

# FMedium\_50\_1\_2\_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 | 17.03 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 1.5 MB/s
Collecting numpy==1.19.2
  Downloading numpy-1.19.2-cp37-cp37m-manylinux2010_x86_64.whl (14.5 MB)
    |                               | 14.5 MB 55.0 MB/s
Collecting regex==2019.12.20
  Downloading regex-2019.12.20-cp37-cp37m-manylinux2010_x86_64.whl (689 kB)
    |                               | 689 kB 50.7 MB/s
Collecting scipy==1.3.2
  Downloading scipy-1.3.2-cp37-cp37m-manylinux1_x86_64.whl (25.2 MB)
    |                               | 25.2 MB 1.5 MB/s
Collecting sentencepiece==0.1.85
  Downloading sentencepiece-0.1.85-cp37-cp37m-manylinux1_x86_64.whl (1.0 MB)
    |                               | 1.0 MB 53.2 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 57.1 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 | 13.05 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 25.7 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 59.5 MB/s
Collecting torch==1.4.0
  Downloading torch-1.4.0-cp37-cp37m-manylinux1_x86_64.whl (753.4 MB)
|                                     | 753.4 MB 7.2 kB/s
Collecting tqdm==4.41.0
  Downloading tqdm-4.41.0-py2.py3-none-any.whl (56 kB)
|                                     | 56 kB 5.3 MB/s
Collecting transformers==3.1.0
  Downloading transformers-3.1.0-py3-none-any.whl (884 kB)
|                                     | 884 kB 68.4 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 52.1 MB/s
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)
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: 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: 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 thinc<7.4.0,>=7.3.0
  Downloading thinc-7.3.1-cp37-cp37m-manylinux1_x86_64.whl (2.2 MB)
    | 2.2 MB 55.8 MB/s
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 catalogue<1.1.0,>=0.0.7
  Downloading catalogue-1.0.0-py2.py3-none-any.whl (7.7 kB)
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: 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)
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: setuptools in /usr/local/lib/python3.7/dist-
packages (from spacy==2.2.3->-r requirements.txt (line 4)) (57.4.0)
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 45.4 MB/s
Collecting srsly<1.1.0,>=0.1.0
  Downloading srsly-1.0.5-cp37-cp37m-manylinux2014_x86_64.whl (184 kB)
    | 184 kB 70.0 MB/s
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 tokenizers==0.8.1.rc2
  Downloading tokenizers-0.8.1rc2-cp37-cp37m-manylinux1_x86_64.whl (3.0 MB)

```

```

| 3.0 MB 54.3 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)
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)
Collecting sacremoses
  Downloading sacremoses-0.0.53.tar.gz (880 kB)
    | 880 kB 67.5 MB/s
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: 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: 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: 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: 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: 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: 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=1449921
sha256=f2c11164224cfa860006ef23e2c60e666c31d419924394cb9e174368fdf2fc76
  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=089d8524b094f355cd9e78c432c8b25a0406a2a32c6eb6555ef85cd714dd5ddd
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: 10085, done.
remote: Counting objects: 100% (201/201), done.
remote: Compressing objects: 100% (138/138), done.
remote: Total 10085 (delta 98), reused 131 (delta 59), pack-reused 9884
Receiving objects: 100% (10085/10085), 14.95 MiB | 10.70 MiB/s, done.
Resolving deltas: 100% (6902/6902), 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-htj_0b2_
Created temporary directory: /tmp/pip-req-tracker-hbn967t5
Initialized build tracking at /tmp/pip-req-tracker-hbn967t5
Created build tracker: /tmp/pip-req-tracker-hbn967t5
Entered build tracker: /tmp/pip-req-tracker-hbn967t5
```

```

Created temporary directory: /tmp/pip-install-2lfurfqi
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-zz2vzxaf
  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-
hbn967t5'
  Running setup.py (path:/tmp/pip-req-build-zz2vzxaf/setup.py) egg_info for
package from file:///content/ditto/apex
  Created temporary directory: /tmp/pip-pip-egg-info-r_zif3q7
  Running command python setup.py egg_info

  torch.__version__ = 1.4.0

  running egg_info
  creating /tmp/pip-pip-egg-info-r_zif3q7/apex.egg-info
  writing /tmp/pip-pip-egg-info-r_zif3q7/apex.egg-info/PKG-INFO
  writing dependency_links to /tmp/pip-pip-egg-info-r_zif3q7/apex.egg-
info/dependency_links.txt
  writing top-level names to /tmp/pip-pip-egg-info-r_zif3q7/apex.egg-
info/top_level.txt
  writing manifest file '/tmp/pip-pip-egg-info-r_zif3q7/apex.egg-
info/SOURCES.txt'
  adding license file 'LICENSE'
  writing manifest file '/tmp/pip-pip-egg-info-r_zif3q7/apex.egg-
info/SOURCES.txt'
  Source in /tmp/pip-req-build-zz2vzxaf 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-hbn967t5'
Created temporary directory: /tmp/pip-unpack-3dtur6oy
Building wheels for collected packages: apex
  Created temporary directory: /tmp/pip-wheel-jdb1celf
  Building wheel for apex (setup.py) ... Destination directory:
/tmp/pip-wheel-jdb1celf
  Running command /usr/bin/python3 -u -c 'import io, os, sys, setuptools,
tokenize; sys.argv[0] = ''"/tmp/pip-req-build-zz2vzxaf/setup.py''';

```



```
__file__='''/tmp/pip-req-build-zz2vzxaf/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-jdb1celf
```

```
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/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/__init__.py -> build/lib/apex/mlp
copying apex/mlp/mlp.py -> build/lib/apex/mlp
creating build/lib/apex/amp
copying apex/amp/opt.py -> build/lib/apex/amp
copying apex/amp/scaler.py -> build/lib/apex/amp
copying apex/amp/frontend.py -> build/lib/apex/amp
copying apex/amp/rnn_compat.py -> build/lib/apex/amp
copying apex/amp/__version__.py -> build/lib/apex/amp
copying apex/amp/utils.py -> build/lib/apex/amp
copying apex/amp/__init__.py -> build/lib/apex/amp
copying apex/amp/amp.py -> build/lib/apex/amp
copying apex/amp/compat.py -> build/lib/apex/amp
copying apex/amp/wrap.py -> build/lib/apex/amp
copying apex/amp/_initialize.py -> build/lib/apex/amp
copying apex/amp/_amp_state.py -> build/lib/apex/amp
copying apex/amp/_process_optimizer.py -> build/lib/apex/amp
copying apex/amp/handle.py -> build/lib/apex/amp
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/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/parallel
```

```

    copying apex/parallel/__init__.py -> build/lib/apex/parallel
    copying apex/parallel/optimized_sync_batchnorm_kernel.py ->
build/lib/apex/parallel
    copying apex/parallel/multiproc.py -> build/lib/apex/parallel
    copying apex/parallel/sync_batchnorm.py -> build/lib/apex/parallel
    copying apex/parallel/LARC.py -> build/lib/apex/parallel
    copying apex/parallel/distributed.py -> build/lib/apex/parallel
    copying apex/parallel/sync_batchnorm_kernel.py -> build/lib/apex/parallel
    copying apex/parallel/optimized_sync_batchnorm.py -> build/lib/apex/parallel
    creating build/lib/apex/optimizers
    copying apex/optimizers/fused_sgd.py -> build/lib/apex/optimizers
    copying apex/optimizers/__init__.py -> build/lib/apex/optimizers
    copying apex/optimizers/fused_adam.py -> build/lib/apex/optimizers
    copying apex/optimizers/fused_mixed_precision_lamb.py ->
build/lib/apex/optimizers
    copying apex/optimizers/fused_adagrad.py -> build/lib/apex/optimizers
    copying apex/optimizers/fused_novograd.py -> build/lib/apex/optimizers
    copying apex/optimizers/fused_lamb.py -> build/lib/apex/optimizers
    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/fp16util.py -> build/lib/apex/fp16_utils
    copying apex/fp16_utils/loss_scaler.py -> build/lib/apex/fp16_utils
    creating build/lib/apex/transformer
    copying apex/transformer/microbatches.py -> build/lib/apex/transformer
    copying apex/transformer/parallel_state.py -> build/lib/apex/transformer
    copying apex/transformer/utils.py -> build/lib/apex/transformer
    copying apex/transformer/__init__.py -> build/lib/apex/transformer
    copying apex/transformer/enums.py -> build/lib/apex/transformer
    copying apex/transformer/log_util.py -> build/lib/apex/transformer
    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/contrib
    copying apex/contrib/__init__.py -> build/lib/apex/contrib
    creating build/lib/apex/amp/lists
    copying apex/amp/lists/__init__.py -> build/lib/apex/amp/lists
    copying apex/amp/lists/tensor_overrides.py -> build/lib/apex/amp/lists
    copying apex/amp/lists/torch_overrides.py -> build/lib/apex/amp/lists
    copying apex/amp/lists/functional_overrides.py -> build/lib/apex/amp/lists
    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/_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/functional
    copying apex/transformer/functional/__init__.py ->
build/lib/apex/transformer/functional
    copying apex/transformer/functional/fused_softmax.py ->
build/lib/apex/transformer/functional
    creating build/lib/apex/transformer/pipeline_parallel
    copying apex/transformer/pipeline_parallel/utils.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/p2p_communication.py ->
build/lib/apex/transformer/pipeline_parallel
    copying apex/transformer/pipeline_parallel/_timers.py ->
build/lib/apex/transformer/pipeline_parallel
    creating build/lib/apex/transformer/layers
    copying apex/transformer/layers/layer_norm.py ->
build/lib/apex/transformer/layers
    copying apex/transformer/layers/__init__.py ->
build/lib/apex/transformer/layers
    creating build/lib/apex/transformer/tensor_parallel
    copying apex/transformer/tensor_parallel/memory.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/__init__.py ->
build/lib/apex/transformer/tensor_parallel
    copying apex/transformer/tensor_parallel/mappings.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/cross_entropy.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/random.py ->
build/lib/apex/transformer/tensor_parallel
    creating build/lib/apex/transformer/testing
    copying apex/transformer/testing/standalone_gpt.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/distributed_test_base.py ->
build/lib/apex/transformer/testing
    copying apex/transformer/testing/arguments.py ->
build/lib/apex/transformer/testing

```

```

    copying apex/transformer/testing/commons.py ->
build/lib/apex/transformer/testing
    copying apex/transformer/testing/global_vars.py ->
build/lib/apex/transformer/testing
    copying apex/transformer/testing/standalone_transformer_lm.py ->
build/lib/apex/transformer/testing
    creating build/lib/apex/transformer/pipeline_parallel/schedules
    copying apex/transformer/pipeline_parallel/schedules/fwd_bwd_no_pipelining.py
-> build/lib/apex/transformer/pipeline_parallel/schedules
    copying apex/transformer/pipeline_parallel/schedules/fwd_bwd_pipelining_with_i
nterleaving.py -> build/lib/apex/transformer/pipeline_parallel/schedules
    copying apex/transformer/pipeline_parallel/schedules/__init__.py ->
build/lib/apex/transformer/pipeline_parallel/schedules
    copying apex/transformer/pipeline_parallel/schedules/fwd_bwd_pipelining_withou
t_interleaving.py -> build/lib/apex/transformer/pipeline_parallel/schedules
    copying apex/transformer/pipeline_parallel/schedules/common.py ->
build/lib/apex/transformer/pipeline_parallel/schedules
    creating build/lib/apex/contrib/bottleneck
    copying apex/contrib/bottleneck/__init__.py ->
build/lib/apex/contrib/bottleneck
    copying apex/contrib/bottleneck/bottleneck_module_test.py ->
build/lib/apex/contrib/bottleneck
    copying apex/contrib/bottleneck/halo_exchangers.py ->
build/lib/apex/contrib/bottleneck
    copying apex/contrib/bottleneck/bottleneck.py ->
build/lib/apex/contrib/bottleneck
    copying apex/contrib/bottleneck/test.py -> build/lib/apex/contrib/bottleneck
    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/groupbn
    copying apex/contrib/groupbn/__init__.py -> build/lib/apex/contrib/groupbn
    copying apex/contrib/groupbn/batch_norm.py -> build/lib/apex/contrib/groupbn
    creating 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/encdec_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_self_multihead_attn_func.py ->
build/lib/apex/contrib/multihead_attn
    copying apex/contrib/multihead_attn/fast_encdec_multihead_attn_func.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/self_multihead_attn.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_func.py ->
build/lib/apex/contrib/multihead_attn
    copying apex/contrib/multihead_attn/self_multihead_attn_func.py ->
build/lib/apex/contrib/multihead_attn
    creating build/lib/apex/contrib/index_mul_2d
    copying apex/contrib/index_mul_2d/__init__.py ->
build/lib/apex/contrib/index_mul_2d
    copying apex/contrib/index_mul_2d/index_mul_2d.py ->
build/lib/apex/contrib/index_mul_2d
    creating build/lib/apex/contrib/optimizers
    copying apex/contrib/optimizers/fp16_optimizer.py ->
build/lib/apex/contrib/optimizers
    copying apex/contrib/optimizers/fused_sgd.py ->
build/lib/apex/contrib/optimizers
    copying apex/contrib/optimizers/__init__.py ->
build/lib/apex/contrib/optimizers
    copying apex/contrib/optimizers/fused_adam.py ->
build/lib/apex/contrib/optimizers
    copying apex/contrib/optimizers/distributed_fused_adam.py ->
build/lib/apex/contrib/optimizers
    copying apex/contrib/optimizers/distributed_fused_lamb.py ->
build/lib/apex/contrib/optimizers
    copying apex/contrib/optimizers/fused_lamb.py ->
build/lib/apex/contrib/optimizers
    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/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/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/asp.py -> build/lib/apex/contrib/sparsity
    copying apex/contrib/sparsity/sparse_masklib.py ->
build/lib/apex/contrib/sparsity
    creating build/lib/apex/contrib/layer_norm
    copying apex/contrib/layer_norm/layer_norm.py ->
build/lib/apex/contrib/layer_norm
    copying apex/contrib/layer_norm/__init__.py ->
build/lib/apex/contrib/layer_norm

```

```

    creating build/lib/apex/contrib/peer_memory
    copying apex/contrib/peer_memory/__init__.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_halo_exchanger_1d.py ->
build/lib/apex/contrib/peer_memory
    copying apex/contrib/peer_memory/peer_memory.py ->
build/lib/apex/contrib/peer_memory
    creating build/lib/apex/contrib/transducer
    copying apex/contrib/transducer/__init__.py ->
build/lib/apex/contrib/transducer
    copying apex/contrib/transducer/transducer.py ->
build/lib/apex/contrib/transducer
    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/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/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
    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/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
    creating build/bdist.linux-x86_64/wheel/apex/mlp
    copying build/lib/apex/mlp/__init__.py ->
build/bdist.linux-x86_64/wheel/apex/mlp
    copying build/lib/apex/mlp/mlp.py -> build/bdist.linux-x86_64/wheel/apex/mlp

```

```

    creating 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/scaler.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/rnn_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
    copying build/lib/apex/amp/utils.py -> build/bdist.linux-x86_64/wheel/apex/amp
    copying build/lib/apex/amp/__init__.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/__init__.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/torch_overrides.py ->
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/amp.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/wrap.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/_amp_state.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
    copying build/lib/apex/amp/handle.py ->
build/bdist.linux-x86_64/wheel/apex/amp
    copying build/lib/apex/__init__.py -> build/bdist.linux-x86_64/wheel/apex
    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
    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/parallel
    copying build/lib/apex/parallel/__init__.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/multiproc.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
    copying build/lib/apex/parallel/LARC.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/sync_batchnorm_kernel.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
    creating 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/__init__.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
    copying build/lib/apex/optimizers/fused_mixed_precision_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/fused_novograd.py ->
build/bdist.linux-x86_64/wheel/apex/optimizers
    copying build/lib/apex/optimizers/fused_lamb.py ->
build/bdist.linux-x86_64/wheel/apex/optimizers
    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/fp16util.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
    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
    copying build/lib/apex/transformer/microbatches.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
  copying build/lib/apex/transformer/parallel_state.py ->
build/bdist.linux-x86_64/wheel/apex/transformer
  copying build/lib/apex/transformer/utils.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/functional
  copying build/lib/apex/transformer/functional/__init__.py ->
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
  creating 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/pipeline_parallel/__init__.py ->
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
  creating
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/fwd_bwd_pipelining_with_interleaving.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/schedules/fwd_bwd_pipelining_without_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/_timers.py ->
build/bdist.linux-x86_64/wheel/apex/transformer/pipeline_parallel
  creating 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/layers/__init__.py ->
build/bdist.linux-x86_64/wheel/apex/transformer/layers
  creating 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/utils.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/mappings.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/cross_entropy.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/random.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/standalone_gpt.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/distributed_test_base.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/commons.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/standalone_transformer_lm.py ->
build/bdist.linux-x86_64/wheel/apex/transformer/testing
    copying build/lib/apex/transformer/enums.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/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
    creating build/bdist.linux-x86_64/wheel/apex/contrib
    creating 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/bottleneck_module_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
  copying build/lib/apex/contrib/bottleneck/bottleneck.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
  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
  copying build/lib/apex/contrib/__init__.py ->
build/bdist.linux-x86_64/wheel/apex/contrib
  creating 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/groupbn/batch_norm.py ->
build/bdist.linux-x86_64/wheel/apex/contrib/groupbn
  creating 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/encdec_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_self_multihead_attn_func.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
  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/self_multihead_attn.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_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
  creating 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
  copying build/lib/apex/contrib/index_mul_2d/index_mul_2d.py ->
build/bdist.linux-x86_64/wheel/apex/contrib/index_mul_2d
  creating 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/fused_sgd.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/fused_adam.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/distributed_fused_lamb.py ->
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
    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/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/sparsity
    copying build/lib/apex/contrib/sparsity/permutation_lib.py ->
build/bdist.linux-x86_64/wheel/apex/contrib/sparsity
    copying build/lib/apex/contrib/sparsity/__init__.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/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/sparse_masklib.py ->
build/bdist.linux-x86_64/wheel/apex/contrib/sparsity
    creating 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
  copying build/lib/apex/contrib/layer_norm/__init__.py ->
build/bdist.linux-x86_64/wheel/apex/contrib/layer_norm
  creating 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
  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_halo_exchanger_1d.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
  creating 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
  copying build/lib/apex/contrib/transducer/transducer.py ->
build/bdist.linux-x86_64/wheel/apex/contrib/transducer
  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/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
  copying build/lib/apex/_autocast_utils.py ->
build/bdist.linux-x86_64/wheel/apex
  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-jdb1celf/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=1cca4568e89ef21aa253d92b38c6a4ded58a057aeb7a7a19f75ec5d9daa4a3d8
Stored in directory: /tmp/pip-ephem-wheel-cache-
htj_0b2_/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-hbn967t5'
/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 32.2 MB/s
Collecting awscli
  Downloading awscli-1.25.47-py3-none-any.whl (3.9 MB)
    |                               | 3.9 MB 48.7 MB/s
Collecting colorama<0.4.5,>=0.2.5
  Downloading colorama-0.4.4-py2.py3-none-any.whl (16 kB)
Collecting s3transfer<0.7.0,>=0.6.0
  Downloading s3transfer-0.6.0-py3-none-any.whl (79 kB)
    |                               | 79 kB 6.7 MB/s
Collecting rsa<4.8,>=3.1.2
  Downloading rsa-4.7.2-py3-none-any.whl (34 kB)
Collecting docutils<0.17,>=0.10
  Downloading docutils-0.16-py2.py3-none-any.whl (548 kB)
    |                               | 548 kB 67.8 MB/s
Collecting boto3==1.27.47
  Downloading boto3-1.27.47-py3-none-any.whl (9.0 MB)
    |                               | 9.0 MB 46.8 MB/s
Requirement already satisfied: PyYAML<5.5,>=3.10 in
/usr/local/lib/python3.7/dist-packages (from awscli) (3.13)
Collecting jmespath<2.0.0,>=0.7.1

```

```
Downloading jmespath-1.0.1-py3-none-any.whl (20 kB)
Requirement already satisfied: python-dateutil<3.0.0,>=2.1 in
/usr/local/lib/python3.7/dist-packages (from botocore==1.27.47->awscli) (2.8.2)
Requirement already satisfied: six>=1.5 in /usr/local/lib/python3.7/dist-
packages (from python-dateutil<3.0.0,>=2.1->botocore==1.27.47->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-1.25.47 botocore-1.27.47 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: 982, done.
remote: Counting objects: 100% (52/52), done.
remote: Compressing objects: 100% (48/48), done.
remote: Total 982 (delta 16), reused 7 (delta 0), pack-reused 930
Receiving objects: 100% (982/982), 251.89 MiB | 17.35 MiB/s, done.
Resolving deltas: 100% (600/600), done.
Checking out files: 100% (1273/1273), 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/FMedium_50_1_2/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/
```

Downloading: 100% 481/481 [00:00<00:00, 505kB/s]

Downloading: 100% 899k/899k [00:00<00:00, 2.39MB/s]

Downloading: 100% 456k/456k [00:00<00:00, 1.80MB/s]

Downloading: 100% 501M/501M [00:10<00:00, 47.1MB/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: 0.4246775209903717
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=0.2692307692307693, f1=0.2666666666666667,
best_f1=0.2666666666666667
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.47920364141464233
/usr/local/lib/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=0.288659793814433, f1=0.2692307692307693,
best_f1=0.2692307692307693
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.41634470224380493
/usr/local/lib/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=0.35443037974683544, f1=0.28571428571428575,
best_f1=0.28571428571428575
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.26893535256385803
/usr/local/lib/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=0.4905660377358491, f1=0.3870967741935483,
best_f1=0.3870967741935483
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.3324672281742096
/usr/local/lib/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=0.5, f1=0.4067796610169491, best_f1=0.4067796610169491
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
warnings.warn("An input tensor was not cuda.")

```

```

step: 0, loss: 0.2114650309085846
/usr/local/lib/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=0.5090909090909091, f1=0.4210526315789474,
best_f1=0.4210526315789474
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.40469202399253845
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 7: dev_f1=0.6111111111111112, f1=0.5128205128205129,
best_f1=0.5128205128205129
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.418935626745224
/usr/local/lib/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=0.7857142857142857, f1=0.588235294117647,
best_f1=0.588235294117647
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.22870619595050812
/usr/local/lib/python3.7/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=0.75, f1=0.6923076923076924, best_f1=0.588235294117647
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")

```



```

step: 0, loss: 0.3659387230873108
/usr/local/lib/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=0.8333333333333333, f1=0.7142857142857143,
best_f1=0.7142857142857143
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.15962573885917664
/usr/local/lib/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=0.8461538461538461, f1=0.6666666666666666,
best_f1=0.6666666666666666
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.08004716038703918
/usr/local/lib/python3.7/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=0.8461538461538461, f1=0.7142857142857143,
best_f1=0.6666666666666666
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.06320689618587494
/usr/local/lib/python3.7/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=0.8333333333333333, f1=0.7142857142857143,
best_f1=0.6666666666666666
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.012955262325704098
/usr/local/lib/python3.7/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=0.8333333333333333, f1=0.7142857142857143,
best_f1=0.6666666666666666
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.022751232609152794
/usr/local/lib/python3.7/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=0.8333333333333333, f1=0.7142857142857143,
best_f1=0.6666666666666666
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, 110472.26it/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 = 0.8666666666666666
real_f1 = 0.8387096774193549
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, 193.51it/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: 0.5749607086181641
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: 0.46850907802581787
step: 20, loss: 0.4881693124771118
step: 30, loss: 0.33247241377830505
step: 40, loss: 0.3667101562023163
step: 50, loss: 0.6017252802848816
step: 60, loss: 0.4771202504634857
step: 70, loss: 0.3134121894836426
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 8192.0
step: 80, loss: 0.30139631032943726
step: 90, loss: 0.1567903310060501
step: 100, loss: 0.1754891276359558
step: 110, loss: 0.10042215883731842
step: 120, loss: 0.019791826605796814
step: 130, loss: 0.016893509775400162
step: 140, loss: 0.016371598467230797
step: 150, loss: 0.2144383192062378
step: 160, loss: 0.029869118705391884
step: 170, loss: 0.0361224040389061
step: 180, loss: 0.00593141233548522
step: 190, loss: 0.026628723368048668
step: 200, loss: 0.046193677932024
step: 210, loss: 0.00944427214562893
step: 220, loss: 0.10953821986913681
step: 230, loss: 0.0028549060225486755
/usr/local/lib/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=0.9474860335195531, f1=0.9503386004514672,
best_f1=0.9503386004514672

```

```

/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.039359938353300095
step: 10, loss: 0.15921860933303833
step: 20, loss: 0.03479389473795891
step: 30, loss: 0.009623296558856964
step: 40, loss: 0.015882203355431557
step: 50, loss: 0.006108604837208986
step: 60, loss: 0.002012949436903
step: 70, loss: 0.010435818694531918
step: 80, loss: 0.016093332320451736
step: 90, loss: 0.01686667464673519
step: 100, loss: 0.004625271074473858
step: 110, loss: 0.0018149052048102021
step: 120, loss: 0.07217506319284439
step: 130, loss: 0.00717606907710433
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 4096.0
step: 140, loss: 0.002941774670034647
step: 150, loss: 0.06971067190170288
step: 160, loss: 0.0019276937237009406
step: 170, loss: 0.008184879086911678
step: 180, loss: 0.0014377129264175892
step: 190, loss: 0.010182122699916363
step: 200, loss: 0.011403673328459263
step: 210, loss: 0.02603006176650524
step: 220, loss: 0.0022311736829578876
step: 230, loss: 0.0010596424108371139
/usr/local/lib/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=0.9865771812080537, f1=0.9841986455981941,
best_f1=0.9841986455981941
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.011491893790662289
step: 10, loss: 0.0018524796469137073
step: 20, loss: 0.02068229392170906
step: 30, loss: 0.00045445625437423587
step: 40, loss: 0.0008231078390963376
step: 50, loss: 0.125575989484787
step: 60, loss: 0.017027901485562325
step: 70, loss: 0.001315944828093052

```

```

step: 80, loss: 0.001280092285014689
step: 90, loss: 0.0013528355630114675
step: 100, loss: 0.0012068775249645114
step: 110, loss: 0.006118736229836941
step: 120, loss: 0.0005890278844162822
step: 130, loss: 0.005621536634862423
step: 140, loss: 0.0233791321516037
step: 150, loss: 0.0050691841170191765
step: 160, loss: 0.0011806858237832785
step: 170, loss: 0.0005101613351143897
step: 180, loss: 0.0013866238296031952
step: 190, loss: 0.11672782152891159
step: 200, loss: 0.034093957394361496
step: 210, loss: 0.0021342942491173744
step: 220, loss: 0.0037592395674437284
step: 230, loss: 0.005575776100158691
/usr/local/lib/python3.7/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=0.9841269841269841, f1=0.9794988610478361,
best_f1=0.9841986455981941
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.002391889924183488
step: 10, loss: 0.0018033395754173398
step: 20, loss: 0.0009095299756154418
step: 30, loss: 0.0004366959328763187
step: 40, loss: 0.0473574623465538
step: 50, loss: 0.0010431077098473907
step: 60, loss: 0.002116978634148836
step: 70, loss: 0.0008783339872024953
step: 80, loss: 0.0003721842367667705
step: 90, loss: 0.000929115922190249
step: 100, loss: 0.0005511687486432493
step: 110, loss: 0.00040994753362610936
step: 120, loss: 0.0007666167221032083
step: 130, loss: 0.013717906549572945
step: 140, loss: 0.001048875623382628
step: 150, loss: 0.0005482014385052025
step: 160, loss: 0.005349561106413603
step: 170, loss: 0.0013123378157615662
step: 180, loss: 0.12456031888723373
step: 190, loss: 0.001657150685787201
step: 200, loss: 0.0013862771447747946
step: 210, loss: 0.00038301452877931297
step: 220, loss: 0.000313088356051594
step: 230, loss: 0.0012603256618604064

```

```

/usr/local/lib/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=0.9888392857142857, f1=0.9854096520763187,
best_f1=0.9854096520763187
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0004122177779208869
step: 10, loss: 0.0010608364827930927
step: 20, loss: 0.0005736601306125522
step: 30, loss: 0.0005463029956445098
step: 40, loss: 0.0016830320237204432
step: 50, loss: 0.0011525126174092293
step: 60, loss: 0.012857920490205288
step: 70, loss: 0.0008982711588032544
step: 80, loss: 0.09284624457359314
step: 90, loss: 0.0416363887488842
step: 100, loss: 0.0005968048353679478
step: 110, loss: 0.0047010802663862705
step: 120, loss: 0.0012454735115170479
step: 130, loss: 0.021749893203377724
step: 140, loss: 0.00023877414059825242
step: 150, loss: 0.07245459407567978
step: 160, loss: 0.05145087465643883
step: 170, loss: 0.006063238251954317
step: 180, loss: 0.05455849692225456
step: 190, loss: 0.15559250116348267
step: 200, loss: 0.009829890914261341
step: 210, loss: 0.022847985848784447
step: 220, loss: 0.0013664179714396596
step: 230, loss: 0.005443754605948925
/usr/local/lib/python3.7/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=0.9820627802690582, f1=0.9819413092550789,
best_f1=0.9854096520763187
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0014138576807454228
step: 10, loss: 0.01813112571835518
step: 20, loss: 0.0018935460830107331
step: 30, loss: 0.0009518121369183064

```

```

step: 40, loss: 0.0004841394838877022
step: 50, loss: 0.00045530337956734
step: 60, loss: 0.008144647814333439
step: 70, loss: 0.0011564072920009494
step: 80, loss: 0.004247407428920269
step: 90, loss: 0.045359767973423004
step: 100, loss: 0.0006385858869180083
step: 110, loss: 0.012242245487868786
step: 120, loss: 0.0005540656275115907
step: 130, loss: 0.000584003864787519
step: 140, loss: 0.00032273941906169057
step: 150, loss: 0.012855594977736473
step: 160, loss: 0.04994453489780426
step: 170, loss: 0.00035542139085009694
step: 180, loss: 0.00012574611173477024
step: 190, loss: 0.00021636734891217202
step: 200, loss: 0.006695536896586418
step: 210, loss: 0.000449707149527967
step: 220, loss: 0.01944686844944954
step: 230, loss: 0.0018084825715050101
/usr/local/lib/python3.7/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=0.987736900780379, f1=0.9810055865921787,
best_f1=0.9854096520763187
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.002641086233779788
step: 10, loss: 0.0013663964346051216
step: 20, loss: 0.0006558587774634361
step: 30, loss: 0.0003140756452921778
step: 40, loss: 0.006934059783816338
step: 50, loss: 0.0008339554769918323
step: 60, loss: 0.0006352144409902394
step: 70, loss: 0.001360758556984365
step: 80, loss: 0.0005287437234073877
step: 90, loss: 0.0009046524064615369
step: 100, loss: 0.0012762820115312934
step: 110, loss: 0.0008001212845556438
step: 120, loss: 0.002484228229150176
step: 130, loss: 0.003765497822314501
step: 140, loss: 0.0007617758237756789
step: 150, loss: 0.011167776770889759
step: 160, loss: 0.00039085137541405857
step: 170, loss: 0.00383256608620286
step: 180, loss: 0.0008226247155107558
step: 190, loss: 0.0003382998111192137

```



```

step: 200, loss: 0.0031323556322604418
step: 210, loss: 0.018907401710748672
step: 220, loss: 0.0035078763030469418
step: 230, loss: 0.0010915284510701895
/usr/local/lib/python3.7/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=0.9843749999999999, f1=0.9709821428571428,
best_f1=0.9854096520763187
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0010825308272615075
step: 10, loss: 0.004788877908140421
step: 20, loss: 0.0030097232665866613
step: 30, loss: 0.004795734770596027
step: 40, loss: 0.002004623180255294
step: 50, loss: 0.07714107632637024
step: 60, loss: 0.0034655805211514235
step: 70, loss: 0.0005395955522544682
step: 80, loss: 0.007911363616585732
step: 90, loss: 0.0017501743277534842
step: 100, loss: 0.0009460980072617531
step: 110, loss: 0.002895552897825837
step: 120, loss: 0.000487611599965021
step: 130, loss: 0.0009787912713363767
step: 140, loss: 0.0004401872574817389
step: 150, loss: 0.06553880125284195
step: 160, loss: 0.019189540296792984
step: 170, loss: 0.043698374181985855
step: 180, loss: 0.0005674316780641675
step: 190, loss: 0.02004486881196499
step: 200, loss: 0.10437820851802826
step: 210, loss: 0.00466080941259861
step: 220, loss: 0.0014410995645448565
step: 230, loss: 0.0011073511559516191
/usr/local/lib/python3.7/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=0.9888392857142857, f1=0.9842696629213483,
best_f1=0.9854096520763187
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.003127569332718849
step: 10, loss: 0.0018240626668557525
step: 20, loss: 0.004555943887680769
step: 30, loss: 0.0018184434156864882

```

```

step: 40, loss: 0.001262512058019638
step: 50, loss: 0.0022873857524245977
step: 60, loss: 0.0009763341513462365
step: 70, loss: 0.03994136303663254
step: 80, loss: 0.0008102519786916673
step: 90, loss: 0.04124089702963829
step: 100, loss: 0.000563468667678535
step: 110, loss: 0.004982004873454571
step: 120, loss: 0.012612714432179928
step: 130, loss: 0.0008648114162497222
step: 140, loss: 0.0008576103718951344
step: 150, loss: 0.000941912061534822
step: 160, loss: 0.11776172369718552
step: 170, loss: 0.0010207444429397583
step: 180, loss: 0.019943414255976677
step: 190, loss: 0.00043674366315826774
step: 200, loss: 0.00043662727694027126
step: 210, loss: 0.018589498475193977
step: 220, loss: 0.000755196379031986
step: 230, loss: 0.0002565741306170821
/usr/local/lib/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=0.9898989898989898, f1=0.9854096520763187,
best_f1=0.9854096520763187
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0011646943166851997
step: 10, loss: 0.0008354673627763987
step: 20, loss: 0.0007002146448940039
step: 30, loss: 0.0003369509940966964
step: 40, loss: 0.002132667461410165
step: 50, loss: 0.00032884610118344426
step: 60, loss: 0.0011027005966752768
step: 70, loss: 0.0029693145770579576
step: 80, loss: 0.0007858749013394117
step: 90, loss: 0.0005348837585188448
step: 100, loss: 0.0004555688938125968
step: 110, loss: 0.0006405137246474624
step: 120, loss: 0.0002933098585344851
step: 130, loss: 0.0006403704173862934
step: 140, loss: 0.00045555937686003745
step: 150, loss: 0.0017310632392764091

```

```

step: 160, loss: 0.000244354538153857
step: 170, loss: 0.00044579882523976266
step: 180, loss: 0.0016552135348320007
step: 190, loss: 0.0007036784663796425
step: 200, loss: 0.0012967863585799932
step: 210, loss: 0.002943645231425762
step: 220, loss: 0.0008318835170939565
step: 230, loss: 0.000266330927843228
/usr/local/lib/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=0.9899665551839464, f1=0.9766925638179801,
best_f1=0.9766925638179801
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00018158349848818034
step: 10, loss: 0.0005067270831204951
step: 20, loss: 0.0004088840796612203
step: 30, loss: 0.0003647050471045077
step: 40, loss: 0.015296838246285915
step: 50, loss: 0.0001514881441835314
step: 60, loss: 0.003173881908878684
step: 70, loss: 0.0002850172168109566
step: 80, loss: 0.0008256282308138907
step: 90, loss: 0.2078460454940796
step: 100, loss: 0.000865673297084868
step: 110, loss: 0.0005826579872518778
step: 120, loss: 0.0007244347361847758
step: 130, loss: 0.0008579139830544591
step: 140, loss: 0.00677891168743372
step: 150, loss: 0.0008423206163570285
step: 160, loss: 0.0012624820228666067
step: 170, loss: 0.0009016256663016975
step: 180, loss: 0.0011807222617790103
step: 190, loss: 0.00080782052827999
step: 200, loss: 0.025475259870290756
step: 210, loss: 0.0010011357953771949
step: 220, loss: 0.0015283881220966578
step: 230, loss: 0.0016879182076081634
/usr/local/lib/python3.7/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=0.9888641425389755, f1=0.9777777777777777,

```

```

best_f1=0.9766925638179801
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0010439383331686258
step: 10, loss: 0.0008032218320295215
step: 20, loss: 0.009155446663498878
step: 30, loss: 0.03918091952800751
step: 40, loss: 0.0014421433443203568
step: 50, loss: 0.007322038058191538
step: 60, loss: 0.0011096384841948748
step: 70, loss: 0.0010311234509572387
step: 80, loss: 0.0005366962868720293
step: 90, loss: 0.0028757101390510798
step: 100, loss: 0.0006357422098517418
step: 110, loss: 0.00040390630601905286
step: 120, loss: 0.0006690098671242595
step: 130, loss: 0.0007641787524335086
step: 140, loss: 0.0018382417038083076
step: 150, loss: 0.0018224369268864393
step: 160, loss: 0.002098604803904891
step: 170, loss: 0.0017038461519405246
step: 180, loss: 0.0010176405776292086
step: 190, loss: 0.0030626626685261726
step: 200, loss: 0.0009553136769682169
step: 210, loss: 0.000653458759188652
step: 220, loss: 0.00499627785757184
step: 230, loss: 0.0009705123375169933
/usr/local/lib/python3.7/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=0.9888392857142857, f1=0.975609756097561,
best_f1=0.9766925638179801
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0009064821642823517
step: 10, loss: 0.0014831981388852
step: 20, loss: 0.0012262002564966679
step: 30, loss: 0.0022044831421226263
step: 40, loss: 0.0011590187205001712
step: 50, loss: 0.004461321979761124
step: 60, loss: 0.0007047895342111588
step: 70, loss: 0.0006067253998480737
step: 80, loss: 0.0012450410285964608
step: 90, loss: 0.0006319336243905127
step: 100, loss: 0.002722573932260275
step: 110, loss: 0.00103367876727134

```

```

step: 120, loss: 0.0010633219499140978
step: 130, loss: 0.0009093428961932659
step: 140, loss: 0.009741341695189476
step: 150, loss: 0.0005354938912205398
step: 160, loss: 0.02215294912457466
step: 170, loss: 0.0012819922994822264
step: 180, loss: 0.038765184581279755
step: 190, loss: 0.0009244114626199007
step: 200, loss: 0.0002377719501964748
step: 210, loss: 0.008508148603141308
step: 220, loss: 0.0008930545882321894
step: 230, loss: 0.0008512748754583299
/usr/local/lib/python3.7/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=0.984304932735426,
best_f1=0.9766925638179801
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0006150684785097837
step: 10, loss: 0.0005661786417476833
step: 20, loss: 0.0008034311467781663
step: 30, loss: 0.0012053161626681685
step: 40, loss: 0.0009978715097531676
step: 50, loss: 0.0004850525874644518
step: 60, loss: 0.0007443802896887064
step: 70, loss: 0.0006105101201683283
step: 80, loss: 0.0005215138662606478
step: 90, loss: 0.002038378268480301
step: 100, loss: 0.0009986805962398648
step: 110, loss: 0.0012081742752343416
step: 120, loss: 0.00024060223950073123
step: 130, loss: 0.0016002599149942398
step: 140, loss: 0.0005862759426236153
step: 150, loss: 0.0002534749801270664
step: 160, loss: 0.0009425480966456234
step: 170, loss: 0.0005900856922380626
step: 180, loss: 0.0005571925430558622
step: 190, loss: 0.0004797503352165222
step: 200, loss: 0.04705464839935303
step: 210, loss: 0.0005548528279177845
step: 220, loss: 0.0008348909905180335
step: 230, loss: 0.0003847593907266855
/usr/local/lib/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=0.9910714285714286, f1=0.9810479375696767,
best_f1=0.9810479375696767
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
```

```
warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.003657794091850519
step: 10, loss: 0.0008247914374805987
step: 20, loss: 0.0014788296539336443
step: 30, loss: 0.0007185635040514171
step: 40, loss: 0.00039466694579459727
step: 50, loss: 0.000491255079396069
step: 60, loss: 0.04100343957543373
step: 70, loss: 0.0008057819213718176
step: 80, loss: 0.0004293776291888207
step: 90, loss: 0.0004183083074167371
step: 100, loss: 0.00039798414218239486
step: 110, loss: 0.0006196595495566726
step: 120, loss: 0.029335565865039825
step: 130, loss: 0.0009146287338808179
step: 140, loss: 0.010814890265464783
step: 150, loss: 0.0008352791192010045
step: 160, loss: 0.007682515308260918
step: 170, loss: 0.0003435051185078919
step: 180, loss: 0.0009881331352517009
step: 190, loss: 0.005649772938340902
step: 200, loss: 0.0011936441296711564
step: 210, loss: 0.14563481509685516
step: 220, loss: 0.0006773225031793118
step: 230, loss: 0.0007147907163016498
/usr/local/lib/python3.7/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=0.9910714285714286, f1=0.9832026875699889,
best_f1=0.9810479375696767
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, 197.42it/s]
load_f1 = 0.9910514541387023
real_f1 = 0.987709497206704
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, 181.24it/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: 0.6195750832557678
```

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: 0.38829389214515686
```

```
step: 20, loss: 0.3180695176124573
```

```
step: 30, loss: 0.2854365408420563
```

```
step: 40, loss: 0.23946513235569
```



step: 50, loss: 0.26806050539016724  
step: 60, loss: 0.2350151240825653  
step: 70, loss: 0.2107764631509781  
step: 80, loss: 0.11545302718877792  
step: 90, loss: 0.1774236559867859  
step: 100, loss: 0.1574803590774536  
step: 110, loss: 0.11532028019428253  
step: 120, loss: 0.0959954783320427  
step: 130, loss: 0.18213242292404175  
step: 140, loss: 0.29258987307548523  
step: 150, loss: 0.10148898512125015  
step: 160, loss: 0.13887456059455872  
step: 170, loss: 0.02861957997083664  
step: 180, loss: 0.10345789045095444  
step: 190, loss: 0.06777550280094147  
step: 200, loss: 0.04701218381524086  
step: 210, loss: 0.05442056804895401  
step: 220, loss: 0.0840514749288559  
step: 230, loss: 0.35393282771110535  
step: 240, loss: 0.07250714302062988  
step: 250, loss: 0.05303320661187172  
step: 260, loss: 0.4171780049800873  
step: 270, loss: 0.17428551614284515  
step: 280, loss: 0.041784901171922684  
step: 290, loss: 0.0592411644756794  
step: 300, loss: 0.10072270780801773  
step: 310, loss: 0.10440503805875778  
step: 320, loss: 0.08218853175640106  
step: 330, loss: 0.08539004623889923  
step: 340, loss: 0.3183068335056305  
step: 350, loss: 0.09827771037817001  
step: 360, loss: 0.05853777006268501  
step: 370, loss: 0.05657683312892914  
step: 380, loss: 0.13621315360069275  
step: 390, loss: 0.010775083675980568  
step: 400, loss: 0.013357091695070267  
step: 410, loss: 0.26296260952949524  
step: 420, loss: 0.006587633863091469  
step: 430, loss: 0.08699800819158554  
step: 440, loss: 0.09320119023323059  
step: 450, loss: 0.01375657320022583  
step: 460, loss: 0.03007686324417591  
step: 470, loss: 0.04058968275785446  
step: 480, loss: 0.13403400778770447  
step: 490, loss: 0.1138634979724884  
step: 500, loss: 0.03143114969134331  
step: 510, loss: 0.044840727001428604  
step: 520, loss: 0.10868675261735916

```

step: 530, loss: 0.04951345548033714
/usr/local/lib/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=0.9397701149425287, f1=0.9375, best_f1=0.9375
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.04418318718671799
step: 10, loss: 0.01077682338654995
step: 20, loss: 0.033372845500707626
step: 30, loss: 0.04482161998748779
step: 40, loss: 0.059944819658994675
step: 50, loss: 0.060861047357320786
step: 60, loss: 0.0239575132727623
step: 70, loss: 0.015069874003529549
step: 80, loss: 0.0025747367180883884
step: 90, loss: 0.02231748215854168
step: 100, loss: 0.09061697125434875
step: 110, loss: 0.00821329653263092
step: 120, loss: 0.17619206011295319
step: 130, loss: 0.0026532390620559454
step: 140, loss: 0.05626025050878525
step: 150, loss: 0.0045733824372291565
step: 160, loss: 0.024099450558423996
step: 170, loss: 0.03793139383196831
step: 180, loss: 0.02706124633550644
step: 190, loss: 0.029204223304986954
step: 200, loss: 0.19023315608501434
step: 210, loss: 0.012335681356489658
step: 220, loss: 0.0034230404999107122
step: 230, loss: 0.009067919105291367
step: 240, loss: 0.023455237969756126
step: 250, loss: 0.022619834169745445
step: 260, loss: 0.058058496564626694
step: 270, loss: 0.016743319109082222
step: 280, loss: 0.0655268132686615
step: 290, loss: 0.01788077875971794
step: 300, loss: 0.025173643603920937
step: 310, loss: 0.015940073877573013
step: 320, loss: 0.055894412100315094
step: 330, loss: 0.07008296996355057
step: 340, loss: 0.14381136000156403
step: 350, loss: 0.0054294653236866

```

```

step: 360, loss: 0.11554902791976929
step: 370, loss: 0.01243690587580204
step: 380, loss: 0.1466854363679886
step: 390, loss: 0.002363581210374832
step: 400, loss: 0.03257618844509125
step: 410, loss: 0.015931615605950356
step: 420, loss: 0.0065810359083116055
step: 430, loss: 0.0863334983587265
step: 440, loss: 0.014184548519551754
step: 450, loss: 0.04397467523813248
step: 460, loss: 0.023547643795609474
step: 470, loss: 0.1465912163257599
step: 480, loss: 0.02190978452563286
step: 490, loss: 0.04171859100461006
step: 500, loss: 0.007143421098589897
step: 510, loss: 0.03536178916692734
step: 520, loss: 0.4011705219745636
step: 530, loss: 0.11365282535552979
/usr/local/lib/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=0.948339483394834, f1=0.9465930018416207,
best_f1=0.9465930018416207
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.09861103445291519
step: 10, loss: 0.045668892562389374
step: 20, loss: 0.003755461424589157
step: 30, loss: 0.017586147412657738
step: 40, loss: 0.07451201975345612
step: 50, loss: 0.13184668123722076
step: 60, loss: 0.026927752420306206
step: 70, loss: 0.007242362480610609
step: 80, loss: 0.0022750271018594503
step: 90, loss: 0.0021483516320586205
step: 100, loss: 0.14340654015541077
step: 110, loss: 0.03573586791753769
step: 120, loss: 0.07263118028640747
step: 130, loss: 0.03601580858230591
step: 140, loss: 0.02191873826086521
step: 150, loss: 0.013111701235175133
step: 160, loss: 0.023603232577443123
step: 170, loss: 0.005901306867599487

```

```

step: 180, loss: 0.0066317967139184475
step: 190, loss: 0.0027615604922175407
step: 200, loss: 0.005343063268810511
step: 210, loss: 0.012098330073058605
step: 220, loss: 0.1714666783809662
step: 230, loss: 0.027462000027298927
step: 240, loss: 0.07877866923809052
step: 250, loss: 0.12283655256032944
step: 260, loss: 0.03712173178792
step: 270, loss: 0.003841045079752803
step: 280, loss: 0.01660451479256153
step: 290, loss: 0.003598500043153763
step: 300, loss: 0.10588007420301437
step: 310, loss: 0.01833556778728962
step: 320, loss: 0.056773193180561066
step: 330, loss: 0.06820520013570786
step: 340, loss: 0.0019566104747354984
step: 350, loss: 0.0635165274143219
step: 360, loss: 0.005522614344954491
step: 370, loss: 0.009383146651089191
step: 380, loss: 0.026748230680823326
step: 390, loss: 0.004793671425431967
step: 400, loss: 0.053563814610242844
step: 410, loss: 0.026450442150235176
step: 420, loss: 0.00498968455940485
step: 430, loss: 0.02735324203968048
step: 440, loss: 0.14162589609622955
step: 450, loss: 0.05119911953806877
step: 460, loss: 0.06776376068592072
step: 470, loss: 0.008375919423997402
step: 480, loss: 0.046370942145586014
step: 490, loss: 0.016290392726659775
step: 500, loss: 0.015726299956440926
step: 510, loss: 0.03823477402329445
step: 520, loss: 0.024492798373103142
step: 530, loss: 0.021350255236029625
/usr/local/lib/python3.7/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=0.936636493568366, f1=0.924618320610687,
best_f1=0.9465930018416207
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.013868561945855618
step: 10, loss: 0.006749461404979229
step: 20, loss: 0.023767784237861633
step: 30, loss: 0.006939297541975975

```

step: 40, loss: 0.01313658058643341  
step: 50, loss: 0.06783848255872726  
step: 60, loss: 0.0048400601372122765  
step: 70, loss: 0.004274267703294754  
step: 80, loss: 0.0010107702109962702  
step: 90, loss: 0.010934142395853996  
step: 100, loss: 0.0007852931739762425  
step: 110, loss: 0.0020359796471893787  
step: 120, loss: 0.0014818159397691488  
step: 130, loss: 0.013647424057126045  
step: 140, loss: 0.01086380984634161  
step: 150, loss: 0.011842883192002773  
step: 160, loss: 0.0020314918365329504  
step: 170, loss: 0.02550419047474861  
step: 180, loss: 0.06183188036084175  
step: 190, loss: 0.0752398744225502  
step: 200, loss: 0.036042820662260056  
step: 210, loss: 0.0006472073728218675  
step: 220, loss: 0.016508957371115685  
step: 230, loss: 0.004695782903581858  
step: 240, loss: 0.0044116792269051075  
step: 250, loss: 0.08387573063373566  
step: 260, loss: 0.0017916675424203277  
step: 270, loss: 0.012840809300541878  
step: 280, loss: 0.0015744552947580814  
step: 290, loss: 0.01141241192817688  
step: 300, loss: 0.001432837569154799  
step: 310, loss: 0.04906873777508736  
step: 320, loss: 0.04061371460556984  
step: 330, loss: 0.02430754527449608  
step: 340, loss: 0.002304907189682126  
step: 350, loss: 0.08500571548938751  
step: 360, loss: 0.009120822884142399  
step: 370, loss: 0.006842937786132097  
step: 380, loss: 0.0014214471448212862  
step: 390, loss: 0.002748647006228566  
step: 400, loss: 0.010783931240439415  
step: 410, loss: 0.004605782683938742  
step: 420, loss: 0.02916596829891205  
step: 430, loss: 0.03008342906832695  
step: 440, loss: 0.026318863034248352  
step: 450, loss: 0.012912783771753311  
step: 460, loss: 0.020026493817567825  
step: 470, loss: 0.001209118403494358  
step: 480, loss: 0.009413805790245533  
step: 490, loss: 0.001186465029604733  
step: 500, loss: 0.04135559871792793  
step: 510, loss: 0.12716490030288696

```

step: 520, loss: 0.12741152942180634
step: 530, loss: 0.11983725428581238
/usr/local/lib/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=0.9495327102803738, f1=0.9496738117427772,
best_f1=0.9496738117427772
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.004828718490898609
step: 10, loss: 0.010370386764407158
step: 20, loss: 0.0016652728663757443
step: 30, loss: 0.0037784716114401817
step: 40, loss: 0.0004705647297669202
step: 50, loss: 0.028088200837373734
step: 60, loss: 0.03197770565748215
step: 70, loss: 0.07989158481359482
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 80, loss: 0.0002706082595977932
step: 90, loss: 0.008069940842688084
step: 100, loss: 0.1347769796848297
step: 110, loss: 0.0017226701602339745
step: 120, loss: 0.06257607787847519
step: 130, loss: 0.009422523900866508
step: 140, loss: 0.007926706224679947
step: 150, loss: 0.0035203187726438046
step: 160, loss: 0.017584064975380898
step: 170, loss: 0.029857521876692772
step: 180, loss: 0.004885590635240078
step: 190, loss: 0.002478365320712328
step: 200, loss: 0.0035888468846678734
step: 210, loss: 0.0026602260768413544
step: 220, loss: 0.009524616412818432
step: 230, loss: 0.004727265797555447
step: 240, loss: 0.035636939108371735
step: 250, loss: 0.06381992995738983
step: 260, loss: 0.0008637248538434505
step: 270, loss: 0.014435887336730957
step: 280, loss: 0.002846041927114129
step: 290, loss: 0.03222327679395676
step: 300, loss: 0.003798709250986576
step: 310, loss: 0.02155149169266224
step: 320, loss: 0.0022985690739005804

```

```

step: 330, loss: 0.005142589099705219
step: 340, loss: 0.00022823632752988487
step: 350, loss: 0.00017041164392139763
step: 360, loss: 0.0009351901244372129
step: 370, loss: 0.0006851719808764756
step: 380, loss: 0.000494864652864635
step: 390, loss: 0.16114583611488342
step: 400, loss: 0.00240715267136693
step: 410, loss: 0.08420846611261368
step: 420, loss: 0.16627348959445953
step: 430, loss: 0.0649028941988945
step: 440, loss: 0.004660909995436668
step: 450, loss: 0.005890071392059326
step: 460, loss: 0.0010647211456671357
step: 470, loss: 0.0047214641235768795
step: 480, loss: 0.015430866740643978
step: 490, loss: 0.00807873997837305
step: 500, loss: 0.001701834611594677
step: 510, loss: 0.0019665330182760954
step: 520, loss: 0.035503000020980835
step: 530, loss: 0.0022471253760159016
/usr/local/lib/python3.7/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=0.947022972339428, f1=0.9471221338324755,
best_f1=0.9496738117427772
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.042691558599472046
step: 10, loss: 0.00021706051484216005
step: 20, loss: 0.0001952366583282128
step: 30, loss: 0.002838411368429661
step: 40, loss: 0.027339160442352295
step: 50, loss: 0.00023792219872120768
step: 60, loss: 0.00046535764704458416
step: 70, loss: 0.00012037142005283386
step: 80, loss: 0.00018496070697437972
step: 90, loss: 0.006512483116239309
step: 100, loss: 0.07109612226486206
step: 110, loss: 0.009273515082895756
step: 120, loss: 0.0966082215309143
step: 130, loss: 0.01470737624913454
step: 140, loss: 0.0015284842811524868
step: 150, loss: 0.0025299149565398693
step: 160, loss: 0.009758105501532555
step: 170, loss: 0.00034765651798807085
step: 180, loss: 0.0018886367324739695

```

```

step: 190, loss: 0.10327029973268509
step: 200, loss: 0.0020483750849962234
step: 210, loss: 0.01196528784930706
step: 220, loss: 0.0008376715704798698
step: 230, loss: 0.007083740551024675
step: 240, loss: 0.0037496881559491158
step: 250, loss: 0.055335719138383865
step: 260, loss: 0.00043686703429557383
step: 270, loss: 0.0017618633573874831
step: 280, loss: 0.008035135455429554
step: 290, loss: 7.739800639683381e-05
step: 300, loss: 0.0018306886777281761
step: 310, loss: 0.011305023916065693
step: 320, loss: 9.427618351764977e-05
step: 330, loss: 0.0006701161619275808
step: 340, loss: 0.00045807810965925455
step: 350, loss: 0.002167181111872196
step: 360, loss: 0.5373721718788147
step: 370, loss: 0.05726729333400726
step: 380, loss: 0.003843617858365178
step: 390, loss: 0.011269516311585903
step: 400, loss: 0.001624679658561945
step: 410, loss: 0.0014548507751896977
step: 420, loss: 0.0027017726097255945
step: 430, loss: 0.00154875498265028
step: 440, loss: 0.000678412732668221
step: 450, loss: 0.2177598476409912
step: 460, loss: 0.001033721142448485
step: 470, loss: 0.002017016988247633
step: 480, loss: 0.0035404085647314787
step: 490, loss: 0.0027604857459664345
step: 500, loss: 0.07179655134677887
step: 510, loss: 0.13966286182403564
step: 520, loss: 0.002893150318413973
step: 530, loss: 0.0018584574572741985
/usr/local/lib/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=0.9499072356215214, f1=0.9480519480519481,
best_f1=0.9480519480519481
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.008115533739328384

```



step: 10, loss: 0.0012356216320767999  
step: 20, loss: 0.0018830305198207498  
step: 30, loss: 0.0013937752228230238  
step: 40, loss: 0.000522955204360187  
step: 50, loss: 0.0036985771730542183  
step: 60, loss: 0.02837488241493702  
step: 70, loss: 0.0010746610350906849  
step: 80, loss: 0.00035958102671429515  
step: 90, loss: 0.0001564528065500781  
step: 100, loss: 0.0032387583050876856  
step: 110, loss: 0.0008433608454652131  
step: 120, loss: 0.00037574514863081276  
step: 130, loss: 0.00015679514035582542  
step: 140, loss: 0.0031298529356718063  
step: 150, loss: 0.008625196292996407  
step: 160, loss: 0.0005136434338055551  
step: 170, loss: 0.03756881132721901  
step: 180, loss: 0.0024655216839164495  
step: 190, loss: 0.010038180276751518  
step: 200, loss: 0.0001875757152447477  
step: 210, loss: 0.0004905639798380435  
step: 220, loss: 0.0011440524831414223  
step: 230, loss: 0.0007233688957057893  
step: 240, loss: 0.011309699155390263  
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 8192.0  
step: 250, loss: 0.0016697775572538376  
step: 260, loss: 0.0013453905703499913  
step: 270, loss: 0.00017885844863485545  
step: 280, loss: 0.00416909996420145  
step: 290, loss: 0.0024572948459535837  
step: 300, loss: 0.00033752553281374276  
step: 310, loss: 0.0006765780854038894  
step: 320, loss: 0.0011396161280572414  
step: 330, loss: 0.0003331278567202389  
step: 340, loss: 0.00037073384737595916  
step: 350, loss: 0.0022324789315462112  
step: 360, loss: 0.0037210409063845873  
step: 370, loss: 0.008884641341865063  
step: 380, loss: 0.011585624888539314  
step: 390, loss: 0.00374084385111928  
step: 400, loss: 0.009017900563776493  
step: 410, loss: 0.0018541074823588133  
step: 420, loss: 0.04876028373837471  
step: 430, loss: 0.0011817277409136295  
step: 440, loss: 0.021017378196120262  
step: 450, loss: 0.00045298662735149264  
step: 460, loss: 0.00019035635341424495  
step: 470, loss: 0.11863642185926437

```

step: 480, loss: 0.007507176138460636
step: 490, loss: 0.004768256098031998
step: 500, loss: 0.0009924032492563128
step: 510, loss: 0.001269471482373774
step: 520, loss: 0.004771859385073185
step: 530, loss: 0.0005451340693980455
/usr/local/lib/python3.7/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=0.9460966542750929, f1=0.9406307977736549,
best_f1=0.9480519480519481
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0017049560556188226
step: 10, loss: 0.0019402399193495512
step: 20, loss: 0.01670052297413349
step: 30, loss: 0.0007115743937902153
step: 40, loss: 7.1160146035254e-05
step: 50, loss: 0.00039753393502905965
step: 60, loss: 0.003645066637545824
step: 70, loss: 0.004084846470504999
step: 80, loss: 0.008193192072212696
step: 90, loss: 0.0019775358960032463
step: 100, loss: 0.00016937602777034044
step: 110, loss: 9.322859114035964e-05
step: 120, loss: 0.003645936492830515
step: 130, loss: 0.0005409144796431065
step: 140, loss: 9.166409290628508e-05
step: 150, loss: 0.0003088700177613646
step: 160, loss: 8.749312837608159e-05
step: 170, loss: 0.1792164146900177
step: 180, loss: 0.0006916804704815149
step: 190, loss: 0.003214142983779311
step: 200, loss: 0.0036865875590592623
step: 210, loss: 0.0754641517996788
step: 220, loss: 0.00041799352038651705
step: 230, loss: 0.02392565831542015
step: 240, loss: 0.013042602688074112
step: 250, loss: 0.02852877415716648
step: 260, loss: 0.00019911346316803247
step: 270, loss: 0.0049168081022799015
step: 280, loss: 0.0656610056757927
step: 290, loss: 0.0005546347238123417
step: 300, loss: 8.142797014443204e-05
step: 310, loss: 0.008461005054414272
step: 320, loss: 0.0008265807409770787
step: 330, loss: 0.0011074991198256612

```

```

step: 340, loss: 0.023165859282016754
step: 350, loss: 5.6719542044447735e-05
step: 360, loss: 0.011618077754974365
step: 370, loss: 0.0644153580069542
step: 380, loss: 0.0001834808790590614
step: 390, loss: 0.000881669286172837
step: 400, loss: 0.002305038506165147
step: 410, loss: 0.042109180241823196
step: 420, loss: 0.0017960109980776906
step: 430, loss: 0.0009667086997069418
step: 440, loss: 0.0008792859152890742
step: 450, loss: 0.000246581737883389
step: 460, loss: 0.0012438218109309673
step: 470, loss: 0.034215714782476425
step: 480, loss: 0.001903533935546875
step: 490, loss: 0.0016529569402337074
step: 500, loss: 0.00011618631106102839
step: 510, loss: 0.00026236107805743814
step: 520, loss: 7.866132364142686e-05
step: 530, loss: 2.894861245295033e-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=0.9459840300610615, f1=0.9437883797827114,
best_f1=0.9480519480519481
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 7.203970017144457e-05
step: 10, loss: 0.0011167289922013879
step: 20, loss: 0.00011762724898289889
step: 30, loss: 0.06953306496143341
step: 40, loss: 4.253941006027162e-05
step: 50, loss: 0.0004283515445422381
step: 60, loss: 0.00030567546491511166
step: 70, loss: 0.00805695727467537
step: 80, loss: 0.0015020680148154497
step: 90, loss: 0.07082222402095795
step: 100, loss: 8.075460937106982e-05
step: 110, loss: 0.0022442550398409367
step: 120, loss: 6.10781426075846e-05
step: 130, loss: 9.047582716448233e-05
step: 140, loss: 6.916038546478376e-05
step: 150, loss: 0.03527715429663658
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 4096.0
step: 160, loss: 0.03421415761113167
step: 170, loss: 0.014774585142731667
step: 180, loss: 0.04330837354063988

```

```

step: 190, loss: 0.0072081321850419044
step: 200, loss: 0.003312890650704503
step: 210, loss: 0.0038483846001327038
step: 220, loss: 0.00044459704076871276
step: 230, loss: 0.0001911245344672352
step: 240, loss: 0.0033181544858962297
step: 250, loss: 0.004341404419392347
step: 260, loss: 4.316939157433808e-05
step: 270, loss: 0.0005373925087042153
step: 280, loss: 0.007333206478506327
step: 290, loss: 3.149231997667812e-05
step: 300, loss: 0.000931727874558419
step: 310, loss: 0.002766784280538559
step: 320, loss: 0.000328754133079201
step: 330, loss: 0.00011229618394281715
step: 340, loss: 8.180601435014978e-05
step: 350, loss: 0.0014451071619987488
step: 360, loss: 0.07142665982246399
step: 370, loss: 5.7183613535016775e-05
step: 380, loss: 0.17247669398784637
step: 390, loss: 1.5090850865817629e-05
step: 400, loss: 0.3672911822795868
step: 410, loss: 0.036210644990205765
step: 420, loss: 0.0009024759056046605
step: 430, loss: 0.003074999898672104
step: 440, loss: 8.341103239217773e-05
step: 450, loss: 0.019735299050807953
step: 460, loss: 0.0004289963690098375
step: 470, loss: 0.0006663970416411757
step: 480, loss: 0.0002876341750379652
step: 490, loss: 0.0004618423117790371
step: 500, loss: 0.004309126641601324
step: 510, loss: 0.00042629503877833486
step: 520, loss: 0.0013471016427502036
step: 530, loss: 0.01011730171740055
/usr/local/lib/python3.7/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=0.9494199535962878, f1=0.9478060046189377,
best_f1=0.9480519480519481
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0020805061794817448
step: 10, loss: 0.0025735311210155487
step: 20, loss: 0.0001216834643855691
step: 30, loss: 0.0011287289671599865
step: 40, loss: 6.453658716054633e-05

```

step: 50, loss: 0.00025596169871278107  
step: 60, loss: 0.0001468222471885383  
step: 70, loss: 8.762624929659069e-05  
step: 80, loss: 7.633084169356152e-05  
step: 90, loss: 9.245140245184302e-05  
step: 100, loss: 0.0019040191546082497  
step: 110, loss: 0.01927047222852707  
step: 120, loss: 5.860928649781272e-05  
step: 130, loss: 0.00013597190263681114  
step: 140, loss: 0.00044964029802940786  
step: 150, loss: 0.0009914110414683819  
step: 160, loss: 0.05610959976911545  
step: 170, loss: 0.00020316033624112606  
step: 180, loss: 0.0006139242323115468  
step: 190, loss: 0.000896351644769311  
step: 200, loss: 0.004673956427723169  
step: 210, loss: 0.009700175374746323  
step: 220, loss: 0.002895104233175516  
step: 230, loss: 0.00011702445044647902  
step: 240, loss: 0.00010437016317155212  
step: 250, loss: 0.000989567837677896  
step: 260, loss: 0.0016534858150407672  
step: 270, loss: 0.00015256059123203158  
step: 280, loss: 0.02145696058869362  
step: 290, loss: 5.1375507609918714e-05  
step: 300, loss: 5.3763902542414144e-05  
step: 310, loss: 0.015637055039405823  
step: 320, loss: 0.0020660965237766504  
step: 330, loss: 0.0018502267776057124  
step: 340, loss: 4.772676766151562e-05  
step: 350, loss: 0.0004614792997017503  
step: 360, loss: 3.0982442694948986e-05  
step: 370, loss: 4.0517159504815936e-05  
step: 380, loss: 0.0007957048364914954  
step: 390, loss: 8.14536542748101e-05  
step: 400, loss: 0.00021855374507140368  
step: 410, loss: 0.002016671933233738  
step: 420, loss: 7.369017839664593e-05  
step: 430, loss: 0.005399645771831274  
step: 440, loss: 8.229049853980541e-05  
step: 450, loss: 0.0001965826959349215  
step: 460, loss: 0.0002570189826656133  
step: 470, loss: 0.003639376489445567  
step: 480, loss: 7.243195432238281e-05  
step: 490, loss: 0.0001455118035664782  
step: 500, loss: 7.640875992365181e-05  
step: 510, loss: 2.9116097721271217e-05  
step: 520, loss: 0.0012259387876838446

```

step: 530, loss: 0.0011454758932814002
/usr/local/lib/python3.7/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=0.9491525423728813, f1=0.9464285714285713,
best_f1=0.9480519480519481
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 2.871389369829558e-05
step: 10, loss: 0.012474087066948414
step: 20, loss: 2.5603323592804372e-05
step: 30, loss: 0.000256739353062585
step: 40, loss: 0.00012815964873880148
step: 50, loss: 8.380485087400302e-05
step: 60, loss: 4.299860665923916e-05
step: 70, loss: 3.135479710181244e-05
step: 80, loss: 0.015768280252814293
step: 90, loss: 0.005002176854759455
step: 100, loss: 0.0005108140758238733
step: 110, loss: 8.942740532802418e-05
step: 120, loss: 0.0010891688289120793
step: 130, loss: 0.08856499195098877
step: 140, loss: 0.00010632543126121163
step: 150, loss: 0.00015606968372594565
step: 160, loss: 0.0031008555088192225
step: 170, loss: 0.00018309388542547822
step: 180, loss: 2.4437407773802988e-05
step: 190, loss: 0.0004334982659202069
step: 200, loss: 9.320990648120642e-05
step: 210, loss: 0.00013577446225099266
step: 220, loss: 0.0015400918200612068
step: 230, loss: 4.454745430848561e-05
step: 240, loss: 0.0008536375244148076
step: 250, loss: 4.7773810365470126e-05
step: 260, loss: 0.05147653818130493
step: 270, loss: 0.0025477283634245396
step: 280, loss: 0.00011880040256073698
step: 290, loss: 0.0002920240513049066
step: 300, loss: 4.054770033690147e-05
step: 310, loss: 9.114912245422602e-05
step: 320, loss: 0.0014982311986386776
step: 330, loss: 2.2660142349195667e-05
step: 340, loss: 0.00022493538563139737
step: 350, loss: 0.00040841830195859075
step: 360, loss: 0.0003702152753248811
step: 370, loss: 0.004468976520001888
step: 380, loss: 0.0012400811538100243

```

```

step: 390, loss: 0.002504673320800066
step: 400, loss: 4.7610643377993256e-05
step: 410, loss: 0.001843355712480843
step: 420, loss: 0.00010015438601840287
step: 430, loss: 0.005465260706841946
step: 440, loss: 6.827187462477013e-05
step: 450, loss: 0.0003563502978067845
step: 460, loss: 0.00010335518163628876
step: 470, loss: 0.0007149773300625384
step: 480, loss: 7.969318539835513e-05
step: 490, loss: 4.9009966460289434e-05
step: 500, loss: 0.001661685761064291
step: 510, loss: 5.5004456953611225e-05
step: 520, loss: 2.694442810025066e-05
step: 530, loss: 0.008940287865698338
/usr/local/lib/python3.7/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=0.9494855004677268, f1=0.9424426766495088,
best_f1=0.9480519480519481
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.002505919197574258
step: 10, loss: 0.030191997066140175
step: 20, loss: 0.0003431527875363827
step: 30, loss: 5.55777114641387e-05
step: 40, loss: 4.478835398913361e-05
step: 50, loss: 2.1241286958684213e-05
step: 60, loss: 0.012205475009977818
step: 70, loss: 2.9179071134421974e-05
step: 80, loss: 8.680135215399787e-05
step: 90, loss: 2.7488098567118868e-05
step: 100, loss: 0.0938107818365097
step: 110, loss: 2.063025203824509e-05
step: 120, loss: 0.0001239758712472394
step: 130, loss: 0.00013503266382031143
step: 140, loss: 0.00013255955127533525
step: 150, loss: 2.217979090346489e-05
step: 160, loss: 5.538838013308123e-05
step: 170, loss: 4.173240449745208e-05
step: 180, loss: 2.272010715387296e-05
step: 190, loss: 0.00022781257575843483
step: 200, loss: 0.0020701370667666197
step: 210, loss: 0.0003390217025298625
step: 220, loss: 2.3986538508324884e-05
step: 230, loss: 5.4424828704213724e-05
step: 240, loss: 0.0006975647993385792

```

```

step: 250, loss: 0.004260603338479996
step: 260, loss: 3.0184084607753903e-05
step: 270, loss: 0.0026077984366565943
step: 280, loss: 9.04009721125476e-05
step: 290, loss: 1.928548408614006e-05
step: 300, loss: 7.392124098259956e-05
step: 310, loss: 3.310101237730123e-05
step: 320, loss: 0.005131850019097328
step: 330, loss: 0.0029288972727954388
step: 340, loss: 0.00015797055675648153
step: 350, loss: 8.352693839697167e-05
step: 360, loss: 0.0009891230147331953
step: 370, loss: 2.673312155820895e-05
step: 380, loss: 0.00012619694462046027
step: 390, loss: 0.024217167869210243
step: 400, loss: 2.69135653070407e-05
step: 410, loss: 0.00016848323866724968
step: 420, loss: 0.0003728090086951852
step: 430, loss: 5.959202826488763e-05
step: 440, loss: 0.0011040396057069302
step: 450, loss: 0.08750869333744049
step: 460, loss: 8.764781523495913e-05
step: 470, loss: 0.03974539786577225
step: 480, loss: 0.011272784322500229
step: 490, loss: 9.27430737647228e-05
step: 500, loss: 0.00021320246742106974
step: 510, loss: 0.0025757434777915478
step: 520, loss: 8.204580808524042e-05
step: 530, loss: 4.074530806974508e-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=0.9507735583684951, f1=0.941509433962264,
best_f1=0.941509433962264
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00045176551793701947
step: 10, loss: 2.4984936317196116e-05
step: 20, loss: 0.0009765371796675026
step: 30, loss: 2.8481839763117023e-05
step: 40, loss: 2.2235761207411997e-05
step: 50, loss: 0.0014021291863173246
step: 60, loss: 0.00015835955855436623

```



step: 70, loss: 5.329161649569869e-05  
step: 80, loss: 0.0006053528049960732  
step: 90, loss: 2.2261679987423122e-05  
step: 100, loss: 2.0216537450323813e-05  
step: 110, loss: 1.8253493180964142e-05  
step: 120, loss: 2.1311827367753722e-05  
step: 130, loss: 1.2412569958542008e-05  
step: 140, loss: 2.2138934582471848e-05  
step: 150, loss: 2.413558650005143e-05  
step: 160, loss: 7.626148726558313e-05  
step: 170, loss: 0.00022264647122938186  
step: 180, loss: 2.7930831492994912e-05  
step: 190, loss: 8.440313104074448e-05  
step: 200, loss: 1.74191445694305e-05  
step: 210, loss: 0.006448476575314999  
step: 220, loss: 1.7668557120487094e-05  
step: 230, loss: 0.00039178592851385474  
step: 240, loss: 0.00021143778576515615  
step: 250, loss: 3.10637078655418e-05  
step: 260, loss: 9.450033394386992e-05  
step: 270, loss: 0.00010759750148281455  
step: 280, loss: 1.556408460601233e-05  
step: 290, loss: 1.4107513379713055e-05  
step: 300, loss: 1.5266061382135376e-05  
step: 310, loss: 4.104997424292378e-05  
step: 320, loss: 2.2696132873534225e-05  
step: 330, loss: 3.816586104221642e-05  
step: 340, loss: 0.00021775650384370238  
step: 350, loss: 1.1466342584753875e-05  
step: 360, loss: 0.05713341385126114  
step: 370, loss: 0.006788663566112518  
step: 380, loss: 1.6927395336097106e-05  
step: 390, loss: 0.031123202294111252  
step: 400, loss: 1.9549965145415626e-05  
step: 410, loss: 3.589072366594337e-05  
step: 420, loss: 1.4275125977292191e-05  
step: 430, loss: 2.1181213014642708e-05  
step: 440, loss: 1.1242826076340862e-05  
step: 450, loss: 1.5556555808871053e-05  
step: 460, loss: 0.00033054148661904037  
step: 470, loss: 0.00022345002798829228  
step: 480, loss: 9.83840163826244e-06  
step: 490, loss: 1.0911277968261857e-05  
step: 500, loss: 1.0400923201814294e-05  
step: 510, loss: 1.6115052858367562e-05  
step: 520, loss: 1.1958065442740917e-05  
step: 530, loss: 1.263605008716695e-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 13: dev_f1=0.951366373320982, f1=0.9437993497445425,
best_f1=0.9437993497445425
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 7.164145790738985e-05
step: 10, loss: 0.0013020102633163333
step: 20, loss: 0.0004918506601825356
step: 30, loss: 2.2331874788505957e-05
step: 40, loss: 3.431926961638965e-05
step: 50, loss: 2.37662679865025e-05
step: 60, loss: 1.295269794354681e-05
step: 70, loss: 0.022755874320864677
step: 80, loss: 9.705981210572645e-05
step: 90, loss: 1.559725023980718e-05
step: 100, loss: 1.6577154383412562e-05
step: 110, loss: 0.000986472237855196
step: 120, loss: 1.2799966498278081e-05
step: 130, loss: 3.320016912766732e-05
step: 140, loss: 2.6105526558239944e-05
step: 150, loss: 1.7407926861778833e-05
step: 160, loss: 1.7102462152251974e-05
step: 170, loss: 3.1880226742941886e-05
step: 180, loss: 9.996697917813435e-05
step: 190, loss: 1.7210400983458385e-05
step: 200, loss: 2.2410078599932604e-05
step: 210, loss: 0.0009061736054718494
step: 220, loss: 2.0267687432351522e-05
step: 230, loss: 1.2524311387096532e-05
step: 240, loss: 3.36046505253762e-05
step: 250, loss: 0.0003122595662716776
step: 260, loss: 4.448561230674386e-05
step: 270, loss: 7.572163303848356e-05
step: 280, loss: 4.797249857801944e-05
step: 290, loss: 1.135830461862497e-05
step: 300, loss: 1.3887670320400503e-05
step: 310, loss: 0.0005803940584883094
step: 320, loss: 1.2766451618517749e-05
step: 330, loss: 0.004119178280234337
step: 340, loss: 5.5622687796130776e-05
step: 350, loss: 0.00013929394481237978
step: 360, loss: 1.4111154996498954e-05

```

```

step: 370, loss: 0.009323655627667904
step: 380, loss: 0.0002437518269289285
step: 390, loss: 1.7738786482368596e-05
step: 400, loss: 0.001496975659392774
step: 410, loss: 1.5273297321982682e-05
step: 420, loss: 1.2379020517983008e-05
step: 430, loss: 0.00017174193635582924
step: 440, loss: 1.8350368918618187e-05
step: 450, loss: 1.3958480849396437e-05
step: 460, loss: 7.530658331234008e-05
step: 470, loss: 1.3109153769619297e-05
step: 480, loss: 2.9775808798149228e-05
step: 490, loss: 1.747114583849907e-05
step: 500, loss: 1.5739115042379126e-05
step: 510, loss: 2.2033022105460986e-05
step: 520, loss: 1.156314647232648e-05
step: 530, loss: 1.3287954971019644e-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=0.952513966480447, f1=0.9417792268281323,
best_f1=0.9417792268281323
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0003997367457486689
step: 10, loss: 2.749753366515506e-05
step: 20, loss: 0.0008294796571135521
step: 30, loss: 9.049909567693248e-05
step: 40, loss: 1.3347565982257947e-05
step: 50, loss: 0.0004737253475468606
step: 60, loss: 1.2598813555086963e-05
step: 70, loss: 2.67893192358315e-05
step: 80, loss: 1.0970893526973668e-05
step: 90, loss: 0.0008743273792788386
step: 100, loss: 8.981619430414867e-06
step: 110, loss: 0.001307241735048592
step: 120, loss: 2.7263780793873593e-05
step: 130, loss: 1.763856562320143e-05
step: 140, loss: 9.044945727509912e-06
step: 150, loss: 1.907618388941046e-05
step: 160, loss: 2.9229770007077605e-05
step: 170, loss: 1.474809232604457e-05
step: 180, loss: 1.1633988833636977e-05

```

```

step: 190, loss: 0.0027404637075960636
step: 200, loss: 0.00018320477101951838
step: 210, loss: 0.00014174042735248804
step: 220, loss: 1.2192707799840719e-05
step: 230, loss: 0.0004523284442257136
step: 240, loss: 1.5638359400327317e-05
step: 250, loss: 9.251337178284302e-05
step: 260, loss: 9.681955816631671e-06
step: 270, loss: 1.430479187547462e-05
step: 280, loss: 1.0467980246176012e-05
step: 290, loss: 2.1556688807322644e-05
step: 300, loss: 9.97998722596094e-06
step: 310, loss: 0.042312659323215485
step: 320, loss: 1.5403722500195727e-05
step: 330, loss: 1.366791912005283e-05
step: 340, loss: 1.6905043594306335e-05
step: 350, loss: 0.007390133570879698
step: 360, loss: 1.526228697912302e-05
step: 370, loss: 1.3369906810112298e-05
step: 380, loss: 1.4211775123840198e-05
step: 390, loss: 1.1626529158093035e-05
step: 400, loss: 0.0004982162499800324
step: 410, loss: 1.443891324015567e-05
step: 420, loss: 0.0005434257909655571
step: 430, loss: 8.754375812713988e-06
step: 440, loss: 1.3656689588970039e-05
step: 450, loss: 8.015103958314285e-05
step: 460, loss: 0.006436201743781567
step: 470, loss: 1.3071832654532045e-05
step: 480, loss: 0.0008459454402327538
step: 490, loss: 2.0011306332889944e-05
step: 500, loss: 1.0453079084982164e-05
step: 510, loss: 1.1037916010536719e-05
step: 520, loss: 1.337352478003595e-05
step: 530, loss: 9.6670428320067e-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.")
/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=0.9530450953045096, f1=0.9437470943747095,
best_f1=0.9437470943747095
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 [00:24, 237.03it/s]
load_f1 = 0.9512761020881672
real_f1 = 0.9496738117427772
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 [00:23, 183.66it/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: 0.48668450117111206

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

step: 10, loss: 0.438909649848938

step: 20, loss: 0.43601876497268677

step: 30, loss: 0.3605321943759918

step: 40, loss: 0.3367757499217987

step: 50, loss: 0.4260593354701996

step: 60, loss: 0.443242609500885

step: 70, loss: 0.30107778310775757

step: 80, loss: 0.3863406181335449

step: 90, loss: 0.22510124742984772

step: 100, loss: 0.26911023259162903

step: 110, loss: 0.279682457447052

step: 120, loss: 0.32042914628982544

step: 130, loss: 0.18761219084262848

step: 140, loss: 0.5689704418182373

step: 150, loss: 0.29663020372390747

step: 160, loss: 0.3015713691711426

step: 170, loss: 0.1050080731511116

step: 180, loss: 0.28945523500442505

step: 190, loss: 0.6445323824882507

step: 200, loss: 0.20184697210788727

step: 210, loss: 0.2694958746433258

/usr/local/lib/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=0.5464285714285715, f1=0.5803921568627451,

best\_f1=0.5803921568627451

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

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

step: 0, loss: 0.28153783082962036

step: 10, loss: 0.07556692510843277

step: 20, loss: 0.37949398159980774

```

step: 30, loss: 0.3090801537036896
step: 40, loss: 0.32320934534072876
step: 50, loss: 0.11541657149791718
step: 60, loss: 0.308791846036911
step: 70, loss: 0.222229465842247
step: 80, loss: 0.24936895072460175
step: 90, loss: 0.20131559669971466
step: 100, loss: 0.3350042700767517
step: 110, loss: 0.45955345034599304
step: 120, loss: 0.13602770864963531
step: 130, loss: 0.21266694366931915
step: 140, loss: 0.12925861775875092
step: 150, loss: 0.26120978593826294
step: 160, loss: 0.08182371407747269
step: 170, loss: 0.26985424757003784
step: 180, loss: 0.1796920746564865
step: 190, loss: 0.22460409998893738
step: 200, loss: 0.14888590574264526
step: 210, loss: 0.13927456736564636
/usr/local/lib/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=0.6045627376425856, f1=0.6356877323420074,
best_f1=0.6356877323420074
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.08643538504838943
step: 10, loss: 0.15192952752113342
step: 20, loss: 0.24178749322891235
step: 30, loss: 0.038194622844457626
step: 40, loss: 0.31044021248817444
step: 50, loss: 0.29293882846832275
step: 60, loss: 0.2435871958732605
step: 70, loss: 0.15499304234981537
step: 80, loss: 0.14503879845142365
step: 90, loss: 0.07094749063253403
step: 100, loss: 0.17166930437088013
step: 110, loss: 0.1008991003036499
step: 120, loss: 0.16244202852249146
step: 130, loss: 0.1095937192440033
step: 140, loss: 0.11374469101428986
step: 150, loss: 0.09566116333007812
step: 160, loss: 0.20154789090156555

```



```

step: 170, loss: 0.202677384018898
step: 180, loss: 0.1478595733642578
step: 190, loss: 0.06534461677074432
step: 200, loss: 0.11632432788610458
step: 210, loss: 0.12982161343097687
/usr/local/lib/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=0.634989200863931, f1=0.6214442013129102,
best_f1=0.6214442013129102
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.08482685685157776
step: 10, loss: 0.088365338742733
step: 20, loss: 0.0895482525229454
step: 30, loss: 0.060315750539302826
step: 40, loss: 0.07297733426094055
step: 50, loss: 0.16792583465576172
step: 60, loss: 0.21748585999011993
step: 70, loss: 0.029178114607930183
step: 80, loss: 0.06591637432575226
step: 90, loss: 0.11871214956045151
step: 100, loss: 0.21012647449970245
step: 110, loss: 0.4207897484302521
step: 120, loss: 0.22940269112586975
step: 130, loss: 0.22966234385967255
step: 140, loss: 0.21946780383586884
step: 150, loss: 0.1207684874534607
step: 160, loss: 0.11267613619565964
step: 170, loss: 0.07182630151510239
step: 180, loss: 0.03344938904047012
step: 190, loss: 0.13416296243667603
step: 200, loss: 0.0502881221473217
step: 210, loss: 0.3598298728466034
/usr/local/lib/python3.7/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=0.6326129666011788, f1=0.668041237113402,
best_f1=0.6214442013129102
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.16849906742572784

```

```

step: 10, loss: 0.03252726420760155
step: 20, loss: 0.14807216823101044
step: 30, loss: 0.041149746626615524
step: 40, loss: 0.09005340933799744
step: 50, loss: 0.1194400042295456
step: 60, loss: 0.1587425172328949
step: 70, loss: 0.2195877730846405
step: 80, loss: 0.08325926959514618
step: 90, loss: 0.02925419807434082
step: 100, loss: 0.011959430761635303
step: 110, loss: 0.09865714609622955
step: 120, loss: 0.12793028354644775
step: 130, loss: 0.10181865841150284
step: 140, loss: 0.3998469412326813
step: 150, loss: 0.13092568516731262
step: 160, loss: 0.04876243695616722
step: 170, loss: 0.02149749919772148
step: 180, loss: 0.1908382773399353
step: 190, loss: 0.12318955361843109
step: 200, loss: 0.10514963418245316
step: 210, loss: 0.1291574090719223
/usr/local/lib/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=0.6490566037735849, f1=0.6615087040618955,
best_f1=0.6615087040618955
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.018432648852467537
step: 10, loss: 0.03269780054688454
step: 20, loss: 0.008607362397015095
step: 30, loss: 0.05452512577176094
step: 40, loss: 0.03389492258429527
step: 50, loss: 0.08506370335817337
step: 60, loss: 0.20780810713768005
step: 70, loss: 0.14202499389648438
step: 80, loss: 0.1195085421204567
step: 90, loss: 0.17919111251831055
step: 100, loss: 0.1342800110578537
step: 110, loss: 0.01891876570880413
step: 120, loss: 0.008612936362624168
step: 130, loss: 0.06720773875713348
step: 140, loss: 0.07732772082090378

```

```

step: 150, loss: 0.11094838380813599
step: 160, loss: 0.06992298364639282
step: 170, loss: 0.11312435567378998
step: 180, loss: 0.04960786551237106
step: 190, loss: 0.026407387107610703
step: 200, loss: 0.18976077437400818
step: 210, loss: 0.02259535901248455
/usr/local/lib/python3.7/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=0.6338797814207651, f1=0.6641509433962264,
best_f1=0.6615087040618955
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.10747488588094711
step: 10, loss: 0.031191643327474594
step: 20, loss: 0.007949508726596832
step: 30, loss: 0.04132949560880661
step: 40, loss: 0.05393214151263237
step: 50, loss: 0.11125777661800385
step: 60, loss: 0.05308928340673447
step: 70, loss: 0.0053015160374343395
step: 80, loss: 0.1729053258895874
step: 90, loss: 0.08445370942354202
step: 100, loss: 0.05927707627415657
step: 110, loss: 0.03955446183681488
step: 120, loss: 0.03363833948969841
step: 130, loss: 0.024731166660785675
step: 140, loss: 0.033305395394563675
step: 150, loss: 0.04278366640210152
step: 160, loss: 0.16981777548789978
step: 170, loss: 0.10761465132236481
step: 180, loss: 0.034319065511226654
step: 190, loss: 0.057131338864564896
step: 200, loss: 0.05010353401303291
step: 210, loss: 0.18575835227966309
/usr/local/lib/python3.7/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=0.6396761133603239, f1=0.6694386694386695,
best_f1=0.6615087040618955
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.09107420593500137
step: 10, loss: 0.017921052873134613
step: 20, loss: 0.013842111453413963

```

```

step: 30, loss: 0.030774561688303947
step: 40, loss: 0.08469988405704498
step: 50, loss: 0.022281933575868607
step: 60, loss: 0.11088753491640091
step: 70, loss: 0.09630400687456131
step: 80, loss: 0.022106921300292015
step: 90, loss: 0.009542075917124748
step: 100, loss: 0.20108966529369354
step: 110, loss: 0.0019705144222825766
step: 120, loss: 0.05965230613946915
step: 130, loss: 0.07845363765954971
step: 140, loss: 0.03661620616912842
step: 150, loss: 0.10203049331903458
step: 160, loss: 0.13471318781375885
step: 170, loss: 0.059851426631212234
step: 180, loss: 0.04301206395030022
step: 190, loss: 0.10827255994081497
step: 200, loss: 0.03160892799496651
step: 210, loss: 0.08308248966932297
/usr/local/lib/python3.7/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=0.6435845213849287, f1=0.6640159045725645,
best_f1=0.6615087040618955
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.051390159875154495
step: 10, loss: 0.037951987236738205
step: 20, loss: 0.17136316001415253
step: 30, loss: 0.005868652369827032
step: 40, loss: 0.055541932582855225
step: 50, loss: 0.13319142162799835
step: 60, loss: 0.15324942767620087
step: 70, loss: 0.06353515386581421
step: 80, loss: 0.11848428845405579
step: 90, loss: 0.01803024299442768
step: 100, loss: 0.009777948260307312
step: 110, loss: 0.013111673295497894
step: 120, loss: 0.19448161125183105
step: 130, loss: 0.06195026636123657
step: 140, loss: 0.2107706069946289
step: 150, loss: 0.016002172604203224
step: 160, loss: 0.07352626323699951
step: 170, loss: 0.0984923467040062
step: 180, loss: 0.02922138385474682
step: 190, loss: 0.006003404967486858
step: 200, loss: 0.00969408918172121

```

```

step: 210, loss: 0.08667650073766708
/usr/local/lib/python3.7/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=0.6440677966101696, f1=0.6752293577981651,
best_f1=0.6615087040618955
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0378078855574131
step: 10, loss: 0.010951886884868145
step: 20, loss: 0.004460430704057217
step: 30, loss: 0.0024596452713012695
step: 40, loss: 0.05461179465055466
step: 50, loss: 0.014455419965088367
step: 60, loss: 0.015217820182442665
step: 70, loss: 0.01641494408249855
step: 80, loss: 0.05355599895119667
step: 90, loss: 0.04924474284052849
step: 100, loss: 0.04096512496471405
step: 110, loss: 0.10200949013233185
step: 120, loss: 0.04041188582777977
step: 130, loss: 0.010258179157972336
step: 140, loss: 0.020111963152885437
step: 150, loss: 0.1370549201965332
step: 160, loss: 0.11004069447517395
step: 170, loss: 0.01823848858475685
step: 180, loss: 0.07922553271055222
step: 190, loss: 0.015268264338374138
step: 200, loss: 0.22104492783546448
step: 210, loss: 0.020941076800227165
/usr/local/lib/python3.7/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=0.6457564575645757, f1=0.6799276672694394,
best_f1=0.6615087040618955
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.02994038537144661
step: 10, loss: 0.013735486194491386
step: 20, loss: 0.006496243178844452
step: 30, loss: 0.013215932063758373
step: 40, loss: 0.0026937234215438366
step: 50, loss: 0.014397809281945229
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 60, loss: 0.1477545201778412
step: 70, loss: 0.002812086371704936

```

```

step: 80, loss: 0.05527050420641899
step: 90, loss: 0.13411131501197815
step: 100, loss: 0.011351427063345909
step: 110, loss: 0.001020799740217626
step: 120, loss: 0.0061863139271736145
step: 130, loss: 0.006646968889981508
step: 140, loss: 0.029056988656520844
step: 150, loss: 0.004628725349903107
step: 160, loss: 0.0027996604330837727
step: 170, loss: 0.0685679242014885
step: 180, loss: 0.004564494825899601
step: 190, loss: 0.08029980212450027
step: 200, loss: 0.004940506536513567
step: 210, loss: 0.1685795783996582
/usr/local/lib/python3.7/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=0.6356275303643724, f1=0.688659793814433,
best_f1=0.6615087040618955
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.014503756538033485
step: 10, loss: 0.02380989119410515
step: 20, loss: 0.08754090219736099
step: 30, loss: 0.04886600002646446
step: 40, loss: 0.0027332021854817867
step: 50, loss: 0.03656753525137901
step: 60, loss: 0.009897866286337376
step: 70, loss: 0.0665193721652031
step: 80, loss: 0.05473249778151512
step: 90, loss: 0.03734053671360016
step: 100, loss: 0.0007257495308294892
step: 110, loss: 0.010914953425526619
step: 120, loss: 0.006229388993233442
step: 130, loss: 0.0017266958020627499
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 8192.0
step: 140, loss: 0.01251225359737873
step: 150, loss: 0.00119341432582587
step: 160, loss: 0.002700187498703599
step: 170, loss: 0.0018081102753058076
step: 180, loss: 0.010999424383044243
step: 190, loss: 0.004964136052876711
step: 200, loss: 0.021333681419491768
step: 210, loss: 0.008539094589650631
/usr/local/lib/python3.7/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=0.6339622641509433, f1=0.6971428571428572,
best_f1=0.6615087040618955
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00907278060913086
step: 10, loss: 0.04291369393467903
step: 20, loss: 0.04182393476366997
step: 30, loss: 0.0019081997452303767
step: 40, loss: 0.00244339881464839
step: 50, loss: 0.027217473834753036
step: 60, loss: 0.010023158974945545
step: 70, loss: 0.09204992651939392
step: 80, loss: 0.07200545072555542
step: 90, loss: 0.044841136783361435
step: 100, loss: 0.0679549127817154
step: 110, loss: 0.07548706978559494
step: 120, loss: 0.06383275240659714
step: 130, loss: 0.05059768259525299
step: 140, loss: 0.006027671974152327
step: 150, loss: 0.0802813321352005
step: 160, loss: 0.06380664557218552
step: 170, loss: 0.029357196763157845
step: 180, loss: 0.007588121108710766
step: 190, loss: 0.023213708773255348
step: 200, loss: 0.01434690784662962
step: 210, loss: 0.07764070481061935
/usr/local/lib/python3.7/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=0.6310517529215358, f1=0.6757215619694398,
best_f1=0.6615087040618955
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.005404351744800806
step: 10, loss: 0.018300889059901237
step: 20, loss: 0.03054545447230339
step: 30, loss: 0.06947214901447296
step: 40, loss: 0.08781830966472626
step: 50, loss: 0.009871447458863258
step: 60, loss: 0.022771628573536873
step: 70, loss: 0.04109417274594307
step: 80, loss: 0.06911454349756241
step: 90, loss: 0.017350200563669205
step: 100, loss: 0.006593918427824974
step: 110, loss: 0.10767147690057755
step: 120, loss: 0.002383989281952381

```

```

step: 130, loss: 0.011523446999490261
step: 140, loss: 0.004574749153107405
step: 150, loss: 0.049129828810691833
step: 160, loss: 0.005410439800471067
step: 170, loss: 0.09271451830863953
step: 180, loss: 0.0025384880136698484
step: 190, loss: 0.031122615560889244
step: 200, loss: 0.00741127273067832
step: 210, loss: 0.0023616233374923468
/usr/local/lib/python3.7/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=0.6365348399246704, f1=0.7088122605363985,
best_f1=0.6615087040618955
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.002663656137883663
step: 10, loss: 0.010268422774970531
step: 20, loss: 0.0018080779118463397
step: 30, loss: 0.016377722844481468
step: 40, loss: 0.010451441630721092
step: 50, loss: 0.009629807434976101
step: 60, loss: 0.0402354821562767
step: 70, loss: 0.003966909367591143
step: 80, loss: 0.03892144933342934
step: 90, loss: 0.002478274516761303
step: 100, loss: 0.008441167883574963
step: 110, loss: 0.06271214783191681
step: 120, loss: 0.0023818297777324915
step: 130, loss: 0.005913670174777508
step: 140, loss: 0.01466827467083931
step: 150, loss: 0.0029394489247351885
step: 160, loss: 0.10005604475736618
step: 170, loss: 0.011428341269493103
step: 180, loss: 0.004604777321219444
step: 190, loss: 0.04766465723514557
step: 200, loss: 0.006014121230691671
step: 210, loss: 0.04559463635087013
/usr/local/lib/python3.7/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=0.638532110091743, f1=0.6987060998151571,
best_f1=0.6615087040618955
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.")
2293it [00:05, 403.70it/s]
load_f1 = 0.642166344294004
real_f1 = 0.6404715127701375
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 [00:23, 187.22it/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: 0.4721939265727997
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: 10, loss: 0.47613510489463806
step: 20, loss: 0.25346338748931885
step: 30, loss: 0.41786324977874756
step: 40, loss: 0.227517232298851
step: 50, loss: 0.32105615735054016
step: 60, loss: 0.43835651874542236
step: 70, loss: 0.4291508197784424
step: 80, loss: 0.1545601338148117
step: 90, loss: 0.31386473774909973
step: 100, loss: 0.4197383522987366
step: 110, loss: 0.24510334432125092
step: 120, loss: 0.3216686546802521
step: 130, loss: 0.31577616930007935
step: 140, loss: 0.1891362965106964
step: 150, loss: 0.3352096974849701
step: 160, loss: 0.2541508078575134
step: 170, loss: 0.3696262836456299
step: 180, loss: 0.15600700676441193
step: 190, loss: 0.15963852405548096
/usr/local/lib/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: 0.38224178552627563
step: 10, loss: 0.3239496350288391
step: 20, loss: 0.6958616375923157
step: 30, loss: 0.25872254371643066

```

```

step: 40, loss: 0.5276517868041992
step: 50, loss: 0.3066784739494324
step: 60, loss: 0.44309791922569275
step: 70, loss: 0.2899143099784851
step: 80, loss: 0.16406014561653137
step: 90, loss: 0.3014567196369171
step: 100, loss: 0.2517034411430359
step: 110, loss: 0.2967916429042816
step: 120, loss: 0.20935609936714172
step: 130, loss: 0.37728720903396606
step: 140, loss: 0.27568215131759644
step: 150, loss: 0.299304723739624
step: 160, loss: 0.33835312724113464
step: 170, loss: 0.1673140674829483
step: 180, loss: 0.16231922805309296
step: 190, loss: 0.2551964521408081
/usr/local/lib/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=0.5444126074498566, f1=0.5301837270341208,
best_f1=0.5301837270341208
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.3474060297012329
step: 10, loss: 0.29429635405540466
step: 20, loss: 0.24400964379310608
step: 30, loss: 0.2088981568813324
step: 40, loss: 0.1496153324842453
step: 50, loss: 0.27692723274230957
step: 60, loss: 0.16164249181747437
step: 70, loss: 0.3652121424674988
step: 80, loss: 0.21840238571166992
step: 90, loss: 0.2762959599494934
step: 100, loss: 0.32488206028938293
step: 110, loss: 0.5348231792449951
step: 120, loss: 0.39158567786216736
step: 130, loss: 0.17641697824001312
step: 140, loss: 0.17310555279254913
step: 150, loss: 0.301565021276474
step: 160, loss: 0.37858906388282776
step: 170, loss: 0.2862355411052704
step: 180, loss: 0.21704131364822388
step: 190, loss: 0.1558499038219452

```

```

/usr/local/lib/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=0.6233766233766234, f1=0.5934065934065934,
best_f1=0.5934065934065934
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.09851516038179398
step: 10, loss: 0.14321371912956238
step: 20, loss: 0.19104906916618347
step: 30, loss: 0.054376598447561264
step: 40, loss: 0.2794499099254608
step: 50, loss: 0.18828585743904114
step: 60, loss: 0.21759343147277832
step: 70, loss: 0.1758642941713333
step: 80, loss: 0.1399543583393097
step: 90, loss: 0.10608511418104172
step: 100, loss: 0.24378976225852966
step: 110, loss: 0.3161776065826416
step: 120, loss: 0.051511798053979874
step: 130, loss: 0.15808390080928802
step: 140, loss: 0.3319988548755646
step: 150, loss: 0.06764258444309235
step: 160, loss: 0.04853322356939316
step: 170, loss: 0.13800500333309174
step: 180, loss: 0.22034338116645813
step: 190, loss: 0.14195463061332703
/usr/local/lib/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=0.6951871657754011, f1=0.7002652519893899,
best_f1=0.7002652519893899
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.25658518075942993
step: 10, loss: 0.2626730799674988
step: 20, loss: 0.044123370200395584
step: 30, loss: 0.13416333496570587

```

```

step: 40, loss: 0.06290885806083679
step: 50, loss: 0.08811171352863312
step: 60, loss: 0.08299019187688828
step: 70, loss: 0.17759570479393005
step: 80, loss: 0.10969599336385727
step: 90, loss: 0.22210289537906647
step: 100, loss: 0.13756507635116577
step: 110, loss: 0.286873996257782
step: 120, loss: 0.17014329135417938
step: 130, loss: 0.4878247082233429
step: 140, loss: 0.105364590883255
step: 150, loss: 0.23506823182106018
step: 160, loss: 0.011187820695340633
step: 170, loss: 0.05111226066946983
step: 180, loss: 0.06764166057109833
step: 190, loss: 0.3330989480018616
/usr/local/lib/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=0.7282608695652173, f1=0.7229551451187335,
best_f1=0.7229551451187335
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.12476036697626114
step: 10, loss: 0.20228920876979828
step: 20, loss: 0.044153325259685516
step: 30, loss: 0.07410579174757004
step: 40, loss: 0.08496959507465363
step: 50, loss: 0.2366665005683899
step: 60, loss: 0.18382571637630463
step: 70, loss: 0.1661592423915863
step: 80, loss: 0.20454946160316467
step: 90, loss: 0.1030678004026413
step: 100, loss: 0.036031439900398254
step: 110, loss: 0.022962328046560287
step: 120, loss: 0.10466384887695312
step: 130, loss: 0.2555747628211975
step: 140, loss: 0.029761094599962234
step: 150, loss: 0.03506653755903244
step: 160, loss: 0.21102330088615417
step: 170, loss: 0.4158933758735657
step: 180, loss: 0.1703430414199829
step: 190, loss: 0.09501216560602188

```

```

/usr/local/lib/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=0.7377521613832854, f1=0.7740112994350282,
best_f1=0.7740112994350282
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.08248136937618256
step: 10, loss: 0.05382460728287697
step: 20, loss: 0.014702262356877327
step: 30, loss: 0.10852986574172974
step: 40, loss: 0.06541217118501663
step: 50, loss: 0.009382733143866062
step: 60, loss: 0.09326282143592834
step: 70, loss: 0.03203771263360977
step: 80, loss: 0.004799552261829376
step: 90, loss: 0.05469565838575363
step: 100, loss: 0.21189376711845398
step: 110, loss: 0.16723360121250153
step: 120, loss: 0.027913354337215424
step: 130, loss: 0.09611400961875916
step: 140, loss: 0.09305035322904587
step: 150, loss: 0.08031883090734482
step: 160, loss: 0.18216653168201447
step: 170, loss: 0.13867859542369843
step: 180, loss: 0.04546016454696655
step: 190, loss: 0.10095618665218353
/usr/local/lib/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=0.7696202531645571, f1=0.7722772277227723,
best_f1=0.7722772277227723
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0963403582572937
step: 10, loss: 0.032376598566770554
step: 20, loss: 0.032395150512456894
step: 30, loss: 0.07958222925662994

```

```

step: 40, loss: 0.10859053581953049
step: 50, loss: 0.046812545508146286
step: 60, loss: 0.04804302006959915
step: 70, loss: 0.03141620010137558
step: 80, loss: 0.11410502344369888
step: 90, loss: 0.10511887818574905
step: 100, loss: 0.004725284408777952
step: 110, loss: 0.026858963072299957
step: 120, loss: 0.00942388828843832
step: 130, loss: 0.023621421307325363
step: 140, loss: 0.03211945667862892
step: 150, loss: 0.07055453956127167
step: 160, loss: 0.033163394778966904
step: 170, loss: 0.012243675999343395
step: 180, loss: 0.10787006467580795
step: 190, loss: 0.002648416208103299
/usr/local/lib/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=0.7936507936507936, f1=0.8031496062992125,
best_f1=0.8031496062992125
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.005718203727155924
step: 10, loss: 0.011232424527406693
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 4096.0
step: 20, loss: 0.01966974511742592
step: 30, loss: 0.012193798087537289
step: 40, loss: 0.23649442195892334
step: 50, loss: 0.04783615469932556
step: 60, loss: 0.039956409484148026
step: 70, loss: 0.002136032097041607
step: 80, loss: 0.1355612725019455
step: 90, loss: 0.02489280514419079
step: 100, loss: 0.04990222677588463
step: 110, loss: 0.17796678841114044
step: 120, loss: 0.03398673236370087
step: 130, loss: 0.03721490874886513
step: 140, loss: 0.10135317593812943
step: 150, loss: 0.013713710941374302
step: 160, loss: 0.011729654856026173
step: 170, loss: 0.039847683161497116
step: 180, loss: 0.021303653717041016

```



```

step: 190, loss: 0.000613425683695823
/usr/local/lib/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=0.8148148148148148, f1=0.8160000000000001,
best_f1=0.8160000000000001
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.008149472996592522
step: 10, loss: 0.0424569770693779
step: 20, loss: 0.002357723657041788
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 2048.0
step: 30, loss: 0.0922163650393486
step: 40, loss: 0.0020353212021291256
step: 50, loss: 0.044193997979164124
step: 60, loss: 0.0006214666645973921
step: 70, loss: 0.0006036697886884212
step: 80, loss: 0.046434223651885986
step: 90, loss: 0.004443855490535498
step: 100, loss: 0.011406171135604382
step: 110, loss: 0.08570490032434464
step: 120, loss: 0.005196318030357361
step: 130, loss: 0.005240065045654774
step: 140, loss: 0.009558452293276787
step: 150, loss: 0.010075921192765236
step: 160, loss: 0.01937168277800083
step: 170, loss: 0.0007690918864682317
step: 180, loss: 0.19611194729804993
step: 190, loss: 0.0022682833950966597
/usr/local/lib/python3.7/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=0.8138297872340426, f1=0.8164383561643836,
best_f1=0.8160000000000001
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.012606888078153133
step: 10, loss: 0.0006535182474181056
step: 20, loss: 0.13319240510463715
step: 30, loss: 0.03486981615424156
step: 40, loss: 0.01694658398628235
step: 50, loss: 0.009635351598262787

```

```

step: 60, loss: 0.01657724380493164
step: 70, loss: 0.004294670652598143
step: 80, loss: 0.09827946126461029
step: 90, loss: 0.01951613463461399
step: 100, loss: 0.014636226929724216
step: 110, loss: 0.21873396635055542
step: 120, loss: 0.005158707499504089
step: 130, loss: 0.03222658112645149
step: 140, loss: 0.0044281380251049995
step: 150, loss: 0.0016411395044997334
step: 160, loss: 0.09091600030660629
step: 170, loss: 0.025587264448404312
step: 180, loss: 0.008525043725967407
step: 190, loss: 0.0636666864156723
/usr/local/lib/python3.7/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=0.8093994778067886, f1=0.8042895442359249,
best_f1=0.8160000000000001
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.006401387043297291
step: 10, loss: 0.0014025520067662
step: 20, loss: 0.004948504269123077
step: 30, loss: 0.049474943429231644
step: 40, loss: 0.0033231317065656185
step: 50, loss: 0.045820944011211395
step: 60, loss: 0.01239522360265255
step: 70, loss: 0.0008924033609218895
step: 80, loss: 0.06228470802307129
step: 90, loss: 0.04352736845612526
step: 100, loss: 0.008031567558646202
step: 110, loss: 0.0644315555691719
step: 120, loss: 0.04582544043660164
step: 130, loss: 0.005049959756433964
step: 140, loss: 0.0011250709649175406
step: 150, loss: 0.007787968032062054
step: 160, loss: 0.0008138103294186294
step: 170, loss: 0.005025146529078484
step: 180, loss: 0.002560446737334132
step: 190, loss: 0.0015454163076356053
/usr/local/lib/python3.7/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=0.8103896103896104, f1=0.8093994778067886,
best_f1=0.8160000000000001
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:

```

```

An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0030791601166129112
step: 10, loss: 0.0007740490254946053
step: 20, loss: 0.0008660836610943079
step: 30, loss: 0.03373381122946739
step: 40, loss: 0.05572792887687683
step: 50, loss: 0.0005784758250229061
step: 60, loss: 0.04760413244366646
step: 70, loss: 0.03188464418053627
step: 80, loss: 0.0016496728640049696
step: 90, loss: 0.00989248976111412
step: 100, loss: 0.0004634282668121159
step: 110, loss: 0.011281769722700119
step: 120, loss: 0.002385897096246481
step: 130, loss: 0.016290796920657158
step: 140, loss: 0.00764966057613492
step: 150, loss: 0.003386370139196515
step: 160, loss: 0.02572176232933998
step: 170, loss: 0.0014000984374433756
step: 180, loss: 0.0013485309900715947
step: 190, loss: 0.19229605793952942
/usr/local/lib/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=0.8164383561643836, f1=0.8314606741573033,
best_f1=0.8314606741573033
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.01285574585199356
step: 10, loss: 0.0061126709915697575
step: 20, loss: 0.001953793689608574
step: 30, loss: 0.008387657813727856
step: 40, loss: 0.0017681436147540808
step: 50, loss: 0.01257795374840498
step: 60, loss: 0.00463897967711091
step: 70, loss: 0.010537656024098396
step: 80, loss: 0.001601331285201013
step: 90, loss: 0.003696844447404146
step: 100, loss: 0.00035799015313386917
step: 110, loss: 0.0007079343195073307
step: 120, loss: 0.008380675688385963
step: 130, loss: 0.0027066529728472233

```

```

step: 140, loss: 0.0002580313594080508
step: 150, loss: 0.005470717791467905
step: 160, loss: 0.0008978981641121209
step: 170, loss: 0.0005249838577583432
step: 180, loss: 0.03112836182117462
step: 190, loss: 0.024622172117233276
/usr/local/lib/python3.7/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=0.8097826086956521, f1=0.8287292817679558,
best_f1=0.8314606741573033
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.011271062307059765
step: 10, loss: 0.002715321956202388
step: 20, loss: 0.008908742107450962
step: 30, loss: 0.024388499557971954
step: 40, loss: 0.0013666542945429683
step: 50, loss: 0.0011275794822722673
step: 60, loss: 0.16013002395629883
step: 70, loss: 0.013943035155534744
step: 80, loss: 0.0014228124637156725
step: 90, loss: 0.006333461031317711
step: 100, loss: 0.002611895091831684
step: 110, loss: 0.10533315688371658
step: 120, loss: 0.0006198053597472608
step: 130, loss: 0.0010392325930297375
step: 140, loss: 0.0013081220677122474
step: 150, loss: 0.00039143391768448055
step: 160, loss: 0.004219063092023134
step: 170, loss: 0.13883104920387268
step: 180, loss: 0.003649107413366437
step: 190, loss: 0.10686155408620834
/usr/local/lib/python3.7/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=0.8081841432225065, f1=0.8272251308900523,
best_f1=0.8314606741573033
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                 : 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 [00:09, 217.26it/s]
load_f1 = 0.7146282973621103
real_f1 = 0.7173396674584323
733it [00:00, 3206.87it/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, 181.21it/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/

```

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: 0.49010276794433594
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 10, loss: 0.5194310545921326
step: 20, loss: 0.319023996591568
step: 30, loss: 0.41302621364593506
step: 40, loss: 0.5559542179107666
step: 50, loss: 0.32027530670166016
step: 60, loss: 0.5609486103057861
step: 70, loss: 0.2999364137649536
step: 80, loss: 0.2550352215766907
step: 90, loss: 0.23540018498897552
step: 100, loss: 0.17442084848880768
step: 110, loss: 0.42830097675323486
step: 120, loss: 0.31859803199768066
step: 130, loss: 0.3481338322162628
step: 140, loss: 0.3469024896621704
step: 150, loss: 0.3355696201324463
step: 160, loss: 0.34595656394958496
step: 170, loss: 0.2735685706138611
/usr/local/lib/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=0.540650406504065, f1=0.540084388185654,
best_f1=0.540084388185654
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.2300741821527481
step: 10, loss: 0.3218590021133423
step: 20, loss: 0.22512437403202057
step: 30, loss: 0.14948469400405884
step: 40, loss: 0.10641095042228699
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: 50, loss: 0.1574273556470871
step: 60, loss: 0.12474176287651062
step: 70, loss: 0.26913151144981384
```

```

step: 80, loss: 0.14030323922634125
step: 90, loss: 0.09151621162891388
step: 100, loss: 0.28927990794181824
step: 110, loss: 0.1668616086244583
step: 120, loss: 0.040234409272670746
step: 130, loss: 0.17619848251342773
step: 140, loss: 0.18451941013336182
step: 150, loss: 0.08058889955282211
step: 160, loss: 0.15982602536678314
step: 170, loss: 0.06398415565490723
/usr/local/lib/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=0.7579462102689486, f1=0.7506053268765133,
best_f1=0.7506053268765133
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.2110593467950821
step: 10, loss: 0.0831843838095665
step: 20, loss: 0.1884152591228485
step: 30, loss: 0.05716912820935249
step: 40, loss: 0.06561034917831421
step: 50, loss: 0.13929779827594757
step: 60, loss: 0.04097763076424599
step: 70, loss: 0.14796826243400574
step: 80, loss: 0.06686688959598541
step: 90, loss: 0.12856528162956238
step: 100, loss: 0.08464375138282776
step: 110, loss: 0.0700530931353569
step: 120, loss: 0.0663740262389183
step: 130, loss: 0.08004765957593918
step: 140, loss: 0.14042213559150696
step: 150, loss: 0.07809337973594666
step: 160, loss: 0.03508491441607475
step: 170, loss: 0.03867239132523537
/usr/local/lib/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=0.8217054263565892, f1=0.8295165394402035,

```



```

best_f1=0.8295165394402035
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.030557166785001755
step: 10, loss: 0.04859202727675438
step: 20, loss: 0.00927011203020811
step: 30, loss: 0.15069977939128876
step: 40, loss: 0.0862884670495987
step: 50, loss: 0.07349174469709396
step: 60, loss: 0.5084205865859985
step: 70, loss: 0.03915085643529892
step: 80, loss: 0.2071535885334015
step: 90, loss: 0.20729470252990723
step: 100, loss: 0.06130681559443474
step: 110, loss: 0.18020124733448029
step: 120, loss: 0.18878725171089172
step: 130, loss: 0.0709327980875969
step: 140, loss: 0.11522123962640762
step: 150, loss: 0.284432590007782
step: 160, loss: 0.004916605539619923
step: 170, loss: 0.10027725994586945
/usr/local/lib/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=0.8436018957345971, f1=0.8590909090909091,
best_f1=0.8590909090909091
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.017453664913773537
step: 10, loss: 0.05779789388179779
step: 20, loss: 0.03557255119085312
step: 30, loss: 0.1445176601409912
step: 40, loss: 0.03377673402428627
step: 50, loss: 0.05423358827829361
step: 60, loss: 0.1432185024023056
step: 70, loss: 0.040925126522779465
step: 80, loss: 0.002480302704498172
step: 90, loss: 0.020396243780851364
step: 100, loss: 0.022584950551390648
step: 110, loss: 0.05448553338646889
step: 120, loss: 0.06779459863901138
step: 130, loss: 0.003299910807982087

```

```

step: 140, loss: 0.038444884121418
step: 150, loss: 0.007811662741005421
step: 160, loss: 0.005026435945183039
step: 170, loss: 0.012619076296687126
/usr/local/lib/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=0.8450000000000001, f1=0.883495145631068,
best_f1=0.883495145631068
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.1529625803232193
step: 10, loss: 0.013217302970588207
step: 20, loss: 0.046746641397476196
step: 30, loss: 0.005098366644233465
step: 40, loss: 0.012020055204629898
step: 50, loss: 0.0015909125795587897
step: 60, loss: 0.0068725175224244595
step: 70, loss: 0.06683771312236786
step: 80, loss: 0.03255106508731842
step: 90, loss: 0.0884452611207962
step: 100, loss: 0.0037789279595017433
step: 110, loss: 0.005167996045202017
step: 120, loss: 0.06876006722450256
step: 130, loss: 0.08997875452041626
step: 140, loss: 0.025086915120482445
step: 150, loss: 0.01202399656176567
step: 160, loss: 0.10346207022666931
step: 170, loss: 0.016099631786346436
/usr/local/lib/python3.7/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=0.8391959798994975, f1=0.8753056234718827,
best_f1=0.883495145631068
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.024241948500275612
step: 10, loss: 0.021630655974149704
step: 20, loss: 0.007023070007562637
step: 30, loss: 0.09008078277111053
step: 40, loss: 0.0019780530128628016
step: 50, loss: 0.0004225738230161369

```

```

step: 60, loss: 0.004767242819070816
step: 70, loss: 0.01414753869175911
step: 80, loss: 0.0908842384815216
step: 90, loss: 0.03426520526409149
step: 100, loss: 0.0008119124686345458
step: 110, loss: 0.026260562241077423
step: 120, loss: 0.010915661230683327
step: 130, loss: 0.004184776917099953
step: 140, loss: 0.011227346025407314
step: 150, loss: 0.07998758554458618
step: 160, loss: 0.02276971936225891
step: 170, loss: 0.019153552129864693
/usr/local/lib/python3.7/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=0.8442211055276382, f1=0.8571428571428571,
best_f1=0.883495145631068
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.013702222146093845
step: 10, loss: 0.024916771799325943
step: 20, loss: 0.03760712593793869
step: 30, loss: 0.00019523689115885645
step: 40, loss: 0.00301631772890687
step: 50, loss: 0.002072283299639821
step: 60, loss: 0.0073394873179495335
step: 70, loss: 0.018223728984594345
step: 80, loss: 0.0006724553531967103
step: 90, loss: 0.03911643102765083
step: 100, loss: 0.0002155398397007957
step: 110, loss: 0.0389416366815567
step: 120, loss: 0.015814045444130898
step: 130, loss: 0.0068559711799025536
step: 140, loss: 0.16510142385959625
step: 150, loss: 0.004153816029429436
step: 160, loss: 0.0016220244579017162
step: 170, loss: 0.04434197023510933
/usr/local/lib/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=0.8578431372549019, f1=0.8699763593380615,
best_f1=0.8699763593380615
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:

```

```

An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.031761638820171356
step: 10, loss: 0.004387764260172844
step: 20, loss: 0.0012959972955286503
step: 30, loss: 0.012796529568731785
step: 40, loss: 0.0018643420189619064
step: 50, loss: 0.0001363919145660475
step: 60, loss: 0.06515432894229889
step: 70, loss: 0.0004413092974573374
step: 80, loss: 0.002926016226410866
step: 90, loss: 0.008174971677362919
step: 100, loss: 0.004400677513331175
step: 110, loss: 0.1276959925889969
step: 120, loss: 0.0014084827853366733
step: 130, loss: 0.0004050096613354981
step: 140, loss: 0.01668861322104931
step: 150, loss: 0.005358417052775621
step: 160, loss: 0.001518524019047618
step: 170, loss: 0.00599901657551527
/usr/local/lib/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=0.8615384615384616, f1=0.8731707317073171,
best_f1=0.8731707317073171
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00436368677765131
step: 10, loss: 0.042021751403808594
step: 20, loss: 0.05720031261444092
step: 30, loss: 0.04112917557358742
step: 40, loss: 0.0011877876240760088
step: 50, loss: 0.014320841059088707
step: 60, loss: 0.0014834951143711805
step: 70, loss: 0.05322641879320145
step: 80, loss: 0.016169728711247444
step: 90, loss: 0.00020494013733696193
step: 100, loss: 0.00668115122243762
step: 110, loss: 0.011103978380560875
step: 120, loss: 0.016562342643737793
step: 130, loss: 0.0002951652277261019
step: 140, loss: 0.00012412064825184643
step: 150, loss: 7.012181595200673e-05

```

```

step: 160, loss: 0.0015419236151501536
step: 170, loss: 0.014142556115984917
/usr/local/lib/python3.7/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=0.8601036269430052, f1=0.8822055137844612,
best_f1=0.8731707317073171
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.021610796451568604
step: 10, loss: 0.0213010311126709
step: 20, loss: 0.0002267305535497144
step: 30, loss: 0.011113767512142658
step: 40, loss: 0.001778175006620586
step: 50, loss: 0.004515563137829304
step: 60, loss: 0.017810989171266556
step: 70, loss: 0.0003743080887943506
step: 80, loss: 0.024238115176558495
step: 90, loss: 0.0005268657114356756
step: 100, loss: 0.004271375015377998
step: 110, loss: 0.00016767725173849612
step: 120, loss: 0.0004965534899383783
step: 130, loss: 0.011350876651704311
step: 140, loss: 0.0014501223340630531
step: 150, loss: 0.00010917047620750964
step: 160, loss: 0.021473353728652
step: 170, loss: 0.006817016750574112
/usr/local/lib/python3.7/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=0.8585607940446649, f1=0.8516746411483254,
best_f1=0.8731707317073171
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.001456680241972208
step: 10, loss: 0.0007332711829803884
step: 20, loss: 0.00018157489830628037
step: 30, loss: 0.03978349640965462
step: 40, loss: 8.98008220246993e-05
step: 50, loss: 0.0003724255075212568
step: 60, loss: 0.015903787687420845
step: 70, loss: 4.6515855501638725e-05
step: 80, loss: 5.612139284494333e-05
step: 90, loss: 0.000758194422814995
step: 100, loss: 0.00011044973507523537
step: 110, loss: 0.0008936421945691109

```

```

step: 120, loss: 0.019272690638899803
step: 130, loss: 0.0025709124747663736
step: 140, loss: 8.907991286832839e-05
step: 150, loss: 0.0601656474173069
step: 160, loss: 0.005934497807174921
step: 170, loss: 0.0003121339250355959
/usr/local/lib/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=0.8647214854111406, f1=0.880597014925373,
best_f1=0.880597014925373
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00014444574480876327
step: 10, loss: 0.0007142752292566001
step: 20, loss: 0.0006081031751818955
step: 30, loss: 0.04111752286553383
step: 40, loss: 0.00022144766990095377
step: 50, loss: 0.003409275086596608
step: 60, loss: 0.002318992745131254
step: 70, loss: 0.024707753211259842
step: 80, loss: 0.00024355655477847904
step: 90, loss: 0.003095713211223483
step: 100, loss: 0.009847949258983135
step: 110, loss: 0.0003170525305904448
step: 120, loss: 0.010560132563114166
step: 130, loss: 0.00014781222853343934
step: 140, loss: 0.062399473041296005
step: 150, loss: 0.0003042052558157593
step: 160, loss: 0.0002893870114348829
step: 170, loss: 0.0065491278655827045
/usr/local/lib/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=0.8894348894348895,
best_f1=0.8894348894348895
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")

```

```

step: 0, loss: 0.025352926924824715
step: 10, loss: 0.004231694620102644
step: 20, loss: 0.0002863703412003815
step: 30, loss: 0.02894543670117855
step: 40, loss: 0.001945263473317027
step: 50, loss: 0.0009616558672860265
step: 60, loss: 0.009578770026564598
step: 70, loss: 0.0033253857400268316
step: 80, loss: 0.0006149326218292117
step: 90, loss: 0.00031783245503902435
step: 100, loss: 0.0007363494951277971
step: 110, loss: 0.020664038136601448
step: 120, loss: 0.00016300789138767868
step: 130, loss: 0.04101790860295296
step: 140, loss: 0.03436487540602684
step: 150, loss: 0.021306538954377174
step: 160, loss: 0.0004397889715619385
step: 170, loss: 0.03285538777709007
/usr/local/lib/python3.7/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=0.8702290076335877, f1=0.8719211822660098,
best_f1=0.8894348894348895
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 3.5824719816446304e-05
step: 10, loss: 0.04534517601132393
step: 20, loss: 0.0011615960393100977
step: 30, loss: 0.0022192385513335466
step: 40, loss: 7.108335557859391e-05
step: 50, loss: 4.9834146921057254e-05
step: 60, loss: 0.0002869473828468472
step: 70, loss: 0.07959194481372833
step: 80, loss: 0.00010441240738146007
step: 90, loss: 0.0194824431091547
step: 100, loss: 0.034344382584095
step: 110, loss: 0.00019019316823687404
step: 120, loss: 7.019969052635133e-05
step: 130, loss: 0.00037740039988420904
step: 140, loss: 0.0322285033762455
step: 150, loss: 0.0006633492303080857
step: 160, loss: 8.424588304478675e-05
step: 170, loss: 0.0020319472532719374
/usr/local/lib/python3.7/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=0.8623376623376624, f1=0.8706467661691543,

```

```
best_f1=0.8894348894348895
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 [00:06, 292.80it/s]
```

```
load_f1 = 0.38048780487804873
```

```
real_f1 = 0.346666666666666673
```

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, 179.81it/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, 429kB/s]
Downloading: 100% 899k/899k [00:01<00:00, 663kB/s]
Downloading: 100% 456k/456k [00:00<00:00, 503kB/s]
Downloading: 100% 501M/501M [00:07<00:00, 65.5MB/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: 0.5590389966964722
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 10, loss: 0.5002449750900269
step: 20, loss: 0.46249833703041077
step: 30, loss: 0.32963353395462036
step: 40, loss: 0.2778749167919159
step: 50, loss: 0.32605987787246704
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 8192.0
step: 60, loss: 0.13435427844524384
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 4096.0
step: 70, loss: 0.14397946000099182
step: 80, loss: 0.23299908638000488
step: 90, loss: 0.18722543120384216
step: 100, loss: 0.20928028225898743
step: 110, loss: 0.15669581294059753
step: 120, loss: 0.1520737260580063
step: 130, loss: 0.07708626985549927
step: 140, loss: 0.014633585698902607
step: 150, loss: 0.22405731678009033
step: 160, loss: 0.08602719753980637
step: 170, loss: 0.20835788547992706
step: 180, loss: 0.20830488204956055
step: 190, loss: 0.04173213988542557
step: 200, loss: 0.05206559598445892
step: 210, loss: 0.07175896316766739
step: 220, loss: 0.09271375089883804
step: 230, loss: 0.040392547845840454

```

```

/usr/local/lib/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=0.9511111111111112, f1=0.9572072072072072,
best_f1=0.9572072072072072
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.008876264095306396
step: 10, loss: 0.16282671689987183
step: 20, loss: 0.06934989243745804
step: 30, loss: 0.046789463609457016
step: 40, loss: 0.05833005532622337
step: 50, loss: 0.11078061163425446
step: 60, loss: 0.0244491808116436
step: 70, loss: 0.01987435668706894
step: 80, loss: 0.06586655229330063
step: 90, loss: 0.016266178339719772
step: 100, loss: 0.011983667500317097
step: 110, loss: 0.009227056987583637
step: 120, loss: 0.009595858864486217
step: 130, loss: 0.014045978896319866
step: 140, loss: 0.03010598011314869
step: 150, loss: 0.10542688518762589
step: 160, loss: 0.04740702360868454
step: 170, loss: 0.001814771443605423
step: 180, loss: 0.009511174634099007
step: 190, loss: 0.008443239144980907
step: 200, loss: 0.026588426902890205
step: 210, loss: 0.01656336709856987
step: 220, loss: 0.039098791778087616
step: 230, loss: 0.0023090827744454145
/usr/local/lib/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=0.961883408071749, f1=0.9590909090909091,
best_f1=0.9590909090909091
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")

```

```

step: 0, loss: 0.05080454424023628
step: 10, loss: 0.003119959030300379
step: 20, loss: 0.0022576828487217426
step: 30, loss: 0.009573020972311497
step: 40, loss: 0.0633724182844162
step: 50, loss: 0.0029667168855667114
step: 60, loss: 0.007985192351043224
step: 70, loss: 0.1474623829126358
step: 80, loss: 0.01662452332675457
step: 90, loss: 0.00894701387733221
step: 100, loss: 0.0388542041182518
step: 110, loss: 0.006835360545665026
step: 120, loss: 0.0016621522372588515
step: 130, loss: 0.0062571982853114605
step: 140, loss: 0.007041057571768761
step: 150, loss: 0.10149605572223663
step: 160, loss: 0.003862140467390418
step: 170, loss: 0.001707739313133061
step: 180, loss: 0.02550848014652729
step: 190, loss: 0.03275108337402344
step: 200, loss: 0.015710093080997467
step: 210, loss: 0.0036216392181813717
step: 220, loss: 0.13284960389137268
step: 230, loss: 0.014392430894076824
/usr/local/lib/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=0.9764309764309763, f1=0.9672316384180792,
best_f1=0.9672316384180792
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.018998317420482635
step: 10, loss: 0.0035267246421426535
step: 20, loss: 0.0069399019703269005
step: 30, loss: 0.0010110011789947748
step: 40, loss: 0.10288895666599274
step: 50, loss: 0.03252794221043587
step: 60, loss: 0.00636560982093215
step: 70, loss: 0.07242341339588165
step: 80, loss: 0.10070707648992538
step: 90, loss: 0.0812988430261612
step: 100, loss: 0.00975632481276989
step: 110, loss: 0.004891911521553993

```

```

step: 120, loss: 0.009860473684966564
step: 130, loss: 0.043512556701898575
step: 140, loss: 0.019150134176015854
step: 150, loss: 0.001956071238964796
step: 160, loss: 0.012039711698889732
step: 170, loss: 0.00418445747345686
step: 180, loss: 0.031163634732365608
step: 190, loss: 0.002646433189511299
step: 200, loss: 0.07090921700000763
step: 210, loss: 0.0812133252620697
step: 220, loss: 0.0033443348947912455
step: 230, loss: 0.003969802055507898
/usr/local/lib/python3.7/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=0.9743016759776536, f1=0.9706546275395034,
best_f1=0.9672316384180792
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0008764320518821478
step: 10, loss: 0.004367002751678228
step: 20, loss: 0.045865487307310104
step: 30, loss: 0.0035576571244746447
step: 40, loss: 0.007287757005542517
step: 50, loss: 0.002448836574330926
step: 60, loss: 0.0025546010583639145
step: 70, loss: 0.0013177931541576982
step: 80, loss: 0.014353899285197258
step: 90, loss: 0.054778847843408585
step: 100, loss: 0.0010106057161465287
step: 110, loss: 0.0008980968268588185
step: 120, loss: 0.0017204777104780078
step: 130, loss: 0.0010007923701778054
step: 140, loss: 0.00251578725874424
step: 150, loss: 0.008259815163910389
step: 160, loss: 0.0007321587763726711
step: 170, loss: 0.002930195303633809
step: 180, loss: 0.005918445996940136
step: 190, loss: 0.12139478325843811
step: 200, loss: 0.003912093583494425
step: 210, loss: 0.0017600596183910966
step: 220, loss: 0.010193523950874805
step: 230, loss: 0.006326672155410051
/usr/local/lib/python3.7/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=0.9761634506242906, f1=0.9597238204833141,

```

```

best_f1=0.9672316384180792
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0034894910641014576
step: 10, loss: 0.0026658466085791588
step: 20, loss: 0.0016642019618302584
step: 30, loss: 0.0006596704479306936
step: 40, loss: 0.00052338105160743
step: 50, loss: 0.0006622977089136839
step: 60, loss: 0.0009859332349151373
step: 70, loss: 0.16875769197940826
step: 80, loss: 0.0019027404487133026
step: 90, loss: 0.0033460943959653378
step: 100, loss: 0.0011739309411495924
step: 110, loss: 0.010459682904183865
step: 120, loss: 0.0012457799166440964
step: 130, loss: 0.0011825573164969683
step: 140, loss: 0.003935706801712513
step: 150, loss: 0.005316666327416897
step: 160, loss: 0.005191100295633078
step: 170, loss: 0.006289668381214142
step: 180, loss: 0.000427121325628832
step: 190, loss: 0.0003306346188765019
step: 200, loss: 0.0015607759123668075
step: 210, loss: 0.0014192754169926047
step: 220, loss: 0.0610097274184227
step: 230, loss: 0.0018149679526686668
/usr/local/lib/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=0.9782857142857143, f1=0.9738339021615473,
best_f1=0.9738339021615473
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.001775477547198534
step: 10, loss: 0.0007527230773121119
step: 20, loss: 0.0012890624348074198
step: 30, loss: 0.0009798755636438727
step: 40, loss: 0.001203481457196176
step: 50, loss: 0.03715460002422333
step: 60, loss: 0.0015592710115015507
step: 70, loss: 0.010675081983208656

```

```

step: 80, loss: 0.015021707862615585
step: 90, loss: 0.005443382542580366
step: 100, loss: 0.003117184154689312
step: 110, loss: 0.0007647420279681683
step: 120, loss: 0.0005544920568354428
step: 130, loss: 0.0030184166971594095
step: 140, loss: 0.0008047962328419089
step: 150, loss: 0.044492945075035095
step: 160, loss: 0.0005738445906899869
step: 170, loss: 0.03457023203372955
step: 180, loss: 0.0010694321244955063
step: 190, loss: 0.0010087746195495129
step: 200, loss: 0.0028301419224590063
step: 210, loss: 0.0019519024062901735
step: 220, loss: 0.0015077515272423625
step: 230, loss: 0.0017423059325665236
/usr/local/lib/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=0.983050847457627, f1=0.9737742303306728,
best_f1=0.9737742303306728
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0024567553773522377
step: 10, loss: 0.003734566504135728
step: 20, loss: 0.009708719328045845
step: 30, loss: 0.0014933435013517737
step: 40, loss: 0.0013504510279744864
step: 50, loss: 0.0021251640282571316
step: 60, loss: 0.0004348848888184875
step: 70, loss: 0.00016084290109574795
step: 80, loss: 0.02360565960407257
step: 90, loss: 0.0005059735849499702
step: 100, loss: 0.008084507659077644
step: 110, loss: 0.0024726546835154295
step: 120, loss: 0.0006037665298208594
step: 130, loss: 0.0020876647904515266
step: 140, loss: 0.0004540120135061443
step: 150, loss: 0.11995416134595871
step: 160, loss: 0.0012721734819933772
step: 170, loss: 0.02769819088280201
step: 180, loss: 0.0007951320731081069
step: 190, loss: 0.20086024701595306

```

```

step: 200, loss: 0.01712074875831604
step: 210, loss: 0.00539805693551898
step: 220, loss: 0.001158643513917923
step: 230, loss: 0.000567821494769305
/usr/local/lib/python3.7/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=0.9753914988814317, f1=0.9751693002257337,
best_f1=0.9737742303306728
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.001512962393462658
step: 10, loss: 0.00835393089801073
step: 20, loss: 0.029500683769583702
step: 30, loss: 0.001689296099357307
step: 40, loss: 0.0006335757207125425
step: 50, loss: 0.0017673051916062832
step: 60, loss: 0.013249123468995094
step: 70, loss: 0.11510413885116577
step: 80, loss: 0.0005430823657661676
step: 90, loss: 0.01674739643931389
step: 100, loss: 0.006926300935447216
step: 110, loss: 0.05998816713690758
step: 120, loss: 0.014749906957149506
step: 130, loss: 0.0009336175280623138
step: 140, loss: 0.0011852113530039787
step: 150, loss: 0.005729489494115114
step: 160, loss: 0.0035600843839347363
step: 170, loss: 0.0002837661886587739
step: 180, loss: 0.0007952682208269835
step: 190, loss: 0.0002298820181749761
step: 200, loss: 0.014588548801839352
step: 210, loss: 0.18189887702465057
step: 220, loss: 0.0012435121461749077
step: 230, loss: 0.0002818311913870275
/usr/local/lib/python3.7/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=0.9785310734463276, f1=0.976054732041049,
best_f1=0.9737742303306728
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0004407633387017995
step: 10, loss: 0.0018396996892988682
step: 20, loss: 0.0008099167025648057
step: 30, loss: 0.00022026384249329567

```



```

step: 40, loss: 0.0008325439994223416
step: 50, loss: 0.00017427460988983512
step: 60, loss: 0.00047756280400790274
step: 70, loss: 0.016363754868507385
step: 80, loss: 0.003825288498774171
step: 90, loss: 0.0004755085101351142
step: 100, loss: 0.00027340915403328836
step: 110, loss: 0.0011966486927121878
step: 120, loss: 0.00836675614118576
step: 130, loss: 0.0009078617440536618
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 4096.0
step: 140, loss: 0.001342790899798274
step: 150, loss: 0.06483083218336105
step: 160, loss: 0.0012474070535972714
step: 170, loss: 0.003342608455568552
step: 180, loss: 0.02375645563006401
step: 190, loss: 0.008231544867157936
step: 200, loss: 0.0006797741516493261
step: 210, loss: 0.020311735570430756
step: 220, loss: 0.0004307323833927512
step: 230, loss: 0.0005108733894303441
/usr/local/lib/python3.7/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=0.9740112994350283, f1=0.9728506787330317,
best_f1=0.9737742303306728
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0005804618122056127
step: 10, loss: 0.000537108804564923
step: 20, loss: 0.0005562698934227228
step: 30, loss: 0.0014927067095413804
step: 40, loss: 0.00016358558787032962
step: 50, loss: 0.00023354623408522457
step: 60, loss: 0.004780137911438942
step: 70, loss: 0.00036760116927325726
step: 80, loss: 0.0005678044981323183
step: 90, loss: 0.14308218657970428
step: 100, loss: 0.0004337129939813167
step: 110, loss: 0.014832425862550735
step: 120, loss: 0.006978536956012249
step: 130, loss: 0.00022309744963422418
step: 140, loss: 0.006642024032771587
step: 150, loss: 0.0006830415804870427
step: 160, loss: 0.004134808201342821
step: 170, loss: 0.0008274375577457249
step: 180, loss: 0.0016250257613137364

```

```

step: 190, loss: 0.00031529925763607025
step: 200, loss: 0.0013804135378450155
step: 210, loss: 0.0005855982308275998
step: 220, loss: 0.0007696808315813541
step: 230, loss: 0.000476013810839504
/usr/local/lib/python3.7/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=0.9752252252252253, f1=0.9762174405436014,
best_f1=0.9737742303306728
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0004944141255691648
step: 10, loss: 0.00019221450202167034
step: 20, loss: 0.010824046097695827
step: 30, loss: 0.0193479023873806
step: 40, loss: 0.0014686400536447763
step: 50, loss: 0.0036108682397753
step: 60, loss: 0.0004758824943564832
step: 70, loss: 0.00015627358516212553
step: 80, loss: 8.744386286707595e-05
step: 90, loss: 0.000926309556234628
step: 100, loss: 0.0002076769305858761
step: 110, loss: 0.00019612156029324979
step: 120, loss: 0.0009168636752292514
step: 130, loss: 0.0002818428329192102
step: 140, loss: 0.000571102078538388
step: 150, loss: 0.0010690436465665698
step: 160, loss: 0.0065476675517857075
step: 170, loss: 0.0007189156021922827
step: 180, loss: 0.0007681918214075267
step: 190, loss: 0.008740263991057873
step: 200, loss: 0.0007593158516101539
step: 210, loss: 0.0005417648935690522
step: 220, loss: 0.011795799247920513
step: 230, loss: 0.000789787620306015
/usr/local/lib/python3.7/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=0.9785794813979707, f1=0.9740112994350283,
best_f1=0.9737742303306728
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0012193990405648947
step: 10, loss: 0.002167463768273592
step: 20, loss: 0.0007755612023174763

```

```

step: 30, loss: 0.00011430019367253408
step: 40, loss: 0.0012664623791351914
step: 50, loss: 0.000929126632399857
step: 60, loss: 0.00021305744303390384
step: 70, loss: 0.0030134047847241163
step: 80, loss: 0.006873562932014465
step: 90, loss: 0.000519985449500382
step: 100, loss: 0.0006969112437218428
step: 110, loss: 0.0008448171429336071
step: 120, loss: 0.0005219014710746706
step: 130, loss: 0.009569162502884865
step: 140, loss: 0.0007654592045582831
step: 150, loss: 0.00029317860025912523
step: 160, loss: 0.005878959316760302
step: 170, loss: 0.00040913833072409034
step: 180, loss: 0.02872474119067192
step: 190, loss: 0.00017126752936746925
step: 200, loss: 5.880032404093072e-05
step: 210, loss: 0.00033293021260760725
step: 220, loss: 0.00015400198753923178
step: 230, loss: 0.0010411542607471347
/usr/local/lib/python3.7/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=0.9796380090497738, f1=0.9738339021615473,
best_f1=0.9737742303306728
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00043155785533599555
step: 10, loss: 0.0001147369112004526
step: 20, loss: 0.0001462110667489469
step: 30, loss: 0.0002742572105489671
step: 40, loss: 0.00019366387277841568
step: 50, loss: 0.00018376664957031608
step: 60, loss: 0.00033081145375035703
step: 70, loss: 0.0002528868499211967
step: 80, loss: 0.00039444788126274943
step: 90, loss: 0.0007410874241031706
step: 100, loss: 0.0002875987847801298
step: 110, loss: 0.000194416381418705
step: 120, loss: 3.49578658642713e-05
step: 130, loss: 0.000449606915935874
step: 140, loss: 0.003178945044055581
step: 150, loss: 0.0005364902317523956
step: 160, loss: 0.0004739846335723996
step: 170, loss: 0.00015884212916716933
step: 180, loss: 0.00016487178800161928

```

```

step: 190, loss: 7.01579701853916e-05
step: 200, loss: 0.0027983002364635468
step: 210, loss: 0.00010831780673470348
step: 220, loss: 0.00017930915055330843
step: 230, loss: 0.00031480463803745806
/usr/local/lib/python3.7/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=0.9785310734463276, f1=0.9762174405436014,
best_f1=0.9737742303306728
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00539482431486249
step: 10, loss: 6.916259007994086e-05
step: 20, loss: 0.00043736951192840934
step: 30, loss: 0.0001726114860503003
step: 40, loss: 4.546922355075367e-05
step: 50, loss: 6.613929872401059e-05
step: 60, loss: 0.05538269504904747
step: 70, loss: 0.00028460242901928723
step: 80, loss: 0.00016855089052114636
step: 90, loss: 0.0001604340213816613
step: 100, loss: 5.046693695476279e-05
step: 110, loss: 0.000112894871563185
step: 120, loss: 0.046775054186582565
step: 130, loss: 0.00017188051424454898
step: 140, loss: 0.012829378247261047
step: 150, loss: 0.00024040078278630972
step: 160, loss: 0.0132053904235363
step: 170, loss: 3.991255653090775e-05
step: 180, loss: 0.00015035834803711623
step: 190, loss: 0.0001528222783235833
step: 200, loss: 0.0005606774939224124
step: 210, loss: 0.0001548241125419736
step: 220, loss: 0.0003376179956831038
step: 230, loss: 6.519549788208678e-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=0.9785310734463276, f1=0.976271186440678,
best_f1=0.9737742303306728
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 [00:15, 156.34it/s]
load_f1 = 0.9853107344632768
real_f1 = 0.9819413092550789
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 [00:30, 144.30it/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 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: 0.6697006821632385

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: 0.42482995986938477  
step: 20, loss: 0.2766985595226288  
step: 30, loss: 0.3048263192176819  
step: 40, loss: 0.2817775011062622  
step: 50, loss: 0.14254218339920044  
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 8192.0  
step: 60, loss: 0.23891796171665192  
step: 70, loss: 0.2215806096792221  
step: 80, loss: 0.08493991196155548  
step: 90, loss: 0.3684939444065094  
step: 100, loss: 0.20191404223442078  
step: 110, loss: 0.16636905074119568  
step: 120, loss: 0.4916318655014038  
step: 130, loss: 0.18801818788051605  
step: 140, loss: 0.21646051108837128  
step: 150, loss: 0.09096415340900421  
step: 160, loss: 0.24299974739551544  
step: 170, loss: 0.10577455163002014  
step: 180, loss: 0.059722695499658585  
step: 190, loss: 0.2353293001651764  
step: 200, loss: 0.09523861855268478  
step: 210, loss: 0.07635896652936935  
step: 220, loss: 0.08411570638418198  
step: 230, loss: 0.23592601716518402  
step: 240, loss: 0.07038753479719162  
step: 250, loss: 0.08807475119829178  
step: 260, loss: 0.2028736174106598  
step: 270, loss: 0.2522951066493988  
step: 280, loss: 0.029201863333582878  
step: 290, loss: 0.24671682715415955  
step: 300, loss: 0.052507344633340836  
step: 310, loss: 0.0980801060795784  
step: 320, loss: 0.10320968180894852  
step: 330, loss: 0.09699789434671402  
step: 340, loss: 0.3705860674381256  
step: 350, loss: 0.05930127948522568  
step: 360, loss: 0.05049298331141472  
step: 370, loss: 0.09354082494974136  
step: 380, loss: 0.0968519002199173  
step: 390, loss: 0.02204907312989235  
step: 400, loss: 0.057815078645944595  
step: 410, loss: 0.17021037638187408  
step: 420, loss: 0.03543499484658241  
step: 430, loss: 0.012207510881125927  
step: 440, loss: 0.044864434748888016  
step: 450, loss: 0.11591032147407532  
step: 460, loss: 0.08917752653360367

```

step: 470, loss: 0.01968301273882389
step: 480, loss: 0.2673661708831787
step: 490, loss: 0.07659497112035751
step: 500, loss: 0.04518282413482666
step: 510, loss: 0.025072526186704636
step: 520, loss: 0.2286616414785385
step: 530, loss: 0.02770482376217842
/usr/local/lib/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=0.9399358092618065, f1=0.9405850091407678,
best_f1=0.9405850091407678
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.11850927770137787
step: 10, loss: 0.11083777993917465
step: 20, loss: 0.1633167862892151
step: 30, loss: 0.026139769703149796
step: 40, loss: 0.1148395910859108
step: 50, loss: 0.03881968930363655
step: 60, loss: 0.014027444645762444
step: 70, loss: 0.033569734543561935
step: 80, loss: 0.07366305589675903
step: 90, loss: 0.02659616805613041
step: 100, loss: 0.25081318616867065
step: 110, loss: 0.016110675409436226
step: 120, loss: 0.1624959409236908
step: 130, loss: 0.005682327318936586
step: 140, loss: 0.03563275560736656
step: 150, loss: 0.07505583018064499
step: 160, loss: 0.022785713896155357
step: 170, loss: 0.0286654494702816
step: 180, loss: 0.04123426973819733
step: 190, loss: 0.04467933624982834
step: 200, loss: 0.0546259842813015
step: 210, loss: 0.03839515894651413
step: 220, loss: 0.001487046480178833
step: 230, loss: 0.08876004070043564
step: 240, loss: 0.09556561708450317
step: 250, loss: 0.02000131830573082
step: 260, loss: 0.04857221990823746
step: 270, loss: 0.05757991969585419
step: 280, loss: 0.14812280237674713

```



```

step: 290, loss: 0.027644777670502663
step: 300, loss: 0.030512256547808647
step: 310, loss: 0.06935875862836838
step: 320, loss: 0.04165484011173248
step: 330, loss: 0.06223583593964577
step: 340, loss: 0.10653398931026459
step: 350, loss: 0.005626088008284569
step: 360, loss: 0.054528020322322845
step: 370, loss: 0.015933044254779816
step: 380, loss: 0.26876136660575867
step: 390, loss: 0.04130322113633156
step: 400, loss: 0.02810104377567768
step: 410, loss: 0.01551960501819849
step: 420, loss: 0.11681908369064331
step: 430, loss: 0.25731751322746277
step: 440, loss: 0.004227197263389826
step: 450, loss: 0.07989617437124252
step: 460, loss: 0.017921220511198044
step: 470, loss: 0.048715267330408096
step: 480, loss: 0.015830496326088905
step: 490, loss: 0.1019844263792038
step: 500, loss: 0.0029289121739566326
step: 510, loss: 0.13924413919448853
step: 520, loss: 0.4036087393760681
step: 530, loss: 0.20844189822673798
/usr/local/lib/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=0.9406934306569343, f1=0.9417652411282984,
best_f1=0.9417652411282984
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.1633431315422058
step: 10, loss: 0.0671287253499031
step: 20, loss: 0.017226798459887505
step: 30, loss: 0.0609397366642952
step: 40, loss: 0.0925477147102356
step: 50, loss: 0.10476009547710419
step: 60, loss: 0.0433499850332737
step: 70, loss: 0.01863253302872181
step: 80, loss: 0.011910736560821533
step: 90, loss: 0.05366196110844612
step: 100, loss: 0.1888040453195572

```

```
step: 110, loss: 0.07603811472654343
step: 120, loss: 0.15573018789291382
step: 130, loss: 0.06559617072343826
step: 140, loss: 0.019029108807444572
step: 150, loss: 0.022139113396406174
step: 160, loss: 0.010360032320022583
step: 170, loss: 0.020111864432692528
step: 180, loss: 0.01346915028989315
step: 190, loss: 0.013483825139701366
step: 200, loss: 0.11957959085702896
step: 210, loss: 0.037187572568655014
step: 220, loss: 0.08251512795686722
step: 230, loss: 0.0856996700167656
step: 240, loss: 0.08140837401151657
step: 250, loss: 0.09853841364383698
step: 260, loss: 0.1761779934167862
step: 270, loss: 0.006857358850538731
step: 280, loss: 0.005285211838781834
step: 290, loss: 0.00586601672694087
step: 300, loss: 0.16299326717853546
step: 310, loss: 0.16263587772846222
step: 320, loss: 0.014376086182892323
step: 330, loss: 0.032322924584150314
step: 340, loss: 0.006190289743244648
step: 350, loss: 0.09204686433076859
step: 360, loss: 0.005800099112093449
step: 370, loss: 0.0993383526802063
step: 380, loss: 0.0018863441655412316
step: 390, loss: 0.0024386306758970022
step: 400, loss: 0.16195346415042877
step: 410, loss: 0.048396412283182144
step: 420, loss: 0.013087151572108269
step: 430, loss: 0.040630556643009186
step: 440, loss: 0.2496170550584793
step: 450, loss: 0.009232385084033012
step: 460, loss: 0.09675264358520508
step: 470, loss: 0.010295290499925613
step: 480, loss: 0.26602959632873535
step: 490, loss: 0.05206288769841194
step: 500, loss: 0.07094442844390869
step: 510, loss: 0.0052596693858504295
step: 520, loss: 0.010203525424003601
step: 530, loss: 0.01845373399555683
/usr/local/lib/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=0.9502762430939227, f1=0.9437269372693727,
best_f1=0.9437269372693727
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.07661855965852737
step: 10, loss: 0.0013236721279099584
step: 20, loss: 0.07374272495508194
step: 30, loss: 0.14805231988430023
step: 40, loss: 0.006854254752397537
step: 50, loss: 0.07737947255373001
step: 60, loss: 0.040915146470069885
step: 70, loss: 0.031095199286937714
step: 80, loss: 0.1242005005478859
step: 90, loss: 0.2078724205493927
step: 100, loss: 0.002045457949861884
step: 110, loss: 0.16446730494499207
step: 120, loss: 0.012053915299475193
step: 130, loss: 0.1137235090136528
step: 140, loss: 0.03104647807776928
step: 150, loss: 0.007266766391694546
step: 160, loss: 0.008552387356758118
step: 170, loss: 0.03852609917521477
step: 180, loss: 0.011150750331580639
step: 190, loss: 0.035624805837869644
step: 200, loss: 0.1507958620786667
step: 210, loss: 0.001832086592912674
step: 220, loss: 0.0058513423427939415
step: 230, loss: 0.02179870940744877
step: 240, loss: 0.011682654730975628
step: 250, loss: 0.08417682349681854
step: 260, loss: 0.005175299942493439
step: 270, loss: 0.09562189131975174
step: 280, loss: 0.029022419825196266
step: 290, loss: 0.05633404105901718
step: 300, loss: 0.002138789277523756
step: 310, loss: 0.01118482742458582
step: 320, loss: 0.11747857183218002
step: 330, loss: 0.04635664448142052
step: 340, loss: 0.02195911668241024
step: 350, loss: 0.10990853607654572
step: 360, loss: 0.02671167068183422
step: 370, loss: 0.003426506882533431
step: 380, loss: 0.028026657178997993
step: 390, loss: 0.0017970781773328781
step: 400, loss: 0.06959256529808044

```

```

step: 410, loss: 0.0012131903786212206
step: 420, loss: 0.004684193059802055
step: 430, loss: 0.05495302006602287
step: 440, loss: 0.00607256731018424
step: 450, loss: 0.05285625159740448
step: 460, loss: 0.1064915582537651
step: 470, loss: 0.011799881234765053
step: 480, loss: 0.0918971598148346
step: 490, loss: 0.009951089508831501
step: 500, loss: 0.02158990502357483
step: 510, loss: 0.1505831778049469
step: 520, loss: 0.02205829694867134
step: 530, loss: 0.13754770159721375
/usr/local/lib/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=0.9542056074766355, f1=0.9447565543071161,
best_f1=0.9447565543071161
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0016945890383794904
step: 10, loss: 0.048739682883024216
step: 20, loss: 0.041346099227666855
step: 30, loss: 0.07946321368217468
step: 40, loss: 0.008394918404519558
step: 50, loss: 0.027042176574468613
step: 60, loss: 0.03857841715216637
step: 70, loss: 0.0012305558193475008
step: 80, loss: 0.0016548385610803962
step: 90, loss: 0.10198512673377991
step: 100, loss: 0.03830462321639061
step: 110, loss: 0.00464896485209465
step: 120, loss: 0.0249002818018198
step: 130, loss: 0.0034557320177555084
step: 140, loss: 0.008202480152249336
step: 150, loss: 0.058819323778152466
step: 160, loss: 0.009572068229317665
step: 170, loss: 0.011615686118602753
step: 180, loss: 0.055878277868032455
step: 190, loss: 0.0698762908577919
step: 200, loss: 0.007105968426913023
step: 210, loss: 0.003605601843446493
step: 220, loss: 0.00932051707059145

```

```
step: 230, loss: 0.0019506351090967655
step: 240, loss: 0.010737366043031216
step: 250, loss: 0.12761788070201874
step: 260, loss: 0.0007994546322152019
step: 270, loss: 0.018653040751814842
step: 280, loss: 0.013206969015300274
step: 290, loss: 0.001188413007184863
step: 300, loss: 0.05841061472892761
step: 310, loss: 0.032105930149555206
step: 320, loss: 0.16082637012004852
step: 330, loss: 0.002248752396553755
step: 340, loss: 0.007726301904767752
step: 350, loss: 0.002841444918885827
step: 360, loss: 0.004654078744351864
step: 370, loss: 0.0008036816143430769
step: 380, loss: 0.0026700496673583984
step: 390, loss: 0.021099692210555077
step: 400, loss: 0.014050963334739208
step: 410, loss: 0.041918765753507614
step: 420, loss: 0.2040967494249344
step: 430, loss: 0.19978344440460205
step: 440, loss: 0.0075347451493144035
step: 450, loss: 0.014611993916332722
step: 460, loss: 0.016975373029708862
step: 470, loss: 0.017224902287125587
step: 480, loss: 0.019139951094985008
step: 490, loss: 0.05787867680191994
step: 500, loss: 0.008362739346921444
step: 510, loss: 0.01947741210460663
step: 520, loss: 0.125391885638237
step: 530, loss: 0.023186059668660164
/usr/local/lib/python3.7/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=0.951048951048951, f1=0.9423791821561338,
best_f1=0.9447565543071161
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.02193440869450569
step: 10, loss: 0.0027963221073150635
step: 20, loss: 0.013721349649131298
step: 30, loss: 0.0008170215296559036
step: 40, loss: 0.003211370436474681
step: 50, loss: 0.0009545133216306567
step: 60, loss: 0.029714742675423622
step: 70, loss: 0.00046315661165863276
step: 80, loss: 0.004436904098838568
```

```
step: 90, loss: 0.058803949505090714
step: 100, loss: 0.014932101592421532
step: 110, loss: 0.0035889367572963238
step: 120, loss: 0.07879437506198883
step: 130, loss: 0.010268568992614746
step: 140, loss: 0.0020059761591255665
step: 150, loss: 0.0032238962594419718
step: 160, loss: 0.09963731467723846
step: 170, loss: 0.04762892797589302
step: 180, loss: 0.00339901982806623
step: 190, loss: 0.113227479159832
step: 200, loss: 0.01951337605714798
step: 210, loss: 0.033852044492959976
step: 220, loss: 0.014307722449302673
step: 230, loss: 0.025862596929073334
step: 240, loss: 0.0018338320078328252
step: 250, loss: 0.050500333309173584
step: 260, loss: 0.0006081422325223684
step: 270, loss: 0.02647169679403305
step: 280, loss: 0.006856406573206186
step: 290, loss: 0.007457926869392395
step: 300, loss: 0.0073790960013866425
step: 310, loss: 0.025386033579707146
step: 320, loss: 0.0002959695120807737
step: 330, loss: 0.005118971690535545
step: 340, loss: 0.0009110938990488648
step: 350, loss: 0.018417201936244965
step: 360, loss: 0.004387013614177704
step: 370, loss: 0.038774292916059494
step: 380, loss: 0.001317375572398305
step: 390, loss: 0.007505771238356829
step: 400, loss: 0.0021684623789042234
step: 410, loss: 0.0005466509610414505
step: 420, loss: 0.012031152844429016
step: 430, loss: 0.002577077364549041
step: 440, loss: 0.023680338636040688
step: 450, loss: 0.1549837589263916
step: 460, loss: 0.003012373112142086
step: 470, loss: 0.0014337071916088462
step: 480, loss: 0.005941456649452448
step: 490, loss: 0.009398638270795345
step: 500, loss: 0.03486188128590584
step: 510, loss: 0.2158873975276947
step: 520, loss: 0.0014942694688215852
step: 530, loss: 0.010053979232907295
/usr/local/lib/python3.7/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=0.939679547596607, f1=0.9284020862968232,  
best_f1=0.9447565543071161  
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:  
An input tensor was not cuda.  
    warnings.warn("An input tensor was not cuda.")  
step: 0, loss: 0.0041841547936201096  
step: 10, loss: 0.0009535756544210017  
step: 20, loss: 0.0025403383187949657  
step: 30, loss: 0.01772012934088707  
step: 40, loss: 0.006639059633016586  
step: 50, loss: 0.03328147530555725  
step: 60, loss: 0.0036344430409371853  
step: 70, loss: 0.0011449980083853006  
step: 80, loss: 0.0040032099932432175  
step: 90, loss: 0.002095466712489724  
step: 100, loss: 0.058484502136707306  
step: 110, loss: 0.00041915549081750214  
step: 120, loss: 0.001046219258569181  
step: 130, loss: 0.00020191351359244436  
step: 140, loss: 0.0009225609828718007  
step: 150, loss: 0.0018878320697695017  
step: 160, loss: 0.00023387865803670138  
step: 170, loss: 0.003782265819609165  
step: 180, loss: 0.0019589969888329506  
step: 190, loss: 0.0038875604514032602  
step: 200, loss: 0.0006782782729715109  
step: 210, loss: 0.009992657229304314  
step: 220, loss: 0.000997658004052937  
step: 230, loss: 0.0007313641835935414  
step: 240, loss: 0.0344170480966568  
step: 250, loss: 0.01780339516699314  
step: 260, loss: 0.03895309939980507  
step: 270, loss: 0.005891980137676001  
step: 280, loss: 0.005473291501402855  
step: 290, loss: 0.00921713188290596  
step: 300, loss: 0.001480131410062313  
step: 310, loss: 0.0014971448108553886  
step: 320, loss: 0.07683499157428741  
step: 330, loss: 0.09624642878770828  
step: 340, loss: 0.0021475821267813444  
step: 350, loss: 0.005516918376088142  
step: 360, loss: 0.010527366772294044  
step: 370, loss: 0.003802514635026455  
step: 380, loss: 0.0520879365503788  
step: 390, loss: 0.0044400948099792  
step: 400, loss: 0.1321883499622345  
step: 410, loss: 0.0018694312311708927  
step: 420, loss: 0.01182466372847557
```

```

step: 430, loss: 0.003454296151176095
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 8192.0
step: 440, loss: 0.10132309794425964
step: 450, loss: 0.00428317254409194
step: 460, loss: 0.0029346062801778316
step: 470, loss: 0.16416846215724945
step: 480, loss: 0.007800453808158636
step: 490, loss: 0.01707109436392784
step: 500, loss: 0.012385190464556217
step: 510, loss: 0.0012228611158207059
step: 520, loss: 0.003531770082190633
step: 530, loss: 0.006098998244851828
/usr/local/lib/python3.7/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=0.9507434944237918, f1=0.9406819243344232,
best_f1=0.9447565543071161
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0019206051947548985
step: 10, loss: 0.010865258984267712
step: 20, loss: 0.005915977992117405
step: 30, loss: 0.008720132522284985
step: 40, loss: 0.0003611368010751903
step: 50, loss: 0.00038205640157684684
step: 60, loss: 0.022662583738565445
step: 70, loss: 0.004961596801877022
step: 80, loss: 0.003559354692697525
step: 90, loss: 0.004144890233874321
step: 100, loss: 0.0036347778514027596
step: 110, loss: 0.0016418533632531762
step: 120, loss: 0.006343544460833073
step: 130, loss: 0.012154273688793182
step: 140, loss: 0.03400922194123268
step: 150, loss: 0.002922773826867342
step: 160, loss: 0.0014243649784475565
step: 170, loss: 0.01806371845304966
step: 180, loss: 0.0018728792201727629
step: 190, loss: 0.0012761109974235296
step: 200, loss: 0.007148296106606722
step: 210, loss: 0.12962785363197327
step: 220, loss: 0.0077002509497106075
step: 230, loss: 0.04222767427563667
step: 240, loss: 0.043305184692144394
step: 250, loss: 0.0015025126049295068
step: 260, loss: 0.0007687482866458595
step: 270, loss: 0.0037894505076110363

```



```

step: 280, loss: 0.0010509940329939127
step: 290, loss: 0.0034970089327543974
step: 300, loss: 0.003382547292858362
step: 310, loss: 0.0014286328805610538
step: 320, loss: 0.0006224092794582248
step: 330, loss: 0.0007020526099950075
step: 340, loss: 0.039972901344299316
step: 350, loss: 0.00014727974485140294
step: 360, loss: 0.00730685843154788
step: 370, loss: 0.00235500605776906
step: 380, loss: 0.0006780849071219563
step: 390, loss: 0.003659845795482397
step: 400, loss: 0.004322448279708624
step: 410, loss: 0.0034815631806850433
step: 420, loss: 0.001974957063794136
step: 430, loss: 0.002990745473653078
step: 440, loss: 0.017144203186035156
step: 450, loss: 0.0023517298977822065
step: 460, loss: 0.019183047115802765
step: 470, loss: 0.13373161852359772
step: 480, loss: 0.007870492525398731
step: 490, loss: 0.03192789480090141
step: 500, loss: 0.0013474165461957455
step: 510, loss: 0.005898824892938137
step: 520, loss: 0.02756776288151741
step: 530, loss: 0.021487072110176086
/usr/local/lib/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=0.9552238805970149, f1=0.9477611940298507,
best_f1=0.9477611940298507
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00071014987770468
step: 10, loss: 0.002311386400833726
step: 20, loss: 0.0004911215510219336
step: 30, loss: 0.07836867868900299
step: 40, loss: 0.012581465765833855
step: 50, loss: 0.0038473522290587425
step: 60, loss: 0.00506492517888546
step: 70, loss: 0.01368271466344595
step: 80, loss: 0.0057622650638222694
step: 90, loss: 0.02972980961203575

```

```
step: 100, loss: 0.001308218576014042
step: 110, loss: 0.0030736951157450676
step: 120, loss: 0.0013617640361189842
step: 130, loss: 0.000869410578161478
step: 140, loss: 0.0006821901770308614
step: 150, loss: 0.04195883125066757
step: 160, loss: 0.0004156018840149045
step: 170, loss: 0.002980210119858384
step: 180, loss: 0.000532334262970835
step: 190, loss: 0.000522100308444351
step: 200, loss: 0.0005691102705895901
step: 210, loss: 0.20318008959293365
step: 220, loss: 0.0024973335675895214
step: 230, loss: 0.0017257985891774297
step: 240, loss: 0.001588917337357998
step: 250, loss: 0.01069992408156395
step: 260, loss: 0.0005297359311953187
step: 270, loss: 0.0009490235825069249
step: 280, loss: 0.00048532593064010143
step: 290, loss: 0.004723472520709038
step: 300, loss: 0.0036212143022567034
step: 310, loss: 0.0008829139405861497
step: 320, loss: 0.0005699383327737451
step: 330, loss: 0.000447666592663154
step: 340, loss: 0.0005850886809639633
step: 350, loss: 0.06718646734952927
step: 360, loss: 0.0016432985430583358
step: 370, loss: 0.16487659513950348
step: 380, loss: 0.0034343767911195755
step: 390, loss: 0.004847930744290352
step: 400, loss: 0.060234468430280685
step: 410, loss: 0.022829430177807808
step: 420, loss: 0.0019362177699804306
step: 430, loss: 0.0011209234362468123
step: 440, loss: 0.00020693267288152128
step: 450, loss: 0.025436710566282272
step: 460, loss: 0.00032058756914921105
step: 470, loss: 0.0007743436144664884
step: 480, loss: 0.0030932778026908636
step: 490, loss: 0.0030832639895379543
step: 500, loss: 0.016558660194277763
step: 510, loss: 0.03168904408812523
step: 520, loss: 0.0030171589460223913
step: 530, loss: 0.043983619660139084
/usr/local/lib/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=0.9553072625698324, f1=0.9401309635173059,

best\_f1=0.9401309635173059

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

An input tensor was not cuda.

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

step: 0, loss: 0.009976224973797798

step: 10, loss: 0.0011306960368528962

step: 20, loss: 0.00015873383381403983

step: 30, loss: 0.030487509444355965

step: 40, loss: 0.0005366219556890428

step: 50, loss: 0.0006892002420499921

step: 60, loss: 0.011690470390021801

step: 70, loss: 0.00016034946020226926

step: 80, loss: 0.0002340311330044642

step: 90, loss: 0.0019962750375270844

step: 100, loss: 0.0010912524303421378

step: 110, loss: 0.007926961407065392

step: 120, loss: 0.00013700684939976782

step: 130, loss: 0.010461152531206608

step: 140, loss: 0.009703652933239937

step: 150, loss: 0.0037412059027701616

step: 160, loss: 0.026477716863155365

step: 170, loss: 0.017896300181746483

step: 180, loss: 0.01488988846540451

step: 190, loss: 0.0001825814979383722

step: 200, loss: 0.0013085838872939348

step: 210, loss: 0.006266804412007332

step: 220, loss: 0.00046296193613670766

step: 230, loss: 0.00643958942964673

step: 240, loss: 0.0029146091546863317

step: 250, loss: 0.004525250289589167

step: 260, loss: 0.006069452501833439

step: 270, loss: 0.013669655658304691

step: 280, loss: 0.010769391432404518

step: 290, loss: 0.0007578355143778026

step: 300, loss: 0.01785266026854515

step: 310, loss: 0.011683288961648941

step: 320, loss: 0.003919489216059446

step: 330, loss: 0.006708613596856594

step: 340, loss: 0.00012721731036435813

step: 350, loss: 0.09346777945756912

step: 360, loss: 0.00021781674877274781

step: 370, loss: 0.0011054769856855273

step: 380, loss: 0.0017648829380050302

step: 390, loss: 0.0004778634465765208

```

step: 400, loss: 0.020897461101412773
step: 410, loss: 0.00025225180434063077
step: 420, loss: 0.00010982620733557269
step: 430, loss: 5.926869926042855e-05
step: 440, loss: 0.00014386739348992705
step: 450, loss: 0.0027737990021705627
step: 460, loss: 0.0032577079255133867
step: 470, loss: 0.00047128243022598326
step: 480, loss: 0.000731186824850738
step: 490, loss: 0.0013191611506044865
step: 500, loss: 0.02535558119416237
step: 510, loss: 0.07476960867643356
step: 520, loss: 0.0777757316827774
step: 530, loss: 0.011967122554779053
/usr/local/lib/python3.7/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=0.9521152952115295, f1=0.9395348837209302,
best_f1=0.9401309635173059
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0001000558040686883
step: 10, loss: 0.0019097620388492942
step: 20, loss: 0.0003651557199191302
step: 30, loss: 0.0027731447480618954
step: 40, loss: 0.000688264612108469
step: 50, loss: 0.0024108279030770063
step: 60, loss: 0.001566292834468186
step: 70, loss: 0.001073856488801539
step: 80, loss: 0.008687414228916168
step: 90, loss: 0.010892585851252079
step: 100, loss: 0.008663121610879898
step: 110, loss: 0.0003427068586461246
step: 120, loss: 0.006297205574810505
step: 130, loss: 0.0009523631306365132
step: 140, loss: 0.00047360730241052806
step: 150, loss: 0.010842467658221722
step: 160, loss: 0.0004847443487960845
step: 170, loss: 0.000852097466122359
step: 180, loss: 0.0002154172252630815
step: 190, loss: 0.005739058367908001
step: 200, loss: 0.0004647981550078839
step: 210, loss: 0.0005375635228119791
step: 220, loss: 0.03737970069050789
step: 230, loss: 0.0016198876546695828
step: 240, loss: 0.0011101925047114491
step: 250, loss: 0.006006523035466671

```

```

step: 260, loss: 0.0010844907956197858
step: 270, loss: 0.003537493757903576
step: 280, loss: 0.0037276397924870253
step: 290, loss: 0.012525452300906181
step: 300, loss: 7.932820153655484e-05
step: 310, loss: 0.08078023046255112
step: 320, loss: 0.0005072678904980421
step: 330, loss: 8.290849655168131e-05
step: 340, loss: 0.002103975508362055
step: 350, loss: 0.0008747795945964754
step: 360, loss: 0.0019410697277635336
step: 370, loss: 0.001390046556480229
step: 380, loss: 0.008472390472888947
step: 390, loss: 0.0019288567127659917
step: 400, loss: 0.0019521935610100627
step: 410, loss: 0.0031845879275351763
step: 420, loss: 0.0020231720991432667
step: 430, loss: 0.002125658793374896
step: 440, loss: 0.00016742463049013168
step: 450, loss: 0.00189426657743752
step: 460, loss: 0.0012698685750365257
step: 470, loss: 0.0017041245009750128
step: 480, loss: 0.0013725502649322152
step: 490, loss: 0.0014767699176445603
step: 500, loss: 0.0010484873782843351
step: 510, loss: 0.0011398263741284609
step: 520, loss: 0.00614657299593091
step: 530, loss: 0.007391444873064756
/usr/local/lib/python3.7/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=0.9513212795549374, f1=0.9428704133766836,
best_f1=0.9401309635173059
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0014422694221138954
step: 10, loss: 0.003238589968532324
step: 20, loss: 0.0615772120654583
step: 30, loss: 0.0016494692536070943
step: 40, loss: 0.0027673018630594015
step: 50, loss: 0.001230378053151071
step: 60, loss: 0.017192212864756584
step: 70, loss: 0.02582237310707569
step: 80, loss: 0.2843347191810608
step: 90, loss: 0.007435736246407032
step: 100, loss: 0.007282155100256205
step: 110, loss: 0.0020361016504466534

```

```

step: 120, loss: 0.0001492744340794161
step: 130, loss: 0.0007107518031261861
step: 140, loss: 0.00016580361989326775
step: 150, loss: 0.0009632413275539875
step: 160, loss: 0.005768784787505865
step: 170, loss: 0.0005120148998685181
step: 180, loss: 0.002480539260432124
step: 190, loss: 0.00032284590997733176
step: 200, loss: 0.0008851474267430604
step: 210, loss: 0.0016027861274778843
step: 220, loss: 0.0004544316907413304
step: 230, loss: 0.002415132476016879
step: 240, loss: 0.0036966681946069
step: 250, loss: 0.22486895322799683
step: 260, loss: 0.006617333739995956
step: 270, loss: 0.005790180992335081
step: 280, loss: 0.001208052970468998
step: 290, loss: 0.002818757202476263
step: 300, loss: 0.001864681369625032
step: 310, loss: 0.006020802538841963
step: 320, loss: 0.007552598137408495
step: 330, loss: 0.000976661453023553
step: 340, loss: 0.0009066408383660018
step: 350, loss: 0.00015589222311973572
step: 360, loss: 0.0002512051141820848
step: 370, loss: 0.0015695232432335615
step: 380, loss: 0.0004541230446193367
step: 390, loss: 0.0392950177192688
step: 400, loss: 0.0023522183764725924
step: 410, loss: 0.0017387184780091047
step: 420, loss: 0.00037689178134314716
step: 430, loss: 0.0004771612584590912
step: 440, loss: 0.0033039022237062454
step: 450, loss: 0.09919843077659607
step: 460, loss: 0.0004816099535673857
step: 470, loss: 0.0014539632247760892
step: 480, loss: 0.0004830543475691229
step: 490, loss: 0.004059716127812862
step: 500, loss: 0.0009145071962848306
step: 510, loss: 0.008727718144655228
step: 520, loss: 0.0034723978023976088
step: 530, loss: 0.002985855797305703
/usr/local/lib/python3.7/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=0.9519813519813519, f1=0.9386991109031353,
best_f1=0.9401309635173059
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:

```

An input tensor was not cuda.

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

```
step: 0, loss: 0.0003120788314845413
step: 10, loss: 0.0004238539841026068
step: 20, loss: 0.0006967085646465421
step: 30, loss: 0.00011697584704961628
step: 40, loss: 0.0002654220152180642
step: 50, loss: 0.0008990766364149749
step: 60, loss: 0.00014387082774192095
step: 70, loss: 0.0027672722935676575
step: 80, loss: 0.010472938418388367
step: 90, loss: 0.0003670849837362766
step: 100, loss: 0.000499542395118624
step: 110, loss: 0.00040804274613037705
step: 120, loss: 0.00011626350897131488
step: 130, loss: 0.0006402336293831468
step: 140, loss: 0.00016748129564803094
step: 150, loss: 0.0013794589322060347
step: 160, loss: 0.0005199673469178379
step: 170, loss: 0.203932985663414
step: 180, loss: 0.0004810614336747676
step: 190, loss: 0.0003682238457258791
step: 200, loss: 0.00023435910406988114
step: 210, loss: 0.0001890283019747585
step: 220, loss: 0.0007740436121821404
step: 230, loss: 0.0003056245041079819
step: 240, loss: 0.01138278841972351
step: 250, loss: 0.00014478107914328575
step: 260, loss: 0.00021528829529415816
step: 270, loss: 0.028225358575582504
step: 280, loss: 0.0015932932728901505
step: 290, loss: 8.797529153525829e-05
step: 300, loss: 0.0003118881140835583
step: 310, loss: 0.000873104901984334
step: 320, loss: 0.0001690835488261655
step: 330, loss: 0.00013560423394665122
step: 340, loss: 0.00016694673104211688
step: 350, loss: 0.0007292545633390546
step: 360, loss: 0.09224891662597656
step: 370, loss: 0.0007569835870526731
step: 380, loss: 0.0006269195000641048
step: 390, loss: 0.0005613019457086921
step: 400, loss: 0.0009742798283696175
step: 410, loss: 0.00014258523879107088
step: 420, loss: 0.00033588253427296877
step: 430, loss: 0.027074959129095078
step: 440, loss: 6.445019243983552e-05
step: 450, loss: 0.0007411550614051521
```

```

step: 460, loss: 0.006918574217706919
step: 470, loss: 0.000506464799400419
step: 480, loss: 0.0019125047838315368
step: 490, loss: 0.0022309215273708105
step: 500, loss: 0.0001435356680303812
step: 510, loss: 0.0022635094355791807
step: 520, loss: 0.00038543675327673554
step: 530, loss: 0.0005996098043397069
/usr/local/lib/python3.7/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=0.9530516431924883, f1=0.9362907031618688,
best_f1=0.9401309635173059
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0003165921662002802
step: 10, loss: 0.00102839688770473
step: 20, loss: 0.001819205703213811
step: 30, loss: 0.001081134076230228
step: 40, loss: 0.0010621313704177737
step: 50, loss: 0.017583578824996948
step: 60, loss: 0.0008672701078467071
step: 70, loss: 0.0006404023733921349
step: 80, loss: 0.00016632424376439303
step: 90, loss: 0.0002555109967943281
step: 100, loss: 0.00043633513269014657
step: 110, loss: 0.0003406973264645785
step: 120, loss: 4.552491009235382e-05
step: 130, loss: 6.163900252431631e-05
step: 140, loss: 0.0012639994965866208
step: 150, loss: 0.0007797794532962143
step: 160, loss: 0.0004686517349909991
step: 170, loss: 0.02062404155731201
step: 180, loss: 0.0002286303206346929
step: 190, loss: 0.001738423714414239
step: 200, loss: 7.890831329859793e-05
step: 210, loss: 0.00018107589858118445
step: 220, loss: 7.630043546669185e-05
step: 230, loss: 0.0001952486636582762
step: 240, loss: 0.0007169988821260631
step: 250, loss: 0.042740095406770706
step: 260, loss: 0.0008523592259734869
step: 270, loss: 0.0006419589044526219
step: 280, loss: 0.00036193677806295455
step: 290, loss: 0.0002608936920296401
step: 300, loss: 9.585882071405649e-05
step: 310, loss: 0.006524840835481882

```



```

step: 320, loss: 5.8439120039111e-05
step: 330, loss: 0.0005194316618144512
step: 340, loss: 0.00046790947089903057
step: 350, loss: 9.413600491825491e-05
step: 360, loss: 0.0007144242408685386
step: 370, loss: 0.0009965515928342938
step: 380, loss: 0.00078967324225232
step: 390, loss: 0.015832198783755302
step: 400, loss: 0.0005470269825309515
step: 410, loss: 9.367903112433851e-05
step: 420, loss: 0.0004584876587614417
step: 430, loss: 0.0009494818514212966
step: 440, loss: 0.0005222384352236986
step: 450, loss: 0.00010116813791682944
step: 460, loss: 0.02519332244992256
step: 470, loss: 0.0011648592771962285
step: 480, loss: 0.0004084585525561124
step: 490, loss: 0.0008827819838188589
step: 500, loss: 0.0013422733172774315
step: 510, loss: 0.04109785705804825
step: 520, loss: 0.0004097700584679842
step: 530, loss: 0.135600745677948
/usr/local/lib/python3.7/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=0.9508804448563485, f1=0.9422632794457274,
best_f1=0.9401309635173059
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 8.40523571241647e-05
step: 10, loss: 0.0011047481093555689
step: 20, loss: 0.0003067019861191511
step: 30, loss: 0.007369996514171362
step: 40, loss: 0.0002269558171974495
step: 50, loss: 0.0005353505839593709
step: 60, loss: 0.0007733466336503625
step: 70, loss: 0.00045332417357712984
step: 80, loss: 0.00014731069677509367
step: 90, loss: 0.0007991233142092824
step: 100, loss: 0.0007062916993163526
step: 110, loss: 0.0004284815222490579
step: 120, loss: 6.37167613604106e-05
step: 130, loss: 6.243526877369732e-05
step: 140, loss: 3.133173959213309e-05
step: 150, loss: 7.890728738857433e-05
step: 160, loss: 0.00010227632446913049
step: 170, loss: 0.00010123423271579668

```

```

step: 180, loss: 0.008243363350629807
step: 190, loss: 0.0008681448525749147
step: 200, loss: 0.0005257417797110975
step: 210, loss: 0.000247410120209679
step: 220, loss: 2.332654366909992e-05
step: 230, loss: 0.07368088513612747
step: 240, loss: 3.570380067685619e-05
step: 250, loss: 2.200853850808926e-05
step: 260, loss: 2.693166970857419e-05
step: 270, loss: 1.5191521924862172e-05
step: 280, loss: 2.5353041564812884e-05
step: 290, loss: 0.003606627229601145
step: 300, loss: 0.00026545338914729655
step: 310, loss: 0.04885741323232651
step: 320, loss: 8.95270422915928e-05
step: 330, loss: 2.7248264814261347e-05
step: 340, loss: 8.852910104906186e-05
step: 350, loss: 0.00028753711376339197
step: 360, loss: 0.0006503452896140516
step: 370, loss: 0.010990628972649574
step: 380, loss: 7.557598291896284e-05
step: 390, loss: 0.0009535133140161633
step: 400, loss: 0.0004516319604590535
step: 410, loss: 4.111522139282897e-05
step: 420, loss: 0.0003444109170231968
step: 430, loss: 2.715970913413912e-05
step: 440, loss: 0.01502978801727295
step: 450, loss: 0.01741507090628147
step: 460, loss: 0.0057188053615391254
step: 470, loss: 3.0436558517976664e-05
step: 480, loss: 0.0015186493983492255
step: 490, loss: 0.005112168844789267
step: 500, loss: 0.015098005533218384
step: 510, loss: 0.0001098588909371756
step: 520, loss: 2.4473678422509693e-05
step: 530, loss: 0.00012455855903681368
/usr/local/lib/python3.7/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=0.9510945505356311, f1=0.9414498141263941,
best_f1=0.9401309635173059
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, 170.91it/s]
load_f1 = 0.9528257823446986
real_f1 = 0.9522024367385192
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, 143.90it/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: 0.44288310408592224
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=0.2916666666666667, f1=0.2692307692307693,
best_f1=0.2692307692307693
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.44488978385925293
/usr/local/lib/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=0.3888888888888889, f1=0.34285714285714286,
best_f1=0.34285714285714286
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.40789881348609924
/usr/local/lib/python3.7/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=0.3373493975903615, f1=0.27184466019417475,
best_f1=0.34285714285714286
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.2661260962486267
/usr/local/lib/python3.7/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=0.32558139534883723, f1=0.27184466019417475,
best_f1=0.34285714285714286
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.3470004200935364
/usr/local/lib/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=0.4444444444444445, f1=0.3181818181818182,
best_f1=0.3181818181818182
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.17063285410404205
/usr/local/lib/python3.7/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=0.42622950819672134, f1=0.2978723404255319,
best_f1=0.3181818181818182
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.4828394651412964
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 7: dev_f1=0.52, f1=0.37333333333333335, best_f1=0.37333333333333335
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.32334354519844055
/usr/local/lib/python3.7/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=0.4827586206896552, f1=0.358974358974359,
best_f1=0.37333333333333335
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:

```

```

An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.3611554503440857
/usr/local/lib/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=0.75, f1=0.48888888888888893, best_f1=0.48888888888888893
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.19394394755363464
/usr/local/lib/python3.7/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=0.6341463414634146, f1=0.4210526315789474,
best_f1=0.48888888888888893
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.20908652245998383
/usr/local/lib/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=0.7857142857142857, f1=0.5116279069767441,
best_f1=0.5116279069767441
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.11156976222991943
/usr/local/lib/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=0.8275862068965518, f1=0.5, best_f1=0.5
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")

```

```

step: 0, loss: 0.07442896068096161
/usr/local/lib/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=0.8571428571428571, f1=0.5116279069767441,
best_f1=0.5116279069767441
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.02772248536348343
/usr/local/lib/python3.7/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=0.8571428571428571, f1=0.5116279069767441,
best_f1=0.5116279069767441
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.10652393847703934
/usr/local/lib/python3.7/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=0.8571428571428571, f1=0.5116279069767441,
best_f1=0.5116279069767441
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, 111831.72it/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 = 0.8571428571428571
real_f1 = 0.8148148148148148
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 [00:26, 164.78it/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/
```

```
Downloading: 100% 481/481 [00:00<00:00, 402kB/s]
Downloading: 100% 899k/899k [00:01<00:00, 808kB/s]
Downloading: 100% 456k/456k [00:00<00:00, 465kB/s]
Downloading: 100% 501M/501M [00:08<00:00, 60.7MB/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: 0.5874584913253784
```

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

```
step: 10, loss: 0.45140376687049866
```

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

```
step: 20, loss: 0.5111132860183716
```

```
step: 30, loss: 0.2837994694709778
```

```
step: 40, loss: 0.41759932041168213
```

```
step: 50, loss: 0.543333888053894
```

```
step: 60, loss: 0.2946127653121948
```

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

```
step: 70, loss: 0.2058243602514267
```

```
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 4096.0
```

```
step: 80, loss: 0.587806761264801
```

```
step: 90, loss: 0.28936052322387695
```

```
step: 100, loss: 0.17613622546195984
```

```

step: 110, loss: 0.11514599621295929
step: 120, loss: 0.015885699540376663
step: 130, loss: 0.005764250177890062
step: 140, loss: 0.015761306509375572
step: 150, loss: 0.06328599900007248
step: 160, loss: 0.023276785388588905
step: 170, loss: 0.07634744793176651
step: 180, loss: 0.11228172481060028
step: 190, loss: 0.019163645803928375
step: 200, loss: 0.050970807671546936
step: 210, loss: 0.012772344052791595
step: 220, loss: 0.16693054139614105
step: 230, loss: 0.016894962638616562
/usr/local/lib/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=0.9562363238512035, f1=0.9646017699115044,
best_f1=0.9646017699115044
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.006803951226174831
step: 10, loss: 0.074605792760849
step: 20, loss: 0.022260962054133415
step: 30, loss: 0.014167780056595802
step: 40, loss: 0.006572495214641094
step: 50, loss: 0.0012891343794763088
step: 60, loss: 0.005202824715524912
step: 70, loss: 0.015608426183462143
step: 80, loss: 0.0038929732982069254
step: 90, loss: 0.0013165355194360018
step: 100, loss: 0.002704617567360401
step: 110, loss: 0.061507850885391235
step: 120, loss: 0.00114253640640527
step: 130, loss: 0.004052772186696529
step: 140, loss: 0.0022636314388364553
step: 150, loss: 0.13831950724124908
step: 160, loss: 0.004431865643709898
step: 170, loss: 0.008387415669858456
step: 180, loss: 0.0015908795176073909
step: 190, loss: 0.03249112516641617
step: 200, loss: 0.0010587162105366588
step: 210, loss: 0.05850614607334137
step: 220, loss: 0.002049278002232313

```

```

step: 230, loss: 0.002775452798232436
/usr/local/lib/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=0.9788182831661093, f1=0.9775784753363228,
best_f1=0.9775784753363228
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.059696175158023834
step: 10, loss: 0.0023161827120929956
step: 20, loss: 0.003911598585546017
step: 30, loss: 0.0008613210520707071
step: 40, loss: 0.0013093145098537207
step: 50, loss: 0.051522932946681976
step: 60, loss: 0.05756384879350662
step: 70, loss: 0.0013612705515697598
step: 80, loss: 0.013513020239770412
step: 90, loss: 0.0009028029162436724
step: 100, loss: 0.0009078781004063785
step: 110, loss: 0.007591970730572939
step: 120, loss: 0.00043979694601148367
step: 130, loss: 0.020047351717948914
step: 140, loss: 0.01637950725853443
step: 150, loss: 0.003738304367288947
step: 160, loss: 0.01176741998642683
step: 170, loss: 0.018737779930233955
step: 180, loss: 0.09030510485172272
step: 190, loss: 0.034087955951690674
step: 200, loss: 0.026625845581293106
step: 210, loss: 0.003592089982703328
step: 220, loss: 0.0061670634895563126
step: 230, loss: 0.019799504429101944
/usr/local/lib/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=0.9820627802690582, f1=0.9786276715410572,
best_f1=0.9786276715410572
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.

```

```

    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00917510874569416
step: 10, loss: 0.006820904091000557
step: 20, loss: 0.0052161565981805325
step: 30, loss: 0.0007594324997626245
step: 40, loss: 0.12963329255580902
step: 50, loss: 0.041292015463113785
step: 60, loss: 0.0018109636148437858
step: 70, loss: 0.003133026883006096
step: 80, loss: 0.0014099081745371222
step: 90, loss: 0.01124762836843729
step: 100, loss: 0.0014029167359694839
step: 110, loss: 0.0005592048401013017
step: 120, loss: 0.00937657617032528
step: 130, loss: 0.019461773335933685
step: 140, loss: 0.0006266534910537302
step: 150, loss: 0.0007796703139320016
step: 160, loss: 0.0011303500505164266
step: 170, loss: 0.11606352776288986
step: 180, loss: 0.12078585475683212
step: 190, loss: 0.003643014468252659
step: 200, loss: 0.023317262530326843
step: 210, loss: 0.0011409125290811062
step: 220, loss: 0.0003858719137497246
step: 230, loss: 0.0033663902431726456
/usr/local/lib/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=0.9876543209876544, 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: 0.0009044855833053589
step: 10, loss: 0.0058761779218912125
step: 20, loss: 0.0010562833631411195
step: 30, loss: 0.0004658415273297578
step: 40, loss: 0.001630284357815981
step: 50, loss: 0.0008075972436927259
step: 60, loss: 0.001390896737575531
step: 70, loss: 0.002583626192063093
step: 80, loss: 0.07010243088006973
step: 90, loss: 0.04627065733075142
step: 100, loss: 0.0006005041650496423

```

```

step: 110, loss: 0.0177837572991848
step: 120, loss: 0.002018429571762681
step: 130, loss: 0.0006459909491240978
step: 140, loss: 0.0011848622234538198
step: 150, loss: 0.10455508530139923
step: 160, loss: 0.0010035211453214288
step: 170, loss: 0.0014041168615221977
step: 180, loss: 0.0025947033427655697
step: 190, loss: 0.00668276147916913
step: 200, loss: 0.001683026785030961
step: 210, loss: 0.017323032021522522
step: 220, loss: 0.0006582676433026791
step: 230, loss: 0.004202922340482473
/usr/local/lib/python3.7/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=0.9865470852017937, f1=0.984304932735426,
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: 0.015424530953168869
step: 10, loss: 0.0010495560709387064
step: 20, loss: 0.0026341050397604704
step: 30, loss: 0.0013133998727425933
step: 40, loss: 0.0005796718178316951
step: 50, loss: 0.0006237680790945888
step: 60, loss: 0.0012657905463129282
step: 70, loss: 0.00034515498555265367
step: 80, loss: 0.0012562331976369023
step: 90, loss: 0.003974954131990671
step: 100, loss: 0.0010540225775912404
step: 110, loss: 0.007259138859808445
step: 120, loss: 0.0004458418698050082
step: 130, loss: 0.0006364120054058731
step: 140, loss: 0.0011242295149713755
step: 150, loss: 0.00030548000358976424
step: 160, loss: 0.04272085428237915
step: 170, loss: 0.0005069730104878545
step: 180, loss: 0.0002049948088824749
step: 190, loss: 0.00036896273377351463
step: 200, loss: 0.002012969460338354
step: 210, loss: 0.06289268285036087
step: 220, loss: 0.015184693969786167
step: 230, loss: 0.0025791707448661327
/usr/local/lib/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=0.988814317673378, f1=0.9799554565701558,
best_f1=0.9799554565701558
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0019305545138195157
step: 10, loss: 0.0007999589433893561
step: 20, loss: 0.0011045782594010234
step: 30, loss: 0.0005673918058164418
step: 40, loss: 0.0018239174969494343
step: 50, loss: 0.0009913714602589607
step: 60, loss: 0.0004955098265781999
step: 70, loss: 0.0005470045143738389
step: 80, loss: 0.0006496031419374049
step: 90, loss: 0.06560084223747253
step: 100, loss: 0.0005230549140833318
step: 110, loss: 0.0013855358120054007
step: 120, loss: 0.0008934124489314854
step: 130, loss: 0.0025507549289613962
step: 140, loss: 0.0019074113806709647
step: 150, loss: 0.007187935058027506
step: 160, loss: 0.004452361259609461
step: 170, loss: 0.0005371003062464297
step: 180, loss: 0.00033585919300094247
step: 190, loss: 0.0009358323877677321
step: 200, loss: 0.0006380395498126745
step: 210, loss: 0.014889808371663094
step: 220, loss: 0.001417951425537467
step: 230, loss: 0.0663522332906723
/usr/local/lib/python3.7/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=0.9876265466816648, f1=0.9810055865921787,
best_f1=0.9799554565701558
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0024885740131139755
step: 10, loss: 0.0029163481667637825
step: 20, loss: 0.002770679770037532
step: 30, loss: 0.001321515184827149
step: 40, loss: 0.0013461180496960878
step: 50, loss: 0.0007934514433145523
step: 60, loss: 0.0009270465816371143

```

```

step: 70, loss: 0.00033038645051419735
step: 80, loss: 0.00047175641520880163
step: 90, loss: 0.0005870468448847532
step: 100, loss: 0.0013269629562273622
step: 110, loss: 0.0016604817938059568
step: 120, loss: 0.0016919111367315054
step: 130, loss: 0.0015916073461994529
step: 140, loss: 0.0008135454263538122
step: 150, loss: 0.014217502437531948
step: 160, loss: 0.00071366922929883
step: 170, loss: 0.010036608204245567
step: 180, loss: 0.0004700912395492196
step: 190, loss: 0.00048751739086583257
step: 200, loss: 0.016871437430381775
step: 210, loss: 0.0049136122688651085
step: 220, loss: 0.0009975965367630124
step: 230, loss: 0.0006426554173231125
/usr/local/lib/python3.7/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=0.9854748603351955, f1=0.9821029082774049,
best_f1=0.9799554565701558
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0004915558383800089
step: 10, loss: 0.0009601548663340509
step: 20, loss: 0.010857946239411831
step: 30, loss: 0.004539794288575649
step: 40, loss: 0.0007869211258366704
step: 50, loss: 0.0013106483966112137
step: 60, loss: 0.00038054437027312815
step: 70, loss: 0.05782892554998398
step: 80, loss: 0.0069616809487342834
step: 90, loss: 0.01567225530743599
step: 100, loss: 0.00031128356931731105
step: 110, loss: 0.0002849942829925567
step: 120, loss: 0.010820221155881882
step: 130, loss: 0.016112273558974266
step: 140, loss: 0.0009254488977603614
step: 150, loss: 0.0005666646757163107
step: 160, loss: 0.0041363914497196674
step: 170, loss: 0.0002446429571136832
step: 180, loss: 0.0004091874579899013
step: 190, loss: 0.00022567635460291058
step: 200, loss: 0.0007837151642888784
step: 210, loss: 0.0016297245165333152
step: 220, loss: 0.0008079588878899813

```



```

step: 230, loss: 0.0006169406697154045
/usr/local/lib/python3.7/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=0.9876819708846584, f1=0.984304932735426,
best_f1=0.9799554565701558
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0009333347552455962
step: 10, loss: 0.0022656794171780348
step: 20, loss: 0.0007425628136843443
step: 30, loss: 0.00014388997806236148
step: 40, loss: 0.0025805195327848196
step: 50, loss: 0.00039016350638121367
step: 60, loss: 0.003582475008442998
step: 70, loss: 0.0005925934528931975
step: 80, loss: 0.0002334695600438863
step: 90, loss: 0.0001642491261009127
step: 100, loss: 0.0001377372827846557
step: 110, loss: 0.012043178081512451
step: 120, loss: 0.00010711365757742897
step: 130, loss: 0.0006215700414031744
step: 140, loss: 0.0007688191253691912
step: 150, loss: 0.0012189527042210102
step: 160, loss: 0.00010614383791107684
step: 170, loss: 0.00040009545045904815
step: 180, loss: 0.0007886952953413129
step: 190, loss: 0.00042850463069044054
step: 200, loss: 0.000860086118336767
step: 210, loss: 0.0028550836723297834
step: 220, loss: 0.016854045912623405
step: 230, loss: 0.00602483656257391
/usr/local/lib/python3.7/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=0.9887892376681614, f1=0.9798657718120806,
best_f1=0.9799554565701558
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00028836651472374797
step: 10, loss: 0.000491812766995281
step: 20, loss: 0.0004026828391943127
step: 30, loss: 0.0002522034919820726
step: 40, loss: 0.00013435645087156445
step: 50, loss: 0.0011775277089327574
step: 60, loss: 0.0030789675656706095

```

```

step: 70, loss: 0.00022508615802507848
step: 80, loss: 0.00042466906597837806
step: 90, loss: 0.14484375715255737
step: 100, loss: 0.0098029263317585
step: 110, loss: 0.0029284716583788395
step: 120, loss: 0.0005468653980642557
step: 130, loss: 0.0008285484509542584
step: 140, loss: 0.0034447633661329746
step: 150, loss: 0.0006832918734289706
step: 160, loss: 0.003081829287111759
step: 170, loss: 0.0009186399984173477
step: 180, loss: 0.0008329841075465083
step: 190, loss: 0.0007878874312154949
step: 200, loss: 0.005823030136525631
step: 210, loss: 0.00099461548961699
step: 220, loss: 0.004990404471755028
step: 230, loss: 0.000813242862932384
/usr/local/lib/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=0.9899216125419933, f1=0.9810479375696767,
best_f1=0.9810479375696767
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0008520496194250882
step: 10, loss: 0.00029252850799821317
step: 20, loss: 0.008399982936680317
step: 30, loss: 0.03408580645918846
step: 40, loss: 0.000898837810382247
step: 50, loss: 0.0010357009014114738
step: 60, loss: 0.0010593285551294684
step: 70, loss: 0.0006649998831562698
step: 80, loss: 0.0004962744424119592
step: 90, loss: 0.001034822198562324
step: 100, loss: 0.00039151209057308733
step: 110, loss: 0.0003955198626499623
step: 120, loss: 0.0006835878593847156
step: 130, loss: 0.00035381974885240197
step: 140, loss: 0.0005538556142710149
step: 150, loss: 0.0005788073176518083
step: 160, loss: 0.0004601191612891853
step: 170, loss: 0.0005728827090933919
step: 180, loss: 0.00042187259532511234

```

```

step: 190, loss: 0.001974481390789151
step: 200, loss: 0.00043280329555273056
step: 210, loss: 0.0008214066037908196
step: 220, loss: 0.002282349858433008
step: 230, loss: 0.0011578588746488094
/usr/local/lib/python3.7/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=0.9888641425389755, f1=0.9777777777777777,
best_f1=0.9810479375696767
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.001415032660588622
step: 10, loss: 0.0008079293766058981
step: 20, loss: 0.0012458949349820614
step: 30, loss: 0.0056993067264556885
step: 40, loss: 0.0007944183889776468
step: 50, loss: 0.0016705194720998406
step: 60, loss: 0.0005046273581683636
step: 70, loss: 0.0033868090249598026
step: 80, loss: 0.005683518014848232
step: 90, loss: 0.0003328614984638989
step: 100, loss: 0.0015750883612781763
step: 110, loss: 0.008142647333443165
step: 120, loss: 0.003253455273807049
step: 130, loss: 0.0003491574025247246
step: 140, loss: 0.00018142179760616273
step: 150, loss: 0.00010769737127702683
step: 160, loss: 0.0019773344974964857
step: 170, loss: 0.0008549299091100693
step: 180, loss: 0.0168751310557127
step: 190, loss: 0.00021640094928443432
step: 200, loss: 0.0027532526291906834
step: 210, loss: 0.00029343910864554346
step: 220, loss: 0.0004948753630742431
step: 230, loss: 0.0006424959283322096
/usr/local/lib/python3.7/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=0.987709497206704, f1=0.978865406006674,
best_f1=0.9810479375696767
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0003617156180553138
step: 10, loss: 0.00012554881686810404
step: 20, loss: 0.0005490662297233939

```

```

step: 30, loss: 0.0004348708607722074
step: 40, loss: 0.0004922565422020853
step: 50, loss: 0.00011045186693081632
step: 60, loss: 0.0007197794620878994
step: 70, loss: 0.009425771422684193
step: 80, loss: 0.0004928044509142637
step: 90, loss: 0.0010958645725622773
step: 100, loss: 0.001381073845550418
step: 110, loss: 0.00045026428415440023
step: 120, loss: 0.0003069079539272934
step: 130, loss: 0.00021582850604318082
step: 140, loss: 0.0001197152232634835
step: 150, loss: 4.738852658192627e-05
step: 160, loss: 0.00024616284645162523
step: 170, loss: 0.00012612405407708138
step: 180, loss: 0.00013694955850951374
step: 190, loss: 0.00011680438910843804
step: 200, loss: 0.00015736426576040685
step: 210, loss: 0.00025287922471761703
step: 220, loss: 0.00043942141928710043
step: 230, loss: 0.0004533810424618423
/usr/local/lib/python3.7/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=0.9888641425389755, f1=0.9778270509977827,
best_f1=0.9810479375696767
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0009813091019168496
step: 10, loss: 0.00014287879457697272
step: 20, loss: 0.0011181074660271406
step: 30, loss: 0.00012931124365422875
step: 40, loss: 6.400729762390256e-05
step: 50, loss: 0.00013288197806105018
step: 60, loss: 0.028200244531035423
step: 70, loss: 0.00011316404561512172
step: 80, loss: 0.0001143789995694533
step: 90, loss: 0.00023222086019814014
step: 100, loss: 7.790594827383757e-05
step: 110, loss: 0.0002048733294941485
step: 120, loss: 0.018364375457167625
step: 130, loss: 0.00037737531238235533
step: 140, loss: 0.01333856675773859
step: 150, loss: 0.0007425449439324439
step: 160, loss: 0.014932889491319656
step: 170, loss: 6.151487468741834e-05
step: 180, loss: 0.0001681731519056484

```

```

step: 190, loss: 0.0003143683716189116
step: 200, loss: 0.0004465229285415262
step: 210, loss: 0.14351888000965118
step: 220, loss: 0.0001278347772313282
step: 230, loss: 0.00015661769430153072
/usr/local/lib/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=0.9899665551839464, f1=0.98, best_f1=0.98
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, 147.59it/s]
load_f1 = 0.9877641824249165
real_f1 = 0.9899665551839464
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, 136.01it/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: 0.6263348460197449
```

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: 0.4973161816596985
```

```
step: 20, loss: 0.2833632230758667
```

```
step: 30, loss: 0.3593776524066925
```

```
step: 40, loss: 0.34269315004348755
```

```
step: 50, loss: 0.39669230580329895
```

```
step: 60, loss: 0.3837755024433136
```

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

```
step: 70, loss: 0.3018675148487091
```

```
step: 80, loss: 0.21441593766212463
```

```
step: 90, loss: 0.17384760081768036
```

```
step: 100, loss: 0.22428159415721893
```

```
step: 110, loss: 0.18347381055355072
```

```
step: 120, loss: 0.11018988490104675
```

```
step: 130, loss: 0.1379026621580124
```

```
step: 140, loss: 0.28511324524879456
```

```
step: 150, loss: 0.1095525473356247
```

```
step: 160, loss: 0.1203758716583252
```

```
step: 170, loss: 0.07205317169427872
```

```
step: 180, loss: 0.07004734873771667
```

```
step: 190, loss: 0.04175862297415733
```

```
step: 200, loss: 0.04107574000954628
```

```
step: 210, loss: 0.04300704970955849
```

```
step: 220, loss: 0.05804172158241272
```

```
step: 230, loss: 0.18291741609573364
```

```
step: 240, loss: 0.06447173655033112
```

```
step: 250, loss: 0.04456693306565285
```

```
step: 260, loss: 0.3131154775619507
```

```
step: 270, loss: 0.2612079679965973
```

```
step: 280, loss: 0.05350590869784355
```

```
step: 290, loss: 0.04354057461023331
```

```
step: 300, loss: 0.05866860970854759
```

```

step: 310, loss: 0.09950673580169678
step: 320, loss: 0.06844452768564224
step: 330, loss: 0.09541264921426773
step: 340, loss: 0.32008102536201477
step: 350, loss: 0.11585213989019394
step: 360, loss: 0.0651768147945404
step: 370, loss: 0.23941285908222198
step: 380, loss: 0.1378161609172821
step: 390, loss: 0.006768961902707815
step: 400, loss: 0.027717486023902893
step: 410, loss: 0.28167030215263367
step: 420, loss: 0.010725093074142933
step: 430, loss: 0.013808106072247028
step: 440, loss: 0.08096048980951309
step: 450, loss: 0.008807875216007233
step: 460, loss: 0.058876264840364456
step: 470, loss: 0.02842910774052143
step: 480, loss: 0.06486668437719345
step: 490, loss: 0.053147703409194946
step: 500, loss: 0.02434937097132206
step: 510, loss: 0.055842600762844086
step: 520, loss: 0.1365514099597931
step: 530, loss: 0.06530742347240448
/usr/local/lib/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=0.93796992481203, f1=0.9250936329588015,
best_f1=0.9250936329588015
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.03944693133234978
step: 10, loss: 0.03317610174417496
step: 20, loss: 0.00678984122350812
step: 30, loss: 0.10393403470516205
step: 40, loss: 0.0744871124625206
step: 50, loss: 0.022533614188432693
step: 60, loss: 0.020579051226377487
step: 70, loss: 0.03416626155376434
step: 80, loss: 0.027936525642871857
step: 90, loss: 0.003809636691585183
step: 100, loss: 0.0933891162276268
step: 110, loss: 0.008617025800049305
step: 120, loss: 0.198206827044487

```



```

step: 130, loss: 0.005584754515439272
step: 140, loss: 0.11160261929035187
step: 150, loss: 0.003612494096159935
step: 160, loss: 0.01600664295256138
step: 170, loss: 0.02976212464272976
step: 180, loss: 0.01697417162358761
step: 190, loss: 0.005604600999504328
step: 200, loss: 0.0733979195356369
step: 210, loss: 0.02063019759953022
step: 220, loss: 0.0016193118644878268
step: 230, loss: 0.035175807774066925
step: 240, loss: 0.08112572133541107
step: 250, loss: 0.012529250234365463
step: 260, loss: 0.010045447386801243
step: 270, loss: 0.04539019986987114
step: 280, loss: 0.07083974033594131
step: 290, loss: 0.00926421582698822
step: 300, loss: 0.03618483617901802
step: 310, loss: 0.009489256888628006
step: 320, loss: 0.008955627679824829
step: 330, loss: 0.06990694999694824
step: 340, loss: 0.0750921294093132
step: 350, loss: 0.0020099012181162834
step: 360, loss: 0.13188712298870087
step: 370, loss: 0.017567291855812073
step: 380, loss: 0.10909925401210785
step: 390, loss: 0.0037479239981621504
step: 400, loss: 0.02608453296124935
step: 410, loss: 0.00699173379689455
step: 420, loss: 0.027912596240639687
step: 430, loss: 0.16620869934558868
step: 440, loss: 0.021188508719205856
step: 450, loss: 0.16266892850399017
step: 460, loss: 0.07063189148902893
step: 470, loss: 0.08947452157735825
step: 480, loss: 0.009427743032574654
step: 490, loss: 0.06880935281515121
step: 500, loss: 0.0030357937794178724
step: 510, loss: 0.011780834756791592
step: 520, loss: 0.3757750391960144
step: 530, loss: 0.08865208178758621
/usr/local/lib/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=0.9475138121546962, f1=0.944954128440367,  
best_f1=0.944954128440367  
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:  
An input tensor was not cuda.  
    warnings.warn("An input tensor was not cuda.")  
step: 0, loss: 0.15370526909828186  
step: 10, loss: 0.03287733718752861  
step: 20, loss: 0.003759226994588971  
step: 30, loss: 0.0248874481767416  
step: 40, loss: 0.008828647434711456  
step: 50, loss: 0.05233617126941681  
step: 60, loss: 0.011494728736579418  
step: 70, loss: 0.003129238961264491  
step: 80, loss: 0.0699615329504013  
step: 90, loss: 0.0013681353302672505  
step: 100, loss: 0.09921572357416153  
step: 110, loss: 0.011414390057325363  
step: 120, loss: 0.08355653285980225  
step: 130, loss: 0.12691199779510498  
step: 140, loss: 0.007877463474869728  
step: 150, loss: 0.00786168035119772  
step: 160, loss: 0.015748558565974236  
step: 170, loss: 0.02879943512380123  
step: 180, loss: 0.00794746819883585  
step: 190, loss: 0.0019380536396056414  
step: 200, loss: 0.01895117573440075  
step: 210, loss: 0.032254740595817566  
step: 220, loss: 0.05944984406232834  
step: 230, loss: 0.014080232009291649  
step: 240, loss: 0.007296237628906965  
step: 250, loss: 0.035379279404878616  
step: 260, loss: 0.017136525362730026  
step: 270, loss: 0.014353403821587563  
step: 280, loss: 0.0036163204349577427  
step: 290, loss: 0.02147928811609745  
step: 300, loss: 0.2593390643596649  
step: 310, loss: 0.06302858144044876  
step: 320, loss: 0.03504505753517151  
step: 330, loss: 0.076636403799057  
step: 340, loss: 0.05632772669196129  
step: 350, loss: 0.0573115311563015  
step: 360, loss: 0.041603267192840576  
step: 370, loss: 0.03347015008330345  
step: 380, loss: 0.006328487768769264  
step: 390, loss: 0.012964210473001003  
step: 400, loss: 0.17149673402309418  
step: 410, loss: 0.07126011699438095  
step: 420, loss: 0.006263755261898041
```

```

step: 430, loss: 0.014844579622149467
step: 440, loss: 0.05665489658713341
step: 450, loss: 0.00945302750915289
step: 460, loss: 0.1355389654636383
step: 470, loss: 0.06507526338100433
step: 480, loss: 0.22667457163333893
step: 490, loss: 0.006806590128690004
step: 500, loss: 0.01163868885487318
step: 510, loss: 0.004002810921519995
step: 520, loss: 0.006253946106880903
step: 530, loss: 0.00917065143585205
/usr/local/lib/python3.7/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=0.940677966101695, f1=0.9347724073205067,
best_f1=0.944954128440367
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.03507433459162712
step: 10, loss: 0.0015883909072726965
step: 20, loss: 0.008514279499650002
step: 30, loss: 0.04627411812543869
step: 40, loss: 0.008653289638459682
step: 50, loss: 0.00857322197407484
step: 60, loss: 0.002305032452568412
step: 70, loss: 0.059984318912029266
step: 80, loss: 0.004187441430985928
step: 90, loss: 0.021083399653434753
step: 100, loss: 0.005515010096132755
step: 110, loss: 0.11356570571660995
step: 120, loss: 0.006510508246719837
step: 130, loss: 0.024743905290961266
step: 140, loss: 0.0038225497119128704
step: 150, loss: 0.00830626580864191
step: 160, loss: 0.0006634774035774171
step: 170, loss: 0.018619215115904808
step: 180, loss: 0.03582095354795456
step: 190, loss: 0.0831371545791626
step: 200, loss: 0.007948186248540878
step: 210, loss: 0.0021317750215530396
step: 220, loss: 0.004434855654835701
step: 230, loss: 0.002757107140496373
step: 240, loss: 0.011614036746323109
step: 250, loss: 0.06693655997514725
step: 260, loss: 0.01793641224503517
step: 270, loss: 0.012915664352476597
step: 280, loss: 0.003219541860744357

```

```

step: 290, loss: 0.017376471310853958
step: 300, loss: 0.0009600566700100899
step: 310, loss: 0.00331440893933177
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 4096.0
step: 320, loss: 0.13199810683727264
step: 330, loss: 0.019244644790887833
step: 340, loss: 0.002560913795605302
step: 350, loss: 0.004322261083871126
step: 360, loss: 0.019219014793634415
step: 370, loss: 0.0007525355322286487
step: 380, loss: 0.0008256121072918177
step: 390, loss: 0.0002805322001222521
step: 400, loss: 0.0016435307916253805
step: 410, loss: 0.0048783449456095695
step: 420, loss: 0.004225256387144327
step: 430, loss: 0.07269539684057236
step: 440, loss: 0.006214291788637638
step: 450, loss: 0.043326299637556076
step: 460, loss: 0.11475323885679245
step: 470, loss: 0.0022102785296738148
step: 480, loss: 0.003802276449277997
step: 490, loss: 0.0013818868901580572
step: 500, loss: 0.044637762010097504
step: 510, loss: 0.07492522150278091
step: 520, loss: 0.008791007101535797
step: 530, loss: 0.035699933767318726
/usr/local/lib/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=0.9485873089393237, f1=0.9456221198156683,
best_f1=0.9456221198156683
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0019358722493052483
step: 10, loss: 0.011355835013091564
step: 20, loss: 0.005152831319719553
step: 30, loss: 0.004713123664259911
step: 40, loss: 0.00031360378488898277
step: 50, loss: 0.0686824768781662
step: 60, loss: 0.003157183527946472
step: 70, loss: 0.01976865716278553
step: 80, loss: 0.003562558675184846
step: 90, loss: 0.12863436341285706

```

```
step: 100, loss: 0.07980790734291077
step: 110, loss: 0.0007455796003341675
step: 120, loss: 0.2623266875743866
step: 130, loss: 0.011044860817492008
step: 140, loss: 0.08915965259075165
step: 150, loss: 0.037766795605421066
step: 160, loss: 0.02775431238114834
step: 170, loss: 0.03895983844995499
step: 180, loss: 0.0046665361151099205
step: 190, loss: 0.0008655605488456786
step: 200, loss: 0.0014702416956424713
step: 210, loss: 0.0012151331175118685
step: 220, loss: 0.004093277268111706
step: 230, loss: 0.007205755449831486
step: 240, loss: 0.004865738097578287
step: 250, loss: 0.05847563594579697
step: 260, loss: 0.0008870845194905996
step: 270, loss: 0.0052370112389326096
step: 280, loss: 0.0027063372544944286
step: 290, loss: 0.0016490034759044647
step: 300, loss: 0.01910649612545967
step: 310, loss: 0.04932970181107521
step: 320, loss: 0.011712828651070595
step: 330, loss: 0.03796180337667465
step: 340, loss: 0.0004703276790678501
step: 350, loss: 0.0009537564474157989
step: 360, loss: 4.652145435102284e-05
step: 370, loss: 8.601158333476633e-05
step: 380, loss: 0.0019393251277506351
step: 390, loss: 0.006169174797832966
step: 400, loss: 0.002991320099681616
step: 410, loss: 0.024220865219831467
step: 420, loss: 0.2687503695487976
step: 430, loss: 0.0007684594020247459
step: 440, loss: 0.0022838895674794912
step: 450, loss: 0.0028489213436841965
step: 460, loss: 0.011061116121709347
step: 470, loss: 0.020732877776026726
step: 480, loss: 0.02561073936522007
step: 490, loss: 0.002688839565962553
step: 500, loss: 0.003667506854981184
step: 510, loss: 0.002444045851007104
step: 520, loss: 0.05305091291666031
step: 530, loss: 0.0038104429841041565
/usr/local/lib/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=0.9542725173210163, f1=0.9443678160919541,

best\_f1=0.9443678160919541

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

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

step: 0, loss: 0.005529404152184725

step: 10, loss: 0.0009775900980457664

step: 20, loss: 0.07224899530410767

step: 30, loss: 0.008879968896508217

step: 40, loss: 0.0013485124800354242

step: 50, loss: 0.0005549369379878044

step: 60, loss: 0.010602624155580997

step: 70, loss: 0.0007424473878927529

step: 80, loss: 0.00035856233444064856

step: 90, loss: 0.0006688113790005445

step: 100, loss: 0.01042536087334156

step: 110, loss: 0.008386376313865185

step: 120, loss: 0.0030543815810233355

step: 130, loss: 0.0006019863649271429

step: 140, loss: 0.0007063047378323972

step: 150, loss: 0.00020947790471836925

step: 160, loss: 0.0014051563339307904

step: 170, loss: 0.0017311120172962546

step: 180, loss: 0.00035303013282828033

step: 190, loss: 0.041197579354047775

step: 200, loss: 0.022210512310266495

step: 210, loss: 0.00810728408396244

step: 220, loss: 0.003652711398899555

step: 230, loss: 0.0022469654213637114

step: 240, loss: 0.00016070243145804852

step: 250, loss: 0.11833884567022324

step: 260, loss: 0.0030890400521457195

step: 270, loss: 0.04233158379793167

step: 280, loss: 0.12402670085430145

step: 290, loss: 0.011711574159562588

step: 300, loss: 0.005866894032806158

step: 310, loss: 0.05014616996049881

step: 320, loss: 0.08516576141119003

step: 330, loss: 0.02823386713862419

step: 340, loss: 0.000533841724973172

step: 350, loss: 0.001213539857417345

step: 360, loss: 0.10150207579135895

step: 370, loss: 0.006982052698731422

step: 380, loss: 0.0008848788565956056

step: 390, loss: 0.020280612632632256

```

step: 400, loss: 0.0031468148808926344
step: 410, loss: 0.000853105157148093
step: 420, loss: 0.0011993101797997952
step: 430, loss: 0.0004442096105776727
step: 440, loss: 0.021312009543180466
step: 450, loss: 0.2910473644733429
step: 460, loss: 0.005771875847131014
step: 470, loss: 0.0019577518105506897
step: 480, loss: 0.005419893190264702
step: 490, loss: 0.008704054169356823
step: 500, loss: 0.015706488862633705
step: 510, loss: 0.2991597056388855
step: 520, loss: 0.0005759337218478322
step: 530, loss: 0.005581557750701904
/usr/local/lib/python3.7/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=0.9496535796766744, f1=0.9401473296500922,
best_f1=0.9443678160919541
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0027270030695945024
step: 10, loss: 0.012207675725221634
step: 20, loss: 0.0009849285706877708
step: 30, loss: 0.004188664723187685
step: 40, loss: 0.002131872344762087
step: 50, loss: 0.000563754525501281
step: 60, loss: 0.0009597527678124607
step: 70, loss: 0.00031458292505703866
step: 80, loss: 0.00022810063092038035
step: 90, loss: 0.005712633021175861
step: 100, loss: 0.00023862195666879416
step: 110, loss: 0.00019706838065758348
step: 120, loss: 0.00044637342216446996
step: 130, loss: 0.0005713290302082896
step: 140, loss: 8.56950573506765e-05
step: 150, loss: 9.473740647081286e-05
step: 160, loss: 0.0008409758447669446
step: 170, loss: 0.006444387137889862
step: 180, loss: 0.006303183734416962
step: 190, loss: 0.0022702652495354414
step: 200, loss: 0.002335494151338935
step: 210, loss: 0.11927134543657303
step: 220, loss: 0.0007150303572416306
step: 230, loss: 0.0003487040230538696
step: 240, loss: 0.0010505920508876443
step: 250, loss: 0.01612623780965805

```

```

step: 260, loss: 0.003886917605996132
step: 270, loss: 0.0009117809822782874
step: 280, loss: 0.0024485911708325148
step: 290, loss: 0.0036626968067139387
step: 300, loss: 0.00017416698392480612
step: 310, loss: 0.00022114213788881898
step: 320, loss: 0.0010113188764080405
step: 330, loss: 0.0007516102632507682
step: 340, loss: 0.00044857640750706196
step: 350, loss: 0.0017814188031479716
step: 360, loss: 0.0035618171095848083
step: 370, loss: 0.004019510000944138
step: 380, loss: 0.0008847266435623169
step: 390, loss: 0.0011753769358620048
step: 400, loss: 0.006187814753502607
step: 410, loss: 0.00022973763407208025
step: 420, loss: 0.18618375062942505
step: 430, loss: 0.0006903507164679468
step: 440, loss: 0.02396620251238346
step: 450, loss: 0.004528460558503866
step: 460, loss: 0.001702727866359055
step: 470, loss: 0.18295322358608246
step: 480, loss: 0.009335954673588276
step: 490, loss: 0.002042989945039153
step: 500, loss: 0.0028280094265937805
step: 510, loss: 0.00028350710636004806
step: 520, loss: 0.0014411471784114838
step: 530, loss: 0.0005457318620756269
/usr/local/lib/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=0.9561975768872321, f1=0.9471243042671614,
best_f1=0.9471243042671614
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0003907237551175058
step: 10, loss: 0.002722564386203885
step: 20, loss: 0.0005056365043856204
step: 30, loss: 0.004366529639810324
step: 40, loss: 0.00012513043475337327
step: 50, loss: 0.00012409340706653893
step: 60, loss: 0.0008139932178892195
step: 70, loss: 0.00011498099775053561

```



```
step: 80, loss: 0.010183071717619896
step: 90, loss: 0.0001082465605577454
step: 100, loss: 0.00038738694274798036
step: 110, loss: 0.0001512881281087175
step: 120, loss: 0.0012851323699578643
step: 130, loss: 0.016764769330620766
step: 140, loss: 5.177991624805145e-05
step: 150, loss: 0.008470816537737846
step: 160, loss: 7.144411210902035e-05
step: 170, loss: 0.2571895718574524
step: 180, loss: 0.0026966682635247707
step: 190, loss: 0.002227236283943057
step: 200, loss: 0.008242227137088776
step: 210, loss: 0.05832676962018013
step: 220, loss: 0.0014456245116889477
step: 230, loss: 0.013917211443185806
step: 240, loss: 0.0013473380822688341
step: 250, loss: 0.0005832742317579687
step: 260, loss: 5.631961903418414e-05
step: 270, loss: 0.00041197205428034067
step: 280, loss: 0.0001339373702649027
step: 290, loss: 0.00016315252287313342
step: 300, loss: 6.930858944542706e-05
step: 310, loss: 0.0015812907367944717
step: 320, loss: 0.007622165139764547
step: 330, loss: 0.08344846963882446
step: 340, loss: 0.004381783772259951
step: 350, loss: 0.000511987425852567
step: 360, loss: 0.09770669788122177
step: 370, loss: 0.08913105726242065
step: 380, loss: 0.001403079368174076
step: 390, loss: 0.07027441263198853
step: 400, loss: 0.0020634003449231386
step: 410, loss: 0.0064681582152843475
step: 420, loss: 0.0003451485827099532
step: 430, loss: 0.0003355229855515063
step: 440, loss: 0.01412061508744955
step: 450, loss: 0.0004759422445204109
step: 460, loss: 0.0005877430667169392
step: 470, loss: 0.06421177089214325
step: 480, loss: 0.0011566831963136792
step: 490, loss: 0.0034572086296975613
step: 500, loss: 0.0004244967130944133
step: 510, loss: 0.0006283335969783366
step: 520, loss: 0.00029219617135822773
step: 530, loss: 0.15796300768852234
/usr/local/lib/python3.7/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=0.9494949494949496, f1=0.943758573388203,
best_f1=0.9471243042671614
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.000363372964784503
step: 10, loss: 0.000489253958221525
step: 20, loss: 0.0003210904833395034
step: 30, loss: 0.08701508492231369
step: 40, loss: 0.0006645518587902188
step: 50, loss: 0.0007633932982571423
step: 60, loss: 0.000688070198521018
step: 70, loss: 0.025727545842528343
step: 80, loss: 0.06157407909631729
step: 90, loss: 0.05511412397027016
step: 100, loss: 0.0007238587131723762
step: 110, loss: 0.009635805152356625
step: 120, loss: 0.006952681113034487
step: 130, loss: 0.0015520998276770115
step: 140, loss: 0.00026591046480461955
step: 150, loss: 0.00042230874532833695
step: 160, loss: 0.0011070739710703492
step: 170, loss: 0.013342779129743576
step: 180, loss: 0.00013487313117366284
step: 190, loss: 5.859455995960161e-05
step: 200, loss: 0.000464351789560169
step: 210, loss: 0.0007214761571958661
step: 220, loss: 0.0008802501833997667
step: 230, loss: 0.0005444721318781376
step: 240, loss: 0.00038075039628893137
step: 250, loss: 0.0004898448241874576
step: 260, loss: 0.00022208041627891362
step: 270, loss: 0.00043375923996791244
step: 280, loss: 0.015320554375648499
step: 290, loss: 0.00018727866699919105
step: 300, loss: 0.0014862969983369112
step: 310, loss: 0.004830300807952881
step: 320, loss: 0.0001668889744905755
step: 330, loss: 0.00043681106762960553
step: 340, loss: 0.00038573029451072216
step: 350, loss: 0.0058491225354373455
step: 360, loss: 0.0006128688110038638
step: 370, loss: 0.000513209670316428
step: 380, loss: 0.000425867474405095
step: 390, loss: 3.646362529252656e-05
step: 400, loss: 0.00020757279708050191
step: 410, loss: 0.00022302249271888286
```

```

step: 420, loss: 0.0001560504169901833
step: 430, loss: 0.0008881050744093955
step: 440, loss: 5.286662053549662e-05
step: 450, loss: 0.002105261664837599
step: 460, loss: 4.6639295760542154e-05
step: 470, loss: 3.620285860961303e-05
step: 480, loss: 4.709874701802619e-05
step: 490, loss: 0.03766992688179016
step: 500, loss: 0.00015215099847409874
step: 510, loss: 0.00010092469165101647
step: 520, loss: 0.014857657253742218
step: 530, loss: 0.01608290523290634
/usr/local/lib/python3.7/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=0.948405253283302, f1=0.9427640763145649,
best_f1=0.9471243042671614
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00047244224697351456
step: 10, loss: 0.001008383696898818
step: 20, loss: 5.515876546269283e-05
step: 30, loss: 0.001858410076238215
step: 40, loss: 4.548938522930257e-05
step: 50, loss: 0.00036529090721160173
step: 60, loss: 0.008571546524763107
step: 70, loss: 0.008730224333703518
step: 80, loss: 0.0026706992648541927
step: 90, loss: 0.0006557951564900577
step: 100, loss: 0.10551457107067108
step: 110, loss: 0.0663413479924202
step: 120, loss: 0.00027669177507050335
step: 130, loss: 0.0003168070106767118
step: 140, loss: 0.0011456306092441082
step: 150, loss: 0.006783670745790005
step: 160, loss: 0.08598891645669937
step: 170, loss: 7.954165630508214e-05
step: 180, loss: 0.00012503853940870613
step: 190, loss: 4.4367119699018076e-05
step: 200, loss: 7.948206621222198e-05
step: 210, loss: 0.003983668051660061
step: 220, loss: 0.0005411064485087991
step: 230, loss: 0.027526361867785454
step: 240, loss: 6.939529703231528e-05
step: 250, loss: 4.992407775716856e-05
step: 260, loss: 0.010155759751796722
step: 270, loss: 5.3700081480201334e-05

```

```

step: 280, loss: 0.003382977331057191
step: 290, loss: 0.0025294062215834856
step: 300, loss: 0.007323050871491432
step: 310, loss: 0.08090304583311081
step: 320, loss: 0.01861836388707161
step: 330, loss: 0.0007352142711170018
step: 340, loss: 0.0001220603589899838
step: 350, loss: 0.0005347086698748171
step: 360, loss: 0.00015788013115525246
step: 370, loss: 0.0011795145692303777
step: 380, loss: 0.0014417872298508883
step: 390, loss: 0.00015524282935075462
step: 400, loss: 7.249390910146758e-05
step: 410, loss: 0.0009415920940227807
step: 420, loss: 4.6451510570477694e-05
step: 430, loss: 8.081840496743098e-05
step: 440, loss: 0.00025062309578061104
step: 450, loss: 9.073725232155994e-05
step: 460, loss: 4.0388404158875346e-05
step: 470, loss: 0.025314049795269966
step: 480, loss: 0.00010482700599823147
step: 490, loss: 0.0003002967860084027
step: 500, loss: 0.00030812682234682143
step: 510, loss: 0.0001381646143272519
step: 520, loss: 0.0003050520608667284
step: 530, loss: 0.017924839630723
/usr/local/lib/python3.7/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=0.9548447789275636, f1=0.9453526389537599,
best_f1=0.9471243042671614
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00021631539857480675
step: 10, loss: 0.00020829896675422788
step: 20, loss: 0.00018456987163517624
step: 30, loss: 0.0001238123222719878
step: 40, loss: 3.283928890596144e-05
step: 50, loss: 0.0006043485482223332
step: 60, loss: 0.00041152103221975267
step: 70, loss: 4.4290729420026764e-05
step: 80, loss: 0.000681036151945591
step: 90, loss: 0.00037473271368071437
step: 100, loss: 5.3687883337261155e-05
step: 110, loss: 3.694135375553742e-05
step: 120, loss: 4.3073003325844184e-05
step: 130, loss: 2.677236989256926e-05

```

```

step: 140, loss: 4.440512566361576e-05
step: 150, loss: 9.504861372988671e-05
step: 160, loss: 0.001206119661219418
step: 170, loss: 3.4592379961395636e-05
step: 180, loss: 3.729837044375017e-05
step: 190, loss: 0.023129304870963097
step: 200, loss: 6.522507464978844e-05
step: 210, loss: 3.8008914998499677e-05
step: 220, loss: 0.031145257875323296
step: 230, loss: 0.006241777911782265
step: 240, loss: 0.0016075781313702464
step: 250, loss: 0.00026353797875344753
step: 260, loss: 0.0006836463580839336
step: 270, loss: 0.002042514504864812
step: 280, loss: 0.1335996687412262
step: 290, loss: 0.0002544767630752176
step: 300, loss: 0.00024206758826039732
step: 310, loss: 0.0002238776214653626
step: 320, loss: 0.0004822692135348916
step: 330, loss: 2.3781209165463224e-05
step: 340, loss: 0.00034354557283222675
step: 350, loss: 0.0004031905555166304
step: 360, loss: 0.00021028699120506644
step: 370, loss: 0.001732174656353891
step: 380, loss: 0.0009627981926314533
step: 390, loss: 0.0018673281883820891
step: 400, loss: 0.00018249715503770858
step: 410, loss: 0.0012991678668186069
step: 420, loss: 0.00015037419507279992
step: 430, loss: 0.0011226708302274346
step: 440, loss: 0.0008390527218580246
step: 450, loss: 0.0008269201498478651
step: 460, loss: 0.0024487001355737448
step: 470, loss: 0.0004767078207805753
step: 480, loss: 0.00017236908024642617
step: 490, loss: 8.037563384277746e-05
step: 500, loss: 0.0005959414411336184
step: 510, loss: 0.0002016128128161654
step: 520, loss: 0.00011640186130534858
step: 530, loss: 0.053541604429483414
/usr/local/lib/python3.7/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=0.9556696220251982, f1=0.9527340129749768,
best_f1=0.9471243042671614
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")

```

step: 0, loss: 3.504462802084163e-05  
step: 10, loss: 0.0005772520089522004  
step: 20, loss: 0.000549144227989018  
step: 30, loss: 0.0017896555364131927  
step: 40, loss: 0.0002343309170100838  
step: 50, loss: 9.485489135840908e-05  
step: 60, loss: 0.000619693542830646  
step: 70, loss: 0.00017899676458910108  
step: 80, loss: 0.10728106647729874  
step: 90, loss: 0.00036632880801334977  
step: 100, loss: 0.0207046065479517  
step: 110, loss: 6.482416938524693e-05  
step: 120, loss: 0.0002720987831708044  
step: 130, loss: 0.0005897378432564437  
step: 140, loss: 0.0003244457475375384  
step: 150, loss: 6.204300734680146e-05  
step: 160, loss: 0.007069696672260761  
step: 170, loss: 0.00011564139276742935  
step: 180, loss: 5.215683995629661e-05  
step: 190, loss: 0.0004924538661725819  
step: 200, loss: 0.0011215978302061558  
step: 210, loss: 0.0003141481429338455  
step: 220, loss: 5.017990406486206e-05  
step: 230, loss: 0.00013076962204650044  
step: 240, loss: 0.017301466315984726  
step: 250, loss: 5.96999452682212e-05  
step: 260, loss: 2.605328700155951e-05  
step: 270, loss: 0.0014245028141885996  
step: 280, loss: 5.871981556992978e-05  
step: 290, loss: 9.659455099608749e-05  
step: 300, loss: 0.0009250009898096323  
step: 310, loss: 4.1459628846496344e-05  
step: 320, loss: 0.0008038975647650659  
step: 330, loss: 0.0014617303386330605  
step: 340, loss: 4.255982639733702e-05  
step: 350, loss: 3.311686668894254e-05  
step: 360, loss: 0.00011108165926998481  
step: 370, loss: 2.1333553377189673e-05  
step: 380, loss: 0.012988357804715633  
step: 390, loss: 0.0014421362429857254  
step: 400, loss: 2.3747690647724085e-05  
step: 410, loss: 0.00026235170662403107  
step: 420, loss: 0.00012125380453653634  
step: 430, loss: 9.182259964291006e-05  
step: 440, loss: 0.002778193447738886  
step: 450, loss: 0.00016536911425646394  
step: 460, loss: 0.00019463409262243658  
step: 470, loss: 0.0006041573942638934

```

step: 480, loss: 0.000785279436968267
step: 490, loss: 3.579805343179032e-05
step: 500, loss: 0.00016190475434996188
step: 510, loss: 0.007029622793197632
step: 520, loss: 0.00019667006563395262
step: 530, loss: 5.6875993323046714e-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=0.9546296296296296, f1=0.9479981592268752,
best_f1=0.9471243042671614
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 5.01203176099807e-05
step: 10, loss: 0.00017383653903380036
step: 20, loss: 4.416376759763807e-05
step: 30, loss: 2.169896470149979e-05
step: 40, loss: 0.00010868276149267331
step: 50, loss: 0.0006799794500693679
step: 60, loss: 0.00023060146486386657
step: 70, loss: 0.00011799567437265068
step: 80, loss: 0.0004472198197618127
step: 90, loss: 9.46956715779379e-05
step: 100, loss: 5.4927724704612046e-05
step: 110, loss: 1.8666760297492146e-05
step: 120, loss: 4.806354991160333e-05
step: 130, loss: 0.0005057569942437112
step: 140, loss: 0.0003330439212732017
step: 150, loss: 0.0006233245367184281
step: 160, loss: 5.684678762918338e-05
step: 170, loss: 0.00013669686450157315
step: 180, loss: 0.008953381329774857
step: 190, loss: 3.6043325962964445e-05
step: 200, loss: 7.440806075464934e-05
step: 210, loss: 3.672325692605227e-05
step: 220, loss: 2.4488888811902143e-05
step: 230, loss: 0.0010336068226024508
step: 240, loss: 0.0005022204713895917
step: 250, loss: 8.895158680388704e-05
step: 260, loss: 5.302633508108556e-05
step: 270, loss: 0.00015978280862327665
step: 280, loss: 0.001561817480251193
step: 290, loss: 1.566058926982805e-05
step: 300, loss: 3.156192906317301e-05
step: 310, loss: 2.5851839382085018e-05
step: 320, loss: 0.0002060453116428107
step: 330, loss: 4.972704118699767e-05

```

```

step: 340, loss: 3.166665555909276e-05
step: 350, loss: 1.9310253264848143e-05
step: 360, loss: 0.10800308734178543
step: 370, loss: 5.607270577456802e-05
step: 380, loss: 2.2491840354632586e-05
step: 390, loss: 2.430657514196355e-05
step: 400, loss: 2.734641748247668e-05
step: 410, loss: 5.270304609439336e-05
step: 420, loss: 0.0003147353127133101
step: 430, loss: 3.5052664316026494e-05
step: 440, loss: 1.1656168680929113e-05
step: 450, loss: 9.031100489664823e-05
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 8192.0
step: 460, loss: 0.005790526047348976
step: 470, loss: 0.0004138429940212518
step: 480, loss: 8.657470971229486e-06
step: 490, loss: 3.2290110539179295e-05
step: 500, loss: 1.4338148503156845e-05
step: 510, loss: 0.00024409868638031185
step: 520, loss: 2.6675248591345735e-05
step: 530, loss: 0.0003953198902308941
/usr/local/lib/python3.7/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=0.9552656104380243, f1=0.9475638051044084,
best_f1=0.9471243042671614
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.001136064762249589
step: 10, loss: 0.0024626206140965223
step: 20, loss: 1.5303161490010098e-05
step: 30, loss: 0.0011347560212016106
step: 40, loss: 6.455412221839651e-05
step: 50, loss: 6.85074701323174e-05
step: 60, loss: 2.8664377168752253e-05
step: 70, loss: 3.26008812407963e-05
step: 80, loss: 0.00017582708096597344
step: 90, loss: 4.829854515264742e-05
step: 100, loss: 3.6288955016061664e-05
step: 110, loss: 3.817009201156907e-05
step: 120, loss: 4.775720299221575e-05
step: 130, loss: 1.8491897208150476e-05
step: 140, loss: 3.455749902059324e-05
step: 150, loss: 0.0005892177578061819
step: 160, loss: 4.526145494310185e-05
step: 170, loss: 1.8793529307004064e-05
step: 180, loss: 2.892926386266481e-05

```



```

step: 190, loss: 0.00014525503502227366
step: 200, loss: 3.2054027542471886e-05
step: 210, loss: 4.229361002217047e-05
step: 220, loss: 4.516686749411747e-05
step: 230, loss: 2.441791002638638e-05
step: 240, loss: 0.0006357752135954797
step: 250, loss: 0.00012931933451909572
step: 260, loss: 2.5587543859728612e-05
step: 270, loss: 0.00034710849286057055
step: 280, loss: 2.5479361283942126e-05
step: 290, loss: 9.886784937407356e-06
step: 300, loss: 4.634396100300364e-05
step: 310, loss: 3.779317921726033e-05
step: 320, loss: 3.6388850276125595e-05
step: 330, loss: 0.0012225676327943802
step: 340, loss: 0.00044268398778513074
step: 350, loss: 1.4520880540658254e-05
step: 360, loss: 2.071179733320605e-05
step: 370, loss: 8.737706957617775e-05
step: 380, loss: 0.0002338382473681122
step: 390, loss: 1.2721538951154798e-05
step: 400, loss: 0.0023848488926887512
step: 410, loss: 9.320549906988163e-06
step: 420, loss: 1.5727604477433488e-05
step: 430, loss: 3.522250335663557e-05
step: 440, loss: 1.892762702482287e-05
step: 450, loss: 1.5012609765108209e-05
step: 460, loss: 0.00018731319869402796
step: 470, loss: 2.9084563720971346e-05
step: 480, loss: 1.450973650207743e-05
step: 490, loss: 4.119515142519958e-05
step: 500, loss: 0.0021394367795437574
step: 510, loss: 0.00018657572218216956
step: 520, loss: 5.393085302785039e-05
step: 530, loss: 0.0003247068088967353
/usr/local/lib/python3.7/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=0.9551820728291317, f1=0.9468779123951537,
best_f1=0.9471243042671614
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 2.9747987355221994e-05
step: 10, loss: 0.0005938876420259476
step: 20, loss: 0.00015338478260673583
step: 30, loss: 0.0017464427510276437
step: 40, loss: 2.6004940082202666e-05

```

step: 50, loss: 0.0001268385531147942  
step: 60, loss: 3.9983140595722944e-05  
step: 70, loss: 2.1628158719977364e-05  
step: 80, loss: 2.5620931410230696e-05  
step: 90, loss: 2.5036215447471477e-05  
step: 100, loss: 1.0803166333062109e-05  
step: 110, loss: 2.8809414288843982e-05  
step: 120, loss: 1.8599836039356887e-05  
step: 130, loss: 1.9646622604341246e-05  
step: 140, loss: 2.601748929009773e-05  
step: 150, loss: 2.3077072910382412e-05  
step: 160, loss: 2.8074720830773003e-05  
step: 170, loss: 2.8716800443362445e-05  
step: 180, loss: 6.191983266035095e-05  
step: 190, loss: 0.0026829584967345  
step: 200, loss: 0.0013407740043476224  
step: 210, loss: 2.2111236830824055e-05  
step: 220, loss: 1.7195488908328116e-05  
step: 230, loss: 0.0003794324875343591  
step: 240, loss: 1.5686850019847043e-05  
step: 250, loss: 1.1335855560901109e-05  
step: 260, loss: 1.3399620911513921e-05  
step: 270, loss: 1.995575803448446e-05  
step: 280, loss: 1.210327900480479e-05  
step: 290, loss: 2.1644676962750964e-05  
step: 300, loss: 1.5172736311797053e-05  
step: 310, loss: 0.027323201298713684  
step: 320, loss: 2.9571718187071383e-05  
step: 330, loss: 2.4623303033877164e-05  
step: 340, loss: 2.3561577108921483e-05  
step: 350, loss: 0.0005803278181701899  
step: 360, loss: 3.7511221307795495e-05  
step: 370, loss: 2.9294513296918012e-05  
step: 380, loss: 0.00010728052438935265  
step: 390, loss: 3.438362182350829e-05  
step: 400, loss: 0.005160119384527206  
step: 410, loss: 8.912247722037137e-05  
step: 420, loss: 1.5247140254359692e-05  
step: 430, loss: 9.585038242221344e-06  
step: 440, loss: 0.0009688138961791992  
step: 450, loss: 0.001367163727991283  
step: 460, loss: 0.0008695179130882025  
step: 470, loss: 1.5247208466462325e-05  
step: 480, loss: 0.002361948136240244  
step: 490, loss: 0.00010925287642749026  
step: 500, loss: 0.0005928142927587032  
step: 510, loss: 4.31550397479441e-05  
step: 520, loss: 6.196799949975684e-05

```

step: 530, loss: 1.577985312906094e-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=0.9556282111163008, f1=0.9468779123951537,
best_f1=0.9471243042671614
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 [00:27, 207.49it/s]
load_f1 = 0.9559573481687529
real_f1 = 0.9536178107606679
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 [00:26, 166.52it/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: 0.5155778527259827
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 10, loss: 0.4549299478530884
step: 20, loss: 0.4210590124130249
step: 30, loss: 0.3467528223991394
step: 40, loss: 0.35602065920829773
step: 50, loss: 0.41927188634872437
step: 60, loss: 0.4874627888202667
step: 70, loss: 0.30477967858314514
step: 80, loss: 0.3670080304145813
step: 90, loss: 0.2171422690153122
step: 100, loss: 0.2738848328590393
step: 110, loss: 0.2638636529445648
step: 120, loss: 0.3125879466533661
step: 130, loss: 0.268472284078598
step: 140, loss: 0.4236496686935425
step: 150, loss: 0.18624068796634674
step: 160, loss: 0.2984656095504761
step: 170, loss: 0.10571030527353287
step: 180, loss: 0.25917384028434753
step: 190, loss: 0.4443119764328003
step: 200, loss: 0.26302197575569153
step: 210, loss: 0.32049360871315
/usr/local/lib/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=0.5580448065173116, f1=0.6021505376344086,
best_f1=0.6021505376344086
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.2562345564365387
step: 10, loss: 0.07473249733448029
step: 20, loss: 0.3290693759918213
step: 30, loss: 0.24858036637306213
step: 40, loss: 0.3928782343864441
step: 50, loss: 0.13339802622795105
step: 60, loss: 0.3200238049030304
step: 70, loss: 0.17974668741226196
step: 80, loss: 0.2757015824317932
step: 90, loss: 0.20416006445884705
step: 100, loss: 0.44972139596939087
step: 110, loss: 0.3422737419605255
step: 120, loss: 0.0958266630768776
step: 130, loss: 0.26221784949302673
step: 140, loss: 0.10715440660715103
step: 150, loss: 0.2984318733215332
step: 160, loss: 0.07618987560272217
step: 170, loss: 0.2587047517299652
step: 180, loss: 0.13406042754650116
step: 190, loss: 0.2423725575208664
step: 200, loss: 0.09407778829336166
step: 210, loss: 0.09281625598669052
/usr/local/lib/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=0.6211180124223603, f1=0.6495049504950495,
best_f1=0.6495049504950495
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.05826778709888458
step: 10, loss: 0.07293475419282913
step: 20, loss: 0.29122474789619446
step: 30, loss: 0.09951261430978775
step: 40, loss: 0.18458281457424164
step: 50, loss: 0.12796422839164734
step: 60, loss: 0.19532734155654907
step: 70, loss: 0.0559188649058342

```

```

step: 80, loss: 0.13691014051437378
step: 90, loss: 0.04730089381337166
step: 100, loss: 0.11861715465784073
step: 110, loss: 0.13230367004871368
step: 120, loss: 0.12355025857686996
step: 130, loss: 0.14565418660640717
step: 140, loss: 0.08644027262926102
step: 150, loss: 0.0984036773443222
step: 160, loss: 0.1656700074672699
step: 170, loss: 0.22207826375961304
step: 180, loss: 0.10968732088804245
step: 190, loss: 0.03209807351231575
step: 200, loss: 0.14467589557170868
step: 210, loss: 0.09719985723495483
/usr/local/lib/python3.7/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=0.6194331983805668, f1=0.6268656716417911,
best_f1=0.6495049504950495
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.05604707449674606
step: 10, loss: 0.0844237357378006
step: 20, loss: 0.1981213390827179
step: 30, loss: 0.11955415457487106
step: 40, loss: 0.02689272351562977
step: 50, loss: 0.10619001090526581
step: 60, loss: 0.09742650389671326
step: 70, loss: 0.08714254200458527
step: 80, loss: 0.08191680163145065
step: 90, loss: 0.028596075251698494
step: 100, loss: 0.14794030785560608
step: 110, loss: 0.2448480874300003
step: 120, loss: 0.17996545135974884
step: 130, loss: 0.21500681340694427
step: 140, loss: 0.17128226161003113
step: 150, loss: 0.19281503558158875
step: 160, loss: 0.0665188729763031
step: 170, loss: 0.06345128268003464
step: 180, loss: 0.024969002231955528
step: 190, loss: 0.19598563015460968
step: 200, loss: 0.1179322823882103
step: 210, loss: 0.28677377104759216
/usr/local/lib/python3.7/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=0.6205607476635514, f1=0.6469428007889546,

```

```

best_f1=0.6495049504950495
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.09146839380264282
step: 10, loss: 0.01393413357436657
step: 20, loss: 0.022417984902858734
step: 30, loss: 0.0343298502266407
step: 40, loss: 0.05592191591858864
step: 50, loss: 0.14227190613746643
step: 60, loss: 0.06748957186937332
step: 70, loss: 0.36031389236450195
step: 80, loss: 0.06186017766594887
step: 90, loss: 0.04791367053985596
step: 100, loss: 0.201520636677742
step: 110, loss: 0.23018178343772888
step: 120, loss: 0.09179428219795227
step: 130, loss: 0.03844389691948891
step: 140, loss: 0.1782013326883316
step: 150, loss: 0.08067953586578369
step: 160, loss: 0.05448570102453232
step: 170, loss: 0.028607646003365517
step: 180, loss: 0.18256275355815887
step: 190, loss: 0.17674468457698822
step: 200, loss: 0.09939305484294891
step: 210, loss: 0.02697477675974369
/usr/local/lib/python3.7/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=0.6195899772209568, f1=0.64, best_f1=0.6495049504950495
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.014090974815189838
step: 10, loss: 0.08234016597270966
step: 20, loss: 0.013842378742992878
step: 30, loss: 0.16951802372932434
step: 40, loss: 0.030039899051189423
step: 50, loss: 0.029522927477955818
step: 60, loss: 0.1345161646604538
step: 70, loss: 0.05797918885946274
step: 80, loss: 0.16371388733386993
step: 90, loss: 0.013280712068080902
step: 100, loss: 0.062232181429862976
step: 110, loss: 0.011154682375490665
step: 120, loss: 0.007420635316520929
step: 130, loss: 0.013572599738836288
step: 140, loss: 0.07603595405817032

```



```

step: 150, loss: 0.04982564598321915
step: 160, loss: 0.09939340502023697
step: 170, loss: 0.0771949291229248
step: 180, loss: 0.08326806128025055
step: 190, loss: 0.0746644139289856
step: 200, loss: 0.2636849582195282
step: 210, loss: 0.0740513727068901
/usr/local/lib/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=0.6330935251798561, f1=0.6791044776119403,
best_f1=0.6791044776119403
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0467936247587204
step: 10, loss: 0.0031366439070552588
step: 20, loss: 0.044848937541246414
step: 30, loss: 0.09173718094825745
step: 40, loss: 0.036165300756692886
step: 50, loss: 0.12875673174858093
step: 60, loss: 0.08605720102787018
step: 70, loss: 0.14492319524288177
step: 80, loss: 0.05187069624662399
step: 90, loss: 0.08592003583908081
step: 100, loss: 0.015270018018782139
step: 110, loss: 0.08661326766014099
step: 120, loss: 0.031904496252536774
step: 130, loss: 0.09759752452373505
step: 140, loss: 0.032417573034763336
step: 150, loss: 0.025669490918517113
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 8192.0
step: 160, loss: 0.202272430062294
step: 170, loss: 0.06953588873147964
step: 180, loss: 0.01935150846838951
step: 190, loss: 0.04871245101094246
step: 200, loss: 0.012845885939896107
step: 210, loss: 0.18136243522167206
/usr/local/lib/python3.7/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=0.606425702811245, f1=0.6598778004073319,
best_f1=0.6791044776119403
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:

```

```

An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.06370304524898529
step: 10, loss: 0.0029892725870013237
step: 20, loss: 0.0020594075322151184
step: 30, loss: 0.004129985347390175
step: 40, loss: 0.010907850228250027
step: 50, loss: 0.0005164567264728248
step: 60, loss: 0.01460691262036562
step: 70, loss: 0.03664705157279968
step: 80, loss: 0.11457548290491104
step: 90, loss: 0.014430524781346321
step: 100, loss: 0.1408519595861435
step: 110, loss: 0.011408142745494843
step: 120, loss: 0.05065353587269783
step: 130, loss: 0.003978252876549959
step: 140, loss: 0.030349375680088997
step: 150, loss: 0.062485065311193466
step: 160, loss: 0.054153501987457275
step: 170, loss: 0.14880135655403137
step: 180, loss: 0.024286920204758644
step: 190, loss: 0.00465440982952714
step: 200, loss: 0.02883940376341343
step: 210, loss: 0.06088022142648697
/usr/local/lib/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=0.6362098138747885, f1=0.6308724832214765,
best_f1=0.6308724832214765
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.03650864213705063
step: 10, loss: 0.022350018844008446
step: 20, loss: 0.11507356911897659
step: 30, loss: 0.003893547225743532
step: 40, loss: 0.00713984714820981
step: 50, loss: 0.04833120480179787
step: 60, loss: 0.006528419908136129
step: 70, loss: 0.003538592718541622
step: 80, loss: 0.43777501583099365
step: 90, loss: 0.02076655812561512
step: 100, loss: 0.06621020287275314
step: 110, loss: 0.030348701402544975

```

```

step: 120, loss: 0.035650692880153656
step: 130, loss: 0.08792688697576523
step: 140, loss: 0.05886368826031685
step: 150, loss: 0.1446310579776764
step: 160, loss: 0.06500979512929916
step: 170, loss: 0.008998574689030647
step: 180, loss: 0.015772460028529167
step: 190, loss: 0.10775304585695267
step: 200, loss: 0.02159402333199978
step: 210, loss: 0.007951818406581879
/usr/local/lib/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=0.6460905349794239, f1=0.6872427983539094,
best_f1=0.6872427983539094
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0007946825935505331
step: 10, loss: 0.014281034469604492
step: 20, loss: 0.06531878560781479
step: 30, loss: 0.001954765524715185
step: 40, loss: 0.006046005990356207
step: 50, loss: 0.03896232321858406
step: 60, loss: 0.0004035198944620788
step: 70, loss: 0.0033294600434601307
step: 80, loss: 0.0022967655677348375
step: 90, loss: 0.013368036597967148
step: 100, loss: 0.0019880631007254124
step: 110, loss: 0.0009769027819857001
step: 120, loss: 0.006812750827521086
step: 130, loss: 0.018560148775577545
step: 140, loss: 0.0013402783079072833
step: 150, loss: 0.0002617717836983502
step: 160, loss: 0.0012007065815851092
step: 170, loss: 0.017706220969557762
step: 180, loss: 0.004200418945401907
step: 190, loss: 0.0012072839308530092
step: 200, loss: 0.02610691823065281
step: 210, loss: 0.037203699350357056
/usr/local/lib/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=0.6487523992322457, f1=0.6691449814126395,
best_f1=0.6691449814126395
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
```

```
warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.07392851263284683
step: 10, loss: 0.0010850524995476007
step: 20, loss: 0.0006368468166328967
step: 30, loss: 0.0035688993521034718
step: 40, loss: 0.023724844679236412
step: 50, loss: 0.03230703994631767
step: 60, loss: 0.003816422773525119
step: 70, loss: 0.0019243961432948709
step: 80, loss: 0.01626993529498577
step: 90, loss: 0.17873656749725342
step: 100, loss: 0.02455108053982258
step: 110, loss: 0.005556050222367048
step: 120, loss: 0.04406498000025749
step: 130, loss: 0.000672699126880616
step: 140, loss: 0.0236563291400671
step: 150, loss: 0.009032146073877811
step: 160, loss: 0.00078672164818272
step: 170, loss: 0.03731634095311165
step: 180, loss: 0.17712868750095367
step: 190, loss: 0.0739326998591423
step: 200, loss: 0.0018096267012879252
step: 210, loss: 0.029065126553177834
/usr/local/lib/python3.7/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=0.6475409836065574, f1=0.6680327868852459,
best_f1=0.6691449814126395
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
```

```
warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0007470162236131728
step: 10, loss: 0.0028912269044667482
step: 20, loss: 0.055608585476875305
step: 30, loss: 0.002365159336477518
step: 40, loss: 0.0010938779450953007
step: 50, loss: 0.007347205653786659
step: 60, loss: 0.0013006675289943814
step: 70, loss: 0.06257250159978867
step: 80, loss: 0.021837159991264343
step: 90, loss: 0.00683048740029335
```

```

step: 100, loss: 0.0013096665497869253
step: 110, loss: 0.06596965342760086
step: 120, loss: 0.0002917995734605938
step: 130, loss: 0.0018416373059153557
step: 140, loss: 0.01808067224919796
step: 150, loss: 8.733119466342032e-05
step: 160, loss: 0.0028132153674960136
step: 170, loss: 0.00174841214902699
step: 180, loss: 0.014916813932359219
step: 190, loss: 0.000921492523048073
step: 200, loss: 0.009275502525269985
step: 210, loss: 0.00025954667944461107
/usr/local/lib/python3.7/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=0.6224066390041494, f1=0.6652892561983471,
best_f1=0.6691449814126395
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0005721098277717829
step: 10, loss: 0.0754278302192688
step: 20, loss: 0.0004934130702167749
step: 30, loss: 0.0006636235048063099
step: 40, loss: 0.0006342585547827184
step: 50, loss: 0.032672978937625885
step: 60, loss: 0.0028503250796347857
step: 70, loss: 0.0022064491640776396
step: 80, loss: 0.05686541274189949
step: 90, loss: 0.004521395545452833
step: 100, loss: 0.010096589103341103
step: 110, loss: 0.16394998133182526
step: 120, loss: 0.001051162602379918
step: 130, loss: 0.000122404468129389
step: 140, loss: 0.0018838188843801618
step: 150, loss: 0.00013624754501506686
step: 160, loss: 0.011759239248931408
step: 170, loss: 0.0004815109714400023
step: 180, loss: 0.015852320939302444
step: 190, loss: 0.03226740285754204
step: 200, loss: 0.0004893321311101317
step: 210, loss: 0.00045045174192637205
/usr/local/lib/python3.7/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=0.6359832635983262, f1=0.6540084388185654,
best_f1=0.6691449814126395
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:

```

```

An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0004535629414021969
step: 10, loss: 0.01225584652274847
step: 20, loss: 0.0012840701965615153
step: 30, loss: 0.012922386638820171
step: 40, loss: 0.0006413118680939078
step: 50, loss: 0.0032047717832028866
step: 60, loss: 0.0038719007279723883
step: 70, loss: 0.0002213678671978414
step: 80, loss: 0.03359135612845421
step: 90, loss: 0.004350944887846708
step: 100, loss: 0.0018070843070745468
step: 110, loss: 0.00015332929615397006
step: 120, loss: 0.00017839166685007513
step: 130, loss: 0.018970375880599022
step: 140, loss: 0.0005133681115694344
step: 150, loss: 0.08049193024635315
step: 160, loss: 0.012999029830098152
step: 170, loss: 0.032794248312711716
step: 180, loss: 0.004714951850473881
step: 190, loss: 0.0031855173874646425
step: 200, loss: 0.020278265699744225
step: 210, loss: 0.05018579587340355
/usr/local/lib/python3.7/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=0.6464646464646464, f1=0.676, best_f1=0.6691449814126395
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0002418714575469494
step: 10, loss: 0.0025451271794736385
step: 20, loss: 0.0021405501756817102
step: 30, loss: 0.018899299204349518
step: 40, loss: 0.0003134789294563234
step: 50, loss: 7.295682735275477e-05
step: 60, loss: 0.014263210818171501
step: 70, loss: 0.019906731322407722
step: 80, loss: 0.0072339195758104324
step: 90, loss: 0.00012581197370309383
step: 100, loss: 0.0006886543706059456
step: 110, loss: 0.02487279288470745
step: 120, loss: 0.0009317626245319843
step: 130, loss: 0.00012559734750539064
step: 140, loss: 0.009036689065396786
step: 150, loss: 0.0017047615256160498
step: 160, loss: 0.009899711236357689

```

```

step: 170, loss: 0.02314736880362034
step: 180, loss: 0.001206891844049096
step: 190, loss: 0.00010162828402826563
step: 200, loss: 0.0025260986294597387
step: 210, loss: 0.02702757529914379
/usr/local/lib/python3.7/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=0.6432989690721649, f1=0.6666666666666666,
best_f1=0.6691449814126395
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 [00:06, 348.50it/s]
load_f1 = 0.6407766990291262
real_f1 = 0.6513409961685824
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 [00:26, 166.77it/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/
```

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: 0.4524528682231903
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 10, loss: 0.4907338619232178
step: 20, loss: 0.27211910486221313
step: 30, loss: 0.39673420786857605
step: 40, loss: 0.30527472496032715
step: 50, loss: 0.28430041670799255
step: 60, loss: 0.4247785806655884
step: 70, loss: 0.42897483706474304
step: 80, loss: 0.16505342721939087
step: 90, loss: 0.29101839661598206
step: 100, loss: 0.4487752616405487
step: 110, loss: 0.2445240467786789
step: 120, loss: 0.3446422517299652
step: 130, loss: 0.31203383207321167
step: 140, loss: 0.1749337911605835
step: 150, loss: 0.29575711488723755
step: 160, loss: 0.22638320922851562
step: 170, loss: 0.34926751255989075
step: 180, loss: 0.16595710813999176
step: 190, loss: 0.19163964688777924
/usr/local/lib/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=0.24048913043478262, f1=0.25844346549192365,

```

```

best_f1=0.25844346549192365
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.32897746562957764
step: 10, loss: 0.29385897517204285
step: 20, loss: 0.7880050539970398
step: 30, loss: 0.14498350024223328
step: 40, loss: 0.3696051239967346
step: 50, loss: 0.35615187883377075
step: 60, loss: 0.24302446842193604
step: 70, loss: 0.31367021799087524
step: 80, loss: 0.04597628116607666
step: 90, loss: 0.20745450258255005
step: 100, loss: 0.1384061872959137
step: 110, loss: 0.1571856141090393
step: 120, loss: 0.04739963263273239
step: 130, loss: 0.1797734647989273
step: 140, loss: 0.25446003675460815
step: 150, loss: 0.2253463715314865
step: 160, loss: 0.3204880952835083
step: 170, loss: 0.04249170795083046
step: 180, loss: 0.03518984094262123
step: 190, loss: 0.07188571244478226
/usr/local/lib/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=0.7861271676300577, f1=0.7988165680473371,
best_f1=0.7988165680473371
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.14519381523132324
step: 10, loss: 0.1213764175772667
step: 20, loss: 0.13455531001091003
step: 30, loss: 0.0952657014131546
step: 40, loss: 0.024308007210493088
step: 50, loss: 0.1472628265619278
step: 60, loss: 0.07685523480176926
step: 70, loss: 0.11076179891824722
step: 80, loss: 0.1212930753827095
step: 90, loss: 0.05316886678338051
step: 100, loss: 0.021097306162118912
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 8192.0

```

```

step: 110, loss: 0.15113277733325958
step: 120, loss: 0.2707178592681885
step: 130, loss: 0.16122573614120483
step: 140, loss: 0.023192953318357468
step: 150, loss: 0.1745539754629135
step: 160, loss: 0.07499146461486816
step: 170, loss: 0.2276066094636917
step: 180, loss: 0.17159993946552277
step: 190, loss: 0.11122038960456848
/usr/local/lib/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=0.8136482939632546, f1=0.8083989501312335,
best_f1=0.8083989501312335
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.022534560412168503
step: 10, loss: 0.1352987289428711
step: 20, loss: 0.040068238973617554
step: 30, loss: 0.013148462399840355
step: 40, loss: 0.14025504887104034
step: 50, loss: 0.01124374009668827
step: 60, loss: 0.04349396377801895
step: 70, loss: 0.02749166637659073
step: 80, loss: 0.008862743154168129
step: 90, loss: 0.028663162142038345
step: 100, loss: 0.15582405030727386
step: 110, loss: 0.14893163740634918
step: 120, loss: 0.08504307270050049
step: 130, loss: 0.0334814116358757
step: 140, loss: 0.1351631134748459
step: 150, loss: 0.009092418476939201
step: 160, loss: 0.01869615912437439
step: 170, loss: 0.13003908097743988
step: 180, loss: 0.03491974622011185
step: 190, loss: 0.01418572012335062
/usr/local/lib/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=0.8324607329842932, f1=0.8494623655913978,
best_f1=0.8494623655913978
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.017391661182045937
step: 10, loss: 0.07477036118507385
step: 20, loss: 0.011140014044940472
step: 30, loss: 0.08745303004980087
step: 40, loss: 0.02060439996421337
step: 50, loss: 0.018678108230233192
step: 60, loss: 0.08724801242351532
step: 70, loss: 0.019251735880970955
step: 80, loss: 0.0014659682055935264
step: 90, loss: 0.2125636637210846
step: 100, loss: 0.17145968973636627
step: 110, loss: 0.09067215025424957
step: 120, loss: 0.0365484394133091
step: 130, loss: 0.2884424924850464
step: 140, loss: 0.029751883819699287
step: 150, loss: 0.2382117211818695
step: 160, loss: 0.04066860303282738
step: 170, loss: 0.10581749677658081
step: 180, loss: 0.016484569758176804
step: 190, loss: 0.013713584281504154
/usr/local/lib/python3.7/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=0.8264462809917356, f1=0.8351648351648351,
best_f1=0.8494623655913978
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.003224472515285015
step: 10, loss: 0.014792942441999912
step: 20, loss: 0.08349142223596573
step: 30, loss: 0.04143908992409706
step: 40, loss: 0.007846100255846977
step: 50, loss: 0.05932077020406723
step: 60, loss: 0.04909524321556091
step: 70, loss: 0.021928641945123672
step: 80, loss: 0.11015064269304276
step: 90, loss: 0.009451083838939667
step: 100, loss: 0.09431928396224976
step: 110, loss: 0.007132353261113167
step: 120, loss: 0.03173824027180672
step: 130, loss: 0.03999459743499756
step: 140, loss: 0.0015092118410393596

```

```

step: 150, loss: 0.008557706139981747
step: 160, loss: 0.05700521171092987
step: 170, loss: 0.036688465625047684
step: 180, loss: 0.03188788890838623
step: 190, loss: 0.003936567343771458
/usr/local/lib/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=0.8369565217391304, f1=0.8337874659400545,
best_f1=0.8337874659400545
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.008806012570858002
step: 10, loss: 0.03935234621167183
step: 20, loss: 0.009343760088086128
step: 30, loss: 0.026387128978967667
step: 40, loss: 0.002684509614482522
step: 50, loss: 0.10411889851093292
step: 60, loss: 0.04248470440506935
step: 70, loss: 0.0035682932939380407
step: 80, loss: 0.00499187596142292
step: 90, loss: 0.0013496403116732836
step: 100, loss: 0.22968028485774994
step: 110, loss: 0.05447835847735405
step: 120, loss: 0.07894329726696014
step: 130, loss: 0.03985690698027611
step: 140, loss: 0.009223251603543758
step: 150, loss: 0.021707020699977875
step: 160, loss: 0.10034465789794922
step: 170, loss: 0.051806677132844925
step: 180, loss: 0.0035703438334167004
step: 190, loss: 0.017221763730049133
/usr/local/lib/python3.7/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=0.8146067415730337, f1=0.8409090909090909,
best_f1=0.8337874659400545
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.050191182643175125
step: 10, loss: 0.01789926551282406
step: 20, loss: 0.0015603273641318083

```

```

step: 30, loss: 0.2533629536628723
step: 40, loss: 0.04932699725031853
step: 50, loss: 0.03417586162686348
step: 60, loss: 0.019867878407239914
step: 70, loss: 0.010573500767350197
step: 80, loss: 0.009977566078305244
step: 90, loss: 0.0009323521517217159
step: 100, loss: 0.0013634904753416777
step: 110, loss: 0.011114557273685932
step: 120, loss: 0.0016037049936130643
step: 130, loss: 0.0017963576829060912
step: 140, loss: 0.013204214163124561
step: 150, loss: 0.002057113917544484
step: 160, loss: 0.0031921351328492165
step: 170, loss: 0.0011507606832310557
step: 180, loss: 0.19511404633522034
step: 190, loss: 0.0023152311332523823
/usr/local/lib/python3.7/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=0.8355795148247979, f1=0.8484848484848484,
best_f1=0.8337874659400545
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0005188150680623949
step: 10, loss: 0.001319206552579999
step: 20, loss: 0.0053605083376169205
step: 30, loss: 0.0010015040170401335
step: 40, loss: 0.12253192067146301
step: 50, loss: 0.01081763580441475
step: 60, loss: 0.04294270649552345
step: 70, loss: 0.003155095735564828
step: 80, loss: 0.006782537326216698
step: 90, loss: 0.011205809190869331
step: 100, loss: 0.0005591197405010462
step: 110, loss: 0.00037901653558947146
step: 120, loss: 0.0013078255578875542
step: 130, loss: 0.11948736757040024
step: 140, loss: 0.01134692132472992
step: 150, loss: 0.0016569209983572364
step: 160, loss: 0.0020877490751445293
step: 170, loss: 0.003609568113461137
step: 180, loss: 0.07667506486177444
step: 190, loss: 0.0006326072034426033
/usr/local/lib/python3.7/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=0.8337595907928389, f1=0.8494623655913978,
best_f1=0.8337874659400545
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.002131679095327854
step: 10, loss: 0.0004263495502527803
step: 20, loss: 0.001966915326192975
step: 30, loss: 0.2815450429916382
step: 40, loss: 0.0026914412155747414
step: 50, loss: 0.0025167304556816816
step: 60, loss: 0.0009985195938497782
step: 70, loss: 0.000855988881085068
step: 80, loss: 0.0015976761933416128
step: 90, loss: 0.0003698104410432279
step: 100, loss: 0.02567587047815323
step: 110, loss: 0.0013943869853392243
step: 120, loss: 0.000962807098403573
step: 130, loss: 0.00796717219054699
step: 140, loss: 0.0005162382731214166
step: 150, loss: 0.00044685660395771265
step: 160, loss: 0.0005195031990297139
step: 170, loss: 0.019152633845806122
step: 180, loss: 0.009658312425017357
step: 190, loss: 0.0004969689762219787
/usr/local/lib/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=0.8378378378378378, f1=0.8324022346368716,
best_f1=0.8324022346368716
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0008024988346733153
step: 10, loss: 0.001802169717848301
step: 20, loss: 0.000561979366466403
step: 30, loss: 0.0017122544813901186
step: 40, loss: 0.0004310287768021226
step: 50, loss: 0.0005994862876832485
step: 60, loss: 0.004688428249210119
step: 70, loss: 0.0006364152650348842
step: 80, loss: 0.003338287118822336
step: 90, loss: 0.0005982921575196087
step: 100, loss: 0.0005334258312359452

```

```

step: 110, loss: 0.23941093683242798
step: 120, loss: 0.0007154257618822157
step: 130, loss: 0.0008126737084239721
step: 140, loss: 0.14616768062114716
step: 150, loss: 0.000703020894434303
step: 160, loss: 0.0016103534726426005
step: 170, loss: 0.002420293167233467
step: 180, loss: 0.002739283721894026
step: 190, loss: 0.0020353677682578564
/usr/local/lib/python3.7/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=0.8238482384823848, f1=0.8222222222222222,
best_f1=0.8324022346368716
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0017361396457999945
step: 10, loss: 0.0012073327088728547
step: 20, loss: 0.0021266101393848658
step: 30, loss: 0.0017217686399817467
step: 40, loss: 0.0006146621308289468
step: 50, loss: 0.0034591336734592915
step: 60, loss: 0.15541177988052368
step: 70, loss: 0.002414047485217452
step: 80, loss: 0.002442194614559412
step: 90, loss: 0.0007013115100562572
step: 100, loss: 0.0008286202792078257
step: 110, loss: 0.010561969131231308
step: 120, loss: 0.003898903261870146
step: 130, loss: 0.001975443447008729
step: 140, loss: 0.0012863982701674104
step: 150, loss: 0.0018359303940087557
step: 160, loss: 0.001220298116095364
step: 170, loss: 0.03301326185464859
step: 180, loss: 0.0005484000430442393
step: 190, loss: 0.0060249860398471355
/usr/local/lib/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=0.847645429362881, f1=0.8441926345609065,
best_f1=0.8441926345609065
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.

```



```

    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0007238429388962686
step: 10, loss: 0.0008522921125404537
step: 20, loss: 0.0010536123299971223
step: 30, loss: 0.007968918420374393
step: 40, loss: 0.0009080563322640955
step: 50, loss: 0.0007999810623005033
step: 60, loss: 0.000802733819000423
step: 70, loss: 0.0029882495291531086
step: 80, loss: 0.0005649300292134285
step: 90, loss: 0.002086290856823325
step: 100, loss: 0.00035952453617937863
step: 110, loss: 0.0005499051185324788
step: 120, loss: 0.0003910073428414762
step: 130, loss: 0.0005928543978370726
step: 140, loss: 0.00135558913461864
step: 150, loss: 0.00040070750401355326
step: 160, loss: 0.0013693647924810648
step: 170, loss: 0.0003493904077913612
step: 180, loss: 0.00047068309504538774
step: 190, loss: 0.23384273052215576
/usr/local/lib/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 13: dev_f1=0.8502673796791445, f1=0.8446866485013624,
best_f1=0.8446866485013624
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0003100736648775637
step: 10, loss: 0.00030023540602996945
step: 20, loss: 0.0010315906256437302
step: 30, loss: 0.0003508291265461594
step: 40, loss: 0.000605689303483814
step: 50, loss: 0.0004288745403755456
step: 60, loss: 0.0006283089169301093
step: 70, loss: 0.0005877601215615869
step: 80, loss: 0.0007920641801320016
step: 90, loss: 0.0003101295151282102
step: 100, loss: 0.00031585185206495225
step: 110, loss: 0.00046427000779658556
step: 120, loss: 0.00042231837869621813
step: 130, loss: 0.0003747475566342473
step: 140, loss: 0.0005883436533622444

```

```

step: 150, loss: 0.000545683316886425
step: 160, loss: 0.0014972533099353313
step: 170, loss: 0.0002384866529610008
step: 180, loss: 0.0003403039008844644
step: 190, loss: 0.0006743537960574031
/usr/local/lib/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=0.8511749347258485, f1=0.8457446808510638,
best_f1=0.8457446808510638
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00040967241511680186
step: 10, loss: 0.00025176480994559824
step: 20, loss: 0.0003250802110414952
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 8192.0
step: 30, loss: 0.0003382269642315805
step: 40, loss: 0.0006465643527917564
step: 50, loss: 0.00115263054613024
step: 60, loss: 0.00030110549414530396
step: 70, loss: 0.0003117382002528757
step: 80, loss: 0.00026170045020990074
step: 90, loss: 0.001055542379617691
step: 100, loss: 0.0006865981267765164
step: 110, loss: 0.00043705798452720046
step: 120, loss: 0.00028607359854504466
step: 130, loss: 0.00035346552613191307
step: 140, loss: 0.0005744096706621349
step: 150, loss: 0.0002357392368139699
step: 160, loss: 0.0003909103979822248
step: 170, loss: 0.0004509420832619071
step: 180, loss: 0.000289744115434587
step: 190, loss: 0.016078295186161995
/usr/local/lib/python3.7/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=0.8511749347258485, f1=0.8533333333333334,
best_f1=0.8457446808510638
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 [00:10, 197.00it/s]
load_f1 = 0.8556149732620322
real_f1 = 0.8486486486486486
733it [00:00, 3406.57it/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 [00:26, 164.10it/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/

```

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: 0.4928053319454193

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

step: 10, loss: 0.4349179267883301

step: 20, loss: 0.33063292503356934

step: 30, loss: 0.40531593561172485

step: 40, loss: 0.5542669296264648

step: 50, loss: 0.31882157921791077

step: 60, loss: 0.5818251371383667

step: 70, loss: 0.3471697270870209

step: 80, loss: 0.2287585288286209

step: 90, loss: 0.20188120007514954

step: 100, loss: 0.17280900478363037

step: 110, loss: 0.41780826449394226

step: 120, loss: 0.31236588954925537

step: 130, loss: 0.3141826391220093

step: 140, loss: 0.4067370891571045

step: 150, loss: 0.31920233368873596

step: 160, loss: 0.398556649684906

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

step: 170, loss: 0.3228471875190735

/usr/local/lib/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=0.1941564561734213, f1=0.1941564561734213,

best\_f1=0.1941564561734213

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

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

step: 0, loss: 0.32559365034103394

step: 10, loss: 0.49822139739990234

step: 20, loss: 0.3404408395290375

step: 30, loss: 0.3393236994743347

step: 40, loss: 0.07340560108423233

step: 50, loss: 0.46622246503829956

step: 60, loss: 0.19561553001403809

step: 70, loss: 0.5166041254997253

```

step: 80, loss: 0.23688791692256927
step: 90, loss: 0.24644088745117188
step: 100, loss: 0.5339024066925049
step: 110, loss: 0.26294073462486267
step: 120, loss: 0.2570662200450897
step: 130, loss: 0.5252088904380798
step: 140, loss: 0.5281322002410889
step: 150, loss: 0.4463338553905487
step: 160, loss: 0.4327203035354614
step: 170, loss: 0.38180431723594666
/usr/local/lib/python3.7/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=0.1941564561734213
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.6099401712417603
step: 10, loss: 0.31569191813468933
step: 20, loss: 0.234957754611969
step: 30, loss: 0.24711769819259644
step: 40, loss: 0.3734431564807892
step: 50, loss: 0.5845890045166016
step: 60, loss: 0.31050196290016174
step: 70, loss: 0.23684018850326538
step: 80, loss: 0.38218942284584045
step: 90, loss: 0.5281054973602295
step: 100, loss: 0.3052397072315216
step: 110, loss: 0.18621434271335602
step: 120, loss: 0.5757997035980225
step: 130, loss: 0.5119825601577759
step: 140, loss: 0.44268882274627686
step: 150, loss: 0.19158978760242462
step: 160, loss: 0.16297093033790588
step: 170, loss: 0.3243035674095154
/usr/local/lib/python3.7/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=0.1941564561734213, f1=0.1941564561734213,
best_f1=0.1941564561734213
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.3867289125919342
step: 10, loss: 0.5340400338172913
step: 20, loss: 0.25765371322631836
step: 30, loss: 0.4281710982322693

```

```

step: 40, loss: 0.24596863985061646
step: 50, loss: 0.3878962993621826
step: 60, loss: 0.6871850490570068
step: 70, loss: 0.3394131064414978
step: 80, loss: 0.4719897210597992
step: 90, loss: 0.31421464681625366
step: 100, loss: 0.38599371910095215
step: 110, loss: 0.4356541633605957
step: 120, loss: 0.4794921875
step: 130, loss: 0.312163770198822
step: 140, loss: 0.2408446967601776
step: 150, loss: 0.680359423160553
step: 160, loss: 0.12367522716522217
step: 170, loss: 0.23533591628074646
/usr/local/lib/python3.7/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=0.1941564561734213
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.378680020570755
step: 10, loss: 0.3231655955314636
step: 20, loss: 0.31657493114471436
step: 30, loss: 0.30738696455955505
step: 40, loss: 0.2504223585128784
step: 50, loss: 0.2397632598876953
step: 60, loss: 0.3061055839061737
step: 70, loss: 0.34375855326652527
step: 80, loss: 0.0818040743470192
step: 90, loss: 0.5570517182350159
step: 100, loss: 0.2815834879875183
step: 110, loss: 0.2552187442779541
step: 120, loss: 0.12900099158287048
step: 130, loss: 0.23849014937877655
step: 140, loss: 0.16912981867790222
step: 150, loss: 0.31275832653045654
step: 160, loss: 0.2439342439174652
step: 170, loss: 0.3759259581565857
/usr/local/lib/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=0.23356231599607455, f1=0.23814541622760801,

```

```

best_f1=0.23814541622760801
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.3059719204902649
step: 10, loss: 0.10533169656991959
step: 20, loss: 0.3565874993801117
step: 30, loss: 0.24812135100364685
step: 40, loss: 0.4786665737628937
step: 50, loss: 0.24954348802566528
step: 60, loss: 0.3985234797000885
step: 70, loss: 0.37315046787261963
step: 80, loss: 0.21431003510951996
step: 90, loss: 0.30784642696380615
step: 100, loss: 0.18381322920322418
step: 110, loss: 0.448807030916214
step: 120, loss: 0.43727946281433105
step: 130, loss: 0.49232590198516846
step: 140, loss: 0.25271695852279663
step: 150, loss: 0.18028776347637177
step: 160, loss: 0.31032586097717285
step: 170, loss: 0.2544368803501129
/usr/local/lib/python3.7/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=0.1941564561734213, f1=0.1941564561734213,
best_f1=0.23814541622760801
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.2379506230354309
step: 10, loss: 0.45694729685783386
step: 20, loss: 0.46228885650634766
step: 30, loss: 0.5348737835884094
step: 40, loss: 0.12381000816822052
step: 50, loss: 0.45130106806755066
step: 60, loss: 0.5584824085235596
step: 70, loss: 0.3721004128456116
step: 80, loss: 0.1702057421207428
step: 90, loss: 0.45268258452415466
step: 100, loss: 0.32235440611839294
step: 110, loss: 0.37733983993530273
step: 120, loss: 0.3715292513370514
step: 130, loss: 0.19262515008449554
step: 140, loss: 0.12275274842977524
step: 150, loss: 0.30880534648895264
step: 160, loss: 0.38340285420417786
step: 170, loss: 0.30660945177078247

```



```

/usr/local/lib/python3.7/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=0.1941564561734213, f1=0.1941564561734213,
best_f1=0.23814541622760801
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.46944504976272583
step: 10, loss: 0.5199652314186096
step: 20, loss: 0.24454271793365479
step: 30, loss: 0.3177546262741089
step: 40, loss: 0.24150945246219635
step: 50, loss: 0.3874492943286896
step: 60, loss: 0.38419806957244873
step: 70, loss: 0.18955938518047333
step: 80, loss: 0.31028884649276733
step: 90, loss: 0.5762788653373718
step: 100, loss: 0.25511854887008667
step: 110, loss: 0.5150011777877808
step: 120, loss: 0.32464152574539185
step: 130, loss: 0.31521978974342346
step: 140, loss: 0.5103778839111328
step: 150, loss: 0.2455732673406601
step: 160, loss: 0.17339982092380524
step: 170, loss: 0.45518940687179565
/usr/local/lib/python3.7/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=0.1941564561734213, f1=0.1941564561734213,
best_f1=0.23814541622760801
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.3797317445278168
step: 10, loss: 0.5951831936836243
step: 20, loss: 0.38609758019447327
step: 30, loss: 0.18192936480045319
step: 40, loss: 0.37623530626296997
step: 50, loss: 0.25748634338378906
step: 60, loss: 0.3784882426261902
step: 70, loss: 0.3056439459323883
step: 80, loss: 0.1691942662000656
step: 90, loss: 0.45615389943122864
step: 100, loss: 0.505968451499939
step: 110, loss: 0.3145592510700226
step: 120, loss: 0.4425269663333893
step: 130, loss: 0.3141370713710785

```

```

step: 140, loss: 0.37432143092155457
step: 150, loss: 0.3696785271167755
step: 160, loss: 0.26204344630241394
step: 170, loss: 0.4262712299823761
/usr/local/lib/python3.7/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=0.1941564561734213, f1=0.1941564561734213,
best_f1=0.23814541622760801
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.4414629340171814
step: 10, loss: 0.3164246678352356
step: 20, loss: 0.24695025384426117
step: 30, loss: 0.6404058933258057
step: 40, loss: 0.37901148200035095
step: 50, loss: 0.4469989538192749
step: 60, loss: 0.5657257437705994
step: 70, loss: 0.2637825906276703
step: 80, loss: 0.2564784288406372
step: 90, loss: 0.2432769536972046
step: 100, loss: 0.2440928965806961
step: 110, loss: 0.31510108709335327
step: 120, loss: 0.30644553899765015
step: 130, loss: 0.17600540816783905
step: 140, loss: 0.2359827160835266
step: 150, loss: 0.3118997812271118
step: 160, loss: 0.2378024309873581
step: 170, loss: 0.5119420886039734
/usr/local/lib/python3.7/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=0.1941564561734213, f1=0.1941564561734213,
best_f1=0.23814541622760801
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.377928227186203
step: 10, loss: 0.30820509791374207
step: 20, loss: 0.1662551313638687
step: 30, loss: 0.5170382261276245
step: 40, loss: 0.23943538963794708
step: 50, loss: 0.23921331763267517
step: 60, loss: 0.3642449378967285
step: 70, loss: 0.3301851749420166
step: 80, loss: 0.187226340174675
step: 90, loss: 0.37177208065986633

```

```

step: 100, loss: 0.5050495862960815
step: 110, loss: 0.37545180320739746
step: 120, loss: 0.37578368186950684
step: 130, loss: 0.37798723578453064
step: 140, loss: 0.506903886795044
step: 150, loss: 0.4390481114387512
step: 160, loss: 0.24898476898670197
step: 170, loss: 0.18357622623443604
/usr/local/lib/python3.7/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=0.1941564561734213, f1=0.1941564561734213,
best_f1=0.23814541622760801
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.3846483528614044
step: 10, loss: 0.3179028630256653
step: 20, loss: 0.2463928908109665
step: 30, loss: 0.30856892466545105
step: 40, loss: 0.31138285994529724
step: 50, loss: 0.3073675334453583
step: 60, loss: 0.3119404911994934
step: 70, loss: 0.17576274275779724
step: 80, loss: 0.2346687763929367
step: 90, loss: 0.2895848751068115
step: 100, loss: 0.2599300742149353
step: 110, loss: 0.44523489475250244
step: 120, loss: 0.3130645751953125
step: 130, loss: 0.3740246891975403
step: 140, loss: 0.30112215876579285
step: 150, loss: 0.2862653434276581
step: 160, loss: 0.7669490575790405
step: 170, loss: 0.37412089109420776
/usr/local/lib/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=0.26315789473684215, f1=0.25, best_f1=0.25
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.3340521454811096
step: 10, loss: 0.220183864235878
step: 20, loss: 0.20277050137519836

```

```

step: 30, loss: 0.4042738974094391
step: 40, loss: 0.40386316180229187
step: 50, loss: 0.20588071644306183
step: 60, loss: 0.2064208984375
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 8192.0
step: 70, loss: 0.3707588315010071
step: 80, loss: 0.1837206929922104
step: 90, loss: 0.2551865577697754
step: 100, loss: 0.19398687779903412
step: 110, loss: 0.4151850640773773
step: 120, loss: 0.2193908393383026
step: 130, loss: 0.2223810851573944
step: 140, loss: 0.4917833209037781
step: 150, loss: 0.3159846067428589
step: 160, loss: 0.31010109186172485
step: 170, loss: 0.29665496945381165
/usr/local/lib/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=0.35491606714628293, f1=0.3270718232044199,
best_f1=0.3270718232044199
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.34222808480262756
step: 10, loss: 0.19693240523338318
step: 20, loss: 0.2968979477882385
step: 30, loss: 0.3656150698661804
step: 40, loss: 0.2769557237625122
step: 50, loss: 0.14439575374126434
step: 60, loss: 0.5129988193511963
step: 70, loss: 0.2033102661371231
step: 80, loss: 0.12573906779289246
step: 90, loss: 0.22990663349628448
step: 100, loss: 0.36983564496040344
step: 110, loss: 0.1635352522134781
step: 120, loss: 0.20766359567642212
step: 130, loss: 0.25195884704589844
step: 140, loss: 0.2898542582988739
step: 150, loss: 0.26189061999320984
step: 160, loss: 0.296896368265152
step: 170, loss: 0.20318573713302612
/usr/local/lib/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=0.6203703703703705, f1=0.5766590389016018,
best_f1=0.5766590389016018
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.044707316905260086
step: 10, loss: 0.2615838050842285
step: 20, loss: 0.18870510160923004
step: 30, loss: 0.4776771366596222
step: 40, loss: 0.0842880979180336
step: 50, loss: 0.06618955731391907
step: 60, loss: 0.1655178964138031
step: 70, loss: 0.1805242896080017
step: 80, loss: 0.06301921606063843
step: 90, loss: 0.15423960983753204
step: 100, loss: 0.37700939178466797
step: 110, loss: 0.22283610701560974
step: 120, loss: 0.17163115739822388
step: 130, loss: 0.2045041024684906
step: 140, loss: 0.09434111416339874
step: 150, loss: 0.0794941857457161
step: 160, loss: 0.07934844493865967
step: 170, loss: 0.1596963107585907
/usr/local/lib/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=0.6559633027522935, f1=0.6272727272727273,
best_f1=0.6272727272727273
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, 236.37it/s]
load_f1 = 0.6535626535626535
real_f1 = 0.6437346437346438
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, 144.04it/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: 0.5938824415206909

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

step: 10, loss: 0.4344732463359833

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

step: 20, loss: 0.5291913747787476

step: 30, loss: 0.3597757816314697

step: 40, loss: 0.34163030982017517

step: 50, loss: 0.5929827690124512

```

step: 60, loss: 0.4895910322666168
step: 70, loss: 0.10770280659198761
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 8192.0
step: 80, loss: 1.4392887353897095
step: 90, loss: 0.2676054835319519
step: 100, loss: 0.20344911515712738
step: 110, loss: 0.2753446698188782
step: 120, loss: 0.23964405059814453
step: 130, loss: 0.03420074284076691
step: 140, loss: 0.03319314867258072
step: 150, loss: 0.1698465347290039
step: 160, loss: 0.30529266595840454
step: 170, loss: 0.16950112581253052
step: 180, loss: 0.21108286082744598
step: 190, loss: 0.06678152084350586
step: 200, loss: 0.1004483550786972
step: 210, loss: 0.13208593428134918
step: 220, loss: 0.05433306843042374
step: 230, loss: 0.004899165127426386
/usr/local/lib/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=0.9542920847268673, f1=0.9635535307517085,
best_f1=0.9635535307517085
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.002245693700388074
step: 10, loss: 0.12693218886852264
step: 20, loss: 0.0356370285153389
step: 30, loss: 0.019969012588262558
step: 40, loss: 0.015283117070794106
step: 50, loss: 0.03440125286579132
step: 60, loss: 0.007714845240116119
step: 70, loss: 0.023798998445272446
step: 80, loss: 0.03721480071544647
step: 90, loss: 0.17699825763702393
step: 100, loss: 0.015402919612824917
step: 110, loss: 0.10658532381057739
step: 120, loss: 0.04078937694430351
step: 130, loss: 0.036639969795942307
step: 140, loss: 0.007676809094846249
step: 150, loss: 0.25826171040534973
step: 160, loss: 0.03473823517560959

```



```

step: 170, loss: 0.0010106794070452452
step: 180, loss: 0.008487689308822155
step: 190, loss: 0.03470682352781296
step: 200, loss: 0.03568645194172859
step: 210, loss: 0.04780500382184982
step: 220, loss: 0.008497631177306175
step: 230, loss: 0.006327452138066292
/usr/local/lib/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=0.9611542730299667, f1=0.961625282167043,
best_f1=0.961625282167043
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.015604588203132153
step: 10, loss: 0.009812377393245697
step: 20, loss: 0.15119704604148865
step: 30, loss: 0.004912353120744228
step: 40, loss: 0.007398274261504412
step: 50, loss: 0.027637140825390816
step: 60, loss: 0.0602424293756485
step: 70, loss: 0.010767459869384766
step: 80, loss: 0.16118605434894562
step: 90, loss: 0.008989744819700718
step: 100, loss: 0.055976491421461105
step: 110, loss: 0.009317193180322647
step: 120, loss: 0.009042806923389435
step: 130, loss: 0.0038800158072263002
step: 140, loss: 0.0025244764983654022
step: 150, loss: 0.11706563085317612
step: 160, loss: 0.019334709271788597
step: 170, loss: 0.005602532997727394
step: 180, loss: 0.021311815828084946
step: 190, loss: 0.005309188272804022
step: 200, loss: 0.031281549483537674
step: 210, loss: 0.02923642471432686
step: 220, loss: 0.11814562231302261
step: 230, loss: 0.022983616217970848
/usr/local/lib/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=0.9752808988764046, f1=0.9624573378839592,
best_f1=0.9624573378839592
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.014348003081977367
step: 10, loss: 0.0022349797654896975
step: 20, loss: 0.0018296042690053582
step: 30, loss: 0.0021317084319889545
step: 40, loss: 0.0336337648332119
step: 50, loss: 0.021363873034715652
step: 60, loss: 0.0027833234053105116
step: 70, loss: 0.06051080301403999
step: 80, loss: 0.13905811309814453
step: 90, loss: 0.21652060747146606
step: 100, loss: 0.057392969727516174
step: 110, loss: 0.0010714296950027347
step: 120, loss: 0.01459306851029396
step: 130, loss: 0.01989440619945526
step: 140, loss: 0.004813509993255138
step: 150, loss: 0.007896467112004757
step: 160, loss: 0.002238625194877386
step: 170, loss: 0.0019910139963030815
step: 180, loss: 0.09023474901914597
step: 190, loss: 0.0011880146339535713
step: 200, loss: 0.13543355464935303
step: 210, loss: 0.006040820851922035
step: 220, loss: 0.0025342742446810007
step: 230, loss: 0.0034970082342624664
/usr/local/lib/python3.7/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=0.9751131221719457, f1=0.9727272727272728,
best_f1=0.9624573378839592
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.005408337339758873
step: 10, loss: 0.0015762008260935545
step: 20, loss: 0.002114564646035433
step: 30, loss: 0.00397872319445014
step: 40, loss: 0.002192205749452114
step: 50, loss: 0.0014171323273330927
step: 60, loss: 0.005165181588381529
step: 70, loss: 0.007668002042919397
step: 80, loss: 0.013527180068194866

```

```

step: 90, loss: 0.019780956208705902
step: 100, loss: 0.004697431344538927
step: 110, loss: 0.0017702047480270267
step: 120, loss: 0.0006836746470071375
step: 130, loss: 0.0028682525735348463
step: 140, loss: 0.003549022600054741
step: 150, loss: 0.010519033297896385
step: 160, loss: 0.0014448069268837571
step: 170, loss: 0.06127874553203583
step: 180, loss: 0.0027902342844754457
step: 190, loss: 0.03431466221809387
step: 200, loss: 0.01706733927130699
step: 210, loss: 0.0026192613877356052
step: 220, loss: 0.03468545153737068
step: 230, loss: 0.010512310080230236
/usr/local/lib/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=0.9755555555555556, f1=0.9753363228699552,
best_f1=0.9753363228699552
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0005810215952806175
step: 10, loss: 0.001228701788932085
step: 20, loss: 0.0010871124686673284
step: 30, loss: 0.0005371551378630102
step: 40, loss: 0.00020156164828222245
step: 50, loss: 0.0003360146947670728
step: 60, loss: 0.012544243596494198
step: 70, loss: 0.04513880982995033
step: 80, loss: 0.000457750225905329
step: 90, loss: 0.035912930965423584
step: 100, loss: 0.00036008298047818244
step: 110, loss: 0.010859716683626175
step: 120, loss: 0.0009840615093708038
step: 130, loss: 0.0024912827648222446
step: 140, loss: 0.00149907183367759
step: 150, loss: 0.00020082674745935947
step: 160, loss: 0.0013273241929709911
step: 170, loss: 0.018406735733151436
step: 180, loss: 0.0005720107001252472
step: 190, loss: 0.0017898220103234053
step: 200, loss: 0.0017371535068377852

```

```

step: 210, loss: 0.0008916400838643312
step: 220, loss: 0.011194156482815742
step: 230, loss: 0.0007877881871536374
/usr/local/lib/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=0.9820627802690582, f1=0.9818594104308391,
best_f1=0.9818594104308391
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0011405491968616843
step: 10, loss: 0.0005382021772675216
step: 20, loss: 0.00022652087500318885
step: 30, loss: 0.0003339348768349737
step: 40, loss: 0.0015395993832498789
step: 50, loss: 0.0015834345249459147
step: 60, loss: 0.0006589479162357748
step: 70, loss: 0.0012703175889328122
step: 80, loss: 0.004100162070244551
step: 90, loss: 0.0009834419470280409
step: 100, loss: 0.001166854053735733
step: 110, loss: 0.0008968852926045656
step: 120, loss: 0.0021537449210882187
step: 130, loss: 0.0055900514125823975
step: 140, loss: 0.0006669766735285521
step: 150, loss: 0.035457100719213486
step: 160, loss: 0.0016224257415160537
step: 170, loss: 0.00042198013397865
step: 180, loss: 0.0003943217743653804
step: 190, loss: 0.0002724641526583582
step: 200, loss: 0.0005502633284777403
step: 210, loss: 0.003246542764827609
step: 220, loss: 0.00037354230880737305
step: 230, loss: 0.0009888748172670603
/usr/local/lib/python3.7/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=0.9778270509977827, f1=0.9753914988814317,
best_f1=0.9818594104308391
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00021744809055235237

```

```

step: 10, loss: 0.0010546285193413496
step: 20, loss: 0.0013599650701507926
step: 30, loss: 0.15195679664611816
step: 40, loss: 0.01786908693611622
step: 50, loss: 0.015297586098313332
step: 60, loss: 0.0019456250593066216
step: 70, loss: 0.0005769685376435518
step: 80, loss: 0.12029172480106354
step: 90, loss: 0.0009158574976027012
step: 100, loss: 0.00029035855550318956
step: 110, loss: 0.001718413201160729
step: 120, loss: 0.0023605271708220243
step: 130, loss: 0.001087366952560842
step: 140, loss: 0.00020536633383017033
step: 150, loss: 0.10987379401922226
step: 160, loss: 0.00507896114140749
step: 170, loss: 0.0514310821890831
step: 180, loss: 0.01723391003906727
step: 190, loss: 0.00748762022703886
step: 200, loss: 0.011952051892876625
step: 210, loss: 0.0018295121844857931
step: 220, loss: 0.0009113281266763806
step: 230, loss: 0.0009805161971598864
/usr/local/lib/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=0.983050847457627, f1=0.9737742303306728,
best_f1=0.9737742303306728
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.005492418073117733
step: 10, loss: 0.0028648825827986
step: 20, loss: 0.00031897996086627245
step: 30, loss: 0.0007523577660322189
step: 40, loss: 0.000310759583953768
step: 50, loss: 0.00047944337711669505
step: 60, loss: 0.00011130112397950143
step: 70, loss: 0.027080891653895378
step: 80, loss: 0.0014067280571907759
step: 90, loss: 0.013270996510982513
step: 100, loss: 0.0004923649248667061
step: 110, loss: 0.000388192042009905
step: 120, loss: 0.014652473852038383

```

```

step: 130, loss: 0.0008035200880840421
step: 140, loss: 0.0002615304256323725
step: 150, loss: 0.00020484364358708262
step: 160, loss: 0.000159167087986134
step: 170, loss: 0.0002028070157393813
step: 180, loss: 0.0004575435013975948
step: 190, loss: 0.0010985435219481587
step: 200, loss: 0.00023947993759065866
step: 210, loss: 0.00047921083751134574
step: 220, loss: 0.0011710900580510497
step: 230, loss: 0.00021309591829776764
/usr/local/lib/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=0.9831649831649831, f1=0.9750566893424036,
best_f1=0.9750566893424036
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0003138341708108783
step: 10, loss: 0.0010961273219436407
step: 20, loss: 0.00030654590227641165
step: 30, loss: 0.00010156352072954178
step: 40, loss: 0.0007349670631811023
step: 50, loss: 0.00012540302122943103
step: 60, loss: 0.0001591600594110787
step: 70, loss: 0.00112350529525429
step: 80, loss: 0.00013804312038701028
step: 90, loss: 0.000148212187923491
step: 100, loss: 0.001646299147978425
step: 110, loss: 0.0012738811783492565
step: 120, loss: 0.0008397971978411078
step: 130, loss: 0.0003490816743578762
step: 140, loss: 0.0007347817881964147
step: 150, loss: 0.00034990248968824744
step: 160, loss: 6.341841071844101e-05
step: 170, loss: 0.00040551283746026456
step: 180, loss: 0.00043851183727383614
step: 190, loss: 0.0001376967557007447
step: 200, loss: 0.0003306887811049819
step: 210, loss: 0.061558809131383896
step: 220, loss: 0.018229439854621887
step: 230, loss: 0.00010615790961310267
/usr/local/lib/python3.7/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=0.9819819819819819, f1=0.979591836734694,
best_f1=0.9750566893424036
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00010282474977429956
step: 10, loss: 0.00022042309865355492
step: 20, loss: 0.00011931249173358083
step: 30, loss: 0.00026072515174746513
step: 40, loss: 4.8071950004668906e-05
step: 50, loss: 7.803809421602637e-05
step: 60, loss: 0.003248602384701371
step: 70, loss: 0.00027242532814852893
step: 80, loss: 0.001438664272427559
step: 90, loss: 0.01568811573088169
step: 100, loss: 0.0001316911802859977
step: 110, loss: 0.002383829327300191
step: 120, loss: 0.00011038028605980799
step: 130, loss: 0.00012826977763324976
step: 140, loss: 0.0026641525328159332
step: 150, loss: 0.0016867013182491064
step: 160, loss: 0.0009056479902938008
step: 170, loss: 0.0004449528641998768
step: 180, loss: 0.0010195012437179685
step: 190, loss: 0.00012887058255728334
step: 200, loss: 0.003366270335391164
step: 210, loss: 0.00013989288709126413
step: 220, loss: 0.0005034170462749898
step: 230, loss: 9.4149487267714e-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=0.9832026875699889, 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: 0.00038362087798304856
step: 10, loss: 0.00012160736514488235
step: 20, loss: 0.012173991650342941
step: 30, loss: 0.033772360533475876
step: 40, loss: 0.00015747110592201352

```

```

step: 50, loss: 0.0011008454021066427
step: 60, loss: 0.0003281717945355922
step: 70, loss: 0.00022071183775551617
step: 80, loss: 0.00010869279503822327
step: 90, loss: 0.003935604821890593
step: 100, loss: 7.89165060268715e-05
step: 110, loss: 7.392139377770945e-05
step: 120, loss: 0.00022278731921687722
step: 130, loss: 0.0003722895635291934
step: 140, loss: 9.689209400676191e-05
step: 150, loss: 0.00010032518184743822
step: 160, loss: 0.00012069158401573077
step: 170, loss: 0.00010923810623353347
step: 180, loss: 0.0001580766256665811
step: 190, loss: 0.0005660004098899662
step: 200, loss: 9.05152628547512e-05
step: 210, loss: 0.00012187920947326347
step: 220, loss: 0.013294676318764687
step: 230, loss: 0.00030917645199224353
/usr/local/lib/python3.7/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=0.9820224719101124, f1=0.9786276715410572,
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: 0.00021453876979649067
step: 10, loss: 0.00012122245971113443
step: 20, loss: 0.00010363764886278659
step: 30, loss: 0.00013207487063482404
step: 40, loss: 0.00013533436867874116
step: 50, loss: 0.00024193670833483338
step: 60, loss: 8.773730223765597e-05
step: 70, loss: 0.00012439244892448187
step: 80, loss: 0.0015393901849165559
step: 90, loss: 0.00015344239363912493
step: 100, loss: 0.00011764196824515238
step: 110, loss: 0.0004009093972854316
step: 120, loss: 0.00011122879368485883
step: 130, loss: 0.00010747146734502167
step: 140, loss: 0.00010668267350411043
step: 150, loss: 0.00038954155752435327
step: 160, loss: 0.002218734472990036
step: 170, loss: 8.266131771961227e-05
step: 180, loss: 0.005815069656819105
step: 190, loss: 0.00010118507634615526
step: 200, loss: 4.7987039579311386e-05

```



```

step: 210, loss: 0.00017224568000528961
step: 220, loss: 0.0001507151173427701
step: 230, loss: 0.00014124641893431544
/usr/local/lib/python3.7/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=0.9820224719101124, f1=0.9751693002257337,
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: 0.00011904740676982328
step: 10, loss: 6.86475177644752e-05
step: 20, loss: 9.787904855329543e-05
step: 30, loss: 0.0003205237735528499
step: 40, loss: 0.00027934799436479807
step: 50, loss: 6.968862726353109e-05
step: 60, loss: 0.0003785028529819101
step: 70, loss: 0.000193051018868573
step: 80, loss: 0.00026253523537889123
step: 90, loss: 0.0012409326154738665
step: 100, loss: 0.00018128554802387953
step: 110, loss: 0.0001008811013889499
step: 120, loss: 4.600878310156986e-05
step: 130, loss: 0.000998022616840899
step: 140, loss: 0.00010726352775236592
step: 150, loss: 5.364017852116376e-05
step: 160, loss: 0.0002214660489698872
step: 170, loss: 0.00016946540563367307
step: 180, loss: 0.00011109936895081773
step: 190, loss: 7.567612192360684e-05
step: 200, loss: 0.00014260775060392916
step: 210, loss: 7.263618317665532e-05
step: 220, loss: 0.00017073970229830593
step: 230, loss: 0.0001987374562304467
/usr/local/lib/python3.7/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=0.9821029082774049, 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: 0.00012729925219900906
step: 10, loss: 0.00019114650785923004
step: 20, loss: 0.0012809338513761759
step: 30, loss: 9.592855349183083e-05
step: 40, loss: 0.00017861739615909755

```

```

step: 50, loss: 7.766341150272638e-05
step: 60, loss: 0.009548150934278965
step: 70, loss: 0.00018352035840507597
step: 80, loss: 0.00020272917754482478
step: 90, loss: 5.7757028116611764e-05
step: 100, loss: 4.963627361576073e-05
step: 110, loss: 0.0001605569414095953
step: 120, loss: 0.03259432688355446
step: 130, loss: 0.00031594003667123616
step: 140, loss: 0.015560006722807884
step: 150, loss: 0.0001630835613468662
step: 160, loss: 0.014262950979173183
step: 170, loss: 4.464119047042914e-05
step: 180, loss: 8.412451279582456e-05
step: 190, loss: 0.00012561213225126266
step: 200, loss: 0.0002634548000060022
step: 210, loss: 0.006528892088681459
step: 220, loss: 6.207792466739193e-05
step: 230, loss: 8.046008588280529e-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=0.9821428571428571, f1=0.9743016759776536,
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 [00:15, 155.67it/s]
load_f1 = 0.9831271091113611
real_f1 = 0.9808773903262092
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, 142.98it/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              : 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: 0.6631510257720947
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: 0.47177523374557495
step: 20, loss: 0.3426883816719055
step: 30, loss: 0.3258879780769348
step: 40, loss: 0.400151789188385
step: 50, loss: 0.7301903963088989
step: 60, loss: 0.3551722466945648
step: 70, loss: 0.42770692706108093
step: 80, loss: 0.5230494737625122
step: 90, loss: 0.42419689893722534
step: 100, loss: 0.5550311207771301
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: 110, loss: 0.17287564277648926
step: 120, loss: 0.41307204961776733
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 2048.0
step: 130, loss: 0.38983041048049927
step: 140, loss: 0.19072894752025604
step: 150, loss: 0.07777298986911774
step: 160, loss: 0.2191697359085083
step: 170, loss: 0.10708991438150406
step: 180, loss: 0.08277495950460434

```

```

step: 190, loss: 0.11903510242700577
step: 200, loss: 0.0673302561044693
step: 210, loss: 0.051074933260679245
step: 220, loss: 0.0993175134062767
step: 230, loss: 0.23800188302993774
step: 240, loss: 0.07323738932609558
step: 250, loss: 0.03546852990984917
step: 260, loss: 0.2488759458065033
step: 270, loss: 0.43791094422340393
step: 280, loss: 0.06684650480747223
step: 290, loss: 0.04484562948346138
step: 300, loss: 0.07493212819099426
step: 310, loss: 0.06775878369808197
step: 320, loss: 0.026020344346761703
step: 330, loss: 0.06580441445112228
step: 340, loss: 0.23236718773841858
step: 350, loss: 0.05738032981753349
step: 360, loss: 0.08351290971040726
step: 370, loss: 0.11178679019212723
step: 380, loss: 0.08410966396331787
step: 390, loss: 0.027698833495378494
step: 400, loss: 0.09989239275455475
step: 410, loss: 0.23884496092796326
step: 420, loss: 0.03703567013144493
step: 430, loss: 0.015448938123881817
step: 440, loss: 0.0386492982506752
step: 450, loss: 0.10136707872152328
step: 460, loss: 0.03800095245242119
step: 470, loss: 0.03865994140505791
step: 480, loss: 0.2167428731918335
step: 490, loss: 0.08855915814638138
step: 500, loss: 0.013617843389511108
step: 510, loss: 0.02724895253777504
step: 520, loss: 0.15517203509807587
step: 530, loss: 0.2753509283065796
/usr/local/lib/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=0.9329529243937231, f1=0.9342230695900858,
best_f1=0.9342230695900858
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.11741320043802261

```

step: 10, loss: 0.037177059799432755  
step: 20, loss: 0.030042795464396477  
step: 30, loss: 0.037504490464925766  
step: 40, loss: 0.01968027465045452  
step: 50, loss: 0.031043076887726784  
step: 60, loss: 0.008911840617656708  
step: 70, loss: 0.03694292530417442  
step: 80, loss: 0.0215435940772295  
step: 90, loss: 0.013324897736310959  
step: 100, loss: 0.172823965549469  
step: 110, loss: 0.007198247127234936  
step: 120, loss: 0.020577656105160713  
step: 130, loss: 0.0042045218870043755  
step: 140, loss: 0.042688049376010895  
step: 150, loss: 0.0223309975117445  
step: 160, loss: 0.08446191251277924  
step: 170, loss: 0.01998022384941578  
step: 180, loss: 0.06287482380867004  
step: 190, loss: 0.017250949516892433  
step: 200, loss: 0.17706921696662903  
step: 210, loss: 0.03353685513138771  
step: 220, loss: 0.0019856984727084637  
step: 230, loss: 0.020695405080914497  
step: 240, loss: 0.14342951774597168  
step: 250, loss: 0.039421163499355316  
step: 260, loss: 0.04823979735374451  
step: 270, loss: 0.00937049649655819  
step: 280, loss: 0.025321755558252335  
step: 290, loss: 0.049613550305366516  
step: 300, loss: 0.04987654462456703  
step: 310, loss: 0.10378684848546982  
step: 320, loss: 0.03331734612584114  
step: 330, loss: 0.06808716058731079  
step: 340, loss: 0.02636868879199028  
step: 350, loss: 0.0012826458550989628  
step: 360, loss: 0.09114451706409454  
step: 370, loss: 0.012522518634796143  
step: 380, loss: 0.1145838126540184  
step: 390, loss: 0.005202979780733585  
step: 400, loss: 0.03594474866986275  
step: 410, loss: 0.031046373769640923  
step: 420, loss: 0.14367258548736572  
step: 430, loss: 0.22288541495800018  
step: 440, loss: 0.027146486565470695  
step: 450, loss: 0.1520404815673828  
step: 460, loss: 0.025053299963474274  
step: 470, loss: 0.02021789737045765  
step: 480, loss: 0.004809098318219185

```

step: 490, loss: 0.04834045097231865
step: 500, loss: 0.009646335616707802
step: 510, loss: 0.06104521453380585
step: 520, loss: 0.3249950110912323
step: 530, loss: 0.10730385780334473
/usr/local/lib/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=0.942008486562942, f1=0.935831381733021,
best_f1=0.935831381733021
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.09492763876914978
step: 10, loss: 0.06386493146419525
step: 20, loss: 0.005874177906662226
step: 30, loss: 0.02630016952753067
step: 40, loss: 0.0800306499004364
step: 50, loss: 0.01171767245978117
step: 60, loss: 0.003278432646766305
step: 70, loss: 0.006444238126277924
step: 80, loss: 0.012089035473763943
step: 90, loss: 0.0017573523800820112
step: 100, loss: 0.045429155230522156
step: 110, loss: 0.018781842663884163
step: 120, loss: 0.21475189924240112
step: 130, loss: 0.07576549798250198
step: 140, loss: 0.02529771439731121
step: 150, loss: 0.01976662687957287
step: 160, loss: 0.0066005331464111805
step: 170, loss: 0.0029781272169202566
step: 180, loss: 0.009653093293309212
step: 190, loss: 0.004499378614127636
step: 200, loss: 0.03257245197892189
step: 210, loss: 0.041036393493413925
step: 220, loss: 0.11555071920156479
step: 230, loss: 0.013496031984686852
step: 240, loss: 0.0787348821759224
step: 250, loss: 0.16280405223369598
step: 260, loss: 0.0257598664611578
step: 270, loss: 0.03332644701004028
step: 280, loss: 0.005068054422736168
step: 290, loss: 0.027704689651727676
step: 300, loss: 0.1994352787733078

```

```

step: 310, loss: 0.13068503141403198
step: 320, loss: 0.01702943816781044
step: 330, loss: 0.045870814472436905
step: 340, loss: 0.008745224215090275
step: 350, loss: 0.2161153256893158
step: 360, loss: 0.02596789412200451
step: 370, loss: 0.03497958555817604
step: 380, loss: 0.001234946888871491
step: 390, loss: 0.002914954209700227
step: 400, loss: 0.12924158573150635
step: 410, loss: 0.04715823009610176
step: 420, loss: 0.011758486740291119
step: 430, loss: 0.030820196494460106
step: 440, loss: 0.23072078824043274
step: 450, loss: 0.055702753365039825
step: 460, loss: 0.07309554517269135
step: 470, loss: 0.02527722530066967
step: 480, loss: 0.05328521877527237
step: 490, loss: 0.04282884672284126
step: 500, loss: 0.010646277107298374
step: 510, loss: 0.06086812913417816
step: 520, loss: 0.008778432384133339
step: 530, loss: 0.008214782923460007
/usr/local/lib/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=0.9446529080675422, f1=0.9397363465160076,
best_f1=0.9397363465160076
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.013798082247376442
step: 10, loss: 0.0014289660612121224
step: 20, loss: 0.049307312816381454
step: 30, loss: 0.11050814390182495
step: 40, loss: 0.009346251375973225
step: 50, loss: 0.036634042859077454
step: 60, loss: 0.0035497320350259542
step: 70, loss: 0.017578493803739548
step: 80, loss: 0.03214704617857933
step: 90, loss: 0.07489942014217377
step: 100, loss: 0.001239457749761641
step: 110, loss: 0.09689787775278091
step: 120, loss: 0.0013953594025224447

```



```

step: 130, loss: 0.02037951350212097
step: 140, loss: 0.012009136378765106
step: 150, loss: 0.030704624950885773
step: 160, loss: 0.030406642705202103
step: 170, loss: 0.02667299658060074
step: 180, loss: 0.0302715040743351
step: 190, loss: 0.015301985666155815
step: 200, loss: 0.047616105526685715
step: 210, loss: 0.0002764402888715267
step: 220, loss: 0.0026403130032122135
step: 230, loss: 0.002883258508518338
step: 240, loss: 0.007472814992070198
step: 250, loss: 0.11454064399003983
step: 260, loss: 0.009659616276621819
step: 270, loss: 0.03713047504425049
step: 280, loss: 0.0028192722238600254
step: 290, loss: 0.05555151030421257
step: 300, loss: 0.004830810707062483
step: 310, loss: 0.0047620791010558605
step: 320, loss: 0.11295825988054276
step: 330, loss: 0.11485271900892258
step: 340, loss: 0.014272892847657204
step: 350, loss: 0.11374890059232712
step: 360, loss: 0.051107484847307205
step: 370, loss: 0.012627721764147282
step: 380, loss: 0.008094137534499168
step: 390, loss: 0.02112513594329357
step: 400, loss: 0.02775552123785019
step: 410, loss: 0.0005523646832443774
step: 420, loss: 0.0012461802689358592
step: 430, loss: 0.0003174465091433376
step: 440, loss: 0.005062011536210775
step: 450, loss: 0.0066785006783902645
step: 460, loss: 0.0221197921782732
step: 470, loss: 0.0018323991680517793
step: 480, loss: 0.024139977991580963
step: 490, loss: 0.052377909421920776
step: 500, loss: 0.04152441397309303
step: 510, loss: 0.08880618959665298
step: 520, loss: 0.008666587062180042
step: 530, loss: 0.049044832587242126
/usr/local/lib/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=0.9498141263940519, f1=0.9473193473193473,  
best_f1=0.9473193473193473  
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:  
An input tensor was not cuda.  
    warnings.warn("An input tensor was not cuda.")  
step: 0, loss: 0.0013290145434439182  
step: 10, loss: 0.005465212743729353  
step: 20, loss: 0.013340276665985584  
step: 30, loss: 0.027282986789941788  
step: 40, loss: 0.0004135115013923496  
step: 50, loss: 0.005080326460301876  
step: 60, loss: 0.004476367495954037  
step: 70, loss: 0.022427095100283623  
step: 80, loss: 0.003404591465368867  
step: 90, loss: 0.034644853323698044  
step: 100, loss: 0.03629877045750618  
step: 110, loss: 0.002940659411251545  
step: 120, loss: 0.24315473437309265  
step: 130, loss: 0.0023439819924533367  
step: 140, loss: 0.0016066976822912693  
step: 150, loss: 0.021519988775253296  
step: 160, loss: 0.00500071095302701  
step: 170, loss: 0.011984122917056084  
step: 180, loss: 0.003253692528232932  
step: 190, loss: 0.0029588595498353243  
step: 200, loss: 0.003561790566891432  
step: 210, loss: 0.0013655063230544329  
step: 220, loss: 0.0329941026866436  
step: 230, loss: 0.0010917647741734982  
step: 240, loss: 0.017646703869104385  
step: 250, loss: 0.15547069907188416  
step: 260, loss: 0.006685519125312567  
step: 270, loss: 0.0032809912227094173  
step: 280, loss: 0.007153820246458054  
step: 290, loss: 0.048815857619047165  
step: 300, loss: 0.0060916622169315815  
step: 310, loss: 0.06033644452691078  
step: 320, loss: 0.010573655366897583  
step: 330, loss: 0.0018753716722130775  
step: 340, loss: 0.017361054196953773  
step: 350, loss: 0.00045872191549278796  
step: 360, loss: 0.005833501927554607  
step: 370, loss: 0.0010522163938730955  
step: 380, loss: 0.0002270216355100274  
step: 390, loss: 0.0027107109781354666  
step: 400, loss: 0.01065738033503294  
step: 410, loss: 0.16572217643260956  
step: 420, loss: 0.14345499873161316
```

```
step: 430, loss: 0.025042623281478882
step: 440, loss: 0.004541922360658646
step: 450, loss: 0.01554618775844574
step: 460, loss: 0.0024080458097159863
step: 470, loss: 0.03968465328216553
step: 480, loss: 0.0275479294359684
step: 490, loss: 0.0019364075269550085
step: 500, loss: 0.010702281259000301
step: 510, loss: 0.0010413742857053876
step: 520, loss: 0.06918098777532578
step: 530, loss: 0.0030176672153174877
/usr/local/lib/python3.7/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=0.9484440315838365, f1=0.9407372841810545,
best_f1=0.9473193473193473
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.012392089702188969
step: 10, loss: 0.0008790647843852639
step: 20, loss: 0.00047428059042431414
step: 30, loss: 0.0008878656663000584
step: 40, loss: 0.0027225506491959095
step: 50, loss: 0.00022200928651727736
step: 60, loss: 0.0028720484115183353
step: 70, loss: 0.0008533311192877591
step: 80, loss: 0.003666578559204936
step: 90, loss: 0.020606625825166702
step: 100, loss: 0.16398361325263977
step: 110, loss: 0.0044881422072649
step: 120, loss: 0.19584739208221436
step: 130, loss: 0.023192783817648888
step: 140, loss: 0.004151188302785158
step: 150, loss: 0.00117881887126714
step: 160, loss: 0.00912188645452261
step: 170, loss: 0.0001632446510484442
step: 180, loss: 0.0009463286842219532
step: 190, loss: 0.12737220525741577
step: 200, loss: 0.005594970658421516
step: 210, loss: 0.014945878647267818
step: 220, loss: 0.025173043832182884
step: 230, loss: 0.008362723514437675
step: 240, loss: 0.0024500482250005007
step: 250, loss: 0.0245102159678936
step: 260, loss: 0.0009090399253182113
step: 270, loss: 0.0022692016791552305
step: 280, loss: 0.001759371254593134
```

```

step: 290, loss: 0.01705995202064514
step: 300, loss: 0.018259292468428612
step: 310, loss: 0.029677700251340866
step: 320, loss: 4.0957678720587865e-05
step: 330, loss: 0.016056755557656288
step: 340, loss: 0.0010372198885306716
step: 350, loss: 0.0014993326039984822
step: 360, loss: 0.010933885350823402
step: 370, loss: 0.008065136149525642
step: 380, loss: 0.002893275348469615
step: 390, loss: 0.001494131051003933
step: 400, loss: 0.021758753806352615
step: 410, loss: 0.0004196331719867885
step: 420, loss: 0.012921527028083801
step: 430, loss: 0.0008424405241385102
step: 440, loss: 0.008473072201013565
step: 450, loss: 0.13727985322475433
step: 460, loss: 0.0027804269921034575
step: 470, loss: 0.0026715544518083334
step: 480, loss: 0.0037504108622670174
step: 490, loss: 0.0015587739180773497
step: 500, loss: 0.00020651589147746563
step: 510, loss: 0.2644132673740387
step: 520, loss: 0.0023704422637820244
step: 530, loss: 0.0019154378678649664
/usr/local/lib/python3.7/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=0.94200187090739, f1=0.9354686764013189,
best_f1=0.9473193473193473
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0037738101091235876
step: 10, loss: 0.001670857542194426
step: 20, loss: 0.006850619800388813
step: 30, loss: 0.011266738176345825
step: 40, loss: 0.006375796627253294
step: 50, loss: 0.0021202918142080307
step: 60, loss: 0.08890272676944733
step: 70, loss: 0.00037468387745320797
step: 80, loss: 0.00034363195300102234
step: 90, loss: 0.0005162478191778064
step: 100, loss: 0.000881291285622865
step: 110, loss: 6.685828702757135e-05
step: 120, loss: 0.020940368995070457
step: 130, loss: 0.0006322991685010493
step: 140, loss: 0.00016880307521205395

```

```

step: 150, loss: 0.02977008745074272
step: 160, loss: 0.00029123705462552607
step: 170, loss: 0.01390268374234438
step: 180, loss: 0.018306532874703407
step: 190, loss: 0.002590804360806942
step: 200, loss: 0.0044114734046161175
step: 210, loss: 0.012627040967345238
step: 220, loss: 0.0003471104719210416
step: 230, loss: 0.000660971796605736
step: 240, loss: 0.00123393046669662
step: 250, loss: 0.0525011382997036
step: 260, loss: 0.0023890738375484943
step: 270, loss: 0.0005840189987793565
step: 280, loss: 0.027425602078437805
step: 290, loss: 0.0014800478238612413
step: 300, loss: 0.00016370137745980173
step: 310, loss: 0.001657145912759006
step: 320, loss: 0.0008810590370558202
step: 330, loss: 0.0031597046181559563
step: 340, loss: 0.003798386547714472
step: 350, loss: 0.0006445776089094579
step: 360, loss: 0.0014434515032917261
step: 370, loss: 0.0059719872660934925
step: 380, loss: 0.0006443258607760072
step: 390, loss: 0.017080310732126236
step: 400, loss: 0.07625246047973633
step: 410, loss: 0.00013957677583675832
step: 420, loss: 0.011643619276583195
step: 430, loss: 0.000516659754794091
step: 440, loss: 0.005214490462094545
step: 450, loss: 0.0019645385909825563
step: 460, loss: 0.033829204738140106
step: 470, loss: 0.13054367899894714
step: 480, loss: 0.007297556381672621
step: 490, loss: 0.1038021370768547
step: 500, loss: 0.0030694182496517897
step: 510, loss: 0.00022378764697350562
step: 520, loss: 0.0035071575548499823
step: 530, loss: 0.0017552707577124238
/usr/local/lib/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=0.9521597770552717, f1=0.9464867380176826,
best_f1=0.9464867380176826

```

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

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

```
step: 0, loss: 0.0009342527482658625  
step: 10, loss: 0.0033070382196456194  
step: 20, loss: 0.0005260029574856162  
step: 30, loss: 0.0005950990598648787  
step: 40, loss: 0.00063158362172544  
step: 50, loss: 0.0013482669601216912  
step: 60, loss: 0.0009306251304224133  
step: 70, loss: 0.0004437978786882013  
step: 80, loss: 0.04592753201723099  
step: 90, loss: 0.00046504929196089506  
step: 100, loss: 0.0003863434540107846  
step: 110, loss: 0.0004118476063013077  
step: 120, loss: 0.0014418604550883174  
step: 130, loss: 0.0001427927054464817  
step: 140, loss: 0.0031610336154699326  
step: 150, loss: 0.00019511398568283767  
step: 160, loss: 0.0001262393780052662  
step: 170, loss: 0.0015734938206151128  
step: 180, loss: 0.0002871699689421803  
step: 190, loss: 0.01123350765556097  
step: 200, loss: 0.0008525022421963513  
step: 210, loss: 0.09746523201465607  
step: 220, loss: 0.00021685297542717308  
step: 230, loss: 0.06529227644205093  
step: 240, loss: 0.0016446999507024884  
step: 250, loss: 0.0010039351182058454  
step: 260, loss: 0.0002549363998696208  
step: 270, loss: 0.006911277770996094  
step: 280, loss: 0.0007572845788672566  
step: 290, loss: 0.0016101045766845345  
step: 300, loss: 0.00034924442297779024  
step: 310, loss: 0.0013306026812642813  
step: 320, loss: 0.0010472690919414163  
step: 330, loss: 0.0013676612870767713  
step: 340, loss: 0.09953443706035614  
step: 350, loss: 0.0004193224594928324  
step: 360, loss: 0.028177538886666298  
step: 370, loss: 0.002031471114605665  
step: 380, loss: 7.237166573759168e-05  
step: 390, loss: 0.0017144945450127125  
step: 400, loss: 0.0016251642955467105  
step: 410, loss: 0.0006582658970728517  
step: 420, loss: 0.00020020786905661225  
step: 430, loss: 0.001270404551178217  
step: 440, loss: 0.003434128360822797
```

```

step: 450, loss: 0.000942566548474133
step: 460, loss: 0.012674743309617043
step: 470, loss: 0.1585671454668045
step: 480, loss: 0.019380580633878708
step: 490, loss: 0.023488158360123634
step: 500, loss: 0.0013240569969639182
step: 510, loss: 0.004268900025635958
step: 520, loss: 0.0002647778019309044
step: 530, loss: 0.002416247734799981
/usr/local/lib/python3.7/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=0.9451360073766714, f1=0.9416126042632066,
best_f1=0.9464867380176826
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00029330188408493996
step: 10, loss: 0.0028728158213198185
step: 20, loss: 0.000138981340569444
step: 30, loss: 0.09799385070800781
step: 40, loss: 0.0012810751795768738
step: 50, loss: 0.0001866332459030673
step: 60, loss: 0.0006658528582192957
step: 70, loss: 0.0024769166484475136
step: 80, loss: 0.0054280818440020084
step: 90, loss: 0.04577948898077011
step: 100, loss: 0.00019327946938574314
step: 110, loss: 0.03247964754700661
step: 120, loss: 0.019028740003705025
step: 130, loss: 0.04394541680812836
step: 140, loss: 0.0019278832478448749
step: 150, loss: 0.024595508351922035
step: 160, loss: 0.0002193194377468899
step: 170, loss: 0.0015647285617887974
step: 180, loss: 0.0002543670707382262
step: 190, loss: 9.711356688058004e-05
step: 200, loss: 0.011477278545498848
step: 210, loss: 0.0004221297276671976
step: 220, loss: 0.0013038819888606668
step: 230, loss: 0.001984015805646777
step: 240, loss: 0.0013548951828852296
step: 250, loss: 0.0017511994810774922
step: 260, loss: 0.0017146007157862186
step: 270, loss: 0.0017759122420102358
step: 280, loss: 0.000521720910910517
step: 290, loss: 0.000131415348732844
step: 300, loss: 0.011189262382686138

```

```

step: 310, loss: 0.001590463682077825
step: 320, loss: 0.00017144737648777664
step: 330, loss: 0.0024736125487834215
step: 340, loss: 0.001820942503400147
step: 350, loss: 0.058346472680568695
step: 360, loss: 0.00034319443511776626
step: 370, loss: 0.00023920492094475776
step: 380, loss: 0.0007703941082581878
step: 390, loss: 0.0002529993071220815
step: 400, loss: 0.12750551104545593
step: 410, loss: 0.00035729468800127506
step: 420, loss: 0.007432352285832167
step: 430, loss: 0.007169617339968681
step: 440, loss: 7.198414095910266e-05
step: 450, loss: 0.0007664108416065574
step: 460, loss: 0.0015664247330278158
step: 470, loss: 0.0004658494144678116
step: 480, loss: 0.001127629424445331
step: 490, loss: 0.0004520677321124822
step: 500, loss: 0.0020705973729491234
step: 510, loss: 0.0008083077264018357
step: 520, loss: 0.0051104044541716576
step: 530, loss: 0.003850232809782028
/usr/local/lib/python3.7/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=0.9506057781919852, f1=0.9434843531060252,
best_f1=0.9464867380176826
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0003380195703357458
step: 10, loss: 0.0015208980767056346
step: 20, loss: 0.0018975718412548304
step: 30, loss: 0.0017015708144754171
step: 40, loss: 0.0007364371558651328
step: 50, loss: 0.002436678158119321
step: 60, loss: 0.005972565617412329
step: 70, loss: 0.0001263960439246148
step: 80, loss: 0.0047851670533418655
step: 90, loss: 4.0150745917344466e-05
step: 100, loss: 0.00017420918447896838
step: 110, loss: 0.011783311143517494
step: 120, loss: 0.0009520664461888373
step: 130, loss: 0.00015215255552902818
step: 140, loss: 0.00010473771544639021
step: 150, loss: 0.0002748311962932348
step: 160, loss: 0.04325684532523155

```



```

step: 170, loss: 0.0005567934131249785
step: 180, loss: 0.00031835175468586385
step: 190, loss: 2.6347279344918206e-05
step: 200, loss: 4.414805880514905e-05
step: 210, loss: 0.0009870269568637013
step: 220, loss: 0.000758049194701016
step: 230, loss: 2.438501178403385e-05
step: 240, loss: 4.319414802012034e-05
step: 250, loss: 0.003805846208706498
step: 260, loss: 0.0015103628393262625
step: 270, loss: 5.4282962082652375e-05
step: 280, loss: 0.03481755033135414
step: 290, loss: 0.00010128906433237717
step: 300, loss: 0.000721063872333616
step: 310, loss: 0.01930287852883339
step: 320, loss: 0.00021739781368523836
step: 330, loss: 0.006806053686887026
step: 340, loss: 7.280929276021197e-05
step: 350, loss: 0.00039852719055488706
step: 360, loss: 5.8624038501875475e-05
step: 370, loss: 0.0259723961353302
step: 380, loss: 0.00034824971226044
step: 390, loss: 0.11684215068817139
step: 400, loss: 0.002479265909641981
step: 410, loss: 0.0016014180146157742
step: 420, loss: 0.00011303351493552327
step: 430, loss: 0.00019492699357215315
step: 440, loss: 8.466611325275153e-05
step: 450, loss: 0.0006085888016968966
step: 460, loss: 0.0006405851454474032
step: 470, loss: 0.01430224534124136
step: 480, loss: 0.0008675642893649638
step: 490, loss: 0.0038884999230504036
step: 500, loss: 0.0015177371678873897
step: 510, loss: 0.00032085395650938153
step: 520, loss: 0.0005258123273961246
step: 530, loss: 0.008069991134107113
/usr/local/lib/python3.7/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=0.9501630181648812, f1=0.9471221338324755,
best_f1=0.9464867380176826
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0004948289133608341
step: 10, loss: 0.00012992705160286278
step: 20, loss: 0.001501885592006147

```

step: 30, loss: 0.00022811256349086761  
step: 40, loss: 7.154076592996716e-05  
step: 50, loss: 0.0005239497986622155  
step: 60, loss: 0.00011300771438982338  
step: 70, loss: 0.00035566333099268377  
step: 80, loss: 0.0009985121432691813  
step: 90, loss: 0.0012269638245925307  
step: 100, loss: 0.000240931156440638  
step: 110, loss: 5.466167567647062e-05  
step: 120, loss: 0.00017036426288541406  
step: 130, loss: 5.656334542436525e-05  
step: 140, loss: 4.9422811571275815e-05  
step: 150, loss: 0.002079392783343792  
step: 160, loss: 0.00029020648798905313  
step: 170, loss: 3.449102587183006e-05  
step: 180, loss: 0.0011587119661271572  
step: 190, loss: 0.14592324197292328  
step: 200, loss: 6.34294847259298e-05  
step: 210, loss: 0.0001668710756348446  
step: 220, loss: 0.005717751104384661  
step: 230, loss: 0.001304592820815742  
step: 240, loss: 0.000522451417054981  
step: 250, loss: 0.0036480792332440615  
step: 260, loss: 0.001417059451341629  
step: 270, loss: 0.0026302188634872437  
step: 280, loss: 0.003880593925714493  
step: 290, loss: 3.3161264582304284e-05  
step: 300, loss: 3.9742251829011366e-05  
step: 310, loss: 0.024966195225715637  
step: 320, loss: 0.00011113385698990896  
step: 330, loss: 2.7681569918058813e-05  
step: 340, loss: 0.003548920853063464  
step: 350, loss: 0.00013795559061691165  
step: 360, loss: 0.0002696604933589697  
step: 370, loss: 0.002125052735209465  
step: 380, loss: 0.0006017002160660923  
step: 390, loss: 0.0009862409206107259  
step: 400, loss: 0.039306577295064926  
step: 410, loss: 0.0018950001103803515  
step: 420, loss: 0.000777191249653697  
step: 430, loss: 0.0012645235983654857  
step: 440, loss: 0.00017229080549441278  
step: 450, loss: 0.000231521378736943  
step: 460, loss: 0.004533964209258556  
step: 470, loss: 0.00034494505962356925  
step: 480, loss: 0.0006135429721325636  
step: 490, loss: 0.004648908041417599  
step: 500, loss: 0.0005149584612809122

```

step: 510, loss: 0.0001736019621603191
step: 520, loss: 0.0002594335819594562
step: 530, loss: 0.02286817878484726
/usr/local/lib/python3.7/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=0.9496944052656323, f1=0.9413953488372093,
best_f1=0.9464867380176826
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0091921491548419
step: 10, loss: 0.0004095975018572062
step: 20, loss: 0.011544519104063511
step: 30, loss: 3.039947296201717e-05
step: 40, loss: 0.0003726030408870429
step: 50, loss: 2.511135426175315e-05
step: 60, loss: 4.719338539871387e-05
step: 70, loss: 6.626672256970778e-05
step: 80, loss: 0.00503385066986084
step: 90, loss: 4.9057947762776166e-05
step: 100, loss: 0.003937715198844671
step: 110, loss: 0.0005039493553340435
step: 120, loss: 7.728115451755002e-05
step: 130, loss: 0.00156929693184793
step: 140, loss: 2.947663415397983e-05
step: 150, loss: 0.0001268247578991577
step: 160, loss: 8.601349691161886e-05
step: 170, loss: 0.00016984695685096085
step: 180, loss: 2.1311643649823964e-05
step: 190, loss: 4.697880649473518e-05
step: 200, loss: 0.0002884110144805163
step: 210, loss: 4.7385834477609023e-05
step: 220, loss: 4.338540747994557e-05
step: 230, loss: 2.8791184377041645e-05
step: 240, loss: 0.0011362670920789242
step: 250, loss: 2.9806617021677084e-05
step: 260, loss: 4.6865960030118003e-05
step: 270, loss: 0.00012161157792434096
step: 280, loss: 0.00036775710759684443
step: 290, loss: 2.3293392587220296e-05
step: 300, loss: 0.00014314422151073813
step: 310, loss: 0.15885916352272034
step: 320, loss: 0.007796788588166237
step: 330, loss: 0.055851049721241
step: 340, loss: 0.0012045593466609716
step: 350, loss: 0.00011653387628030032
step: 360, loss: 0.0001321962772635743

```

```

step: 370, loss: 0.0006296260980889201
step: 380, loss: 0.0003986964584328234
step: 390, loss: 0.004584207199513912
step: 400, loss: 2.2786758563597687e-05
step: 410, loss: 5.7923673011828214e-05
step: 420, loss: 0.000269654905423522
step: 430, loss: 0.009690248407423496
step: 440, loss: 4.078376878169365e-05
step: 450, loss: 0.00021975785784889013
step: 460, loss: 0.00012635134044103324
step: 470, loss: 0.0010468311375007033
step: 480, loss: 0.00045617198338732123
step: 490, loss: 4.359573722467758e-05
step: 500, loss: 9.248782589565963e-05
step: 510, loss: 0.002812970196828246
step: 520, loss: 0.002568027237430215
step: 530, loss: 7.89305631769821e-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=0.95004712535344, f1=0.9441052137153593,
best_f1=0.9464867380176826
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.012446236796677113
step: 10, loss: 0.00011112452921224758
step: 20, loss: 5.313123983796686e-05
step: 30, loss: 1.4848676073597744e-05
step: 40, loss: 2.4566645151935518e-05
step: 50, loss: 0.0006025693146511912
step: 60, loss: 0.00010244004806736484
step: 70, loss: 0.0012210299028083682
step: 80, loss: 7.282780279638246e-05
step: 90, loss: 3.126484079984948e-05
step: 100, loss: 0.0011443570256233215
step: 110, loss: 1.2311938917264342e-05
step: 120, loss: 1.996699211304076e-05
step: 130, loss: 6.753401248715818e-05
step: 140, loss: 0.00010981885861838236
step: 150, loss: 0.00011168952187290415
step: 160, loss: 3.9168677176348865e-05
step: 170, loss: 3.13156378979329e-05
step: 180, loss: 2.2582054953090847e-05
step: 190, loss: 0.00034895483986474574
step: 200, loss: 2.2399128283723257e-05
step: 210, loss: 0.0001229621411766857
step: 220, loss: 2.115092138410546e-05

```

```

step: 230, loss: 2.4679249690962024e-05
step: 240, loss: 0.001257053343579173
step: 250, loss: 3.888268838636577e-05
step: 260, loss: 0.00014701149484608322
step: 270, loss: 0.0001218021716340445
step: 280, loss: 0.008684376254677773
step: 290, loss: 0.0025675962679088116
step: 300, loss: 1.9341163351782598e-05
step: 310, loss: 9.641284123063087e-05
step: 320, loss: 8.386738772969693e-05
step: 330, loss: 1.9061975763179362e-05
step: 340, loss: 2.433607733109966e-05
step: 350, loss: 9.123138625000138e-06
step: 360, loss: 0.07255271822214127
step: 370, loss: 0.00019951301510445774
step: 380, loss: 2.3833197701605968e-05
step: 390, loss: 1.8596158042782918e-05
step: 400, loss: 3.43524370691739e-05
step: 410, loss: 0.0248210858553648
step: 420, loss: 0.002154873451218009
step: 430, loss: 0.0003095695283263922
step: 440, loss: 2.3143633370636962e-05
step: 450, loss: 0.016674183309078217
step: 460, loss: 0.0006924314657226205
step: 470, loss: 0.0008834194159135222
step: 480, loss: 1.2918983884446789e-05
step: 490, loss: 0.006347615737468004
step: 500, loss: 1.9597604477894492e-05
step: 510, loss: 0.00017393435700796545
step: 520, loss: 7.610849570482969e-05
step: 530, loss: 0.00012907024938613176
/usr/local/lib/python3.7/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=0.9482596425211665, f1=0.942615239887112,
best_f1=0.9464867380176826
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0014112950302660465
step: 10, loss: 5.517063254956156e-05
step: 20, loss: 8.692011761013418e-05
step: 30, loss: 0.00014675605052616447
step: 40, loss: 0.00011914374772459269
step: 50, loss: 0.0030389949679374695
step: 60, loss: 5.730077828047797e-05
step: 70, loss: 0.00011124341835966334
step: 80, loss: 0.00010184250277234241

```

```
step: 90, loss: 1.5221112334984355e-05
step: 100, loss: 9.695412882138044e-05
step: 110, loss: 4.149025335209444e-05
step: 120, loss: 1.893891931104008e-05
step: 130, loss: 1.5929050277918577e-05
step: 140, loss: 5.214691918808967e-05
step: 150, loss: 7.192888733698055e-05
step: 160, loss: 2.2686155716655776e-05
step: 170, loss: 0.0007830546819604933
step: 180, loss: 7.622817065566778e-05
step: 190, loss: 3.6377092328621075e-05
step: 200, loss: 2.3576660169055685e-05
step: 210, loss: 7.289170025615022e-05
step: 220, loss: 4.634939614334144e-05
step: 230, loss: 3.9163973269751295e-05
step: 240, loss: 0.0001641005219426006
step: 250, loss: 0.001957208849489689
step: 260, loss: 2.5829316655290313e-05
step: 270, loss: 0.00047806446673348546
step: 280, loss: 1.3518771993403789e-05
step: 290, loss: 2.8810871299356222e-05
step: 300, loss: 7.227327296277508e-05
step: 310, loss: 1.9630662791314535e-05
step: 320, loss: 2.8578329875017516e-05
step: 330, loss: 2.6790097763296217e-05
step: 340, loss: 0.0001607692101970315
step: 350, loss: 1.9801966118393466e-05
step: 360, loss: 1.5537807485088706e-05
step: 370, loss: 0.002788426587358117
step: 380, loss: 0.00011607866326812655
step: 390, loss: 0.0002534578670747578
step: 400, loss: 0.0016377083957195282
step: 410, loss: 1.002462795440806e-05
step: 420, loss: 1.7355714589939453e-05
step: 430, loss: 0.0003766241134144366
step: 440, loss: 5.889981912332587e-05
step: 450, loss: 2.5351451768074185e-05
step: 460, loss: 0.020973041653633118
step: 470, loss: 1.9117420379188843e-05
step: 480, loss: 1.4237628420232795e-05
step: 490, loss: 5.042196426074952e-05
step: 500, loss: 0.061101283878088
step: 510, loss: 0.007580767851322889
step: 520, loss: 4.244318915880285e-05
step: 530, loss: 0.0012121936306357384
/usr/local/lib/python3.7/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=0.9504672897196261, f1=0.9453015427769986,  
best_f1=0.9464867380176826  
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:  
An input tensor was not cuda.  
    warnings.warn("An input tensor was not cuda.")  
step: 0, loss: 2.2022391931386665e-05  
step: 10, loss: 0.0006574745057150722  
step: 20, loss: 0.0001231592905241996  
step: 30, loss: 0.00485201133415103  
step: 40, loss: 2.1046740585006773e-05  
step: 50, loss: 0.00020327542733866721  
step: 60, loss: 0.0013512088917195797  
step: 70, loss: 1.9899300241377205e-05  
step: 80, loss: 1.2971291653229855e-05  
step: 90, loss: 2.544150993344374e-05  
step: 100, loss: 7.670313607377466e-06  
step: 110, loss: 0.00040068128146231174  
step: 120, loss: 1.2743970728479326e-05  
step: 130, loss: 3.3577824069652706e-05  
step: 140, loss: 1.5686789993196726e-05  
step: 150, loss: 2.6645355319487862e-05  
step: 160, loss: 2.362384293519426e-05  
step: 170, loss: 1.3030736226937734e-05  
step: 180, loss: 1.92553416127339e-05  
step: 190, loss: 4.263272421667352e-05  
step: 200, loss: 0.00010143612598767504  
step: 210, loss: 2.0386498363222927e-05  
step: 220, loss: 1.8812079360941425e-05  
step: 230, loss: 0.01337635051459074  
step: 240, loss: 2.2301870558294468e-05  
step: 250, loss: 8.083809916570317e-06  
step: 260, loss: 9.842104191193357e-06  
step: 270, loss: 2.223037154180929e-05  
step: 280, loss: 1.1600402103795204e-05  
step: 290, loss: 9.179003427561838e-06  
step: 300, loss: 1.4937862943043001e-05  
step: 310, loss: 0.007150289136916399  
step: 320, loss: 2.000023414439056e-05  
step: 330, loss: 0.00031311606289818883  
step: 340, loss: 2.8708847821690142e-05  
step: 350, loss: 2.011559445236344e-05  
step: 360, loss: 0.00029084988636896014  
step: 370, loss: 0.0002430541644571349  
step: 380, loss: 4.416201409185305e-05  
step: 390, loss: 3.922516407328658e-05  
step: 400, loss: 2.4876228053472005e-05  
step: 410, loss: 3.7512047128984705e-05  
step: 420, loss: 1.3485228009813e-05
```

```

step: 430, loss: 8.072620403254405e-06
step: 440, loss: 1.131354110839311e-05
step: 450, loss: 0.005462229251861572
step: 460, loss: 0.00018487458874005824
step: 470, loss: 1.2174056791991461e-05
step: 480, loss: 0.0013142938259989023
step: 490, loss: 8.730898116482422e-05
step: 500, loss: 0.00011440498201409355
step: 510, loss: 1.194305150420405e-05
step: 520, loss: 5.734053047490306e-05
step: 530, loss: 1.4651153833256103e-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=0.9506517690875234, f1=0.9433085501858736,
best_f1=0.9464867380176826
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, 171.46it/s]
load_f1 = 0.9522041763341067

```



```

real_f1 = 0.9521152952115295
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, 143.19it/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 1.2 MB/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: 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: 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: 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: 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: 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: 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: 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: 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: 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: 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: 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: 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: 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: 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: 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=b8ee3362a0d56ed0b4cfb15d1ec1d0fb67fd413bfb9e8e0af0f3b3fe66437b9e
  Stored in directory: /tmp/pip-ephem-wheel-cache-1t9kalor/wheels/11/95/ba/2c36c
c368c0bd339b44a791c2c1881a1fb714b78c29a4cb8f5
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/
```

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: 0.4464835822582245
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=0.2666666666666667, f1=0.2666666666666667,
best_f1=0.2666666666666667
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.3882683515548706
/usr/local/lib/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=0.37837837837837834, f1=0.26262626262626265,
best_f1=0.26262626262626265

```

```

/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.40121328830718994
/usr/local/lib/python3.7/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=0.2692307692307693, f1=0.2692307692307693,
best_f1=0.26262626262626265
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.3524555265903473
/usr/local/lib/python3.7/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=0.2692307692307693, f1=0.2692307692307693,
best_f1=0.26262626262626265
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.30721378326416016
/usr/local/lib/python3.7/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=0.2692307692307693, f1=0.2692307692307693,
best_f1=0.26262626262626265
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.2896624803543091
/usr/local/lib/python3.7/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=0.2692307692307693, f1=0.2692307692307693,
best_f1=0.26262626262626265
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.6813454627990723
/usr/local/lib/python3.7/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=0.2692307692307693, f1=0.2692307692307693,
best_f1=0.26262626262626265
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")

```

```

step: 0, loss: 0.5024064183235168
/usr/local/lib/python3.7/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=0.2692307692307693, f1=0.2692307692307693,
best_f1=0.26262626262626265
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.37582382559776306
/usr/local/lib/python3.7/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=0.2692307692307693, f1=0.2692307692307693,
best_f1=0.26262626262626265
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.48656052350997925
/usr/local/lib/python3.7/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=0.27184466019417475, f1=0.2692307692307693,
best_f1=0.26262626262626265
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.506280243396759
/usr/local/lib/python3.7/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=0.28, f1=0.2692307692307693, best_f1=0.26262626262626265
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.44152024388313293
/usr/local/lib/python3.7/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=0.2916666666666667, f1=0.2692307692307693,
best_f1=0.26262626262626265
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.4162924289703369
/usr/local/lib/python3.7/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=0.288659793814433, f1=0.2692307692307693,
best_f1=0.26262626262626265
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.35893523693084717
/usr/local/lib/python3.7/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=0.288659793814433, f1=0.2692307692307693,
best_f1=0.26262626262626265
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.40308016538619995
/usr/local/lib/python3.7/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=0.288659793814433, f1=0.2692307692307693,
best_f1=0.26262626262626265
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, 125925.99it/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 = 0.3333333333333337
real_f1 = 0.30434782608695654
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 [00:20, 216.17it/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/

```

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.")
step: 0, loss: 0.642425000667572
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: 0.5017406344413757
step: 20, loss: 0.5103674530982971
step: 30, loss: 0.26089468598365784
step: 40, loss: 0.3818041682243347
step: 50, loss: 0.668758749961853
step: 60, loss: 0.4791797995567322
step: 70, loss: 0.37628480792045593
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: 0.3796873092651367
step: 90, loss: 0.0978928804397583
step: 100, loss: 0.15923985838890076
step: 110, loss: 0.08144884556531906
step: 120, loss: 0.10255473852157593
step: 130, loss: 0.024221794679760933
step: 140, loss: 0.014062696136534214
step: 150, loss: 0.1743423491716385
step: 160, loss: 0.012653644196689129
step: 170, loss: 0.1481861025094986
step: 180, loss: 0.07103492319583893
step: 190, loss: 0.023944532498717308

```

```

step: 200, loss: 0.03059191256761551
step: 210, loss: 0.013501688838005066
step: 220, loss: 0.050308212637901306
step: 230, loss: 0.007186539005488157
/usr/local/lib/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=0.9743589743589743, f1=0.9709821428571428,
best_f1=0.9709821428571428
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.013825001195073128
step: 10, loss: 0.06415978074073792
step: 20, loss: 0.010028814896941185
step: 30, loss: 0.007613536436110735
step: 40, loss: 0.043118804693222046
step: 50, loss: 0.001932749291881919
step: 60, loss: 0.0065463874489068985
step: 70, loss: 0.007705240044742823
step: 80, loss: 0.002561191562563181
step: 90, loss: 0.01105353981256485
step: 100, loss: 0.04220879450440407
step: 110, loss: 0.019783522933721542
step: 120, loss: 0.0012342592235654593
step: 130, loss: 0.054859939962625504
step: 140, loss: 0.0016032621497288346
step: 150, loss: 0.05695808306336403
step: 160, loss: 0.09995504468679428
step: 170, loss: 0.017613552510738373
step: 180, loss: 0.005648380611091852
step: 190, loss: 0.026212861761450768
step: 200, loss: 0.0038839494809508324
step: 210, loss: 0.02364102192223072
step: 220, loss: 0.002109998371452093
step: 230, loss: 0.0017560218693688512
/usr/local/lib/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=0.9819004524886877, f1=0.9759999999999999,

```

```

best_f1=0.9759999999999999
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00975214596837759
step: 10, loss: 0.012295710854232311
step: 20, loss: 0.0023515603970736265
step: 30, loss: 0.024020466953516006
step: 40, loss: 0.01965656504034996
step: 50, loss: 0.0684603899717331
step: 60, loss: 0.0012475531548261642
step: 70, loss: 0.0008373182499781251
step: 80, loss: 0.0003917946887668222
step: 90, loss: 0.005513705778867006
step: 100, loss: 0.0009700815426185727
step: 110, loss: 0.007343923673033714
step: 120, loss: 0.0002742062497418374
step: 130, loss: 0.003801067592576146
step: 140, loss: 0.000866077549289912
step: 150, loss: 0.006922523491084576
step: 160, loss: 0.0066132089123129845
step: 170, loss: 0.007128482218831778
step: 180, loss: 0.008291102945804596
step: 190, loss: 0.02285112626850605
step: 200, loss: 0.01119040697813034
step: 210, loss: 0.003248992608860135
step: 220, loss: 0.005212882999330759
step: 230, loss: 0.006613322999328375
/usr/local/lib/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=0.9841628959276018, f1=0.9817767653758542,
best_f1=0.9817767653758542
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.007583811413496733
step: 10, loss: 0.001373698585666716
step: 20, loss: 0.0017009630100801587
step: 30, loss: 0.0011821433436125517
step: 40, loss: 0.1014334037899971
step: 50, loss: 0.009383843280375004
step: 60, loss: 0.011903120204806328
step: 70, loss: 0.007983568124473095

```

```

step: 80, loss: 0.0007251294446177781
step: 90, loss: 0.0034628533758223057
step: 100, loss: 0.001198501675389707
step: 110, loss: 0.0007024130900390446
step: 120, loss: 0.0016079447232186794
step: 130, loss: 0.00957517884671688
step: 140, loss: 0.0031223262194544077
step: 150, loss: 0.0003345843288116157
step: 160, loss: 0.000890760391484946
step: 170, loss: 0.003190930699929595
step: 180, loss: 0.17122255265712738
step: 190, loss: 0.0060203587636351585
step: 200, loss: 0.0019636410288512707
step: 210, loss: 0.0006054681143723428
step: 220, loss: 0.0006189013365656137
step: 230, loss: 0.004168540704995394
/usr/local/lib/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=0.9898762654668166, f1=0.9854096520763187,
best_f1=0.9854096520763187
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0011390057625249028
step: 10, loss: 0.0019692606292665005
step: 20, loss: 0.0012602244969457388
step: 30, loss: 0.0007584878476336598
step: 40, loss: 0.0012641869252547622
step: 50, loss: 0.009406768716871738
step: 60, loss: 0.32990872859954834
step: 70, loss: 0.001613251632079482
step: 80, loss: 0.08294341713190079
step: 90, loss: 0.03838430345058441
step: 100, loss: 0.0004959430661983788
step: 110, loss: 0.04047269746661186
step: 120, loss: 0.004568095784634352
step: 130, loss: 0.0011925563449040055
step: 140, loss: 0.003099958412349224
step: 150, loss: 0.15344658493995667
step: 160, loss: 0.0014114173827692866
step: 170, loss: 0.0035601069685071707
step: 180, loss: 0.0134193766862154
step: 190, loss: 0.010896381922066212

```

```

step: 200, loss: 0.0009666422847658396
step: 210, loss: 0.007130152080208063
step: 220, loss: 0.0010894632432609797
step: 230, loss: 0.01655147783458233
/usr/local/lib/python3.7/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=0.9854748603351955, f1=0.9887133182844244,
best_f1=0.9854096520763187
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.001039210706949234
step: 10, loss: 0.0013619466917589307
step: 20, loss: 0.0013389494270086288
step: 30, loss: 0.0008213629480451345
step: 40, loss: 0.00041643521399237216
step: 50, loss: 0.0015926542691886425
step: 60, loss: 0.0008094740333035588
step: 70, loss: 0.00041037038317881525
step: 80, loss: 0.0016776411794126034
step: 90, loss: 0.013464372605085373
step: 100, loss: 0.0007202615379355848
step: 110, loss: 0.020474523305892944
step: 120, loss: 0.0004686861939262599
step: 130, loss: 0.0004612489137798548
step: 140, loss: 0.0005610320949926972
step: 150, loss: 0.00010268326877849177
step: 160, loss: 0.030357204377651215
step: 170, loss: 0.000409951142501086
step: 180, loss: 0.0016846274957060814
step: 190, loss: 0.0005928885657340288
step: 200, loss: 0.004244954325258732
step: 210, loss: 0.0031360520515590906
step: 220, loss: 0.0661851167678833
step: 230, loss: 0.0009360030526295304
/usr/local/lib/python3.7/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=0.9887892376681614, f1=0.9876265466816648,
best_f1=0.9854096520763187
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0014375094324350357
step: 10, loss: 0.004966686014086008
step: 20, loss: 0.0004405365907587111
step: 30, loss: 0.0009448733180761337

```

```

step: 40, loss: 0.0023988732136785984
step: 50, loss: 0.0006069926312193274
step: 60, loss: 0.0006414172239601612
step: 70, loss: 0.0017863760003820062
step: 80, loss: 0.0003319327952340245
step: 90, loss: 0.0002732092689257115
step: 100, loss: 0.000467059260699898
step: 110, loss: 0.0007066680118441582
step: 120, loss: 0.0016971349250525236
step: 130, loss: 0.004045458976179361
step: 140, loss: 0.0004324638575781137
step: 150, loss: 0.018132738769054413
step: 160, loss: 0.0004243534349370748
step: 170, loss: 0.0005134929670020938
step: 180, loss: 0.0002461016410961747
step: 190, loss: 0.00039906188612803817
step: 200, loss: 0.02335965260863304
step: 210, loss: 0.00021230966376606375
step: 220, loss: 0.0006113977869972587
step: 230, loss: 0.0007983265677466989
/usr/local/lib/python3.7/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=0.9865470852017937, f1=0.9898534385569334,
best_f1=0.9854096520763187
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0007676382083445787
step: 10, loss: 0.007765226531773806
step: 20, loss: 0.002901819534599781
step: 30, loss: 0.0006829558406025171
step: 40, loss: 0.00051949976477772
step: 50, loss: 0.001476836041547358
step: 60, loss: 0.0003574073198251426
step: 70, loss: 0.0001952205057023093
step: 80, loss: 0.0004320907173678279
step: 90, loss: 0.00043643455137498677
step: 100, loss: 0.0006240651709958911
step: 110, loss: 0.0005139712593518198
step: 120, loss: 0.0022956125903874636
step: 130, loss: 0.0008265624055638909
step: 140, loss: 0.0003466866328381002
step: 150, loss: 0.11724685877561569
step: 160, loss: 0.022033432498574257
step: 170, loss: 0.03855443000793457
step: 180, loss: 0.0005812786403112113
step: 190, loss: 0.014773928560316563

```

```

step: 200, loss: 0.007046978920698166
step: 210, loss: 0.0012361397966742516
step: 220, loss: 0.0005695613799616694
step: 230, loss: 0.000590082781855017
/usr/local/lib/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=0.9898989898989898, f1=0.9831271091113611,
best_f1=0.9831271091113611
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.000425478327088058
step: 10, loss: 0.008282983675599098
step: 20, loss: 0.0006905407644808292
step: 30, loss: 0.0013126143021509051
step: 40, loss: 0.0007154716877266765
step: 50, loss: 0.0008065300062298775
step: 60, loss: 0.0018369709141552448
step: 70, loss: 0.04417150467634201
step: 80, loss: 0.0001325635239481926
step: 90, loss: 0.010572846047580242
step: 100, loss: 0.00022088302648626268
step: 110, loss: 0.00017490806931164116
step: 120, loss: 0.01065025757998228
step: 130, loss: 0.0005525926244445145
step: 140, loss: 0.00045215917634777725
step: 150, loss: 0.0010679508559405804
step: 160, loss: 0.00137861049734056
step: 170, loss: 0.0004948487039655447
step: 180, loss: 0.0021537276916205883
step: 190, loss: 0.0002769882557913661
step: 200, loss: 0.0004119597433600575
step: 210, loss: 0.0036184233613312244
step: 220, loss: 0.0005307663232088089
step: 230, loss: 0.000341481645591557
/usr/local/lib/python3.7/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=0.9854748603351955, f1=0.9854096520763187,
best_f1=0.9831271091113611
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")

```

```

step: 0, loss: 0.0009103349293582141
step: 10, loss: 0.0020700758323073387
step: 20, loss: 0.0005596668925136328
step: 30, loss: 0.00038873523590154946
step: 40, loss: 0.0014280484756454825
step: 50, loss: 0.00031027282238937914
step: 60, loss: 0.00041427984251640737
step: 70, loss: 0.0009211570140905678
step: 80, loss: 0.0002195816778112203
step: 90, loss: 0.0002469484752509743
step: 100, loss: 0.0001652469945838675
step: 110, loss: 0.00028491736156865954
step: 120, loss: 0.0001770404342096299
step: 130, loss: 0.0005420050583779812
step: 140, loss: 0.00045483617577701807
step: 150, loss: 0.002367848763242364
step: 160, loss: 0.00021355622448027134
step: 170, loss: 0.00039031117921695113
step: 180, loss: 0.005482246167957783
step: 190, loss: 0.0006904475740157068
step: 200, loss: 0.0006232013693079352
step: 210, loss: 0.0008714560535736382
step: 220, loss: 0.0001760378945618868
step: 230, loss: 0.0002215367858298123
/usr/local/lib/python3.7/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=0.987709497206704, f1=0.9865771812080537,
best_f1=0.9831271091113611
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00019753210654016584
step: 10, loss: 0.0036594851408153772
step: 20, loss: 0.0003764512948691845
step: 30, loss: 0.0002448961604386568
step: 40, loss: 9.657604095991701e-05
step: 50, loss: 0.00014799443306401372
step: 60, loss: 0.0070287734270095825
step: 70, loss: 0.00014100289263296872
step: 80, loss: 0.0006586425588466227
step: 90, loss: 0.0048077162355184555
step: 100, loss: 0.0023100809194147587
step: 110, loss: 0.0022427989169955254
step: 120, loss: 0.0006817023968324065
step: 130, loss: 0.0004948003916069865
step: 140, loss: 0.009164690040051937
step: 150, loss: 0.0015668851556256413

```



```

step: 160, loss: 0.03281458094716072
step: 170, loss: 0.01885787770152092
step: 180, loss: 0.000667267304379493
step: 190, loss: 0.00044515958870761096
step: 200, loss: 0.005036630667746067
step: 210, loss: 0.0003425592731218785
step: 220, loss: 0.0013879985781386495
step: 230, loss: 0.0004942481173202395
/usr/local/lib/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=0.992108229988726, f1=0.9887133182844244,
best_f1=0.9887133182844244
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.000283715664409101
step: 10, loss: 0.0001677164837019518
step: 20, loss: 0.013975896872580051
step: 30, loss: 0.02996145375072956
step: 40, loss: 0.0008043526322580874
step: 50, loss: 0.0006920191226527095
step: 60, loss: 0.001203193562105298
step: 70, loss: 0.0007264676969498396
step: 80, loss: 0.00016816990682855248
step: 90, loss: 0.002199603710323572
step: 100, loss: 0.00019984458049293607
step: 110, loss: 0.00014716126315761358
step: 120, loss: 0.00018890330102294683
step: 130, loss: 0.00012017120752716437
step: 140, loss: 0.000460757699329406
step: 150, loss: 0.0002493517240509391
step: 160, loss: 0.006999438162893057
step: 170, loss: 0.00020266325736884028
step: 180, loss: 0.0001579675736138597
step: 190, loss: 0.004355217795819044
step: 200, loss: 0.00016674326616339386
step: 210, loss: 0.00027442339342087507
step: 220, loss: 0.006711889524012804
step: 230, loss: 0.0017209643265232444
/usr/local/lib/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=0.9921436588103255, f1=0.984304932735426,
best_f1=0.984304932735426
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
```

```
warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0013659364776685834
step: 10, loss: 0.00031585790566168725
step: 20, loss: 0.0008338844636455178
step: 30, loss: 0.0004189653554931283
step: 40, loss: 0.0011647335486486554
step: 50, loss: 0.007552823517471552
step: 60, loss: 0.00042577803833410144
step: 70, loss: 0.0005084963631816208
step: 80, loss: 0.0004337663121987134
step: 90, loss: 0.0001558627118356526
step: 100, loss: 0.0015805287985131145
step: 110, loss: 0.000893113377969712
step: 120, loss: 0.00020512549963314086
step: 130, loss: 0.00020589045016095042
step: 140, loss: 0.00012355307990219444
step: 150, loss: 6.376688543241471e-05
step: 160, loss: 0.00022254962823353708
step: 170, loss: 0.00014739031030330807
step: 180, loss: 0.009847572073340416
step: 190, loss: 0.0002571415388956666
step: 200, loss: 3.867768100462854e-05
step: 210, loss: 0.005773370619863272
step: 220, loss: 0.00011620506847975776
step: 230, loss: 0.00016839485033415258
/usr/local/lib/python3.7/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=0.9910313901345291, f1=0.9854423292273236,
best_f1=0.984304932735426
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
```

```
warnings.warn("An input tensor was not cuda.")
step: 0, loss: 7.854962314013392e-05
step: 10, loss: 9.133284038398415e-05
step: 20, loss: 0.00011773577716667205
step: 30, loss: 0.00016900579794310033
step: 40, loss: 0.000107829604530707
step: 50, loss: 6.38978453935124e-05
step: 60, loss: 9.689160651760176e-05
step: 70, loss: 0.00015548487135674804
```

```

step: 80, loss: 0.00010986236884491518
step: 90, loss: 0.0004798285663127899
step: 100, loss: 0.0002060107362922281
step: 110, loss: 0.00029075221391394734
step: 120, loss: 6.016324186930433e-05
step: 130, loss: 0.0003563356585800648
step: 140, loss: 7.093021122273058e-05
step: 150, loss: 5.261049227556214e-05
step: 160, loss: 0.0009601161582395434
step: 170, loss: 0.00013744279567617923
step: 180, loss: 0.00012511477689258754
step: 190, loss: 7.48715756344609e-05
step: 200, loss: 0.0008335883030667901
step: 210, loss: 0.00010969580034725368
step: 220, loss: 0.00014754087897017598
step: 230, loss: 8.156183321261778e-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=0.9910313901345291, f1=0.9854096520763187,
best_f1=0.984304932735426
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0002877477090805769
step: 10, loss: 0.00013701235002372414
step: 20, loss: 0.0011643999023362994
step: 30, loss: 8.244549826486036e-05
step: 40, loss: 5.779182174592279e-05
step: 50, loss: 7.02483594068326e-05
step: 60, loss: 0.06856506317853928
step: 70, loss: 0.00011189009092049673
step: 80, loss: 6.217105692485347e-05
step: 90, loss: 5.350727224140428e-05
step: 100, loss: 0.00011758998152799904
step: 110, loss: 0.000198557652765885
step: 120, loss: 0.015836775302886963
step: 130, loss: 5.674786734743975e-05
step: 140, loss: 0.01004179660230875
step: 150, loss: 0.00010817413567565382
step: 160, loss: 0.00956691149622202
step: 170, loss: 3.553222632035613e-05
step: 180, loss: 0.0005517103127203882
step: 190, loss: 0.00253272638656199
step: 200, loss: 9.619855700293556e-05
step: 210, loss: 0.0431685633957386
step: 220, loss: 0.0001009062398225069
step: 230, loss: 0.00010814137931447476

```

```

/usr/local/lib/python3.7/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=0.9898534385569334,
best_f1=0.984304932735426
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, 218.45it/s]
load_f1 = 0.9943630214205187
real_f1 = 0.9898534385569334
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, 199.36it/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.")
step: 0, loss: 0.6089612245559692
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: 0.4016510248184204
step: 20, loss: 0.31712397933006287
step: 30, loss: 0.372112900018692
step: 40, loss: 0.34108850359916687
step: 50, loss: 0.08926969766616821
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 8192.0
step: 60, loss: 0.13151559233665466
step: 70, loss: 0.24695783853530884
step: 80, loss: 0.11330589652061462
step: 90, loss: 0.18827638030052185
step: 100, loss: 0.1722632348537445
step: 110, loss: 0.14077606797218323
step: 120, loss: 0.14251145720481873
step: 130, loss: 0.259803831577301
step: 140, loss: 0.3016637861728668
step: 150, loss: 0.1206800788640976
step: 160, loss: 0.2311892807483673
step: 170, loss: 0.06388650834560394
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 4096.0
step: 180, loss: 0.021601304411888123
step: 190, loss: 0.12284837663173676
step: 200, loss: 0.03699744865298271
step: 210, loss: 0.03647492825984955
step: 220, loss: 0.07315949350595474
step: 230, loss: 0.1511138677597046
step: 240, loss: 0.02041328325867653
step: 250, loss: 0.11886411905288696
step: 260, loss: 0.2537340819835663
step: 270, loss: 0.2752493619918823
step: 280, loss: 0.045938290655612946
step: 290, loss: 0.13688188791275024
step: 300, loss: 0.09582068026065826

```

```

step: 310, loss: 0.11906357854604721
step: 320, loss: 0.0617513544857502
step: 330, loss: 0.077440045773983
step: 340, loss: 0.356437623500824
step: 350, loss: 0.0876178964972496
step: 360, loss: 0.10020691901445389
step: 370, loss: 0.08274967968463898
step: 380, loss: 0.09306110441684723
step: 390, loss: 0.0078098708763718605
step: 400, loss: 0.057483673095703125
step: 410, loss: 0.2621709406375885
step: 420, loss: 0.04580499976873398
step: 430, loss: 0.06847318261861801
step: 440, loss: 0.04297994449734688
step: 450, loss: 0.011385372839868069
step: 460, loss: 0.01477331854403019
step: 470, loss: 0.014120325446128845
step: 480, loss: 0.10617470741271973
step: 490, loss: 0.12974843382835388
step: 500, loss: 0.03273019939661026
step: 510, loss: 0.018962595611810684
step: 520, loss: 0.2411423921585083
step: 530, loss: 0.030779512599110603
/usr/local/lib/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=0.9343735658558971, f1=0.9372423270728356,
best_f1=0.9372423270728356
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0972883552312851
step: 10, loss: 0.0316658578813076
step: 20, loss: 0.04629868268966675
step: 30, loss: 0.11915736645460129
step: 40, loss: 0.0741962268948555
step: 50, loss: 0.030807267874479294
step: 60, loss: 0.029400324448943138
step: 70, loss: 0.040451645851135254
step: 80, loss: 0.024475038051605225
step: 90, loss: 0.01082471664994955
step: 100, loss: 0.15096919238567352
step: 110, loss: 0.013290183618664742
step: 120, loss: 0.2682948410511017

```

```

step: 130, loss: 0.033416200429201126
step: 140, loss: 0.05682038515806198
step: 150, loss: 0.022475028410553932
step: 160, loss: 0.006962081417441368
step: 170, loss: 0.02391129918396473
step: 180, loss: 0.03983677923679352
step: 190, loss: 0.03798910602927208
step: 200, loss: 0.19492414593696594
step: 210, loss: 0.03684442117810249
step: 220, loss: 0.0010576056083664298
step: 230, loss: 0.046231575310230255
step: 240, loss: 0.12334544956684113
step: 250, loss: 0.09374584257602692
step: 260, loss: 0.05963636934757233
step: 270, loss: 0.009003497660160065
step: 280, loss: 0.03202337026596069
step: 290, loss: 0.014480789192020893
step: 300, loss: 0.011458178982138634
step: 310, loss: 0.04805297777056694
step: 320, loss: 0.026231421157717705
step: 330, loss: 0.05190826952457428
step: 340, loss: 0.0670299082994461
step: 350, loss: 0.0025494894944131374
step: 360, loss: 0.02439262717962265
step: 370, loss: 0.017327435314655304
step: 380, loss: 0.13336478173732758
step: 390, loss: 0.007565770763903856
step: 400, loss: 0.07867047190666199
step: 410, loss: 0.01929369382560253
step: 420, loss: 0.009878741577267647
step: 430, loss: 0.1641150563955307
step: 440, loss: 0.08742599189281464
step: 450, loss: 0.03763822093605995
step: 460, loss: 0.056781619787216187
step: 470, loss: 0.034485358744859695
step: 480, loss: 0.02490108646452427
step: 490, loss: 0.17578420042991638
step: 500, loss: 0.008340057916939259
step: 510, loss: 0.0219767726957798
step: 520, loss: 0.35816752910614014
step: 530, loss: 0.13343289494514465
/usr/local/lib/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=0.9436947417403444, f1=0.9404096834264433,  
best_f1=0.9404096834264433  
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:  
An input tensor was not cuda.  
    warnings.warn("An input tensor was not cuda.")  
step: 0, loss: 0.14337767660617828  
step: 10, loss: 0.02861062064766884  
step: 20, loss: 0.0037493640556931496  
step: 30, loss: 0.06681866943836212  
step: 40, loss: 0.02986249141395092  
step: 50, loss: 0.06075010448694229  
step: 60, loss: 0.013322013430297375  
step: 70, loss: 0.007422365248203278  
step: 80, loss: 0.020761841908097267  
step: 90, loss: 0.016984155401587486  
step: 100, loss: 0.24530573189258575  
step: 110, loss: 0.022914189845323563  
step: 120, loss: 0.13962498307228088  
step: 130, loss: 0.015649668872356415  
step: 140, loss: 0.003879171796143055  
step: 150, loss: 0.03833119571208954  
step: 160, loss: 0.02237282507121563  
step: 170, loss: 0.023375531658530235  
step: 180, loss: 0.0168803371489048  
step: 190, loss: 0.0061256168410182  
step: 200, loss: 0.01817593351006508  
step: 210, loss: 0.08632976561784744  
step: 220, loss: 0.1375247985124588  
step: 230, loss: 0.04292796924710274  
step: 240, loss: 0.07236658781766891  
step: 250, loss: 0.04711786285042763  
step: 260, loss: 0.10006499290466309  
step: 270, loss: 0.036699019372463226  
step: 280, loss: 0.02499477006494999  
step: 290, loss: 0.006564737297594547  
step: 300, loss: 0.13059619069099426  
step: 310, loss: 0.03486108407378197  
step: 320, loss: 0.031602129340171814  
step: 330, loss: 0.039549026638269424  
step: 340, loss: 0.014266104437410831  
step: 350, loss: 0.051969900727272034  
step: 360, loss: 0.0030975532718002796  
step: 370, loss: 0.02464665286242962  
step: 380, loss: 0.07051780819892883  
step: 390, loss: 0.019598081707954407  
step: 400, loss: 0.2434217780828476  
step: 410, loss: 0.0759233608841896  
step: 420, loss: 0.006586997304111719
```

```

step: 430, loss: 0.014667063020169735
step: 440, loss: 0.22918832302093506
step: 450, loss: 0.062224168330430984
step: 460, loss: 0.041979677975177765
step: 470, loss: 0.08932089805603027
step: 480, loss: 0.12551918625831604
step: 490, loss: 0.007557259406894445
step: 500, loss: 0.05375609174370766
step: 510, loss: 0.05619954317808151
step: 520, loss: 0.005751168355345726
step: 530, loss: 0.004565411247313023
/usr/local/lib/python3.7/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=0.9433962264150944, f1=0.9399527186761228,
best_f1=0.9404096834264433
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.024954143911600113
step: 10, loss: 0.0011167444754391909
step: 20, loss: 0.044858019798994064
step: 30, loss: 0.020094314590096474
step: 40, loss: 0.041448045521974564
step: 50, loss: 0.12475957721471786
step: 60, loss: 0.0032713774126023054
step: 70, loss: 0.02288908138871193
step: 80, loss: 0.01990930177271366
step: 90, loss: 0.013755979016423225
step: 100, loss: 0.004922434221953154
step: 110, loss: 0.0671735405921936
step: 120, loss: 0.0024781590327620506
step: 130, loss: 0.007642771117389202
step: 140, loss: 0.005945139564573765
step: 150, loss: 0.006982822436839342
step: 160, loss: 0.04271283000707626
step: 170, loss: 0.016866430640220642
step: 180, loss: 0.13958178460597992
step: 190, loss: 0.05831108242273331
step: 200, loss: 0.05553830415010452
step: 210, loss: 0.0029633580707013607
step: 220, loss: 0.011327115818858147
step: 230, loss: 0.01319810375571251
step: 240, loss: 0.008945898152887821
step: 250, loss: 0.09204365313053131
step: 260, loss: 0.0016071901191025972
step: 270, loss: 0.11499883234500885
step: 280, loss: 0.003190105315297842

```

```

step: 290, loss: 0.005142547190189362
step: 300, loss: 0.003231661394238472
step: 310, loss: 0.0069673736579716206
step: 320, loss: 0.08667398244142532
step: 330, loss: 0.05328160151839256
step: 340, loss: 0.001869620056822896
step: 350, loss: 0.012089907191693783
step: 360, loss: 0.007795896381139755
step: 370, loss: 0.0020241448655724525
step: 380, loss: 0.007504928857088089
step: 390, loss: 0.003974901977926493
step: 400, loss: 0.002857252024114132
step: 410, loss: 0.0018647804390639067
step: 420, loss: 0.0036382453981786966
step: 430, loss: 0.0411624051630497
step: 440, loss: 0.003761813510209322
step: 450, loss: 0.022352585569024086
step: 460, loss: 0.014157080091536045
step: 470, loss: 0.0028673307970166206
step: 480, loss: 0.012196121737360954
step: 490, loss: 0.009077120572328568
step: 500, loss: 0.023836927488446236
step: 510, loss: 0.20543009042739868
step: 520, loss: 0.007039960473775864
step: 530, loss: 0.10835722833871841
/usr/local/lib/python3.7/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=0.9383974062065772, f1=0.9400369003690037,
best_f1=0.9404096834264433
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0015874895034357905
step: 10, loss: 0.010793488472700119
step: 20, loss: 0.0031174037139862776
step: 30, loss: 0.005681683775037527
step: 40, loss: 0.0011616022093221545
step: 50, loss: 0.03284046798944473
step: 60, loss: 0.005954340565949678
step: 70, loss: 0.0022264081053435802
step: 80, loss: 0.000887762289494276
step: 90, loss: 0.13189764320850372
step: 100, loss: 0.02387685887515545
step: 110, loss: 0.020900703966617584
step: 120, loss: 0.18705390393733978
step: 130, loss: 0.027182800695300102
step: 140, loss: 0.0030806618742644787

```

```

step: 150, loss: 0.0037601622752845287
step: 160, loss: 0.003608164144679904
step: 170, loss: 0.03950970619916916
step: 180, loss: 0.011919633485376835
step: 190, loss: 0.01010801550000906
step: 200, loss: 0.042239148169755936
step: 210, loss: 0.0027351106982678175
step: 220, loss: 0.0036225812509655952
step: 230, loss: 0.002401405945420265
step: 240, loss: 0.024915920570492744
step: 250, loss: 0.039857909083366394
step: 260, loss: 0.000571323384065181
step: 270, loss: 0.013778243213891983
step: 280, loss: 0.01629810221493244
step: 290, loss: 0.010015456937253475
step: 300, loss: 0.0313040055334568
step: 310, loss: 0.08394245058298111
step: 320, loss: 0.011238766834139824
step: 330, loss: 0.0036923668812960386
step: 340, loss: 0.005118487402796745
step: 350, loss: 0.0008642902248539031
step: 360, loss: 0.004011403303593397
step: 370, loss: 0.000465579389128834
step: 380, loss: 0.00022423797054216266
step: 390, loss: 0.00038616755045950413
step: 400, loss: 0.08269092440605164
step: 410, loss: 0.017430784180760384
step: 420, loss: 0.22986245155334473
step: 430, loss: 0.023962536826729774
step: 440, loss: 0.0009949398227036
step: 450, loss: 0.01313697174191475
step: 460, loss: 0.009588822722434998
step: 470, loss: 0.13237397372722626
step: 480, loss: 0.03584165871143341
step: 490, loss: 0.03109557367861271
step: 500, loss: 0.0063224975019693375
step: 510, loss: 0.001752092270180583
step: 520, loss: 0.028596477583050728
step: 530, loss: 0.056013643741607666
/usr/local/lib/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=0.9441860465116279, f1=0.9459459459459459,
best_f1=0.9459459459459459

```

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

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

```
step: 0, loss: 0.027730001136660576  
step: 10, loss: 0.0010029241675511003  
step: 20, loss: 0.0020240293815732002  
step: 30, loss: 0.0022503535728901625  
step: 40, loss: 0.0018492323579266667  
step: 50, loss: 0.0002623196633066982  
step: 60, loss: 0.000926628359593451  
step: 70, loss: 0.00017194345127791166  
step: 80, loss: 0.00035989086609333754  
step: 90, loss: 0.020938169211149216  
step: 100, loss: 0.005359154660254717  
step: 110, loss: 0.0020932441111654043  
step: 120, loss: 0.006843896582722664  
step: 130, loss: 0.02937605232000351  
step: 140, loss: 0.0015153357526287436  
step: 150, loss: 0.0006472717504948378  
step: 160, loss: 0.03017158806324005  
step: 170, loss: 0.0009434616658836603  
step: 180, loss: 0.0011555773671716452  
step: 190, loss: 0.14617039263248444  
step: 200, loss: 0.019945688545703888  
step: 210, loss: 0.004170006141066551  
step: 220, loss: 0.003073420375585556  
step: 230, loss: 0.020184693858027458  
step: 240, loss: 0.0031038899905979633  
step: 250, loss: 0.033016253262758255  
step: 260, loss: 0.0008701165788806975  
step: 270, loss: 0.0012432244839146733  
step: 280, loss: 0.08259213715791702  
step: 290, loss: 0.0015887421322986484  
step: 300, loss: 0.002627661218866706  
step: 310, loss: 0.001544304541312158  
step: 320, loss: 0.0009005212923511863  
step: 330, loss: 0.00020416968618519604  
step: 340, loss: 0.00042720482451841235  
step: 350, loss: 0.009860367514193058  
step: 360, loss: 0.009875779040157795  
step: 370, loss: 0.0054396395571529865  
step: 380, loss: 0.003170388052240014  
step: 390, loss: 0.0017430087318643928  
step: 400, loss: 0.0008122275467030704  
step: 410, loss: 0.0013594237389042974  
step: 420, loss: 0.0017618769779801369  
step: 430, loss: 0.0383133664727211  
step: 440, loss: 0.00476841302588582
```

```
step: 450, loss: 0.151334747672081
step: 460, loss: 0.0026172969955950975
step: 470, loss: 0.001639084191992879
step: 480, loss: 0.007662672083824873
step: 490, loss: 0.0027061845175921917
step: 500, loss: 0.0028598057106137276
step: 510, loss: 0.20600421726703644
step: 520, loss: 0.0040522669441998005
step: 530, loss: 0.0046122511848807335
/usr/local/lib/python3.7/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=0.9432029795158287, f1=0.940795559666975,
best_f1=0.9459459459459459
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0051259128376841545
step: 10, loss: 0.002780412556603551
step: 20, loss: 0.007880650460720062
step: 30, loss: 0.0036861563567072153
step: 40, loss: 0.0022542388178408146
step: 50, loss: 0.004514184780418873
step: 60, loss: 0.03051207959651947
step: 70, loss: 0.0013830700190737844
step: 80, loss: 0.0008888870943337679
step: 90, loss: 0.00038524382398463786
step: 100, loss: 0.0006251009763218462
step: 110, loss: 0.0002523771545384079
step: 120, loss: 0.0004433807043824345
step: 130, loss: 0.00012028997298330069
step: 140, loss: 0.0005535971140488982
step: 150, loss: 0.010228010825812817
step: 160, loss: 0.0059777782298624516
step: 170, loss: 0.0033657292369753122
step: 180, loss: 0.09715139120817184
step: 190, loss: 0.0017219781875610352
step: 200, loss: 0.00038920415681786835
step: 210, loss: 0.002092239446938038
step: 220, loss: 0.000694401329383254
step: 230, loss: 0.0012954944977536798
step: 240, loss: 0.043064992874860764
step: 250, loss: 0.04129005968570709
step: 260, loss: 0.0023542619310319424
step: 270, loss: 0.0005683128256350756
step: 280, loss: 0.004024135880172253
step: 290, loss: 0.002661231206730008
step: 300, loss: 0.005510171875357628
```

```

step: 310, loss: 0.0007091195438988507
step: 320, loss: 0.0018499352736398578
step: 330, loss: 0.009631969965994358
step: 340, loss: 0.001002300065010786
step: 350, loss: 0.010273615829646587
step: 360, loss: 0.08686962723731995
step: 370, loss: 0.09962131828069687
step: 380, loss: 0.02392657846212387
step: 390, loss: 0.002891409443691373
step: 400, loss: 0.007562188897281885
step: 410, loss: 0.0003456571139395237
step: 420, loss: 0.026990283280611038
step: 430, loss: 0.011140036396682262
step: 440, loss: 0.0063084447756409645
step: 450, loss: 0.0021313356701284647
step: 460, loss: 0.0001628733880352229
step: 470, loss: 0.10077625513076782
step: 480, loss: 0.011078369803726673
step: 490, loss: 0.0010809262748807669
step: 500, loss: 0.0006910167867317796
step: 510, loss: 0.0001496447075624019
step: 520, loss: 0.09142327308654785
step: 530, loss: 0.0006702787359245121
/usr/local/lib/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=0.9490740740740741, f1=0.9365446966188051,
best_f1=0.9365446966188051
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.000932642724364996
step: 10, loss: 0.014652889221906662
step: 20, loss: 0.021390248090028763
step: 30, loss: 0.0009695531334728003
step: 40, loss: 0.0004427703970577568
step: 50, loss: 0.0010849999962374568
step: 60, loss: 0.0005626128404401243
step: 70, loss: 0.0006735817296430469
step: 80, loss: 0.0004348883521743119
step: 90, loss: 0.0035218987613916397
step: 100, loss: 0.0002583850873634219
step: 110, loss: 0.00014288508100435138
step: 120, loss: 0.0002589219366200268

```

```

step: 130, loss: 0.0002381915837759152
step: 140, loss: 0.00015853032527957112
step: 150, loss: 0.0008261111797764897
step: 160, loss: 0.00011663122131722048
step: 170, loss: 0.0021721196826547384
step: 180, loss: 0.0007750015356577933
step: 190, loss: 0.001663114526309073
step: 200, loss: 0.007908567786216736
step: 210, loss: 0.1423148810863495
step: 220, loss: 0.0020215471740812063
step: 230, loss: 0.011040530167520046
step: 240, loss: 0.0031894019339233637
step: 250, loss: 0.0021885009482502937
step: 260, loss: 0.0009186223614960909
step: 270, loss: 0.001152166398242116
step: 280, loss: 0.0013692151987925172
step: 290, loss: 0.0003481886233203113
step: 300, loss: 0.00010423688945593312
step: 310, loss: 0.00022196384088601917
step: 320, loss: 0.00029474368784576654
step: 330, loss: 0.0582464300096035
step: 340, loss: 0.09289698302745819
step: 350, loss: 0.00469879200682044
step: 360, loss: 0.004297854844480753
step: 370, loss: 0.007460962515324354
step: 380, loss: 0.00047640452976338565
step: 390, loss: 0.0024372972548007965
step: 400, loss: 0.0014916092623025179
step: 410, loss: 0.0009466689662076533
step: 420, loss: 0.0004290770157240331
step: 430, loss: 0.006396720185875893
step: 440, loss: 0.006073642522096634
step: 450, loss: 0.001780266291461885
step: 460, loss: 0.00101346499286592
step: 470, loss: 0.24781650304794312
step: 480, loss: 0.007225663401186466
step: 490, loss: 0.0052864584140479565
step: 500, loss: 0.07084193080663681
step: 510, loss: 0.0011537447571754456
step: 520, loss: 0.0005910302279517055
step: 530, loss: 0.01038189698010683
/usr/local/lib/python3.7/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=0.9490268767377201, f1=0.9460083064143978,
best_f1=0.9365446966188051
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.

```



```
warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0013654688373208046
step: 10, loss: 0.0042361728847026825
step: 20, loss: 0.00032978522358462214
step: 30, loss: 0.09450706094503403
step: 40, loss: 0.001988018164411187
step: 50, loss: 0.0014558570692315698
step: 60, loss: 0.005543766543269157
step: 70, loss: 0.00649621756747365
step: 80, loss: 0.0048213740810751915
step: 90, loss: 0.02882956527173519
step: 100, loss: 0.0007187425508163869
step: 110, loss: 0.0075227306224405766
step: 120, loss: 0.0005896607181057334
step: 130, loss: 0.00035219627898186445
step: 140, loss: 0.000331234245095402
step: 150, loss: 0.0006696748896501958
step: 160, loss: 0.0004739435389637947
step: 170, loss: 0.008558874949812889
step: 180, loss: 0.00014940112305339426
step: 190, loss: 0.0004516390326898545
step: 200, loss: 0.00015169694961514324
step: 210, loss: 0.0007835079450160265
step: 220, loss: 0.0005790751310996711
step: 230, loss: 0.0006691170856356621
step: 240, loss: 0.0012815133668482304
step: 250, loss: 0.03988127410411835
step: 260, loss: 0.0013586211716756225
step: 270, loss: 0.0031863730400800705
step: 280, loss: 0.032487835735082626
step: 290, loss: 0.0002299973857589066
step: 300, loss: 0.00215065386146307
step: 310, loss: 0.006178102921694517
step: 320, loss: 0.0005901475669816136
step: 330, loss: 0.0017599075799807906
step: 340, loss: 0.005171274766325951
step: 350, loss: 0.06132163852453232
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 8192.0
step: 360, loss: 0.003724693786352873
step: 370, loss: 0.0020674136467278004
step: 380, loss: 0.0005881093093194067
step: 390, loss: 0.00782088190317154
step: 400, loss: 0.0004850148980040103
step: 410, loss: 0.000537665793672204
step: 420, loss: 0.0005395965999923646
step: 430, loss: 0.0005092878709547222
step: 440, loss: 9.102689364226535e-05
step: 450, loss: 0.003227694658562541
```

```
step: 460, loss: 0.0032601249404251575
step: 470, loss: 0.0006130889523774385
step: 480, loss: 0.0006655485485680401
step: 490, loss: 0.00088942312868312
step: 500, loss: 0.0018806083826348186
step: 510, loss: 0.0035992173943668604
step: 520, loss: 0.010503503493964672
step: 530, loss: 0.0032467786222696304
/usr/local/lib/python3.7/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=0.9461966604823748, f1=0.9439555349698935,
best_f1=0.9365446966188051
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0008057790109887719
step: 10, loss: 0.0005867477739229798
step: 20, loss: 0.015679210424423218
step: 30, loss: 0.0013476255116984248
step: 40, loss: 0.0034700389951467514
step: 50, loss: 0.0011726301163434982
step: 60, loss: 0.00025851192185655236
step: 70, loss: 0.000540428445674479
step: 80, loss: 0.0005714800790883601
step: 90, loss: 0.0005844431580044329
step: 100, loss: 0.0013156337663531303
step: 110, loss: 0.014932222664356232
step: 120, loss: 0.00015237851766869426
step: 130, loss: 0.011869003996253014
step: 140, loss: 0.0002952419454231858
step: 150, loss: 0.00048708554822951555
step: 160, loss: 0.08077961206436157
step: 170, loss: 0.006189426872879267
step: 180, loss: 0.0007740459404885769
step: 190, loss: 0.0001334466942353174
step: 200, loss: 0.0002972946094814688
step: 210, loss: 0.009283735416829586
step: 220, loss: 0.0011981434654444456
step: 230, loss: 0.00038043843233026564
step: 240, loss: 0.000693033856805414
step: 250, loss: 0.0002524142328184098
step: 260, loss: 0.0008327924879267812
step: 270, loss: 0.0004448821709956974
step: 280, loss: 0.016293341293931007
step: 290, loss: 0.0035308152437210083
step: 300, loss: 0.011249348521232605
step: 310, loss: 0.0285439845174551
```

```

step: 320, loss: 0.0023083891719579697
step: 330, loss: 0.0014003815595060587
step: 340, loss: 6.0025024140486494e-05
step: 350, loss: 0.0004916936159133911
step: 360, loss: 7.383363845292479e-05
step: 370, loss: 0.0006360887200571597
step: 380, loss: 0.00026735535357147455
step: 390, loss: 0.00014494518109131604
step: 400, loss: 0.0002489531470928341
step: 410, loss: 0.00025896611623466015
step: 420, loss: 5.4185729823075235e-05
step: 430, loss: 0.0011101111304014921
step: 440, loss: 0.0012645383831113577
step: 450, loss: 0.0005774840246886015
step: 460, loss: 0.00021810819453094155
step: 470, loss: 0.0007116320775821805
step: 480, loss: 0.00013383050099946558
step: 490, loss: 0.03424801677465439
step: 500, loss: 0.000845833623316139
step: 510, loss: 0.0003200212959200144
step: 520, loss: 0.007259285543113947
step: 530, loss: 0.001687456271611154
/usr/local/lib/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=0.9559099437148217, f1=0.9451476793248945,
best_f1=0.9451476793248945
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.002526551252231002
step: 10, loss: 0.0003889760992024094
step: 20, loss: 0.00023121933918446302
step: 30, loss: 0.00022327392071019858
step: 40, loss: 0.0001136160281021148
step: 50, loss: 9.494777623331174e-05
step: 60, loss: 7.754279795335606e-05
step: 70, loss: 0.00799601711332798
step: 80, loss: 0.0001141631364589557
step: 90, loss: 0.0004133519541937858
step: 100, loss: 0.0006734919152222574
step: 110, loss: 0.0004093205789104104
step: 120, loss: 0.00020717098959721625
step: 130, loss: 0.0001841755147324875

```

```

step: 140, loss: 7.021163037279621e-05
step: 150, loss: 0.004263550043106079
step: 160, loss: 0.04631773754954338
step: 170, loss: 0.0002320078492630273
step: 180, loss: 0.0007078485796228051
step: 190, loss: 0.012806098908185959
step: 200, loss: 0.00015793894999660552
step: 210, loss: 0.00022495724260807037
step: 220, loss: 0.01759902946650982
step: 230, loss: 7.609061140101403e-05
step: 240, loss: 0.0034493200946599245
step: 250, loss: 0.00031327197211794555
step: 260, loss: 0.000759469170589
step: 270, loss: 0.0030973409302532673
step: 280, loss: 0.001569350017234683
step: 290, loss: 0.004438537172973156
step: 300, loss: 0.00020780088379979134
step: 310, loss: 0.010525580495595932
step: 320, loss: 0.0003038956201635301
step: 330, loss: 0.0003938609443139285
step: 340, loss: 0.0007320090662688017
step: 350, loss: 0.011606894433498383
step: 360, loss: 0.0033932740334421396
step: 370, loss: 0.0016285997116938233
step: 380, loss: 0.002255247673019767
step: 390, loss: 0.00031730299815535545
step: 400, loss: 0.0038606866728514433
step: 410, loss: 0.0015621044440194964
step: 420, loss: 0.0006005085306242108
step: 430, loss: 0.0022311778739094734
step: 440, loss: 0.00030986653291620314
step: 450, loss: 0.0005935419467277825
step: 460, loss: 0.00040541135240346193
step: 470, loss: 0.0007750296499580145
step: 480, loss: 0.0006148662650957704
step: 490, loss: 0.001951193786226213
step: 500, loss: 0.0009467832860536873
step: 510, loss: 0.012675965204834938
step: 520, loss: 0.0028960651252418756
step: 530, loss: 0.006072956137359142
/usr/local/lib/python3.7/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=0.9524253731343284, f1=0.9361305361305362,
best_f1=0.9451476793248945
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")

```

step: 0, loss: 0.00046989330439828336  
step: 10, loss: 0.00025482228375039995  
step: 20, loss: 0.0004381704202387482  
step: 30, loss: 6.06147623329889e-05  
step: 40, loss: 0.00011781398643506691  
step: 50, loss: 0.0004548810247797519  
step: 60, loss: 9.196218161378056e-05  
step: 70, loss: 0.001957214204594493  
step: 80, loss: 0.0005391589365899563  
step: 90, loss: 0.00030514979152940214  
step: 100, loss: 0.19653815031051636  
step: 110, loss: 0.0016945756506174803  
step: 120, loss: 0.0034504348877817392  
step: 130, loss: 0.0006271819001995027  
step: 140, loss: 0.0003280468808952719  
step: 150, loss: 7.942577940411866e-05  
step: 160, loss: 0.0007953417371027172  
step: 170, loss: 0.0001354240084765479  
step: 180, loss: 0.00012133840937167406  
step: 190, loss: 0.00014862576790619642  
step: 200, loss: 0.004273626487702131  
step: 210, loss: 0.00038500470691360533  
step: 220, loss: 0.00020302488701418042  
step: 230, loss: 0.00010490958084119484  
step: 240, loss: 0.0009187753894366324  
step: 250, loss: 7.427745003951713e-05  
step: 260, loss: 0.0001789149537216872  
step: 270, loss: 0.0007522394298575819  
step: 280, loss: 0.00015133156557567418  
step: 290, loss: 0.00010093392484122887  
step: 300, loss: 0.00021585194917861372  
step: 310, loss: 0.007106791716068983  
step: 320, loss: 0.0001828910899348557  
step: 330, loss: 0.0007987716817297041  
step: 340, loss: 0.001271153916604817  
step: 350, loss: 0.00013121233496349305  
step: 360, loss: 0.00011225731577724218  
step: 370, loss: 0.00011145209282403812  
step: 380, loss: 0.00012458219134714454  
step: 390, loss: 0.0014147255569696426  
step: 400, loss: 0.00010249249316984788  
step: 410, loss: 0.00043131865095347166  
step: 420, loss: 0.00025884638307616115  
step: 430, loss: 0.0008074079523794353  
step: 440, loss: 0.005629927385598421  
step: 450, loss: 0.0004341899766586721  
step: 460, loss: 0.000223902243305929  
step: 470, loss: 0.004696945194154978

```

step: 480, loss: 0.007911836728453636
step: 490, loss: 0.0002817553759086877
step: 500, loss: 4.760422234539874e-05
step: 510, loss: 0.0012824343284592032
step: 520, loss: 0.0010365542257204652
step: 530, loss: 8.395557961193845e-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=0.9544408651633686, f1=0.944649446494465,
best_f1=0.9451476793248945
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 4.807809091289528e-05
step: 10, loss: 0.00013843612396158278
step: 20, loss: 5.214686825638637e-05
step: 30, loss: 3.5600245610112324e-05
step: 40, loss: 0.00012020234862575307
step: 50, loss: 0.0007354166591539979
step: 60, loss: 0.00012188324035378173
step: 70, loss: 0.00015662013902328908
step: 80, loss: 0.00015086514758877456
step: 90, loss: 0.00012435272219590843
step: 100, loss: 9.563824278302491e-05
step: 110, loss: 8.157412958098575e-05
step: 120, loss: 8.669335511513054e-05
step: 130, loss: 0.00013254574150778353
step: 140, loss: 9.143104398390278e-05
step: 150, loss: 0.004373867996037006
step: 160, loss: 0.0004996961215510964
step: 170, loss: 0.0015142099000513554
step: 180, loss: 8.26673349365592e-05
step: 190, loss: 0.00023607671028003097
step: 200, loss: 0.00019094179151579738
step: 210, loss: 0.00021188758546486497
step: 220, loss: 6.0860824305564165e-05
step: 230, loss: 0.001048530451953411
step: 240, loss: 0.0023516230285167694
step: 250, loss: 0.00015002096188254654
step: 260, loss: 0.021919531747698784
step: 270, loss: 0.00022942826035432518
step: 280, loss: 7.54507418605499e-05
step: 290, loss: 7.90016638347879e-05
step: 300, loss: 0.0354212187230587
step: 310, loss: 0.00011167916818521917
step: 320, loss: 0.00010242947610095143
step: 330, loss: 0.0007350342930294573

```

```

step: 340, loss: 9.28804074646905e-05
step: 350, loss: 0.0007301230798475444
step: 360, loss: 0.07772528380155563
step: 370, loss: 0.026941582560539246
step: 380, loss: 0.00038404547376558185
step: 390, loss: 0.00014420058869291097
step: 400, loss: 0.00014804632519371808
step: 410, loss: 0.00021330261370167136
step: 420, loss: 0.00010690921772038564
step: 430, loss: 0.0009671669686213136
step: 440, loss: 5.823792889714241e-05
step: 450, loss: 0.00011336689203744754
step: 460, loss: 0.0005856252973899245
step: 470, loss: 0.0028226415161043406
step: 480, loss: 3.263584949309006e-05
step: 490, loss: 0.0001699943677522242
step: 500, loss: 0.0003416196268517524
step: 510, loss: 0.00039307516999542713
step: 520, loss: 6.970463437028229e-05
step: 530, loss: 0.00018117961008101702
/usr/local/lib/python3.7/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=0.9533146591970122, f1=0.9463369108726085,
best_f1=0.9451476793248945
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00024880070122890174
step: 10, loss: 0.00019306466856505722
step: 20, loss: 3.9476173697039485e-05
step: 30, loss: 0.002406616695225239
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 8192.0
step: 40, loss: 6.222397496458143e-05
step: 50, loss: 0.005626068450510502
step: 60, loss: 5.537055403692648e-05
step: 70, loss: 5.391345621319488e-05
step: 80, loss: 0.00017224899784196168
step: 90, loss: 4.537480344879441e-05
step: 100, loss: 8.576217805966735e-05
step: 110, loss: 4.7146448196144775e-05
step: 120, loss: 4.0208262362284586e-05
step: 130, loss: 5.063408752903342e-05
step: 140, loss: 0.00031048653181642294
step: 150, loss: 0.0003833411610685289
step: 160, loss: 5.021879405830987e-05
step: 170, loss: 0.00012244886602275074
step: 180, loss: 7.994702900759876e-05

```

```

step: 190, loss: 0.0003200224309694022
step: 200, loss: 0.00012917126878164709
step: 210, loss: 7.50578474253416e-05
step: 220, loss: 3.635916073108092e-05
step: 230, loss: 4.921030267723836e-05
step: 240, loss: 4.062103835167363e-05
step: 250, loss: 0.00038533174665644765
step: 260, loss: 4.887238537776284e-05
step: 270, loss: 0.0006442677695304155
step: 280, loss: 0.00010783699690364301
step: 290, loss: 3.528840534272604e-05
step: 300, loss: 0.00013871790724806488
step: 310, loss: 0.0003323113196529448
step: 320, loss: 8.817019988782704e-05
step: 330, loss: 0.0033119446597993374
step: 340, loss: 7.194100908236578e-05
step: 350, loss: 2.7204290745430626e-05
step: 360, loss: 4.2393749026814476e-05
step: 370, loss: 0.0005207439535297453
step: 380, loss: 0.0018338047666475177
step: 390, loss: 0.004203092772513628
step: 400, loss: 0.0026678605936467648
step: 410, loss: 2.3306258299271576e-05
step: 420, loss: 0.00010771519009722397
step: 430, loss: 0.0007240719860419631
step: 440, loss: 0.00028009136440232396
step: 450, loss: 3.682517490233295e-05
step: 460, loss: 0.00043287183507345617
step: 470, loss: 4.665169399231672e-05
step: 480, loss: 5.087145109428093e-05
step: 490, loss: 0.002382913837209344
step: 500, loss: 0.0003030076331924647
step: 510, loss: 0.00030936990515328944
step: 520, loss: 4.182835255051032e-05
step: 530, loss: 0.00013785381452180445
/usr/local/lib/python3.7/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=0.9541029207232267, f1=0.9476124246638852,
best_f1=0.9451476793248945
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 9.323858830612153e-05
step: 10, loss: 7.091977749951184e-05
step: 20, loss: 0.0007313434034585953
step: 30, loss: 0.007978050038218498
step: 40, loss: 3.855051181744784e-05

```



step: 50, loss: 0.004538755398243666  
step: 60, loss: 0.0012263285461813211  
step: 70, loss: 4.762950266012922e-05  
step: 80, loss: 5.223086918704212e-05  
step: 90, loss: 5.3880597988609225e-05  
step: 100, loss: 3.982616544817574e-05  
step: 110, loss: 0.00024237690377049148  
step: 120, loss: 8.682003681315109e-05  
step: 130, loss: 5.5739361414453015e-05  
step: 140, loss: 5.055499786976725e-05  
step: 150, loss: 6.221527291927487e-05  
step: 160, loss: 0.00013637995289172977  
step: 170, loss: 1.9720499039976858e-05  
step: 180, loss: 0.00033333527971990407  
step: 190, loss: 0.004386389162391424  
step: 200, loss: 0.0006553312414325774  
step: 210, loss: 0.00023238245921675116  
step: 220, loss: 4.6574485168093815e-05  
step: 230, loss: 4.6905588533263654e-05  
step: 240, loss: 6.799285620218143e-05  
step: 250, loss: 2.2573580281459726e-05  
step: 260, loss: 2.654081617947668e-05  
step: 270, loss: 0.0008523553842678666  
step: 280, loss: 7.203916902653873e-05  
step: 290, loss: 6.319824751699343e-05  
step: 300, loss: 2.3486754798796028e-05  
step: 310, loss: 0.04995815083384514  
step: 320, loss: 4.727295890916139e-05  
step: 330, loss: 0.00010463776561664417  
step: 340, loss: 5.317356044542976e-05  
step: 350, loss: 0.001258639502339065  
step: 360, loss: 4.145497950958088e-05  
step: 370, loss: 8.33231897559017e-05  
step: 380, loss: 5.0286722398595884e-05  
step: 390, loss: 0.0004725692269857973  
step: 400, loss: 0.0022329941857606173  
step: 410, loss: 0.00016578730719629675  
step: 420, loss: 2.8250611649127677e-05  
step: 430, loss: 1.676653300819453e-05  
step: 440, loss: 6.444154132623225e-05  
step: 450, loss: 0.0014136980753391981  
step: 460, loss: 0.0002189625083701685  
step: 470, loss: 3.066466160817072e-05  
step: 480, loss: 0.0022803654428571463  
step: 490, loss: 4.1245311876991764e-05  
step: 500, loss: 0.00042567020864225924  
step: 510, loss: 3.882153396261856e-05  
step: 520, loss: 7.593755435664207e-05

```

step: 530, loss: 2.499830225133337e-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=0.9547785547785548, f1=0.9471221338324755,
best_f1=0.9451476793248945
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 [00:22, 251.77it/s]
load_f1 = 0.9544186046511628
real_f1 = 0.951366373320982
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, 208.28it/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: 0.5256560444831848
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 10, loss: 0.5046546459197998
step: 20, loss: 0.41702619194984436
step: 30, loss: 0.41729527711868286
step: 40, loss: 0.32436081767082214
step: 50, loss: 0.4694766402244568
step: 60, loss: 0.44174709916114807
step: 70, loss: 0.3558689057826996
step: 80, loss: 0.387315034866333
step: 90, loss: 0.26610028743743896
step: 100, loss: 0.22170816361904144
step: 110, loss: 0.2703292965888977
step: 120, loss: 0.49102187156677246
step: 130, loss: 0.2604563534259796
step: 140, loss: 0.4658019244670868
step: 150, loss: 0.3830311596393585
step: 160, loss: 0.5066589713096619
step: 170, loss: 0.24289780855178833
step: 180, loss: 0.38947343826293945
step: 190, loss: 0.6326875686645508
step: 200, loss: 0.4032491445541382
step: 210, loss: 0.5167878270149231
/usr/local/lib/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: 0.3795267939567566
step: 10, loss: 0.20477545261383057
step: 20, loss: 0.4814799726009369
step: 30, loss: 0.47475650906562805
step: 40, loss: 0.43443700671195984
step: 50, loss: 0.23990701138973236
step: 60, loss: 0.30496153235435486
step: 70, loss: 0.4190770089626312
step: 80, loss: 0.4001544117927551
step: 90, loss: 0.45558497309684753
step: 100, loss: 0.49704673886299133
step: 110, loss: 0.38014549016952515
step: 120, loss: 0.25123053789138794
step: 130, loss: 0.16730180382728577
step: 140, loss: 0.24157042801380157
step: 150, loss: 0.44513070583343506
step: 160, loss: 0.15692442655563354
step: 170, loss: 0.6010239720344543
step: 180, loss: 0.3245784044265747
step: 190, loss: 0.30898019671440125
step: 200, loss: 0.15316303074359894
step: 210, loss: 0.31064727902412415
/usr/local/lib/python3.7/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=0.18519984170953702, 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: 0.25453004240989685
step: 10, loss: 0.23234710097312927
step: 20, loss: 0.4750593304634094
step: 30, loss: 0.25556230545043945
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 8192.0
step: 40, loss: 0.46512728929519653
step: 50, loss: 0.505531907081604
step: 60, loss: 0.5045008659362793

```

```

step: 70, loss: 0.18389621376991272
step: 80, loss: 0.477409303188324
step: 90, loss: 0.25102514028549194
step: 100, loss: 0.3845066428184509
step: 110, loss: 0.2589631974697113
step: 120, loss: 0.2504101097583771
step: 130, loss: 0.15915681421756744
step: 140, loss: 0.36631473898887634
step: 150, loss: 0.3260229825973511
step: 160, loss: 0.2048550248146057
step: 170, loss: 0.3828948140144348
step: 180, loss: 0.24429501593112946
step: 190, loss: 0.166031613945961
step: 200, loss: 0.23814575374126434
step: 210, loss: 0.269571453332901
/usr/local/lib/python3.7/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=0.18519984170953702, 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: 0.3219221830368042
step: 10, loss: 0.3103790283203125
step: 20, loss: 0.32134369015693665
step: 30, loss: 0.2388373762369156
step: 40, loss: 0.23847505450248718
step: 50, loss: 0.241471529006958
step: 60, loss: 0.493104487657547
step: 70, loss: 0.2515726387500763
step: 80, loss: 0.24514512717723846
step: 90, loss: 0.47127532958984375
step: 100, loss: 0.37451601028442383
step: 110, loss: 0.5925753116607666
step: 120, loss: 0.35345786809921265
step: 130, loss: 0.6620739102363586
step: 140, loss: 0.5068488717079163
step: 150, loss: 0.388830304145813
step: 160, loss: 0.3318195044994354
step: 170, loss: 0.17553310096263885
step: 180, loss: 0.07519423961639404
step: 190, loss: 0.16405481100082397
step: 200, loss: 0.316340833902359
step: 210, loss: 0.3979664444923401
/usr/local/lib/python3.7/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=0.18519984170953702, 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: 0.3838878870010376
step: 10, loss: 0.3028911054134369
step: 20, loss: 0.26162728667259216
step: 30, loss: 0.210226908326149
step: 40, loss: 0.4070267379283905
step: 50, loss: 0.41737720370292664
step: 60, loss: 0.3614498972892761
step: 70, loss: 0.23278431594371796
step: 80, loss: 0.3515598475933075
step: 90, loss: 0.4606224000453949
step: 100, loss: 0.23534250259399414
step: 110, loss: 0.16430963575839996
step: 120, loss: 0.24305568635463715
step: 130, loss: 0.27656394243240356
step: 140, loss: 0.5465062856674194
step: 150, loss: 0.3270464837551117
step: 160, loss: 0.2447238266468048
step: 170, loss: 0.38959574699401855
step: 180, loss: 0.24110472202301025
step: 190, loss: 0.5077227354049683
step: 200, loss: 0.3623637557029724
step: 210, loss: 0.2357654869556427
/usr/local/lib/python3.7/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=0.18519984170953702, 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: 0.15676714479923248
step: 10, loss: 0.34317460656166077
step: 20, loss: 0.3134663701057434
step: 30, loss: 0.2398395836353302
step: 40, loss: 0.2715475261211395
step: 50, loss: 0.5287100672721863
step: 60, loss: 0.2296798825263977
step: 70, loss: 0.3795630633831024
step: 80, loss: 0.2594280242919922
step: 90, loss: 0.2961295247077942
step: 100, loss: 0.3499734103679657
step: 110, loss: 0.3208400309085846
step: 120, loss: 0.2070361077785492

```

```

step: 130, loss: 0.21928708255290985
step: 140, loss: 0.37782102823257446
step: 150, loss: 0.18218636512756348
step: 160, loss: 0.19533614814281464
step: 170, loss: 0.3731187880039215
step: 180, loss: 0.4076756238937378
step: 190, loss: 0.33035409450531006
step: 200, loss: 0.27843624353408813
step: 210, loss: 0.30935221910476685
/usr/local/lib/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=0.28797468354430383, f1=0.28892455858747995,
best_f1=0.28892455858747995
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.43738850951194763
step: 10, loss: 0.27474358677864075
step: 20, loss: 0.3851282596588135
step: 30, loss: 0.20236268639564514
step: 40, loss: 0.23509439826011658
step: 50, loss: 0.3198070228099823
step: 60, loss: 0.2715081572532654
step: 70, loss: 0.23963260650634766
step: 80, loss: 0.42127588391304016
step: 90, loss: 0.3719950318336487
step: 100, loss: 0.39628615975379944
step: 110, loss: 0.2812681496143341
step: 120, loss: 0.2726941406726837
step: 130, loss: 0.4120584726333618
step: 140, loss: 0.3174207806587219
step: 150, loss: 0.2946838438510895
step: 160, loss: 0.5371550917625427
step: 170, loss: 0.5601440072059631
step: 180, loss: 0.20696474611759186
step: 190, loss: 0.2968573272228241
step: 200, loss: 0.29638367891311646
step: 210, loss: 0.29332712292671204
/usr/local/lib/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=0.3028169014084507, f1=0.2998236331569665,
best_f1=0.2998236331569665
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.5124892592430115
step: 10, loss: 0.36760661005973816
step: 20, loss: 0.19725148379802704
step: 30, loss: 0.2140347957611084
step: 40, loss: 0.24670933187007904
step: 50, loss: 0.14087416231632233
step: 60, loss: 0.18135428428649902
step: 70, loss: 0.34444424510002136
step: 80, loss: 0.26108816266059875
step: 90, loss: 0.33581453561782837
step: 100, loss: 0.5297002792358398
step: 110, loss: 0.2961474359035492
step: 120, loss: 0.30617567896842957
step: 130, loss: 0.11538037657737732
step: 140, loss: 0.3177794814109802
step: 150, loss: 0.532988429069519
step: 160, loss: 0.40076714754104614
step: 170, loss: 0.4823712706565857
step: 180, loss: 0.21328765153884888
step: 190, loss: 0.1865435689687729
step: 200, loss: 0.3850748538970947
step: 210, loss: 0.34763476252555847
/usr/local/lib/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=0.32232496697490093, f1=0.3289817232375979,
best_f1=0.3289817232375979
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.4705325961112976
step: 10, loss: 0.19493147730827332
step: 20, loss: 0.3726638853549957
step: 30, loss: 0.13524852693080902
step: 40, loss: 0.12701833248138428
step: 50, loss: 0.4922756552696228
step: 60, loss: 0.15215672552585602

```

```

step: 70, loss: 0.3523675203323364
step: 80, loss: 0.2666992247104645
step: 90, loss: 0.3218724727630615
step: 100, loss: 0.3832259774208069
step: 110, loss: 0.4873022437095642
step: 120, loss: 0.4201198220252991
step: 130, loss: 0.20984207093715668
step: 140, loss: 0.47306570410728455
step: 150, loss: 0.11819647997617722
step: 160, loss: 0.3191278576850891
step: 170, loss: 0.27250632643699646
step: 180, loss: 0.4602178931236267
step: 190, loss: 0.2265533208847046
step: 200, loss: 0.2479306310415268
step: 210, loss: 0.2392304688692093
/usr/local/lib/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=0.3401506996770721, f1=0.3307607497243661,
best_f1=0.3307607497243661
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.23555728793144226
step: 10, loss: 0.44004198908805847
step: 20, loss: 0.238517165184021
step: 30, loss: 0.21025805175304413
step: 40, loss: 0.3485206961631775
step: 50, loss: 0.38766616582870483
step: 60, loss: 0.15969495475292206
step: 70, loss: 0.3868013322353363
step: 80, loss: 0.13976719975471497
step: 90, loss: 0.3735732138156891
step: 100, loss: 0.41184669733047485
step: 110, loss: 0.13727697730064392
step: 120, loss: 0.34281864762306213
step: 130, loss: 0.1684083789587021
step: 140, loss: 0.3075522184371948
step: 150, loss: 0.2079710215330124
step: 160, loss: 0.2439766824245453
step: 170, loss: 0.1823650449514389
step: 180, loss: 0.28792330622673035
step: 190, loss: 0.17614398896694183
step: 200, loss: 0.5507678389549255

```

```

step: 210, loss: 0.20344777405261993
/usr/local/lib/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=0.3529411764705882, f1=0.3934426229508197,
best_f1=0.3934426229508197
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.21694974601268768
step: 10, loss: 0.3366050124168396
step: 20, loss: 0.2545435428619385
step: 30, loss: 0.18532642722129822
step: 40, loss: 0.39878466725349426
step: 50, loss: 0.39481401443481445
step: 60, loss: 0.23125791549682617
step: 70, loss: 0.1884276419878006
step: 80, loss: 0.29818829894065857
step: 90, loss: 0.3882178068161011
step: 100, loss: 0.40720197558403015
step: 110, loss: 0.448878675699234
step: 120, loss: 0.31517890095710754
step: 130, loss: 0.1864912062883377
step: 140, loss: 0.2854679524898529
step: 150, loss: 0.28265607357025146
step: 160, loss: 0.13969376683235168
step: 170, loss: 0.21194937825202942
step: 180, loss: 0.21708986163139343
step: 190, loss: 0.4463649392127991
step: 200, loss: 0.16457191109657288
step: 210, loss: 0.3442610502243042
/usr/local/lib/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=0.3695652173913043, f1=0.371571072319202,
best_f1=0.371571072319202
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.23244182765483856

```

```

step: 10, loss: 0.25150302052497864
step: 20, loss: 0.4044765532016754
step: 30, loss: 0.24890559911727905
step: 40, loss: 0.11383330076932907
step: 50, loss: 0.48832404613494873
step: 60, loss: 0.10398933291435242
step: 70, loss: 0.3661366105079651
step: 80, loss: 0.27194029092788696
step: 90, loss: 0.44080016016960144
step: 100, loss: 0.05651477724313736
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 8192.0
step: 110, loss: 0.25310537219047546
step: 120, loss: 0.15455995500087738
step: 130, loss: 0.4096345901489258
step: 140, loss: 0.42885246872901917
step: 150, loss: 0.08249706774950027
step: 160, loss: 0.15154163539409637
step: 170, loss: 0.16611728072166443
step: 180, loss: 0.32904887199401855
step: 190, loss: 0.34893715381622314
step: 200, loss: 0.09276235848665237
step: 210, loss: 0.12013549357652664
/usr/local/lib/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=0.4716981132075472, f1=0.5440313111545988,
best_f1=0.5440313111545988
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.1787831038236618
step: 10, loss: 0.10371901839971542
step: 20, loss: 0.1079833060503006
step: 30, loss: 0.08989761769771576
step: 40, loss: 0.16405589878559113
step: 50, loss: 0.2167740911245346
step: 60, loss: 0.38423651456832886
step: 70, loss: 0.1750698834657669
step: 80, loss: 0.2915131151676178
step: 90, loss: 0.35258209705352783
step: 100, loss: 0.2089955061674118
step: 110, loss: 0.20067116618156433
step: 120, loss: 0.19702701270580292
step: 130, loss: 0.13981249928474426

```

```

step: 140, loss: 0.15195083618164062
step: 150, loss: 0.17570459842681885
step: 160, loss: 0.2536356747150421
step: 170, loss: 0.09684058278799057
step: 180, loss: 0.1500312089920044
step: 190, loss: 0.17817027866840363
step: 200, loss: 0.2225930392742157
step: 210, loss: 0.304394006729126
/usr/local/lib/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=0.5621414913957935, f1=0.6242990654205607,
best_f1=0.6242990654205607
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.10010382533073425
step: 10, loss: 0.17755407094955444
step: 20, loss: 0.29133495688438416
step: 30, loss: 0.07378628849983215
step: 40, loss: 0.26404738426208496
step: 50, loss: 0.15320442616939545
step: 60, loss: 0.20733290910720825
step: 70, loss: 0.18909591436386108
step: 80, loss: 0.12395182996988297
step: 90, loss: 0.06848897784948349
step: 100, loss: 0.12066859006881714
step: 110, loss: 0.22648069262504578
step: 120, loss: 0.08058451861143112
step: 130, loss: 0.07471349090337753
step: 140, loss: 0.19448965787887573
step: 150, loss: 0.1707676649093628
step: 160, loss: 0.08018394559621811
step: 170, loss: 0.17534787952899933
step: 180, loss: 0.12549225986003876
step: 190, loss: 0.24002479016780853
step: 200, loss: 0.0587749183177948
step: 210, loss: 0.21132014691829681
/usr/local/lib/python3.7/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=0.5466448445171849, f1=0.598705501618123,
best_f1=0.6242990654205607
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:

```

```

An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.11314454674720764
step: 10, loss: 0.15695719420909882
step: 20, loss: 0.38505110144615173
step: 30, loss: 0.09841132164001465
step: 40, loss: 0.24590028822422028
step: 50, loss: 0.0736171305179596
step: 60, loss: 0.18609566986560822
step: 70, loss: 0.21023496985435486
step: 80, loss: 0.19693569839000702
step: 90, loss: 0.2136196792125702
step: 100, loss: 0.1466730237007141
step: 110, loss: 0.16490691900253296
step: 120, loss: 0.33058595657348633
step: 130, loss: 0.10386569052934647
step: 140, loss: 0.26794424653053284
step: 150, loss: 0.2414316087961197
step: 160, loss: 0.3738042414188385
step: 170, loss: 0.15438510477542877
step: 180, loss: 0.17529620230197906
step: 190, loss: 0.14733977615833282
step: 200, loss: 0.08047953993082047
step: 210, loss: 0.24369680881500244
/usr/local/lib/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=0.5625, f1=0.5954692556634305, best_f1=0.5954692556634305
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.")
2293it [00:04, 535.04it/s]
load_f1 = 0.5357142857142857
real_f1 = 0.5273833671399594
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 [00:20, 210.98it/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 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: 0.4949153959751129



```

Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 10, loss: 0.5073251724243164
step: 20, loss: 0.2675057649612427
step: 30, loss: 0.3595905303955078
step: 40, loss: 0.252562016248703
step: 50, loss: 0.34240806102752686
step: 60, loss: 0.48125511407852173
step: 70, loss: 0.4460850656032562
step: 80, loss: 0.17734165489673615
step: 90, loss: 0.3094157576560974
step: 100, loss: 0.4194526970386505
step: 110, loss: 0.248111754655838
step: 120, loss: 0.31536978483200073
step: 130, loss: 0.3110385835170746
step: 140, loss: 0.1658197045326233
step: 150, loss: 0.2971802055835724
step: 160, loss: 0.22114045917987823
step: 170, loss: 0.4590890407562256
step: 180, loss: 0.13520756363868713
step: 190, loss: 0.12907493114471436
/usr/local/lib/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=0.5612472160356347, f1=0.5785876993166287,
best_f1=0.5785876993166287
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.3277832269668579
step: 10, loss: 0.3266008496284485
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 8192.0
step: 20, loss: 0.8541995286941528
step: 30, loss: 0.15774479508399963
step: 40, loss: 0.29823967814445496
step: 50, loss: 0.3550439178943634
step: 60, loss: 0.123057521879673
step: 70, loss: 0.28412431478500366
step: 80, loss: 0.05308227241039276
step: 90, loss: 0.11014040559530258
step: 100, loss: 0.15302957594394684
step: 110, loss: 0.21917416155338287
step: 120, loss: 0.05570093169808388
step: 130, loss: 0.1251225620508194
step: 140, loss: 0.20552265644073486

```

```

step: 150, loss: 0.24295878410339355
step: 160, loss: 0.1895386129617691
step: 170, loss: 0.038265421986579895
step: 180, loss: 0.08991725742816925
step: 190, loss: 0.05657688155770302
/usr/local/lib/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=0.7419354838709677, f1=0.7731092436974789,
best_f1=0.7731092436974789
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.16619235277175903
step: 10, loss: 0.14552336931228638
step: 20, loss: 0.10445481538772583
step: 30, loss: 0.02164752408862114
step: 40, loss: 0.1167815700173378
step: 50, loss: 0.14244021475315094
step: 60, loss: 0.019945835694670677
step: 70, loss: 0.2335360050201416
step: 80, loss: 0.25774726271629333
step: 90, loss: 0.11031216382980347
step: 100, loss: 0.07831598073244095
step: 110, loss: 0.20851899683475494
step: 120, loss: 0.2086750864982605
step: 130, loss: 0.07572148740291595
step: 140, loss: 0.0626305416226387
step: 150, loss: 0.16141168773174286
step: 160, loss: 0.04211818799376488
step: 170, loss: 0.29638680815696716
step: 180, loss: 0.13413850963115692
step: 190, loss: 0.08928581327199936
/usr/local/lib/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=0.7759036144578314, f1=0.7852193995381062,
best_f1=0.7852193995381062
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.

```

```

    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.023135900497436523
step: 10, loss: 0.02447669580578804
step: 20, loss: 0.08667381852865219
step: 30, loss: 0.13765409588813782
step: 40, loss: 0.22274374961853027
step: 50, loss: 0.07785145193338394
step: 60, loss: 0.0438796691596508
step: 70, loss: 0.0685327798128128
step: 80, loss: 0.017795991152524948
step: 90, loss: 0.03248830512166023
step: 100, loss: 0.16374874114990234
step: 110, loss: 0.1930035650730133
step: 120, loss: 0.05853065103292465
step: 130, loss: 0.09795606881380081
step: 140, loss: 0.2138664573431015
step: 150, loss: 0.013411096297204494
step: 160, loss: 0.02548850327730179
step: 170, loss: 0.07941025495529175
step: 180, loss: 0.12135205417871475
step: 190, loss: 0.04758143424987793
/usr/local/lib/python3.7/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=0.7658536585365853, f1=0.769607843137255,
best_f1=0.7852193995381062
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.2339976280927658
step: 10, loss: 0.14797788858413696
step: 20, loss: 0.05784142017364502
step: 30, loss: 0.028381144627928734
step: 40, loss: 0.02248363196849823
step: 50, loss: 0.018782738596200943
step: 60, loss: 0.16574031114578247
step: 70, loss: 0.1257181167602539
step: 80, loss: 0.01322004571557045
step: 90, loss: 0.23733848333358765
step: 100, loss: 0.13438546657562256
step: 110, loss: 0.11342757195234299
step: 120, loss: 0.03839763253927231
step: 130, loss: 0.3085111677646637
step: 140, loss: 0.017781995236873627
step: 150, loss: 0.2774507701396942
step: 160, loss: 0.012199895456433296
step: 170, loss: 0.021687636151909828
step: 180, loss: 0.035476818680763245

```

```

step: 190, loss: 0.02089606784284115
/usr/local/lib/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=0.7989556135770235, f1=0.7916666666666666,
best_f1=0.7916666666666666
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.01760876178741455
step: 10, loss: 0.0654076337814331
step: 20, loss: 0.003914030268788338
step: 30, loss: 0.10583187639713287
step: 40, loss: 0.012281776405870914
step: 50, loss: 0.02503105439245701
step: 60, loss: 0.05736769735813141
step: 70, loss: 0.04406176880002022
step: 80, loss: 0.07139932364225388
step: 90, loss: 0.053164951503276825
step: 100, loss: 0.02393747679889202
step: 110, loss: 0.024076830595731735
step: 120, loss: 0.01650143973529339
step: 130, loss: 0.10519401729106903
step: 140, loss: 0.0035142956767231226
step: 150, loss: 0.024227125570178032
step: 160, loss: 0.1916133612394333
step: 170, loss: 0.1701589971780777
step: 180, loss: 0.07530733942985535
step: 190, loss: 0.016203319653868675
/usr/local/lib/python3.7/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=0.7921760391198045, f1=0.8077858880778588,
best_f1=0.7916666666666666
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.020294692367315292
step: 10, loss: 0.07334255427122116
step: 20, loss: 0.003911056090146303
step: 30, loss: 0.034917935729026794
step: 40, loss: 0.06039974093437195
step: 50, loss: 0.00678511057049036
step: 60, loss: 0.006766526959836483

```

```

step: 70, loss: 0.004499088507145643
step: 80, loss: 0.0034701991826295853
step: 90, loss: 0.0017976914532482624
step: 100, loss: 0.20394732058048248
step: 110, loss: 0.10717611759901047
step: 120, loss: 0.07891780883073807
step: 130, loss: 0.007952417246997356
step: 140, loss: 0.016374222934246063
step: 150, loss: 0.004293127451092005
step: 160, loss: 0.08445660024881363
step: 170, loss: 0.0062193856574594975
step: 180, loss: 0.01470264419913292
step: 190, loss: 0.02327943779528141
/usr/local/lib/python3.7/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=0.7919799498746868, f1=0.8113695090439276,
best_f1=0.7916666666666666
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.02707202173769474
step: 10, loss: 0.004506107419729233
step: 20, loss: 0.0026872309390455484
step: 30, loss: 0.1276949942111969
step: 40, loss: 0.04784363508224487
step: 50, loss: 0.00913923792541027
step: 60, loss: 0.032974209636449814
step: 70, loss: 0.005538120865821838
step: 80, loss: 0.08477449417114258
step: 90, loss: 0.0025869577657431364
step: 100, loss: 0.00623583747074008
step: 110, loss: 0.09691162407398224
step: 120, loss: 0.0035041950177401304
step: 130, loss: 0.0015474739484488964
step: 140, loss: 0.0037249766755849123
step: 150, loss: 0.14563453197479248
step: 160, loss: 0.0067129735834896564
step: 170, loss: 0.0008382229716517031
step: 180, loss: 0.030733970925211906
step: 190, loss: 0.0024357398506253958
/usr/local/lib/python3.7/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=0.7979539641943734, f1=0.795969773299748,
best_f1=0.7916666666666666
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.

```

```

    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.012073909863829613
step: 10, loss: 0.0006984818028286099
step: 20, loss: 0.0033177933655679226
step: 30, loss: 0.002865065587684512
step: 40, loss: 0.030016181990504265
step: 50, loss: 0.029109599068760872
step: 60, loss: 0.011706414632499218
step: 70, loss: 0.0026022084057331085
step: 80, loss: 0.014934813603758812
step: 90, loss: 0.0012583571951836348
step: 100, loss: 0.0007392395636998117
step: 110, loss: 0.0005864903214387596
step: 120, loss: 0.0037463465705513954
step: 130, loss: 0.0329907163977623
step: 140, loss: 0.056884896010160446
step: 150, loss: 0.0021395611111074686
step: 160, loss: 0.0009391973144374788
step: 170, loss: 0.005628280341625214
step: 180, loss: 0.00456096651032567
step: 190, loss: 0.00043179019121453166
/usr/local/lib/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 9: dev_f1=0.8169761273209548, f1=0.8288770053475937,
best_f1=0.8288770053475937
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.003491143463179469
step: 10, loss: 0.0010782465105876327
step: 20, loss: 0.0009495048434473574
step: 30, loss: 0.17523346841335297
step: 40, loss: 0.0014192801900207996
step: 50, loss: 0.004005652852356434
step: 60, loss: 0.0004690038622356951
step: 70, loss: 0.0003731635515578091
step: 80, loss: 0.001387840136885643
step: 90, loss: 0.0013374964473769069
step: 100, loss: 0.000494323845487088
step: 110, loss: 0.0004965000553056598
step: 120, loss: 0.0009014259558171034
step: 130, loss: 0.00042120704893022776
step: 140, loss: 0.013596657663583755

```

```

step: 150, loss: 0.00026634763344191015
step: 160, loss: 0.0005734710139222443
step: 170, loss: 0.04200499877333641
step: 180, loss: 0.005664633121341467
step: 190, loss: 0.00046727684093639255
/usr/local/lib/python3.7/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=0.7806122448979592, f1=0.8226221079691518,
best_f1=0.8288770053475937
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0033152818214148283
step: 10, loss: 0.0007566402782686055
step: 20, loss: 0.0020806898828595877
step: 30, loss: 0.002761928364634514
step: 40, loss: 0.0005963582661934197
step: 50, loss: 0.005902057513594627
step: 60, loss: 0.001393725979141891
step: 70, loss: 0.0033867903985083103
step: 80, loss: 0.003818897297605872
step: 90, loss: 0.00779363326728344
step: 100, loss: 0.0005073242355138063
step: 110, loss: 0.18932074308395386
step: 120, loss: 0.0006035424885340035
step: 130, loss: 0.0008600241271778941
step: 140, loss: 0.003433970035985112
step: 150, loss: 0.03476765751838684
step: 160, loss: 0.002485094126313925
step: 170, loss: 0.002461736323311925
step: 180, loss: 0.002592360833659768
step: 190, loss: 0.0009215666796080768
/usr/local/lib/python3.7/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=0.8129675810473815, f1=0.8, best_f1=0.8288770053475937
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0006717428914271295
step: 10, loss: 0.02279203198850155
step: 20, loss: 0.0004831039987038821
step: 30, loss: 0.0007741758599877357
step: 40, loss: 0.00022451297263614833
step: 50, loss: 0.025252828374505043
step: 60, loss: 0.013665142469108105
step: 70, loss: 0.0007288993801921606

```

```

step: 80, loss: 0.0005817125202156603
step: 90, loss: 0.0009856689721345901
step: 100, loss: 0.000423167715780437
step: 110, loss: 0.002485520439222455
step: 120, loss: 0.000249385746428743
step: 130, loss: 0.0016085215611383319
step: 140, loss: 0.0022897496819496155
step: 150, loss: 0.0006236285553313792
step: 160, loss: 0.0004061888321302831
step: 170, loss: 0.0006144448998384178
step: 180, loss: 0.00069764896761626
step: 190, loss: 0.0001978341315407306
/usr/local/lib/python3.7/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=0.7977839335180055, f1=0.8179271708683474,
best_f1=0.8288770053475937
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.002590187592431903
step: 10, loss: 0.0003519994206726551
step: 20, loss: 0.0007068563718348742
step: 30, loss: 0.00721993762999773
step: 40, loss: 0.0007721178117208183
step: 50, loss: 0.0017163777956739068
step: 60, loss: 0.0003762743144761771
step: 70, loss: 0.009433058090507984
step: 80, loss: 0.000892349926289171
step: 90, loss: 0.0008665493223816156
step: 100, loss: 0.001096282503567636
step: 110, loss: 0.0014639555010944605
step: 120, loss: 0.0009448418277315795
step: 130, loss: 0.011330514214932919
step: 140, loss: 0.0013083290541544557
step: 150, loss: 0.0003681464877445251
step: 160, loss: 0.0006124814390204847
step: 170, loss: 0.0001845379447331652
step: 180, loss: 0.0005127131589688361
step: 190, loss: 0.16450481116771698
/usr/local/lib/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=0.8179419525065963, f1=0.8213333333333334,

```



```

best_f1=0.8213333333333334
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0021828741300851107
step: 10, loss: 0.0002674744464457035
step: 20, loss: 0.0004329598159529269
step: 30, loss: 0.002014914294704795
step: 40, loss: 0.000401488650823012
step: 50, loss: 0.00031912472331896424
step: 60, loss: 0.00272843474522233
step: 70, loss: 0.0011103413999080658
step: 80, loss: 0.0007336512207984924
step: 90, loss: 0.000192279985640198
step: 100, loss: 0.003073622239753604
step: 110, loss: 0.001912764273583889
step: 120, loss: 0.002083786064758897
step: 130, loss: 0.0002698293246794492
step: 140, loss: 0.00027859595138579607
step: 150, loss: 0.00043768397881649435
step: 160, loss: 0.0003184577217325568
step: 170, loss: 0.00015021003491710871
step: 180, loss: 0.0006455505499616265
step: 190, loss: 0.0007668031612411141
/usr/local/lib/python3.7/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=0.8146067415730337, f1=0.8441926345609065,
best_f1=0.8213333333333334
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0011312311980873346
step: 10, loss: 0.0007883823709562421
step: 20, loss: 0.0011630072258412838
step: 30, loss: 0.002433845540508628
step: 40, loss: 0.0007944984245114028
step: 50, loss: 0.00018288730643689632
step: 60, loss: 0.00016971949662547559
step: 70, loss: 0.0006101452745497227
step: 80, loss: 0.00016183160187210888
step: 90, loss: 0.000895194592885673
step: 100, loss: 0.000349442329024896
step: 110, loss: 0.00239246035926044
step: 120, loss: 0.00013891000708099455
step: 130, loss: 0.00031283075804822147
step: 140, loss: 0.008531270548701286
step: 150, loss: 0.00019069199333898723

```

```

step: 160, loss: 0.00027313243481330574
step: 170, loss: 0.0010509842541068792
step: 180, loss: 0.0012855792883783579
step: 190, loss: 0.005933270324021578
/usr/local/lib/python3.7/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=0.8125, f1=0.8376068376068376, best_f1=0.8213333333333334
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 [00:08, 255.70it/s]
load_f1 = 0.837696335078534
real_f1 = 0.810126582278481
733it [00:00, 3277.73it/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, 203.63it/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/
```

```
Downloading: 100% 481/481 [00:00<00:00, 431kB/s]
Downloading: 100% 899k/899k [00:01<00:00, 682kB/s]
Downloading: 100% 456k/456k [00:00<00:00, 512kB/s]
Downloading: 100% 501M/501M [00:07<00:00, 71.1MB/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: 0.5162197351455688
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 10, loss: 0.4652726650238037
step: 20, loss: 0.30483829975128174
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 8192.0
step: 30, loss: 0.43703436851501465
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 4096.0
step: 40, loss: 0.5278929471969604
step: 50, loss: 0.35351303219795227
step: 60, loss: 0.6001684665679932
step: 70, loss: 0.30587223172187805
step: 80, loss: 0.24077089130878448
step: 90, loss: 0.19923202693462372
step: 100, loss: 0.17677664756774902
step: 110, loss: 0.4411058723926544
step: 120, loss: 0.29203489422798157
step: 130, loss: 0.2932400703430176
step: 140, loss: 0.38842636346817017
step: 150, loss: 0.29728078842163086

```

```

step: 160, loss: 0.3779364228248596
step: 170, loss: 0.34423866868019104
/usr/local/lib/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=0.1941564561734213, f1=0.1941564561734213,
best_f1=0.1941564561734213
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.32432037591934204
step: 10, loss: 0.4972870945930481
step: 20, loss: 0.31009721755981445
step: 30, loss: 0.3265571892261505
step: 40, loss: 0.08676432073116302
step: 50, loss: 0.45236432552337646
step: 60, loss: 0.1898670196533203
step: 70, loss: 0.5097609758377075
step: 80, loss: 0.24591916799545288
step: 90, loss: 0.24997738003730774
step: 100, loss: 0.5153307318687439
step: 110, loss: 0.2572435140609741
step: 120, loss: 0.251792311668396
step: 130, loss: 0.5463679432868958
step: 140, loss: 0.5243008732795715
step: 150, loss: 0.44542914628982544
step: 160, loss: 0.44386211037635803
step: 170, loss: 0.3793767988681793
/usr/local/lib/python3.7/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=0.1941564561734213
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.5914474129676819
step: 10, loss: 0.2885126769542694
step: 20, loss: 0.23146218061447144
step: 30, loss: 0.24863308668136597
step: 40, loss: 0.37419214844703674
step: 50, loss: 0.5847966074943542
step: 60, loss: 0.30148956179618835
step: 70, loss: 0.23383578658103943

```

```

step: 80, loss: 0.3909297585487366
step: 90, loss: 0.5434029698371887
step: 100, loss: 0.30677416920661926
step: 110, loss: 0.18601615726947784
step: 120, loss: 0.5472939610481262
step: 130, loss: 0.4982450008392334
step: 140, loss: 0.44023215770721436
step: 150, loss: 0.20079508423805237
step: 160, loss: 0.16690853238105774
step: 170, loss: 0.3026820421218872
/usr/local/lib/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=0.19424799622819425, f1=0.19443133553563002,
best_f1=0.19443133553563002
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.362222820520401
step: 10, loss: 0.519442617893219
step: 20, loss: 0.25359755754470825
step: 30, loss: 0.46089547872543335
step: 40, loss: 0.25572240352630615
step: 50, loss: 0.38359886407852173
step: 60, loss: 0.7035394906997681
step: 70, loss: 0.325758159160614
step: 80, loss: 0.44848328828811646
step: 90, loss: 0.30660468339920044
step: 100, loss: 0.39245107769966125
step: 110, loss: 0.4312525987625122
step: 120, loss: 0.5008854866027832
step: 130, loss: 0.3065096437931061
step: 140, loss: 0.25602373480796814
step: 150, loss: 0.7247604727745056
step: 160, loss: 0.12260399758815765
step: 170, loss: 0.24507850408554077
/usr/local/lib/python3.7/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=0.19443133553563002
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")

```

```

step: 0, loss: 0.3816032409667969
step: 10, loss: 0.31026172637939453
step: 20, loss: 0.3258618116378784
step: 30, loss: 0.3154430389404297
step: 40, loss: 0.23139777779579163
step: 50, loss: 0.2716657221317291
step: 60, loss: 0.30525949597358704
step: 70, loss: 0.3322230577468872
step: 80, loss: 0.06729047745466232
step: 90, loss: 0.566967248916626
step: 100, loss: 0.2481723129749298
step: 110, loss: 0.26505163311958313
step: 120, loss: 0.14227613806724548
step: 130, loss: 0.24959850311279297
step: 140, loss: 0.17272265255451202
step: 150, loss: 0.30118030309677124
step: 160, loss: 0.25390753149986267
step: 170, loss: 0.3412156105041504
/usr/local/lib/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=0.625, f1=0.6453089244851259, best_f1=0.6453089244851259
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.268867164850235
step: 10, loss: 0.12049839645624161
step: 20, loss: 0.3748992681503296
step: 30, loss: 0.22136712074279785
step: 40, loss: 0.29520341753959656
step: 50, loss: 0.13439461588859558
step: 60, loss: 0.22730869054794312
step: 70, loss: 0.39962878823280334
step: 80, loss: 0.11782456189393997
step: 90, loss: 0.21884785592556
step: 100, loss: 0.18687704205513
step: 110, loss: 0.1725548952817917
step: 120, loss: 0.270090252161026
step: 130, loss: 0.2504620850086212
step: 140, loss: 0.2843702435493469
step: 150, loss: 0.055475201457738876
step: 160, loss: 0.22379732131958008
step: 170, loss: 0.15027618408203125
/usr/local/lib/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=0.7037037037037038, f1=0.6810176125244618,
best_f1=0.6810176125244618
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.17860601842403412
step: 10, loss: 0.23518387973308563
step: 20, loss: 0.29857516288757324
step: 30, loss: 0.08689091354608536
step: 40, loss: 0.0029645864851772785
step: 50, loss: 0.06717459112405777
step: 60, loss: 0.07899274677038193
step: 70, loss: 0.14463907480239868
step: 80, loss: 0.1070052832365036
step: 90, loss: 0.09890471398830414
step: 100, loss: 0.024168971925973892
step: 110, loss: 0.08370909839868546
step: 120, loss: 0.08570743352174759
step: 130, loss: 0.0980653464794159
step: 140, loss: 0.08111893385648727
step: 150, loss: 0.16866149008274078
step: 160, loss: 0.18641069531440735
step: 170, loss: 0.07440294325351715
/usr/local/lib/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=0.7706855791962176, f1=0.7636363636363637,
best_f1=0.7636363636363637
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.12209781259298325
step: 10, loss: 0.1634446531534195
step: 20, loss: 0.15514086186885834
step: 30, loss: 0.02502894029021263
step: 40, loss: 0.11745693534612656
step: 50, loss: 0.02096601203083992
step: 60, loss: 0.20637395977973938

```



```

step: 70, loss: 0.16787883639335632
step: 80, loss: 0.021062323823571205
step: 90, loss: 0.16484200954437256
step: 100, loss: 0.04107052460312843
step: 110, loss: 0.16952787339687347
step: 120, loss: 0.3910244405269623
step: 130, loss: 0.09841425716876984
step: 140, loss: 0.10421693325042725
step: 150, loss: 0.03143533319234848
step: 160, loss: 0.06439757347106934
step: 170, loss: 0.1524340957403183
/usr/local/lib/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=0.7969924812030076, f1=0.7943262411347518,
best_f1=0.7943262411347518
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.09916597604751587
step: 10, loss: 0.08272336423397064
step: 20, loss: 0.06447123736143112
step: 30, loss: 0.043498728424310684
step: 40, loss: 0.1130584180355072
step: 50, loss: 0.017899710685014725
step: 60, loss: 0.062117800116539
step: 70, loss: 0.07761258631944656
step: 80, loss: 0.05651164799928665
step: 90, loss: 0.11434605717658997
step: 100, loss: 0.10725453495979309
step: 110, loss: 0.15610450506210327
step: 120, loss: 0.12068069726228714
step: 130, loss: 0.019737184047698975
step: 140, loss: 0.16818338632583618
step: 150, loss: 0.23829616606235504
step: 160, loss: 0.030575282871723175
step: 170, loss: 0.06312304735183716
/usr/local/lib/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=0.8186528497409326, f1=0.8316831683168316,
best_f1=0.8316831683168316
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.11165445297956467
step: 10, loss: 0.0352398082613945
step: 20, loss: 0.12394160032272339
step: 30, loss: 0.07508540153503418
step: 40, loss: 0.07967433333396912
step: 50, loss: 0.09316997975111008
step: 60, loss: 0.09098467975854874
step: 70, loss: 0.06480751931667328
step: 80, loss: 0.053598128259181976
step: 90, loss: 0.007535899057984352
step: 100, loss: 0.045659389346838
step: 110, loss: 0.04771190136671066
step: 120, loss: 0.06126842647790909
step: 130, loss: 0.07948238402605057
step: 140, loss: 0.015116500668227673
step: 150, loss: 0.018015213310718536
step: 160, loss: 0.005798035766929388
step: 170, loss: 0.05094923451542854
/usr/local/lib/python3.7/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=0.8115183246073299, f1=0.8333333333333333,
best_f1=0.8316831683168316
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.08161529153585434
step: 10, loss: 0.03275800868868828
step: 20, loss: 0.03258025273680687
step: 30, loss: 0.0677364394068718
step: 40, loss: 0.061610084027051926
step: 50, loss: 0.17916244268417358
step: 60, loss: 0.1519109010696411
step: 70, loss: 0.043660905212163925
step: 80, loss: 0.0319613479077816
step: 90, loss: 0.033204834908246994
step: 100, loss: 0.19831311702728271
step: 110, loss: 0.02264401875436306
step: 120, loss: 0.07905546575784683
step: 130, loss: 0.03879585489630699
step: 140, loss: 0.02614648826420307
step: 150, loss: 0.0028408404905349016
step: 160, loss: 0.04955313354730606

```

```

step: 170, loss: 0.18424341082572937
/usr/local/lib/python3.7/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=0.8177083333333334, f1=0.834567901234568,
best_f1=0.8316831683168316
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.04055623337626457
step: 10, loss: 0.024435164406895638
step: 20, loss: 0.02495514042675495
step: 30, loss: 0.1902477890253067
step: 40, loss: 0.004899326711893082
step: 50, loss: 0.02622396871447563
step: 60, loss: 0.02074681967496872
step: 70, loss: 0.031187547370791435
step: 80, loss: 0.0037156809121370316
step: 90, loss: 0.0755162462592125
step: 100, loss: 0.008354008197784424
step: 110, loss: 0.04058603197336197
step: 120, loss: 0.059135787189006805
step: 130, loss: 0.10565447807312012
step: 140, loss: 0.003861530451104045
step: 150, loss: 0.011128529906272888
step: 160, loss: 0.0388069748878479
step: 170, loss: 0.29752302169799805
/usr/local/lib/python3.7/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=0.8116710875331565, f1=0.8287841191066997,
best_f1=0.8316831683168316
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.033762142062187195
step: 10, loss: 0.15921112895011902
step: 20, loss: 0.018473612144589424
step: 30, loss: 0.073427714407444
step: 40, loss: 0.005408929660916328
step: 50, loss: 0.11541611701250076
step: 60, loss: 0.05802937224507332
step: 70, loss: 0.0928603783249855
step: 80, loss: 0.0018540788441896439
step: 90, loss: 0.08347893506288528
step: 100, loss: 0.06818325817584991
step: 110, loss: 0.0063418010249733925
step: 120, loss: 0.05754885450005531

```

```

step: 130, loss: 0.004739714786410332
step: 140, loss: 0.05755268782377243
step: 150, loss: 0.030726799741387367
step: 160, loss: 0.0030800416134297848
step: 170, loss: 0.009004872292280197
/usr/local/lib/python3.7/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=0.8159203980099503, f1=0.8314087759815242,
best_f1=0.8316831683168316
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.008624750189483166
step: 10, loss: 0.010562204755842686
step: 20, loss: 0.013126160949468613
step: 30, loss: 0.03540001064538956
step: 40, loss: 0.004660783801227808
step: 50, loss: 0.0011097171809524298
step: 60, loss: 0.0050000352784991264
step: 70, loss: 0.01055472157895565
step: 80, loss: 0.0008165116887539625
step: 90, loss: 0.001579601550474763
step: 100, loss: 0.009006007574498653
step: 110, loss: 0.013462717644870281
step: 120, loss: 0.09499290585517883
step: 130, loss: 0.06939482688903809
step: 140, loss: 0.1051851361989975
step: 150, loss: 0.05967523902654648
step: 160, loss: 0.004395695868879557
step: 170, loss: 0.06392832100391388
/usr/local/lib/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=0.8191489361702128, f1=0.8341708542713568,
best_f1=0.8341708542713568
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.011337080970406532
step: 10, loss: 0.01037656795233488
step: 20, loss: 0.0185487512499094
step: 30, loss: 0.024104349315166473
step: 40, loss: 0.010847964324057102

```

```

step: 50, loss: 0.0014947388553991914
step: 60, loss: 0.007714853156358004
step: 70, loss: 0.10122599452733994
step: 80, loss: 0.010199243202805519
step: 90, loss: 0.010936705395579338
step: 100, loss: 0.05064341798424721
step: 110, loss: 0.023164810612797737
step: 120, loss: 0.06038697063922882
step: 130, loss: 0.02297557145357132
step: 140, loss: 0.02878340519964695
step: 150, loss: 0.0007763314060866833
step: 160, loss: 0.0016004386125132442
step: 170, loss: 0.003147713840007782
/usr/local/lib/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=0.8203753351206435, f1=0.8226221079691517,
best_f1=0.8226221079691517
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 [00:07, 244.90it/s]
load_f1 = 0.7658536585365854
real_f1 = 0.7482014388489209
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 [00:29, 149.80it/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 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: 0.6428719162940979
```

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: 0.5104607939720154
```

```
step: 20, loss: 0.5695582032203674
```

```
step: 30, loss: 0.2913742661476135
```

```
step: 40, loss: 0.3620990514755249
```

```
step: 50, loss: 0.6036851406097412
```

```
step: 60, loss: 0.5134280323982239
```

```
step: 70, loss: 0.4286361634731293
```

```
step: 80, loss: 0.5946087837219238
```

```
step: 90, loss: 0.45967039465904236
```

```
step: 100, loss: 0.4587825536727905
```

```
step: 110, loss: 0.6426102519035339
```

```
step: 120, loss: 0.4088066518306732
```

```
step: 130, loss: 0.3278732895851135
```

```
step: 140, loss: 0.3905588388442993
```

```
step: 150, loss: 0.578973650932312
```

```
step: 160, loss: 0.49971523880958557
```

```

step: 170, loss: 0.47726520895957947
step: 180, loss: 0.3984268307685852
step: 190, loss: 0.2904142439365387
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 8192.0
step: 200, loss: 0.3940889835357666
step: 210, loss: 0.49132266640663147
step: 220, loss: 0.6176818013191223
step: 230, loss: 0.35527467727661133
/usr/local/lib/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=0.3044223517312307, f1=0.3044223517312307,
best_f1=0.3044223517312307
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.4869082570075989
step: 10, loss: 0.524720311164856
step: 20, loss: 0.5583420395851135
step: 30, loss: 0.5179626941680908
step: 40, loss: 0.537490725517273
step: 50, loss: 0.3879796862602234
step: 60, loss: 0.2601166069507599
step: 70, loss: 0.1309613436460495
step: 80, loss: 0.030187496915459633
step: 90, loss: 0.14508892595767975
step: 100, loss: 0.0512714646756649
step: 110, loss: 0.1000601276755333
step: 120, loss: 0.09886672347784042
step: 130, loss: 0.1940009593963623
step: 140, loss: 0.024306731298565865
step: 150, loss: 0.029667820781469345
step: 160, loss: 0.028672700747847557
step: 170, loss: 0.012035577557981014
step: 180, loss: 0.010086273774504662
step: 190, loss: 0.024762552231550217
step: 200, loss: 0.12739530205726624
step: 210, loss: 0.03373342007398605
step: 220, loss: 0.0030221203342080116
step: 230, loss: 0.0035402686335146427
/usr/local/lib/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=0.9496080627099663, f1=0.9524886877828055,

best\_f1=0.9524886877828055

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

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

step: 0, loss: 0.08365448564291

step: 10, loss: 0.0484897680580616

step: 20, loss: 0.14606156945228577

step: 30, loss: 0.015981215983629227

step: 40, loss: 0.04396842420101166

step: 50, loss: 0.011377297341823578

step: 60, loss: 0.04024071618914604

step: 70, loss: 0.0801975429058075

step: 80, loss: 0.06314227730035782

step: 90, loss: 0.06252217292785645

step: 100, loss: 0.030802996829152107

step: 110, loss: 0.019675131887197495

step: 120, loss: 0.0015189606929197907

step: 130, loss: 0.05527089536190033

step: 140, loss: 0.004025152884423733

step: 150, loss: 0.154102623462677

step: 160, loss: 0.009134032763540745

step: 170, loss: 0.005037502385675907

step: 180, loss: 0.013968913815915585

step: 190, loss: 0.04223497584462166

step: 200, loss: 0.004354025237262249

step: 210, loss: 0.0038539497181773186

step: 220, loss: 0.0685005709528923

step: 230, loss: 0.04631125181913376

/usr/local/lib/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=0.9686800894854586, f1=0.963882618510158,

best\_f1=0.963882618510158

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

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

step: 0, loss: 0.022808117792010307

step: 10, loss: 0.007249094545841217

step: 20, loss: 0.037905484437942505

step: 30, loss: 0.0025787476915866137

```

step: 40, loss: 0.11276113241910934
step: 50, loss: 0.028411485254764557
step: 60, loss: 0.008857461623847485
step: 70, loss: 0.047926463186740875
step: 80, loss: 0.010615545324981213
step: 90, loss: 0.0175553597509861
step: 100, loss: 0.024119926616549492
step: 110, loss: 0.002574666403234005
step: 120, loss: 0.049165356904268265
step: 130, loss: 0.04241384565830231
step: 140, loss: 0.004982438404113054
step: 150, loss: 0.011149407364428043
step: 160, loss: 0.012040266767144203
step: 170, loss: 0.009022966958582401
step: 180, loss: 0.15100277960300446
step: 190, loss: 0.004938112106174231
step: 200, loss: 0.137749582529068
step: 210, loss: 0.11785215884447098
step: 220, loss: 0.002203962067142129
step: 230, loss: 0.004323951900005341
/usr/local/lib/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=0.9689578713968958, 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: 0.004142193589359522
step: 10, loss: 0.007311354856938124
step: 20, loss: 0.01651945523917675
step: 30, loss: 0.005639621987938881
step: 40, loss: 0.0035942448303103447
step: 50, loss: 0.006566951517015696
step: 60, loss: 0.024736052379012108
step: 70, loss: 0.0012392106000334024
step: 80, loss: 0.03918059170246124
step: 90, loss: 0.24571700394153595
step: 100, loss: 0.0038186851888895035
step: 110, loss: 0.006878652144223452
step: 120, loss: 0.010423251427710056
step: 130, loss: 0.002060945611447096
step: 140, loss: 0.20732718706130981
step: 150, loss: 0.008295971900224686

```

```

step: 160, loss: 0.0026826029643416405
step: 170, loss: 0.006520366296172142
step: 180, loss: 0.0026734210550785065
step: 190, loss: 0.08401115238666534
step: 200, loss: 0.035709962248802185
step: 210, loss: 0.07346872985363007
step: 220, loss: 0.0011857874924317002
step: 230, loss: 0.012093371711671352
/usr/local/lib/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=0.9693654266958425, f1=0.9618320610687023,
best_f1=0.9618320610687023
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0011451573809608817
step: 10, loss: 0.003083331510424614
step: 20, loss: 0.001636508502997458
step: 30, loss: 0.0014495160430669785
step: 40, loss: 0.0005296161398291588
step: 50, loss: 0.0005907418089918792
step: 60, loss: 0.0010428800014778972
step: 70, loss: 0.1342984437942505
step: 80, loss: 0.0025658723898231983
step: 90, loss: 0.0020947065204381943
step: 100, loss: 0.0069968970492482185
step: 110, loss: 0.0904054194688797
step: 120, loss: 0.0036300141364336014
step: 130, loss: 0.0023139268159866333
step: 140, loss: 0.0018246641848236322
step: 150, loss: 0.0010062218643724918
step: 160, loss: 0.003084801137447357
step: 170, loss: 0.008459839969873428
step: 180, loss: 0.003065630095079541
step: 190, loss: 0.0031543858349323273
step: 200, loss: 0.010898304171860218
step: 210, loss: 0.0023484507109969854
step: 220, loss: 0.0044947960413992405
step: 230, loss: 0.0010110782459378242
/usr/local/lib/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=0.978675645342312, f1=0.9752252252252253,
best_f1=0.9752252252252253
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
```

```
warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0012286331038922071
step: 10, loss: 0.0008829715079627931
step: 20, loss: 0.008323049172759056
step: 30, loss: 0.0003655800537671894
step: 40, loss: 0.0006017001578584313
step: 50, loss: 0.0008269454119727015
step: 60, loss: 0.0008932240889407694
step: 70, loss: 0.0018135039135813713
step: 80, loss: 0.0007139253430068493
step: 90, loss: 0.05458999052643776
step: 100, loss: 0.00038398877950385213
step: 110, loss: 0.00025359413120895624
step: 120, loss: 0.0007990880985744298
step: 130, loss: 0.003990247845649719
step: 140, loss: 0.0027435917872935534
step: 150, loss: 0.022369874641299248
step: 160, loss: 0.0033257801551371813
step: 170, loss: 0.001556918490678072
step: 180, loss: 0.0005262380582280457
step: 190, loss: 0.0011414639884606004
step: 200, loss: 0.0014956739032641053
step: 210, loss: 0.017754197120666504
step: 220, loss: 0.00885135680437088
step: 230, loss: 0.004218411631882191
/usr/local/lib/python3.7/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=0.9753363228699552, f1=0.9718785151856018,
best_f1=0.9752252252252253
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
```

```
warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0014380983775481582
step: 10, loss: 0.0055404589511454105
step: 20, loss: 0.004692419897764921
step: 30, loss: 0.0011298615718260407
step: 40, loss: 0.00143915053922683
step: 50, loss: 0.0007032316061668098
step: 60, loss: 0.00043803558219224215
step: 70, loss: 0.003925844095647335
```

```

step: 80, loss: 0.028674324974417686
step: 90, loss: 0.0031111487187445164
step: 100, loss: 0.0008546631434001029
step: 110, loss: 0.0009150591213256121
step: 120, loss: 0.0006658090278506279
step: 130, loss: 0.0027102851308882236
step: 140, loss: 0.0013762887101620436
step: 150, loss: 0.22670066356658936
step: 160, loss: 0.0011558022815734148
step: 170, loss: 0.11508937180042267
step: 180, loss: 0.002232031896710396
step: 190, loss: 0.02351432479918003
step: 200, loss: 0.010662727057933807
step: 210, loss: 0.0020042159594595432
step: 220, loss: 0.004585167393088341
step: 230, loss: 0.0012475962284952402
/usr/local/lib/python3.7/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=0.9786276715410572, f1=0.9752252252252253,
best_f1=0.9752252252252253
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.000589379109442234
step: 10, loss: 0.0017885713605210185
step: 20, loss: 0.00036936436663381755
step: 30, loss: 0.0010469343978911638
step: 40, loss: 0.0006707963184453547
step: 50, loss: 0.0020221476443111897
step: 60, loss: 0.0004670898197218776
step: 70, loss: 0.052483294159173965
step: 80, loss: 0.00017396641487721354
step: 90, loss: 0.003373504616320133
step: 100, loss: 0.0006214799359440804
step: 110, loss: 0.047022659331560135
step: 120, loss: 0.017768345773220062
step: 130, loss: 0.0009865663014352322
step: 140, loss: 0.0011084151919931173
step: 150, loss: 0.002189673949033022
step: 160, loss: 0.0009067679056897759
step: 170, loss: 0.00034030008828267455
step: 180, loss: 0.0003808217588812113
step: 190, loss: 0.00024203417706303298
step: 200, loss: 0.003368593053892255
step: 210, loss: 0.0020797057077288628
step: 220, loss: 0.0005536901298910379
step: 230, loss: 0.0006665863329544663

```

```

/usr/local/lib/python3.7/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=0.9766925638179801, f1=0.9799107142857142,
best_f1=0.9752252252252253
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0004850992118008435
step: 10, loss: 0.0006815621163696051
step: 20, loss: 0.0010128193534910679
step: 30, loss: 0.00019710618653334677
step: 40, loss: 0.000981309567578137
step: 50, loss: 0.00025376654230058193
step: 60, loss: 0.001214535441249609
step: 70, loss: 0.004026125650852919
step: 80, loss: 0.0009550382965244353
step: 90, loss: 0.0005548358312807977
step: 100, loss: 0.0002159569994546473
step: 110, loss: 0.007302450947463512
step: 120, loss: 0.00045171717647463083
step: 130, loss: 0.01136760413646698
step: 140, loss: 0.0003226593544241041
step: 150, loss: 0.0003015224647242576
step: 160, loss: 5.0700968131422997e-05
step: 170, loss: 0.00025401872699148953
step: 180, loss: 0.004697847180068493
step: 190, loss: 0.001032471191138029
step: 200, loss: 0.0015313458861783147
step: 210, loss: 0.010522786527872086
step: 220, loss: 0.004408301319926977
step: 230, loss: 0.000586775888223201
/usr/local/lib/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=0.9809203142536477, f1=0.9764837625979844,
best_f1=0.9764837625979844
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 9.747833973960951e-05
step: 10, loss: 0.0004909625858999789
step: 20, loss: 0.0003142375499010086
step: 30, loss: 0.0004180852265562862

```

```

step: 40, loss: 4.200398689135909e-05
step: 50, loss: 0.0022062521893531084
step: 60, loss: 0.010743875056505203
step: 70, loss: 0.0004935219185426831
step: 80, loss: 0.0005578084383159876
step: 90, loss: 0.23084521293640137
step: 100, loss: 0.0005806732806377113
step: 110, loss: 0.0005426335264928639
step: 120, loss: 0.00036627304507419467
step: 130, loss: 0.00027346808928996325
step: 140, loss: 0.0027787250000983477
step: 150, loss: 0.0004783739277627319
step: 160, loss: 0.0006006642361171544
step: 170, loss: 0.0005132118239998817
step: 180, loss: 0.0004866914823651314
step: 190, loss: 0.0001810014946386218
step: 200, loss: 0.0008996751275844872
step: 210, loss: 0.00017075489449780434
step: 220, loss: 0.00025140694924630225
step: 230, loss: 0.00024020786804612726
/usr/local/lib/python3.7/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=0.9797297297297298, f1=0.9774774774774775,
best_f1=0.9764837625979844
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0001477016630815342
step: 10, loss: 0.00017594904056750238
step: 20, loss: 0.0005607671337202191
step: 30, loss: 0.0072531793266534805
step: 40, loss: 0.0007916066097095609
step: 50, loss: 0.0003712291654665023
step: 60, loss: 0.0002293606667080894
step: 70, loss: 9.263752872357145e-05
step: 80, loss: 5.251328184385784e-05
step: 90, loss: 0.001764380489476025
step: 100, loss: 0.00010748958447948098
step: 110, loss: 7.661453128093854e-05
step: 120, loss: 0.0009440597495995462
step: 130, loss: 0.00039931462379172444
step: 140, loss: 0.0003076926514040679
step: 150, loss: 0.0002768302510958165
step: 160, loss: 0.0008750429842621088
step: 170, loss: 0.0007479758933186531
step: 180, loss: 0.0005158359417691827
step: 190, loss: 0.0007547999266535044

```

```

step: 200, loss: 0.000220646194065921
step: 210, loss: 0.012041430920362473
step: 220, loss: 0.008551264181733131
step: 230, loss: 0.001090615289285779
/usr/local/lib/python3.7/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=0.9797752808988766, f1=0.9786276715410572,
best_f1=0.9764837625979844
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00414681201800704
step: 10, loss: 0.0012317619984969497
step: 20, loss: 0.0004984884290024638
step: 30, loss: 0.00016239157412201166
step: 40, loss: 0.001957314321771264
step: 50, loss: 0.002482266165316105
step: 60, loss: 0.00020922553085256368
step: 70, loss: 0.0024156083818525076
step: 80, loss: 0.0031320438720285892
step: 90, loss: 0.0002862394612748176
step: 100, loss: 0.0002478977548889816
step: 110, loss: 0.00462883897125721
step: 120, loss: 0.000526243238709867
step: 130, loss: 0.00017498880333732814
step: 140, loss: 0.0001646863529458642
step: 150, loss: 0.00012171210983069614
step: 160, loss: 0.0004068997805006802
step: 170, loss: 0.0005819122307002544
step: 180, loss: 0.009709845297038555
step: 190, loss: 0.00031647837022319436
step: 200, loss: 3.9525646570837125e-05
step: 210, loss: 0.0003022134187631309
step: 220, loss: 0.00038914495962671936
step: 230, loss: 0.0005692846025340259
/usr/local/lib/python3.7/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=0.9718785151856018, f1=0.976324689966178,
best_f1=0.9764837625979844
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.003432160010561347
step: 10, loss: 0.00019743003940675408
step: 20, loss: 0.00018364496645517647
step: 30, loss: 0.0004864833317697048

```



```

step: 40, loss: 0.00025746162282302976
step: 50, loss: 0.00010644095164025202
step: 60, loss: 0.00011787388939410448
step: 70, loss: 6.886095070512965e-05
step: 80, loss: 0.0006498168804682791
step: 90, loss: 0.004555969499051571
step: 100, loss: 0.009916134178638458
step: 110, loss: 0.0003030056250281632
step: 120, loss: 7.49875107430853e-05
step: 130, loss: 0.004201244562864304
step: 140, loss: 0.0001964486436918378
step: 150, loss: 0.00029504860867746174
step: 160, loss: 0.0003678000357467681
step: 170, loss: 0.0003209953720215708
step: 180, loss: 0.00012812558270525187
step: 190, loss: 6.462214514613152e-05
step: 200, loss: 0.00026556081138551235
step: 210, loss: 0.00010317224223399535
step: 220, loss: 0.0001287998748011887
step: 230, loss: 7.032915891613811e-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=0.9786276715410572, f1=0.9797752808988766,
best_f1=0.9764837625979844
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00017790075798984617
step: 10, loss: 0.0001526531996205449
step: 20, loss: 0.0002153721870854497
step: 30, loss: 0.00010920627391897142
step: 40, loss: 9.001060243463144e-05
step: 50, loss: 6.772070628358051e-05
step: 60, loss: 0.0385112538933754
step: 70, loss: 0.00034621459781192243
step: 80, loss: 0.0003349040634930134
step: 90, loss: 0.0001721002336125821
step: 100, loss: 9.749310265760869e-05
step: 110, loss: 0.0003608346451073885
step: 120, loss: 0.0006750551401637495
step: 130, loss: 0.00010599695087876171
step: 140, loss: 0.00036941509461030364
step: 150, loss: 0.00023000742658041418
step: 160, loss: 0.007344904821366072
step: 170, loss: 8.83882530615665e-05
step: 180, loss: 0.0001057112094713375
step: 190, loss: 0.0037163496017456055

```

```

step: 200, loss: 0.0003467377391643822
step: 210, loss: 0.0012332132318988442
step: 220, loss: 7.765449117869139e-05
step: 230, loss: 0.00010413404379505664
/usr/local/lib/python3.7/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=0.9797297297297298, f1=0.9797752808988766,
best_f1=0.9764837625979844
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, 156.70it/s]
load_f1 = 0.9820224719101124
real_f1 = 0.978675645342312
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 [00:29, 149.80it/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: 0.620115339756012
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: 0.4311714470386505
step: 20, loss: 0.31964048743247986
step: 30, loss: 0.37118226289749146
step: 40, loss: 0.4153235852718353
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 8192.0
step: 50, loss: 0.873443603515625
step: 60, loss: 0.36087894439697266
step: 70, loss: 0.4752439856529236
step: 80, loss: 0.5118817687034607
step: 90, loss: 0.45755240321159363
step: 100, loss: 0.5601931810379028
step: 110, loss: 0.45906680822372437
step: 120, loss: 0.6785778999328613
step: 130, loss: 0.5236913561820984
step: 140, loss: 0.3765549659729004
step: 150, loss: 0.4394179582595825
step: 160, loss: 0.47050443291664124
step: 170, loss: 0.5487673878669739
step: 180, loss: 0.31128355860710144
step: 190, loss: 0.34980064630508423
step: 200, loss: 0.21310926973819733
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 4096.0
step: 210, loss: 0.4138714373111725
step: 220, loss: 0.3942685127258301
step: 230, loss: 0.2746538519859314
step: 240, loss: 0.20844529569149017
step: 250, loss: 0.18245738744735718
step: 260, loss: 0.388115793466568
step: 270, loss: 0.1498035043478012
step: 280, loss: 0.24851514399051666
step: 290, loss: 0.29960882663726807
step: 300, loss: 0.30501237511634827

```

```

step: 310, loss: 0.20415829122066498
step: 320, loss: 0.1586577147245407
step: 330, loss: 0.1949453055858612
step: 340, loss: 0.49898141622543335
step: 350, loss: 0.42232924699783325
step: 360, loss: 0.06626772880554199
step: 370, loss: 0.2484072744846344
step: 380, loss: 0.27292001247406006
step: 390, loss: 0.08321573585271835
step: 400, loss: 0.09386128187179565
step: 410, loss: 0.2615313231945038
step: 420, loss: 0.08552486449480057
step: 430, loss: 0.27120092511177063
step: 440, loss: 0.12420050799846649
step: 450, loss: 0.09447978436946869
step: 460, loss: 0.2026575654745102
step: 470, loss: 0.10242994129657745
step: 480, loss: 0.33627262711524963
step: 490, loss: 0.2877534329891205
step: 500, loss: 0.12446290254592896
step: 510, loss: 0.06208726763725281
step: 520, loss: 0.14762504398822784
step: 530, loss: 0.20638899505138397
/usr/local/lib/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=0.8873873873873874, f1=0.8899865289627301,
best_f1=0.8899865289627301
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.11062570661306381
step: 10, loss: 0.07747289538383484
step: 20, loss: 0.09498060494661331
step: 30, loss: 0.15240968763828278
step: 40, loss: 0.21682508289813995
step: 50, loss: 0.19594542682170868
step: 60, loss: 0.09502038359642029
step: 70, loss: 0.09779280424118042
step: 80, loss: 0.07480575144290924
step: 90, loss: 0.04827239736914635
step: 100, loss: 0.3128368556499481
step: 110, loss: 0.13457246124744415
step: 120, loss: 0.041529517620801926

```

```

step: 130, loss: 0.009177293628454208
step: 140, loss: 0.19070592522621155
step: 150, loss: 0.046350881457328796
step: 160, loss: 0.06986777484416962
step: 170, loss: 0.06222371384501457
step: 180, loss: 0.1551697850227356
step: 190, loss: 0.08850813657045364
step: 200, loss: 0.3604755997657776
step: 210, loss: 0.041048564016819
step: 220, loss: 0.044967252761125565
step: 230, loss: 0.05399801954627037
step: 240, loss: 0.17503654956817627
step: 250, loss: 0.0753093957901001
step: 260, loss: 0.20157799124717712
step: 270, loss: 0.011128302663564682
step: 280, loss: 0.13672053813934326
step: 290, loss: 0.06535913795232773
step: 300, loss: 0.07954385876655579
step: 310, loss: 0.1119384691119194
step: 320, loss: 0.08763442188501358
step: 330, loss: 0.07240278273820877
step: 340, loss: 0.16294580698013306
step: 350, loss: 0.0019352161325514317
step: 360, loss: 0.19842633605003357
step: 370, loss: 0.03051593154668808
step: 380, loss: 0.15420982241630554
step: 390, loss: 0.06096306070685387
step: 400, loss: 0.10515051335096359
step: 410, loss: 0.034721747040748596
step: 420, loss: 0.047930944710969925
step: 430, loss: 0.08128444105386734
step: 440, loss: 0.004836108069866896
step: 450, loss: 0.0630631148815155
step: 460, loss: 0.09784143418073654
step: 470, loss: 0.08435256779193878
step: 480, loss: 0.03853129222989082
step: 490, loss: 0.09751105308532715
step: 500, loss: 0.01757550984621048
step: 510, loss: 0.06390825659036636
step: 520, loss: 0.32448071241378784
step: 530, loss: 0.18457920849323273
/usr/local/lib/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=0.9262564584311883, f1=0.9243856332703214,  
best_f1=0.9243856332703214  
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:  
An input tensor was not cuda.  
    warnings.warn("An input tensor was not cuda.")  
step: 0, loss: 0.1903192549943924  
step: 10, loss: 0.09118395298719406  
step: 20, loss: 0.018595103174448013  
step: 30, loss: 0.03843051567673683  
step: 40, loss: 0.059064775705337524  
step: 50, loss: 0.09321951866149902  
step: 60, loss: 0.042590558528900146  
step: 70, loss: 0.033908661454916  
step: 80, loss: 0.05000559613108635  
step: 90, loss: 0.012194342911243439  
step: 100, loss: 0.06109674274921417  
step: 110, loss: 0.033359345048666  
step: 120, loss: 0.29891330003738403  
step: 130, loss: 0.022645827382802963  
step: 140, loss: 0.03667449206113815  
step: 150, loss: 0.02195161022245884  
step: 160, loss: 0.01576647348701954  
step: 170, loss: 0.003685307689011097  
step: 180, loss: 0.0029649685602635145  
step: 190, loss: 0.013046114705502987  
step: 200, loss: 0.02414325810968876  
step: 210, loss: 0.04126037657260895  
step: 220, loss: 0.1918879598379135  
step: 230, loss: 0.020340463146567345  
step: 240, loss: 0.1917167454957962  
step: 250, loss: 0.31999117136001587  
step: 260, loss: 0.07838902622461319  
step: 270, loss: 0.037262026220560074  
step: 280, loss: 0.014096815139055252  
step: 290, loss: 0.010436318814754486  
step: 300, loss: 0.12024743854999542  
step: 310, loss: 0.1187124028801918  
step: 320, loss: 0.010587264783680439  
step: 330, loss: 0.0205099955201149  
step: 340, loss: 0.02322164922952652  
step: 350, loss: 0.08004424721002579  
step: 360, loss: 0.02680172026157379  
step: 370, loss: 0.08683045208454132  
step: 380, loss: 0.013790959492325783  
step: 390, loss: 0.03671788051724434  
step: 400, loss: 0.22413082420825958  
step: 410, loss: 0.19509388506412506  
step: 420, loss: 0.17105814814567566
```

```

step: 430, loss: 0.14976248145103455
step: 440, loss: 0.1711910516023636
step: 450, loss: 0.0277034230530262
step: 460, loss: 0.11984537541866302
step: 470, loss: 0.09056513756513596
step: 480, loss: 0.05698157101869583
step: 490, loss: 0.03825126588344574
step: 500, loss: 0.0403473936021328
step: 510, loss: 0.04397078603506088
step: 520, loss: 0.01773415505886078
step: 530, loss: 0.036605775356292725
/usr/local/lib/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=0.9364055299539169, f1=0.9338805289557685,
best_f1=0.9338805289557685
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.021809175610542297
step: 10, loss: 0.008349471725523472
step: 20, loss: 0.10056591033935547
step: 30, loss: 0.1217704564332962
step: 40, loss: 0.11159799993038177
step: 50, loss: 0.0997292697429657
step: 60, loss: 0.015915581956505775
step: 70, loss: 0.07169957458972931
step: 80, loss: 0.15728643536567688
step: 90, loss: 0.08188606053590775
step: 100, loss: 0.0032182075083255768
step: 110, loss: 0.15593191981315613
step: 120, loss: 0.0367002971470356
step: 130, loss: 0.07446243613958359
step: 140, loss: 0.04683218151330948
step: 150, loss: 0.021707482635974884
step: 160, loss: 0.020772811025381088
step: 170, loss: 0.03198597580194473
step: 180, loss: 0.04885441064834595
step: 190, loss: 0.15189315378665924
step: 200, loss: 0.10126112401485443
step: 210, loss: 0.01151012908667326
step: 220, loss: 0.007686942350119352
step: 230, loss: 0.0365670807659626
step: 240, loss: 0.1606374979019165

```



```

step: 250, loss: 0.07137513160705566
step: 260, loss: 0.008190568536520004
step: 270, loss: 0.14485877752304077
step: 280, loss: 0.009584451094269753
step: 290, loss: 0.1351378709077835
step: 300, loss: 0.006096694152802229
step: 310, loss: 0.009592791087925434
step: 320, loss: 0.1704380214214325
step: 330, loss: 0.04238790273666382
step: 340, loss: 0.04042026400566101
step: 350, loss: 0.10411839932203293
step: 360, loss: 0.08602417260408401
step: 370, loss: 0.01646536774933338
step: 380, loss: 0.007233472540974617
step: 390, loss: 0.005138942506164312
step: 400, loss: 0.08137886226177216
step: 410, loss: 0.00313959876075387
step: 420, loss: 0.016772283241152763
step: 430, loss: 0.010598125867545605
step: 440, loss: 0.014788525179028511
step: 450, loss: 0.08730486035346985
step: 460, loss: 0.027043547481298447
step: 470, loss: 0.008525418117642403
step: 480, loss: 0.017309950664639473
step: 490, loss: 0.0066529554314911366
step: 500, loss: 0.03479648754000664
step: 510, loss: 0.06600208580493927
step: 520, loss: 0.013158275745809078
step: 530, loss: 0.2868507504463196
/usr/local/lib/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=0.9417163836622304, f1=0.9415554532903819,
best_f1=0.9415554532903819
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.016591105610132217
step: 10, loss: 0.013090177439153194
step: 20, loss: 0.0028352041263133287
step: 30, loss: 0.08954416215419769
step: 40, loss: 0.00231374753639102
step: 50, loss: 0.10049070417881012
step: 60, loss: 0.06274169683456421

```

step: 70, loss: 0.014099126681685448  
step: 80, loss: 0.01468626968562603  
step: 90, loss: 0.13205227255821228  
step: 100, loss: 0.12837809324264526  
step: 110, loss: 0.050516288727521896  
step: 120, loss: 0.16950589418411255  
step: 130, loss: 0.012788519263267517  
step: 140, loss: 0.01332942582666874  
step: 150, loss: 0.006788263563066721  
step: 160, loss: 0.03826943412423134  
step: 170, loss: 0.017883814871311188  
step: 180, loss: 0.0045899031683802605  
step: 190, loss: 0.009321051649749279  
step: 200, loss: 0.10454417765140533  
step: 210, loss: 0.004181544296443462  
step: 220, loss: 0.02439878135919571  
step: 230, loss: 0.00548263406381011  
step: 240, loss: 0.026324747130274773  
step: 250, loss: 0.11446328461170197  
step: 260, loss: 0.004196288529783487  
step: 270, loss: 0.01038012932986021  
step: 280, loss: 0.06164684146642685  
step: 290, loss: 0.010100960731506348  
step: 300, loss: 0.15399505198001862  
step: 310, loss: 0.01975315436720848  
step: 320, loss: 0.10164867341518402  
step: 330, loss: 0.007025393191725016  
step: 340, loss: 0.03393206372857094  
step: 350, loss: 0.002754047978669405  
step: 360, loss: 0.032935552299022675  
step: 370, loss: 0.003494834993034601  
step: 380, loss: 0.0010895247105509043  
step: 390, loss: 0.19983386993408203  
step: 400, loss: 0.0086705032736063  
step: 410, loss: 0.0904918909072876  
step: 420, loss: 0.29228121042251587  
step: 430, loss: 0.03949315845966339  
step: 440, loss: 0.017964787781238556  
step: 450, loss: 0.015166228637099266  
step: 460, loss: 0.04770543798804283  
step: 470, loss: 0.03410873934626579  
step: 480, loss: 0.003875946393236518  
step: 490, loss: 0.10742808133363724  
step: 500, loss: 0.12592282891273499  
step: 510, loss: 0.019307613372802734  
step: 520, loss: 0.14442600309848785  
step: 530, loss: 0.03303230181336403  
/usr/local/lib/python3.7/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=0.935602575896964, f1=0.9293954776188279,
best_f1=0.9415554532903819
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.07583492994308472
step: 10, loss: 0.001313654938712716
step: 20, loss: 0.08604860305786133
step: 30, loss: 0.0009305658750236034
step: 40, loss: 0.02747304178774357
step: 50, loss: 0.005726574454456568
step: 60, loss: 0.0083832498639822
step: 70, loss: 0.028390642255544662
step: 80, loss: 0.0016997656784951687
step: 90, loss: 0.1895502507686615
step: 100, loss: 0.01884690299630165
step: 110, loss: 0.0490533821284771
step: 120, loss: 0.10714489966630936
step: 130, loss: 0.006592228543013334
step: 140, loss: 0.012318931519985199
step: 150, loss: 0.014669078402221203
step: 160, loss: 0.17168529331684113
step: 170, loss: 0.002844876842573285
step: 180, loss: 0.04938913509249687
step: 190, loss: 0.08938241004943848
step: 200, loss: 0.02964102104306221
step: 210, loss: 0.01346464641392231
step: 220, loss: 0.01240122877061367
step: 230, loss: 0.003172086551785469
step: 240, loss: 0.006138928234577179
step: 250, loss: 0.09828823804855347
step: 260, loss: 0.013245129957795143
step: 270, loss: 0.003712789388373494
step: 280, loss: 0.3612387180328369
step: 290, loss: 0.03085186704993248
step: 300, loss: 0.1044096127152443
step: 310, loss: 0.10366939753293991
step: 320, loss: 0.060692399740219116
step: 330, loss: 0.060437895357608795
step: 340, loss: 0.0015753858024254441
step: 350, loss: 0.03836807608604431
step: 360, loss: 0.016767559573054314
step: 370, loss: 0.0029313932172954082
step: 380, loss: 0.009838273748755455
step: 390, loss: 0.00489457743242383
step: 400, loss: 0.004069631453603506
```

```

step: 410, loss: 0.0010654703946784139
step: 420, loss: 0.10477600246667862
step: 430, loss: 0.027260569855570793
step: 440, loss: 0.005856361240148544
step: 450, loss: 0.30807021260261536
step: 460, loss: 0.0063782064244151115
step: 470, loss: 0.017063230276107788
step: 480, loss: 0.016350559890270233
step: 490, loss: 0.017181817442178726
step: 500, loss: 0.00832400657236576
step: 510, loss: 0.18241478502750397
step: 520, loss: 0.0018365789437666535
step: 530, loss: 0.05445875972509384
/usr/local/lib/python3.7/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=0.9395604395604396, f1=0.9358560221504384,
best_f1=0.9415554532903819
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0033834530040621758
step: 10, loss: 0.003300223732367158
step: 20, loss: 0.012116929516196251
step: 30, loss: 0.013275359757244587
step: 40, loss: 0.004163894336670637
step: 50, loss: 0.003690431360155344
step: 60, loss: 0.004850556142628193
step: 70, loss: 0.002012822777032852
step: 80, loss: 0.0010235507506877184
step: 90, loss: 0.0003651638689916581
step: 100, loss: 0.0010031969286501408
step: 110, loss: 0.0010051121935248375
step: 120, loss: 0.005091742612421513
step: 130, loss: 0.0012376424856483936
step: 140, loss: 0.0011677740840241313
step: 150, loss: 0.0076930648647248745
step: 160, loss: 0.000865665206220001
step: 170, loss: 0.0022765775211155415
step: 180, loss: 0.0017651169328019023
step: 190, loss: 0.05549779161810875
step: 200, loss: 0.001246499246917665
step: 210, loss: 0.03533460199832916
step: 220, loss: 0.002543083392083645
step: 230, loss: 0.004120625089854002
step: 240, loss: 0.02907324768602848
step: 250, loss: 0.022296089679002762
step: 260, loss: 0.015981871634721756

```

```

step: 270, loss: 0.0037525747902691364
step: 280, loss: 0.0061273882165551186
step: 290, loss: 0.007186051458120346
step: 300, loss: 0.003355608554556966
step: 310, loss: 0.0013554181205108762
step: 320, loss: 0.015397903509438038
step: 330, loss: 0.04188456013798714
step: 340, loss: 0.0036382062826305628
step: 350, loss: 0.014701643027365208
step: 360, loss: 0.010465804487466812
step: 370, loss: 0.003199332393705845
step: 380, loss: 0.0045487224124372005
step: 390, loss: 0.01719762571156025
step: 400, loss: 0.018688786774873734
step: 410, loss: 0.003584792371839285
step: 420, loss: 0.03142545372247696
step: 430, loss: 0.012355919927358627
step: 440, loss: 0.02696460671722889
step: 450, loss: 0.012600237503647804
step: 460, loss: 0.004672457929700613
step: 470, loss: 0.1627059131860733
step: 480, loss: 0.0014130235649645329
step: 490, loss: 0.04284512624144554
step: 500, loss: 0.0066431015729904175
step: 510, loss: 0.00040187835111282766
step: 520, loss: 0.0013381782919168472
step: 530, loss: 0.0079928208142519
/usr/local/lib/python3.7/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=0.9410664172123478, f1=0.9367924528301887,
best_f1=0.9415554532903819
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0011279018362984061
step: 10, loss: 0.0014962858986109495
step: 20, loss: 0.0008744727238081396
step: 30, loss: 0.007873956114053726
step: 40, loss: 0.0007362234173342586
step: 50, loss: 0.018825329840183258
step: 60, loss: 0.0006893634563311934
step: 70, loss: 0.032606981694698334
step: 80, loss: 0.00454552099108696
step: 90, loss: 0.0006607325631193817
step: 100, loss: 0.0015633394941687584
step: 110, loss: 0.0020495187491178513
step: 120, loss: 0.0007348290528170764

```

```

step: 130, loss: 0.009361767210066319
step: 140, loss: 0.00017747646779753268
step: 150, loss: 0.00302584795281291
step: 160, loss: 0.004616708494722843
step: 170, loss: 0.025346355512738228
step: 180, loss: 0.003849806496873498
step: 190, loss: 0.07483373582363129
step: 200, loss: 0.011856251396238804
step: 210, loss: 0.031346652656793594
step: 220, loss: 0.001578828669153154
step: 230, loss: 0.10314050316810608
step: 240, loss: 0.0012338242959231138
step: 250, loss: 0.022657720372080803
step: 260, loss: 0.004154559690505266
step: 270, loss: 0.008133409544825554
step: 280, loss: 0.00048491384950466454
step: 290, loss: 0.005407269112765789
step: 300, loss: 0.0024490824434906244
step: 310, loss: 0.008500608615577221
step: 320, loss: 0.010218190960586071
step: 330, loss: 0.003172549419105053
step: 340, loss: 0.20896556973457336
step: 350, loss: 0.0010341386077925563
step: 360, loss: 0.042944394052028656
step: 370, loss: 0.2518560290336609
step: 380, loss: 0.0015871860086917877
step: 390, loss: 0.054266925901174545
step: 400, loss: 0.0015840455889701843
step: 410, loss: 0.0033459002152085304
step: 420, loss: 0.003892812877893448
step: 430, loss: 0.008917039260268211
step: 440, loss: 0.037978313863277435
step: 450, loss: 0.0009825811721384525
step: 460, loss: 0.013316957280039787
step: 470, loss: 0.04744335636496544
step: 480, loss: 0.030284063890576363
step: 490, loss: 0.007468405179679394
step: 500, loss: 0.001100599986082315
step: 510, loss: 0.012705945409834385
step: 520, loss: 0.0007815302815288305
step: 530, loss: 0.0014081599656492472
/usr/local/lib/python3.7/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=0.9389454209065681, f1=0.9343200740055504,
best_f1=0.9415554532903819
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.

```

```
warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0003078277222812176
step: 10, loss: 0.005519408732652664
step: 20, loss: 0.0010497808689251542
step: 30, loss: 0.08004510402679443
step: 40, loss: 0.00658252090215683
step: 50, loss: 0.002523270668461919
step: 60, loss: 0.004355372861027718
step: 70, loss: 0.017767934128642082
step: 80, loss: 0.19054678082466125
step: 90, loss: 0.02074185572564602
step: 100, loss: 0.011609543114900589
step: 110, loss: 0.10284744203090668
step: 120, loss: 0.01012219674885273
step: 130, loss: 0.006944979541003704
step: 140, loss: 0.006306488532572985
step: 150, loss: 0.001864197081886232
step: 160, loss: 0.0014631727244704962
step: 170, loss: 0.014303985983133316
step: 180, loss: 0.004249137826263905
step: 190, loss: 0.005100009962916374
step: 200, loss: 0.0008990836795419455
step: 210, loss: 0.00442300271242857
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 8192.0
step: 220, loss: 0.039825085550546646
step: 230, loss: 0.00766243739053607
step: 240, loss: 0.0015307711437344551
step: 250, loss: 0.036166705191135406
step: 260, loss: 0.002482626587152481
step: 270, loss: 0.002911878051236272
step: 280, loss: 0.0019110421417281032
step: 290, loss: 0.0014446568675339222
step: 300, loss: 0.0409519299864769
step: 310, loss: 0.0852041095495224
step: 320, loss: 0.0040350970812141895
step: 330, loss: 0.0011423738906159997
step: 340, loss: 0.027526946738362312
step: 350, loss: 0.021209094673395157
step: 360, loss: 0.0009392725769430399
step: 370, loss: 0.018767068162560463
step: 380, loss: 0.0006128295208327472
step: 390, loss: 0.0002824528201017529
step: 400, loss: 0.0006379141705110669
step: 410, loss: 0.0036814615596085787
step: 420, loss: 0.05024314671754837
step: 430, loss: 0.0009043617756105959
step: 440, loss: 0.0013052936410531402
step: 450, loss: 0.07552976906299591
```

```

step: 460, loss: 0.0011985008604824543
step: 470, loss: 0.08134125918149948
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 4096.0
step: 480, loss: 0.009641336277127266
step: 490, loss: 0.05077467858791351
step: 500, loss: 0.0008256454020738602
step: 510, loss: 0.16296496987342834
step: 520, loss: 0.016931304708123207
step: 530, loss: 0.10626773536205292
/usr/local/lib/python3.7/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=0.937962962962963, f1=0.9324074074074072,
best_f1=0.9415554532903819
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.04906793683767319
step: 10, loss: 0.00287756253965199
step: 20, loss: 0.007539424579590559
step: 30, loss: 0.06392665207386017
step: 40, loss: 0.0004820904287043959
step: 50, loss: 0.003284689737483859
step: 60, loss: 0.004142592195421457
step: 70, loss: 0.00250441487878561
step: 80, loss: 0.00048749265260994434
step: 90, loss: 0.0006269093719311059
step: 100, loss: 0.0034182171802967787
step: 110, loss: 0.006890783552080393
step: 120, loss: 0.0004560283268801868
step: 130, loss: 0.0005315706948749721
step: 140, loss: 0.0027985484339296818
step: 150, loss: 0.00786300003528595
step: 160, loss: 0.00862416997551918
step: 170, loss: 0.006375395692884922
step: 180, loss: 0.03938954323530197
step: 190, loss: 0.004402631893754005
step: 200, loss: 0.00559247238561511
step: 210, loss: 0.006996903568506241
step: 220, loss: 0.001382417161948979
step: 230, loss: 0.0015742782270535827
step: 240, loss: 0.004708945285528898
step: 250, loss: 0.003986488562077284
step: 260, loss: 0.01679806411266327
step: 270, loss: 0.001586925471201539
step: 280, loss: 0.06794743239879608
step: 290, loss: 0.0008239576127380133
step: 300, loss: 0.0026861780788749456

```



```

step: 310, loss: 0.012498142197728157
step: 320, loss: 0.008243697695434093
step: 330, loss: 0.00961526483297348
step: 340, loss: 0.013387657701969147
step: 350, loss: 0.007882033474743366
step: 360, loss: 0.0006312616169452667
step: 370, loss: 0.004110563546419144
step: 380, loss: 0.0025605657137930393
step: 390, loss: 0.0017255195416510105
step: 400, loss: 0.00261926488019526
step: 410, loss: 0.002268081996589899
step: 420, loss: 0.0016267335740849376
step: 430, loss: 0.0011677963193506002
step: 440, loss: 0.00458911806344986
step: 450, loss: 0.011162519454956055
step: 460, loss: 0.003583453595638275
step: 470, loss: 0.005406326614320278
step: 480, loss: 0.004217417445033789
step: 490, loss: 0.0007011371199041605
step: 500, loss: 0.014673775061964989
step: 510, loss: 0.0006202482036314905
step: 520, loss: 0.047305796295404434
step: 530, loss: 0.016169199720025063
/usr/local/lib/python3.7/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=0.9413953488372093, f1=0.9393090569561158,
best_f1=0.9415554532903819
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0005823746905662119
step: 10, loss: 0.000523843860719353
step: 20, loss: 0.0005126986070536077
step: 30, loss: 0.0006940575549378991
step: 40, loss: 0.00037494118441827595
step: 50, loss: 0.0003723041736520827
step: 60, loss: 0.00023922241234686226
step: 70, loss: 0.0010346429189667106
step: 80, loss: 0.005117656663060188
step: 90, loss: 0.0006841076537966728
step: 100, loss: 0.001912760315462947
step: 110, loss: 0.0014341720379889011
step: 120, loss: 0.0013383063487708569
step: 130, loss: 0.0007065167301334441
step: 140, loss: 0.000672797323204577
step: 150, loss: 0.0006408043554984033
step: 160, loss: 0.0004737004346679896

```

```

step: 170, loss: 0.0004101844097021967
step: 180, loss: 0.0007849276880733669
step: 190, loss: 0.0010126701090484858
step: 200, loss: 0.0006976303993724287
step: 210, loss: 0.005955908913165331
step: 220, loss: 0.051632050424814224
step: 230, loss: 0.0009289452573284507
step: 240, loss: 0.016429666429758072
step: 250, loss: 0.000794425664935261
step: 260, loss: 0.012510204687714577
step: 270, loss: 0.017591720446944237
step: 280, loss: 0.0046396818943321705
step: 290, loss: 0.0004156610812060535
step: 300, loss: 0.0032960507087409496
step: 310, loss: 0.1351713389158249
step: 320, loss: 0.0034330347552895546
step: 330, loss: 0.00519280880689621
step: 340, loss: 0.12164704501628876
step: 350, loss: 0.003222484840080142
step: 360, loss: 0.03741230070590973
step: 370, loss: 0.002262558788061142
step: 380, loss: 0.005474877078086138
step: 390, loss: 0.023247040808200836
step: 400, loss: 0.03988344594836235
step: 410, loss: 0.0004442122299224138
step: 420, loss: 0.00027656639576889575
step: 430, loss: 0.01256740465760231
step: 440, loss: 0.0014406356494873762
step: 450, loss: 0.012605019845068455
step: 460, loss: 0.13426043093204498
step: 470, loss: 0.0024661605712026358
step: 480, loss: 0.0013525960966944695
step: 490, loss: 0.024168651551008224
step: 500, loss: 0.0016133161261677742
step: 510, loss: 0.16740287840366364
step: 520, loss: 0.007021278142929077
step: 530, loss: 0.015561302192509174
/usr/local/lib/python3.7/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=0.9366130558183537, f1=0.9292161520190024,
best_f1=0.9415554532903819
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0016689012991264462
step: 10, loss: 0.0010071925353258848
step: 20, loss: 0.003251082729548216

```

step: 30, loss: 0.003356020199134946  
step: 40, loss: 0.0023511196486651897  
step: 50, loss: 0.0006876944680698216  
step: 60, loss: 0.0011000022059306502  
step: 70, loss: 0.0021100633312016726  
step: 80, loss: 0.1886577308177948  
step: 90, loss: 0.007049834355711937  
step: 100, loss: 0.18128031492233276  
step: 110, loss: 0.001168899005278945  
step: 120, loss: 0.0024238084442913532  
step: 130, loss: 0.0024719845969229937  
step: 140, loss: 0.0015911223599687219  
step: 150, loss: 0.0027475778479129076  
step: 160, loss: 0.00204587634652853  
step: 170, loss: 0.0009773229248821735  
step: 180, loss: 0.0011672666296362877  
step: 190, loss: 0.0015337271615862846  
step: 200, loss: 0.000917085213586688  
step: 210, loss: 0.008529379963874817  
step: 220, loss: 0.0013468759134411812  
step: 230, loss: 0.0009833935182541609  
step: 240, loss: 0.00830388069152832  
step: 250, loss: 0.0014681443572044373  
step: 260, loss: 0.0009922375902533531  
step: 270, loss: 0.002077815355733037  
step: 280, loss: 0.0014631027588620782  
step: 290, loss: 0.0044363937340676785  
step: 300, loss: 0.03583824634552002  
step: 310, loss: 0.0008683057385496795  
step: 320, loss: 0.04128827527165413  
step: 330, loss: 0.002451017964631319  
step: 340, loss: 0.0013475617161020637  
step: 350, loss: 0.0003412748337723315  
step: 360, loss: 0.008140305057168007  
step: 370, loss: 0.00034883455373346806  
step: 380, loss: 0.0006348313763737679  
step: 390, loss: 0.0036149185616523027  
step: 400, loss: 0.0003921906172763556  
step: 410, loss: 0.0037175181787461042  
step: 420, loss: 0.0008244586060754955  
step: 430, loss: 0.051635753363370895  
step: 440, loss: 0.0009634579182602465  
step: 450, loss: 0.000694962334819138  
step: 460, loss: 0.0011996423127129674  
step: 470, loss: 0.0013294161763042212  
step: 480, loss: 0.0006359133985824883  
step: 490, loss: 0.001525519648566842  
step: 500, loss: 0.003011277411133051

```

step: 510, loss: 0.0019832453690469265
step: 520, loss: 0.0022234790958464146
step: 530, loss: 0.002360390033572912
/usr/local/lib/python3.7/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=0.9391955098222639, f1=0.9378794955628212,
best_f1=0.9415554532903819
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0012053760001435876
step: 10, loss: 0.0006678706267848611
step: 20, loss: 0.0008358751074410975
step: 30, loss: 0.0006113637937232852
step: 40, loss: 0.0013092011213302612
step: 50, loss: 0.002098648576065898
step: 60, loss: 0.0005612092209048569
step: 70, loss: 0.0008079408435150981
step: 80, loss: 0.001194517477415502
step: 90, loss: 0.00802051555365324
step: 100, loss: 0.0003848467895295471
step: 110, loss: 0.00023320868785958737
step: 120, loss: 0.0010269915219396353
step: 130, loss: 0.00029709673253819346
step: 140, loss: 0.0004071070288773626
step: 150, loss: 0.0007041766657494009
step: 160, loss: 0.0012171607231721282
step: 170, loss: 0.016963399946689606
step: 180, loss: 0.0005066000740043819
step: 190, loss: 0.00040693345363251865
step: 200, loss: 0.0003729159652721137
step: 210, loss: 0.0007765774498693645
step: 220, loss: 0.0005180197185836732
step: 230, loss: 0.0009623540681786835
step: 240, loss: 0.044996533542871475
step: 250, loss: 0.0008605021284893155
step: 260, loss: 0.0007920475327409804
step: 270, loss: 0.05069535970687866
step: 280, loss: 0.07792457193136215
step: 290, loss: 0.0019220003159716725
step: 300, loss: 0.0007225267472676933
step: 310, loss: 0.0005473473574966192
step: 320, loss: 0.0004746364429593086
step: 330, loss: 0.0008849063306115568
step: 340, loss: 0.000749381142668426
step: 350, loss: 0.0003407746844459325
step: 360, loss: 0.08996599167585373

```

```
step: 370, loss: 0.002232619561254978
step: 380, loss: 0.00048222194891422987
step: 390, loss: 0.0011339070042595267
step: 400, loss: 0.0013628537999466062
step: 410, loss: 0.009482457302510738
step: 420, loss: 0.0022071211133152246
step: 430, loss: 0.0010441811755299568
step: 440, loss: 0.0003862013982143253
step: 450, loss: 0.0007953918538987637
step: 460, loss: 0.0028176961932331324
step: 470, loss: 0.0008862917311489582
step: 480, loss: 0.00023212826636154205
step: 490, loss: 0.0005622722092084587
step: 500, loss: 0.00167655770201236
step: 510, loss: 0.0004302806919440627
step: 520, loss: 0.0008292876882478595
step: 530, loss: 0.0004408714594319463
/usr/local/lib/python3.7/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=0.9401947148817803, f1=0.9401947148817803,
best_f1=0.9415554532903819
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0005725410883314908
step: 10, loss: 0.00044340186286717653
step: 20, loss: 0.0003778923419304192
step: 30, loss: 0.0010669329203665257
step: 40, loss: 0.000803831615485251
step: 50, loss: 0.18976500630378723
step: 60, loss: 0.0011795349419116974
step: 70, loss: 0.002899534534662962
step: 80, loss: 0.0012996889417991042
step: 90, loss: 0.00047620595432817936
step: 100, loss: 0.0006650651339441538
step: 110, loss: 0.0007223517750389874
step: 120, loss: 0.0006801376584917307
step: 130, loss: 0.0007377103320322931
step: 140, loss: 0.0012169118272140622
step: 150, loss: 0.0010893334401771426
step: 160, loss: 0.0009089026134461164
step: 170, loss: 0.002542782574892044
step: 180, loss: 0.0009990675607696176
step: 190, loss: 0.0022442208137363195
step: 200, loss: 0.0007688028272241354
step: 210, loss: 0.10068739950656891
step: 220, loss: 0.00036619516322389245
```

```

step: 230, loss: 0.0004349799710325897
step: 240, loss: 0.0005866821156814694
step: 250, loss: 0.0010163176339119673
step: 260, loss: 0.00045101065188646317
step: 270, loss: 0.0013106069527566433
step: 280, loss: 0.00048140756553038955
step: 290, loss: 0.00043335993541404605
step: 300, loss: 0.0006685142870992422
step: 310, loss: 0.0006816714303568006
step: 320, loss: 0.0005884146085008979
step: 330, loss: 0.0012344351271167397
step: 340, loss: 0.000622995721641928
step: 350, loss: 0.0005560719291679561
step: 360, loss: 0.0004761535965371877
step: 370, loss: 0.002262966940179467
step: 380, loss: 0.0006786042358726263
step: 390, loss: 0.018310274928808212
step: 400, loss: 0.0018252965528517962
step: 410, loss: 0.00022550803259946406
step: 420, loss: 0.0004490474530030042
step: 430, loss: 0.0007082626689225435
step: 440, loss: 0.001370370271615684
step: 450, loss: 0.0004953920724801719
step: 460, loss: 0.06643135100603104
step: 470, loss: 0.0006735132774338126
step: 480, loss: 0.0005584003520198166
step: 490, loss: 0.0053179520182311535
step: 500, loss: 0.0026392294093966484
step: 510, loss: 0.02076626382768154
step: 520, loss: 0.0005557896220125258
step: 530, loss: 0.0009311999310739338
/usr/local/lib/python3.7/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=0.9398148148148149, f1=0.9361702127659576,
best_f1=0.9415554532903819
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0003984327195212245
step: 10, loss: 0.001379869063384831
step: 20, loss: 0.0009068229701370001
step: 30, loss: 0.019328102469444275
step: 40, loss: 0.000528152973856777
step: 50, loss: 0.0031309095211327076
step: 60, loss: 0.013129675760865211
step: 70, loss: 0.0004537225468084216
step: 80, loss: 0.0005824796389788389

```

```
step: 90, loss: 0.0003354967338964343
step: 100, loss: 0.00024723971728235483
step: 110, loss: 0.0009809760376811028
step: 120, loss: 0.0003289235755801201
step: 130, loss: 0.0003876377595588565
step: 140, loss: 0.0003281343379057944
step: 150, loss: 0.000885112676769495
step: 160, loss: 0.0006801928975619376
step: 170, loss: 0.0002490688639227301
step: 180, loss: 0.0008758906042203307
step: 190, loss: 0.0005710478872060776
step: 200, loss: 0.0006326287984848022
step: 210, loss: 0.00046181725338101387
step: 220, loss: 0.0008845261763781309
step: 230, loss: 0.17080196738243103
step: 240, loss: 0.0008331039571203291
step: 250, loss: 0.00035915596527047455
step: 260, loss: 0.00047308707144111395
step: 270, loss: 0.0007313016685657203
step: 280, loss: 0.0004095419717486948
step: 290, loss: 0.004815749824047089
step: 300, loss: 0.00047927876585163176
step: 310, loss: 0.0563792809844017
step: 320, loss: 0.0009088709484785795
step: 330, loss: 0.0007530645816586912
step: 340, loss: 0.0010717385448515415
step: 350, loss: 0.0008264859789051116
step: 360, loss: 0.012479661963880062
step: 370, loss: 0.0011615701951086521
step: 380, loss: 0.012671885080635548
step: 390, loss: 0.003195052267983556
step: 400, loss: 0.008401889353990555
step: 410, loss: 0.0005757251055911183
step: 420, loss: 0.0005271135596558452
step: 430, loss: 0.00020574375230353326
step: 440, loss: 0.00045562098966911435
step: 450, loss: 0.054988522082567215
step: 460, loss: 0.0008192822569981217
step: 470, loss: 0.0004178787348791957
step: 480, loss: 0.007287952117621899
step: 490, loss: 0.004349939059466124
step: 500, loss: 0.0005822041421197355
step: 510, loss: 0.0007644628640264273
step: 520, loss: 0.0005267994711175561
step: 530, loss: 0.0008426797576248646
/usr/local/lib/python3.7/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=0.9408502772643254, f1=0.9367205542725173,
best_f1=0.9415554532903819
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, 177.62it/s]
```

```
load_f1 = 0.9377901578458682
```

```
real_f1 = 0.9299242424242424
```

```
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 [00:28, 152.59it/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: 0.43686917424201965
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=0.3943661971830986, f1=0.2826086956521739,
best_f1=0.2826086956521739
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.4203495681285858
/usr/local/lib/python3.7/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=0.345679012345679, f1=0.27184466019417475,

```

```

best_f1=0.2826086956521739
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.3736698031425476
/usr/local/lib/python3.7/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=0.375, f1=0.3728813559322034, best_f1=0.2826086956521739
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.25810903310775757
/usr/local/lib/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=0.5106382978723404, f1=0.3823529411764706,
best_f1=0.3823529411764706
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.3034130930900574
/usr/local/lib/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=0.5714285714285715, f1=0.3943661971830986,
best_f1=0.3943661971830986
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.2988949418067932
/usr/local/lib/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=0.5777777777777778, f1=0.45901639344262296,
best_f1=0.45901639344262296

```

```

/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.4438624978065491
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 7: dev_f1=0.8, f1=0.7000000000000001, best_f1=0.7000000000000001
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.37235701084136963
/usr/local/lib/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=0.8571428571428571, f1=0.6666666666666666,
best_f1=0.6666666666666666
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.15649902820587158
/usr/local/lib/python3.7/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=0.8125000000000001, f1=0.6511627906976745,
best_f1=0.6666666666666666
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.16246235370635986
/usr/local/lib/python3.7/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=0.8125000000000001, f1=0.7000000000000001,
best_f1=0.6666666666666666
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.14130587875843048

```

```

/usr/local/lib/python3.7/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=0.8387096774193549, f1=0.7368421052631579,
best_f1=0.6666666666666666
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.09526408463716507
/usr/local/lib/python3.7/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=0.8125000000000001, f1=0.7000000000000001,
best_f1=0.6666666666666666
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.03637118265032768
/usr/local/lib/python3.7/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=0.8125000000000001, f1=0.7000000000000001,
best_f1=0.6666666666666666
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.029821613803505898
/usr/local/lib/python3.7/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=0.8125000000000001, f1=0.7000000000000001,
best_f1=0.6666666666666666
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.04043501615524292
/usr/local/lib/python3.7/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=0.8125000000000001, f1=0.7000000000000001,
best_f1=0.6666666666666666
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, 125018.56it/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 = 0.7142857142857143
real_f1 = 0.7096774193548386
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 [00:29, 151.28it/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: 0.601436972618103
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 32768.0
step: 10, loss: 0.438076376914978
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 20, loss: 0.49649977684020996
step: 30, loss: 0.34297990798950195
step: 40, loss: 0.341727077960968
step: 50, loss: 0.5359809398651123
step: 60, loss: 0.4142535626888275
step: 70, loss: 0.1777476817369461
step: 80, loss: 0.3532622158527374
step: 90, loss: 0.2232976257801056
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 8192.0
step: 100, loss: 0.12028683722019196
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 4096.0
step: 110, loss: 0.1012941300868988
step: 120, loss: 0.10007330775260925
step: 130, loss: 0.02151414193212986
step: 140, loss: 0.033402200788259506
step: 150, loss: 0.20620006322860718
step: 160, loss: 0.00897417776286602
step: 170, loss: 0.07680325955152512
step: 180, loss: 0.044045403599739075
step: 190, loss: 0.03369678556919098
step: 200, loss: 0.0826323851943016
step: 210, loss: 0.06161903962492943
step: 220, loss: 0.1648285686969757
step: 230, loss: 0.021617693826556206
/usr/local/lib/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=0.9527027027027027, f1=0.9515219842164601,
best_f1=0.9515219842164601
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0341973602771759
step: 10, loss: 0.1507985144853592
step: 20, loss: 0.01922895386815071
step: 30, loss: 0.05873505398631096
step: 40, loss: 0.03399426490068436
step: 50, loss: 0.013065723702311516
step: 60, loss: 0.055080752819776535
step: 70, loss: 0.06769652664661407

```



```

step: 80, loss: 0.01068596076220274
step: 90, loss: 0.009098054841160774
step: 100, loss: 0.003219601698219776
step: 110, loss: 0.0016680205008015037
step: 120, loss: 0.10647735744714737
step: 130, loss: 0.0833936408162117
step: 140, loss: 0.010867338627576828
step: 150, loss: 0.07515305280685425
step: 160, loss: 0.007468657102435827
step: 170, loss: 0.0069259218871593475
step: 180, loss: 0.0019741731230169535
step: 190, loss: 0.00836250837892294
step: 200, loss: 0.0017911904724314809
step: 210, loss: 0.043774332851171494
step: 220, loss: 0.0032900446094572544
step: 230, loss: 0.00191565474960953
/usr/local/lib/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=0.9819413092550789, f1=0.9783352337514253,
best_f1=0.9783352337514253
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.021435877308249474
step: 10, loss: 0.013371615670621395
step: 20, loss: 0.002932582050561905
step: 30, loss: 0.00734708970412612
step: 40, loss: 0.001492952462285757
step: 50, loss: 0.21232353150844574
step: 60, loss: 0.003890708787366748
step: 70, loss: 0.0022064168006181717
step: 80, loss: 0.08829393982887268
step: 90, loss: 0.003784741275012493
step: 100, loss: 0.001993450103327632
step: 110, loss: 0.010266194120049477
step: 120, loss: 0.019342990592122078
step: 130, loss: 0.09725482016801834
step: 140, loss: 0.004704549442976713
step: 150, loss: 0.0256541445851326
step: 160, loss: 0.0022519396152347326
step: 170, loss: 0.00056865019723773
step: 180, loss: 0.00254240189678967
step: 190, loss: 0.022883053869009018

```

```

step: 200, loss: 0.01701510325074196
step: 210, loss: 0.0019769780337810516
step: 220, loss: 0.004315390717238188
step: 230, loss: 0.005524769425392151
/usr/local/lib/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=0.9853768278965129, f1=0.9853107344632768,
best_f1=0.9853107344632768
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0031364860478788614
step: 10, loss: 0.006892392877489328
step: 20, loss: 0.019987359642982483
step: 30, loss: 0.00163277518004179
step: 40, loss: 0.064165860414505
step: 50, loss: 0.004708753898739815
step: 60, loss: 0.004159587901085615
step: 70, loss: 0.013584217987954617
step: 80, loss: 0.0020517325028777122
step: 90, loss: 0.019308796152472496
step: 100, loss: 0.0014202488819137216
step: 110, loss: 0.0008548916666768491
step: 120, loss: 0.017783280462026596
step: 130, loss: 0.022060109302401543
step: 140, loss: 0.004942999221384525
step: 150, loss: 0.00036611707764677703
step: 160, loss: 0.0011471848702058196
step: 170, loss: 0.0060946084558963776
step: 180, loss: 0.16200527548789978
step: 190, loss: 0.0060466593131423
step: 200, loss: 0.007718529086560011
step: 210, loss: 0.0019657891243696213
step: 220, loss: 0.0016628705197945237
step: 230, loss: 0.009454834274947643
/usr/local/lib/python3.7/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=0.9797752808988766, f1=0.9751131221719457,
best_f1=0.9853107344632768
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")

```

```

step: 0, loss: 0.0008188868523575366
step: 10, loss: 0.0036372686736285686
step: 20, loss: 0.0034532155841588974
step: 30, loss: 0.0009713476174511015
step: 40, loss: 0.0017691768007352948
step: 50, loss: 0.0030763985123485327
step: 60, loss: 0.10631097853183746
step: 70, loss: 0.10106409341096878
step: 80, loss: 0.06382296979427338
step: 90, loss: 0.040915172547101974
step: 100, loss: 0.0008599931024946272
step: 110, loss: 0.016456928104162216
step: 120, loss: 0.0012922856258228421
step: 130, loss: 0.0010909802513197064
step: 140, loss: 0.0017977155512198806
step: 150, loss: 0.025217385962605476
step: 160, loss: 0.001615031622350216
step: 170, loss: 0.005953832529485226
step: 180, loss: 0.02786378376185894
step: 190, loss: 0.0220523402094841
step: 200, loss: 0.018297597765922546
step: 210, loss: 0.01199384219944477
step: 220, loss: 0.004687993787229061
step: 230, loss: 0.012709815986454487
/usr/local/lib/python3.7/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=0.983277591973244, f1=0.9843400447427293,
best_f1=0.9853107344632768
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.000945588864851743
step: 10, loss: 0.02790111117064953
step: 20, loss: 0.0015001478604972363
step: 30, loss: 0.0008111781207844615
step: 40, loss: 0.0004514936590567231
step: 50, loss: 0.0009109880775213242
step: 60, loss: 0.003415618324652314
step: 70, loss: 0.006854836829006672
step: 80, loss: 0.00849850382655859
step: 90, loss: 0.009211290627717972
step: 100, loss: 0.0007982142269611359
step: 110, loss: 0.0690777450799942
step: 120, loss: 0.0014174483949318528
step: 130, loss: 0.007997419685125351
step: 140, loss: 0.0007289572968147695
step: 150, loss: 0.0009961440227925777

```

```

step: 160, loss: 0.026615289971232414
step: 170, loss: 0.01155075617134571
step: 180, loss: 0.001476111588999629
step: 190, loss: 0.0025494364090263844
step: 200, loss: 0.004254000727087259
step: 210, loss: 0.02038438804447651
step: 220, loss: 0.05148817226290703
step: 230, loss: 0.003176068188622594
/usr/local/lib/python3.7/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=0.9830890642615557, f1=0.976054732041049,
best_f1=0.9853107344632768
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.003814094001427293
step: 10, loss: 0.0009830944472923875
step: 20, loss: 0.0005213734111748636
step: 30, loss: 0.0006966931978240609
step: 40, loss: 0.0029010269790887833
step: 50, loss: 0.002641932340338826
step: 60, loss: 0.0005059046088717878
step: 70, loss: 0.0014130532508715987
step: 80, loss: 0.0012669936986640096
step: 90, loss: 0.04545332118868828
step: 100, loss: 0.0011427513090893626
step: 110, loss: 0.0026788462419062853
step: 120, loss: 0.0038386215455830097
step: 130, loss: 0.0043459320440888405
step: 140, loss: 0.000896155193913728
step: 150, loss: 0.027404244989156723
step: 160, loss: 0.0011177826672792435
step: 170, loss: 0.023799598217010498
step: 180, loss: 0.0005410680314525962
step: 190, loss: 0.006193355657160282
step: 200, loss: 0.01069674827158451
step: 210, loss: 0.01706995628774166
step: 220, loss: 0.00039975775871425867
step: 230, loss: 0.0004880241467617452
/usr/local/lib/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=0.9866369710467707, f1=0.9821428571428571,

```

```

best_f1=0.9821428571428571
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0006508505321107805
step: 10, loss: 0.0064972504042088985
step: 20, loss: 0.002296508988365531
step: 30, loss: 0.0064212060533463955
step: 40, loss: 0.0015996466390788555
step: 50, loss: 0.0018677779007703066
step: 60, loss: 0.0007823217893019319
step: 70, loss: 0.00020107986347284168
step: 80, loss: 0.0008340264903381467
step: 90, loss: 0.0007843866478651762
step: 100, loss: 0.018373508006334305
step: 110, loss: 0.000495673215482384
step: 120, loss: 0.000768486235756427
step: 130, loss: 0.0026725674979388714
step: 140, loss: 0.0006584412767551839
step: 150, loss: 0.09081470221281052
step: 160, loss: 0.017400117591023445
step: 170, loss: 0.03162234276533127
step: 180, loss: 0.0006535037537105381
step: 190, loss: 0.05729709193110466
step: 200, loss: 0.025675207376480103
step: 210, loss: 0.0032567856833338737
step: 220, loss: 0.003949559759348631
step: 230, loss: 0.00039063376607373357
/usr/local/lib/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=0.9876543209876544, 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: 0.03715415298938751
step: 10, loss: 0.0028109776321798563
step: 20, loss: 0.002743246266618371
step: 30, loss: 0.0018276766641065478
step: 40, loss: 0.008672214113175869
step: 50, loss: 0.0017756689339876175
step: 60, loss: 0.0010865116491913795
step: 70, loss: 0.2042893022298813

```

```

step: 80, loss: 0.004293039441108704
step: 90, loss: 0.0068870424292981625
step: 100, loss: 0.0005076596862636507
step: 110, loss: 0.0004555476480163634
step: 120, loss: 0.014038858003914356
step: 130, loss: 0.0007709752535447478
step: 140, loss: 0.0010096885962411761
step: 150, loss: 0.0015745348064228892
step: 160, loss: 0.007854324765503407
step: 170, loss: 0.0002579969877842814
step: 180, loss: 0.0007386141223832965
step: 190, loss: 0.0002623525506351143
step: 200, loss: 0.0007638201350346208
step: 210, loss: 0.001573361805640161
step: 220, loss: 0.0009297685464844108
step: 230, loss: 0.0010098294587805867
/usr/local/lib/python3.7/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=0.9876543209876544, 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: 0.0008185406913980842
step: 10, loss: 0.0010524599347263575
step: 20, loss: 0.000683176564052701
step: 30, loss: 0.00022073967556934804
step: 40, loss: 0.002787386067211628
step: 50, loss: 0.0001962301175808534
step: 60, loss: 0.0010282143484801054
step: 70, loss: 0.0011245274217799306
step: 80, loss: 0.0003850744978990406
step: 90, loss: 0.000617283396422863
step: 100, loss: 0.00022502342471852899
step: 110, loss: 0.04433847591280937
step: 120, loss: 0.00017904682317748666
step: 130, loss: 0.0003715687489602715
step: 140, loss: 0.002391585148870945
step: 150, loss: 0.0010426640510559082
step: 160, loss: 7.873111462686211e-05
step: 170, loss: 0.00029463093960657716
step: 180, loss: 0.001989183248952031
step: 190, loss: 0.001084622461348772
step: 200, loss: 0.0012138671008870006
step: 210, loss: 0.0042327274568378925
step: 220, loss: 0.03871351480484009
step: 230, loss: 0.0007156369392760098

```

```

/usr/local/lib/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=0.9887640449438202, f1=0.9841986455981941,
best_f1=0.9841986455981941
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.000320831430144608
step: 10, loss: 0.0005528887850232422
step: 20, loss: 0.0003988258249592036
step: 30, loss: 0.0010666688904166222
step: 40, loss: 5.7037985243368894e-05
step: 50, loss: 0.00019516557222232223
step: 60, loss: 0.011149386875331402
step: 70, loss: 0.0001831393310567364
step: 80, loss: 0.034148313105106354
step: 90, loss: 0.18987905979156494
step: 100, loss: 0.0006712324684485793
step: 110, loss: 0.002592929406091571
step: 120, loss: 0.0006136226584203541
step: 130, loss: 0.0003511887334752828
step: 140, loss: 0.0021084609907120466
step: 150, loss: 0.0006460124859586358
step: 160, loss: 0.0017138270195573568
step: 170, loss: 0.000745861092582345
step: 180, loss: 0.004192701540887356
step: 190, loss: 0.0006881873123347759
step: 200, loss: 0.0052424268797039986
step: 210, loss: 0.0008760349592193961
step: 220, loss: 0.03435801342129707
step: 230, loss: 0.00098376942332834
/usr/local/lib/python3.7/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=0.9876819708846584, f1=0.9865168539325843,
best_f1=0.9841986455981941
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0012395662488415837
step: 10, loss: 0.00043616723269224167
step: 20, loss: 0.019214246422052383
step: 30, loss: 0.013919529505074024

```

```

step: 40, loss: 0.002620579907670617
step: 50, loss: 0.0020379803609102964
step: 60, loss: 0.001431406126357615
step: 70, loss: 0.0016249152831733227
step: 80, loss: 0.0005633291439153254
step: 90, loss: 0.002284703077748418
step: 100, loss: 0.00047561427345499396
step: 110, loss: 0.0003077318542636931
step: 120, loss: 0.000613208394497633
step: 130, loss: 0.0004718206764664501
step: 140, loss: 0.0009639980271458626
step: 150, loss: 0.0007547828718088567
step: 160, loss: 0.0022861575707793236
step: 170, loss: 0.0007839511963538826
step: 180, loss: 0.0005501165287569165
step: 190, loss: 0.006864774040877819
step: 200, loss: 0.0004032420401927084
step: 210, loss: 0.006938362028449774
step: 220, loss: 0.008938937447965145
step: 230, loss: 0.0007575266645289958
/usr/local/lib/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=0.9887892376681614, f1=0.9841986455981941,
best_f1=0.9841986455981941
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.002165335463359952
step: 10, loss: 0.0007073446759022772
step: 20, loss: 0.0011911054607480764
step: 30, loss: 0.005970723927021027
step: 40, loss: 0.0007538119680248201
step: 50, loss: 0.00723624462261796
step: 60, loss: 0.0003951282415073365
step: 70, loss: 0.028048550710082054
step: 80, loss: 0.0016286105383187532
step: 90, loss: 0.0009622075594961643
step: 100, loss: 0.0058470843359827995
step: 110, loss: 0.0010559860384091735
step: 120, loss: 0.004100009799003601
step: 130, loss: 0.0005801782826893032
step: 140, loss: 0.0005985188181512058
step: 150, loss: 0.005566069856286049

```



```

step: 160, loss: 0.0012505323393270373
step: 170, loss: 0.0013845707289874554
step: 180, loss: 0.026494689285755157
step: 190, loss: 0.0012888303026556969
step: 200, loss: 0.00015976183931343257
step: 210, loss: 0.0010490627028048038
step: 220, loss: 0.0005887798615731299
step: 230, loss: 0.000607528374530375
/usr/local/lib/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=0.9898762654668166, f1=0.9853107344632768,
best_f1=0.9853107344632768
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0004006966482847929
step: 10, loss: 0.0005075340159237385
step: 20, loss: 0.0005587634514085948
step: 30, loss: 0.0012180466437712312
step: 40, loss: 0.0015400117263197899
step: 50, loss: 0.0004640059487428516
step: 60, loss: 0.00041065734694711864
step: 70, loss: 0.0007803748594596982
step: 80, loss: 0.0006839905981905758
step: 90, loss: 0.004675634671002626
step: 100, loss: 0.0016163254622370005
step: 110, loss: 0.012905655428767204
step: 120, loss: 0.0001480287901358679
step: 130, loss: 0.00103627925273031
step: 140, loss: 0.00044894011807627976
step: 150, loss: 0.00028247429872862995
step: 160, loss: 0.0009209493873640895
step: 170, loss: 0.0005593475070782006
step: 180, loss: 0.00038031325675547123
step: 190, loss: 0.0004829212848562747
step: 200, loss: 0.0005914190551266074
step: 210, loss: 0.0003386717871762812
step: 220, loss: 0.0010005729272961617
step: 230, loss: 0.0003631417639553547
/usr/local/lib/python3.7/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=0.9854748603351955, f1=0.9865168539325843,

```

```

best_f1=0.9853107344632768
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.004545886069536209
step: 10, loss: 0.000571529264561832
step: 20, loss: 0.0009373924694955349
step: 30, loss: 0.00032586450106464326
step: 40, loss: 0.00017783220391720533
step: 50, loss: 0.0002621747553348541
step: 60, loss: 0.04307000711560249
step: 70, loss: 0.0005605124752037227
step: 80, loss: 0.0009366576559841633
step: 90, loss: 0.0005454298225231469
step: 100, loss: 0.00018407913739793003
step: 110, loss: 0.0005265597719699144
step: 120, loss: 0.16194123029708862
step: 130, loss: 0.0006033480749465525
step: 140, loss: 0.007320030592381954
step: 150, loss: 0.0006643315427936614
step: 160, loss: 0.0044508823193609715
step: 170, loss: 0.00031144253443926573
step: 180, loss: 0.0008411650196649134
step: 190, loss: 0.0005973249790258706
step: 200, loss: 0.0017487986478954554
step: 210, loss: 0.08493306487798691
step: 220, loss: 0.0020809208508580923
step: 230, loss: 0.0005402371753007174
/usr/local/lib/python3.7/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=0.9876265466816648, f1=0.9830124575311437,
best_f1=0.9853107344632768
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 [00:15, 156.54it/s]
load_f1 = 0.9888392857142857
real_f1 = 0.9865771812080537
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 [00:30, 144.75it/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/

```

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: 0.6235836744308472
```

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: 0.4949053227901459
```

```
step: 20, loss: 0.3340202569961548
```

```
step: 30, loss: 0.4121380150318146
```

```
step: 40, loss: 0.3581545650959015
```

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

step: 50, loss: 0.6965097188949585  
step: 60, loss: 0.24889063835144043  
step: 70, loss: 0.24465960264205933  
step: 80, loss: 0.09544917196035385  
step: 90, loss: 0.1156267300248146  
step: 100, loss: 0.23837728798389435  
step: 110, loss: 0.15472526848316193  
step: 120, loss: 0.11126413941383362  
step: 130, loss: 0.1844794750213623  
step: 140, loss: 0.3023279011249542  
step: 150, loss: 0.08560624718666077  
step: 160, loss: 0.10514306277036667  
step: 170, loss: 0.04337035492062569  
step: 180, loss: 0.10512343049049377  
step: 190, loss: 0.20227575302124023  
step: 200, loss: 0.07045090943574905  
step: 210, loss: 0.05319429188966751  
step: 220, loss: 0.03645522892475128  
step: 230, loss: 0.21052931249141693  
step: 240, loss: 0.09561100602149963  
step: 250, loss: 0.014085793867707253  
step: 260, loss: 0.4553062915802002  
step: 270, loss: 0.15290024876594543  
step: 280, loss: 0.048270802944898605  
step: 290, loss: 0.10857444256544113  
step: 300, loss: 0.09599959850311279  
step: 310, loss: 0.218829944729805  
step: 320, loss: 0.1403767168521881  
step: 330, loss: 0.110983707010746  
step: 340, loss: 0.32631105184555054  
step: 350, loss: 0.10203611850738525  
step: 360, loss: 0.10357322543859482  
step: 370, loss: 0.022863758727908134  
step: 380, loss: 0.16043059527873993  
step: 390, loss: 0.018944796174764633  
step: 400, loss: 0.06757479161024094  
step: 410, loss: 0.2513650953769684  
step: 420, loss: 0.025812076404690742  
step: 430, loss: 0.06081555783748627  
step: 440, loss: 0.10213037580251694  
step: 450, loss: 0.09652339667081833  
step: 460, loss: 0.04755681753158569  
step: 470, loss: 0.044860854744911194  
step: 480, loss: 0.11536581814289093  
step: 490, loss: 0.1033681109547615  
step: 500, loss: 0.16103078424930573  
step: 510, loss: 0.06694255769252777  
step: 520, loss: 0.10467365384101868

```

step: 530, loss: 0.010547861456871033
/usr/local/lib/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=0.935258500232883, f1=0.9306197964847363,
best_f1=0.9306197964847363
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.09582428634166718
step: 10, loss: 0.018332544714212418
step: 20, loss: 0.021919649094343185
step: 30, loss: 0.1496172547340393
step: 40, loss: 0.1462516337633133
step: 50, loss: 0.044466353952884674
step: 60, loss: 0.04049420729279518
step: 70, loss: 0.019451886415481567
step: 80, loss: 0.04054459556937218
step: 90, loss: 0.08869384974241257
step: 100, loss: 0.05370495095849037
step: 110, loss: 0.06632164865732193
step: 120, loss: 0.20177429914474487
step: 130, loss: 0.008638659492135048
step: 140, loss: 0.08073996752500534
step: 150, loss: 0.007055394817143679
step: 160, loss: 0.030512487515807152
step: 170, loss: 0.20308759808540344
step: 180, loss: 0.05874735489487648
step: 190, loss: 0.042556680738925934
step: 200, loss: 0.26576703786849976
step: 210, loss: 0.06991517543792725
step: 220, loss: 0.003947316203266382
step: 230, loss: 0.004117238335311413
step: 240, loss: 0.10693518817424774
step: 250, loss: 0.015532131306827068
step: 260, loss: 0.05641206353902817
step: 270, loss: 0.02098824456334114
step: 280, loss: 0.03441765159368515
step: 290, loss: 0.08671742677688599
step: 300, loss: 0.032240044325590134
step: 310, loss: 0.033461686223745346
step: 320, loss: 0.028369951993227005
step: 330, loss: 0.1982576996088028
step: 340, loss: 0.2154167890548706

```

```

step: 350, loss: 0.005421780049800873
step: 360, loss: 0.05652381107211113
step: 370, loss: 0.01398501917719841
step: 380, loss: 0.11026860028505325
step: 390, loss: 0.003193051554262638
step: 400, loss: 0.07814624160528183
step: 410, loss: 0.01275002770125866
step: 420, loss: 0.00949656218290329
step: 430, loss: 0.027875907719135284
step: 440, loss: 0.003884970908984542
step: 450, loss: 0.15890249609947205
step: 460, loss: 0.06907449662685394
step: 470, loss: 0.0959114357829094
step: 480, loss: 0.006042566616088152
step: 490, loss: 0.14660440385341644
step: 500, loss: 0.02451857179403305
step: 510, loss: 0.016116462647914886
step: 520, loss: 0.4197255074977875
step: 530, loss: 0.08540182560682297
/usr/local/lib/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=0.945387792565397, f1=0.9414434861552429,
best_f1=0.9414434861552429
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.05260690674185753
step: 10, loss: 0.0632961094379425
step: 20, loss: 0.005261018872261047
step: 30, loss: 0.02502172254025936
step: 40, loss: 0.028407253324985504
step: 50, loss: 0.10066527128219604
step: 60, loss: 0.009313737042248249
step: 70, loss: 0.019796079024672508
step: 80, loss: 0.15216825902462006
step: 90, loss: 0.003626737743616104
step: 100, loss: 0.06222368776798248
step: 110, loss: 0.03705534338951111
step: 120, loss: 0.10895058512687683
step: 130, loss: 0.13897910714149475
step: 140, loss: 0.006940280552953482
step: 150, loss: 0.20764833688735962
step: 160, loss: 0.03386836126446724

```

```

step: 170, loss: 0.007794264703989029
step: 180, loss: 0.047894712537527084
step: 190, loss: 0.003980153240263462
step: 200, loss: 0.010016162879765034
step: 210, loss: 0.041795071214437485
step: 220, loss: 0.1832878589630127
step: 230, loss: 0.03317577391862869
step: 240, loss: 0.09732095152139664
step: 250, loss: 0.0949598029255867
step: 260, loss: 0.14024777710437775
step: 270, loss: 0.01759517751634121
step: 280, loss: 0.0073523554019629955
step: 290, loss: 0.012108864262700081
step: 300, loss: 0.09706930816173553
step: 310, loss: 0.025597859174013138
step: 320, loss: 0.15774889290332794
step: 330, loss: 0.0009175438899546862
step: 340, loss: 0.018206311389803886
step: 350, loss: 0.0507844015955925
step: 360, loss: 0.12619304656982422
step: 370, loss: 0.018044201657176018
step: 380, loss: 0.12240937352180481
step: 390, loss: 0.021274380385875702
step: 400, loss: 0.28800106048583984
step: 410, loss: 0.06016786769032478
step: 420, loss: 0.021100779995322227
step: 430, loss: 0.03653917461633682
step: 440, loss: 0.191927969455719
step: 450, loss: 0.02814873866736889
step: 460, loss: 0.09389020502567291
step: 470, loss: 0.11802579462528229
step: 480, loss: 0.08167220652103424
step: 490, loss: 0.04813679680228233
step: 500, loss: 0.017613140866160393
step: 510, loss: 0.018917903304100037
step: 520, loss: 0.007866577245295048
step: 530, loss: 0.0075886426493525505
/usr/local/lib/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=0.9493908153701968, f1=0.9411764705882353,
best_f1=0.9411764705882353
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.

```



```
warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.028023866936564445
step: 10, loss: 0.0024997955188155174
step: 20, loss: 0.07087399065494537
step: 30, loss: 0.0022688601166009903
step: 40, loss: 0.01610160991549492
step: 50, loss: 0.045915354043245316
step: 60, loss: 0.005387059412896633
step: 70, loss: 0.08782260119915009
step: 80, loss: 0.0449925996363163
step: 90, loss: 0.004318659659475088
step: 100, loss: 0.002604636363685131
step: 110, loss: 0.04165353253483772
step: 120, loss: 0.006599645595997572
step: 130, loss: 0.06432680040597916
step: 140, loss: 0.02352091670036316
step: 150, loss: 0.04949076473712921
step: 160, loss: 0.005528190638870001
step: 170, loss: 0.04989666864275932
step: 180, loss: 0.11802433431148529
step: 190, loss: 0.08080728352069855
step: 200, loss: 0.09884238243103027
step: 210, loss: 0.030457496643066406
step: 220, loss: 0.017180971801280975
step: 230, loss: 0.043381791561841965
step: 240, loss: 0.0018419591942802072
step: 250, loss: 0.14692561328411102
step: 260, loss: 0.0054425448179244995
step: 270, loss: 0.040234703570604324
step: 280, loss: 0.011392883025109768
step: 290, loss: 0.009281409904360771
step: 300, loss: 0.00291931489482522
step: 310, loss: 0.001789191970601678
step: 320, loss: 0.1697242558002472
step: 330, loss: 0.08307495713233948
step: 340, loss: 0.0006609621341340244
step: 350, loss: 0.08502323180437088
step: 360, loss: 0.09184720367193222
step: 370, loss: 0.025057155638933182
step: 380, loss: 0.007504351902753115
step: 390, loss: 0.0009105968056246638
step: 400, loss: 0.019754279404878616
step: 410, loss: 0.0023252139799296856
step: 420, loss: 0.013550798408687115
step: 430, loss: 0.028574461117386818
step: 440, loss: 0.019593805074691772
step: 450, loss: 0.01832343451678753
step: 460, loss: 0.04798406362533569
```

```

step: 470, loss: 0.004980182740837336
step: 480, loss: 0.0021088537760078907
step: 490, loss: 0.001919249538332224
step: 500, loss: 0.014987965114414692
step: 510, loss: 0.019153447821736336
step: 520, loss: 0.09914856404066086
step: 530, loss: 0.048940956592559814
/usr/local/lib/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=0.9510166358595193, f1=0.9436555199267065,
best_f1=0.9436555199267065
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.02266533672809601
step: 10, loss: 0.009817478246986866
step: 20, loss: 0.00254686176776886
step: 30, loss: 0.02472858875989914
step: 40, loss: 0.004479541443288326
step: 50, loss: 0.025629179552197456
step: 60, loss: 0.07385016977787018
step: 70, loss: 0.020811205729842186
step: 80, loss: 0.011165948584675789
step: 90, loss: 0.20665138959884644
step: 100, loss: 0.07086271047592163
step: 110, loss: 0.016911599785089493
step: 120, loss: 0.20884905755519867
step: 130, loss: 0.017984818667173386
step: 140, loss: 0.11472906917333603
step: 150, loss: 0.012680836021900177
step: 160, loss: 0.032800011336803436
step: 170, loss: 0.17019322514533997
step: 180, loss: 0.005546013358980417
step: 190, loss: 0.005796527024358511
step: 200, loss: 0.11048237234354019
step: 210, loss: 0.004921692423522472
step: 220, loss: 0.005955337546765804
step: 230, loss: 0.0008949417970143259
step: 240, loss: 0.0015959113370627165
step: 250, loss: 0.07869889587163925
step: 260, loss: 0.011438580229878426
step: 270, loss: 0.004259144421666861
step: 280, loss: 0.025299008935689926

```

```

step: 290, loss: 0.07353639602661133
step: 300, loss: 0.17018449306488037
step: 310, loss: 0.12413433194160461
step: 320, loss: 0.10103917121887207
step: 330, loss: 0.005605505313724279
step: 340, loss: 0.007850767113268375
step: 350, loss: 0.0010093750897794962
step: 360, loss: 0.0025097569450736046
step: 370, loss: 0.000961722747888416
step: 380, loss: 0.0002694054855965078
step: 390, loss: 0.028902972117066383
step: 400, loss: 0.02707536704838276
step: 410, loss: 0.031101934611797333
step: 420, loss: 0.164592906832695
step: 430, loss: 0.012056000530719757
step: 440, loss: 0.005521767772734165
step: 450, loss: 0.007034240290522575
step: 460, loss: 0.009679294191300869
step: 470, loss: 0.006590495351701975
step: 480, loss: 0.25284844636917114
step: 490, loss: 0.03138194978237152
step: 500, loss: 0.008004941046237946
step: 510, loss: 0.03136828541755676
step: 520, loss: 0.2608211636543274
step: 530, loss: 0.007124943193048239
/usr/local/lib/python3.7/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=0.9452181987000929, f1=0.9441624365482235,
best_f1=0.9436555199267065
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.021157963201403618
step: 10, loss: 0.0011301658814772964
step: 20, loss: 0.001694376114755869
step: 30, loss: 0.0010107861598953605
step: 40, loss: 0.11088994890451431
step: 50, loss: 0.0014327903045341372
step: 60, loss: 0.02135009877383709
step: 70, loss: 0.004161171615123749
step: 80, loss: 0.001410163240507245
step: 90, loss: 0.011088227853178978
step: 100, loss: 0.009389967657625675
step: 110, loss: 0.0007488572737202048
step: 120, loss: 0.004124436527490616
step: 130, loss: 0.007475285325199366
step: 140, loss: 0.008461706340312958

```

```

step: 150, loss: 0.0006228543934412301
step: 160, loss: 0.12607845664024353
step: 170, loss: 0.0030604840721935034
step: 180, loss: 0.0029731320682913065
step: 190, loss: 0.02307809330523014
step: 200, loss: 0.0012019178830087185
step: 210, loss: 0.007180306129157543
step: 220, loss: 0.008365863002836704
step: 230, loss: 0.05858977511525154
step: 240, loss: 0.004042036831378937
step: 250, loss: 0.07252518087625504
step: 260, loss: 0.0018563789781183004
step: 270, loss: 0.007205397356301546
step: 280, loss: 0.005075590219348669
step: 290, loss: 0.033137135207653046
step: 300, loss: 0.005095894914120436
step: 310, loss: 0.09291652590036392
step: 320, loss: 0.022174958139657974
step: 330, loss: 0.011602725833654404
step: 340, loss: 0.0009759392123669386
step: 350, loss: 0.0036624413914978504
step: 360, loss: 0.005359257105737925
step: 370, loss: 0.03751003369688988
step: 380, loss: 0.03056665137410164
step: 390, loss: 0.0019015532452613115
step: 400, loss: 0.0010952157899737358
step: 410, loss: 0.007587744854390621
step: 420, loss: 0.009405074641108513
step: 430, loss: 0.02358388900756836
step: 440, loss: 0.0017113022040575743
step: 450, loss: 0.2108297049999237
step: 460, loss: 0.015197362750768661
step: 470, loss: 0.030763709917664528
step: 480, loss: 0.0017920583486557007
step: 490, loss: 0.12949055433273315
step: 500, loss: 0.001414356753230095
step: 510, loss: 0.21860335767269135
step: 520, loss: 0.0020564058795571327
step: 530, loss: 0.00880943238735199
/usr/local/lib/python3.7/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=0.9445234708392602, f1=0.9360902255639098,
best_f1=0.9436555199267065
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.010279174894094467

```

step: 10, loss: 0.0024592143017798662  
step: 20, loss: 0.0011889757588505745  
step: 30, loss: 0.0015884889289736748  
step: 40, loss: 0.0021037752740085125  
step: 50, loss: 0.007994696497917175  
step: 60, loss: 0.0007136373315006495  
step: 70, loss: 0.01040969043970108  
step: 80, loss: 0.002080831676721573  
step: 90, loss: 0.0034299970138818026  
step: 100, loss: 0.004238254390656948  
step: 110, loss: 0.0001396243751514703  
step: 120, loss: 0.0037690112367272377  
step: 130, loss: 0.0024839674588292837  
step: 140, loss: 0.04041123762726784  
step: 150, loss: 0.0037307031452655792  
step: 160, loss: 0.0002804319665301591  
step: 170, loss: 0.006065718829631805  
step: 180, loss: 0.009348534978926182  
step: 190, loss: 0.03468746691942215  
step: 200, loss: 0.001135934959165752  
step: 210, loss: 0.05332978069782257  
step: 220, loss: 0.01075077336281538  
step: 230, loss: 0.0016394013073295355  
step: 240, loss: 0.06976805627346039  
step: 250, loss: 0.021069487556815147  
step: 260, loss: 0.01440854649990797  
step: 270, loss: 0.011119052767753601  
step: 280, loss: 0.010085185058414936  
step: 290, loss: 0.01460309885442257  
step: 300, loss: 0.00036515164538286626  
step: 310, loss: 0.0025708256289362907  
step: 320, loss: 0.047289565205574036  
step: 330, loss: 0.0010710892966017127  
step: 340, loss: 0.001059570349752903  
step: 350, loss: 0.0032098705414682627  
step: 360, loss: 0.06967556476593018  
step: 370, loss: 0.06507119536399841  
step: 380, loss: 0.01658480055630207  
step: 390, loss: 0.002820270834490657  
step: 400, loss: 0.03501391410827637  
step: 410, loss: 4.703840386355296e-05  
step: 420, loss: 0.0034527124371379614  
step: 430, loss: 0.02587594836950302  
step: 440, loss: 0.05131955444812775  
step: 450, loss: 0.018963390961289406  
step: 460, loss: 0.007889987900853157  
step: 470, loss: 0.1426657736301422  
step: 480, loss: 0.002230652840808034

```

step: 490, loss: 0.10332522541284561
step: 500, loss: 0.0016081830253824592
step: 510, loss: 0.0015366899315267801
step: 520, loss: 0.00297342031262815
step: 530, loss: 0.010443910956382751
/usr/local/lib/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=0.9513108614232209, f1=0.95, best_f1=0.95
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.000772496045101434
step: 10, loss: 0.0043389322236180305
step: 20, loss: 0.02406550571322441
step: 30, loss: 0.0007350765517912805
step: 40, loss: 0.000801900343503803
step: 50, loss: 0.002097371267154813
step: 60, loss: 0.0008880756213329732
step: 70, loss: 0.0008062673732638359
step: 80, loss: 0.0641806572675705
step: 90, loss: 0.0016698699910193682
step: 100, loss: 0.0018006074242293835
step: 110, loss: 0.013004492968320847
step: 120, loss: 0.008821860887110233
step: 130, loss: 0.00029076123610138893
step: 140, loss: 0.07359228283166885
step: 150, loss: 0.09565090388059616
step: 160, loss: 0.0007748173666186631
step: 170, loss: 0.14982785284519196
step: 180, loss: 0.0017470415914431214
step: 190, loss: 0.01755860447883606
step: 200, loss: 0.013243254274129868
step: 210, loss: 0.06592364609241486
step: 220, loss: 0.00042951543582603335
step: 230, loss: 0.019570855423808098
step: 240, loss: 0.0074874162673950195
step: 250, loss: 0.0009507669601589441
step: 260, loss: 4.835258368984796e-05
step: 270, loss: 0.00043049734085798264
step: 280, loss: 0.009898432530462742
step: 290, loss: 0.0005415361956693232
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 300, loss: 3.1452032999368384e-05

```

```

step: 310, loss: 0.0009855936514213681
step: 320, loss: 0.0023305302020162344
step: 330, loss: 0.012019430287182331
step: 340, loss: 0.06425531953573227
step: 350, loss: 0.0002159438154194504
step: 360, loss: 0.012928225100040436
step: 370, loss: 0.0330202542245388
step: 380, loss: 0.000526130199432373
step: 390, loss: 0.001809529378078878
step: 400, loss: 0.0008437116630375385
step: 410, loss: 0.000482113566249609
step: 420, loss: 0.000675915100146085
step: 430, loss: 0.005288252141326666
step: 440, loss: 0.005705762188881636
step: 450, loss: 0.00039834174094721675
step: 460, loss: 0.00036809747689403594
step: 470, loss: 0.08715634793043137
step: 480, loss: 0.004937636200338602
step: 490, loss: 0.0028464121278375387
step: 500, loss: 0.0047203670255839825
step: 510, loss: 0.005644166376441717
step: 520, loss: 6.775084330001846e-05
step: 530, loss: 0.008831147104501724
/usr/local/lib/python3.7/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=0.9462563160312356, f1=0.9408553230209282, best_f1=0.95
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.004548340570181608
step: 10, loss: 0.0062348004430532455
step: 20, loss: 0.07622257620096207
step: 30, loss: 0.0896143987774849
step: 40, loss: 0.0012209531851112843
step: 50, loss: 0.00011625434854067862
step: 60, loss: 0.00013885054795537144
step: 70, loss: 0.011355586349964142
step: 80, loss: 0.0026114131323993206
step: 90, loss: 0.10451869666576385
step: 100, loss: 0.0021872995421290398
step: 110, loss: 0.009515291079878807
step: 120, loss: 0.007536483462899923
step: 130, loss: 0.0031119631603360176
step: 140, loss: 0.0006420818390324712
step: 150, loss: 0.0009867135668173432
step: 160, loss: 0.015688972547650337
step: 170, loss: 0.010277907364070415

```

```

step: 180, loss: 0.0011454442283138633
step: 190, loss: 0.0009532590629532933
step: 200, loss: 0.0033748962450772524
step: 210, loss: 0.01511395163834095
step: 220, loss: 0.013437763787806034
step: 230, loss: 0.005444212816655636
step: 240, loss: 0.0038297364953905344
step: 250, loss: 0.018725676462054253
step: 260, loss: 0.0030791384633630514
step: 270, loss: 0.0005099251866340637
step: 280, loss: 0.0013514540623873472
step: 290, loss: 0.00013859830505680293
step: 300, loss: 0.006061120890080929
step: 310, loss: 0.03535695746541023
step: 320, loss: 0.0005842375103384256
step: 330, loss: 0.0002999693970195949
step: 340, loss: 0.016353148967027664
step: 350, loss: 0.012911507859826088
step: 360, loss: 0.043553270399570465
step: 370, loss: 0.018377266824245453
step: 380, loss: 2.3956719815032557e-05
step: 390, loss: 0.006325086113065481
step: 400, loss: 0.0006746512372046709
step: 410, loss: 0.0021374376956373453
step: 420, loss: 0.002210944192484021
step: 430, loss: 0.00012068988871760666
step: 440, loss: 1.804508974601049e-05
step: 450, loss: 0.0006520616007037461
step: 460, loss: 4.9015794502338395e-05
step: 470, loss: 0.000318643928039819
step: 480, loss: 5.685196447302587e-05
step: 490, loss: 0.0010527550475671887
step: 500, loss: 0.024060113355517387
step: 510, loss: 0.0013213410275056958
step: 520, loss: 0.002006059978157282
step: 530, loss: 0.0016742937732487917
/usr/local/lib/python3.7/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=0.9468283582089553, f1=0.9455216989843028, best_f1=0.95
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0058599053882062435
step: 10, loss: 0.0005327855469658971
step: 20, loss: 0.00025167909916490316
step: 30, loss: 0.007295406423509121
step: 40, loss: 0.030244935303926468

```



step: 50, loss: 0.0016971547156572342  
step: 60, loss: 0.0021999890450388193  
step: 70, loss: 7.353570981649682e-05  
step: 80, loss: 0.00040820121648721397  
step: 90, loss: 0.00021291121083777398  
step: 100, loss: 0.00840213242918253  
step: 110, loss: 0.0019243784481659532  
step: 120, loss: 0.00030165465432219207  
step: 130, loss: 0.037605687975883484  
step: 140, loss: 0.0001875099987955764  
step: 150, loss: 3.41322083841078e-05  
step: 160, loss: 0.09757696837186813  
step: 170, loss: 0.0004987125284969807  
step: 180, loss: 0.0036615480203181505  
step: 190, loss: 0.00020964100258424878  
step: 200, loss: 0.0005173460813239217  
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 8192.0  
step: 210, loss: 0.005257163662463427  
step: 220, loss: 0.003874304937198758  
step: 230, loss: 0.0008407766581512988  
step: 240, loss: 0.00199851649813354  
step: 250, loss: 0.000438264716649428  
step: 260, loss: 0.003949700854718685  
step: 270, loss: 0.00062279321718961  
step: 280, loss: 0.045356158167123795  
step: 290, loss: 0.0013152190949767828  
step: 300, loss: 0.0003746877482626587  
step: 310, loss: 0.004370194859802723  
step: 320, loss: 0.0031296424567699432  
step: 330, loss: 0.00557226687669754  
step: 340, loss: 0.0003653440799098462  
step: 350, loss: 8.872504986356944e-05  
step: 360, loss: 3.5528406442608684e-05  
step: 370, loss: 0.035682160407304764  
step: 380, loss: 0.005738281179219484  
step: 390, loss: 0.00011243780318181962  
step: 400, loss: 0.0002716686576604843  
step: 410, loss: 0.009931320324540138  
step: 420, loss: 0.00016985024558380246  
step: 430, loss: 0.003015211783349514  
step: 440, loss: 3.809203553828411e-05  
step: 450, loss: 0.000335186836309731  
step: 460, loss: 0.00587307708337903  
step: 470, loss: 0.09667254239320755  
step: 480, loss: 0.00549622206017375  
step: 490, loss: 0.0004741852171719074  
step: 500, loss: 0.0018278028583154082  
step: 510, loss: 0.0003354644577484578

```

step: 520, loss: 0.009797856211662292
step: 530, loss: 0.010814250446856022
/usr/local/lib/python3.7/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=0.9457221711131555, f1=0.939517962710323, best_f1=0.95
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.003175294492393732
step: 10, loss: 0.043146781623363495
step: 20, loss: 0.0018075520638376474
step: 30, loss: 0.0004297386039979756
step: 40, loss: 0.0003414057719055563
step: 50, loss: 7.661127165192738e-05
step: 60, loss: 0.0003225121763534844
step: 70, loss: 0.00019964342936873436
step: 80, loss: 0.00016730561037547886
step: 90, loss: 9.833327931119129e-05
step: 100, loss: 9.032276284415275e-05
step: 110, loss: 0.0022703446447849274
step: 120, loss: 0.008223219774663448
step: 130, loss: 0.0008745890227146447
step: 140, loss: 0.0006019773427397013
step: 150, loss: 0.004774518311023712
step: 160, loss: 0.001484146574512124
step: 170, loss: 0.006059739738702774
step: 180, loss: 2.044402208412066e-05
step: 190, loss: 0.001624459051527083
step: 200, loss: 0.0013023227220401168
step: 210, loss: 0.0024124435149133205
step: 220, loss: 0.02365308813750744
step: 230, loss: 6.967597437324002e-05
step: 240, loss: 0.003758892184123397
step: 250, loss: 8.547628385713324e-05
step: 260, loss: 0.00022178140352480114
step: 270, loss: 0.000288918090518564
step: 280, loss: 0.003853085683658719
step: 290, loss: 0.0005326664540916681
step: 300, loss: 0.0007826388464309275
step: 310, loss: 0.18291479349136353
step: 320, loss: 0.0011756322346627712
step: 330, loss: 2.5525227101752535e-05
step: 340, loss: 0.001550432643853128
step: 350, loss: 0.0006128768436610699
step: 360, loss: 0.0022007657680660486
step: 370, loss: 0.004709990229457617
step: 380, loss: 0.0014528505271300673

```

```
step: 390, loss: 0.002252202946692705
step: 400, loss: 0.0014675408601760864
step: 410, loss: 0.00020364332885947078
step: 420, loss: 0.002036611782386899
step: 430, loss: 0.0010564300464466214
step: 440, loss: 0.04959004744887352
step: 450, loss: 0.00023292495461646467
step: 460, loss: 0.03767430782318115
step: 470, loss: 0.00036945455940440297
step: 480, loss: 0.0002740486233960837
step: 490, loss: 0.027629978954792023
step: 500, loss: 0.00343683990649879
step: 510, loss: 0.0041128406301140785
step: 520, loss: 0.0004601321998052299
step: 530, loss: 0.012972353026270866
/usr/local/lib/python3.7/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=0.9465861588481189, f1=0.9425393883225208, best_f1=0.95
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.005972345359623432
step: 10, loss: 0.02638235129415989
step: 20, loss: 0.027058426290750504
step: 30, loss: 0.06820325553417206
step: 40, loss: 0.0014013029867783189
step: 50, loss: 5.621005402645096e-05
step: 60, loss: 0.003664308227598667
step: 70, loss: 0.00016602958203293383
step: 80, loss: 0.0015956986462697387
step: 90, loss: 0.004346094094216824
step: 100, loss: 0.03731871396303177
step: 110, loss: 0.007837495766580105
step: 120, loss: 0.0001260935969185084
step: 130, loss: 0.0013189062010496855
step: 140, loss: 1.6357420463464223e-05
step: 150, loss: 4.307659764890559e-05
step: 160, loss: 0.0010279933921992779
step: 170, loss: 0.0010151407914236188
step: 180, loss: 0.0004741016891784966
step: 190, loss: 0.00029375177109614015
step: 200, loss: 0.0054284436628222466
step: 210, loss: 0.0002439113159198314
step: 220, loss: 6.923756882315502e-05
step: 230, loss: 0.0006436521653085947
step: 240, loss: 0.0008425275445915759
step: 250, loss: 2.5191795430146158e-05
```

```

step: 260, loss: 2.119550117640756e-05
step: 270, loss: 0.001718074199743569
step: 280, loss: 1.9463732314761728e-05
step: 290, loss: 5.3990013839211315e-05
step: 300, loss: 9.951119864126667e-05
step: 310, loss: 0.00011793010344263166
step: 320, loss: 0.0015667888801544905
step: 330, loss: 0.010615733452141285
step: 340, loss: 0.002060387749224901
step: 350, loss: 0.00013641898112837225
step: 360, loss: 0.00014340992493089288
step: 370, loss: 0.00041257188422605395
step: 380, loss: 0.00013947866682428867
step: 390, loss: 0.0015071934321895242
step: 400, loss: 0.0007506146212108433
step: 410, loss: 0.00024080014554783702
step: 420, loss: 0.030215125530958176
step: 430, loss: 0.00268577691167593
step: 440, loss: 0.07936178892850876
step: 450, loss: 0.09339699149131775
step: 460, loss: 0.003335206536576152
step: 470, loss: 0.0019072560826316476
step: 480, loss: 0.010502557270228863
step: 490, loss: 0.00015257063205353916
step: 500, loss: 7.67539459047839e-05
step: 510, loss: 0.00026743128546513617
step: 520, loss: 0.0020840608049184084
step: 530, loss: 0.0011899442179128528
/usr/local/lib/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=0.9527374824520355, f1=0.9456572224802601,
best_f1=0.9456572224802601
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.002130664186552167
step: 10, loss: 0.0007848880486562848
step: 20, loss: 4.457665636437014e-05
step: 30, loss: 5.281100675347261e-05
step: 40, loss: 0.003620416158810258
step: 50, loss: 0.0007877820753492415
step: 60, loss: 7.903968798927963e-05
step: 70, loss: 0.002000214299187064

```

step: 80, loss: 0.002225875621661544  
step: 90, loss: 0.002908776979893446  
step: 100, loss: 9.38984812819399e-05  
step: 110, loss: 8.879297820385545e-05  
step: 120, loss: 2.8637852665269747e-05  
step: 130, loss: 0.00023128892644308507  
step: 140, loss: 0.00015243800589814782  
step: 150, loss: 0.0005635959678329527  
step: 160, loss: 0.005108403507620096  
step: 170, loss: 0.002676059491932392  
step: 180, loss: 0.009639966301620007  
step: 190, loss: 0.00020087792654521763  
step: 200, loss: 0.0018255740869790316  
step: 210, loss: 2.48506275966065e-05  
step: 220, loss: 0.0003786521847359836  
step: 230, loss: 0.00047906310646794736  
step: 240, loss: 0.00025518645998090506  
step: 250, loss: 7.939319766592234e-05  
step: 260, loss: 0.014623979106545448  
step: 270, loss: 0.0013089502463117242  
step: 280, loss: 0.00012471203808672726  
step: 290, loss: 0.0011271955445408821  
step: 300, loss: 0.00016514110029675066  
step: 310, loss: 0.00015885189350228757  
step: 320, loss: 2.1162655684747733e-05  
step: 330, loss: 9.475507977185771e-05  
step: 340, loss: 8.44743917696178e-05  
step: 350, loss: 0.0009531527175568044  
step: 360, loss: 0.05459558963775635  
step: 370, loss: 0.009688341058790684  
step: 380, loss: 8.056867955019698e-05  
step: 390, loss: 0.00229735323227942  
step: 400, loss: 0.00011439144145697355  
step: 410, loss: 0.0454443022608757  
step: 420, loss: 0.017566030845046043  
step: 430, loss: 0.00018304074183106422  
step: 440, loss: 0.00013004337961319834  
step: 450, loss: 0.0001737626298563555  
step: 460, loss: 0.012643427588045597  
step: 470, loss: 0.008930277079343796  
step: 480, loss: 4.281242217984982e-05  
step: 490, loss: 0.00015091565728653222  
step: 500, loss: 0.012858616188168526  
step: 510, loss: 0.00012264770339243114  
step: 520, loss: 5.822117964271456e-05  
step: 530, loss: 6.013858364894986e-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=0.9482185273159144, f1=0.940570893776322,
best_f1=0.9456572224802601
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0003288751468062401
step: 10, loss: 0.0012420808197930455
step: 20, loss: 2.9650340366060846e-05
step: 30, loss: 0.00018484247266314924
step: 40, loss: 5.3659689001506194e-05
step: 50, loss: 0.0011508205207064748
step: 60, loss: 0.000313875381834805
step: 70, loss: 4.5187811338109896e-05
step: 80, loss: 0.0004439519834704697
step: 90, loss: 2.9098122467985377e-05
step: 100, loss: 0.00045101181603968143
step: 110, loss: 0.0004503658274188638
step: 120, loss: 1.244587929249974e-05
step: 130, loss: 3.596628812374547e-05
step: 140, loss: 0.004680579528212547
step: 150, loss: 0.036050863564014435
step: 160, loss: 7.938810449559242e-05
step: 170, loss: 4.4942586100660264e-05
step: 180, loss: 0.001097927917726338
step: 190, loss: 6.496047717519104e-05
step: 200, loss: 0.14441686868667603
step: 210, loss: 0.0009187874966301024
step: 220, loss: 0.0002108602930093184
step: 230, loss: 6.583008507732302e-05
step: 240, loss: 0.022205207496881485
step: 250, loss: 0.000302773347357288
step: 260, loss: 0.00018084126350004226
step: 270, loss: 0.0008468671003356576
step: 280, loss: 0.0006038869614712894
step: 290, loss: 2.4935779947554693e-05
step: 300, loss: 1.2535383575595915e-05
step: 310, loss: 0.025947585701942444
step: 320, loss: 1.9891887859557755e-05
step: 330, loss: 0.0007280200952664018
step: 340, loss: 0.0002268410025862977
step: 350, loss: 0.00032602716237306595
step: 360, loss: 0.00022379684378392994
step: 370, loss: 0.006703819613903761
step: 380, loss: 0.11919849365949631
step: 390, loss: 0.00032885049586184323
step: 400, loss: 0.017314041033387184
step: 410, loss: 5.935007720836438e-05
```

```
step: 420, loss: 0.00013864960055798292
step: 430, loss: 0.014630105346441269
step: 440, loss: 0.020160747691988945
step: 450, loss: 5.4985743190627545e-05
step: 460, loss: 0.0195735115557909
step: 470, loss: 6.582149944733828e-05
step: 480, loss: 0.002849870827049017
step: 490, loss: 9.245659748557955e-05
step: 500, loss: 0.0002114035887643695
step: 510, loss: 0.004911710508167744
step: 520, loss: 0.00039928779006004333
step: 530, loss: 0.00036606844514608383
/usr/local/lib/python3.7/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=0.9499766245909304, f1=0.9440591770688858,
best_f1=0.9456572224802601
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00013022871280554682
step: 10, loss: 0.00020498427329584956
step: 20, loss: 0.001351278042420745
step: 30, loss: 0.011283784173429012
step: 40, loss: 3.689081859192811e-05
step: 50, loss: 0.02101592719554901
step: 60, loss: 0.00020105403382331133
step: 70, loss: 8.387610432691872e-05
step: 80, loss: 0.0012229534331709146
step: 90, loss: 0.0001582368859089911
step: 100, loss: 2.4991399186546914e-05
step: 110, loss: 0.011053268797695637
step: 120, loss: 9.843255247687921e-05
step: 130, loss: 0.0034037348814308643
step: 140, loss: 2.332864642085042e-05
step: 150, loss: 0.014923185110092163
step: 160, loss: 0.002458878792822361
step: 170, loss: 6.326310540316626e-05
step: 180, loss: 0.002097364515066147
step: 190, loss: 0.003370001446455717
step: 200, loss: 0.001014654990285635
step: 210, loss: 3.70833913621027e-05
step: 220, loss: 0.12194394320249557
step: 230, loss: 0.00041136686922982335
step: 240, loss: 0.004681771155446768
step: 250, loss: 0.00012061349116265774
step: 260, loss: 7.167577859945595e-05
step: 270, loss: 1.173443433799548e-05
```

```

step: 280, loss: 1.3317597222339828e-05
step: 290, loss: 0.00018675060709938407
step: 300, loss: 0.00022080967028159648
step: 310, loss: 0.030027145519852638
step: 320, loss: 9.15553537197411e-05
step: 330, loss: 5.950398917775601e-05
step: 340, loss: 0.00013411005784291774
step: 350, loss: 0.008476442657411098
step: 360, loss: 0.00041608759784139693
step: 370, loss: 0.005606547929346561
step: 380, loss: 0.002771536586806178
step: 390, loss: 0.00010235852096229792
step: 400, loss: 0.006005843635648489
step: 410, loss: 0.00015219573106151074
step: 420, loss: 0.011139729991555214
step: 430, loss: 1.340321159659652e-05
step: 440, loss: 0.00022849300876259804
step: 450, loss: 0.0018463494488969445
step: 460, loss: 0.013735978864133358
step: 470, loss: 0.00021382630802690983
step: 480, loss: 0.011131701059639454
step: 490, loss: 0.00011117690883111209
step: 500, loss: 0.000841280969325453
step: 510, loss: 0.00044048551353625953
step: 520, loss: 0.00020855346519965678
step: 530, loss: 0.034896865487098694
/usr/local/lib/python3.7/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=0.950681070925317, f1=0.9438515081206497,
best_f1=0.9456572224802601
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, 171.34it/s]
load_f1 = 0.9533146591970122
real_f1 = 0.9517564402810305
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, 144.35it/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: 0.5086256265640259
step: 10, loss: 0.41612404584884644
step: 20, loss: 0.4577934443950653
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 30, loss: 0.31788212060928345
step: 40, loss: 0.3502870500087738
step: 50, loss: 0.47081008553504944
step: 60, loss: 0.5017970204353333
step: 70, loss: 0.3349350094795227
step: 80, loss: 0.38371917605400085
step: 90, loss: 0.23197206854820251
step: 100, loss: 0.295478880405426
step: 110, loss: 0.28868332505226135
step: 120, loss: 0.40088552236557007
step: 130, loss: 0.26275891065597534
step: 140, loss: 0.3985206186771393
step: 150, loss: 0.14914818108081818
step: 160, loss: 0.48064538836479187
step: 170, loss: 0.19103705883026123
step: 180, loss: 0.3250727355480194
step: 190, loss: 0.4355223476886749
step: 200, loss: 0.2202732264995575
step: 210, loss: 0.21418413519859314
/usr/local/lib/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=0.5007974481658691, f1=0.5333333333333332,
best_f1=0.5333333333333332
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.2833678126335144
step: 10, loss: 0.13413883745670319
step: 20, loss: 0.4240352511405945
step: 30, loss: 0.2806291878223419
step: 40, loss: 0.34423211216926575
step: 50, loss: 0.2329908013343811
step: 60, loss: 0.3889244496822357
step: 70, loss: 0.16056710481643677
step: 80, loss: 0.20059970021247864
step: 90, loss: 0.15279065072536469
step: 100, loss: 0.4524977505207062
step: 110, loss: 0.3548129200935364
step: 120, loss: 0.12562336027622223

```

```

step: 130, loss: 0.2517614960670471
step: 140, loss: 0.15000911056995392
step: 150, loss: 0.32492268085479736
step: 160, loss: 0.09802205860614777
step: 170, loss: 0.22028961777687073
step: 180, loss: 0.1638701856136322
step: 190, loss: 0.2911703884601593
step: 200, loss: 0.11103343218564987
step: 210, loss: 0.1968298852443695
/usr/local/lib/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=0.592734225621415, f1=0.6188679245283019,
best_f1=0.6188679245283019
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.10362720489501953
step: 10, loss: 0.09486681967973709
step: 20, loss: 0.24764369428157806
step: 30, loss: 0.05118637531995773
step: 40, loss: 0.24419090151786804
step: 50, loss: 0.20432782173156738
step: 60, loss: 0.20156818628311157
step: 70, loss: 0.08160605281591415
step: 80, loss: 0.17485590279102325
step: 90, loss: 0.07744058966636658
step: 100, loss: 0.27071821689605713
step: 110, loss: 0.06637992709875107
step: 120, loss: 0.14198853075504303
step: 130, loss: 0.15163016319274902
step: 140, loss: 0.08196346461772919
step: 150, loss: 0.1833205223083496
step: 160, loss: 0.14225947856903076
step: 170, loss: 0.30042317509651184
step: 180, loss: 0.07738704234361649
step: 190, loss: 0.023746401071548462
step: 200, loss: 0.16656482219696045
step: 210, loss: 0.15955860912799835
/usr/local/lib/python3.7/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=0.5826235093696762, f1=0.6388888888888888,
best_f1=0.6188679245283019

```

```

/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.1673239916563034
step: 10, loss: 0.13271470367908478
step: 20, loss: 0.11635757982730865
step: 30, loss: 0.14262953400611877
step: 40, loss: 0.04938328638672829
step: 50, loss: 0.10683782398700714
step: 60, loss: 0.2054523527622223
step: 70, loss: 0.14159086346626282
step: 80, loss: 0.09451833367347717
step: 90, loss: 0.07069920748472214
step: 100, loss: 0.26775339245796204
step: 110, loss: 0.558018684387207
step: 120, loss: 0.19120392203330994
step: 130, loss: 0.3959060311317444
step: 140, loss: 0.3310644030570984
step: 150, loss: 0.13553908467292786
step: 160, loss: 0.13064563274383545
step: 170, loss: 0.1727769523859024
step: 180, loss: 0.07855533063411713
step: 190, loss: 0.166444793343544
step: 200, loss: 0.06627125293016434
step: 210, loss: 0.4382694959640503
/usr/local/lib/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=0.6232741617357003, f1=0.6571428571428571,
best_f1=0.6571428571428571
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.1620785892009735
step: 10, loss: 0.092039093375206
step: 20, loss: 0.13871155679225922
step: 30, loss: 0.03257761523127556
step: 40, loss: 0.11371558904647827
step: 50, loss: 0.12984120845794678
step: 60, loss: 0.08277779817581177
step: 70, loss: 0.22598829865455627
step: 80, loss: 0.07315058261156082
step: 90, loss: 0.13462108373641968
step: 100, loss: 0.005983374547213316

```

```

step: 110, loss: 0.1624094396829605
step: 120, loss: 0.1283838152885437
step: 130, loss: 0.10824644565582275
step: 140, loss: 0.20408964157104492
step: 150, loss: 0.07820635288953781
step: 160, loss: 0.05931814759969711
step: 170, loss: 0.04567905142903328
step: 180, loss: 0.18310587108135223
step: 190, loss: 0.24858027696609497
step: 200, loss: 0.24896202981472015
step: 210, loss: 0.07983982563018799
/usr/local/lib/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=0.6345811051693404, f1=0.6815642458100558,
best_f1=0.6815642458100558
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.025912277400493622
step: 10, loss: 0.06983970105648041
step: 20, loss: 0.023040207102894783
step: 30, loss: 0.15669141709804535
step: 40, loss: 0.05259590595960617
step: 50, loss: 0.13275399804115295
step: 60, loss: 0.09776223450899124
step: 70, loss: 0.11145090311765671
step: 80, loss: 0.03064206801354885
step: 90, loss: 0.03117545135319233
step: 100, loss: 0.0526144914329052
step: 110, loss: 0.15812356770038605
step: 120, loss: 0.009191041812300682
step: 130, loss: 0.02315087430179119
step: 140, loss: 0.040473662316799164
step: 150, loss: 0.05249851197004318
step: 160, loss: 0.021952936425805092
step: 170, loss: 0.03985228389501572
step: 180, loss: 0.04401343688368797
step: 190, loss: 0.0768163874745369
step: 200, loss: 0.10867326706647873
step: 210, loss: 0.09738139808177948
/usr/local/lib/python3.7/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=0.6335174953959485, f1=0.6979362101313321,
best_f1=0.6815642458100558
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.09554770588874817
step: 10, loss: 0.012215590104460716
step: 20, loss: 0.033800676465034485
step: 30, loss: 0.10156688094139099
step: 40, loss: 0.047386862337589264
step: 50, loss: 0.0759609118103981
step: 60, loss: 0.0591922253370285
step: 70, loss: 0.09162511676549911
step: 80, loss: 0.07628366351127625
step: 90, loss: 0.037615928798913956
step: 100, loss: 0.09471079707145691
step: 110, loss: 0.35481172800064087
step: 120, loss: 0.0949360802769661
step: 130, loss: 0.07895660400390625
step: 140, loss: 0.025700535625219345
step: 150, loss: 0.0634094625711441
step: 160, loss: 0.24449579417705536
step: 170, loss: 0.08612587302923203
step: 180, loss: 0.01004744041711092
step: 190, loss: 0.07219813019037247
step: 200, loss: 0.0110324053093791
step: 210, loss: 0.2258138656616211
/usr/local/lib/python3.7/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=0.6295585412667946, f1=0.6826347305389221,
best_f1=0.6815642458100558
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.10889778286218643
step: 10, loss: 0.011621021665632725
step: 20, loss: 0.020405253395438194
step: 30, loss: 0.24676208198070526
step: 40, loss: 0.10603626072406769
step: 50, loss: 0.006709471810609102
step: 60, loss: 0.1749172955751419
step: 70, loss: 0.1559273898601532
step: 80, loss: 0.1289476454257965
step: 90, loss: 0.17774082720279694
step: 100, loss: 0.17350465059280396
step: 110, loss: 0.024716539308428764
step: 120, loss: 0.053465597331523895

```

```

step: 130, loss: 0.029750101268291473
step: 140, loss: 0.03949691355228424
step: 150, loss: 0.08861776441335678
step: 160, loss: 0.14479117095470428
step: 170, loss: 0.14755110442638397
step: 180, loss: 0.04679175093770027
step: 190, loss: 0.08343423157930374
step: 200, loss: 0.06349232792854309
step: 210, loss: 0.08800263702869415
/usr/local/lib/python3.7/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=0.6297709923664123, f1=0.6806083650190115,
best_f1=0.6815642458100558
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.028116805478930473
step: 10, loss: 0.08546948432922363
step: 20, loss: 0.03428782522678375
step: 30, loss: 0.07241374254226685
step: 40, loss: 0.008911800570786
step: 50, loss: 0.0729726180434227
step: 60, loss: 0.014441462233662605
step: 70, loss: 0.07282520085573196
step: 80, loss: 0.034875158220529556
step: 90, loss: 0.027835527434945107
step: 100, loss: 0.010870789177715778
step: 110, loss: 0.028193030506372452
step: 120, loss: 0.035957515239715576
step: 130, loss: 0.08026159554719925
step: 140, loss: 0.058614734560251236
step: 150, loss: 0.03103623166680336
step: 160, loss: 0.0801723375916481
step: 170, loss: 0.029368281364440918
step: 180, loss: 0.03246400132775307
step: 190, loss: 0.020919786766171455
step: 200, loss: 0.06380905210971832
step: 210, loss: 0.0765429213643074
/usr/local/lib/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=0.6451612903225807, f1=0.6777546777546777,
best_f1=0.6777546777546777

```



```

/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.030230725184082985
step: 10, loss: 0.024766797199845314
step: 20, loss: 0.0038660261780023575
step: 30, loss: 0.0022501302883028984
step: 40, loss: 0.02080022543668747
step: 50, loss: 0.05003904923796654
step: 60, loss: 0.09997828304767609
step: 70, loss: 0.02751813642680645
step: 80, loss: 0.025449078530073166
step: 90, loss: 0.09174223989248276
step: 100, loss: 0.015344563871622086
step: 110, loss: 0.014450764283537865
step: 120, loss: 0.07420209795236588
step: 130, loss: 0.05687015876173973
step: 140, loss: 0.00574226351454854
step: 150, loss: 0.017070546746253967
step: 160, loss: 0.034510258585214615
step: 170, loss: 0.023218927904963493
step: 180, loss: 0.012997672893106937
step: 190, loss: 0.009633286856114864
step: 200, loss: 0.039117030799388885
step: 210, loss: 0.10332268476486206
/usr/local/lib/python3.7/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=0.6294964028776978, f1=0.6937269372693727,
best_f1=0.6777546777546777
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.049014050513505936
step: 10, loss: 0.051695529371500015
step: 20, loss: 0.04718824476003647
step: 30, loss: 0.045065999031066895
step: 40, loss: 0.01085655577480793
step: 50, loss: 0.03385894000530243
step: 60, loss: 0.023677000775933266
step: 70, loss: 0.029623888432979584
step: 80, loss: 0.010548214428126812
step: 90, loss: 0.15955793857574463
step: 100, loss: 0.16685663163661957
step: 110, loss: 0.08332715928554535
step: 120, loss: 0.06588653475046158
step: 130, loss: 0.02110329456627369
step: 140, loss: 0.26502475142478943

```

```

step: 150, loss: 0.013838821090757847
step: 160, loss: 0.015101809054613113
step: 170, loss: 0.011301213875412941
step: 180, loss: 0.00975011009722948
step: 190, loss: 0.031968098133802414
step: 200, loss: 0.013224445283412933
step: 210, loss: 0.08711009472608566
/usr/local/lib/python3.7/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=0.6397058823529412, f1=0.6781609195402298,
best_f1=0.6777546777546777
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.005535597912967205
step: 10, loss: 0.013644412159919739
step: 20, loss: 0.3573288917541504
step: 30, loss: 0.016289444640278816
step: 40, loss: 0.008597902953624725
step: 50, loss: 0.0031407459173351526
step: 60, loss: 0.006734529975801706
step: 70, loss: 0.002622544765472412
step: 80, loss: 0.014879901893436909
step: 90, loss: 0.01530485786497593
step: 100, loss: 0.0012714904733002186
step: 110, loss: 0.018779275938868523
step: 120, loss: 0.0013781852321699262
step: 130, loss: 0.004616851452738047
step: 140, loss: 0.06133987009525299
step: 150, loss: 0.0026671062223613262
step: 160, loss: 0.018505502492189407
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 170, loss: 0.007756715174764395
step: 180, loss: 0.05522908270359039
step: 190, loss: 0.015857690945267677
step: 200, loss: 0.004970876034349203
step: 210, loss: 0.006349790841341019
/usr/local/lib/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=0.646643109540636, f1=0.6810035842293907,
best_f1=0.6810035842293907
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:

```

```

An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.004346881993114948
step: 10, loss: 0.0555773600935936
step: 20, loss: 0.0036798056680709124
step: 30, loss: 0.0006894786492921412
step: 40, loss: 0.05571551248431206
step: 50, loss: 0.08007440716028214
step: 60, loss: 0.011134816333651543
step: 70, loss: 0.0034691286273300648
step: 80, loss: 0.027584929019212723
step: 90, loss: 0.09131988137960434
step: 100, loss: 0.08931994438171387
step: 110, loss: 0.06100132688879967
step: 120, loss: 0.005594199988991022
step: 130, loss: 0.0065441434271633625
step: 140, loss: 0.013626376166939735
step: 150, loss: 0.00111361313611269
step: 160, loss: 0.026249175891280174
step: 170, loss: 0.019527478143572807
step: 180, loss: 0.01952027529478073
step: 190, loss: 0.023280788213014603
step: 200, loss: 0.041204459965229034
step: 210, loss: 0.003424506401643157
/usr/local/lib/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=0.655367231638418, f1=0.6879699248120301,
best_f1=0.6879699248120301
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0066306209191679955
step: 10, loss: 0.032840121537446976
step: 20, loss: 0.004523843992501497
step: 30, loss: 0.01375288050621748
step: 40, loss: 0.0031594315078109503
step: 50, loss: 0.0038564251735806465
step: 60, loss: 0.023565182462334633
step: 70, loss: 0.0021316497586667538
step: 80, loss: 0.10048118233680725
step: 90, loss: 0.0774931088089943
step: 100, loss: 0.00571187399327755
step: 110, loss: 0.0026479989755898714

```

```

step: 120, loss: 0.005517073441296816
step: 130, loss: 0.011539044789969921
step: 140, loss: 0.014050778932869434
step: 150, loss: 0.27494707703590393
step: 160, loss: 0.001387935713864863
step: 170, loss: 0.037699829787015915
step: 180, loss: 0.007669058162719011
step: 190, loss: 0.03411960229277611
step: 200, loss: 0.07352373749017715
step: 210, loss: 0.0125635527074337
/usr/local/lib/python3.7/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=0.6431226765799257, f1=0.6966292134831461,
best_f1=0.6879699248120301
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.010348054580390453
step: 10, loss: 0.024934416636824608
step: 20, loss: 0.022256184369325638
step: 30, loss: 0.00603345176205039
step: 40, loss: 0.0028380309231579304
step: 50, loss: 0.004381721839308739
step: 60, loss: 0.0015780660323798656
step: 70, loss: 0.04114622250199318
step: 80, loss: 0.09721468389034271
step: 90, loss: 0.005019229371100664
step: 100, loss: 0.012900205329060555
step: 110, loss: 0.03318193554878235
step: 120, loss: 0.03394200652837753
step: 130, loss: 0.005671001970767975
step: 140, loss: 0.13387693464756012
step: 150, loss: 0.0058832657523453236
step: 160, loss: 0.008789795450866222
step: 170, loss: 0.0031691037584096193
step: 180, loss: 0.01616487093269825
step: 190, loss: 0.0015557543374598026
step: 200, loss: 0.0011582609731703997
step: 210, loss: 0.05251605808734894
/usr/local/lib/python3.7/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=0.6450381679389313, f1=0.6835937499999999,
best_f1=0.6879699248120301
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.")
2293it [00:09, 241.85it/s]
load_f1 = 0.6506469500924215
real_f1 = 0.6407407407407407
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, 143.51it/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: 0.46696722507476807
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 10, loss: 0.5223172903060913
step: 20, loss: 0.24930629134178162
step: 30, loss: 0.3845055401325226
step: 40, loss: 0.23351410031318665
step: 50, loss: 0.31782785058021545
step: 60, loss: 0.4599210023880005
step: 70, loss: 0.4105948805809021
step: 80, loss: 0.15496882796287537
step: 90, loss: 0.3158591091632843
step: 100, loss: 0.4332602024078369
step: 110, loss: 0.25697416067123413
step: 120, loss: 0.36120298504829407
step: 130, loss: 0.3413851261138916
step: 140, loss: 0.19719186425209045
step: 150, loss: 0.30288997292518616
step: 160, loss: 0.2397376298904419
step: 170, loss: 0.3707473874092102
step: 180, loss: 0.16259802877902985
step: 190, loss: 0.16540496051311493
/usr/local/lib/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: 0.3863462507724762

```

```

step: 10, loss: 0.3113236427307129
step: 20, loss: 0.7209707498550415
step: 30, loss: 0.20440295338630676
step: 40, loss: 0.3835618495941162
step: 50, loss: 0.31374743580818176
step: 60, loss: 0.228581503033638
step: 70, loss: 0.23896078765392303
step: 80, loss: 0.04016793891787529
step: 90, loss: 0.13847675919532776
step: 100, loss: 0.16638565063476562
step: 110, loss: 0.11875394731760025
step: 120, loss: 0.180443674325943
step: 130, loss: 0.19483977556228638
step: 140, loss: 0.27253416180610657
step: 150, loss: 0.17995043098926544
step: 160, loss: 0.24568909406661987
step: 170, loss: 0.02513565495610237
step: 180, loss: 0.060027290135622025
step: 190, loss: 0.1934477835893631
/usr/local/lib/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=0.7704918032786885, f1=0.7837837837837839,
best_f1=0.7837837837837839
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.18263676762580872
step: 10, loss: 0.14082524180412292
step: 20, loss: 0.19196194410324097
step: 30, loss: 0.0111743975430727
step: 40, loss: 0.06504873186349869
step: 50, loss: 0.0681341215968132
step: 60, loss: 0.07182851433753967
step: 70, loss: 0.09554020315408707
step: 80, loss: 0.24174344539642334
step: 90, loss: 0.11490752547979355
step: 100, loss: 0.10932406783103943
step: 110, loss: 0.31550511717796326
step: 120, loss: 0.16521939635276794
step: 130, loss: 0.06895964592695236
step: 140, loss: 0.08161978423595428
step: 150, loss: 0.2245696783065796
step: 160, loss: 0.04435766488313675

```



```

step: 170, loss: 0.24203850328922272
step: 180, loss: 0.16754846274852753
step: 190, loss: 0.06236737221479416
/usr/local/lib/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=0.7989130434782609, f1=0.8116710875331565,
best_f1=0.8116710875331565
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.008709488436579704
step: 10, loss: 0.03711690381169319
step: 20, loss: 0.03105505369603634
step: 30, loss: 0.007215695921331644
step: 40, loss: 0.15381385385990143
step: 50, loss: 0.0792425349354744
step: 60, loss: 0.0408676341176033
step: 70, loss: 0.08146657794713974
step: 80, loss: 0.03099505417048931
step: 90, loss: 0.06680808961391449
step: 100, loss: 0.16874663531780243
step: 110, loss: 0.11272131651639938
step: 120, loss: 0.07462608814239502
step: 130, loss: 0.11983326077461243
step: 140, loss: 0.16394075751304626
step: 150, loss: 0.002871517790481448
step: 160, loss: 0.04602091759443283
step: 170, loss: 0.07033819705247879
step: 180, loss: 0.023726776242256165
step: 190, loss: 0.006071147508919239
/usr/local/lib/python3.7/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=0.7696476964769648, f1=0.7978436657681941,
best_f1=0.8116710875331565
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.27108606696128845
step: 10, loss: 0.10122024267911911
step: 20, loss: 0.006466829217970371
step: 30, loss: 0.04024134576320648
step: 40, loss: 0.00347076915204525

```

```

step: 50, loss: 0.04779309779405594
step: 60, loss: 0.17122752964496613
step: 70, loss: 0.12552428245544434
step: 80, loss: 0.06797832250595093
step: 90, loss: 0.16178052127361298
step: 100, loss: 0.05709342285990715
step: 110, loss: 0.050933029502630234
step: 120, loss: 0.02942153438925743
step: 130, loss: 0.2437765896320343
step: 140, loss: 0.015143364667892456
step: 150, loss: 0.1303461641073227
step: 160, loss: 0.08576948940753937
step: 170, loss: 0.014477788470685482
step: 180, loss: 0.026930196210741997
step: 190, loss: 0.013188430108129978
/usr/local/lib/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=0.802030456852792, f1=0.785, best_f1=0.785
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.008347606286406517
step: 10, loss: 0.01576179824769497
step: 20, loss: 0.19869111478328705
step: 30, loss: 0.053692009299993515
step: 40, loss: 0.017388703301548958
step: 50, loss: 0.02389562502503395
step: 60, loss: 0.022359240800142288
step: 70, loss: 0.09221960604190826
step: 80, loss: 0.08077005296945572
step: 90, loss: 0.1616876870393753
step: 100, loss: 0.07957003265619278
step: 110, loss: 0.004353655967861414
step: 120, loss: 0.021693546324968338
step: 130, loss: 0.1288551241159439
step: 140, loss: 0.010649819858372211
step: 150, loss: 0.007614356465637684
step: 160, loss: 0.08007282763719559
step: 170, loss: 0.08015541732311249
step: 180, loss: 0.0174605343490839
step: 190, loss: 0.008439463563263416
/usr/local/lib/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=0.8153846153846155, f1=0.8040712468193384,
best_f1=0.8040712468193384
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.006419401615858078
step: 10, loss: 0.044581279158592224
step: 20, loss: 0.0019963327795267105
step: 30, loss: 0.0896696224808693
step: 40, loss: 0.008117806166410446
step: 50, loss: 0.0007579482044093311
step: 60, loss: 0.004185533616691828
step: 70, loss: 0.001863638637587428
step: 80, loss: 0.0007001321064308286
step: 90, loss: 0.016353171318769455
step: 100, loss: 0.26052626967430115
step: 110, loss: 0.04222693294286728
step: 120, loss: 0.01647946983575821
step: 130, loss: 0.010208300314843655
step: 140, loss: 0.01830548420548439
step: 150, loss: 0.050125084817409515
step: 160, loss: 0.007082927506417036
step: 170, loss: 0.1996358186006546
step: 180, loss: 0.022503923624753952
step: 190, loss: 0.04495585337281227
/usr/local/lib/python3.7/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=0.8055555555555556, f1=0.8247978436657681,
best_f1=0.8040712468193384
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.02118837833404541
step: 10, loss: 0.011107198894023895
step: 20, loss: 0.0024966890923678875
step: 30, loss: 0.003667006967589259
step: 40, loss: 0.0430489182472229
step: 50, loss: 0.1652243733406067
step: 60, loss: 0.1621386706829071
step: 70, loss: 0.06441491842269897
step: 80, loss: 0.05932936072349548
step: 90, loss: 0.06195119395852089

```

```

step: 100, loss: 0.0037224143743515015
step: 110, loss: 0.03113877959549427
step: 120, loss: 0.12038569152355194
step: 130, loss: 0.0071009485982358456
step: 140, loss: 0.022149566560983658
step: 150, loss: 0.06692028790712357
step: 160, loss: 0.001554093207232654
step: 170, loss: 0.001410702825523913
step: 180, loss: 0.1767594814300537
step: 190, loss: 0.005693000741302967
/usr/local/lib/python3.7/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=0.8030303030303031, f1=0.8174807197943443,
best_f1=0.8040712468193384
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.020675053820014
step: 10, loss: 0.004207057878375053
step: 20, loss: 0.004140826407819986
step: 30, loss: 0.006679190322756767
step: 40, loss: 0.033071115612983704
step: 50, loss: 0.016891557723283768
step: 60, loss: 0.0020112693309783936
step: 70, loss: 0.0011672850232571363
step: 80, loss: 0.01631266623735428
step: 90, loss: 0.0021055482793599367
step: 100, loss: 0.001630711485631764
step: 110, loss: 0.00036810286110267043
step: 120, loss: 0.24578548967838287
step: 130, loss: 0.025451447814702988
step: 140, loss: 0.04346829280257225
step: 150, loss: 0.001430911011993885
step: 160, loss: 0.17255684733390808
step: 170, loss: 0.003294250462204218
step: 180, loss: 0.009233874268829823
step: 190, loss: 0.00045531996875070035
/usr/local/lib/python3.7/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=0.8126649076517151, f1=0.8010752688172043,
best_f1=0.8040712468193384
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00915826577693224
step: 10, loss: 0.001334281638264656

```

```

step: 20, loss: 0.0031534337904304266
step: 30, loss: 0.0060263448394834995
step: 40, loss: 0.0034292778000235558
step: 50, loss: 0.002332995180040598
step: 60, loss: 0.08085235208272934
step: 70, loss: 0.0022318908013403416
step: 80, loss: 0.032988496124744415
step: 90, loss: 0.0029022949747741222
step: 100, loss: 0.0011080458061769605
step: 110, loss: 0.005872105713933706
step: 120, loss: 0.014985660091042519
step: 130, loss: 0.0005453511839732528
step: 140, loss: 0.081398144364357
step: 150, loss: 0.006149685941636562
step: 160, loss: 0.0033001736737787724
step: 170, loss: 0.0007656418601982296
step: 180, loss: 0.0025102966465055943
step: 190, loss: 0.005271937698125839
/usr/local/lib/python3.7/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=0.805, f1=0.7949367088607595, best_f1=0.8040712468193384
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.008985031396150589
step: 10, loss: 0.0009015848627313972
step: 20, loss: 0.0004721497534774244
step: 30, loss: 0.07941319048404694
step: 40, loss: 0.0011005300329998136
step: 50, loss: 0.0006767690065316856
step: 60, loss: 0.0018021779833361506
step: 70, loss: 0.0015381646808236837
step: 80, loss: 0.002774093532934785
step: 90, loss: 0.001959946472197771
step: 100, loss: 0.006926009897142649
step: 110, loss: 0.11670556664466858
step: 120, loss: 0.0014998794067651033
step: 130, loss: 0.02736133337020874
step: 140, loss: 0.0017731406260281801
step: 150, loss: 0.0018685257527977228
step: 160, loss: 0.0008127427427098155
step: 170, loss: 0.019686002284288406
step: 180, loss: 0.001339695998467505
step: 190, loss: 0.0011456158244982362
/usr/local/lib/python3.7/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=0.8131868131868132, f1=0.8142076502732241,
best_f1=0.8040712468193384
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.001189405913464725
step: 10, loss: 0.0213627889752388
step: 20, loss: 0.002918063197284937
step: 30, loss: 0.05156689137220383
step: 40, loss: 0.0017168591730296612
step: 50, loss: 0.007530090399086475
step: 60, loss: 0.006122680846601725
step: 70, loss: 0.010824852623045444
step: 80, loss: 0.04146299883723259
step: 90, loss: 0.0015591385308653116
step: 100, loss: 0.008015461266040802
step: 110, loss: 0.017383961006999016
step: 120, loss: 0.0015899017453193665
step: 130, loss: 0.008983165957033634
step: 140, loss: 0.0015756525099277496
step: 150, loss: 0.005281913094222546
step: 160, loss: 0.0017682545585557818
step: 170, loss: 0.0029624488670378923
step: 180, loss: 0.0006672506569884717
step: 190, loss: 0.0005567196058109403
/usr/local/lib/python3.7/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=0.8125, f1=0.8083989501312335, best_f1=0.8040712468193384
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0005179121508263052
step: 10, loss: 0.002682323567569256
step: 20, loss: 0.0010137977078557014
step: 30, loss: 0.008747152984142303
step: 40, loss: 0.0030052028596401215
step: 50, loss: 0.06583605706691742
step: 60, loss: 0.03290007263422012
step: 70, loss: 0.0023626491893082857
step: 80, loss: 0.000919554615393281
step: 90, loss: 0.00039817410288378596
step: 100, loss: 0.00043479312444105744
step: 110, loss: 0.0018136478029191494
step: 120, loss: 0.001190091366879642
step: 130, loss: 0.0006838831468485296
step: 140, loss: 0.0005286234663799405
step: 150, loss: 0.0004565089475363493

```

```

step: 160, loss: 0.0004426779050845653
step: 170, loss: 0.07488042861223221
step: 180, loss: 0.0005265618674457073
step: 190, loss: 0.12432362884283066
/usr/local/lib/python3.7/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=0.8150134048257373, f1=0.8096514745308311,
best_f1=0.8040712468193384
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0010790068190544844
step: 10, loss: 0.04676591604948044
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 20, loss: 0.0016569752478972077
step: 30, loss: 0.004561539739370346
step: 40, loss: 0.006627959199249744
step: 50, loss: 0.0008230841485783458
step: 60, loss: 0.0005560893332585692
step: 70, loss: 0.0025147427804768085
step: 80, loss: 0.00210075662471354
step: 90, loss: 0.00039628471131436527
step: 100, loss: 0.0003150800766889006
step: 110, loss: 0.0001331991225015372
step: 120, loss: 0.0008365956018678844
step: 130, loss: 0.005047863349318504
step: 140, loss: 0.0007824301719665527
step: 150, loss: 0.013677949085831642
step: 160, loss: 0.008287813514471054
step: 170, loss: 0.004704525228589773
step: 180, loss: 0.004881961736828089
step: 190, loss: 0.00247213919647038
/usr/local/lib/python3.7/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=0.8108108108108106, f1=0.8219178082191781,
best_f1=0.8040712468193384
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.003964603878557682
step: 10, loss: 0.0011767070973291993
step: 20, loss: 0.002154207555577159
step: 30, loss: 0.0007602862897329032
step: 40, loss: 0.0014341695932671428
step: 50, loss: 0.0006939075537957251
step: 60, loss: 0.001653047394938767

```

```

step: 70, loss: 0.004817640408873558
step: 80, loss: 0.0007841151673346758
step: 90, loss: 0.047339677810668945
step: 100, loss: 0.00046488255611620843
step: 110, loss: 0.0017339319456368685
step: 120, loss: 0.008439003489911556
step: 130, loss: 0.0032167029567062855
step: 140, loss: 0.0013194250641390681
step: 150, loss: 0.0007945906836539507
step: 160, loss: 0.0017711755353957415
step: 170, loss: 0.03402891382575035
step: 180, loss: 0.000878617400303483
step: 190, loss: 0.005759032443165779
/usr/local/lib/python3.7/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=0.8148148148148148, f1=0.8150134048257373,
best_f1=0.8040712468193384
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.")

```



```

2049it [00:12, 158.17it/s]
load_f1 = 0.7855153203342619
real_f1 = 0.784741144414169
733it [00:00, 3465.43it/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, 143.78it/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 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: 0.46702054142951965
```

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

```
step: 10, loss: 0.43212762475013733
```

```
step: 20, loss: 0.31524422764778137
```

```
step: 30, loss: 0.4114842116832733
```

```
step: 40, loss: 0.511871874332428
```

```
step: 50, loss: 0.3734758794307709
```

```
step: 60, loss: 0.5921560525894165
```

```
step: 70, loss: 0.31557798385620117
```

```
step: 80, loss: 0.1997348517179489
```

```
step: 90, loss: 0.22200199961662292
```

```
step: 100, loss: 0.15738524496555328
```

```

step: 110, loss: 0.43344467878341675
step: 120, loss: 0.3240932524204254
step: 130, loss: 0.33221378922462463
step: 140, loss: 0.38768717646598816
step: 150, loss: 0.3094812035560608
step: 160, loss: 0.40023067593574524
step: 170, loss: 0.33208024501800537
/usr/local/lib/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=0.24628820960698694, f1=0.25426944971537,
best_f1=0.25426944971537
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.29884111881256104
step: 10, loss: 0.42173853516578674
step: 20, loss: 0.3173537850379944
step: 30, loss: 0.2943190932273865
step: 40, loss: 0.05133942887187004
step: 50, loss: 0.3960120379924774
step: 60, loss: 0.16536539793014526
step: 70, loss: 0.4387056529521942
step: 80, loss: 0.1861044019460678
step: 90, loss: 0.2298939973115921
step: 100, loss: 0.2857889235019684
step: 110, loss: 0.2252410501241684
step: 120, loss: 0.07917921245098114
step: 130, loss: 0.1416466236114502
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 8192.0
step: 140, loss: 0.35057002305984497
step: 150, loss: 0.08079265058040619
step: 160, loss: 0.22037875652313232
step: 170, loss: 0.09758653491735458
/usr/local/lib/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=0.7714987714987714, f1=0.7632183908045976,
best_f1=0.7632183908045976
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:

```

```

An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.5072116255760193
step: 10, loss: 0.057101164013147354
step: 20, loss: 0.04974523186683655
step: 30, loss: 0.10113365203142166
step: 40, loss: 0.09448252618312836
step: 50, loss: 0.16401304304599762
step: 60, loss: 0.12359840422868729
step: 70, loss: 0.13256308436393738
step: 80, loss: 0.09701098501682281
step: 90, loss: 0.41328099370002747
step: 100, loss: 0.07533479481935501
step: 110, loss: 0.057480618357658386
step: 120, loss: 0.06908250600099564
step: 130, loss: 0.16402913630008698
step: 140, loss: 0.0789942815899849
step: 150, loss: 0.11252420395612717
step: 160, loss: 0.09630647301673889
step: 170, loss: 0.0644557997584343
/usr/local/lib/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=0.8155844155844156, f1=0.827250608272506,
best_f1=0.827250608272506
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.14784865081310272
step: 10, loss: 0.17995993793010712
step: 20, loss: 0.05379597842693329
step: 30, loss: 0.1926402449607849
step: 40, loss: 0.11940898001194
step: 50, loss: 0.0674111545085907
step: 60, loss: 0.13845562934875488
step: 70, loss: 0.0013957304181531072
step: 80, loss: 0.22567501664161682
step: 90, loss: 0.30848047137260437
step: 100, loss: 0.1236807331442833
step: 110, loss: 0.11216096580028534
step: 120, loss: 0.20351049304008484
step: 130, loss: 0.264702171087265
step: 140, loss: 0.14445461332798004
step: 150, loss: 0.15337009727954865

```

```

step: 160, loss: 0.004743530414998531
step: 170, loss: 0.032914336770772934
/usr/local/lib/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=0.8287037037037036, f1=0.8232662192393736,
best_f1=0.8232662192393736
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.11446776986122131
step: 10, loss: 0.03455202281475067
step: 20, loss: 0.05249044671654701
step: 30, loss: 0.02644798718392849
step: 40, loss: 0.07812058925628662
step: 50, loss: 0.051957931369543076
step: 60, loss: 0.17039409279823303
step: 70, loss: 0.007698417641222477
step: 80, loss: 0.00706544378772378
step: 90, loss: 0.025460200384259224
step: 100, loss: 0.010366717353463173
step: 110, loss: 0.04673406854271889
step: 120, loss: 0.1064995527267456
step: 130, loss: 0.011769838631153107
step: 140, loss: 0.1462889015674591
step: 150, loss: 0.04579703137278557
step: 160, loss: 0.006617054343223572
step: 170, loss: 0.006122846622020006
/usr/local/lib/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=0.8599033816425121, f1=0.8689320388349513,
best_f1=0.8689320388349513
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.18298955261707306
step: 10, loss: 0.017616644501686096
step: 20, loss: 0.08549655228853226
step: 30, loss: 0.004092440940439701

```

```

step: 40, loss: 0.009209509007632732
step: 50, loss: 0.07580183446407318
step: 60, loss: 0.024612585082650185
step: 70, loss: 0.02185145765542984
step: 80, loss: 0.02352973259985447
step: 90, loss: 0.16129498183727264
step: 100, loss: 0.0034964478109031916
step: 110, loss: 0.04800403118133545
step: 120, loss: 0.04561310261487961
step: 130, loss: 0.2084832489490509
step: 140, loss: 0.02567823976278305
step: 150, loss: 0.025315308943390846
step: 160, loss: 0.034760359674692154
step: 170, loss: 0.02918240800499916
/usr/local/lib/python3.7/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=0.8529411764705882, f1=0.8815165876777251,
best_f1=0.8689320388349513
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.03321656957268715
step: 10, loss: 0.043316252529621124
step: 20, loss: 0.011849420145154
step: 30, loss: 0.08497006446123123
step: 40, loss: 0.003312258282676339
step: 50, loss: 0.0012185962405055761
step: 60, loss: 0.09586720913648605
step: 70, loss: 0.024712838232517242
step: 80, loss: 0.05731869488954544
step: 90, loss: 0.012879472225904465
step: 100, loss: 0.02097908779978752
step: 110, loss: 0.01208537444472313
step: 120, loss: 0.0270168948918581
step: 130, loss: 0.05763527378439903
step: 140, loss: 0.03725564479827881
step: 150, loss: 0.0822284072637558
step: 160, loss: 0.06665146350860596
step: 170, loss: 0.03240028768777847
/usr/local/lib/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=0.8725490196078431, f1=0.8805620608899297,

```

```

best_f1=0.8805620608899297
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.02026655152440071
step: 10, loss: 0.12615245580673218
step: 20, loss: 0.026357464492321014
step: 30, loss: 0.0008285861113108695
step: 40, loss: 0.04733579605817795
step: 50, loss: 0.010808417573571205
step: 60, loss: 0.006797115784138441
step: 70, loss: 0.005133844446390867
step: 80, loss: 0.00033281449577771127
step: 90, loss: 0.023938676342368126
step: 100, loss: 0.0007415566360577941
step: 110, loss: 0.06897298246622086
step: 120, loss: 0.18442879617214203
step: 130, loss: 0.01657160557806492
step: 140, loss: 0.03934534266591072
step: 150, loss: 0.012430130504071712
step: 160, loss: 0.0013489816337823868
step: 170, loss: 0.0940350592136383
/usr/local/lib/python3.7/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=0.86, f1=0.8909952606635072, best_f1=0.8805620608899297
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.03754720091819763
step: 10, loss: 0.007230158895254135
step: 20, loss: 0.0008941606502048671
step: 30, loss: 0.00308199692517519
step: 40, loss: 0.13937322795391083
step: 50, loss: 0.00502235209569335
step: 60, loss: 0.017397835850715637
step: 70, loss: 0.006290370598435402
step: 80, loss: 0.06116149201989174
step: 90, loss: 0.01778327487409115
step: 100, loss: 0.012199979275465012
step: 110, loss: 0.07646695524454117
step: 120, loss: 0.04166563227772713
step: 130, loss: 0.0021077252458781004
step: 140, loss: 0.01642446592450142
step: 150, loss: 0.10485411435365677
step: 160, loss: 0.00486669922247529
step: 170, loss: 0.026890795677900314
/usr/local/lib/python3.7/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=0.8622448979591837, f1=0.8977556109725686,
best_f1=0.8805620608899297
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0038567918818444014
step: 10, loss: 0.007039605174213648
step: 20, loss: 0.07893287390470505
step: 30, loss: 0.05030197277665138
step: 40, loss: 0.006075473967939615
step: 50, loss: 0.03747643902897835
step: 60, loss: 0.009204705245792866
step: 70, loss: 0.011469081044197083
step: 80, loss: 0.009055336937308311
step: 90, loss: 0.010066596791148186
step: 100, loss: 0.01177978329360485
step: 110, loss: 0.016942018643021584
step: 120, loss: 0.0019268247997388244
step: 130, loss: 0.002131915418431163
step: 140, loss: 0.0011363404337316751
step: 150, loss: 0.0004738192947115749
step: 160, loss: 0.0003957203880418092
step: 170, loss: 0.019883664324879646
/usr/local/lib/python3.7/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=0.8671679197994987, f1=0.8926829268292683,
best_f1=0.8805620608899297
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.006896683014929295
step: 10, loss: 0.011870814487338066
step: 20, loss: 0.010909120552241802
step: 30, loss: 0.008415794931352139
step: 40, loss: 0.00796441175043583
step: 50, loss: 0.008839775808155537
step: 60, loss: 0.057856831699609756
step: 70, loss: 0.0021301216911524534
step: 80, loss: 0.016220588237047195
step: 90, loss: 0.0026962023694068193
step: 100, loss: 0.016339661553502083
step: 110, loss: 0.07667868584394455
step: 120, loss: 0.011627831496298313
step: 130, loss: 0.007740046828985214
step: 140, loss: 0.0025987846311181784

```



```

step: 150, loss: 0.00013201843830756843
step: 160, loss: 0.0029072873294353485
step: 170, loss: 0.06981299072504044
/usr/local/lib/python3.7/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=0.8506329113924049, f1=0.8786407766990291,
best_f1=0.8805620608899297
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.04290195554494858
step: 10, loss: 0.0014528670581057668
step: 20, loss: 0.00729254400357604
step: 30, loss: 0.048522476106882095
step: 40, loss: 0.0020746837835758924
step: 50, loss: 0.004453095141798258
step: 60, loss: 0.019412271678447723
step: 70, loss: 0.00039251145790331066
step: 80, loss: 0.00022203702246770263
step: 90, loss: 0.0023347786627709866
step: 100, loss: 0.0004926255205646157
step: 110, loss: 0.034367818385362625
step: 120, loss: 0.03613396733999252
step: 130, loss: 0.2125980406999588
step: 140, loss: 0.0003191028372384608
step: 150, loss: 0.021869447082281113
step: 160, loss: 0.06568457186222076
step: 170, loss: 0.002245865296572447
/usr/local/lib/python3.7/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=0.8720626631853786, f1=0.8989898989898989,
best_f1=0.8805620608899297
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.008640487678349018
step: 10, loss: 0.004779363051056862
step: 20, loss: 0.004425869323313236
step: 30, loss: 0.04601927474141121
step: 40, loss: 0.054583799093961716
step: 50, loss: 0.018265748396515846
step: 60, loss: 0.0017535793595016003
step: 70, loss: 0.07917200773954391
step: 80, loss: 0.0003798638063017279
step: 90, loss: 0.02081366814672947
step: 100, loss: 0.022054651752114296

```

```

step: 110, loss: 0.0016345684416592121
step: 120, loss: 0.02627178467810154
step: 130, loss: 0.012339258566498756
step: 140, loss: 0.020243804901838303
step: 150, loss: 0.0003415604878682643
step: 160, loss: 0.0004081043298356235
step: 170, loss: 0.0016612531617283821
/usr/local/lib/python3.7/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=0.854922279792746, f1=0.8905472636815919,
best_f1=0.8805620608899297
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.009834669530391693
step: 10, loss: 0.0008931048214435577
step: 20, loss: 0.004807529505342245
step: 30, loss: 0.01282070018351078
step: 40, loss: 0.011171282269060612
step: 50, loss: 0.0017796439351513982
step: 60, loss: 0.006344193127006292
step: 70, loss: 0.006566516123712063
step: 80, loss: 0.01904691569507122
step: 90, loss: 0.0004782379255630076
step: 100, loss: 0.09982948005199432
step: 110, loss: 0.029451783746480942
step: 120, loss: 0.0020483999978750944
step: 130, loss: 0.07858239859342575
step: 140, loss: 0.01542894821614027
step: 150, loss: 0.01669330894947052
step: 160, loss: 0.0010910759447142482
step: 170, loss: 0.0360715314745903
/usr/local/lib/python3.7/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=0.8640776699029126, f1=0.8658823529411764,
best_f1=0.8805620608899297
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00045429336023516953
step: 10, loss: 0.0013478926848620176
step: 20, loss: 0.2600279748439789
step: 30, loss: 0.003836559597402811
step: 40, loss: 0.00044029654236510396
step: 50, loss: 0.0004140314122196287
step: 60, loss: 0.011927993036806583

```

```

step: 70, loss: 0.12843897938728333
step: 80, loss: 0.0005036680959165096
step: 90, loss: 0.002589755691587925
step: 100, loss: 0.031446296721696854
step: 110, loss: 0.0005056431982666254
step: 120, loss: 0.0006990960100665689
step: 130, loss: 0.00037476845318451524
step: 140, loss: 0.002819425193592906
step: 150, loss: 0.001456850441172719
step: 160, loss: 0.00104794732760638
step: 170, loss: 0.0013098722556605935
/usr/local/lib/python3.7/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=0.8634146341463416, f1=0.8741092636579572,
best_f1=0.8805620608899297
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.")
1916it [00:08, 235.54it/s]
load_f1 = 0.4487369985141158

```

```

real_f1 = 0.4253521126760563
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, 144.57it/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/

```

```

Downloading: 100% 481/481 [00:00<00:00, 404kB/s]
Downloading: 100% 899k/899k [00:01<00:00, 795kB/s]

```

Downloading: 100% 456k/456k [00:01<00:00, 403kB/s]  
Downloading: 100% 501M/501M [00:06<00:00, 73.1MB/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.")

step: 0, loss: 0.6029409170150757

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: 0.4525490999221802

step: 20, loss: 0.49699318408966064

step: 30, loss: 0.3406045138835907

step: 40, loss: 0.33596473932266235

step: 50, loss: 0.47067373991012573

step: 60, loss: 0.07124824076890945

step: 70, loss: 0.06522604078054428

step: 80, loss: 0.04914839193224907

step: 90, loss: 0.105853371322155

step: 100, loss: 0.23052705824375153

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

step: 110, loss: 0.07264204323291779

step: 120, loss: 0.027242165058851242

step: 130, loss: 0.17191459238529205

step: 140, loss: 0.0107550248503685

step: 150, loss: 0.1220327615737915

step: 160, loss: 0.028470950201153755

```

step: 170, loss: 0.07560119032859802
step: 180, loss: 0.21842187643051147
step: 190, loss: 0.017669538035988808
step: 200, loss: 0.03317601978778839
step: 210, loss: 0.288647323846817
step: 220, loss: 0.09639342874288559
step: 230, loss: 0.006854568608105183
/usr/local/lib/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=0.9690949227373068, f1=0.9673790776152981,
best_f1=0.9673790776152981
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.009777556173503399
step: 10, loss: 0.10090525448322296
step: 20, loss: 0.05571822449564934
step: 30, loss: 0.014020557515323162
step: 40, loss: 0.029889367520809174
step: 50, loss: 0.0033335413318127394
step: 60, loss: 0.007266242057085037
step: 70, loss: 0.009220811538398266
step: 80, loss: 0.10143912583589554
step: 90, loss: 0.04078512638807297
step: 100, loss: 0.003594406880438328
step: 110, loss: 0.04511532559990883
step: 120, loss: 0.03424602746963501
step: 130, loss: 0.030947711318731308
step: 140, loss: 0.005883266683667898
step: 150, loss: 0.0846845805644989
step: 160, loss: 0.048740457743406296
step: 170, loss: 0.001252524321898818
step: 180, loss: 0.018357649445533752
step: 190, loss: 0.0037145738024264574
step: 200, loss: 0.032419923692941666
step: 210, loss: 0.015783078968524933
step: 220, loss: 0.0056887115351855755
step: 230, loss: 0.0020908862352371216
/usr/local/lib/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=0.9700996677740864, f1=0.9698324022346367,
best_f1=0.9698324022346367
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.02803192101418972
step: 10, loss: 0.055741190910339355
step: 20, loss: 0.034613415598869324
step: 30, loss: 0.017135318368673325
step: 40, loss: 0.08839885145425797
step: 50, loss: 0.013625426217913628
step: 60, loss: 0.00865261536091566
step: 70, loss: 0.005898303352296352
step: 80, loss: 0.1224130243062973
step: 90, loss: 0.01098044402897358
step: 100, loss: 0.005069551058113575
step: 110, loss: 0.0167231522500515
step: 120, loss: 0.00707939313724637
step: 130, loss: 0.009630982764065266
step: 140, loss: 0.0050214440561831
step: 150, loss: 0.08417918533086777
step: 160, loss: 0.0013009726535528898
step: 170, loss: 0.0008833333849906921
step: 180, loss: 0.010821714997291565
step: 190, loss: 0.009182856418192387
step: 200, loss: 0.010008787736296654
step: 210, loss: 0.002235592110082507
step: 220, loss: 0.11520547419786453
step: 230, loss: 0.014505358412861824
/usr/local/lib/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=0.9755555555555556, f1=0.9699666295884317,
best_f1=0.9699666295884317
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.016011981293559074
step: 10, loss: 0.006189489737153053
step: 20, loss: 0.010099017061293125
step: 30, loss: 0.0010146981803700328
step: 40, loss: 0.04963235929608345

```

```

step: 50, loss: 0.04744022712111473
step: 60, loss: 0.0011708062374964356
step: 70, loss: 0.020251775160431862
step: 80, loss: 0.057670775800943375
step: 90, loss: 0.10048602521419525
step: 100, loss: 0.02096804976463318
step: 110, loss: 0.001782967709004879
step: 120, loss: 0.027184145525097847
step: 130, loss: 0.014467140659689903
step: 140, loss: 0.007745147682726383
step: 150, loss: 0.001658713212236762
step: 160, loss: 0.008007261902093887
step: 170, loss: 0.000576779420953244
step: 180, loss: 0.1394500881433487
step: 190, loss: 0.004811698570847511
step: 200, loss: 0.05687907710671425
step: 210, loss: 0.07579611241817474
step: 220, loss: 0.017964735627174377
step: 230, loss: 0.004844917915761471
/usr/local/lib/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=0.9767441860465117, f1=0.9723756906077348,
best_f1=0.9723756906077348
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0011431712191551924
step: 10, loss: 0.005538526922464371
step: 20, loss: 0.02080768346786499
step: 30, loss: 0.002440342679619789
step: 40, loss: 0.0006971890688873827
step: 50, loss: 0.001393481739796698
step: 60, loss: 0.0004062712541781366
step: 70, loss: 0.004742393270134926
step: 80, loss: 0.028818916529417038
step: 90, loss: 0.1275143027305603
step: 100, loss: 0.0008402207749895751
step: 110, loss: 0.04008299484848976
step: 120, loss: 0.001130227348767221
step: 130, loss: 0.0005125732277520001
step: 140, loss: 0.0033881859853863716
step: 150, loss: 0.007427656091749668
step: 160, loss: 0.012874292209744453

```



```

step: 170, loss: 0.0022480112966150045
step: 180, loss: 0.007183664944022894
step: 190, loss: 0.203715980052948
step: 200, loss: 0.005425562616437674
step: 210, loss: 0.01585371047258377
step: 220, loss: 0.005920805037021637
step: 230, loss: 0.015024008229374886
/usr/local/lib/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=0.980963045912654, f1=0.9796839729119639,
best_f1=0.9796839729119639
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0009670207509770989
step: 10, loss: 0.00953495129942894
step: 20, loss: 0.005000092554837465
step: 30, loss: 0.0018335779896005988
step: 40, loss: 0.0004391554684843868
step: 50, loss: 0.00042973330710083246
step: 60, loss: 0.0003775745281018317
step: 70, loss: 0.12463563680648804
step: 80, loss: 0.0013281282735988498
step: 90, loss: 0.006905474700033665
step: 100, loss: 0.00114630616735667
step: 110, loss: 0.0030999062582850456
step: 120, loss: 0.0006576831801794469
step: 130, loss: 0.007723875343799591
step: 140, loss: 0.0024883055593818426
step: 150, loss: 0.002198756206780672
step: 160, loss: 0.0037129665724933147
step: 170, loss: 0.0004272288060747087
step: 180, loss: 0.00034931403934024274
step: 190, loss: 0.0006032600649632514
step: 200, loss: 0.0007971973391249776
step: 210, loss: 0.01033781748265028
step: 220, loss: 0.0031644843984395266
step: 230, loss: 0.0007460821070708334
/usr/local/lib/python3.7/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=0.9809203142536477, f1=0.9699666295884317,
best_f1=0.9796839729119639

```

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

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

step: 0, loss: 0.0009591256966814399  
step: 10, loss: 0.0005017781513743103  
step: 20, loss: 0.021455027163028717  
step: 30, loss: 0.0005108578479848802  
step: 40, loss: 0.0009030138608068228  
step: 50, loss: 0.0034112134017050266  
step: 60, loss: 0.0007213001372292638  
step: 70, loss: 0.0006213239976204932  
step: 80, loss: 0.003919612616300583  
step: 90, loss: 0.0007272300426848233  
step: 100, loss: 0.08060546964406967  
step: 110, loss: 0.004856579005718231  
step: 120, loss: 0.0009638870833441615  
step: 130, loss: 0.003619919531047344  
step: 140, loss: 0.0008702790946699679  
step: 150, loss: 0.01179694663733244  
step: 160, loss: 0.0014002377865836024  
step: 170, loss: 0.0031083389185369015  
step: 180, loss: 0.0003153184661641717  
step: 190, loss: 0.00013995503832120448  
step: 200, loss: 0.022265180945396423  
step: 210, loss: 0.0014297803863883018  
step: 220, loss: 0.0002574908430688083  
step: 230, loss: 0.004008038435131311

/usr/local/lib/python3.7/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=0.9730941704035874, f1=0.9642058165548099,  
best\_f1=0.9796839729119639

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

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

step: 0, loss: 0.0002949887712020427  
step: 10, loss: 0.021472115069627762  
step: 20, loss: 0.001247613807208836  
step: 30, loss: 0.0004635625227820128  
step: 40, loss: 0.000900403072591871  
step: 50, loss: 0.0010729653295129538  
step: 60, loss: 0.0014529806794598699  
step: 70, loss: 0.001092495396733284  
step: 80, loss: 0.021014899015426636  
step: 90, loss: 0.0029837272595614195  
step: 100, loss: 0.0007513427408412099  
step: 110, loss: 0.0012258231872692704  
step: 120, loss: 0.0019871562253683805

```

step: 130, loss: 0.05302342399954796
step: 140, loss: 0.0004836528387386352
step: 150, loss: 0.1698484867811203
step: 160, loss: 0.0015513699036091566
step: 170, loss: 0.006984334904700518
step: 180, loss: 0.0011089255567640066
step: 190, loss: 0.003232580842450261
step: 200, loss: 0.0029940796084702015
step: 210, loss: 0.0021565889474004507
step: 220, loss: 0.0005137945991009474
step: 230, loss: 0.00022311304928734899
/usr/local/lib/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=0.9831649831649831, f1=0.9831649831649831,
best_f1=0.9831649831649831
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0017204009927809238
step: 10, loss: 0.0005560991121456027
step: 20, loss: 0.000722876749932766
step: 30, loss: 0.0012642014771699905
step: 40, loss: 0.003075425513088703
step: 50, loss: 0.0008559454581700265
step: 60, loss: 0.00019109617278445512
step: 70, loss: 0.03882889449596405
step: 80, loss: 0.0004537806671578437
step: 90, loss: 0.005362921394407749
step: 100, loss: 0.00021796040527988225
step: 110, loss: 0.0001959790533874184
step: 120, loss: 0.03330694139003754
step: 130, loss: 0.002309044823050499
step: 140, loss: 0.00011449614248704165
step: 150, loss: 0.00019312751828692853
step: 160, loss: 0.00012186180538265035
step: 170, loss: 5.0190741603728384e-05
step: 180, loss: 0.0007595678325742483
step: 190, loss: 5.8051515225088224e-05
step: 200, loss: 7.580640522064641e-05
step: 210, loss: 0.0015759498346596956
step: 220, loss: 0.00020349562691990286
step: 230, loss: 0.0006511099054478109
/usr/local/lib/python3.7/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=0.9820224719101124, f1=0.9775280898876404,
best_f1=0.9831649831649831
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00032076984643936157
step: 10, loss: 0.0001855433511082083
step: 20, loss: 0.000348919682437554
step: 30, loss: 5.745440284954384e-05
step: 40, loss: 0.013020376674830914
step: 50, loss: 6.204542296472937e-05
step: 60, loss: 5.917880844208412e-05
step: 70, loss: 0.003419417655095458
step: 80, loss: 0.0001292603265028447
step: 90, loss: 0.00011156481195939705
step: 100, loss: 0.00035523815313354135
step: 110, loss: 0.0005850454326719046
step: 120, loss: 0.010843967087566853
step: 130, loss: 0.0002792346349451691
step: 140, loss: 0.00012242680531926453
step: 150, loss: 0.00024633921566419303
step: 160, loss: 5.2432114898692816e-05
step: 170, loss: 0.000315526791382581
step: 180, loss: 0.03772986680269241
step: 190, loss: 0.00034204169060103595
step: 200, loss: 0.0001518077915534377
step: 210, loss: 0.048837751150131226
step: 220, loss: 0.001952730817720294
step: 230, loss: 7.600565731991082e-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=0.9820224719101124, f1=0.9787234042553192,
best_f1=0.9831649831649831
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 7.234704389702529e-05
step: 10, loss: 0.00019648371380753815
step: 20, loss: 0.01885083317756653
step: 30, loss: 0.010779124684631824
step: 40, loss: 0.0003961570910178125
step: 50, loss: 0.0001042397998389788
step: 60, loss: 0.002124491613358259
step: 70, loss: 0.004246925935149193
step: 80, loss: 0.00022942645591683686

```

```

step: 90, loss: 0.21009008586406708
step: 100, loss: 0.00022699045075569302
step: 110, loss: 0.00047683980665169656
step: 120, loss: 0.0002816602645907551
step: 130, loss: 0.00020541799312923104
step: 140, loss: 0.0026388908736407757
step: 150, loss: 0.00028944460791535676
step: 160, loss: 0.0011749346740543842
step: 170, loss: 0.00029727473156526685
step: 180, loss: 0.00015211898426059633
step: 190, loss: 0.00010284734889864922
step: 200, loss: 0.000710601918399334
step: 210, loss: 0.0002587584313005209
step: 220, loss: 0.0009695907356217504
step: 230, loss: 0.00020419416250661016
/usr/local/lib/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=0.9842696629213483, 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: 0.0001188965470646508
step: 10, loss: 5.6152195611502975e-05
step: 20, loss: 0.0003830956411547959
step: 30, loss: 0.02676595374941826
step: 40, loss: 0.0242752842605114
step: 50, loss: 0.00020288201631046832
step: 60, loss: 0.0002589320356491953
step: 70, loss: 9.300607052864507e-05
step: 80, loss: 5.7460434618406e-05
step: 90, loss: 0.0003615952155087143
step: 100, loss: 7.22004406270571e-05
step: 110, loss: 0.00011722334602382034
step: 120, loss: 0.0002900389372371137
step: 130, loss: 0.00029658753192052245
step: 140, loss: 0.00016735473764128983
step: 150, loss: 0.0001447979302611202
step: 160, loss: 0.0013361723395064473
step: 170, loss: 0.0020499706733971834
step: 180, loss: 0.00019115947361569852
step: 190, loss: 0.0009479449363425374
step: 200, loss: 5.799436621600762e-05

```

```

step: 210, loss: 0.007118409965187311
step: 220, loss: 0.005465859081596136
step: 230, loss: 0.00016300761490128934
/usr/local/lib/python3.7/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=0.9842696629213483, f1=0.980963045912654,
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: 7.98258333816193e-05
step: 10, loss: 0.00018067171913571656
step: 20, loss: 0.0015938645228743553
step: 30, loss: 3.6708668631035835e-05
step: 40, loss: 0.0008027477306313813
step: 50, loss: 0.0017399401403963566
step: 60, loss: 9.268642315873876e-05
step: 70, loss: 0.05522292107343674
step: 80, loss: 0.00016850225802045316
step: 90, loss: 4.7079611249500886e-05
step: 100, loss: 8.443260594503954e-05
step: 110, loss: 0.0044048139825463295
step: 120, loss: 7.111661398084834e-05
step: 130, loss: 8.196304406737909e-05
step: 140, loss: 0.0027332636527717113
step: 150, loss: 4.724492828245275e-05
step: 160, loss: 0.017187079414725304
step: 170, loss: 0.0002803397364914417
step: 180, loss: 0.007453208323568106
step: 190, loss: 0.0003212160081602633
step: 200, loss: 0.00010833463602466509
step: 210, loss: 0.00036953395465388894
step: 220, loss: 0.000245588191319257
step: 230, loss: 7.688700134167448e-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=0.9831649831649831, f1=0.9798657718120806,
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: 7.521810766775161e-05
step: 10, loss: 3.771914998651482e-05
step: 20, loss: 0.00010488511907169595
step: 30, loss: 0.00019077349861618131
step: 40, loss: 4.537441418506205e-05

```

```

step: 50, loss: 3.587586616049521e-05
step: 60, loss: 6.185044185258448e-05
step: 70, loss: 5.757052713306621e-05
step: 80, loss: 7.456001185346395e-05
step: 90, loss: 0.00015716379857622087
step: 100, loss: 0.001025811186991632
step: 110, loss: 0.00010063155787065625
step: 120, loss: 4.3281452235532925e-05
step: 130, loss: 0.00031963762012310326
step: 140, loss: 0.0031930580735206604
step: 150, loss: 6.047669376130216e-05
step: 160, loss: 8.716571028344333e-05
step: 170, loss: 0.0002698382595553994
step: 180, loss: 0.00012010219506919384
step: 190, loss: 0.0001946143456734717
step: 200, loss: 0.00014575180830433965
step: 210, loss: 0.00018850521882995963
step: 220, loss: 0.004015225451439619
step: 230, loss: 5.0714781536953524e-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=0.9830890642615557, f1=0.9809203142536477,
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: 0.0018409028416499496
step: 10, loss: 0.0001542306854389608
step: 20, loss: 0.00029874240863136947
step: 30, loss: 4.684249506681226e-05
step: 40, loss: 3.4252698242198676e-05
step: 50, loss: 4.932699084747583e-05
step: 60, loss: 0.0239032544195652
step: 70, loss: 0.0012622610665857792
step: 80, loss: 5.6480981584172696e-05
step: 90, loss: 3.874441244988702e-05
step: 100, loss: 2.7249478080193512e-05
step: 110, loss: 0.0062121194787323475
step: 120, loss: 0.02157861553132534
step: 130, loss: 7.028587424429134e-05
step: 140, loss: 0.00055761105613783
step: 150, loss: 5.6928667618194595e-05
step: 160, loss: 0.0003412714577279985
step: 170, loss: 2.574057725723833e-05
step: 180, loss: 0.00019093864830210805
step: 190, loss: 5.43282221769914e-05
step: 200, loss: 0.0002752675791271031

```

```

step: 210, loss: 0.0013603395782411098
step: 220, loss: 4.447377068572678e-05
step: 230, loss: 0.0003380323469173163
/usr/local/lib/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=0.9843400447427293, f1=0.9787709497206705,
best_f1=0.9787709497206705
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, 152.39it/s]
load_f1 = 0.9831649831649831
real_f1 = 0.9820627802690582
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, 139.38it/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/
```

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                 : 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: 0.6016530990600586
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: 0.5408405065536499
step: 20, loss: 0.28653261065483093
step: 30, loss: 0.35872137546539307
step: 40, loss: 0.41295287013053894
step: 50, loss: 0.6875376105308533
step: 60, loss: 0.3477814197540283
step: 70, loss: 0.40336477756500244
step: 80, loss: 0.4122371971607208
step: 90, loss: 0.1399577409029007
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 8192.0
step: 100, loss: 0.4617884159088135
step: 110, loss: 0.22702954709529877
step: 120, loss: 0.2952024042606354
step: 130, loss: 0.4424716532230377
step: 140, loss: 0.18787793815135956
step: 150, loss: 0.0972527340054512
step: 160, loss: 0.2083234190940857
step: 170, loss: 0.0806606113910675
step: 180, loss: 0.07299027591943741
step: 190, loss: 0.22406448423862457
step: 200, loss: 0.08400201052427292
step: 210, loss: 0.08374384045600891
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 4096.0
step: 220, loss: 0.1012105718255043
step: 230, loss: 0.23348629474639893
step: 240, loss: 0.05722294747829437
step: 250, loss: 0.009426192380487919

```

```

step: 260, loss: 0.20769068598747253
step: 270, loss: 0.3662005364894867
step: 280, loss: 0.0879683718085289
step: 290, loss: 0.17358660697937012
step: 300, loss: 0.08575130254030228
step: 310, loss: 0.05794616788625717
step: 320, loss: 0.08948354423046112
step: 330, loss: 0.054789766669273376
step: 340, loss: 0.2938365042209625
step: 350, loss: 0.1354418694972992
step: 360, loss: 0.009953401051461697
step: 370, loss: 0.1263837218284607
step: 380, loss: 0.09944333136081696
step: 390, loss: 0.017864318564534187
step: 400, loss: 0.021069254726171494
step: 410, loss: 0.2636454701423645
step: 420, loss: 0.00951365940272808
step: 430, loss: 0.029558170586824417
step: 440, loss: 0.07601604610681534
step: 450, loss: 0.22567126154899597
step: 460, loss: 0.1728018820285797
step: 470, loss: 0.0456177219748497
step: 480, loss: 0.27847304940223694
step: 490, loss: 0.1206713393330574
step: 500, loss: 0.09007709473371506
step: 510, loss: 0.03325798735022545
step: 520, loss: 0.19587314128875732
step: 530, loss: 0.048381321132183075
/usr/local/lib/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=0.9298653042266605, f1=0.933826931975937,
best_f1=0.933826931975937
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.05784802883863449
step: 10, loss: 0.06369255483150482
step: 20, loss: 0.02335163950920105
step: 30, loss: 0.12659195065498352
step: 40, loss: 0.030319146811962128
step: 50, loss: 0.24631290137767792
step: 60, loss: 0.06275836378335953
step: 70, loss: 0.05508187413215637

```

```
step: 80, loss: 0.07162082195281982
step: 90, loss: 0.009767288342118263
step: 100, loss: 0.23254084587097168
step: 110, loss: 0.037632569670677185
step: 120, loss: 0.04403890296816826
step: 130, loss: 0.01886136457324028
step: 140, loss: 0.05292225629091263
step: 150, loss: 0.03667594864964485
step: 160, loss: 0.02141730859875679
step: 170, loss: 0.03861039876937866
step: 180, loss: 0.03241823986172676
step: 190, loss: 0.05772625282406807
step: 200, loss: 0.26222914457321167
step: 210, loss: 0.048020586371421814
step: 220, loss: 0.002724266843870282
step: 230, loss: 0.05773240327835083
step: 240, loss: 0.13335634768009186
step: 250, loss: 0.09444093704223633
step: 260, loss: 0.30314022302627563
step: 270, loss: 0.09897645562887192
step: 280, loss: 0.03409251570701599
step: 290, loss: 0.03282182663679123
step: 300, loss: 0.06005137786269188
step: 310, loss: 0.1350526660680771
step: 320, loss: 0.045950569212436676
step: 330, loss: 0.03982783108949661
step: 340, loss: 0.0903383195400238
step: 350, loss: 0.001642780378460884
step: 360, loss: 0.163279727101326
step: 370, loss: 0.022730210795998573
step: 380, loss: 0.10548748821020126
step: 390, loss: 0.005326990969479084
step: 400, loss: 0.06807754188776016
step: 410, loss: 0.0716806948184967
step: 420, loss: 0.07305973023176193
step: 430, loss: 0.16609054803848267
step: 440, loss: 0.0029850462451577187
step: 450, loss: 0.09967481344938278
step: 460, loss: 0.050001710653305054
step: 470, loss: 0.03572541102766991
step: 480, loss: 0.006130714900791645
step: 490, loss: 0.07009946554899216
step: 500, loss: 0.02521761879324913
step: 510, loss: 0.05420561507344246
step: 520, loss: 0.3604614734649658
step: 530, loss: 0.11859466880559921
/usr/local/lib/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=0.9479553903345725, f1=0.9455560725919032,
best_f1=0.9455560725919032
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.10970021039247513
step: 10, loss: 0.11173954606056213
step: 20, loss: 0.023615362122654915
step: 30, loss: 0.03223143890500069
step: 40, loss: 0.2618955075740814
step: 50, loss: 0.0665571317076683
step: 60, loss: 0.010638369247317314
step: 70, loss: 0.0035402653738856316
step: 80, loss: 0.1855698972940445
step: 90, loss: 0.015388491563498974
step: 100, loss: 0.020752031356096268
step: 110, loss: 0.015686361119151115
step: 120, loss: 0.20281414687633514
step: 130, loss: 0.044531047344207764
step: 140, loss: 0.03794972226023674
step: 150, loss: 0.04908537119626999
step: 160, loss: 0.021808039397001266
step: 170, loss: 0.03843977674841881
step: 180, loss: 0.05240754410624504
step: 190, loss: 0.01150195486843586
step: 200, loss: 0.06906747072935104
step: 210, loss: 0.05263938382267952
step: 220, loss: 0.08989366888999939
step: 230, loss: 0.06610200554132462
step: 240, loss: 0.12612955272197723
step: 250, loss: 0.2356761246919632
step: 260, loss: 0.08307333290576935
step: 270, loss: 0.01218412071466446
step: 280, loss: 0.09309328347444534
step: 290, loss: 0.031319763511419296
step: 300, loss: 0.10742229968309402
step: 310, loss: 0.15073467791080475
step: 320, loss: 0.01354017574340105
step: 330, loss: 0.03713750094175339
step: 340, loss: 0.012175904586911201
step: 350, loss: 0.050819069147109985
step: 360, loss: 0.0076941740699112415
step: 370, loss: 0.028160452842712402

```

```

step: 380, loss: 0.010294970124959946
step: 390, loss: 0.06840109080076218
step: 400, loss: 0.042622555047273636
step: 410, loss: 0.06737484782934189
step: 420, loss: 0.08204139024019241
step: 430, loss: 0.048282936215400696
step: 440, loss: 0.2508247494697571
step: 450, loss: 0.1101219654083252
step: 460, loss: 0.03355337306857109
step: 470, loss: 0.031527355313301086
step: 480, loss: 0.25817447900772095
step: 490, loss: 0.04609078913927078
step: 500, loss: 0.009882249869406223
step: 510, loss: 0.019901009276509285
step: 520, loss: 0.02001138962805271
step: 530, loss: 0.023192306980490685
/usr/local/lib/python3.7/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=0.9452503509592888, f1=0.9411764705882353,
best_f1=0.9455560725919032
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0359351709485054
step: 10, loss: 0.026728352531790733
step: 20, loss: 0.042070422321558
step: 30, loss: 0.04014396667480469
step: 40, loss: 0.013134544715285301
step: 50, loss: 0.06681036949157715
step: 60, loss: 0.023093638941645622
step: 70, loss: 0.06435362249612808
step: 80, loss: 0.10143382847309113
step: 90, loss: 0.12168386578559875
step: 100, loss: 0.01207798719406128
step: 110, loss: 0.07741712033748627
step: 120, loss: 0.0010100057115778327
step: 130, loss: 0.0035036751069128513
step: 140, loss: 0.04455578327178955
step: 150, loss: 0.014379926957190037
step: 160, loss: 0.006537744775414467
step: 170, loss: 0.038175102323293686
step: 180, loss: 0.07914018630981445
step: 190, loss: 0.043242208659648895
step: 200, loss: 0.022722844034433365
step: 210, loss: 0.004958782810717821
step: 220, loss: 0.2164851576089859
step: 230, loss: 0.015869295224547386

```

```

step: 240, loss: 0.009892456233501434
step: 250, loss: 0.09961816668510437
step: 260, loss: 0.0019455589354038239
step: 270, loss: 0.07434170693159103
step: 280, loss: 0.003291697707027197
step: 290, loss: 0.013788674026727676
step: 300, loss: 0.0012893383391201496
step: 310, loss: 0.002504087286069989
step: 320, loss: 0.18358655273914337
step: 330, loss: 0.03900253027677536
step: 340, loss: 0.010082261636853218
step: 350, loss: 0.08754180371761322
step: 360, loss: 0.04675056412816048
step: 370, loss: 0.009363174438476562
step: 380, loss: 0.008976919576525688
step: 390, loss: 0.0018760983366519213
step: 400, loss: 0.04307993873953819
step: 410, loss: 0.013529636897146702
step: 420, loss: 0.010325999930500984
step: 430, loss: 0.01909567043185234
step: 440, loss: 0.018484922125935555
step: 450, loss: 0.028417831286787987
step: 460, loss: 0.009586052969098091
step: 470, loss: 0.0013525920221582055
step: 480, loss: 0.09407349675893784
step: 490, loss: 0.024726012721657753
step: 500, loss: 0.36926764249801636
step: 510, loss: 0.0732816606760025
step: 520, loss: 0.01948152668774128
step: 530, loss: 0.35449549555778503
/usr/local/lib/python3.7/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=0.9452181987000929, f1=0.9416666666666667,
best_f1=0.9455560725919032
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.006918288301676512
step: 10, loss: 0.004789547063410282
step: 20, loss: 0.013705424033105373
step: 30, loss: 0.023515785112977028
step: 40, loss: 0.012093986384570599
step: 50, loss: 0.00210145628079772
step: 60, loss: 0.006938346195966005
step: 70, loss: 0.020094215869903564
step: 80, loss: 0.02079654112458229
step: 90, loss: 0.05142612010240555

```

```
step: 100, loss: 0.018386930227279663
step: 110, loss: 0.0006750907632522285
step: 120, loss: 0.23073774576187134
step: 130, loss: 0.027179211378097534
step: 140, loss: 0.020720748230814934
step: 150, loss: 0.05981963127851486
step: 160, loss: 0.008472807705402374
step: 170, loss: 0.015034377574920654
step: 180, loss: 0.015584232285618782
step: 190, loss: 0.001842675032094121
step: 200, loss: 0.0036947159096598625
step: 210, loss: 0.02128755860030651
step: 220, loss: 0.04454192891716957
step: 230, loss: 0.0028236545622348785
step: 240, loss: 0.030380673706531525
step: 250, loss: 0.08878667652606964
step: 260, loss: 0.00379247241653502
step: 270, loss: 0.001432347111403942
step: 280, loss: 0.009857376106083393
step: 290, loss: 0.00976146012544632
step: 300, loss: 0.08936517685651779
step: 310, loss: 0.05314638838171959
step: 320, loss: 0.0058546108193695545
step: 330, loss: 0.0038691714871674776
step: 340, loss: 0.011121603660285473
step: 350, loss: 0.003881945740431547
step: 360, loss: 0.0003224439569748938
step: 370, loss: 0.00034039950696751475
step: 380, loss: 0.014573422260582447
step: 390, loss: 0.0012833827640861273
step: 400, loss: 0.017811022698879242
step: 410, loss: 0.07852821052074432
step: 420, loss: 0.21845190227031708
step: 430, loss: 0.07336302101612091
step: 440, loss: 0.006760946940630674
step: 450, loss: 0.010712899267673492
step: 460, loss: 0.025615738704800606
step: 470, loss: 0.0891864225268364
step: 480, loss: 0.014612526632845402
step: 490, loss: 0.0962894931435585
step: 500, loss: 0.06243037059903145
step: 510, loss: 0.0194003414362669
step: 520, loss: 0.08815788477659225
step: 530, loss: 0.030127136036753654
/usr/local/lib/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=0.9500233535730968, f1=0.941286989196806,
best_f1=0.941286989196806
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
```

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

```
step: 0, loss: 0.043854255229234695
step: 10, loss: 0.0549180768430233
step: 20, loss: 0.01309348177164793
step: 30, loss: 0.0020928129088133574
step: 40, loss: 0.00034825364127755165
step: 50, loss: 0.001978342654183507
step: 60, loss: 0.0002814355830196291
step: 70, loss: 0.0013180356472730637
step: 80, loss: 0.02076425589621067
step: 90, loss: 0.10781989246606827
step: 100, loss: 0.011818824335932732
step: 110, loss: 0.015303621999919415
step: 120, loss: 0.008018726482987404
step: 130, loss: 0.0027926114853471518
step: 140, loss: 0.004379339516162872
step: 150, loss: 0.0007390871760435402
step: 160, loss: 0.03754456713795662
step: 170, loss: 0.004415706731379032
step: 180, loss: 0.002336923498660326
step: 190, loss: 0.02492458000779152
step: 200, loss: 0.03368731588125229
step: 210, loss: 0.01011132262647152
step: 220, loss: 0.005515509285032749
step: 230, loss: 0.04169200733304024
step: 240, loss: 0.0011298198951408267
step: 250, loss: 0.11767568439245224
step: 260, loss: 0.0012717292411252856
step: 270, loss: 0.012609170749783516
step: 280, loss: 0.001657032291404903
step: 290, loss: 0.0011653626570478082
step: 300, loss: 0.013148698955774307
step: 310, loss: 0.12386675179004669
step: 320, loss: 0.0001459649356547743
step: 330, loss: 0.02099902741611004
step: 340, loss: 0.005090456455945969
step: 350, loss: 0.01282986905425787
step: 360, loss: 0.04420290142297745
step: 370, loss: 0.002986110746860504
step: 380, loss: 0.0002478230162523687
step: 390, loss: 0.0026343849021941423
```

```

step: 400, loss: 0.004384488333016634
step: 410, loss: 0.002423298079520464
step: 420, loss: 0.020233260467648506
step: 430, loss: 0.020891990512609482
step: 440, loss: 0.0038720250595360994
step: 450, loss: 0.17996086180210114
step: 460, loss: 0.002331545576453209
step: 470, loss: 0.010562616400420666
step: 480, loss: 0.007144487462937832
step: 490, loss: 0.13781823217868805
step: 500, loss: 0.0006835209205746651
step: 510, loss: 0.2591519057750702
step: 520, loss: 0.0035544000566005707
step: 530, loss: 0.004062245599925518
/usr/local/lib/python3.7/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=0.9453162149310508, f1=0.9410085632730733,
best_f1=0.941286989196806
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.005555413663387299
step: 10, loss: 0.0009964030468836427
step: 20, loss: 0.013423729687929153
step: 30, loss: 0.030157960951328278
step: 40, loss: 0.009430057369172573
step: 50, loss: 0.008510874584317207
step: 60, loss: 0.00800918135792017
step: 70, loss: 0.0018311613239347935
step: 80, loss: 0.004161683842539787
step: 90, loss: 0.0002657197474036366
step: 100, loss: 0.005602334160357714
step: 110, loss: 0.0006791893974877894
step: 120, loss: 0.2753038704395294
step: 130, loss: 0.008400698192417622
step: 140, loss: 0.003291537519544363
step: 150, loss: 0.0025641778483986855
step: 160, loss: 0.016405753791332245
step: 170, loss: 0.0011561751598492265
step: 180, loss: 0.04825247824192047
step: 190, loss: 0.021337559446692467
step: 200, loss: 0.0019745659083127975
step: 210, loss: 0.0029578381218016148
step: 220, loss: 0.008158196695148945
step: 230, loss: 0.002637979108840227
step: 240, loss: 0.00414246367290616
step: 250, loss: 0.007896088063716888

```

```

step: 260, loss: 0.009378033690154552
step: 270, loss: 0.012813379056751728
step: 280, loss: 0.0041433777660131454
step: 290, loss: 0.12041941285133362
step: 300, loss: 0.002391788177192211
step: 310, loss: 0.001387760043144226
step: 320, loss: 0.02337006852030754
step: 330, loss: 0.0006564960349351168
step: 340, loss: 0.017182407900691032
step: 350, loss: 0.004663480445742607
step: 360, loss: 0.010886318981647491
step: 370, loss: 0.011584371328353882
step: 380, loss: 0.000515387044288218
step: 390, loss: 0.007614050060510635
step: 400, loss: 0.037287428975105286
step: 410, loss: 0.00016042821516748518
step: 420, loss: 0.0149949686601758
step: 430, loss: 0.0013773980317637324
step: 440, loss: 0.0004788915393874049
step: 450, loss: 0.025299161672592163
step: 460, loss: 0.002676486037671566
step: 470, loss: 0.26857906579971313
step: 480, loss: 0.0016864617355167866
step: 490, loss: 0.005268121603876352
step: 500, loss: 0.007861284539103508
step: 510, loss: 0.0014681603061035275
step: 520, loss: 0.0020440183579921722
step: 530, loss: 0.032150767743587494
/usr/local/lib/python3.7/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=0.9402427637721755, f1=0.9342105263157895,
best_f1=0.941286989196806
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0008746355306357145
step: 10, loss: 0.0004877602041233331
step: 20, loss: 0.0007298318669199944
step: 30, loss: 0.0018322198884561658
step: 40, loss: 0.0004972138558514416
step: 50, loss: 0.004275182262063026
step: 60, loss: 0.0002898193197324872
step: 70, loss: 0.0016361285233870149
step: 80, loss: 0.022310560569167137
step: 90, loss: 0.013630124740302563
step: 100, loss: 0.007846484892070293
step: 110, loss: 0.0003398441185709089

```

```
step: 120, loss: 0.011694865301251411
step: 130, loss: 0.0018024482997134328
step: 140, loss: 0.000515044666826725
step: 150, loss: 0.001080166082829237
step: 160, loss: 0.002987568499520421
step: 170, loss: 0.05419242009520531
step: 180, loss: 0.0015480410074815154
step: 190, loss: 0.0123860202729702
step: 200, loss: 0.004001634661108255
step: 210, loss: 0.05849594622850418
step: 220, loss: 0.0031122767832130194
step: 230, loss: 0.08431331068277359
step: 240, loss: 0.11878979951143265
step: 250, loss: 0.0005712403799407184
step: 260, loss: 0.0020281802862882614
step: 270, loss: 0.027611156925559044
step: 280, loss: 0.002180828945711255
step: 290, loss: 0.0022865913342684507
step: 300, loss: 0.00034951846464537084
step: 310, loss: 0.0014648103388026357
step: 320, loss: 0.0029751486144959927
step: 330, loss: 0.0013221633853390813
step: 340, loss: 0.06280297040939331
step: 350, loss: 0.0006355607183650136
step: 360, loss: 0.08707458525896072
step: 370, loss: 0.02056816592812538
step: 380, loss: 0.0015494738472625613
step: 390, loss: 0.007509011309593916
step: 400, loss: 0.05490861460566521
step: 410, loss: 0.11420269310474396
step: 420, loss: 0.0001302588643739
step: 430, loss: 0.005537188146263361
step: 440, loss: 0.017094528302550316
step: 450, loss: 0.003361625364050269
step: 460, loss: 0.016246983781456947
step: 470, loss: 0.08182860910892487
step: 480, loss: 0.009460925124585629
step: 490, loss: 0.07191771268844604
step: 500, loss: 0.001842243131250143
step: 510, loss: 0.02706834115087986
step: 520, loss: 0.0003515793941915035
step: 530, loss: 0.0004505084070842713
/usr/local/lib/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=0.950187969924812, f1=0.9418439716312057,
best_f1=0.9418439716312057
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0008860734524205327
step: 10, loss: 0.004002880770713091
step: 20, loss: 0.0015995565336197615
step: 30, loss: 0.07992997020483017
step: 40, loss: 0.01506322156637907
step: 50, loss: 0.001921455143019557
step: 60, loss: 0.0009824992157518864
step: 70, loss: 0.0005767598049715161
step: 80, loss: 0.014030566439032555
step: 90, loss: 0.04659184068441391
step: 100, loss: 0.0026293464470654726
step: 110, loss: 0.002846753690391779
step: 120, loss: 0.006421966943889856
step: 130, loss: 0.003273813519626856
step: 140, loss: 0.007793476339429617
step: 150, loss: 0.002493505598977208
step: 160, loss: 0.0025967333931475878
step: 170, loss: 0.001107400981709361
step: 180, loss: 0.009270939975976944
step: 190, loss: 0.0025500894989818335
step: 200, loss: 0.0008827592828311026
step: 210, loss: 0.006448649801313877
step: 220, loss: 0.09797403216362
step: 230, loss: 0.000415524875279516
step: 240, loss: 0.002384789753705263
step: 250, loss: 0.0009487204952165484
step: 260, loss: 0.021943707019090652
step: 270, loss: 0.009214391000568867
step: 280, loss: 0.0008350482094101608
step: 290, loss: 0.0007018441101536155
step: 300, loss: 0.007238366641104221
step: 310, loss: 0.0029286423232406378
step: 320, loss: 0.0011196386767551303
step: 330, loss: 0.0018926130142062902
step: 340, loss: 0.006761026568710804
step: 350, loss: 0.020973296836018562
step: 360, loss: 0.03646313026547432
step: 370, loss: 0.002031535841524601
step: 380, loss: 0.00106630299706012
step: 390, loss: 0.0001816151780076325
step: 400, loss: 0.031359512358903885
step: 410, loss: 0.0008129371562972665
```

```

step: 420, loss: 0.0011846349807456136
step: 430, loss: 0.0027269686106592417
step: 440, loss: 9.34306881390512e-05
step: 450, loss: 0.029804568737745285
step: 460, loss: 0.0030003751162439585
step: 470, loss: 0.026970289647579193
step: 480, loss: 0.003604867262765765
step: 490, loss: 0.0010238748509436846
step: 500, loss: 0.028632057830691338
step: 510, loss: 0.10528609901666641
step: 520, loss: 0.012703250162303448
step: 530, loss: 0.040222782641649246
/usr/local/lib/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=0.9527777777777777, f1=0.9469767441860464,
best_f1=0.9469767441860464
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.008707325905561447
step: 10, loss: 0.001683301292359829
step: 20, loss: 9.56230578594841e-05
step: 30, loss: 0.007446676027029753
step: 40, loss: 0.0015721881063655019
step: 50, loss: 0.0005267660599201918
step: 60, loss: 0.005307563580572605
step: 70, loss: 0.03282836452126503
step: 80, loss: 0.00012854867964051664
step: 90, loss: 0.004134104121476412
step: 100, loss: 0.02804214134812355
step: 110, loss: 0.02552664466202259
step: 120, loss: 0.0021253919694572687
step: 130, loss: 0.006635759491473436
step: 140, loss: 0.0013896750751882792
step: 150, loss: 0.003103662049397826
step: 160, loss: 0.026078738272190094
step: 170, loss: 0.014943505637347698
step: 180, loss: 0.02329309470951557
step: 190, loss: 7.488032133551314e-05
step: 200, loss: 0.0011645944323390722
step: 210, loss: 0.0014243398327380419
step: 220, loss: 0.0007537057972513139
step: 230, loss: 0.005350407678633928

```

```

step: 240, loss: 0.00035411881981417537
step: 250, loss: 0.0009598026517778635
step: 260, loss: 0.0011736978776752949
step: 270, loss: 4.6813394874334335e-05
step: 280, loss: 0.0748748630285263
step: 290, loss: 0.0002711203705985099
step: 300, loss: 0.0042902217246592045
step: 310, loss: 0.0023204258177429438
step: 320, loss: 0.061923619359731674
step: 330, loss: 0.015371795743703842
step: 340, loss: 0.0014038638910278678
step: 350, loss: 0.05940777063369751
step: 360, loss: 0.00017801392823457718
step: 370, loss: 0.0012813986977562308
step: 380, loss: 0.0004449300467967987
step: 390, loss: 0.00877027865499258
step: 400, loss: 0.004513143561780453
step: 410, loss: 0.001803395338356495
step: 420, loss: 9.14255651878193e-05
step: 430, loss: 0.01376900915056467
step: 440, loss: 0.00021173048298805952
step: 450, loss: 0.017020365223288536
step: 460, loss: 0.0003269918670412153
step: 470, loss: 0.004391095135360956
step: 480, loss: 0.0002873440971598029
step: 490, loss: 0.0058835819363594055
step: 500, loss: 0.009304195642471313
step: 510, loss: 0.0003911121457349509
step: 520, loss: 0.002476766239851713
step: 530, loss: 0.010669856332242489
/usr/local/lib/python3.7/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=0.951851851851852, f1=0.9458082445576655,
best_f1=0.9469767441860464
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0006220873910933733
step: 10, loss: 0.0005455577047541738
step: 20, loss: 8.424015686614439e-05
step: 30, loss: 0.00015521436580456793
step: 40, loss: 0.00042618034058250487
step: 50, loss: 0.00012750618043355644
step: 60, loss: 0.0004109530127607286
step: 70, loss: 0.0001407250383635983
step: 80, loss: 0.007809455972164869
step: 90, loss: 0.00152855203486979

```

```
step: 100, loss: 0.010218752548098564
step: 110, loss: 0.00014783612277824432
step: 120, loss: 0.001217442099004984
step: 130, loss: 0.00010385778296040371
step: 140, loss: 0.00011118053225800395
step: 150, loss: 0.00013456668239086866
step: 160, loss: 0.0011762662325054407
step: 170, loss: 0.003271055407822132
step: 180, loss: 4.090202855877578e-05
step: 190, loss: 0.007583215832710266
step: 200, loss: 0.001232219859957695
step: 210, loss: 8.195603732019663e-05
step: 220, loss: 0.05285891145467758
step: 230, loss: 0.0017947423039004207
step: 240, loss: 0.0007421479094773531
step: 250, loss: 0.00019150493608321995
step: 260, loss: 0.001250464585609734
step: 270, loss: 0.006090317387133837
step: 280, loss: 0.0006934432312846184
step: 290, loss: 0.0006638043560087681
step: 300, loss: 3.751200347323902e-05
step: 310, loss: 0.0030944321770220995
step: 320, loss: 0.0019199518719688058
step: 330, loss: 0.0003969414974562824
step: 340, loss: 0.005832551047205925
step: 350, loss: 0.004020940978080034
step: 360, loss: 0.00017588349874131382
step: 370, loss: 0.0011278066085651517
step: 380, loss: 0.0016767096240073442
step: 390, loss: 0.0011163491290062666
step: 400, loss: 0.0021817830856889486
step: 410, loss: 0.001760201994329691
step: 420, loss: 0.01602233201265335
step: 430, loss: 0.015472856350243092
step: 440, loss: 0.00015566287038382143
step: 450, loss: 0.002190578728914261
step: 460, loss: 0.003143091220408678
step: 470, loss: 0.0009275175398215652
step: 480, loss: 0.0008460283279418945
step: 490, loss: 0.07175615429878235
step: 500, loss: 0.00066947570303455
step: 510, loss: 0.0002052946074400097
step: 520, loss: 0.0009446527110412717
step: 530, loss: 0.007274319417774677
/usr/local/lib/python3.7/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=0.951048951048951, f1=0.9454545454545454,
```



```
best_f1=0.9469767441860464
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.006881266832351685
step: 10, loss: 0.0015425619203597307
step: 20, loss: 0.005845093633979559
step: 30, loss: 5.5393986258422956e-05
step: 40, loss: 0.0067327567376196384
step: 50, loss: 0.00717034051194787
step: 60, loss: 2.8430758902686648e-05
step: 70, loss: 0.0006138192256912589
step: 80, loss: 0.29764485359191895
step: 90, loss: 0.004131283611059189
step: 100, loss: 0.06776203960180283
step: 110, loss: 0.001397539977915585
step: 120, loss: 0.0009998325258493423
step: 130, loss: 0.0032675613183528185
step: 140, loss: 0.0003475412377156317
step: 150, loss: 0.058091968297958374
step: 160, loss: 0.0017860678490251303
step: 170, loss: 0.00031096735619939864
step: 180, loss: 0.0003809505724348128
step: 190, loss: 0.018621359020471573
step: 200, loss: 0.0006118231685832143
step: 210, loss: 0.001108266063965857
step: 220, loss: 0.0005734594888053834
step: 230, loss: 0.029112551361322403
step: 240, loss: 0.001098955748602748
step: 250, loss: 0.00012672781303990632
step: 260, loss: 0.0007082909578457475
step: 270, loss: 0.0006672080489806831
step: 280, loss: 0.0012799871619790792
step: 290, loss: 7.595965871587396e-05
step: 300, loss: 0.0014760105405002832
step: 310, loss: 0.000369786168448627
step: 320, loss: 0.21305634081363678
step: 330, loss: 0.004410632885992527
step: 340, loss: 0.0026179328560829163
step: 350, loss: 0.0004993161419406533
step: 360, loss: 0.0022885575890541077
step: 370, loss: 0.0004593775956891477
step: 380, loss: 0.0009006606414914131
step: 390, loss: 0.014219379983842373
step: 400, loss: 0.0004971189773641527
step: 410, loss: 0.0006418824195861816
step: 420, loss: 0.0008689446258358657
step: 430, loss: 0.0005928716855123639
```

```
step: 440, loss: 0.0019902470521628857
step: 450, loss: 0.00057133415248245
step: 460, loss: 0.00014841335359960794
step: 470, loss: 0.001940669841133058
step: 480, loss: 0.0037825063336640596
step: 490, loss: 0.00225664465688169
step: 500, loss: 0.0005140199209563434
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 510, loss: 0.006364952772855759
step: 520, loss: 0.0014259335584938526
step: 530, loss: 0.0004474818706512451
/usr/local/lib/python3.7/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=0.948339483394834, f1=0.9426456984273821,
best_f1=0.9469767441860464
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0033618221059441566
step: 10, loss: 0.0002762152871582657
step: 20, loss: 0.0002758036716841161
step: 30, loss: 0.0007317520794458687
step: 40, loss: 0.0005944709409959614
step: 50, loss: 0.0019438257440924644
step: 60, loss: 0.00028361400472931564
step: 70, loss: 0.0005441776011139154
step: 80, loss: 0.0046615442261099815
step: 90, loss: 0.0014098136452957988
step: 100, loss: 0.00011108448234153911
step: 110, loss: 5.64622096135281e-05
step: 120, loss: 0.0003374250081833452
step: 130, loss: 0.0007592585170641541
step: 140, loss: 0.00027547471108846366
step: 150, loss: 0.001108434284105897
step: 160, loss: 0.011103007942438126
step: 170, loss: 0.00186661456245929
step: 180, loss: 0.00212098122574389
step: 190, loss: 4.128316868445836e-05
step: 200, loss: 5.562596197705716e-05
step: 210, loss: 4.6442528400802985e-05
step: 220, loss: 0.02166741155087948
step: 230, loss: 0.00027848215540871024
step: 240, loss: 0.002491488354280591
step: 250, loss: 0.0012350452598184347
step: 260, loss: 0.11679515987634659
step: 270, loss: 0.00011211973469471559
step: 280, loss: 0.0001266501349164173
```

```

step: 290, loss: 0.0003446451446507126
step: 300, loss: 1.690140925347805e-05
step: 310, loss: 0.005304827820509672
step: 320, loss: 0.00010957181075355038
step: 330, loss: 0.0007471635472029448
step: 340, loss: 0.10618622601032257
step: 350, loss: 0.00013619671517517418
step: 360, loss: 0.12331455945968628
step: 370, loss: 0.00038178550312295556
step: 380, loss: 0.00017766088421922177
step: 390, loss: 0.00029959093080833554
step: 400, loss: 0.027294989675283432
step: 410, loss: 7.774818368488923e-05
step: 420, loss: 0.00012903234164696187
step: 430, loss: 0.0002716419694479555
step: 440, loss: 3.016240407305304e-05
step: 450, loss: 0.09091512113809586
step: 460, loss: 0.019313951954245567
step: 470, loss: 0.002042289823293686
step: 480, loss: 1.647983845032286e-05
step: 490, loss: 0.00020879674411844462
step: 500, loss: 0.0002294548903591931
step: 510, loss: 0.0005496318335644901
step: 520, loss: 0.0012295338092371821
step: 530, loss: 8.553521911380813e-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=0.9508041627246925, f1=0.9430740037950663,
best_f1=0.9469767441860464
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0004180096439085901
step: 10, loss: 0.00023069503367878497
step: 20, loss: 0.0002184257609769702
step: 30, loss: 0.0014226097846403718
step: 40, loss: 0.00031757925171405077
step: 50, loss: 0.016930969431996346
step: 60, loss: 6.306605064310133e-05
step: 70, loss: 0.001852954039350152
step: 80, loss: 0.00021708168787881732
step: 90, loss: 6.840629794169217e-05
step: 100, loss: 0.000841341505292803
step: 110, loss: 0.00029643491143360734
step: 120, loss: 1.4766698768653441e-05
step: 130, loss: 7.6185395300854e-05
step: 140, loss: 0.00537909334525466

```

```

step: 150, loss: 0.01537015289068222
step: 160, loss: 8.834696200210601e-05
step: 170, loss: 0.0007590291206724942
step: 180, loss: 0.00020423824025783688
step: 190, loss: 0.0019363496685400605
step: 200, loss: 6.642203516094014e-05
step: 210, loss: 0.006545147858560085
step: 220, loss: 0.0002684843202587217
step: 230, loss: 0.0007981043308973312
step: 240, loss: 0.02576333098113537
step: 250, loss: 0.0019871287513524294
step: 260, loss: 0.0028307815082371235
step: 270, loss: 0.0011756251333281398
step: 280, loss: 0.00020629620121326298
step: 290, loss: 0.00012292570318095386
step: 300, loss: 0.0005088479374535382
step: 310, loss: 0.0009560619364492595
step: 320, loss: 0.0004425456572789699
step: 330, loss: 0.001917450106702745
step: 340, loss: 0.0005201316671445966
step: 350, loss: 0.00013782037422060966
step: 360, loss: 0.00011755935702240095
step: 370, loss: 0.07307565957307816
step: 380, loss: 0.0009464796748943627
step: 390, loss: 0.006490449886769056
step: 400, loss: 0.0028616918716579676
step: 410, loss: 0.00013481736823450774
step: 420, loss: 0.0008938032551668584
step: 430, loss: 0.012774602510035038
step: 440, loss: 5.5824020819272846e-05
step: 450, loss: 7.203660061350092e-05
step: 460, loss: 0.08575161546468735
step: 470, loss: 1.7221518646692857e-05
step: 480, loss: 0.0007638730457983911
step: 490, loss: 0.0012852151412516832
step: 500, loss: 0.0030509456992149353
step: 510, loss: 0.04613456875085831
step: 520, loss: 0.013317476958036423
step: 530, loss: 0.0016738693229854107
/usr/local/lib/python3.7/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=0.9517564402810305, f1=0.947417840375587,
best_f1=0.9469767441860464
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0004222436691634357

```

step: 10, loss: 0.004627940244972706  
step: 20, loss: 0.0005342259537428617  
step: 30, loss: 0.005306811071932316  
step: 40, loss: 8.27194526209496e-05  
step: 50, loss: 0.0066964286379516125  
step: 60, loss: 0.001435297541320324  
step: 70, loss: 0.00044237170368433  
step: 80, loss: 0.0006843424052931368  
step: 90, loss: 0.0030790637247264385  
step: 100, loss: 8.658464503241703e-05  
step: 110, loss: 0.027746178209781647  
step: 120, loss: 0.0001402635534759611  
step: 130, loss: 0.006967463530600071  
step: 140, loss: 5.143204907653853e-05  
step: 150, loss: 0.0015041421866044402  
step: 160, loss: 0.0004850762488786131  
step: 170, loss: 6.934606790309772e-05  
step: 180, loss: 0.0008032795740291476  
step: 190, loss: 0.0033397350925952196  
step: 200, loss: 0.000947357271797955  
step: 210, loss: 0.001360374386422336  
step: 220, loss: 0.0015230294084176421  
step: 230, loss: 0.17688605189323425  
step: 240, loss: 0.00010334992111893371  
step: 250, loss: 0.0005162360030226409  
step: 260, loss: 1.8905371689470485e-05  
step: 270, loss: 1.9162380340276286e-05  
step: 280, loss: 3.488634683890268e-05  
step: 290, loss: 0.000692423083819449  
step: 300, loss: 0.0006359717808663845  
step: 310, loss: 0.0005335371824912727  
step: 320, loss: 9.516167483525351e-05  
step: 330, loss: 0.0006620539352297783  
step: 340, loss: 0.0002591015072539449  
step: 350, loss: 0.005798682104796171  
step: 360, loss: 0.0005548717454075813  
step: 370, loss: 0.01261928305029869  
step: 380, loss: 0.001973051344975829  
step: 390, loss: 0.0003705014241859317  
step: 400, loss: 0.002709152176976204  
step: 410, loss: 0.00035543672856874764  
step: 420, loss: 0.0005356731126084924  
step: 430, loss: 1.7686741557554342e-05  
step: 440, loss: 0.001336113316938281  
step: 450, loss: 0.0013364125043153763  
step: 460, loss: 0.0006142942584119737  
step: 470, loss: 5.720312401535921e-05  
step: 480, loss: 0.001149865915067494

```

step: 490, loss: 0.002732646418735385
step: 500, loss: 0.0005100140697322786
step: 510, loss: 0.04236208647489548
step: 520, loss: 3.2743886549724266e-05
step: 530, loss: 0.0005925345467403531
/usr/local/lib/python3.7/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=0.9512308406874129, f1=0.9441340782122905,
best_f1=0.9469767441860464
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, 169.89it/s]
load_f1 = 0.9527777777777777
real_f1 = 0.9526022304832714
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, 142.49it/s]

```