0.09122481942176819
step: 50, loss: 0.42692574858665466
step: 60, loss: 0.18935143947601318
step: 70, loss: 0.4955059289932251
step: 80, loss: 0.2408725768327713
step: 90, loss: 0.24837106466293335
step: 100, loss: 0.5260030031204224
step: 110, loss: 0.2662883400917053
step: 120, loss: 0.25172358751296997
step: 130, loss: 0.5489847660064697
step: 140, loss: 0.5067744255065918
step: 150, loss: 0.44404250383377075
step: 160, loss: 0.4455161988735199
step: 170, loss: 0.38234445452690125
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
epoch 2: dev_f1=0.1941564561734213, f1=0.1941564561734213,
best_f1=0.2007060010085729
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.5930290222167969
step: 10, loss: 0.3385988175868988
step: 20, loss: 0.22250832617282867
step: 30, loss: 0.2551211416721344
step: 40, loss: 0.38785019516944885
step: 50, loss: 0.5534290671348572
step: 60, loss: 0.30147385597229004
step: 70, loss: 0.24351878464221954
step: 80, loss: 0.3729287087917328
step: 90, loss: 0.4977056384086609
step: 100, loss: 0.3330325782299042
step: 110, loss: 0.1782858967781067
step: 120, loss: 0.6390545964241028
step: 130, loss: 0.5272359251976013
step: 140, loss: 0.4400160014629364
step: 150, loss: 0.19316202402114868
step: 160, loss: 0.16675668954849243
step: 170, loss: 0.31894510984420776
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.

```

```

    warnings.warn("An input tensor was not cuda.")
epoch 3: dev_f1=0.1949834358731661, f1=0.19400855920114124,
best_f1=0.2007060010085729
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.3822176158428192
step: 10, loss: 0.5273535847663879
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 8192.0
step: 20, loss: 0.2343563586473465
step: 30, loss: 0.4509004056453705
step: 40, loss: 0.25385430455207825
step: 50, loss: 0.3955070674419403
step: 60, loss: 0.7383049130439758
step: 70, loss: 0.3291999101638794
step: 80, loss: 0.45313024520874023
step: 90, loss: 0.3016875088214874
step: 100, loss: 0.3824690282344818
step: 110, loss: 0.46354565024375916
step: 120, loss: 0.49559706449508667
step: 130, loss: 0.32221734523773193
step: 140, loss: 0.24901294708251953
step: 150, loss: 0.7329615354537964
step: 160, loss: 0.12890329957008362
step: 170, loss: 0.25039342045783997
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 4: dev_f1=0.1941564561734213, f1=0.1941564561734213,
best_f1=0.2007060010085729
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.3811734914779663
step: 10, loss: 0.3194943964481354
step: 20, loss: 0.31113094091415405
step: 30, loss: 0.31968986988067627
step: 40, loss: 0.24785156548023224
step: 50, loss: 0.25620999932289124
step: 60, loss: 0.3136289119720459
step: 70, loss: 0.31677955389022827
step: 80, loss: 0.07970760762691498
step: 90, loss: 0.6209622025489807
step: 100, loss: 0.2537771463394165
step: 110, loss: 0.25516343116760254
step: 120, loss: 0.13063152134418488
step: 130, loss: 0.24325546622276306
step: 140, loss: 0.16107885539531708

```

```

step: 150, loss: 0.32003939151763916
step: 160, loss: 0.2403990477323532
step: 170, loss: 0.376262903213501
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 5: dev_f1=0.19542421353670164, f1=0.19549592716818398,
best_f1=0.2007060010085729
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.31524932384490967
step: 10, loss: 0.10836049914360046
step: 20, loss: 0.3799661695957184
step: 30, loss: 0.25355082750320435
step: 40, loss: 0.428436815738678
step: 50, loss: 0.2537284791469574
step: 60, loss: 0.37419506907463074
step: 70, loss: 0.3936058282852173
step: 80, loss: 0.20876803994178772
step: 90, loss: 0.30620077252388
step: 100, loss: 0.15136952698230743
step: 110, loss: 0.4004453420639038
step: 120, loss: 0.4212333559989929
step: 130, loss: 0.49293017387390137
step: 140, loss: 0.23254305124282837
step: 150, loss: 0.1752634048461914
step: 160, loss: 0.3063866198062897
step: 170, loss: 0.2457888126373291
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
    warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 6: dev_f1=0.209009009009009, f1=0.2265415549597855,
best_f1=0.2265415549597855
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.23876133561134338
step: 10, loss: 0.46717581152915955
step: 20, loss: 0.4717448353767395
step: 30, loss: 0.5127902626991272
step: 40, loss: 0.13287591934204102
step: 50, loss: 0.4159441590309143
step: 60, loss: 0.5471561551094055

```

```

step: 70, loss: 0.3695981204509735
step: 80, loss: 0.18710307776927948
step: 90, loss: 0.45860105752944946
step: 100, loss: 0.28380581736564636
step: 110, loss: 0.39087486267089844
step: 120, loss: 0.35233238339424133
step: 130, loss: 0.1865856796503067
step: 140, loss: 0.08784130960702896
step: 150, loss: 0.24462038278579712
step: 160, loss: 0.22051827609539032
step: 170, loss: 0.32074543833732605
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
  warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 7: dev_f1=0.6128364389233955, f1=0.601980198019802,
best_f1=0.601980198019802
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.3669475317001343
step: 10, loss: 0.3614099621772766
step: 20, loss: 0.1599818617105484
step: 30, loss: 0.15306171774864197
step: 40, loss: 0.1169828325510025
step: 50, loss: 0.0933879092335701
step: 60, loss: 0.30468612909317017
step: 70, loss: 0.11185882985591888
step: 80, loss: 0.11284645646810532
step: 90, loss: 0.17966625094413757
step: 100, loss: 0.0629194900393486
step: 110, loss: 0.2442605048418045
step: 120, loss: 0.31968146562576294
step: 130, loss: 0.0648743063211441
step: 140, loss: 0.12067398428916931
step: 150, loss: 0.02338379994034767
step: 160, loss: 0.05561099946498871
step: 170, loss: 0.18981894850730896
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
  warnings.warn(SAVE_STATE_WARNING, UserWarning)

```

```

epoch 8: dev_f1=0.7941176470588236, f1=0.8123515439429927,
best_f1=0.8123515439429927
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.3320409059524536
step: 10, loss: 0.14315231144428253
step: 20, loss: 0.10705137997865677
step: 30, loss: 0.1995948851108551
step: 40, loss: 0.25174516439437866
step: 50, loss: 0.02260177955031395
step: 60, loss: 0.057503990828990936
step: 70, loss: 0.024834366515278816
step: 80, loss: 0.08147601783275604
step: 90, loss: 0.13496404886245728
step: 100, loss: 0.12880975008010864
step: 110, loss: 0.14305861294269562
step: 120, loss: 0.10285967588424683
step: 130, loss: 0.08777198195457458
step: 140, loss: 0.06739120930433273
step: 150, loss: 0.34180474281311035
step: 160, loss: 0.015917977318167686
step: 170, loss: 0.1878332942724228
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
    warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 9: dev_f1=0.8329297820823245, f1=0.8558139534883721,
best_f1=0.8558139534883721
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.06452548503875732
step: 10, loss: 0.027984503656625748
step: 20, loss: 0.18690703809261322
step: 30, loss: 0.08238504827022552
step: 40, loss: 0.05038336664438248
step: 50, loss: 0.11636849492788315
step: 60, loss: 0.14455854892730713
step: 70, loss: 0.09947232156991959
step: 80, loss: 0.03622153773903847
step: 90, loss: 0.027318844571709633
step: 100, loss: 0.09743087738752365
step: 110, loss: 0.03719521686434746
step: 120, loss: 0.043341491371393204

```

```

step: 130, loss: 0.017820267006754875
step: 140, loss: 0.012555791065096855
step: 150, loss: 0.003578639356419444
step: 160, loss: 0.0064825029112398624
step: 170, loss: 0.14530403912067413
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
    warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 10: dev_f1=0.836272040302267, f1=0.8674698795180722,
best_f1=0.8674698795180722
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.08584115654230118
step: 10, loss: 0.020658276975154877
step: 20, loss: 0.022005824372172356
step: 30, loss: 0.055643677711486816
step: 40, loss: 0.006887335330247879
step: 50, loss: 0.18688447773456573
step: 60, loss: 0.2149103730916977
step: 70, loss: 0.014452117495238781
step: 80, loss: 0.03642712160944939
step: 90, loss: 0.02133745513856411
step: 100, loss: 0.05225011333823204
step: 110, loss: 0.03186111897230148
step: 120, loss: 0.03985797241330147
step: 130, loss: 0.026647573336958885
step: 140, loss: 0.05719887465238571
step: 150, loss: 0.00331934099085629
step: 160, loss: 0.09764960408210754
step: 170, loss: 0.11709108203649521
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
    warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 11: dev_f1=0.8606356968215159, f1=0.8915094339622642,
best_f1=0.8915094339622642
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.007520296610891819

```

```

step: 10, loss: 0.03330305218696594
step: 20, loss: 0.011044727638363838
step: 30, loss: 0.16303086280822754
step: 40, loss: 0.006220315583050251
step: 50, loss: 0.012635422870516777
step: 60, loss: 0.041781771928071976
step: 70, loss: 0.0028826631605625153
step: 80, loss: 0.0025279123801738024
step: 90, loss: 0.13253891468048096
step: 100, loss: 0.02026541344821453
step: 110, loss: 0.11761848628520966
step: 120, loss: 0.04622986540198326
step: 130, loss: 0.04956376925110817
step: 140, loss: 0.004138560499995947
step: 150, loss: 0.02362770587205887
step: 160, loss: 0.0813630148768425
step: 170, loss: 0.2755770981311798
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
epoch 12: dev_f1=0.8491879350348029, f1=0.874439461883408,
best_f1=0.8915094339622642
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0132129592821002
step: 10, loss: 0.005496396217495203
step: 20, loss: 0.01040680706501007
step: 30, loss: 0.03189444914460182
step: 40, loss: 0.01966434344649315
step: 50, loss: 0.1325387805700302
step: 60, loss: 0.06525904685258865
step: 70, loss: 0.1945759654045105
step: 80, loss: 0.00770606892183423
step: 90, loss: 0.022335771471261978
step: 100, loss: 0.04954719543457031
step: 110, loss: 0.005934963934123516
step: 120, loss: 0.01502192672342062
step: 130, loss: 0.004507388919591904
step: 140, loss: 0.034749194979667664
step: 150, loss: 0.015971196815371513
step: 160, loss: 0.005732514429837465
step: 170, loss: 0.004116678610444069
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
epoch 13: dev_f1=0.8524590163934426, f1=0.8727272727272728,
best_f1=0.8915094339622642

```



```

/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.004056384786963463
step: 10, loss: 0.0014258126029744744
step: 20, loss: 0.04477419704198837
step: 30, loss: 0.04062935709953308
step: 40, loss: 0.10801924020051956
step: 50, loss: 0.0013226609444245696
step: 60, loss: 0.01130054984241724
step: 70, loss: 0.00952674075961113
step: 80, loss: 0.014219122007489204
step: 90, loss: 0.000984029844403267
step: 100, loss: 0.010857627727091312
step: 110, loss: 0.004258237779140472
step: 120, loss: 0.11327812820672989
step: 130, loss: 0.16197098791599274
step: 140, loss: 0.09570369124412537
step: 150, loss: 0.040029995143413544
step: 160, loss: 0.01171029731631279
step: 170, loss: 0.03545937314629555
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 14: dev_f1=0.8528678304239402, f1=0.8764568764568764,
best_f1=0.8915094339622642
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0013147584395483136
step: 10, loss: 0.02274971269071102
step: 20, loss: 0.00494047487154603
step: 30, loss: 0.004700468387454748
step: 40, loss: 0.0034243683330714703
step: 50, loss: 0.005003156140446663
step: 60, loss: 0.008709670975804329
step: 70, loss: 0.097342848777771
step: 80, loss: 0.020830141380429268
step: 90, loss: 0.0053145973943173885
step: 100, loss: 0.049388233572244644
step: 110, loss: 0.00127299758605659
step: 120, loss: 0.004265972413122654
step: 130, loss: 0.002257556188851595
step: 140, loss: 0.011978501453995705
step: 150, loss: 0.0011513022473081946
step: 160, loss: 0.004632806405425072
step: 170, loss: 0.021547917276620865
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:

```

```

An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
epoch 15: dev_f1=0.8543209876543211, f1=0.8842592592592592,
best_f1=0.8915094339622642
mv: cannot stat '*_dev.pt': No such file or directory
Selected optimization level O2:  FP16 training with FP32 batchnorm and FP32
master weights.

Defaults for this optimization level are:
enabled                : True
opt_level              : O2
cast_model_type        : torch.float16
patch_torch_functions  : False
keep_batchnorm_fp32    : True
master_weights         : True
loss_scale             : dynamic
Processing user overrides (additional kwargs that are not None)...
After processing overrides, optimization options are:
enabled                : True
opt_level              : O2
cast_model_type        : torch.float16
patch_torch_functions  : False
keep_batchnorm_fp32    : True
master_weights         : True
loss_scale             : dynamic
Warning:  multi_tensor_applier fused unscale kernel is unavailable, possibly
because apex was installed without --cuda_ext --cpp_ext. Using Python fallback.
Original ImportError was: ModuleNotFoundError("No module named 'amp_C'")
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
1916it [00:05, 325.90it/s]
load_f1 = 0.43704775687409547
real_f1 = 0.39838492597577385
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
4398it [00:21, 200.66it/s]

```

## 4.3 DITTO DIRTY

### 4.3.1 DBLP-ACM - Running the matcher

```
[ ]: # Step 1:
# train the wdc model with the save_model flag on
# it will produce two files *_dev.pt and *_test.pt which are the
# best checkpoints on the dev set and the test set.
!CUDA_VISIBLE_DEVICES=0 python train_ditto.py \
  --task Dirty/DBLP-ACM \
  --batch_size 32 \
  --max_len 256 \
  --lr 3e-5 \
  --n_epochs 15 \
  --finetuning \
  --lm roberta \
  --fp16 \
  --save_model \
  --dk general \
  --dk product \
  --summarize \
  --da entry_swap \
  --da attr_del

# Step 2: pick a checkpoint, rename it to wdc_all_small.pt
!mv *_dev.pt checkpoints/DBLP-ACM.pt

# Step 3: run the matcher
!CUDA_VISIBLE_DEVICES=0 python matcher.py \
  --task Dirty/DBLP-ACM \
  --input_path input/input_small.jsonl \
  --output_path output/output_small.jsonl \
  --lm roberta \
  --use_gpu \
  --fp16 \
  --checkpoint_path checkpoints/
```

Selected optimization level 02: FP16 training with FP32 batchnorm and FP32 master weights.

Defaults for this optimization level are:

enabled	: True
opt_level	: 02
cast_model_type	: torch.float16
patch_torch_functions	: False
keep_batchnorm_fp32	: True

```

master_weights      : True
loss_scale          : dynamic
Processing user overrides (additional kwargs that are not None)...
After processing overrides, optimization options are:
enabled             : True
opt_level           : 02
cast_model_type     : torch.float16
patch_torch_functions : False
keep_batchnorm_fp32 : True
master_weights      : True
loss_scale          : dynamic
Warning: multi_tensor_applier fused unscale kernel is unavailable, possibly
because apex was installed without --cuda_ext --cpp_ext. Using Python fallback.
Original ImportError was: ModuleNotFoundError("No module named 'amp_C'")
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.5604655146598816
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 32768.0
step: 10, loss: 0.44155898690223694
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 8192.0
step: 20, loss: 0.4850066006183624
step: 30, loss: 0.26345282793045044
step: 40, loss: 0.26380816102027893
step: 50, loss: 0.2208133041858673
step: 60, loss: 0.40293383598327637
step: 70, loss: 0.17102986574172974
step: 80, loss: 0.045636728405952454
step: 90, loss: 0.1598510891199112
step: 100, loss: 0.05628705769777298
step: 110, loss: 0.1842304915189743
step: 120, loss: 0.12066038697957993
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 4096.0
step: 130, loss: 0.015418944880366325
step: 140, loss: 0.0312710776925087
step: 150, loss: 0.20664604008197784
step: 160, loss: 0.023088136687874794
step: 170, loss: 0.11032677441835403
step: 180, loss: 0.19614581763744354
step: 190, loss: 0.05103830620646477
step: 200, loss: 0.024302611127495766
step: 210, loss: 0.01050816010683775
step: 220, loss: 0.04749374836683273
step: 230, loss: 0.006191050633788109
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")

```

```

/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
    warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 1: dev_f1=0.9608062709966406, f1=0.9593679458239278,
best_f1=0.9593679458239278
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.009693896397948265
step: 10, loss: 0.15550681948661804
step: 20, loss: 0.025387465953826904
step: 30, loss: 0.03458377346396446
step: 40, loss: 0.07833989709615707
step: 50, loss: 0.004456744994968176
step: 60, loss: 0.010629555210471153
step: 70, loss: 0.030499916523694992
step: 80, loss: 0.008840023539960384
step: 90, loss: 0.0014728512614965439
step: 100, loss: 0.0022210206370800734
step: 110, loss: 0.04077187553048134
step: 120, loss: 0.01728583313524723
step: 130, loss: 0.0563645213842392
step: 140, loss: 0.002719819312915206
step: 150, loss: 0.1795995533466339
step: 160, loss: 0.0676671713590622
step: 170, loss: 0.007313624955713749
step: 180, loss: 0.013985726051032543
step: 190, loss: 0.0020147531758993864
step: 200, loss: 0.027372486889362335
step: 210, loss: 0.01760026253759861
step: 220, loss: 0.008153047412633896
step: 230, loss: 0.0024296455085277557
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
    warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 2: dev_f1=0.9643652561247216, f1=0.9697648376259798,
best_f1=0.9697648376259798
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0051164510659873486
step: 10, loss: 0.004622053820639849
step: 20, loss: 0.0031387144699692726

```

```

step: 30, loss: 0.012064206413924694
step: 40, loss: 0.1493174433708191
step: 50, loss: 0.07066220045089722
step: 60, loss: 0.014064355753362179
step: 70, loss: 0.006792484782636166
step: 80, loss: 0.04231393337249756
step: 90, loss: 0.009074791334569454
step: 100, loss: 0.010554932989180088
step: 110, loss: 0.002141887554898858
step: 120, loss: 0.0004425575607456267
step: 130, loss: 0.0036259854678064585
step: 140, loss: 0.010037275031208992
step: 150, loss: 0.13071461021900177
step: 160, loss: 0.00599752739071846
step: 170, loss: 0.0017001694068312645
step: 180, loss: 0.012306639924645424
step: 190, loss: 0.02279580570757389
step: 200, loss: 0.005683114752173424
step: 210, loss: 0.003488339949399233
step: 220, loss: 0.027084803208708763
step: 230, loss: 0.021402504295110703
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
    warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 3: dev_f1=0.9741282339707535, f1=0.9628796400449944,
best_f1=0.9628796400449944
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.016740402206778526
step: 10, loss: 0.005479384679347277
step: 20, loss: 0.0021663277875632048
step: 30, loss: 0.004580992739647627
step: 40, loss: 0.13021603226661682
step: 50, loss: 0.04490169137716293
step: 60, loss: 0.013526789844036102
step: 70, loss: 0.009316968731582165
step: 80, loss: 0.1419290006160736
step: 90, loss: 0.06996852159500122
step: 100, loss: 0.013005489483475685
step: 110, loss: 0.010195587761700153
step: 120, loss: 0.024508262053132057
step: 130, loss: 0.007191097363829613
step: 140, loss: 0.002849624492228031

```

```

step: 150, loss: 0.004346015397459269
step: 160, loss: 0.0031491487752646208
step: 170, loss: 0.004667105618864298
step: 180, loss: 0.04773574322462082
step: 190, loss: 0.0014224312035366893
step: 200, loss: 0.07917966693639755
step: 210, loss: 0.009708140976727009
step: 220, loss: 0.0008208858780562878
step: 230, loss: 0.0033925131428986788
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
  warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 4: dev_f1=0.9744160177975528, f1=0.9754464285714286,
best_f1=0.9754464285714286
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00048056856030598283
step: 10, loss: 0.01473875343799591
step: 20, loss: 0.09613033384084702
step: 30, loss: 0.0005630368832498789
step: 40, loss: 0.0007702155853621662
step: 50, loss: 0.0010895704617723823
step: 60, loss: 0.013549638912081718
step: 70, loss: 0.002605857327580452
step: 80, loss: 0.05041225254535675
step: 90, loss: 0.24920654296875
step: 100, loss: 0.0024762549437582493
step: 110, loss: 0.007226001005619764
step: 120, loss: 0.0039813159964978695
step: 130, loss: 0.0004763590113725513
step: 140, loss: 0.005109606310725212
step: 150, loss: 0.0035994411446154118
step: 160, loss: 0.0004874104051850736
step: 170, loss: 0.0011234278790652752
step: 180, loss: 0.014381607063114643
step: 190, loss: 0.29255449771881104
step: 200, loss: 0.036056965589523315
step: 210, loss: 0.02426859177649021
step: 220, loss: 0.002011216012760997
step: 230, loss: 0.004103220533579588
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")

```

```

/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
  warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 5: dev_f1=0.9797752808988766, f1=0.9718785151856018,
best_f1=0.9718785151856018
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0012047523632645607
step: 10, loss: 0.0032427161931991577
step: 20, loss: 0.001660690177232027
step: 30, loss: 0.001039917697198689
step: 40, loss: 0.0005161739536561072
step: 50, loss: 0.00191054109018296
step: 60, loss: 0.0015539351152256131
step: 70, loss: 0.008807758800685406
step: 80, loss: 0.010915341787040234
step: 90, loss: 0.011715386994183064
step: 100, loss: 0.0016041084891185164
step: 110, loss: 0.030889974907040596
step: 120, loss: 0.0017331859562546015
step: 130, loss: 0.001093982020393014
step: 140, loss: 0.0009726134012453258
step: 150, loss: 0.001231342670507729
step: 160, loss: 0.0058329240418970585
step: 170, loss: 0.008969184011220932
step: 180, loss: 0.04828621819615364
step: 190, loss: 0.0009134724386967719
step: 200, loss: 0.04693649336695671
step: 210, loss: 0.0007276318501681089
step: 220, loss: 0.016849059611558914
step: 230, loss: 0.0018192788120359182
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
epoch 6: dev_f1=0.9755555555555556, f1=0.9688888888888889,
best_f1=0.9718785151856018
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0023993304930627346
step: 10, loss: 0.0015079863369464874
step: 20, loss: 0.0009149947436526418
step: 30, loss: 0.0005429945886135101
step: 40, loss: 0.007433871738612652
step: 50, loss: 0.0029576525557786226
step: 60, loss: 0.0018483547028154135

```



```

step: 70, loss: 0.0009479836444370449
step: 80, loss: 0.0007487251423299313
step: 90, loss: 0.0005440957029350102
step: 100, loss: 0.0006121896440163255
step: 110, loss: 0.0003760808613151312
step: 120, loss: 0.001261114259250462
step: 130, loss: 0.0015721531817689538
step: 140, loss: 0.00038148200837895274
step: 150, loss: 0.12506109476089478
step: 160, loss: 0.000781141803599894
step: 170, loss: 0.00493573909625411
step: 180, loss: 0.0038187357131391764
step: 190, loss: 0.01914774253964424
step: 200, loss: 0.10593722015619278
step: 210, loss: 0.0035329051315784454
step: 220, loss: 0.000793333922047168
step: 230, loss: 0.0021171383559703827
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
  warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 7: dev_f1=0.9799554565701558, f1=0.9776286353467561,
best_f1=0.9776286353467561
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.002468792023137212
step: 10, loss: 0.0119217773899436
step: 20, loss: 0.005613015033304691
step: 30, loss: 0.001566971535794437
step: 40, loss: 0.0015715170884504914
step: 50, loss: 0.0016527536790817976
step: 60, loss: 0.0010642512934282422
step: 70, loss: 0.00018693438323680311
step: 80, loss: 0.03755110874772072
step: 90, loss: 0.042740464210510254
step: 100, loss: 0.0036558336578309536
step: 110, loss: 0.00607500271871686
step: 120, loss: 0.004354516509920359
step: 130, loss: 0.006866576615720987
step: 140, loss: 0.0008081796113401651
step: 150, loss: 0.18942570686340332
step: 160, loss: 0.0016437058802694082
step: 170, loss: 0.009529653936624527
step: 180, loss: 0.0009480990702286363

```

```

step: 190, loss: 0.017317384481430054
step: 200, loss: 0.031700558960437775
step: 210, loss: 0.003713486948981881
step: 220, loss: 0.0010072439908981323
step: 230, loss: 0.0016692206263542175
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 8: dev_f1=0.9777777777777777, f1=0.9666666666666666,
best_f1=0.9776286353467561
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.001971782185137272
step: 10, loss: 0.0011228213552385569
step: 20, loss: 0.0008160012657754123
step: 30, loss: 0.0008894751663319767
step: 40, loss: 0.00200673914514482
step: 50, loss: 0.001059988746419549
step: 60, loss: 0.00033747495035640895
step: 70, loss: 0.03786543011665344
step: 80, loss: 0.00018331597675569355
step: 90, loss: 0.03668602555990219
step: 100, loss: 0.0017032879404723644
step: 110, loss: 0.0005238184239715338
step: 120, loss: 0.031239891424775124
step: 130, loss: 0.05162188038229942
step: 140, loss: 0.010473598726093769
step: 150, loss: 0.001355274929665029
step: 160, loss: 0.02082233503460884
step: 170, loss: 0.02269759029150009
step: 180, loss: 0.0019512329017743468
step: 190, loss: 0.00044229914783500135
step: 200, loss: 0.0007889981498010457
step: 210, loss: 0.06331387907266617
step: 220, loss: 0.0011567999608814716
step: 230, loss: 0.0005014659836888313
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 9: dev_f1=0.977728285077951, f1=0.9755555555555556,
best_f1=0.9776286353467561
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0010763490572571754
step: 10, loss: 0.0006406124448403716
step: 20, loss: 0.0016636732034385204

```

```

step: 30, loss: 0.00029212009394541383
step: 40, loss: 0.0023193215020000935
step: 50, loss: 0.0014880355447530746
step: 60, loss: 0.0009611163404770195
step: 70, loss: 0.04547448083758354
step: 80, loss: 0.00026804531808011234
step: 90, loss: 0.00028773644589819014
step: 100, loss: 0.00026521197287365794
step: 110, loss: 0.009015422314405441
step: 120, loss: 0.0012376190861687064
step: 130, loss: 0.017474940046668053
step: 140, loss: 0.0003120263572782278
step: 150, loss: 0.0009072759421542287
step: 160, loss: 6.367938476614654e-05
step: 170, loss: 0.00011217990686418489
step: 180, loss: 0.00038300309097394347
step: 190, loss: 0.0007597979856655002
step: 200, loss: 0.0010470778215676546
step: 210, loss: 0.0009111303370445967
step: 220, loss: 0.0020153014920651913
step: 230, loss: 0.00030996656278148293
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 10: dev_f1=0.9785794813979707, f1=0.971815107102593,
best_f1=0.9776286353467561
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0002462579868733883
step: 10, loss: 0.00027131193201057613
step: 20, loss: 0.0003099836176261306
step: 30, loss: 0.0004466791287995875
step: 40, loss: 4.4989617890678346e-05
step: 50, loss: 0.00018521107267588377
step: 60, loss: 0.012760237790644169
step: 70, loss: 0.010893789120018482
step: 80, loss: 0.0026917948853224516
step: 90, loss: 0.0287100151181221
step: 100, loss: 0.00030156836146488786
step: 110, loss: 0.0003196708858013153
step: 120, loss: 0.00021249879500828683
step: 130, loss: 0.00015431226347573102
step: 140, loss: 0.0003607547259889543
step: 150, loss: 0.0003403624868951738
step: 160, loss: 0.036112550646066666
step: 170, loss: 0.022407291457057
step: 180, loss: 0.0009385033044964075

```

```

step: 190, loss: 0.0003512483963277191
step: 200, loss: 0.003119303844869137
step: 210, loss: 0.00029050151351839304
step: 220, loss: 0.0025608253199607134
step: 230, loss: 0.0001750810770317912
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
    warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 11: dev_f1=0.9831271091113611, f1=0.976271186440678,
best_f1=0.976271186440678
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0006615154561586678
step: 10, loss: 0.00017569995543453842
step: 20, loss: 0.012513450346887112
step: 30, loss: 0.03940140828490257
step: 40, loss: 0.0004978753859177232
step: 50, loss: 0.012510428205132484
step: 60, loss: 0.0007570068701170385
step: 70, loss: 0.00014898442896082997
step: 80, loss: 8.270146645372733e-05
step: 90, loss: 0.016144314780831337
step: 100, loss: 0.00011324584193062037
step: 110, loss: 0.00013258906255941838
step: 120, loss: 0.0007562178070656955
step: 130, loss: 0.000272261822829023
step: 140, loss: 0.0003367222670931369
step: 150, loss: 0.001760050654411316
step: 160, loss: 0.00012175177835160866
step: 170, loss: 0.003616618923842907
step: 180, loss: 0.00013745531032327563
step: 190, loss: 0.0016638691304251552
step: 200, loss: 0.00012350019824225456
step: 210, loss: 0.016444755718111992
step: 220, loss: 0.008045397698879242
step: 230, loss: 0.004238884896039963
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 12: dev_f1=0.9810479375696767, f1=0.972129319955407,
best_f1=0.976271186440678
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.

```

```

    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0001444649533368647
step: 10, loss: 0.0047078123316168785
step: 20, loss: 0.0005672434344887733
step: 30, loss: 0.0022584276739507914
step: 40, loss: 0.0012125704670324922
step: 50, loss: 0.0011656106216832995
step: 60, loss: 0.00013803088222630322
step: 70, loss: 0.00022432512196246535
step: 80, loss: 0.00011989012273261324
step: 90, loss: 0.00010621463297866285
step: 100, loss: 0.0011071875924244523
step: 110, loss: 0.00018717649800237268
step: 120, loss: 0.00013497543113771826
step: 130, loss: 0.00016013970889616758
step: 140, loss: 0.0027229334227740765
step: 150, loss: 0.00035374151775613427
step: 160, loss: 0.002951872767880559
step: 170, loss: 0.00011960035044467077
step: 180, loss: 0.0343315452337265
step: 190, loss: 0.00013432191917672753
step: 200, loss: 6.651553849224001e-05
step: 210, loss: 0.000402487232349813
step: 220, loss: 0.0001748968061292544
step: 230, loss: 0.0010507258120924234
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 13: dev_f1=0.98, f1=0.9723145071982282, best_f1=0.976271186440678
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00012584452633745968
step: 10, loss: 0.00014289742102846503
step: 20, loss: 0.006953591015189886
step: 30, loss: 0.0009037387790158391
step: 40, loss: 0.0001177926897071302
step: 50, loss: 6.705065607093275e-05
step: 60, loss: 9.418915578862652e-05
step: 70, loss: 0.00013728455814998597
step: 80, loss: 8.340519707417116e-05
step: 90, loss: 0.00014885349082760513
step: 100, loss: 0.00010853326239157468
step: 110, loss: 0.00012573698768392205
step: 120, loss: 3.7740923289675266e-05
step: 130, loss: 0.004748076666146517
step: 140, loss: 0.03180813789367676
step: 150, loss: 5.044598947279155e-05

```

```

step: 160, loss: 0.0001033717198879458
step: 170, loss: 0.000526965653989464
step: 180, loss: 8.663705375511199e-05
step: 190, loss: 0.02085316739976406
step: 200, loss: 0.00033143977634608746
step: 210, loss: 0.001096611493267119
step: 220, loss: 0.00010682098218239844
step: 230, loss: 0.00031477806624025106
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 14: dev_f1=0.9774774774774775, f1=0.9707865168539327,
best_f1=0.976271186440678
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.026981350034475327
step: 10, loss: 0.00010808306251419708
step: 20, loss: 0.0006557669839821756
step: 30, loss: 0.00011526630260050297
step: 40, loss: 5.94982884649653e-05
step: 50, loss: 0.000103351405414287
step: 60, loss: 0.04138602688908577
step: 70, loss: 0.00020965203293599188
step: 80, loss: 0.0001909898710437119
step: 90, loss: 9.68158055911772e-05
step: 100, loss: 5.2127972594462335e-05
step: 110, loss: 9.628752013668418e-05
step: 120, loss: 0.08749013394117355
step: 130, loss: 0.00012010705540888011
step: 140, loss: 0.008503561839461327
step: 150, loss: 0.000204018026124686
step: 160, loss: 0.007629952859133482
step: 170, loss: 5.3826432122150436e-05
step: 180, loss: 0.002227665623649955
step: 190, loss: 0.015906602144241333
step: 200, loss: 0.012998891994357109
step: 210, loss: 0.04744216799736023
step: 220, loss: 0.00010549715807428584
step: 230, loss: 0.0001509303692728281
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 15: dev_f1=0.9785310734463276, f1=0.971815107102593,
best_f1=0.976271186440678
mv: cannot stat '*_dev.pt': No such file or directory
Selected optimization level O2:  FP16 training with FP32 batchnorm and FP32
master weights.

```

```

Defaults for this optimization level are:
enabled                : True
opt_level              : 02
cast_model_type        : torch.float16
patch_torch_functions  : False
keep_batchnorm_fp32    : True
master_weights         : True
loss_scale             : dynamic
Processing user overrides (additional kwargs that are not None)...
After processing overrides, optimization options are:
enabled                : True
opt_level              : 02
cast_model_type        : torch.float16
patch_torch_functions  : False
keep_batchnorm_fp32    : True
master_weights         : True
loss_scale             : dynamic
Warning: multi_tensor_applier fused unscale kernel is unavailable, possibly
because apex was installed without --cuda_ext --cpp_ext. Using Python fallback.
Original ImportError was: ModuleNotFoundError("No module named 'amp_C'")
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
2473it [00:11, 214.57it/s]
load_f1 = 0.9776286353467561
real_f1 = 0.9776785714285714
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
4398it [00:22, 196.60it/s]

```

### 4.3.2 DBLP-GoogleScholar - Running the matcher

```

[ ]: # Step 1:
# train the wdc model with the save_model flag on
# it will produce two files *_dev.pt and *_test.pt which are the
# best checkpoints on the dev set and the test set.
!CUDA_VISIBLE_DEVICES=0 python train_ditto.py \
--task Dirty/DBLP-GoogleScholar \
--batch_size 32 \
--max_len 256 \
--lr 3e-5 \

```

```

--n_epochs 15 \
--finetuning \
--lm roberta \
--fp16 \
--save_model \
--dk general \
--dk product \
--summarize \
--da entry_swap \
--da attr_del

```

*# Step 2: pick a checkpoint, rename it to wdc\_all\_small.pt*

```
!mv *_dev.pt checkpoints/DBLP-GoogleScholar/.pt
```

*# Step 3: run the matcher*

```

!CUDA_VISIBLE_DEVICES=0 python matcher.py \
--task Dirty/DBLP-GoogleScholar \
--input_path input/input_small.jsonl \
--output_path output/output_small.jsonl \
--lm roberta \
--use_gpu \
--fp16 \
--checkpoint_path checkpoints/

```

Downloading: 100% 481/481 [00:00<00:00, 362kB/s]

Downloading: 100% 899k/899k [00:01<00:00, 809kB/s]

Downloading: 100% 456k/456k [00:01<00:00, 401kB/s]

Downloading: 100% 501M/501M [00:08<00:00, 61.8MB/s]

Selected optimization level 02: FP16 training with FP32 batchnorm and FP32 master weights.

Defaults for this optimization level are:

```

enabled           : True
opt_level         : 02
cast_model_type   : torch.float16
patch_torch_functions : False
keep_batchnorm_fp32 : True
master_weights    : True
loss_scale        : dynamic

```

Processing user overrides (additional kwargs that are not None)...

After processing overrides, optimization options are:

```

enabled           : True
opt_level         : 02
cast_model_type   : torch.float16
patch_torch_functions : False
keep_batchnorm_fp32 : True

```



```

master_weights      : True
loss_scale          : dynamic
Warning: multi_tensor_applier fused unscale kernel is unavailable, possibly
because apex was installed without --cuda_ext --cpp_ext. Using Python fallback.
Original ImportError was: ModuleNotFoundError("No module named 'amp_C'")
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 32768.0
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:136:
UserWarning: Detected call of `lr_scheduler.step()` before `optimizer.step()`.
In PyTorch 1.1.0 and later, you should call them in the opposite order:
`optimizer.step()` before `lr_scheduler.step()`. Failure to do this will result
in PyTorch skipping the first value of the learning rate schedule. See more
details at https://pytorch.org/docs/stable/optim.html#how-to-adjust-learning-
rate
    "https://pytorch.org/docs/stable/optim.html#how-to-adjust-learning-rate",
UserWarning)
step: 0, loss: 0.7184592485427856
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 10, loss: 0.44172826409339905
step: 20, loss: 0.30677053332328796
step: 30, loss: 0.3128829598426819
step: 40, loss: 0.37262651324272156
step: 50, loss: 0.6099631190299988
step: 60, loss: 0.24366886913776398
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 8192.0
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 4096.0
step: 70, loss: 0.43374425172805786
step: 80, loss: 0.12683895230293274
step: 90, loss: 0.3747903108596802
step: 100, loss: 0.21929039061069489
step: 110, loss: 0.10437599569559097
step: 120, loss: 0.08709374070167542
step: 130, loss: 0.12041877955198288
step: 140, loss: 0.2806977927684784
step: 150, loss: 0.1329231858253479
step: 160, loss: 0.058210548013448715
step: 170, loss: 0.04193515703082085
step: 180, loss: 0.006843035575002432
step: 190, loss: 0.11514091491699219
step: 200, loss: 0.029814423993229866
step: 210, loss: 0.068008653819561
step: 220, loss: 0.021108662709593773
step: 230, loss: 0.21528102457523346
step: 240, loss: 0.00837776716798544
step: 250, loss: 0.06067739054560661
step: 260, loss: 0.12266325950622559

```

```

step: 270, loss: 0.4444204866886139
step: 280, loss: 0.035308435559272766
step: 290, loss: 0.1366531103849411
step: 300, loss: 0.03001370280981064
step: 310, loss: 0.04079974442720413
step: 320, loss: 0.0957999899983406
step: 330, loss: 0.08065049350261688
step: 340, loss: 0.48377665877342224
step: 350, loss: 0.16512036323547363
step: 360, loss: 0.01262720488011837
step: 370, loss: 0.07536574453115463
step: 380, loss: 0.1828906238079071
step: 390, loss: 0.051443323493003845
step: 400, loss: 0.08344889432191849
step: 410, loss: 0.28255271911621094
step: 420, loss: 0.024501748383045197
step: 430, loss: 0.0298654492944479
step: 440, loss: 0.02135683223605156
step: 450, loss: 0.2194194495677948
step: 460, loss: 0.049025699496269226
step: 470, loss: 0.03678993880748749
step: 480, loss: 0.21754193305969238
step: 490, loss: 0.09960325062274933
step: 500, loss: 0.021474627777934074
step: 510, loss: 0.028711147606372833
step: 520, loss: 0.07827781140804291
step: 530, loss: 0.01336915697902441
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
  warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 1: dev_f1=0.93790546802595, f1=0.9406392694063926,
best_f1=0.9406392694063926
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.06600295007228851
step: 10, loss: 0.05812999606132507
step: 20, loss: 0.01161210983991623
step: 30, loss: 0.07531194388866425
step: 40, loss: 0.06060177832841873
step: 50, loss: 0.06757470220327377
step: 60, loss: 0.024195201694965363
step: 70, loss: 0.0480007603764534
step: 80, loss: 0.07367745041847229

```

```
step: 90, loss: 0.006134992931038141
step: 100, loss: 0.13208001852035522
step: 110, loss: 0.00851086713373661
step: 120, loss: 0.035621386021375656
step: 130, loss: 0.007225413341075182
step: 140, loss: 0.066670261323452
step: 150, loss: 0.01685498282313347
step: 160, loss: 0.028809741139411926
step: 170, loss: 0.02547890692949295
step: 180, loss: 0.047808509320020676
step: 190, loss: 0.03756645321846008
step: 200, loss: 0.25765034556388855
step: 210, loss: 0.016896948218345642
step: 220, loss: 0.002308238297700882
step: 230, loss: 0.09256727248430252
step: 240, loss: 0.0756845697760582
step: 250, loss: 0.013674027286469936
step: 260, loss: 0.0608845092356205
step: 270, loss: 0.003687777789309621
step: 280, loss: 0.05104535073041916
step: 290, loss: 0.057491544634103775
step: 300, loss: 0.02967149391770363
step: 310, loss: 0.06434717774391174
step: 320, loss: 0.014127710834145546
step: 330, loss: 0.14563488960266113
step: 340, loss: 0.015762509778141975
step: 350, loss: 0.0005535827949643135
step: 360, loss: 0.10775962471961975
step: 370, loss: 0.006825475487858057
step: 380, loss: 0.10752834379673004
step: 390, loss: 0.004674466326832771
step: 400, loss: 0.04553304612636566
step: 410, loss: 0.03001563996076584
step: 420, loss: 0.012241825461387634
step: 430, loss: 0.1255439817905426
step: 440, loss: 0.005589818116277456
step: 450, loss: 0.017322614789009094
step: 460, loss: 0.039557941257953644
step: 470, loss: 0.03644019737839699
step: 480, loss: 0.0020316969603300095
step: 490, loss: 0.03135434165596962
step: 500, loss: 0.013352807611227036
step: 510, loss: 0.010999638587236404
step: 520, loss: 0.2536555230617523
step: 530, loss: 0.05290210247039795
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
```

```

/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
  warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 2: dev_f1=0.9528172240036646, f1=0.9451553930530164,
best_f1=0.9451553930530164
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.12052368372678757
step: 10, loss: 0.043969545513391495
step: 20, loss: 0.012336114421486855
step: 30, loss: 0.02633410505950451
step: 40, loss: 0.1619599312543869
step: 50, loss: 0.026583340018987656
step: 60, loss: 0.022637147456407547
step: 70, loss: 0.0035065473057329655
step: 80, loss: 0.1468943953514099
step: 90, loss: 0.02440583147108555
step: 100, loss: 0.027447707951068878
step: 110, loss: 0.0057309274561703205
step: 120, loss: 0.13546666502952576
step: 130, loss: 0.10416033118963242
step: 140, loss: 0.019514938816428185
step: 150, loss: 0.09389611333608627
step: 160, loss: 0.01405302993953228
step: 170, loss: 0.002869116608053446
step: 180, loss: 0.06135757267475128
step: 190, loss: 0.00352878263220191
step: 200, loss: 0.04519542679190636
step: 210, loss: 0.03555573150515556
step: 220, loss: 0.10244055092334747
step: 230, loss: 0.02893819659948349
step: 240, loss: 0.04152009263634682
step: 250, loss: 0.06372220069169998
step: 260, loss: 0.09018898755311966
step: 270, loss: 0.007348001003265381
step: 280, loss: 0.007833942770957947
step: 290, loss: 0.01965988613665104
step: 300, loss: 0.12017461657524109
step: 310, loss: 0.03758212924003601
step: 320, loss: 0.011467383243143559
step: 330, loss: 0.03821853548288345
step: 340, loss: 0.01901034638285637
step: 350, loss: 0.16083134710788727
step: 360, loss: 0.02060527540743351
step: 370, loss: 0.040400244295597076
step: 380, loss: 0.00969027727842331

```

```

step: 390, loss: 0.00981468427926302
step: 400, loss: 0.13159075379371643
step: 410, loss: 0.03597744181752205
step: 420, loss: 0.011342300102114677
step: 430, loss: 0.02257881872355938
step: 440, loss: 0.06924257427453995
step: 450, loss: 0.021320784464478493
step: 460, loss: 0.04494604095816612
step: 470, loss: 0.04666099324822426
step: 480, loss: 0.07070634514093399
step: 490, loss: 0.015943067148327827
step: 500, loss: 0.010513370856642723
step: 510, loss: 0.01799081079661846
step: 520, loss: 0.014503615908324718
step: 530, loss: 0.0029386940877884626
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
    warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 3: dev_f1=0.954060324825986, f1=0.9548206800186306,
best_f1=0.9548206800186306
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0064709861762821674
step: 10, loss: 0.007207596208900213
step: 20, loss: 0.04716787114739418
step: 30, loss: 0.10338205844163895
step: 40, loss: 0.005328549537807703
step: 50, loss: 0.031215209513902664
step: 60, loss: 0.011759867891669273
step: 70, loss: 0.003083360381424427
step: 80, loss: 0.13518092036247253
step: 90, loss: 0.12595981359481812
step: 100, loss: 0.014914611354470253
step: 110, loss: 0.07905411720275879
step: 120, loss: 0.0038353216368705034
step: 130, loss: 0.032451316714286804
step: 140, loss: 0.0673760399222374
step: 150, loss: 0.02391761727631092
step: 160, loss: 0.012844772078096867
step: 170, loss: 0.016970358788967133
step: 180, loss: 0.04381484165787697
step: 190, loss: 0.2501792907714844
step: 200, loss: 0.060954105108976364

```

```

step: 210, loss: 0.0010477469768375158
step: 220, loss: 0.0026837224140763283
step: 230, loss: 0.009823462925851345
step: 240, loss: 0.02938971109688282
step: 250, loss: 0.06171730160713196
step: 260, loss: 0.0018924959003925323
step: 270, loss: 0.06822988390922546
step: 280, loss: 0.005751635879278183
step: 290, loss: 0.04290067031979561
step: 300, loss: 0.0058183507062494755
step: 310, loss: 0.004233103711158037
step: 320, loss: 0.05699579417705536
step: 330, loss: 0.015663083642721176
step: 340, loss: 0.003123285947367549
step: 350, loss: 0.2184174805879593
step: 360, loss: 0.03867006301879883
step: 370, loss: 0.002845662645995617
step: 380, loss: 0.0028416099958121777
step: 390, loss: 0.00021452843793667853
step: 400, loss: 0.018878737464547157
step: 410, loss: 0.003323888871818781
step: 420, loss: 0.016294535249471664
step: 430, loss: 0.0034157861955463886
step: 440, loss: 0.007655350491404533
step: 450, loss: 0.013896823860704899
step: 460, loss: 0.00501150032505393
step: 470, loss: 0.0012153908610343933
step: 480, loss: 0.0020255260169506073
step: 490, loss: 0.002996128983795643
step: 500, loss: 0.051145270466804504
step: 510, loss: 0.04011933133006096
step: 520, loss: 0.04520272836089134
step: 530, loss: 0.08799539506435394
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
    warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 4: dev_f1=0.958891454965358, f1=0.9542302357836338,
best_f1=0.9542302357836338
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0038605525624006987
step: 10, loss: 0.01232614740729332
step: 20, loss: 0.01573917455971241

```

step: 30, loss: 0.037895672023296356  
step: 40, loss: 0.0014963089488446712  
step: 50, loss: 0.021054215729236603  
step: 60, loss: 0.03255026414990425  
step: 70, loss: 0.015437000431120396  
step: 80, loss: 0.030341671779751778  
step: 90, loss: 0.06804543733596802  
step: 100, loss: 0.08549012988805771  
step: 110, loss: 0.016083799302577972  
step: 120, loss: 0.11623214185237885  
step: 130, loss: 0.0104984100908041  
step: 140, loss: 0.04031507298350334  
step: 150, loss: 0.017709173262119293  
step: 160, loss: 0.019746100530028343  
step: 170, loss: 0.04144442081451416  
step: 180, loss: 0.012221201322972775  
step: 190, loss: 0.0060996548272669315  
step: 200, loss: 0.008698237128555775  
step: 210, loss: 0.0002403192047495395  
step: 220, loss: 0.004049630369991064  
step: 230, loss: 0.008632569573819637  
step: 240, loss: 0.003300553187727928  
step: 250, loss: 0.15885311365127563  
step: 260, loss: 0.004261228255927563  
step: 270, loss: 0.0182918980717659  
step: 280, loss: 0.005608962848782539  
step: 290, loss: 0.0070995911955833435  
step: 300, loss: 0.028752028942108154  
step: 310, loss: 0.16887806355953217  
step: 320, loss: 0.016299335286021233  
step: 330, loss: 0.005830602254718542  
step: 340, loss: 0.0026317923329770565  
step: 350, loss: 0.0016414523124694824  
step: 360, loss: 6.832745566498488e-05  
step: 370, loss: 0.00012413485092110932  
step: 380, loss: 0.00015351959154941142  
step: 390, loss: 0.005522002000361681  
step: 400, loss: 0.06843914836645126  
step: 410, loss: 0.07063174992799759  
step: 420, loss: 0.21542447805404663  
step: 430, loss: 0.08241183310747147  
step: 440, loss: 0.0005808962741866708  
step: 450, loss: 0.01826796866953373  
step: 460, loss: 0.02175351232290268  
step: 470, loss: 0.03888111189007759  
step: 480, loss: 0.003980396315455437  
step: 490, loss: 0.15274111926555634  
step: 500, loss: 0.0523979552090168

```

step: 510, loss: 0.07610027492046356
step: 520, loss: 0.02300518937408924
step: 530, loss: 0.05452800914645195
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 5: dev_f1=0.9547596606974552, f1=0.946773433820066,
best_f1=0.9542302357836338
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.013047179207205772
step: 10, loss: 0.0074784583412110806
step: 20, loss: 0.0029708256479352713
step: 30, loss: 0.0007146992138586938
step: 40, loss: 0.0006229663849808276
step: 50, loss: 0.00042033300269395113
step: 60, loss: 0.011300287209451199
step: 70, loss: 0.0053926208056509495
step: 80, loss: 0.0003719555679708719
step: 90, loss: 0.0034812763333320618
step: 100, loss: 0.005061126314103603
step: 110, loss: 0.00347664556466043
step: 120, loss: 0.024944165721535683
step: 130, loss: 0.02053925022482872
step: 140, loss: 0.00903218425810337
step: 150, loss: 0.0033964132890105247
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 4096.0
step: 160, loss: 0.06107741966843605
step: 170, loss: 0.0008292743004858494
step: 180, loss: 0.004590541590005159
step: 190, loss: 0.020975343883037567
step: 200, loss: 0.03766445443034172
step: 210, loss: 0.01266125962138176
step: 220, loss: 0.0029956733342260122
step: 230, loss: 0.028614027425646782
step: 240, loss: 0.0012901189038529992
step: 250, loss: 0.17010506987571716
step: 260, loss: 0.001665326301008463
step: 270, loss: 0.03398897871375084
step: 280, loss: 0.21252334117889404
step: 290, loss: 0.01498398557305336
step: 300, loss: 0.04688074812293053
step: 310, loss: 0.13006113469600677
step: 320, loss: 0.007175024598836899
step: 330, loss: 0.0003917287103831768
step: 340, loss: 0.00018671169527806342
step: 350, loss: 0.005625689402222633

```



```

step: 360, loss: 0.02097342722117901
step: 370, loss: 0.009506139904260635
step: 380, loss: 0.000884358654730022
step: 390, loss: 0.000841766654048115
step: 400, loss: 0.03140394017100334
step: 410, loss: 0.00011644762707874179
step: 420, loss: 0.010244268923997879
step: 430, loss: 0.00033729441929608583
step: 440, loss: 0.00015927555796224624
step: 450, loss: 0.11320704221725464
step: 460, loss: 0.003128021489828825
step: 470, loss: 0.0051229288801550865
step: 480, loss: 0.01105088833719492
step: 490, loss: 0.014711637049913406
step: 500, loss: 0.005886551458388567
step: 510, loss: 0.07253848761320114
step: 520, loss: 0.0006994776194915175
step: 530, loss: 0.001306309481151402
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 6: dev_f1=0.9585081585081584, f1=0.9501165501165502,
best_f1=0.9542302357836338
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00837091077119112
step: 10, loss: 0.02129383012652397
step: 20, loss: 0.004557020962238312
step: 30, loss: 0.013359930366277695
step: 40, loss: 0.0049863760359585285
step: 50, loss: 0.004818301647901535
step: 60, loss: 0.0017297479789704084
step: 70, loss: 0.033134445548057556
step: 80, loss: 0.0004636826051864773
step: 90, loss: 0.001523482846096158
step: 100, loss: 0.0007819464663043618
step: 110, loss: 0.00013089382264297456
step: 120, loss: 0.0003133115533273667
step: 130, loss: 0.0007986948476172984
step: 140, loss: 0.0010870882542803884
step: 150, loss: 0.01395981851965189
step: 160, loss: 0.000518164481036365
step: 170, loss: 0.02474963665008545
step: 180, loss: 0.011856612749397755
step: 190, loss: 0.012111039832234383
step: 200, loss: 0.00034934081486426294
step: 210, loss: 0.002191677689552307

```

```

step: 220, loss: 0.051815539598464966
step: 230, loss: 0.00129488087259233
step: 240, loss: 0.0032797178719192743
step: 250, loss: 0.07905413210391998
step: 260, loss: 0.13298454880714417
step: 270, loss: 0.005446027964353561
step: 280, loss: 0.0071474541909992695
step: 290, loss: 0.024256503209471703
step: 300, loss: 0.00022165404516272247
step: 310, loss: 0.000769158941693604
step: 320, loss: 0.031745780259370804
step: 330, loss: 0.0028276790399104357
step: 340, loss: 0.0006160767516121268
step: 350, loss: 0.0004021882195957005
step: 360, loss: 0.05699755251407623
step: 370, loss: 0.006578715518116951
step: 380, loss: 0.0020898268558084965
step: 390, loss: 0.0054898555390536785
step: 400, loss: 0.0028938271570950747
step: 410, loss: 0.0035632799845188856
step: 420, loss: 0.02191292867064476
step: 430, loss: 0.003347294172272086
step: 440, loss: 0.0006946888752281666
step: 450, loss: 0.015999065712094307
step: 460, loss: 0.03530074656009674
step: 470, loss: 0.18140701949596405
step: 480, loss: 0.005645327735692263
step: 490, loss: 0.004902991931885481
step: 500, loss: 0.010214779525995255
step: 510, loss: 0.000616520585026592
step: 520, loss: 0.0006859801360405982
step: 530, loss: 0.00033297823392786086
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
epoch 7: dev_f1=0.9552098066949553, f1=0.947417840375587,
best_f1=0.9542302357836338
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0013212040066719055
step: 10, loss: 0.0010601825779303908
step: 20, loss: 0.0009368265746161342
step: 30, loss: 0.0003726765571627766
step: 40, loss: 0.00011457215441623703
step: 50, loss: 0.0003302383702248335
step: 60, loss: 0.00027320615481585264
step: 70, loss: 0.0008105913293547928

```

```
step: 80, loss: 0.0023350431583821774
step: 90, loss: 0.00012596572923939675
step: 100, loss: 0.005079344380646944
step: 110, loss: 0.0002956337411887944
step: 120, loss: 0.013929177075624466
step: 130, loss: 0.0004589864402078092
step: 140, loss: 0.0001459386112401262
step: 150, loss: 0.006796679925173521
step: 160, loss: 0.0002831815800163895
step: 170, loss: 0.08579052239656448
step: 180, loss: 0.00018012372311204672
step: 190, loss: 0.022041410207748413
step: 200, loss: 0.0079684779047966
step: 210, loss: 0.02121076174080372
step: 220, loss: 0.0005131200887262821
step: 230, loss: 0.11260386556386948
step: 240, loss: 0.022455591708421707
step: 250, loss: 0.013062422163784504
step: 260, loss: 0.00034917693119496107
step: 270, loss: 0.00213979952968657
step: 280, loss: 9.485684859100729e-05
step: 290, loss: 0.05853739008307457
step: 300, loss: 0.0001554071350255981
step: 310, loss: 0.020241662859916687
step: 320, loss: 0.0004409824905451387
step: 330, loss: 0.0001519373618066311
step: 340, loss: 0.055927131325006485
step: 350, loss: 0.0011135400272905827
step: 360, loss: 0.052589111030101776
step: 370, loss: 0.02309800125658512
step: 380, loss: 0.0005482321139425039
step: 390, loss: 0.004373207688331604
step: 400, loss: 0.006840356159955263
step: 410, loss: 0.005514539312571287
step: 420, loss: 0.0007716730469837785
step: 430, loss: 0.010029067285358906
step: 440, loss: 0.18229664862155914
step: 450, loss: 0.001122659770771861
step: 460, loss: 0.006605980917811394
step: 470, loss: 0.03385964035987854
step: 480, loss: 0.0025663760025054216
step: 490, loss: 0.0007572518661618233
step: 500, loss: 0.0002083620202029124
step: 510, loss: 0.002814711071550846
step: 520, loss: 0.00018849092884920537
step: 530, loss: 0.005833880044519901
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
```

```
warnings.warn("An input tensor was not cuda.")
epoch 8: dev_f1=0.9496268656716418, f1=0.9460465116279071,
best_f1=0.9542302357836338
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0004459635238163173
step: 10, loss: 0.029038017615675926
step: 20, loss: 0.0011964672012254596
step: 30, loss: 0.062026116997003555
step: 40, loss: 0.0008914134232327342
step: 50, loss: 0.00021298901992850006
step: 60, loss: 0.0006592852296307683
step: 70, loss: 0.021519631147384644
step: 80, loss: 0.04663478210568428
step: 90, loss: 0.05568462982773781
step: 100, loss: 0.0015513473190367222
step: 110, loss: 0.01180457416921854
step: 120, loss: 0.0016059302724897861
step: 130, loss: 0.01593189686536789
step: 140, loss: 0.0135953389108181
step: 150, loss: 0.10300803184509277
step: 160, loss: 0.0006803692085668445
step: 170, loss: 0.002851678291335702
step: 180, loss: 0.006748998537659645
step: 190, loss: 0.019186895340681076
step: 200, loss: 0.00012314885680098087
step: 210, loss: 0.0005393525352701545
step: 220, loss: 0.0037428722716867924
step: 230, loss: 0.0033929478377103806
step: 240, loss: 0.0016578231006860733
step: 250, loss: 0.011788906529545784
step: 260, loss: 0.0001508953282609582
step: 270, loss: 0.0037867792416363955
step: 280, loss: 0.0013423803029581904
step: 290, loss: 0.00013109931023791432
step: 300, loss: 2.6676116249291226e-05
step: 310, loss: 0.0239297803491354
step: 320, loss: 4.521420123637654e-05
step: 330, loss: 0.004031172953546047
step: 340, loss: 0.0021188880782574415
step: 350, loss: 0.06684368848800659
step: 360, loss: 0.0011006054701283574
step: 370, loss: 0.003972445614635944
step: 380, loss: 0.003247408661991358
step: 390, loss: 0.009684067219495773
step: 400, loss: 0.01459359098225832
step: 410, loss: 0.013910344801843166
```

```
step: 420, loss: 0.0019715754315257072
step: 430, loss: 0.1479337215423584
step: 440, loss: 0.00036067888140678406
step: 450, loss: 0.0067850383929908276
step: 460, loss: 0.0016695233061909676
step: 470, loss: 0.010910986922681332
step: 480, loss: 0.0027001856360584497
step: 490, loss: 0.02067626640200615
step: 500, loss: 0.0005849571898579597
step: 510, loss: 0.0033647119998931885
step: 520, loss: 0.0020132744684815407
step: 530, loss: 0.023096727207303047
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
epoch 9: dev_f1=0.9548447789275636, f1=0.9471221338324755,
best_f1=0.9542302357836338
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0005591843510046601
step: 10, loss: 0.00040139947668649256
step: 20, loss: 8.877511572791263e-05
step: 30, loss: 0.0007580303936265409
step: 40, loss: 0.00020181597210466862
step: 50, loss: 5.0403818022459745e-05
step: 60, loss: 0.008355457335710526
step: 70, loss: 0.014888295903801918
step: 80, loss: 0.01765272207558155
step: 90, loss: 0.00033333979081362486
step: 100, loss: 0.006259043235331774
step: 110, loss: 0.012287669815123081
step: 120, loss: 0.0018062859307974577
step: 130, loss: 0.0027095519471913576
step: 140, loss: 4.8488276661373675e-05
step: 150, loss: 9.277532808482647e-05
step: 160, loss: 0.023482711985707283
step: 170, loss: 0.002314928686246276
step: 180, loss: 0.0897517129778862
step: 190, loss: 0.0016432907432317734
step: 200, loss: 0.006250279955565929
step: 210, loss: 0.022393442690372467
step: 220, loss: 0.0013935982715338469
step: 230, loss: 0.0020063098054379225
step: 240, loss: 0.002302485518157482
step: 250, loss: 0.0004454689333215356
step: 260, loss: 0.004081275779753923
step: 270, loss: 0.0027704115491360426
```

```

step: 280, loss: 0.028383780270814896
step: 290, loss: 0.000741257332265377
step: 300, loss: 0.00408603809773922
step: 310, loss: 0.01330000814050436
step: 320, loss: 0.009392848238348961
step: 330, loss: 0.008065510541200638
step: 340, loss: 0.00026455370243638754
step: 350, loss: 0.0008300073095597327
step: 360, loss: 0.0008045885479077697
step: 370, loss: 0.047309234738349915
step: 380, loss: 0.0015829304466024041
step: 390, loss: 0.0001418121246388182
step: 400, loss: 4.4875887397211045e-05
step: 410, loss: 0.0024873886723071337
step: 420, loss: 1.949416946445126e-05
step: 430, loss: 3.253854447393678e-05
step: 440, loss: 0.00031598153873346746
step: 450, loss: 0.00820738822221756
step: 460, loss: 0.0007990964804776013
step: 470, loss: 0.007492485921829939
step: 480, loss: 0.005034476052969694
step: 490, loss: 0.006316180806607008
step: 500, loss: 0.0010165675776079297
step: 510, loss: 0.0006622286164201796
step: 520, loss: 3.745503272511996e-05
step: 530, loss: 0.028662260621786118
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
epoch 10: dev_f1=0.9583528310715957, f1=0.9548627268496975,
best_f1=0.9542302357836338
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0010763726895675063
step: 10, loss: 1.6677751773386262e-05
step: 20, loss: 0.000206079421332106
step: 30, loss: 1.5910520232864656e-05
step: 40, loss: 8.315791637869552e-05
step: 50, loss: 3.3391253964509815e-05
step: 60, loss: 0.0001140195454354398
step: 70, loss: 5.3082170779816806e-05
step: 80, loss: 2.7962491003563628e-05
step: 90, loss: 0.0006115604192018509
step: 100, loss: 0.0010892658028751612
step: 110, loss: 2.0660076188505627e-05
step: 120, loss: 0.010129988193511963
step: 130, loss: 6.953419506317005e-05

```

```

step: 140, loss: 8.556987449992448e-05
step: 150, loss: 0.0006527575897052884
step: 160, loss: 0.0015574793796986341
step: 170, loss: 0.02118952013552189
step: 180, loss: 1.9106175386696123e-05
step: 190, loss: 0.00127256172709167
step: 200, loss: 7.419486792059615e-05
step: 210, loss: 2.0480574676184915e-05
step: 220, loss: 0.00041457617771811783
step: 230, loss: 0.0006745827849954367
step: 240, loss: 0.007147268392145634
step: 250, loss: 1.6063047951320186e-05
step: 260, loss: 0.004060852807015181
step: 270, loss: 0.000984812737442553
step: 280, loss: 0.00021445186575874686
step: 290, loss: 0.009163813665509224
step: 300, loss: 6.031220436852891e-06
step: 310, loss: 0.08544608205555687
step: 320, loss: 0.0009534055134281516
step: 330, loss: 7.783833279972896e-05
step: 340, loss: 0.04271091893315315
step: 350, loss: 0.00030404888093471527
step: 360, loss: 0.0009135605650953948
step: 370, loss: 0.01317763701081276
step: 380, loss: 0.0011434579500928521
step: 390, loss: 0.0026707416400313377
step: 400, loss: 4.881954373558983e-05
step: 410, loss: 0.0007116035558283329
step: 420, loss: 0.001003554556518793
step: 430, loss: 0.0038226039614528418
step: 440, loss: 0.0006941870669834316
step: 450, loss: 0.006098269950598478
step: 460, loss: 0.01733282208442688
step: 470, loss: 0.0027154856361448765
step: 480, loss: 0.02066938579082489
step: 490, loss: 0.0012417715042829514
step: 500, loss: 0.00023529166355729103
step: 510, loss: 0.0026118189562112093
step: 520, loss: 0.0016506362007930875
step: 530, loss: 0.01178065873682499
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
epoch 11: dev_f1=0.9530887134231305, f1=0.9519408502772643,
best_f1=0.9542302357836338
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")

```

step: 0, loss: 0.01787511818110943  
step: 10, loss: 0.0018646661192178726  
step: 20, loss: 0.0029903261456638575  
step: 30, loss: 6.214211316546425e-05  
step: 40, loss: 0.0006612979923374951  
step: 50, loss: 0.03890007734298706  
step: 60, loss: 0.0006389091722667217  
step: 70, loss: 0.003954425919800997  
step: 80, loss: 0.0001629946636967361  
step: 90, loss: 0.0040755197405815125  
step: 100, loss: 0.012465997599065304  
step: 110, loss: 0.0005564512102864683  
step: 120, loss: 0.000612023111898452  
step: 130, loss: 0.00013643187412526459  
step: 140, loss: 2.4399958419962786e-05  
step: 150, loss: 0.009112892672419548  
step: 160, loss: 0.0003705244744196534  
step: 170, loss: 0.00033691574935801327  
step: 180, loss: 2.5499135517748073e-05  
step: 190, loss: 0.00105486495885998  
step: 200, loss: 0.0033500283025205135  
step: 210, loss: 0.00015231494035106152  
step: 220, loss: 0.003485976718366146  
step: 230, loss: 0.00010347264469601214  
step: 240, loss: 0.023261768743395805  
step: 250, loss: 4.788856313098222e-05  
step: 260, loss: 0.0001389469689456746  
step: 270, loss: 0.011860496364533901  
step: 280, loss: 3.5212542570661753e-05  
step: 290, loss: 0.00296963844448328  
step: 300, loss: 0.001606250531040132  
step: 310, loss: 0.011722084134817123  
step: 320, loss: 0.010859710164368153  
step: 330, loss: 0.003919024486094713  
step: 340, loss: 0.018900811672210693  
step: 350, loss: 3.7771245843032375e-05  
step: 360, loss: 0.0029551684856414795  
step: 370, loss: 0.0006601848290301859  
step: 380, loss: 0.00030306371627375484  
step: 390, loss: 0.0004131880996283144  
step: 400, loss: 0.00032425494282506406  
step: 410, loss: 0.0011723325587809086  
step: 420, loss: 0.0016053941799327731  
step: 430, loss: 0.00044122125837020576  
step: 440, loss: 0.0024674010928720236  
step: 450, loss: 0.003071896731853485  
step: 460, loss: 0.00021447626932058483  
step: 470, loss: 0.008109279908239841



```

step: 480, loss: 0.008104861713945866
step: 490, loss: 0.0003042286552954465
step: 500, loss: 0.00033967988565564156
step: 510, loss: 0.020429324358701706
step: 520, loss: 0.0026269392110407352
step: 530, loss: 0.0007686792523600161
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
    warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 12: dev_f1=0.9602618045815802, f1=0.9507434944237918,
best_f1=0.9507434944237918
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00010439189645694569
step: 10, loss: 0.00012291391612961888
step: 20, loss: 0.00028624097467400134
step: 30, loss: 4.625465953722596e-05
step: 40, loss: 0.015565079636871815
step: 50, loss: 0.0010629595490172505
step: 60, loss: 0.002755924826487899
step: 70, loss: 0.001105902367271483
step: 80, loss: 0.0063237389549613
step: 90, loss: 0.0003987993113696575
step: 100, loss: 0.0003381053393241018
step: 110, loss: 8.646654896438122e-05
step: 120, loss: 9.269882866647094e-05
step: 130, loss: 0.00013164235861040652
step: 140, loss: 0.0022363823372870684
step: 150, loss: 0.0023962287232279778
step: 160, loss: 0.032037924975156784
step: 170, loss: 0.0006354136276058853
step: 180, loss: 0.0003351061895955354
step: 190, loss: 0.0014066807925701141
step: 200, loss: 3.596715396270156e-05
step: 210, loss: 0.002536286134272814
step: 220, loss: 0.0012009780621156096
step: 230, loss: 0.00041437847539782524
step: 240, loss: 0.0032994714565575123
step: 250, loss: 0.00020654805121012032
step: 260, loss: 0.002414499409496784
step: 270, loss: 0.005626623518764973
step: 280, loss: 4.058566628373228e-05
step: 290, loss: 1.963920658454299e-05

```

```

step: 300, loss: 1.2736648386635352e-05
step: 310, loss: 7.883125363150612e-05
step: 320, loss: 1.4181967344484292e-05
step: 330, loss: 0.00018944840121548623
step: 340, loss: 0.00869692675769329
step: 350, loss: 0.015242439694702625
step: 360, loss: 0.08336086571216583
step: 370, loss: 0.005451134406030178
step: 380, loss: 6.695480988128111e-05
step: 390, loss: 0.000705618062056601
step: 400, loss: 0.004350417759269476
step: 410, loss: 0.003364653093740344
step: 420, loss: 0.0003895737463608384
step: 430, loss: 0.0003253319300711155
step: 440, loss: 0.00012047508789692074
step: 450, loss: 0.00010242194548482075
step: 460, loss: 0.00012078231520717964
step: 470, loss: 0.006607299670577049
step: 480, loss: 2.2115693354862742e-05
step: 490, loss: 0.0001850239495979622
step: 500, loss: 0.0009828959591686726
step: 510, loss: 8.487635932397097e-05
step: 520, loss: 6.0372309235390276e-05
step: 530, loss: 0.0001214350268128328
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
epoch 13: dev_f1=0.9583718778908418, f1=0.9537333944113604,
best_f1=0.9507434944237918
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00045268412213772535
step: 10, loss: 0.0035132234916090965
step: 20, loss: 0.0001380346220685169
step: 30, loss: 0.0010457801399752498
step: 40, loss: 0.0002923332212958485
step: 50, loss: 0.0011087325401604176
step: 60, loss: 0.01065845973789692
step: 70, loss: 0.002673657378181815
step: 80, loss: 0.00011633960821200162
step: 90, loss: 1.936290937010199e-05
step: 100, loss: 0.00016516796313226223
step: 110, loss: 4.5991091610630974e-05
step: 120, loss: 1.102677470044e-05
step: 130, loss: 1.506102125858888e-05
step: 140, loss: 0.005833739414811134
step: 150, loss: 0.001193888601846993

```

```

step: 160, loss: 0.0008023512782528996
step: 170, loss: 0.004272558260709047
step: 180, loss: 0.004925974644720554
step: 190, loss: 0.000373811402823776
step: 200, loss: 0.00012245902325958014
step: 210, loss: 0.0004234718217048794
step: 220, loss: 0.00019562151283025742
step: 230, loss: 2.817696258716751e-05
step: 240, loss: 0.0017362793441861868
step: 250, loss: 0.0005281332996673882
step: 260, loss: 0.0009675284964032471
step: 270, loss: 0.0037613138556480408
step: 280, loss: 0.0001325865014223382
step: 290, loss: 9.350919572170824e-05
step: 300, loss: 0.0017902494873851538
step: 310, loss: 0.00025266504962928593
step: 320, loss: 0.00012109697127016261
step: 330, loss: 0.0021429192274808884
step: 340, loss: 0.001005443511530757
step: 350, loss: 0.00011822285159723833
step: 360, loss: 0.0015470385551452637
step: 370, loss: 0.0007514248718507588
step: 380, loss: 0.0002502397692296654
step: 390, loss: 0.0026690270751714706
step: 400, loss: 0.004692866466939449
step: 410, loss: 0.00034604297252371907
step: 420, loss: 0.00013998153735883534
step: 430, loss: 0.005590761546045542
step: 440, loss: 0.0005869594751857221
step: 450, loss: 0.0022633839398622513
step: 460, loss: 0.02241344563663006
step: 470, loss: 3.9227932575158775e-05
step: 480, loss: 0.0004408984095789492
step: 490, loss: 0.005409381352365017
step: 500, loss: 0.0023881313391029835
step: 510, loss: 0.028887975960969925
step: 520, loss: 0.0003731794422492385
step: 530, loss: 0.00048572293599136174
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 14: dev_f1=0.958955223880597, f1=0.9537037037037037,
best_f1=0.9507434944237918
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0006499439477920532
step: 10, loss: 0.00031446636421605945

```

step: 20, loss: 0.00016845835489220917  
step: 30, loss: 0.0024721031077206135  
step: 40, loss: 0.006301965564489365  
step: 50, loss: 0.0011474336497485638  
step: 60, loss: 0.01056100707501173  
step: 70, loss: 2.0719084204756655e-05  
step: 80, loss: 1.9080465790466405e-05  
step: 90, loss: 0.0001493448653491214  
step: 100, loss: 1.7225482224603184e-05  
step: 110, loss: 0.006132818758487701  
step: 120, loss: 2.7049711206927896e-05  
step: 130, loss: 2.226159631391056e-05  
step: 140, loss: 2.3337046513915993e-05  
step: 150, loss: 2.240557296318002e-05  
step: 160, loss: 7.556217315141112e-05  
step: 170, loss: 6.325059075606987e-05  
step: 180, loss: 0.0007943622767925262  
step: 190, loss: 0.0018191100098192692  
step: 200, loss: 0.0014799507334828377  
step: 210, loss: 0.0008157408447004855  
step: 220, loss: 0.0020637442357838154  
step: 230, loss: 4.039338818984106e-05  
step: 240, loss: 0.00015213096048682928  
step: 250, loss: 5.073358988738619e-05  
step: 260, loss: 3.758318780455738e-05  
step: 270, loss: 2.4410515834460966e-05  
step: 280, loss: 0.00021025209571234882  
step: 290, loss: 0.0014387102564796805  
step: 300, loss: 6.572194979526103e-05  
step: 310, loss: 0.017544645816087723  
step: 320, loss: 0.00012800763943232596  
step: 330, loss: 0.00011548939073691145  
step: 340, loss: 0.00016550166765227914  
step: 350, loss: 0.009481027722358704  
step: 360, loss: 3.473985998425633e-05  
step: 370, loss: 0.0009073467808775604  
step: 380, loss: 7.585387356812134e-05  
step: 390, loss: 6.883282912895083e-05  
step: 400, loss: 0.002279818756505847  
step: 410, loss: 3.8685517210979015e-05  
step: 420, loss: 5.719956607208587e-05  
step: 430, loss: 3.51681919710245e-05  
step: 440, loss: 0.009653199464082718  
step: 450, loss: 0.010808511637151241  
step: 460, loss: 0.00123012310359627  
step: 470, loss: 6.836171814939007e-05  
step: 480, loss: 0.0024434933438897133  
step: 490, loss: 0.008897574618458748

```

step: 500, loss: 0.0009494246914982796
step: 510, loss: 0.00015745563723612577
step: 520, loss: 1.602954216650687e-05
step: 530, loss: 0.004944357089698315
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 15: dev_f1=0.9598506069094305, f1=0.9541454377026402,
best_f1=0.9507434944237918
mv: cannot stat '*_dev.pt': No such file or directory
Selected optimization level 02:  FP16 training with FP32 batchnorm and FP32
master weights.

Defaults for this optimization level are:
enabled                : True
opt_level              : 02
cast_model_type        : torch.float16
patch_torch_functions  : False
keep_batchnorm_fp32    : True
master_weights         : True
loss_scale             : dynamic
Processing user overrides (additional kwargs that are not None)...
After processing overrides, optimization options are:
enabled                : True
opt_level              : 02
cast_model_type        : torch.float16
patch_torch_functions  : False
keep_batchnorm_fp32    : True
master_weights         : True
loss_scale             : dynamic
Warning:  multi_tensor_applier fused unscale kernel is unavailable, possibly
because apex was installed without --cuda_ext --cpp_ext. Using Python fallback.
Original ImportError was: ModuleNotFoundError("No module named 'amp_C'")
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
Oit [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
5742it [00:33, 169.53it/s]
load_f1 = 0.9568884723523898
real_f1 = 0.9551401869158879
Oit [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
4398it [00:31, 140.33it/s]

```