DMedium_10_3_5_bert

August 13, 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/Snippext_public
     %cd Snippext_public
     !pip install -r requirements.txt
     %cd ..

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

```
Cloning into 'ditto'…
remote: Enumerating objects: 291, done.
remote: Total 291 (delta 0), reused 0 (delta 0), pack-reused 291
Receiving objects: 100% (291/291), 26.87 MiB | 16.91 MiB/s, done.
Resolving deltas: 100% (142/142), done.
Checking out files: 100% (128/128), 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 2.3 MB/s
Collecting numpy==1.19.2
  Downloading numpy-1.19.2-cp37-cp37m-manylinux2010_x86_64.whl (14.5 MB)
     |                      | 14.5 MB 48.5 MB/s
Collecting regex==2019.12.20
  Downloading regex-2019.12.20-cp37-cp37m-manylinux2010_x86_64.whl (689 kB)
     |                      | 689 kB 54.7 MB/s
Collecting scipy==1.3.2
  Downloading scipy-1.3.2-cp37-cp37m-manylinux1_x86_64.whl (25.2 MB)
     |                      | 25.2 MB 1.3 MB/s
Collecting sentencepiece==0.1.85
  Downloading sentencepiece-0.1.85-cp37-cp37m-manylinux1_x86_64.whl (1.0 MB)
     |                      | 1.0 MB 59.4 MB/s
Requirement already satisfied: sklearn==0.0 in
```

1

/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 47.3 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 | 7.97 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 2.7 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 64.7 MB/s
Collecting torch==1.4.0
  Downloading torch-1.4.0-cp37-cp37m-manylinux1_x86_64.whl (753.4 MB)
     | | 753.4 MB 7.1 kB/s
Collecting tqdm==4.41.0
  Downloading tqdm-4.41.0-py2.py3-none-any.whl (56 kB)
     | | 56 kB 4.6 MB/s
Collecting transformers==3.1.0
  Downloading transformers-3.1.0-py3-none-any.whl (884 kB)
     | | 884 kB 52.1 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 5.9 MB/s
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: 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: 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 plac<1.2.0,>=0.9.6
  Downloading plac-1.1.3-py2.py3-none-any.whl (20 kB)
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 catalogue<1.1.0,>=0.0.7
  Downloading catalogue-1.0.0-py2.py3-none-any.whl (7.7 kB)
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)
Requirement already satisfied: preshed<3.1.0,>=3.0.2 in
/usr/local/lib/python3.7/dist-packages (from spacy==2.2.3->-r requirements.txt
(line 4)) (3.0.6)
Collecting srsly<1.1.0,>=0.1.0
  Downloading srsly-1.0.5-cp37-cp37m-manylinux2014_x86_64.whl (184 kB)
     |                    | 184 kB 68.7 MB/s
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 46.3 MB/s
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 48.9 MB/s
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: 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)
Requirement already satisfied: packaging in /usr/local/lib/python3.7/dist-
packages (from transformers==3.1.0->-r requirements.txt (line 11)) (21.3)
Collecting tokenizers==0.8.1.rc2
```

```
    Downloading tokenizers-0.8.1rc2-cp37-cp37m-manylinux1_x86_64.whl (3.0 MB)
        |                              | 3.0 MB 53.0 MB/s
Collecting sacremoses
  Downloading sacremoses-0.0.53.tar.gz (880 kB)
        |                              | 880 kB 64.5 MB/s
Requirement already satisfied: filelock in /usr/local/lib/python3.7/dist-
packages (from transformers==3.1.0->-r requirements.txt (line 11)) (3.7.1)
Requirement already satisfied: importlib-metadata>=0.20 in
/usr/local/lib/python3.7/dist-packages (from
catalogue<1.1.0,>=0.0.7->spacy==2.2.3->-r requirements.txt (line 4)) (4.12.0)
Requirement already satisfied: 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: 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: 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: 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: pyparsing!=3.0.5,>=2.0.2 in
/usr/local/lib/python3.7/dist-packages (from packaging->transformers==3.1.0->-r
requirements.txt (line 11)) (3.0.9)
Requirement already satisfied: click in /usr/local/lib/python3.7/dist-packages
(from sacremoses->transformers==3.1.0->-r requirements.txt (line 11)) (7.1.2)
Requirement already satisfied: joblib in /usr/local/lib/python3.7/dist-packages
(from sacremoses->transformers==3.1.0->-r requirements.txt (line 11)) (1.1.0)
Requirement already satisfied: threadpoolctl>=2.0.0 in
/usr/local/lib/python3.7/dist-packages (from scikit-learn->sklearn==0.0->-r
requirements.txt (line 6)) (3.1.0)
Building wheels for collected packages: nltk, sacremoses
  Building wheel for nltk (setup.py) … done
  Created wheel for nltk: filename=nltk-3.4.5-py3-none-any.whl size=1449925
sha256=915453515f69141dcb9901c62d94585070b15aab1e0b51db24eb600edaf6a06f
  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=203f67017add305cd7cb634afb477e68b42dc21e16848ba78853e28c05d2a984
  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% (139/139), done.
remote: Total 10085 (delta 98), reused 130 (delta 58), pack-reused 9884
Receiving objects: 100% (10085/10085), 14.95 MiB | 25.43 MiB/s, done.
Resolving deltas: 100% (6903/6903), 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 writeable
Created temporary directory: /tmp/pip-ephem-wheel-cache-8rwq7uf0
Created temporary directory: /tmp/pip-req-tracker-qhoyqkj4
Initialized build tracking at /tmp/pip-req-tracker-qhoyqkj4
Created build tracker: /tmp/pip-req-tracker-qhoyqkj4
Entered build tracker: /tmp/pip-req-tracker-qhoyqkj4
```

```
Created temporary directory: /tmp/pip-install-ou7ne1rg
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-4im6w0t0
  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-
qhoyqkj4'
    Running setup.py (path:/tmp/pip-req-build-4im6w0t0/setup.py) egg_info for
package from file:///content/ditto/apex
    Created temporary directory: /tmp/pip-pip-egg-info-8ac51vf1
    Running command python setup.py egg_info


    torch.__version__  = 1.4.0


    running egg_info
    creating /tmp/pip-pip-egg-info-8ac51vf1/apex.egg-info
    writing /tmp/pip-pip-egg-info-8ac51vf1/apex.egg-info/PKG-INFO
    writing dependency_links to /tmp/pip-pip-egg-info-8ac51vf1/apex.egg-
info/dependency_links.txt
    writing top-level names to /tmp/pip-pip-egg-info-8ac51vf1/apex.egg-
info/top_level.txt
    writing manifest file '/tmp/pip-pip-egg-info-8ac51vf1/apex.egg-
info/SOURCES.txt'
    adding license file 'LICENSE'
    writing manifest file '/tmp/pip-pip-egg-info-8ac51vf1/apex.egg-
info/SOURCES.txt'
  Source in /tmp/pip-req-build-4im6w0t0 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-qhoyqkj4'
Created temporary directory: /tmp/pip-unpack-qhqreldo
Building wheels for collected packages: apex
  Created temporary directory: /tmp/pip-wheel-jwo8dnkb
  Building wheel for apex (setup.py) …   Destination directory:
/tmp/pip-wheel-jwo8dnkb
  Running command /usr/bin/python3 -u -c 'import io, os, sys, setuptools,
tokenize; sys.argv[0] = '"'"'/tmp/pip-req-build-4im6w0t0/setup.py'"'"';
```

```
__file__='"'"'/tmp/pip-req-build-4im6w0t0/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-jwo8dnkb


  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/_batchsampler.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/_batchsampler.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_pipelin
ing_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_pipelin
ing_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-jwo8dnkb/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/_batchsampler.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=74f9761f72b0edf794838a9a21862337102c92f14094d6a917b68020fde0f2aa
  Stored in directory: /tmp/pip-ephem-wheel-cache-8rwq7uf0/wheels/6b/22/a2/16655
26ee3c3061243260fd522525eca31398e04bfa5ad7e1f
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-qhoyqkj4'
/content/ditto

```python
# 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 2.7 MB/s
Collecting awscli
  Downloading awscli-1.25.47-py3-none-any.whl (3.9 MB)
     | | 3.9 MB 20.6 MB/s
Collecting rsa<4.8,>=3.1.2
  Downloading rsa-4.7.2-py3-none-any.whl (34 kB)
Requirement already satisfied: PyYAML<5.5,>=3.10 in
/usr/local/lib/python3.7/dist-packages (from awscli) (3.13)
Collecting s3transfer<0.7.0,>=0.6.0
  Downloading s3transfer-0.6.0-py3-none-any.whl (79 kB)
     | | 79 kB 6.4 MB/s
Collecting colorama<0.4.5,>=0.2.5
  Downloading colorama-0.4.4-py2.py3-none-any.whl (16 kB)
Collecting botocore==1.27.47
  Downloading botocore-1.27.47-py3-none-any.whl (9.0 MB)
     | | 9.0 MB 35.3 MB/s
Collecting docutils<0.17,>=0.10
  Downloading docutils-0.16-py2.py3-none-any.whl (548 kB)
     | | 548 kB 45.0 MB/s
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)
Collecting jmespath<2.0.0,>=0.7.1
  Downloading jmespath-1.0.1-py3-none-any.whl (20 kB)
Requirement already satisfied: six>=1.5 in /usr/local/lib/python3.7/dist-
packages (from python-dateutil<3.0.0,>=2.1->botocore==1.27.47->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: numpy in /usr/local/lib/python3.7/dist-
packages (from torch==1.7.1) (1.19.2)
Requirement already satisfied: typing-extensions in
/usr/local/lib/python3.7/dist-packages (from torch==1.7.1) (4.1.1)
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 | 19.20 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/DMedium_10_3_5/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 bert-base-uncased \
        --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 bert-base-uncased \
        --use_gpu \
        --fp16 \
        --checkpoint_path checkpoints/
```

```
Downloading: 100% 433/433 [00:00<00:00, 422kB/s]
Downloading: 100% 232k/232k [00:00<00:00, 2.52MB/s]
Downloading: 100% 440M/440M [00:08<00:00, 49.3MB/s]
Selected optimization level O2:  FP16 training with FP32 batchnorm and FP32
master weights.

Defaults for this optimization level are:
enabled                : True
opt_level              : O2
```

```
cast_model_type        : torch.float16
patch_torch_functions  : False
keep_batchnorm_fp32    : True
master_weights         : True
loss_scale             : dynamic
Processing user overrides (additional kwargs that are not None)…
After processing overrides, optimization options are:
enabled                : True
opt_level              : O2
cast_model_type        : torch.float16
patch_torch_functions  : False
keep_batchnorm_fp32    : True
master_weights         : True
loss_scale             : dynamic
Warning:  multi_tensor_applier fused unscale kernel is unavailable, possibly
because apex was installed without --cuda_ext --cpp_ext. Using Python fallback.
Original ImportError was: ModuleNotFoundError("No module named 'amp_C'")
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
Gradient overflow.  Skipping step, loss scaler 0 reducing loss scale to 32768.0
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:136:
UserWarning: Detected call of `lr_scheduler.step()` before `optimizer.step()`.
In PyTorch 1.1.0 and later, you should call them in the opposite order:
`optimizer.step()` before `lr_scheduler.step()`.  Failure to do this will result
in PyTorch skipping the first value of the learning rate schedule. See more
details at https://pytorch.org/docs/stable/optim.html#how-to-adjust-learning-
rate
  "https://pytorch.org/docs/stable/optim.html#how-to-adjust-learning-rate",
UserWarning)
step: 0, loss: 0.5640740990638733
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.4782608695652174, f1=0.4285714285714286,
best_f1=0.4285714285714286
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.5096163153648376
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
```

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=0.5263157894736842, f1=0.36363636363636365,
best_f1=0.36363636363636365
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.5595760941505432
/usr/local/lib/python3.7/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=0.5, f1=0.5925925925925927, best_f1=0.36363636363636365
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.2257394641637802
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.56, f1=0.5714285714285714, best_f1=0.5714285714285714
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.25573891401290894
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.7999999999999999, f1=0.7333333333333334,
best_f1=0.7333333333333334
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.32583025097846985
/usr/local/lib/python3.7/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=0.7741935483870968, f1=0.8275862068965518,
best_f1=0.7333333333333334
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:

```
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.12048496305942535
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.888888888888889, f1=0.8275862068965518,
best_f1=0.8275862068965518
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.08019805699586868
/usr/local/lib/python3.7/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=0.7333333333333334, f1=0.8571428571428571,
best_f1=0.8275862068965518
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.004988952074199915
/usr/local/lib/python3.7/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=0.8571428571428571, f1=0.8461538461538461,
best_f1=0.8275862068965518
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.016349472105503082
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9032258064516129, f1=0.8484848484848484,
best_f1=0.8484848484848484
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.006149995140731335
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
```

```
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
epoch 11: dev_f1=0.896551724137931, f1=0.8275862068965518,
best_f1=0.8484848484848484
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.006262524519115686
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
epoch 12: dev_f1=0.8750000000000001, f1=0.8, best_f1=0.8484848484848484
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.008986786007881165
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
epoch 13: dev_f1=0.8750000000000001, f1=0.8, best_f1=0.8484848484848484
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.005184875335544348
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
epoch 14: dev_f1=0.8750000000000001, f1=0.8, best_f1=0.8484848484848484
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.004854745231568813
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
epoch 15: dev_f1=0.8750000000000001, f1=0.8, best_f1=0.8484848484848484
mv: cannot stat '*_dev.pt': No such file or directory
Selected optimization level O2:  FP16 training with FP32 batchnorm and FP32
master weights.

Defaults for this optimization level are:
enabled                : True
opt_level              : 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.")
91it [00:00, 112723.47it/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 = 0.7777777777777778
real_f1 = 0.7777777777777778
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 [00:15, 278.04it/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 bert-base-uncased \
        --fp16 \
        --save_model \
        --da entry_swap  \
        --da span_shuffle \
        --summarize
```

```python
# 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 bert-base-uncased \
  --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: 0.6228980422019958
Gradient overflow.  Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 10, loss: 0.5863699913024902
step: 20, loss: 0.30607908964157104
step: 30, loss: 0.1094709187746048
step: 40, loss: 0.3555612564086914
step: 50, loss: 0.02586987242102623
step: 60, loss: 0.04791754111647606
step: 70, loss: 0.013000697828829288
step: 80, loss: 0.07372020184993744
step: 90, loss: 0.16054457426071167
step: 100, loss: 0.005453249905258417
step: 110, loss: 0.1510896533727646
step: 120, loss: 0.015591434203088284
step: 130, loss: 0.011086992919445038
step: 140, loss: 0.0047528063878417015
step: 150, loss: 0.05179345980286598
step: 160, loss: 0.02156897820532322
step: 170, loss: 0.07825727760791779
step: 180, loss: 0.04700244590640068
step: 190, loss: 0.021764077246189117
step: 200, loss: 0.010646590031683445
step: 210, loss: 0.0022854504641145468
step: 220, loss: 0.0028285605367273092
step: 230, loss: 0.07367614656686783
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9808773903262092, f1=0.9761092150170648,
best_f1=0.9761092150170648
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.011377292685210705
step: 10, loss: 0.0011840376537293196
step: 20, loss: 0.10645114630460739
step: 30, loss: 0.17216558754444122
step: 40, loss: 0.05012442544102669
step: 50, loss: 0.01083909161388874
step: 60, loss: 0.004339596722275019
step: 70, loss: 0.25144267082214355
step: 80, loss: 0.04766112193465233
```

```
step: 90, loss: 0.015946106985211372
step: 100, loss: 0.04882505163550377
step: 110, loss: 0.004094211850315332
step: 120, loss: 0.002087813103571534
step: 130, loss: 0.10734773427248001
step: 140, loss: 0.003047225996851921
step: 150, loss: 0.03866511583328247
step: 160, loss: 0.005615469068288803
step: 170, loss: 0.0017206012271344662
step: 180, loss: 0.002817912260070443
step: 190, loss: 0.002804357325658202
step: 200, loss: 0.006099838763475418
step: 210, loss: 0.0009519438608549535
step: 220, loss: 0.04516240581870079
step: 230, loss: 0.0017594879027456045
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9887387387387387, f1=0.9809203142536477,
best_f1=0.9809203142536477
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.010605240240693092
step: 10, loss: 0.07387708127498627
step: 20, loss: 0.00078320357715711
step: 30, loss: 0.006043223664164543
step: 40, loss: 0.14414824545383453
step: 50, loss: 0.005000881850719452
step: 60, loss: 0.0025397534482181072
step: 70, loss: 0.0048161037266254425
step: 80, loss: 0.00821332074701786
step: 90, loss: 0.08766929060220718
step: 100, loss: 0.002490755869075656
step: 110, loss: 0.003835116745904088
step: 120, loss: 0.0039279223419725895
step: 130, loss: 0.0007320567383430898
step: 140, loss: 0.010188527405261993
step: 150, loss: 0.022683117538690567
step: 160, loss: 0.02370462194085121
step: 170, loss: 0.00730894086882472
step: 180, loss: 0.03371215984225273
step: 190, loss: 0.004540520720183849
step: 200, loss: 0.006955225020647049
```

```
step: 210, loss: 0.0024727811105549335
step: 220, loss: 0.0006486361380666494
step: 230, loss: 0.0006984625360928476
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9898762654668166, f1=0.9819004524886877,
best_f1=0.9819004524886877
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0015741082606837153
step: 10, loss: 0.0004965653060935438
step: 20, loss: 0.0005792625015601516
step: 30, loss: 0.0007146396674215794
step: 40, loss: 0.11712441593408585
step: 50, loss: 0.0009360643452964723
step: 60, loss: 0.002142357872799039
step: 70, loss: 0.00047619215911254287
step: 80, loss: 0.0021774331107735634
step: 90, loss: 0.0012047667987644672
step: 100, loss: 0.0006502058240585029
step: 110, loss: 0.0014660254819318652
step: 120, loss: 0.06659480929374695
step: 130, loss: 0.0178818441927433
step: 140, loss: 0.021686818450689316
step: 150, loss: 0.11093389987945557
step: 160, loss: 0.05681394785642624
step: 170, loss: 0.012401141226291656
step: 180, loss: 0.000998098636046052
step: 190, loss: 0.005468618590384722
step: 200, loss: 0.0018905228935182095
step: 210, loss: 0.16360729932785034
step: 220, loss: 0.0005826420383527875
step: 230, loss: 0.08657993376255035
/usr/local/lib/python3.7/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=0.9853768278965129, f1=0.979591836734694,
best_f1=0.9819004524886877
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0007089207065291703
```

```
step: 10, loss: 0.0005815931945107877
step: 20, loss: 0.0019848623778671026
step: 30, loss: 0.0003410937206353992
step: 40, loss: 0.0007139823865145445
step: 50, loss: 0.0013432074338197708
step: 60, loss: 0.015767551958560944
step: 70, loss: 0.0018171954434365034
step: 80, loss: 0.0006604054942727089
step: 90, loss: 0.0016027081292122602
step: 100, loss: 0.015156090259552002
step: 110, loss: 0.00031482859048992395
step: 120, loss: 0.00014477409422397614
step: 130, loss: 0.0011760396882891655
step: 140, loss: 0.003026468912139535
step: 150, loss: 0.0019346874905750155
step: 160, loss: 0.000435057794675231
step: 170, loss: 0.008343076333403587
step: 180, loss: 0.0007998895598575473
step: 190, loss: 0.13563895225524902
step: 200, loss: 0.0036242324858903885
step: 210, loss: 0.0011066069127991796
step: 220, loss: 0.0031985114328563213
step: 230, loss: 0.0005012442707084119
/usr/local/lib/python3.7/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=0.9876543209876544, f1=0.9796839729119639,
best_f1=0.9819004524886877
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0005576357361860573
step: 10, loss: 0.0005351788713596761
step: 20, loss: 0.002163889352232218
step: 30, loss: 0.0005643994663842022
step: 40, loss: 0.0063793552108109
step: 50, loss: 0.0007198374369181693
step: 60, loss: 0.00026824252563528717
step: 70, loss: 0.0013319378485903144
step: 80, loss: 0.0008599652210250497
step: 90, loss: 0.0003324495046399534
step: 100, loss: 0.040858104825019836
step: 110, loss: 0.00034865227644331753
step: 120, loss: 0.0002696793235372752
step: 130, loss: 0.0020672152750194073
step: 140, loss: 0.003127743722870946
step: 150, loss: 0.00021218118490065125
step: 160, loss: 0.002769381506368518
```

```
step: 170, loss: 0.0002846438728738576
step: 180, loss: 0.10501813143491745
step: 190, loss: 0.002540266839787364
step: 200, loss: 0.007461454253643751
step: 210, loss: 0.0010275250533595681
step: 220, loss: 0.0003874943940900266
step: 230, loss: 0.05343835800886154
/usr/local/lib/python3.7/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=0.9865168539325843, f1=0.9784335981838819,
best_f1=0.9819004524886877
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.004498402122408152
step: 10, loss: 0.0015145970974117517
step: 20, loss: 0.00022525583335664123
step: 30, loss: 0.0002746133250184357
step: 40, loss: 0.00016437664453405887
step: 50, loss: 0.0002633488620631397
step: 60, loss: 0.009661387652158737
step: 70, loss: 0.020703161135315895
step: 80, loss: 0.00013684276200365275
step: 90, loss: 8.224636258091778e-05
step: 100, loss: 0.00013390240201260895
step: 110, loss: 0.0002596180129330605
step: 120, loss: 0.00018831485067494214
step: 130, loss: 0.0006743965204805136
step: 140, loss: 0.00020107897580601275
step: 150, loss: 0.006872196681797504
step: 160, loss: 0.07121862471103668
step: 170, loss: 0.004555300809442997
step: 180, loss: 0.0023133528884500265
step: 190, loss: 0.00020077158114872873
step: 200, loss: 0.04413405805826187
step: 210, loss: 7.450749399140477e-05
step: 220, loss: 0.008629337884485722
step: 230, loss: 0.005541052203625441
/usr/local/lib/python3.7/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=0.9887387387387387, f1=0.9875706214689265,
best_f1=0.9819004524886877
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00015455052198376507
```

```
step: 10, loss: 0.000242917041759938
step: 20, loss: 8.816029730951414e-05
step: 30, loss: 0.0009819166734814644
step: 40, loss: 0.00025413575349375606
step: 50, loss: 0.0023024824913591146
step: 60, loss: 0.00011177505075465888
step: 70, loss: 0.0001902145304484293
step: 80, loss: 0.0003281212120782584
step: 90, loss: 4.707573680207133e-05
step: 100, loss: 0.0001813489943742752
step: 110, loss: 0.005771099589765072
step: 120, loss: 0.00015660320059396327
step: 130, loss: 0.0006481692544184625
step: 140, loss: 7.078477938193828e-05
step: 150, loss: 0.00013676160597242415
step: 160, loss: 0.0001546582643641159
step: 170, loss: 7.767870556563139e-05
step: 180, loss: 7.74846485001035e-05
step: 190, loss: 8.374649769393727e-05
step: 200, loss: 0.00019476505985949188
step: 210, loss: 0.00011158034210675396
step: 220, loss: 0.00036972720408812165
step: 230, loss: 0.0007705431198701262
/usr/local/lib/python3.7/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=0.9864864864864865, f1=0.9830124575311437,
best_f1=0.9819004524886877
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 9.337574010714889e-05
step: 10, loss: 0.001028841477818787
step: 20, loss: 0.0002316754253115505
step: 30, loss: 5.8120200264966115e-05
step: 40, loss: 0.07008083909749985
step: 50, loss: 6.0914462665095925e-05
step: 60, loss: 7.242970605148003e-05
step: 70, loss: 0.00019373033137526363
step: 80, loss: 6.498175935121253e-05
step: 90, loss: 7.786645437590778e-05
step: 100, loss: 0.00030130412778817117
step: 110, loss: 4.996409188606776e-05
step: 120, loss: 4.2790699808392674e-05
step: 130, loss: 8.721372432773933e-05
step: 140, loss: 4.358742808108218e-05
step: 150, loss: 0.0008197219576686621
step: 160, loss: 0.00010616983490763232
```

```
step: 170, loss: 0.00021400328841991723
step: 180, loss: 0.003443666733801365
step: 190, loss: 0.00020360626513138413
step: 200, loss: 6.101193866925314e-05
step: 210, loss: 0.0014394031604751945
step: 220, loss: 5.876241993973963e-05
step: 230, loss: 0.0001449597766622901
/usr/local/lib/python3.7/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=0.987598647125141, f1=0.9807037457434733,
best_f1=0.9819004524886877
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00012863671872764826
step: 10, loss: 0.0001271922083105892
step: 20, loss: 0.00039408382144756615
step: 30, loss: 9.804198634810746e-05
step: 40, loss: 6.662847590632737e-05
step: 50, loss: 0.00019493498257361352
step: 60, loss: 0.0005249571986496449
step: 70, loss: 0.00023090967442840338
step: 80, loss: 6.589612166862935e-05
step: 90, loss: 0.00021688378183171153
step: 100, loss: 6.994591967668384e-05
step: 110, loss: 5.671608232660219e-05
step: 120, loss: 0.00013184110866859555
step: 130, loss: 6.306399154709652e-05
step: 140, loss: 0.026478173211216927
step: 150, loss: 0.015274678356945515
step: 160, loss: 2.8855340133304708e-05
step: 170, loss: 7.933800225146115e-05
step: 180, loss: 0.00010248775652144104
step: 190, loss: 0.033564191311597824
step: 200, loss: 0.00019520451314747334
step: 210, loss: 7.315121911233291e-05
step: 220, loss: 5.5894724937388673e-05
step: 230, loss: 8.545270247850567e-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=0.987598647125141, f1=0.9818181818181818,
best_f1=0.9819004524886877
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 8.794292079983279e-05
```

```
step: 10, loss: 6.07870751991868e-05
step: 20, loss: 0.0023742879275232553
step: 30, loss: 0.020979158580303192
step: 40, loss: 6.362090061884373e-05
step: 50, loss: 0.00011377176269888878
step: 60, loss: 8.013677143026143e-05
step: 70, loss: 9.556925215292722e-05
step: 80, loss: 4.305018956074491e-05
step: 90, loss: 3.9939946873346344e-05
step: 100, loss: 4.580301174428314e-05
step: 110, loss: 0.012285598553717136
step: 120, loss: 2.9156992241041735e-05
step: 130, loss: 3.540876059560105e-05
step: 140, loss: 5.396997948992066e-05
step: 150, loss: 0.03156280145049095
step: 160, loss: 3.692729296744801e-05
step: 170, loss: 0.024380775168538094
step: 180, loss: 4.450655251275748e-05
step: 190, loss: 4.729364809463732e-05
step: 200, loss: 0.0005974808591417968
step: 210, loss: 2.7365293135517277e-05
step: 220, loss: 5.19970380992163e-05
step: 230, loss: 9.147506352746859e-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 11: dev_f1=0.9887387387387387, f1=0.9807037457434733,
best_f1=0.9819004524886877
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 4.0905753849074244e-05
step: 10, loss: 3.7989066186128184e-05
step: 20, loss: 3.6617777368519455e-05
step: 30, loss: 0.00017392009613104165
step: 40, loss: 0.00015063695900607854
step: 50, loss: 0.00010349326475989074
step: 60, loss: 0.0001330486120423302
step: 70, loss: 5.1678973250091076e-05
step: 80, loss: 8.281988993985578e-05
step: 90, loss: 3.450248914305121e-05
step: 100, loss: 2.598698847577907e-05
step: 110, loss: 3.962780465371907e-05
step: 120, loss: 3.736672078957781e-05
step: 130, loss: 3.91217581636738e-05
step: 140, loss: 0.017129182815551758
step: 150, loss: 0.00012604934454429895
step: 160, loss: 3.472231037449092e-05
```

```
step: 170, loss: 3.9508871850557625e-05
step: 180, loss: 0.021861644461750984
step: 190, loss: 0.00010258064139634371
step: 200, loss: 2.1751502572442405e-05
step: 210, loss: 3.4025673812720925e-05
step: 220, loss: 3.743345951079391e-05
step: 230, loss: 0.02278188243508339
/usr/local/lib/python3.7/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=0.9887387387387387, f1=0.9807037457434733,
best_f1=0.9819004524886877
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 9.53065391513519e-05
step: 10, loss: 9.440714347874746e-05
step: 20, loss: 7.388195808744058e-05
step: 30, loss: 4.8880010581342503e-05
step: 40, loss: 5.199880979489535e-05
step: 50, loss: 3.8436137401731685e-05
step: 60, loss: 4.274185994290747e-05
step: 70, loss: 4.235434971633367e-05
step: 80, loss: 8.503066055709496e-05
step: 90, loss: 4.052539588883519e-05
step: 100, loss: 3.7286372389644384e-05
step: 110, loss: 0.023902863264083862
step: 120, loss: 0.03346480056643486
step: 130, loss: 7.683052535867319e-05
step: 140, loss: 4.259293564246036e-05
step: 150, loss: 5.564148159464821e-05
step: 160, loss: 6.737699004588649e-05
step: 170, loss: 7.51329425838776e-05
step: 180, loss: 0.00010095291509060189
step: 190, loss: 0.00015187131066340953
step: 200, loss: 8.511371561326087e-05
step: 210, loss: 2.5104063752223738e-05
step: 220, loss: 6.273540202528238e-05
step: 230, loss: 0.0003457744314800948
/usr/local/lib/python3.7/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=0.987598647125141, f1=0.9807037457434733,
best_f1=0.9819004524886877
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 3.510568058118224e-05
```

```
step: 10, loss: 0.015102985315024853
step: 20, loss: 7.897005707491189e-05
step: 30, loss: 0.00011271794210188091
step: 40, loss: 0.0002492136263754219
step: 50, loss: 4.639626058633439e-05
step: 60, loss: 5.198983126319945e-05
step: 70, loss: 3.875985566992313e-05
step: 80, loss: 4.036937389173545e-05
step: 90, loss: 5.72421595279593e-05
step: 100, loss: 3.987680975114927e-05
step: 110, loss: 0.00010532675514696166
step: 120, loss: 2.420635792077519e-05
step: 130, loss: 4.38632741861511e-05
step: 140, loss: 0.00016585046250838786
step: 150, loss: 4.641794294002466e-05
step: 160, loss: 0.004285008180886507
step: 170, loss: 2.4955177650554106e-05
step: 180, loss: 0.00017087427841033787
step: 190, loss: 3.88455628126394e-05
step: 200, loss: 3.985387593274936e-05
step: 210, loss: 6.056418351363391e-05
step: 220, loss: 6.788458995288238e-05
step: 230, loss: 9.42218757700175e-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=0.987598647125141, f1=0.9807037457434733,
best_f1=0.9819004524886877
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 2.871711694751866e-05
step: 10, loss: 2.236241562059149e-05
step: 20, loss: 5.321148273651488e-05
step: 30, loss: 8.611248631495982e-05
step: 40, loss: 8.011708268895745e-05
step: 50, loss: 0.048464108258485794
step: 60, loss: 0.00011985125456703827
step: 70, loss: 0.0002250395336886868
step: 80, loss: 4.478227856452577e-05
step: 90, loss: 5.2095023420406505e-05
step: 100, loss: 5.3555799240712076e-05
step: 110, loss: 3.7138313928153366e-05
step: 120, loss: 6.805254088249058e-05
step: 130, loss: 0.024560105055570602
step: 140, loss: 2.8251866751816124e-05
step: 150, loss: 0.00011864562111441046
step: 160, loss: 3.314972491352819e-05
```

```
step: 170, loss: 3.010279397130944e-05
step: 180, loss: 0.00011744382936740294
step: 190, loss: 8.522695861756802e-05
step: 200, loss: 5.297021198202856e-05
step: 210, loss: 7.309284410439432e-05
step: 220, loss: 7.567807915620506e-05
step: 230, loss: 4.058438935317099e-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=0.9886877828054299, f1=0.9807037457434733,
best_f1=0.9819004524886877
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.")
2473it [00:11, 211.82it/s]
load_f1 = 0.9876819708846584
real_f1 = 0.9887640449438202
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 [00:16, 261.10it/s]
```

1.1.3 DBLP-GoogleScholar - Running the matcher

```python
# 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 bert-base-uncased \
  --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 bert-base-uncased \
  --use_gpu \
  --fp16 \
  --checkpoint_path checkpoints/
```

```
Downloading: 100% 433/433 [00:00<00:00, 370kB/s]
Downloading: 100% 232k/232k [00:00<00:00, 256kB/s]
Downloading: 100% 440M/440M [00:07<00:00, 60.7MB/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
```

47

```
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: 0.6126914024353027
step: 10, loss: 0.54026776552200032
step: 20, loss: 0.468113511800766
step: 30, loss: 0.04900440573692322
step: 40, loss: 0.10789372026920319
step: 50, loss: 0.18530017137527466
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 60, loss: 0.03122362308204174
step: 70, loss: 0.1570146530866623
step: 80, loss: 0.0830511674284935
step: 90, loss: 0.23337377607822418
step: 100, loss: 0.02696705237030983
step: 110, loss: 0.06327388435602188
step: 120, loss: 0.05342787131667137
step: 130, loss: 0.10768258571624756
step: 140, loss: 0.03528009355068207
step: 150, loss: 0.031661272048950195
step: 160, loss: 0.016096582636237144
step: 170, loss: 0.1347726583480835
step: 180, loss: 0.1444343626499176
```

```
step: 190, loss: 0.01397034153342247
step: 200, loss: 0.13563118875026703
step: 210, loss: 0.18261854350566864
step: 220, loss: 0.23281723260879517
step: 230, loss: 0.14403918385505676
step: 240, loss: 0.061004023998975754
step: 250, loss: 0.023994604125618935
step: 260, loss: 0.0384732224047184
step: 270, loss: 0.030562521889805794
step: 280, loss: 0.06900054961442947
step: 290, loss: 0.21588750183582306
step: 300, loss: 0.05907373130321503
step: 310, loss: 0.3186620771884918
step: 320, loss: 0.10006927698850632
step: 330, loss: 0.026922035962343216
step: 340, loss: 0.03725241869688034
step: 350, loss: 0.010210404172539711
step: 360, loss: 0.0167580246925354
step: 370, loss: 0.037335287779569626
step: 380, loss: 0.014531533233821392
step: 390, loss: 0.0943903774023056
step: 400, loss: 0.19441036880016327
step: 410, loss: 0.06568671762943268
step: 420, loss: 0.018938174471259117
step: 430, loss: 0.16228154301643372
step: 440, loss: 0.019714433699846268
step: 450, loss: 0.0034755882807075977
step: 460, loss: 0.004007184877991676
step: 470, loss: 0.13278087973594666
step: 480, loss: 0.026872573420405388
step: 490, loss: 0.04736391082406044
step: 500, loss: 0.05719708651304245
step: 510, loss: 0.07511918991804123
step: 520, loss: 0.03014962747693062
step: 530, loss: 0.0027137678116559982
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9290681502086231, f1=0.9308755760368663,
best_f1=0.9308755760368663
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.16470134258270264
```

```
step: 10, loss: 0.10433731228113174
step: 20, loss: 0.00901168305426836
step: 30, loss: 0.004355961456894875
step: 40, loss: 0.05468469113111496
step: 50, loss: 0.12420389801263809
step: 60, loss: 0.02467755787074566
step: 70, loss: 0.03254658356308937
step: 80, loss: 0.06583093106746674
step: 90, loss: 0.009970592334866524
step: 100, loss: 0.005523313768208027
step: 110, loss: 0.0066283163614571095
step: 120, loss: 0.1602526158094406
step: 130, loss: 0.13755853474140167
step: 140, loss: 0.00808892771601677
step: 150, loss: 0.05775017663836479
step: 160, loss: 0.008292906917631626
step: 170, loss: 0.011437583714723587
step: 180, loss: 0.030482245609164238
step: 190, loss: 0.05791756138205528
step: 200, loss: 0.019875988364219666
step: 210, loss: 0.06949193775653839
step: 220, loss: 0.13286960124969482
step: 230, loss: 0.018204577267169952
step: 240, loss: 0.11544477194547653
step: 250, loss: 0.048200685530900955
step: 260, loss: 0.005207365844398737
step: 270, loss: 0.20782440900802612
step: 280, loss: 0.02910631336271763
step: 290, loss: 0.025183653458952904
step: 300, loss: 0.1516650915145874
step: 310, loss: 0.016077708452939987
step: 320, loss: 0.12677335739135742
step: 330, loss: 0.08828000724315643
step: 340, loss: 0.006435933988541365
step: 350, loss: 0.0029837889596819878
step: 360, loss: 0.10500706732273102
step: 370, loss: 0.14313173294067383
step: 380, loss: 0.07010339945554733
step: 390, loss: 0.04659273475408554
step: 400, loss: 0.04559575766324997
step: 410, loss: 0.007899819873273373
step: 420, loss: 0.045738041400909424
step: 430, loss: 0.010459942743182182
step: 440, loss: 0.29136115312576294
step: 450, loss: 0.06278116255998611
step: 460, loss: 0.008417867124080658
step: 470, loss: 0.048007313162088394
step: 480, loss: 0.23535068333148956
```

```
step: 490, loss: 0.022634508088231087
step: 500, loss: 0.22416643798351288
step: 510, loss: 0.0141258854418993
step: 520, loss: 0.044912390410900116
step: 530, loss: 0.0026600584387779236
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9427516158818098, f1=0.9358974358974358,
best_f1=0.9358974358974358
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.19042807817459106
step: 10, loss: 0.10025180876255035
step: 20, loss: 0.07881336659193039
step: 30, loss: 0.11123448610305786
step: 40, loss: 0.014192407950758934
step: 50, loss: 0.10941998660564423
step: 60, loss: 0.05134594812989235
step: 70, loss: 0.002770595485344529
step: 80, loss: 0.002983263460919261
step: 90, loss: 0.013311812654137611
step: 100, loss: 0.021001363173127174
step: 110, loss: 0.001544963219203055
step: 120, loss: 0.0023753503337750248
step: 130, loss: 0.005271228961646557
step: 140, loss: 0.07630261778831482
step: 150, loss: 0.01743313856422901
step: 160, loss: 0.005410072859376669
step: 170, loss: 0.07347887009382248
step: 180, loss: 0.007170042954385281
step: 190, loss: 0.006059164647012949
step: 200, loss: 0.004919389262795448
step: 210, loss: 0.0854024812579155
step: 220, loss: 0.12235096096992493
step: 230, loss: 0.19792912900447845
step: 240, loss: 0.009513192810118198
step: 250, loss: 0.0349949486553669
step: 260, loss: 0.01756730116903782
step: 270, loss: 0.009405415505170822
step: 280, loss: 0.20107115805149078
step: 290, loss: 0.002050023525953293
step: 300, loss: 0.0289093516767025
```

```
step: 310, loss: 0.010493369773030281
step: 320, loss: 0.006351838819682598
step: 330, loss: 0.0011608557542786002
step: 340, loss: 0.00307755870744586
step: 350, loss: 0.0020633742678910494
step: 360, loss: 0.027407029643654823
step: 370, loss: 0.0206987876445055
step: 380, loss: 0.046781327575445175
step: 390, loss: 0.005782168358564377
step: 400, loss: 0.05122540518641472
step: 410, loss: 0.001871950109489262
step: 420, loss: 0.20396870374679565
step: 430, loss: 0.01149191614240408
step: 440, loss: 0.004463013727217913
step: 450, loss: 0.04572093486785889
step: 460, loss: 0.03594893962144852
step: 470, loss: 0.12075477093458176
step: 480, loss: 0.003158523002639413
step: 490, loss: 0.007328635547310114
step: 500, loss: 0.0042362292229718447
step: 510, loss: 0.00239096162840724
step: 520, loss: 0.15758922696113586
step: 530, loss: 0.022772926837205887
/usr/local/lib/python3.7/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=0.9424964936886395, f1=0.9298653042266605,
best_f1=0.9358974358974358
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00439460389316082
step: 10, loss: 0.006696674972772598
step: 20, loss: 0.004430629778653383
step: 30, loss: 0.0034219662193208933
step: 40, loss: 0.008619488216936588
step: 50, loss: 0.004716451279819012
step: 60, loss: 0.001704710186459124
step: 70, loss: 0.0012918857391923666
step: 80, loss: 0.026410652324557304
step: 90, loss: 0.04263043776154518
step: 100, loss: 0.011339865624904633
step: 110, loss: 0.035997211933135986
step: 120, loss: 0.0003051131498068571
step: 130, loss: 0.005666709039360285
step: 140, loss: 0.004689618945121765
step: 150, loss: 0.0013733998639509082
step: 160, loss: 0.0017052431358024478
```

```
step: 170, loss: 0.0005983520532026887
step: 180, loss: 0.028864873573184013
step: 190, loss: 0.049972668290138245
step: 200, loss: 0.007843556813895702
step: 210, loss: 0.02965567074716091
step: 220, loss: 0.0022311238572001457
step: 230, loss: 0.05489347130060196
step: 240, loss: 0.0027021223213523626
step: 250, loss: 0.047352537512779236
step: 260, loss: 0.0015182873466983438
step: 270, loss: 0.01174167636781931
step: 280, loss: 0.032747361809015274
step: 290, loss: 0.023130401968955994
step: 300, loss: 0.00114855554420501
step: 310, loss: 0.002995306858792901
step: 320, loss: 0.0007468751864507794
step: 330, loss: 0.0024861767888069153
step: 340, loss: 0.013651836663484573
step: 350, loss: 0.08907318860292435
step: 360, loss: 0.07902108877897263
step: 370, loss: 0.009764429181814194
step: 380, loss: 0.003113477723672986
step: 390, loss: 0.0003579351759981364
step: 400, loss: 0.02502894215285778
step: 410, loss: 0.0003441376320552081
step: 420, loss: 0.010035932995378971
step: 430, loss: 0.04025642201304436
step: 440, loss: 0.018507469445466995
step: 450, loss: 0.012015493586659431
step: 460, loss: 0.06231765076518059
step: 470, loss: 0.0630878433585167
step: 480, loss: 0.0933866873383522
step: 490, loss: 0.04368791729211807
step: 500, loss: 0.015693917870521545
step: 510, loss: 0.006034651305526495
step: 520, loss: 0.022187838330864906
step: 530, loss: 0.020674560219049454
/usr/local/lib/python3.7/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=0.9417206290471785, f1=0.9385321100917432,
best_f1=0.9358974358974358
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.027133313938975334
step: 10, loss: 0.0501541793346405
step: 20, loss: 0.014446211978793144
```

```
step: 30, loss: 0.014249171130359173
step: 40, loss: 0.0030805328860878944
step: 50, loss: 0.0022533147130161524
step: 60, loss: 0.028709422796964645
step: 70, loss: 0.0037158646155148745
step: 80, loss: 0.0012136220466345549
step: 90, loss: 0.001561089651659131
step: 100, loss: 0.023567968979477882
step: 110, loss: 0.00029149098554626107
step: 120, loss: 0.002623210661113262
step: 130, loss: 0.0016026379307731986
step: 140, loss: 0.0006484212353825569
step: 150, loss: 0.004117886070162058
step: 160, loss: 0.00048345542745664716
step: 170, loss: 0.014032945968210697
step: 180, loss: 0.0003970974648837
step: 190, loss: 0.007056704722344875
step: 200, loss: 0.012542922049760818
step: 210, loss: 0.021248135715723038
step: 220, loss: 0.0014264601049944758
step: 230, loss: 0.007729977369308472
step: 240, loss: 0.0030954862013459206
step: 250, loss: 0.002454508328810334
step: 260, loss: 0.007502687629312277
step: 270, loss: 0.0006018176209181547
step: 280, loss: 0.0007609904860146344
step: 290, loss: 0.031175540760159492
step: 300, loss: 0.00616080779582262
step: 310, loss: 0.0013811326352879405
step: 320, loss: 0.062115151435136795
step: 330, loss: 0.019157573580741882
step: 340, loss: 0.003741256892681122
step: 350, loss: 0.00561300665140152
step: 360, loss: 0.010348053649067879
step: 370, loss: 0.027382491156458855
step: 380, loss: 0.0019198922673240304
step: 390, loss: 0.0001955386105692014
step: 400, loss: 0.001548839383758045
step: 410, loss: 0.004681967664510012
step: 420, loss: 0.008733535185456276
step: 430, loss: 0.0008863236871547997
step: 440, loss: 0.06002369523048401
step: 450, loss: 0.011215785518288612
step: 460, loss: 0.0005613012472167611
step: 470, loss: 0.02108694612979889
step: 480, loss: 0.00033846774022094905
step: 490, loss: 0.0002605675545055419
step: 500, loss: 0.00107770471367985
```

```
step: 510, loss: 0.03559482470154762
step: 520, loss: 0.03853047266602516
step: 530, loss: 0.0005707246600650251
/usr/local/lib/python3.7/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=0.9339578454332552, f1=0.9286376274328081,
best_f1=0.9358974358974358
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00947534292936325
step: 10, loss: 0.004601853899657726
step: 20, loss: 0.06074771285057068
step: 30, loss: 0.0010695800883695483
step: 40, loss: 0.0006468745414167643
step: 50, loss: 0.033441949635744095
step: 60, loss: 0.0012812333879992366
step: 70, loss: 0.0009619094198569655
step: 80, loss: 0.004524452146142721
step: 90, loss: 0.005964800715446472
step: 100, loss: 0.0018746725982055068
step: 110, loss: 0.024855200201272964
step: 120, loss: 0.000998830539174378
step: 130, loss: 0.004614308010786772
step: 140, loss: 0.0019042800413444638
step: 150, loss: 0.0003633900487329811
step: 160, loss: 0.004733152221888304
step: 170, loss: 0.0008646278292872012
step: 180, loss: 0.005259213037788868
step: 190, loss: 0.0003430464712437242
step: 200, loss: 0.0050543793477118015
step: 210, loss: 0.0044583831913769245
step: 220, loss: 0.011813939549028873
step: 230, loss: 0.008972346782684326
step: 240, loss: 0.02715512178838253
step: 250, loss: 0.0003594631270971149
step: 260, loss: 0.0002840524248313159
step: 270, loss: 0.064564548432827
step: 280, loss: 0.0017310590483248234
step: 290, loss: 0.0005452170153148472
step: 300, loss: 0.021494299173355103
step: 310, loss: 0.009627155028283596
step: 320, loss: 0.0022081234492361546
step: 330, loss: 0.04672866687178612
step: 340, loss: 0.0820426344871521
step: 350, loss: 0.00475970096886158
step: 360, loss: 0.018384601920843124
```

```
step: 370, loss: 0.0009163650684058666
step: 380, loss: 0.00041851852438412607
step: 390, loss: 0.015254190191626549
step: 400, loss: 5.429358498076908e-05
step: 410, loss: 0.0002892209740821272
step: 420, loss: 0.0036695031449198723
step: 430, loss: 0.00016828252410050482
step: 440, loss: 0.00039342796662822366
step: 450, loss: 0.0005035824724473059
step: 460, loss: 0.00028734520310536027
step: 470, loss: 0.002348958281800151
step: 480, loss: 0.002258558291941881
step: 490, loss: 0.007482689339667559
step: 500, loss: 0.0032315445132553577
step: 510, loss: 0.0009451280348002911
step: 520, loss: 0.0012233753222972155
step: 530, loss: 0.09461146593093872
/usr/local/lib/python3.7/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=0.9348837209302325, f1=0.9293954776188279,
best_f1=0.9358974358974358
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0007734455284662545
step: 10, loss: 0.00023815319582354277
step: 20, loss: 0.0001328185317106545
step: 30, loss: 0.002177858492359519
step: 40, loss: 0.0005235529388301075
step: 50, loss: 0.0033789328299462795
step: 60, loss: 0.0005265395739115775
step: 70, loss: 0.03690299391746521
step: 80, loss: 0.10081176459789276
step: 90, loss: 0.0006870166980661452
step: 100, loss: 0.0004206752637401223
step: 110, loss: 0.0016060739289969206
step: 120, loss: 0.000213076506042853
step: 130, loss: 0.016521550714969635
step: 140, loss: 0.027473904192447662
step: 150, loss: 0.00021506792108993977
step: 160, loss: 5.4203366744332016e-05
step: 170, loss: 0.0005073924548923969
step: 180, loss: 9.357331873616204e-05
step: 190, loss: 0.011743343435227871
step: 200, loss: 0.00041495473124086857
step: 210, loss: 0.002596869133412838
step: 220, loss: 0.02219618484377861
```

```
step: 230, loss: 0.024682506918907166
step: 240, loss: 0.00018663030641619116
step: 250, loss: 0.0005920403054915369
step: 260, loss: 0.0011854668846353889
step: 270, loss: 0.004194135777652264
step: 280, loss: 0.0109592629596591
step: 290, loss: 0.00020993873476982117
step: 300, loss: 0.0016563701210543513
step: 310, loss: 0.0005385936819948256
step: 320, loss: 0.0001226418826263398
step: 330, loss: 0.028524739667773247
step: 340, loss: 0.02384384348988533
step: 350, loss: 0.0001417706225765869
step: 360, loss: 0.00046726653818041086
step: 370, loss: 0.0001387690717820078
step: 380, loss: 0.006724480073899031
step: 390, loss: 0.00037142637302167714
step: 400, loss: 0.050647567957639694
step: 410, loss: 0.000641263322904706
step: 420, loss: 0.00026412747683934870
step: 430, loss: 4.6243083488661796e-05
step: 440, loss: 0.0026103772688657045
step: 450, loss: 0.0001257206458831206
step: 460, loss: 0.00048130759387277067
step: 470, loss: 0.002466554520651698
step: 480, loss: 0.12621493637561798
step: 490, loss: 0.0005449372692964971
step: 500, loss: 0.00010110207222169265
step: 510, loss: 0.13165263831615448
step: 520, loss: 0.0019822297617793083
step: 530, loss: 0.04207827150821686
/usr/local/lib/python3.7/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=0.9354536950420954, f1=0.9333954354913834,
best_f1=0.9358974358974358
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.005759984254837036
step: 10, loss: 0.029931064695119858
step: 20, loss: 0.0005466046277433634
step: 30, loss: 0.0006992521230131388
step: 40, loss: 0.000577468890696764
step: 50, loss: 0.0009198893676511943
step: 60, loss: 0.0005452500772662461
step: 70, loss: 0.00018365166033618152
step: 80, loss: 0.00026199218700639904
```

```
step: 90, loss: 0.002723563928157091
step: 100, loss: 8.283270290121436e-05
step: 110, loss: 5.472078191814944e-05
step: 120, loss: 0.16236519813537598
step: 130, loss: 0.0003389510966371745
step: 140, loss: 0.00038289109943434596
step: 150, loss: 7.34229979570955e-05
step: 160, loss: 0.0013085752725601196
step: 170, loss: 0.00905402097851038
step: 180, loss: 0.0005465145222842693
step: 190, loss: 0.001322379568591714
step: 200, loss: 0.00011883122351719067
step: 210, loss: 0.0008667406509630382
step: 220, loss: 0.00016144388064276427
step: 230, loss: 0.0008764386875554919
step: 240, loss: 0.0008360145729966462
step: 250, loss: 6.423232844099402e-05
step: 260, loss: 0.00012213065929245204
step: 270, loss: 0.0006380114355124533
step: 280, loss: 0.005680836271494627
step: 290, loss: 0.001143038272857666
step: 300, loss: 0.03390984982252121
step: 310, loss: 0.008233853615820408
step: 320, loss: 0.01147383451461792
step: 330, loss: 0.0003423707385081798
step: 340, loss: 0.00030769180739298463
step: 350, loss: 0.004466481506824493
step: 360, loss: 0.00022744861780665815
step: 370, loss: 0.006224595010280609
step: 380, loss: 0.006315531674772501
step: 390, loss: 0.12721122801303864
step: 400, loss: 0.00014550378546118736
step: 410, loss: 5.1009643357247114e-05
step: 420, loss: 0.0016301858704537153
step: 430, loss: 0.05541117861866951
step: 440, loss: 0.011907807551324368
step: 450, loss: 0.0011057284427806735
Gradient overflow.  Skipping step, loss scaler 0 reducing loss scale to 32768.0
step: 460, loss: 0.00023649170179851353
step: 470, loss: 4.695031020673923e-05
step: 480, loss: 3.146300878142938e-05
step: 490, loss: 0.0003483303589746356
step: 500, loss: 0.0021673727314919233
step: 510, loss: 0.0001238165277754888
step: 520, loss: 0.0005160917644388974
step: 530, loss: 0.00051989313215017322
/usr/local/lib/python3.7/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=0.9360112097150864, f1=0.9348729792147806,
best_f1=0.9358974358974358
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0008355825557373464
step: 10, loss: 0.00022499554324895144
step: 20, loss: 0.00013223293353803456
step: 30, loss: 0.00010986456618411466
step: 40, loss: 0.00019821416935883462
step: 50, loss: 0.0013377632712945342
step: 60, loss: 0.00010886448581004515
step: 70, loss: 0.00014847249258309603
step: 80, loss: 0.0012717755744233727
step: 90, loss: 0.0001654146908549592
step: 100, loss: 4.533704850473441e-05
step: 110, loss: 2.7670854251482524e-05
step: 120, loss: 6.897923594806343e-05
step: 130, loss: 0.00015064141189213842
step: 140, loss: 0.0001832240232033655
step: 150, loss: 5.3567946451948956e-05
step: 160, loss: 0.0014471530448645353
step: 170, loss: 0.0009369614999741316
step: 180, loss: 3.559575998224318e-05
step: 190, loss: 7.321361772483215e-05
step: 200, loss: 9.173843136522919e-05
step: 210, loss: 0.00019836572755593807
step: 220, loss: 0.0035806167870759964
step: 230, loss: 2.425113598292228e-05
step: 240, loss: 0.00015114441339392215
step: 250, loss: 0.0003723755944520235
step: 260, loss: 0.0009226992842741311
step: 270, loss: 2.0146098904660903e-05
step: 280, loss: 0.0011049192398786545
step: 290, loss: 0.0729079395532608
step: 300, loss: 3.68739347322844e-05
step: 310, loss: 0.01257374882698059
step: 320, loss: 7.502979860873893e-05
step: 330, loss: 3.7076395528856665e-05
step: 340, loss: 2.837436113622971e-05
step: 350, loss: 0.00023923757544253021
step: 360, loss: 2.089476038236171e-05
step: 370, loss: 0.011564546264708042
step: 380, loss: 2.7454727387521416e-05
step: 390, loss: 3.0924766178941354e-05
step: 400, loss: 0.0003237284254282713
step: 410, loss: 2.3695680283708498e-05
```

```
step: 420, loss: 2.152805245714262e-05
step: 430, loss: 0.0001070157450158149
step: 440, loss: 0.001521856989711523
step: 450, loss: 2.6657427952159196e-05
step: 460, loss: 5.615750706056133e-05
step: 470, loss: 1.5497043932555243e-05
step: 480, loss: 0.0001615441287867725
step: 490, loss: 0.000830135599244386
step: 500, loss: 0.0005817008204758167
step: 510, loss: 0.0021958258002996445
step: 520, loss: 0.00010679013939807191
step: 530, loss: 0.006799535825848579
/usr/local/lib/python3.7/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=0.9382022471910113, f1=0.9323378441437238,
best_f1=0.9358974358974358
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 7.309911597985774e-05
step: 10, loss: 8.133828669087961e-05
step: 20, loss: 2.414305527054239e-05
step: 30, loss: 2.7566220524022356e-05
step: 40, loss: 6.466188642662019e-05
step: 50, loss: 5.734670048695989e-05
step: 60, loss: 0.0005265343352220953
step: 70, loss: 3.496962381177582e-05
step: 80, loss: 2.5510200430289842e-05
step: 90, loss: 2.5972211005864665e-05
step: 100, loss: 0.01587485335767269
step: 110, loss: 3.1441624742001295e-05
step: 120, loss: 0.00011856194760184735
step: 130, loss: 0.00010291244689142331
step: 140, loss: 0.019523916766047478
step: 150, loss: 8.720646292204037e-05
step: 160, loss: 0.0008008089498616755
step: 170, loss: 0.0007165406132116914
step: 180, loss: 6.733618647558615e-05
step: 190, loss: 0.0014627125347033143
step: 200, loss: 0.004393062554299831
step: 210, loss: 0.0003648198035079986
step: 220, loss: 0.014111100696027279
step: 230, loss: 0.017463358119130135
step: 240, loss: 8.10562924016267e-05
step: 250, loss: 2.7309288270771503e-05
step: 260, loss: 0.00010195457434747368
step: 270, loss: 4.936775440000929e-05
```

```
step: 280, loss: 0.00036187394289299846
step: 290, loss: 4.7538567741867155e-05
step: 300, loss: 0.00017981822020374238
step: 310, loss: 0.0001732778619043529
step: 320, loss: 0.00017528259195387363
step: 330, loss: 0.007129355799406767
step: 340, loss: 0.005115174688398838
step: 350, loss: 0.0002692163980100304
step: 360, loss: 0.00012863404117524624
step: 370, loss: 0.0002655237331055105
step: 380, loss: 0.005311053246259689
step: 390, loss: 8.120511483866721e-05
step: 400, loss: 6.660798680968583e-05
step: 410, loss: 4.0673512557987124e-05
step: 420, loss: 5.350595893105492e-05
step: 430, loss: 2.0060344468220137e-05
step: 440, loss: 9.360406693303958e-05
step: 450, loss: 3.1595085602020845e-05
step: 460, loss: 5.2191055146977305e-05
step: 470, loss: 0.0017207274213433266
step: 480, loss: 7.250547059811652e-05
step: 490, loss: 0.0009060962474904954
step: 500, loss: 0.0026563284918665886
step: 510, loss: 4.86390563310124e-05
step: 520, loss: 0.00019348584464751184
step: 530, loss: 0.0001533656322862953
/usr/local/lib/python3.7/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=0.9387755102040817, f1=0.9383624655013799,
best_f1=0.9358974358974358
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 7.262855069711804e-05
step: 10, loss: 0.00038091649184934795
step: 20, loss: 0.00014379702042788267
step: 30, loss: 0.0001652142673265189
step: 40, loss: 0.00010924050002358854
step: 50, loss: 0.0001415468577761203
step: 60, loss: 0.01290160696953535
step: 70, loss: 0.00012441458238754421
step: 80, loss: 3.810310226981528e-05
step: 90, loss: 3.572598143364303e-05
step: 100, loss: 0.0022954640444368124
step: 110, loss: 2.9764316423097625e-05
step: 120, loss: 5.203060572966933e-05
step: 130, loss: 0.0005254442221485078
```

```
step: 140, loss: 0.0001384083880111575
step: 150, loss: 2.1561419998761266e-05
step: 160, loss: 5.3743973694508895e-05
step: 170, loss: 6.727849540766329e-05
step: 180, loss: 3.072468098253012e-05
step: 190, loss: 0.0011059100506827235
step: 200, loss: 0.000933182833250612
step: 210, loss: 5.5529610108351335e-05
step: 220, loss: 0.0006514700944535434
step: 230, loss: 1.9672927010105923e-05
step: 240, loss: 2.2142607122077607e-05
step: 250, loss: 1.6733791198930703e-05
step: 260, loss: 1.920356044138316e-05
step: 270, loss: 0.002679127501323819
step: 280, loss: 2.2556276235263795e-05
step: 290, loss: 4.122330210520886e-05
step: 300, loss: 5.577989577432163e-05
step: 310, loss: 6.761207623640075e-05
step: 320, loss: 3.630356150097214e-05
step: 330, loss: 0.00019177155627403408
step: 340, loss: 2.1572524929069914e-05
step: 350, loss: 3.0090666768955998e-05
step: 360, loss: 0.0006781063857488334
step: 370, loss: 0.0008142524748109281
step: 380, loss: 2.4042408767854795e-05
step: 390, loss: 8.033720951061696e-05
step: 400, loss: 4.7686411562608555e-05
step: 410, loss: 2.9022949092905037e-05
step: 420, loss: 4.026004171464592e-05
step: 430, loss: 2.2634416382061318e-05
step: 440, loss: 0.00011226977949263528
step: 450, loss: 5.080740083940327e-05
step: 460, loss: 3.840774661512114e-05
step: 470, loss: 4.977414573659189e-05
step: 480, loss: 7.525023829657584e-05
step: 490, loss: 0.0024311181623488665
step: 500, loss: 3.270935485488735e-05
step: 510, loss: 0.00014729905524291098
step: 520, loss: 2.4496841433574446e-05
step: 530, loss: 2.3594438971485943e-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 11: dev_f1=0.9347111319868483, f1=0.9298000929800094,
best_f1=0.9358974358974358
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
```

```
step: 0, loss: 2.4559903977205977e-05
step: 10, loss: 2.0190571376588196e-05
step: 20, loss: 3.382137947482988e-05
step: 30, loss: 0.0005193030228838325
step: 40, loss: 3.482522879494354e-05
step: 50, loss: 0.06438743323087692
step: 60, loss: 7.441382331307977e-05
step: 70, loss: 0.00010185163409914821
step: 80, loss: 2.182946991524659e-05
step: 90, loss: 0.00010293370723957196
step: 100, loss: 0.0007599650998599827
step: 110, loss: 0.00012926376075483859
step: 120, loss: 1.728498682496138e-05
step: 130, loss: 7.207638554973528e-05
step: 140, loss: 1.9453091226750985e-05
step: 150, loss: 1.7367021428071894e-05
step: 160, loss: 5.388204226619564e-05
step: 170, loss: 4.0742248529568315e-05
step: 180, loss: 0.0005858806543983519
step: 190, loss: 4.6369561459869146e-05
step: 200, loss: 0.0007444637594744563
step: 210, loss: 0.0002035116485785693
step: 220, loss: 0.0018821409903466702
step: 230, loss: 1.3392295841185842e-05
step: 240, loss: 0.00010437277524033561
step: 250, loss: 2.667214175744448e-05
step: 260, loss: 2.346871406189166e-05
step: 270, loss: 4.285392424208112e-05
step: 280, loss: 0.001984104048460722
step: 290, loss: 0.0006853451486676931
step: 300, loss: 0.03859507292509079
step: 310, loss: 8.386038098251447e-05
step: 320, loss: 1.5314466509153135e-05
step: 330, loss: 6.468353967648e-05
step: 340, loss: 0.000882055377587676
step: 350, loss: 2.5163668396999128e-05
step: 360, loss: 0.0021876762621104717
step: 370, loss: 8.394069300265983e-05
step: 380, loss: 1.7948264940059744e-05
step: 390, loss: 0.18559151887893677
step: 400, loss: 4.826001531910151e-05
step: 410, loss: 0.00019405802595429122
step: 420, loss: 0.0008801458170637488
step: 430, loss: 0.015222040005028248
step: 440, loss: 0.00026946773868985474
step: 450, loss: 4.601033651852049e-05
step: 460, loss: 6.518858572235331e-05
step: 470, loss: 4.401839760248549e-05
```

```
step: 480, loss: 0.0015405818121507764
step: 490, loss: 9.638497431296855e-05
step: 500, loss: 9.04811458894983e-05
step: 510, loss: 0.0001501324150012806
step: 520, loss: 5.729968688683584e-05
step: 530, loss: 0.0003161555214319378
/usr/local/lib/python3.7/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=0.9393364928909952, f1=0.9340196537201684,
best_f1=0.9358974358974358
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 1.7720847608870827e-05
step: 10, loss: 1.8041100702248514e-05
step: 20, loss: 2.588336428743787e-05
step: 30, loss: 3.371136335772462e-05
step: 40, loss: 2.2205589630175382e-05
step: 50, loss: 0.00041822082130238414
step: 60, loss: 0.0003134091093670577
step: 70, loss: 0.0007419074536301196
step: 80, loss: 0.00020964909344911575
step: 90, loss: 0.02470560558140278
step: 100, loss: 3.734287020051852e-05
step: 110, loss: 0.00010821627074619755
step: 120, loss: 0.0002478087553754449
step: 130, loss: 4.986815474694595e-05
step: 140, loss: 0.001130991498939693
step: 150, loss: 0.003005096223205328
step: 160, loss: 0.0001014044537441805
step: 170, loss: 4.490140054258518e-05
step: 180, loss: 4.425729639478959e-05
step: 190, loss: 1.751232412061654e-05
step: 200, loss: 1.7151021893369034e-05
step: 210, loss: 8.030104072531685e-05
step: 220, loss: 2.5367004127474502e-05
step: 230, loss: 1.2423743100953288e-05
step: 240, loss: 4.125635314267129e-05
step: 250, loss: 4.0970524423755705e-05
step: 260, loss: 0.00018653615552466363
step: 270, loss: 3.2946347346296534e-05
step: 280, loss: 1.7411766748409718e-05
step: 290, loss: 0.000323774351272732
step: 300, loss: 3.655127147794701e-05
step: 310, loss: 2.6160612833336927e-05
step: 320, loss: 3.869570355163887e-05
step: 330, loss: 2.7579430025070906e-05
```

```
step: 340, loss: 1.5172874554991722e-05
step: 350, loss: 1.3697771464649122e-05
step: 360, loss: 1.5444818927790038e-05
step: 370, loss: 0.001929566846229136
step: 380, loss: 3.5342294722795486e-05
step: 390, loss: 0.0008451023022644222
step: 400, loss: 1.7653845134191215e-05
step: 410, loss: 9.949513332685456e-05
step: 420, loss: 1.648785655561369e-05
step: 430, loss: 0.0016127958660945296
step: 440, loss: 3.969619865529239e-05
step: 450, loss: 2.4306302293553017e-05
step: 460, loss: 3.2572363124927506e-05
step: 470, loss: 1.6286525351461023e-05
step: 480, loss: 2.8948414183105342e-05
step: 490, loss: 9.311053872806951e-05
step: 500, loss: 1.1786699360527564e-05
step: 510, loss: 1.7620281141716987e-05
step: 520, loss: 7.123462273739278e-05
step: 530, loss: 1.2133178643125575e-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=0.9403475810239549, f1=0.9335192933519293,
best_f1=0.9358974358974358
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 1.581360811542254e-05
step: 10, loss: 1.1671251741063315e-05
step: 20, loss: 1.2494523616624065e-05
step: 30, loss: 1.58285074576626984e-05
step: 40, loss: 2.6257819627062418e-05
step: 50, loss: 3.117509186267853e-05
step: 60, loss: 1.5586147128487937e-05
step: 70, loss: 1.690869430603925e-05
step: 80, loss: 1.591424188518431e-05
step: 90, loss: 2.3147778847487643e-05
step: 100, loss: 0.004102020990103483
step: 110, loss: 2.220085298176855e-05
step: 120, loss: 6.634197052335367e-05
step: 130, loss: 0.018001476302742958
step: 140, loss: 1.4830094187345821e-05
step: 150, loss: 1.0974625183735043e-05
step: 160, loss: 7.797498255968094e-05
step: 170, loss: 2.733737710514106e-05
step: 180, loss: 8.751584391575307e-05
step: 190, loss: 2.4531764211133122e-05
```

```
step: 200, loss: 6.179964111652225e-05
step: 210, loss: 1.4338497749122325e-05
step: 220, loss: 1.4502382327918895e-05
step: 230, loss: 1.9959643395850435e-05
step: 240, loss: 1.204376167152077e-05
step: 250, loss: 1.4282605661719572e-05
step: 260, loss: 6.403660518117249e-05
step: 270, loss: 2.9768691092613153e-05
step: 280, loss: 1.2300803064135835e-05
step: 290, loss: 1.645809788897168e-05
step: 300, loss: 2.1639620172209106e-05
step: 310, loss: 0.00030891262576915324
step: 320, loss: 3.9067777834134176e-05
step: 330, loss: 2.546489668020513e-05
step: 340, loss: 1.79444468813017e-05
step: 350, loss: 2.1438259864225984e-05
step: 360, loss: 0.00024509686045348644
step: 370, loss: 0.000733400578610599
step: 380, loss: 1.4819008356425911e-05
step: 390, loss: 0.0260605551302433
step: 400, loss: 2.871639298973605e-05
step: 410, loss: 1.3399715498962905e-05
step: 420, loss: 5.55237966182176e-05
step: 430, loss: 1.746758607623633e-05
step: 440, loss: 8.581849397160113e-05
Gradient overflow.  Skipping step, loss scaler 0 reducing loss scale to 32768.0
step: 450, loss: 0.008189959451556206
step: 460, loss: 2.0357672838144936e-05
step: 470, loss: 1.142166365752928e-05
step: 480, loss: 1.331030853179982e-05
step: 490, loss: 1.2636073734029196e-05
step: 500, loss: 1.1064024874940515e-05
step: 510, loss: 0.010185821913182735
step: 520, loss: 2.5587662094039842e-05
step: 530, loss: 0.001273882808163762
/usr/local/lib/python3.7/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=0.9404536862003782, f1=0.9336448598130841,
best_f1=0.9358974358974358
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 1.1652616194623988e-05
step: 10, loss: 1.4923270100553054e-05
step: 20, loss: 2.143812525901012e-05
step: 30, loss: 8.716288721188903e-05
step: 40, loss: 0.03714234009385109
```

```
step: 50, loss: 0.00013751228107139468
step: 60, loss: 1.1317348253214732e-05
step: 70, loss: 6.584537914022803e-05
step: 80, loss: 1.310916013608221e-05
step: 90, loss: 1.4059090972295962e-05
step: 100, loss: 6.256495544221252e-05
step: 110, loss: 4.521431401371956e-05
step: 120, loss: 0.000700093456543982
step: 130, loss: 0.04693368077278137
step: 140, loss: 2.5971035938709974e-05
step: 150, loss: 1.6335126929334365e-05
step: 160, loss: 1.5012683434179053e-05
step: 170, loss: 1.1175778126926161e-05
step: 180, loss: 1.2218815754749812e-05
step: 190, loss: 2.9402026484603994e-05
step: 200, loss: 0.00014450807066168636
step: 210, loss: 0.00019451050320640206
step: 220, loss: 1.1350863132975064e-05
step: 230, loss: 2.5415811251150444e-05
step: 240, loss: 1.7236467101611197e-05
step: 250, loss: 4.313247700338252e-05
step: 260, loss: 2.5878065571305342e-05
step: 270, loss: 1.756801975716371e-05
step: 280, loss: 2.5663528504082933e-05
step: 290, loss: 1.1332220310578123e-05
step: 300, loss: 1.4856087545922492e-05
step: 310, loss: 2.923184001701884e-05
step: 320, loss: 0.00012242489901836962
step: 330, loss: 3.865330290864222e-05
step: 340, loss: 1.0814434062922373e-05
step: 350, loss: 1.6978936400846578e-05
step: 360, loss: 3.233482493669726e-05
step: 370, loss: 1.0549953913141508e-05
step: 380, loss: 1.5221335161186289e-05
step: 390, loss: 2.0040653907926753e-05
step: 400, loss: 1.7899519662023522e-05
step: 410, loss: 1.3936127288616262e-05
step: 420, loss: 6.111847324064001e-05
step: 430, loss: 1.7404327081749216e-05
step: 440, loss: 0.00012003934534732252
step: 450, loss: 1.94859821931459e-05
step: 460, loss: 0.0001281368313357234
step: 470, loss: 0.008537035435438156
step: 480, loss: 1.0341334018448833e-05
step: 490, loss: 1.1224219633731991e-05
step: 500, loss: 5.999854329274967e-05
step: 510, loss: 5.2645817049779e-05
step: 520, loss: 1.1719654139596969e-05
```

step: 530, loss: 1.648025499889627e-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=0.9388523047977423, f1=0.9320297951582867,
best_f1=0.9358974358974358
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 [00:28, 198.11it/s]
load_f1 = 0.940795559666975
real_f1 = 0.9400369003690037
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, 190.95it/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 bert-base-uncased \
       --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 bert-base-uncased \
       --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: 0.5460110306739807
step: 10, loss: 0.3640379011631012
step: 20, loss: 0.38557374477386475
step: 30, loss: 0.3200463652610779
step: 40, loss: 0.17705665528774261
step: 50, loss: 0.4138333797454834
step: 60, loss: 0.3495092988014221
step: 70, loss: 0.20973734557628632
step: 80, loss: 0.2612484097480774
step: 90, loss: 0.3215775787830353
step: 100, loss: 0.5610376596450806
step: 110, loss: 0.24223178625106812
step: 120, loss: 0.31590598821640015
step: 130, loss: 0.22225825488567352
step: 140, loss: 0.13823223114013672
step: 150, loss: 0.14714591205120087
step: 160, loss: 0.30325013399124146
step: 170, loss: 0.28464218974113464
step: 180, loss: 0.06281592696905136
step: 190, loss: 0.18835780024528503
step: 200, loss: 0.22519442439079285
step: 210, loss: 0.1982792317867279
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
```

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=0.6539923954372623, f1=0.6819047619047619,
best_f1=0.6819047619047619
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.07427912950515747
step: 10, loss: 0.26146015524864197
step: 20, loss: 0.20621486008167267
step: 30, loss: 0.18790943920612335
step: 40, loss: 0.1430908441543579
step: 50, loss: 0.13172025978565216
step: 60, loss: 0.32401812076568604
step: 70, loss: 0.11444240808486938
step: 80, loss: 0.11838476359844208
step: 90, loss: 0.12966375052928925
step: 100, loss: 0.02194887027144432
step: 110, loss: 0.1906481236219406
step: 120, loss: 0.1337033063173294
step: 130, loss: 0.02349781058728695
step: 140, loss: 0.1641443818807602
step: 150, loss: 0.17309239506721497
step: 160, loss: 0.18990522623062134
step: 170, loss: 0.10621516406536102
step: 180, loss: 0.14439161121845245
step: 190, loss: 0.08148373663425446
step: 200, loss: 0.05172054097056389
step: 210, loss: 0.2188519686460495
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.6779661016949152, f1=0.7052023121387283,
best_f1=0.7052023121387283
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.07846951484680176
step: 10, loss: 0.1051056832075119
step: 20, loss: 0.0718395859003067
step: 30, loss: 0.2322656810283661
step: 40, loss: 0.12750861048698425
step: 50, loss: 0.035229433327913284

```
step: 60, loss: 0.16098646819591522
step: 70, loss: 0.09904292225837708
step: 80, loss: 0.2345113307237625
step: 90, loss: 0.14285115897655487
step: 100, loss: 0.09136670082807541
step: 110, loss: 0.1823520064353943
step: 120, loss: 0.12232840806245804
step: 130, loss: 0.12842801213264465
step: 140, loss: 0.1435440331697464
step: 150, loss: 0.20759905874729156
step: 160, loss: 0.014913223683834076
step: 170, loss: 0.10252367705106735
step: 180, loss: 0.04339880868792534
step: 190, loss: 0.25809401273727417
step: 200, loss: 0.04074728116393089
step: 210, loss: 0.10077787935733795
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.6852589641434264, f1=0.6800804828973844,
best_f1=0.6800804828973844
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.2867015302181244
step: 10, loss: 0.07729137688875198
step: 20, loss: 0.1713157743215561
step: 30, loss: 0.07907741516828537
step: 40, loss: 0.05741015449166298
step: 50, loss: 0.28890261054039
step: 60, loss: 0.1830299347639084
step: 70, loss: 0.19520939886569977
step: 80, loss: 0.18467283248901367
step: 90, loss: 0.0421462282538414
step: 100, loss: 0.23554624617099762
step: 110, loss: 0.16803649067878723
step: 120, loss: 0.06807751953601837
step: 130, loss: 0.22234056890010834
step: 140, loss: 0.21941722929477692
step: 150, loss: 0.12331709265708923
step: 160, loss: 0.03649255260825157
step: 170, loss: 0.12596359848976135
step: 180, loss: 0.23122838139533997
step: 190, loss: 0.03737068921327591
```

```
step: 200, loss: 0.45732468366622925
step: 210, loss: 0.15000581741333008
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.7035573122529645, f1=0.6811881188118812,
best_f1=0.6811881188118812
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.10171671956777573
step: 10, loss: 0.028143087401986122
step: 20, loss: 0.22986458241939545
step: 30, loss: 0.1058935672044754
step: 40, loss: 0.07174821197986603
step: 50, loss: 0.10862641036510468
step: 60, loss: 0.203245148062706
step: 70, loss: 0.1147703155875206
step: 80, loss: 0.08001385629177094
step: 90, loss: 0.12532895803451538
step: 100, loss: 0.02025780640542507
step: 110, loss: 0.27055442333221436
step: 120, loss: 0.13001364469528198
step: 130, loss: 0.07908331602811813
step: 140, loss: 0.05927503481507301
step: 150, loss: 0.2591136395931244
step: 160, loss: 0.08920928835868835
step: 170, loss: 0.07699290663003922
step: 180, loss: 0.027257554233074188
step: 190, loss: 0.021941982209682465
step: 200, loss: 0.08688194304704666
step: 210, loss: 0.017693115398287773
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.7099391480730223, f1=0.674698795180723,
best_f1=0.674698795180723
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
```

```
step: 0, loss: 0.050902061995792389
step: 10, loss: 0.050597675144672394
step: 20, loss: 0.022069135680794716
step: 30, loss: 0.003705403069034219
step: 40, loss: 0.007740612607449293
step: 50, loss: 0.0803874284029007
step: 60, loss: 0.13684354722499847
step: 70, loss: 0.0435134582221508
step: 80, loss: 0.15811839699745178
step: 90, loss: 0.1932021975517273
step: 100, loss: 0.009985497722135067
step: 110, loss: 0.06163354963064194
step: 120, loss: 0.12373057007789612
step: 130, loss: 0.09795457869768143
step: 140, loss: 0.07236781716346741
step: 150, loss: 0.030005119740962982
step: 160, loss: 0.02981763519346714
step: 170, loss: 0.211735799908638
step: 180, loss: 0.10200728476047516
step: 190, loss: 0.1428057849407196
step: 200, loss: 0.009075342677533627
step: 210, loss: 0.08426512777805328
/usr/local/lib/python3.7/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=0.6895161290322581, f1=0.6827309236947792,
best_f1=0.674698795180723
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.04055974632501602
step: 10, loss: 0.04148866981267929
step: 20, loss: 0.006609252654016018
step: 30, loss: 0.02916661463677883
step: 40, loss: 0.09240681678056717
step: 50, loss: 0.1086752712726593
step: 60, loss: 0.026204394176602364
step: 70, loss: 0.005446212366223335
step: 80, loss: 0.049453362822532654
step: 90, loss: 0.03244739770889282
step: 100, loss: 0.005402076989412308
step: 110, loss: 0.10385506600141525
step: 120, loss: 0.08952292054891586
step: 130, loss: 0.16519513726234436
step: 140, loss: 0.02919781766831875
step: 150, loss: 0.011475283652544022
step: 160, loss: 0.0687779113650322
step: 170, loss: 0.03219043090939522
```

```
step: 180, loss: 0.10969164967536926
step: 190, loss: 0.13864856958389282
step: 200, loss: 0.028296858072280884
step: 210, loss: 0.05896434187889099
/usr/local/lib/python3.7/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=0.6992481203007519, f1=0.6765799256505576,
best_f1=0.674698795180723
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.08829942345619202
step: 10, loss: 0.07809695601463318
step: 20, loss: 0.011327533051371574
step: 30, loss: 0.07729370892047882
step: 40, loss: 0.059367381036281586
step: 50, loss: 0.13742542266845703
step: 60, loss: 0.18593141436576843
step: 70, loss: 0.024722177535295486
step: 80, loss: 0.07872379571199417
step: 90, loss: 0.0568871907889843
step: 100, loss: 0.09185666590929031
step: 110, loss: 0.11116853356361389
step: 120, loss: 0.03900628909468651
step: 130, loss: 0.006590620614588261
step: 140, loss: 0.031745586544275284
step: 150, loss: 0.006725140381604433
step: 160, loss: 0.025930000469088554
Gradient overflow.  Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 170, loss: 0.028039205819368362
step: 180, loss: 0.05684515833854675
step: 190, loss: 0.013810480013489723
step: 200, loss: 0.015137090347707272
step: 210, loss: 0.25603026151657104
/usr/local/lib/python3.7/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=0.696, f1=0.697029702970297, best_f1=0.674698795180723
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0033837640658020973
step: 10, loss: 0.029602201655507088
step: 20, loss: 0.007716941647231579
step: 30, loss: 0.02074120193719864
step: 40, loss: 0.05881250649690628
step: 50, loss: 0.19398312270641327
```

```
step: 60, loss: 0.1432693898677826
step: 70, loss: 0.12687554955482483
step: 80, loss: 0.0245358869433403
step: 90, loss: 0.00124840578064322247
step: 100, loss: 0.020483000203967094
step: 110, loss: 0.012955144047737122
step: 120, loss: 0.019589394330978394
step: 130, loss: 0.09674002230167389
step: 140, loss: 0.043878763914108276
step: 150, loss: 0.03161392733454704
step: 160, loss: 0.0022610013838857412
step: 170, loss: 0.01930256001651287
step: 180, loss: 0.014774334616959095
step: 190, loss: 0.0008339303894899786
step: 200, loss: 0.06994052231311798
step: 210, loss: 0.018406150862574577
/usr/local/lib/python3.7/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=0.6720647773279351, f1=0.6680327868852459,
best_f1=0.674698795180723
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.04963570833206177
step: 10, loss: 0.13766159117221832
step: 20, loss: 0.007140354719012976
step: 30, loss: 0.08685290068387985
step: 40, loss: 0.0010464085498824716
step: 50, loss: 0.014028976671397686
step: 60, loss: 0.05703551322221756
step: 70, loss: 0.16602793335914612
step: 80, loss: 0.058530911803245544
step: 90, loss: 0.1051250547170639
step: 100, loss: 0.06448351591825485
step: 110, loss: 0.0019111970905214548
step: 120, loss: 0.01657973602414131
step: 130, loss: 0.04611348733305931
step: 140, loss: 0.005275181494653225
step: 150, loss: 0.05492795258760452
step: 160, loss: 0.10450175404548645
step: 170, loss: 0.008550608530640602
step: 180, loss: 0.043130047619342804
step: 190, loss: 0.2928697466850281
step: 200, loss: 0.04026957228779793
step: 210, loss: 0.04697244614362717
/usr/local/lib/python3.7/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=0.6915113871635611, f1=0.6818181818181819,
best_f1=0.674698795180723
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.04855935275554657
step: 10, loss: 0.03713934123516083
step: 20, loss: 0.057616036385297775
step: 30, loss: 0.0029454207979142666
step: 40, loss: 0.028485752642154694
step: 50, loss: 0.029838865622878075
step: 60, loss: 0.050202395766973495
step: 70, loss: 0.003365657525137067
step: 80, loss: 0.07429825514554977
step: 90, loss: 0.018998514860868454
step: 100, loss: 0.20468126237392426
step: 110, loss: 0.04424005746841431
step: 120, loss: 0.017086008563637733
step: 130, loss: 0.015433981083333492
step: 140, loss: 0.007049860432744026
step: 150, loss: 0.03592013567686081
step: 160, loss: 0.016304880380630493
step: 170, loss: 0.0023198402486741543
step: 180, loss: 0.0017704698257148266
step: 190, loss: 0.06504325568675995
step: 200, loss: 0.0013662214623764157
step: 210, loss: 0.004096821416169405
/usr/local/lib/python3.7/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=0.6604127579737337, f1=0.6783625730994153,
best_f1=0.674698795180723
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.06822383403778076
step: 10, loss: 0.047326356172561646
step: 20, loss: 0.03380564600229263
step: 30, loss: 0.04519183561205864
step: 40, loss: 0.01958766020834446
step: 50, loss: 0.0023678282741457224
step: 60, loss: 0.014476804994046688
step: 70, loss: 0.006783766206353903
step: 80, loss: 0.0064787482842803
step: 90, loss: 0.0730719193816185
step: 100, loss: 0.011650768108665943
step: 110, loss: 0.00471073854714632
```

```
step: 120, loss: 0.00239170272834599
step: 130, loss: 0.013794654048979282
step: 140, loss: 0.018376894295215607
step: 150, loss: 0.04575144127011299
step: 160, loss: 0.05398211255669594
step: 170, loss: 0.01932297833263874
step: 180, loss: 0.01130414754152298
step: 190, loss: 0.004152459558099508
step: 200, loss: 0.004593872930854559
step: 210, loss: 0.030747972428798676
/usr/local/lib/python3.7/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=0.6749482401656315, f1=0.6666666666666666,
best_f1=0.674698795180723
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.028550049290060997
step: 10, loss: 0.0013405540958046913
step: 20, loss: 0.08097144216299057
step: 30, loss: 0.25377705693244934
step: 40, loss: 0.09989720582962036
step: 50, loss: 0.0039873430505394936
step: 60, loss: 0.024640027433633804
step: 70, loss: 0.10107827186584473
step: 80, loss: 0.02193531207740307
step: 90, loss: 0.000872860080562532
step: 100, loss: 0.09232308715581894
step: 110, loss: 0.0038806325756013393
step: 120, loss: 0.0318298377096653
step: 130, loss: 0.0061177900061011314
step: 140, loss: 0.036169059574604034
step: 150, loss: 0.005896552931517363
step: 160, loss: 0.00803544744849205
step: 170, loss: 0.010572453960776329
step: 180, loss: 0.18136487901210785
step: 190, loss: 0.004217945039272308
step: 200, loss: 0.0018376258667558432
step: 210, loss: 0.01575513184070587
/usr/local/lib/python3.7/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=0.6878727634194831, f1=0.6720647773279351,
best_f1=0.674698795180723
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
```

```
step: 0, loss: 0.0019961893558502197
step: 10, loss: 0.017384210601449013
step: 20, loss: 0.0975879430770874
step: 30, loss: 0.01453519519418478
step: 40, loss: 0.00259379064664244465
step: 50, loss: 0.038933444768190384
step: 60, loss: 0.009425967931747437
step: 70, loss: 0.004380682948976755
step: 80, loss: 0.01962720789015293
step: 90, loss: 0.28139054775238037
step: 100, loss: 0.00554626202210784
step: 110, loss: 0.010361350141465664
step: 120, loss: 0.042718555778265
step: 130, loss: 0.026853490620851517
step: 140, loss: 0.02252880111336708
step: 150, loss: 0.0029357830062508583
step: 160, loss: 0.010521702468395233
step: 170, loss: 0.03647460415959358
step: 180, loss: 0.0035917069762945175
step: 190, loss: 0.004632844123989344
step: 200, loss: 0.0030912859365344048
step: 210, loss: 0.024338852614164352
/usr/local/lib/python3.7/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=0.6782273603082851, f1=0.671875, best_f1=0.674698795180723
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.005284283775836229
step: 10, loss: 0.0009890757501125336
step: 20, loss: 0.006231535691767931
step: 30, loss: 0.0389404371380806
step: 40, loss: 0.003501209430396557
step: 50, loss: 0.011801994405686855
step: 60, loss: 0.005087288562208414
step: 70, loss: 0.005880544893443584844
step: 80, loss: 0.0011002932442352176
step: 90, loss: 0.012507778592407703
step: 100, loss: 0.00395980104804039
step: 110, loss: 0.026293661445379257
step: 120, loss: 0.00794249027967453
step: 130, loss: 0.07599946111440659
step: 140, loss: 0.0035379838664084673
step: 150, loss: 0.0013969545252621174
step: 160, loss: 0.025451717898249626
step: 170, loss: 0.002121566329151392
step: 180, loss: 0.0009701864910311997
```

```
step: 190, loss: 0.0028855896089226007
step: 200, loss: 0.009082087315618992
step: 210, loss: 0.016096312552690506
/usr/local/lib/python3.7/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=0.67578125, f1=0.6666666666666666, best_f1=0.674698795180723
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.")
2293it [00:04, 467.66it/s]
load_f1 = 0.6975806451612904
real_f1 = 0.6881287726358148
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 [00:17, 250.41it/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 bert-base-uncased \
  --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 bert-base-uncased \
  --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: 0.5863348841667175
step: 10, loss: 0.36719444394111633
step: 20, loss: 0.2808379828929901
step: 30, loss: 0.4149381220340729
step: 40, loss: 0.4123986065387726
step: 50, loss: 0.283336877822876
step: 60, loss: 0.19170859456062317
step: 70, loss: 0.2671651542186737
step: 80, loss: 0.23234212398529053
step: 90, loss: 0.2686989903450012
step: 100, loss: 0.2523171305656433
step: 110, loss: 0.39307674765586853
step: 120, loss: 0.037775252014398575
step: 130, loss: 0.09541941434144974
step: 140, loss: 0.11081825941801071
step: 150, loss: 0.13763533532619476
step: 160, loss: 0.05665842071175575
step: 170, loss: 0.1655033379793167
step: 180, loss: 0.0316956602036953
step: 190, loss: 0.12065298110246658
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.7052896725440806, f1=0.7282051282051282,
best_f1=0.7282051282051282
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.14337432384490967
step: 10, loss: 0.0383516326546669
step: 20, loss: 0.1577739417552948
step: 30, loss: 0.13050086796283722
step: 40, loss: 0.136930450797081
step: 50, loss: 0.12288717180490494
step: 60, loss: 0.43379244208335876
step: 70, loss: 0.10225459188222885
step: 80, loss: 0.10163073241710663
step: 90, loss: 0.13978393375873566
step: 100, loss: 0.009211502969264984
step: 110, loss: 0.04561059921979904
step: 120, loss: 0.14400553703308105
step: 130, loss: 0.05973486974835396
step: 140, loss: 0.08390423655509949
step: 150, loss: 0.13584911823272705
step: 160, loss: 0.025293663144111633
step: 170, loss: 0.14221717417240143
step: 180, loss: 0.272377610206604
step: 190, loss: 0.022600259631872177
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.7885714285714286, f1=0.7999999999999999,
best_f1=0.7999999999999999
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.06338358670473099
step: 10, loss: 0.34959420561790466
step: 20, loss: 0.032400909811258316
step: 30, loss: 0.11705099791288376
step: 40, loss: 0.07433856278657913
step: 50, loss: 0.1720314919948578
step: 60, loss: 0.11407586187124252
step: 70, loss: 0.22069209814071655
step: 80, loss: 0.03798311576247215
step: 90, loss: 0.12273147702217102
```

```
step: 100, loss: 0.026897137984633446
step: 110, loss: 0.032313939183950424
step: 120, loss: 0.02946445718407631
step: 130, loss: 0.005897526163607836
step: 140, loss: 0.013948960229754448
step: 150, loss: 0.109466053545475
step: 160, loss: 0.08083123713731766
step: 170, loss: 0.127328559756279
step: 180, loss: 0.03218990936875343
step: 190, loss: 0.06883582472801208
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.8042895442359249, f1=0.8216216216216217,
best_f1=0.8216216216216217
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.024944698438048363
step: 10, loss: 0.11612502485513687
step: 20, loss: 0.02769279107451439
step: 30, loss: 0.006437323521822691
step: 40, loss: 0.006287685129791498
step: 50, loss: 0.014398363418877125
step: 60, loss: 0.0669625848531723
step: 70, loss: 0.02134626731276512
step: 80, loss: 0.08039671182632446
step: 90, loss: 0.008957872167229652
step: 100, loss: 0.07409115880727768
step: 110, loss: 0.0014459789963439107
step: 120, loss: 0.06643760949373245
step: 130, loss: 0.11857735365629196
step: 140, loss: 0.027325909584760666
step: 150, loss: 0.010603741742670536
step: 160, loss: 0.003124062204733491
step: 170, loss: 0.1760624796152115
step: 180, loss: 0.018527289852499962
step: 190, loss: 0.052404824644327164
/usr/local/lib/python3.7/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=0.8031496062992125, f1=0.7648578811369509,
best_f1=0.8216216216216217
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
```

```
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.12191174924373627
step: 10, loss: 0.027622591704130173
step: 20, loss: 0.011495121754705906
step: 30, loss: 0.0357002392411232
step: 40, loss: 0.20701080560684204
step: 50, loss: 0.07788628339767456
step: 60, loss: 0.12398643046617508
step: 70, loss: 0.027227899059653282
step: 80, loss: 0.003119734115898609
step: 90, loss: 0.006592325400561094
step: 100, loss: 0.002146015642210841
step: 110, loss: 0.002570433309301734
step: 120, loss: 0.0019475817680358887
step: 130, loss: 0.038686059415340424
step: 140, loss: 0.009533567354083061
step: 150, loss: 0.1544882357120514
step: 160, loss: 0.20964553952217102
step: 170, loss: 0.010844715870916843
step: 180, loss: 0.058047857135534286
step: 190, loss: 0.12381460517644882
/usr/local/lib/python3.7/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=0.7888040712468193, f1=0.7719298245614035,
best_f1=0.8216216216216217
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.008434156887233257
step: 10, loss: 0.1966938078403473
step: 20, loss: 0.00402115471661090085
step: 30, loss: 0.0010808712104335427
step: 40, loss: 0.06296687573194504
step: 50, loss: 0.016001274809241295
step: 60, loss: 0.012607386335730553
step: 70, loss: 0.0145558537915349
step: 80, loss: 0.03392740339040756
Gradient overflow.  Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 90, loss: 0.007304161787033081
step: 100, loss: 0.0027104979380965233
step: 110, loss: 0.010309840552508831
step: 120, loss: 0.04039593040943146
step: 130, loss: 0.0033841419499367476
step: 140, loss: 0.08252058178186417
step: 150, loss: 0.00112355244345963
step: 160, loss: 0.006836860906332731
```

```
step: 170, loss: 0.012529334984719753
step: 180, loss: 0.002208793070167303
step: 190, loss: 0.06509532779455185
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.8088642659279778, f1=0.7967914438502673,
best_f1=0.7967914438502673
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0017602334264665842
step: 10, loss: 0.0012402830179780722
step: 20, loss: 0.16751225292682648
step: 30, loss: 0.011175879277288914
step: 40, loss: 0.004125806037336588
step: 50, loss: 0.03386402875185013
step: 60, loss: 0.2181219458580017
step: 70, loss: 0.00309959295108914375
step: 80, loss: 0.006915769539773464
step: 90, loss: 0.020644504576921463
step: 100, loss: 0.00037453623372130096
step: 110, loss: 0.04211104288697243
step: 120, loss: 0.008097061887383461
step: 130, loss: 0.0012472844682633877
step: 140, loss: 0.0029775933362543583
step: 150, loss: 0.01819540373980999
step: 160, loss: 0.06871362030506134
step: 170, loss: 0.004258759319782257
step: 180, loss: 0.0014769192785024643
step: 190, loss: 0.0023346843663603067
/usr/local/lib/python3.7/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=0.7967032967032966, f1=0.7880434782608696,
best_f1=0.7967914438502673
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.003444471163675189
step: 10, loss: 0.0027870042249560356
step: 20, loss: 0.16861699521541595
step: 30, loss: 0.01047882903367281
step: 40, loss: 0.0030063132289797068
```

```
step: 50, loss: 0.0014851322630420327
step: 60, loss: 0.0009774441132321954
step: 70, loss: 0.05389348417520523
step: 80, loss: 0.0005209450027905405
step: 90, loss: 0.05968094617128372
step: 100, loss: 0.00372248119674623
step: 110, loss: 0.0015918747521936893
step: 120, loss: 0.01903715543448925
step: 130, loss: 0.043780550360679626
step: 140, loss: 0.0048003895208239555
step: 150, loss: 0.014294222928583622
step: 160, loss: 0.003174350829795003
step: 170, loss: 0.0020541115663945675
step: 180, loss: 0.013504471629858017
step: 190, loss: 0.00554483849555254
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.8103896103896104, f1=0.7948051948051948,
best_f1=0.7948051948051948
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0012671143049374223
step: 10, loss: 0.04523082077503204
step: 20, loss: 0.0017400607466697693
step: 30, loss: 0.0044248090125620365
step: 40, loss: 0.000816446146927774
step: 50, loss: 0.09060851484537125
step: 60, loss: 0.0006074737175367773
step: 70, loss: 0.0005116761312820017
step: 80, loss: 0.0016094736056402326
step: 90, loss: 0.034241221845149994
step: 100, loss: 0.0019181654788553715
step: 110, loss: 0.00936796236783266
step: 120, loss: 0.001299688359722495
step: 130, loss: 0.00070259312633425
step: 140, loss: 0.022694319486618042
step: 150, loss: 0.001341380411759019
step: 160, loss: 0.025770826265215874
step: 170, loss: 0.003401722526177764
step: 180, loss: 0.002815653569996357
step: 190, loss: 0.001213572919368744
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
```

```
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.8184281842818427, f1=0.8087431693989071,
best_f1=0.8087431693989071
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.001786579261533916
step: 10, loss: 0.07044008374214172
step: 20, loss: 0.004325034096837044
step: 30, loss: 0.003706501331180334
step: 40, loss: 0.019371170550584793
step: 50, loss: 0.0009484820766374469
step: 60, loss: 0.011258131824433804
step: 70, loss: 0.0019577222706720233
step: 80, loss: 0.002422815188765526
step: 90, loss: 0.0012219113996252418
step: 100, loss: 0.016109993681311607
step: 110, loss: 0.04983460158109665
step: 120, loss: 0.10796018689870834
step: 130, loss: 0.0013222881825640798
step: 140, loss: 0.005334322340786457
step: 150, loss: 0.0014415071345865726
step: 160, loss: 0.03097892552614212
step: 170, loss: 0.001813960145227611
step: 180, loss: 0.0017415813636034727
step: 190, loss: 0.0015784410061314702
/usr/local/lib/python3.7/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=0.8136482939632546, f1=0.8083989501312335,
best_f1=0.8087431693989071
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.000976820127107203
step: 10, loss: 0.0017297740560024977
step: 20, loss: 0.015570908784866333
step: 30, loss: 0.03148031234741211
step: 40, loss: 0.000352498289430514
step: 50, loss: 0.0006074673729017377
step: 60, loss: 0.0011515399673953652
step: 70, loss: 0.0033501535654067993
step: 80, loss: 0.0005768518312834203
```

```
step: 90, loss: 0.00033280940260738134
step: 100, loss: 0.0063935332000255585
step: 110, loss: 0.0006498200818896294
step: 120, loss: 0.000700960517860949
step: 130, loss: 0.0008233832777477801
step: 140, loss: 0.0016656082589179277
step: 150, loss: 0.000462591735413298
step: 160, loss: 0.006232441868633032
step: 170, loss: 0.0006632785080000758
step: 180, loss: 0.003730416763573885
step: 190, loss: 0.004310243763029575
/usr/local/lib/python3.7/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=0.8070175438596492, f1=0.8223350253807106,
best_f1=0.8087431693989071
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0005789405549876392
step: 10, loss: 0.0022402540780603886
step: 20, loss: 0.0007700963760726154
step: 30, loss: 0.0003837059484794736
step: 40, loss: 0.0022261515259742737
step: 50, loss: 0.011634394526481628
step: 60, loss: 0.0006709602894261479
step: 70, loss: 0.037935446947813034
step: 80, loss: 0.0008645522175356746
step: 90, loss: 0.0004095335316378623
step: 100, loss: 0.00032777138403616846
step: 110, loss: 0.00068328419001773
step: 120, loss: 0.0015380216063931584
step: 130, loss: 0.0008774831658229232
step: 140, loss: 0.023356324061751366
step: 150, loss: 0.0010922375367954373
step: 160, loss: 0.0008180485456250608
step: 170, loss: 0.0017802031943574548
step: 180, loss: 0.0020912194158881903
step: 190, loss: 0.0006197789334692061
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.8188976377952756, f1=0.8073878627968338,
best_f1=0.8073878627968338
```

```
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0022536034230014283
step: 10, loss: 0.0037637220229953527
step: 20, loss: 0.0015495785046368837
step: 30, loss: 0.0011416713241487741
step: 40, loss: 0.0009505274356342852
step: 50, loss: 0.03564697876572609
step: 60, loss: 0.0015775869833305478
step: 70, loss: 0.0038602440617978573
step: 80, loss: 0.0031725293956696987
step: 90, loss: 0.0028974025044590235
step: 100, loss: 0.0013501200592145324
step: 110, loss: 0.0005852851900272071
step: 120, loss: 0.011359594762325287
step: 130, loss: 0.00038467784179374576
step: 140, loss: 0.0012980567989870906
step: 150, loss: 0.0007756119011901319
step: 160, loss: 0.0016228166641667485
step: 170, loss: 0.0004982429090887308
step: 180, loss: 0.0015824951697140932
step: 190, loss: 0.005967116914689541
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.8253968253968254, f1=0.8106666666666666,
best_f1=0.8106666666666666
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0038795163854956627
step: 10, loss: 0.0010554904583841562
step: 20, loss: 0.0002915211080107838
step: 30, loss: 0.001013193977996707
step: 40, loss: 0.003027983708307147
step: 50, loss: 0.0007099007489159703
step: 60, loss: 0.08579505234956741
step: 70, loss: 0.0007833486306481063
step: 80, loss: 0.0022650519385933876
step: 90, loss: 0.004383643623441458
step: 100, loss: 0.006480396259576082
step: 110, loss: 0.048109084367752075
step: 120, loss: 0.03705503046512604
```

```
step: 130, loss: 0.0037940458860248327
step: 140, loss: 0.000668252760078758
step: 150, loss: 0.0009469513315707445
step: 160, loss: 0.011536970734596252
step: 170, loss: 0.10653608292341232
step: 180, loss: 0.004057825077325106
step: 190, loss: 0.001079733483493328
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.8320000000000001, f1=0.8128342245989304,
best_f1=0.8128342245989304
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0038182553835213184
step: 10, loss: 0.0008416124037466943
step: 20, loss: 0.002149177948012948
step: 30, loss: 0.0012866926845163107
step: 40, loss: 0.0004670748603530228
step: 50, loss: 0.003149968571960926
step: 60, loss: 0.0009478137362748384
step: 70, loss: 0.0011384247336536646
step: 80, loss: 0.0019959311466664076
step: 90, loss: 0.0009325107675977051
step: 100, loss: 0.000826692848931998
step: 110, loss: 0.0012162926141172647
step: 120, loss: 0.0018391249468550086
step: 130, loss: 0.0006646986585110426
step: 140, loss: 0.029006563127040863
step: 150, loss: 0.0017820593202486634
step: 160, loss: 0.010929107666015625
step: 170, loss: 0.0008678120793774724
step: 180, loss: 0.0008574177627451718
step: 190, loss: 0.13083042204380035
/usr/local/lib/python3.7/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=0.8279569892473119, f1=0.8184281842818427,
best_f1=0.8128342245989304
mv: cannot stat '*_dev.pt': No such file or directory
Selected optimization level O2:  FP16 training with FP32 batchnorm and FP32
master weights.
```

```
Defaults for this optimization level are:
enabled                : True
opt_level              : O2
cast_model_type        : torch.float16
patch_torch_functions  : False
keep_batchnorm_fp32     : True
master_weights          : True
loss_scale              : dynamic
Processing user overrides (additional kwargs that are not None)…
After processing overrides, optimization options are:
enabled                : True
opt_level              : O2
cast_model_type        : torch.float16
patch_torch_functions  : False
keep_batchnorm_fp32     : True
master_weights          : True
loss_scale              : dynamic
Warning:  multi_tensor_applier fused unscale kernel is unavailable, possibly
because apex was installed without --cuda_ext --cpp_ext. Using Python fallback.
Original ImportError was: ModuleNotFoundError("No module named 'amp_C'")
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
2049it [00:08, 234.55it/s]
load_f1 = 0.7525773195876289
real_f1 = 0.6987951807228916
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 [00:16, 259.95it/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 bert-base-uncased \
    --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 bert-base-uncased \
    --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: 0.6161351799964905
step: 10, loss: 0.3771972060203552
step: 20, loss: 0.2929035425186157
step: 30, loss: 0.3960883319377899
step: 40, loss: 0.2657056152820587
step: 50, loss: 0.2588040828704834
step: 60, loss: 0.2468966245651245
step: 70, loss: 0.35691922903060913
step: 80, loss: 0.367149293422699
step: 90, loss: 0.24240654706954956
step: 100, loss: 0.19538556039333344
step: 110, loss: 0.22437922656536102
step: 120, loss: 0.17751477658748627
step: 130, loss: 0.023072484880685806
step: 140, loss: 0.07158875465393066
step: 150, loss: 0.2852569818496704
step: 160, loss: 0.13901107013225555
step: 170, loss: 0.27880918979644775
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.7574257425742574, f1=0.7303921568627451,
best_f1=0.7303921568627451
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.03112310729920864
step: 10, loss: 0.17344270646572113
step: 20, loss: 0.04922337085008621
step: 30, loss: 0.15818728506565094
step: 40, loss: 0.12795612215995789
step: 50, loss: 0.06808970868587494

```
step: 60, loss: 0.16136465966701508
step: 70, loss: 0.06251000612974167
step: 80, loss: 0.022791776806116104
step: 90, loss: 0.218578040599823
step: 100, loss: 0.21618632972240448
step: 110, loss: 0.07609671354293823
step: 120, loss: 0.05584296956658363
step: 130, loss: 0.05387517064809799
step: 140, loss: 0.3202720284461975
step: 150, loss: 0.15689127147197723
step: 160, loss: 0.12183349579572678
step: 170, loss: 0.04670441895723343
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.812206572769953, f1=0.7934272300469484,
best_f1=0.7934272300469484
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.1102205365896225
step: 10, loss: 0.04202934354543686
step: 20, loss: 0.01655176840722561
step: 30, loss: 0.11652398109436035
step: 40, loss: 0.058949198573827744
step: 50, loss: 0.19922301173210144
step: 60, loss: 0.07201080769300461
step: 70, loss: 0.06148970127105713
step: 80, loss: 0.037860725075006485
step: 90, loss: 0.0638519749045372
step: 100, loss: 0.0014333612052723765
step: 110, loss: 0.0440671443939209
step: 120, loss: 0.13122093677520752
step: 130, loss: 0.13080385327339172
step: 140, loss: 0.11479058861732483
step: 150, loss: 0.011300114914774895
step: 160, loss: 0.03273747116327286
step: 170, loss: 0.05341845378279686
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.8316831683168316, f1=0.8232445520581113,
best_f1=0.8232445520581113
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.03228550776839256
step: 10, loss: 0.009719460271298885
step: 20, loss: 0.0015901743900030851
step: 30, loss: 0.08329065144062042
step: 40, loss: 0.002968223299831152
step: 50, loss: 0.046448152512311935
step: 60, loss: 0.17650921642780304
step: 70, loss: 0.010933252051472664
step: 80, loss: 0.14375154674053192
step: 90, loss: 0.05631589889526367
step: 100, loss: 0.20671147108078003
step: 110, loss: 0.09203403443098068
step: 120, loss: 0.009909788146615028
step: 130, loss: 0.012697913683950901
step: 140, loss: 0.04340779036283493
step: 150, loss: 0.12650710344314575
step: 160, loss: 0.04213138669729233
step: 170, loss: 0.02165766805410385
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.8369829683698297, f1=0.8129330254041571,
best_f1=0.8129330254041571
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.051570579409599304
step: 10, loss: 0.018780067563056946
step: 20, loss: 0.034185245633125305
step: 30, loss: 0.020762693136930466
step: 40, loss: 0.018236752599477768
step: 50, loss: 0.012874462641775608
step: 60, loss: 0.01376788504421711
step: 70, loss: 0.09232843667268753
step: 80, loss: 0.05027405172586441
step: 90, loss: 0.021878782659769058
step: 100, loss: 0.009818745777010918
step: 110, loss: 0.013513852842152119
```

```
step: 120, loss: 0.013012184761464596
step: 130, loss: 0.023350976407527924
step: 140, loss: 0.043398793786764145
step: 150, loss: 0.04778027534484863
step: 160, loss: 0.061197079718112946
Gradient overflow.  Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 170, loss: 0.019385050982236862
/usr/local/lib/python3.7/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=0.8363636363636364, f1=0.8181818181818181,
best_f1=0.8129330254041571
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.003182914573699236
step: 10, loss: 0.11796750873327255
step: 20, loss: 0.06253071874380112
step: 30, loss: 0.010079636238515377
step: 40, loss: 0.03480485454201698
step: 50, loss: 0.09468475729227066
step: 60, loss: 0.01721484400331974
step: 70, loss: 0.0367811918258667
step: 80, loss: 0.0035045347176492214
step: 90, loss: 0.025548089295625687
step: 100, loss: 0.001048882957547903
step: 110, loss: 0.002597035840153694
step: 120, loss: 0.005008351057767868
step: 130, loss: 0.0429978109896183
step: 140, loss: 0.003866314422339201
step: 150, loss: 0.04989969730377197
step: 160, loss: 0.07806938886642456
step: 170, loss: 0.005324242170900106
/usr/local/lib/python3.7/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=0.8201058201058201, f1=0.8174807197943444,
best_f1=0.8129330254041571
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0014212167588993907
step: 10, loss: 0.0017490005120635033
step: 20, loss: 0.007514874916523695
step: 30, loss: 0.02463478036224842
step: 40, loss: 0.004279595334082842
step: 50, loss: 0.03873978182673454
step: 60, loss: 0.001408133888617158
```

```
step: 70, loss: 0.0007796110585331917
step: 80, loss: 0.010728541761636734
step: 90, loss: 0.00027332588797435164
step: 100, loss: 0.0015414234949275851
step: 110, loss: 0.02113087847828865
step: 120, loss: 0.011931680142879486
step: 130, loss: 0.13105171918869019
step: 140, loss: 0.0010927343973889947
step: 150, loss: 0.0015707820421084762
step: 160, loss: 0.0005197270074859262
step: 170, loss: 0.016324061900377274
/usr/local/lib/python3.7/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=0.8155844155844156, f1=0.8134715025906737,
best_f1=0.8129330254041571
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.16222675144672394
step: 10, loss: 0.0012776171788573265
step: 20, loss: 0.0010246215388178825
step: 30, loss: 0.008396302349865437
step: 40, loss: 0.0003472567186690867
step: 50, loss: 0.0008117278339341283
step: 60, loss: 0.0018030498176813126
step: 70, loss: 0.0019673246424645185
step: 80, loss: 0.009222162887454033
step: 90, loss: 0.0023272973485291004
step: 100, loss: 0.12085769325494766
step: 110, loss: 0.09512405842542648
step: 120, loss: 0.0008188595529645681
step: 130, loss: 0.0008745656232349575
step: 140, loss: 0.0010531700681895018
step: 150, loss: 0.008320217952132225
step: 160, loss: 0.008464349433779716
step: 170, loss: 0.000836511084344238
/usr/local/lib/python3.7/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=0.8072916666666666, f1=0.7969543147208121,
best_f1=0.8129330254041571
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0003043690521735698
step: 10, loss: 0.009672467596828938
step: 20, loss: 0.0011109731858596206
```

```
step: 30, loss: 0.0099861295893788834
step: 40, loss: 0.0014839344657957554
step: 50, loss: 0.00062633302639126778
Gradient overflow.  Skipping step, loss scaler 0 reducing loss scale to 8192.0
step: 60, loss: 0.00838529970496893
step: 70, loss: 0.18702952563762665
step: 80, loss: 0.03422393277287483
step: 90, loss: 0.0822332426905632
step: 100, loss: 0.029060132801532745
step: 110, loss: 0.009540011174976826
step: 120, loss: 0.018633518368005753
step: 130, loss: 0.0343262143433094
step: 140, loss: 0.0017065084539353848
step: 150, loss: 0.0014674450503662229
step: 160, loss: 0.006578455679118633
step: 170, loss: 0.0380181148648262
/usr/local/lib/python3.7/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=0.8184019370460048, f1=0.8113207547169812,
best_f1=0.8129330254041571
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.04326873645186424
step: 10, loss: 0.00015787180745974183
step: 20, loss: 0.051748890429735184
step: 30, loss: 0.008997778408229351
step: 40, loss: 0.00011624390026554465
step: 50, loss: 0.17262552678585052
step: 60, loss: 0.00018328802252653986
step: 70, loss: 0.0017491886392235756
step: 80, loss: 0.004454418085515499
step: 90, loss: 0.06744886934757233
step: 100, loss: 0.001197486068122089
step: 110, loss: 0.006633310578763485
step: 120, loss: 0.0023265157360583544
step: 130, loss: 0.009207779541611671
step: 140, loss: 0.0037734273355454206
step: 150, loss: 0.021151622757315636
step: 160, loss: 0.001689208555035293
step: 170, loss: 0.001036317553371191
/usr/local/lib/python3.7/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=0.8203883495145631, f1=0.8, best_f1=0.8129330254041571
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
```

```
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.04515211656689644
step: 10, loss: 0.010617751628160477
step: 20, loss: 0.00020087591838091612
step: 30, loss: 0.00015727621212136
step: 40, loss: 0.0005380847142077982
step: 50, loss: 0.004784903954714537
step: 60, loss: 0.002420746488496661
step: 70, loss: 0.00011761269706767052
step: 80, loss: 0.00030422789859585464
step: 90, loss: 0.004978337325155735
step: 100, loss: 0.0009404636221006513
step: 110, loss: 0.03592713549733162
step: 120, loss: 0.0014014277840033174
step: 130, loss: 0.00020036657224409282
step: 140, loss: 0.0008130468777380884
step: 150, loss: 0.0128707280382514
step: 160, loss: 0.0015073306858539581
step: 170, loss: 0.00228407746180892
/usr/local/lib/python3.7/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=0.8211586901763224, f1=0.7971014492753623,
best_f1=0.8129330254041571
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.004751875530928373
step: 10, loss: 0.0005871678586117923
step: 20, loss: 0.0002678988385014236
step: 30, loss: 0.00016327921184711158
step: 40, loss: 0.0002783870149869472
step: 50, loss: 0.00013603994739241898
step: 60, loss: 0.00019041210180148482
step: 70, loss: 0.04788631945848465
step: 80, loss: 9.88426327239722e-05
step: 90, loss: 0.005939730908721685
step: 100, loss: 0.0002250838588224724
step: 110, loss: 0.00023023271933197975
step: 120, loss: 0.06071215495467186
step: 130, loss: 0.005258081946521997 5
step: 140, loss: 0.0028118968475610018
step: 150, loss: 0.0016077762702479959
step: 160, loss: 0.00017984426813200116
step: 170, loss: 0.00016920457710511982
/usr/local/lib/python3.7/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=0.8157894736842105, f1=0.7949367088607596,
best_f1=0.8129330254041571
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.02444756217300892
step: 10, loss: 0.0010728344786912203
step: 20, loss: 0.00020279565069358796
step: 30, loss: 0.00046663818648085
step: 40, loss: 0.00025566943804733455
step: 50, loss: 0.0005802965024486184
step: 60, loss: 0.0007531403098255396
step: 70, loss: 0.0002972372167278081
step: 80, loss: 0.00018428079783916473
step: 90, loss: 0.0003850380307994783
step: 100, loss: 0.0002818226639647037
step: 110, loss: 0.0003974397259298712
step: 120, loss: 0.042294442653656006
step: 130, loss: 0.00012994515418540686
step: 140, loss: 0.00012900651199743152
step: 150, loss: 0.04766729101538658
step: 160, loss: 0.00033674505539238453
step: 170, loss: 0.004154559224843979
/usr/local/lib/python3.7/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=0.8142493638676845, f1=0.7941176470588236,
best_f1=0.8129330254041571
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.000335999095113948
step: 10, loss: 0.0001297970156883821
step: 20, loss: 0.0005137301050126553
step: 30, loss: 0.0007824734784662724
step: 40, loss: 0.00015044429164845496
step: 50, loss: 0.000120554119348526
step: 60, loss: 7.899254705989733e-05
step: 70, loss: 0.0007737100822851062
step: 80, loss: 0.000628872134257108
step: 90, loss: 0.0016268916660919785
step: 100, loss: 0.007363116834312677
step: 110, loss: 0.0019394161645323038
step: 120, loss: 0.01666201464831829
step: 130, loss: 0.033323001116514206
step: 140, loss: 8.721756603335962e-05
step: 150, loss: 0.00029534229543060064
step: 160, loss: 9.345356374979019e-05
```

step: 170, loss: 0.0013663643039762974
/usr/local/lib/python3.7/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=0.8162162162162162, f1=0.793733681462141,
best_f1=0.8129330254041571
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0001383692870149389
step: 10, loss: 0.0010683056898415089
step: 20, loss: 0.02758750505745411
step: 30, loss: 0.002419780008494854
step: 40, loss: 0.0018143674824386835
step: 50, loss: 0.0006946674548089504
step: 60, loss: 0.006658510304987431
step: 70, loss: 0.00010276463581249118
step: 80, loss: 0.04137443006038666
step: 90, loss: 0.005365755409002304
step: 100, loss: 0.0009237434132955968
step: 110, loss: 0.0001783553307177499
step: 120, loss: 0.0006922970642335713
step: 130, loss: 0.0011897677322849631
step: 140, loss: 0.0002692330745048821
step: 150, loss: 0.00045432886690832675
step: 160, loss: 0.00013559585204347968
step: 170, loss: 0.00042637038859538734
/usr/local/lib/python3.7/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=0.8164383561643836, f1=0.804289544235925,
best_f1=0.8129330254041571
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, 275.55it/s]
load_f1 = 0.3864915572232645
real_f1 = 0.36462093862815886
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 [00:17, 249.11it/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 bert-base-uncased \
        --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 bert-base-uncased \
  --use_gpu \
  --fp16 \
  --checkpoint_path checkpoints/
```

Downloading: 100% 433/433 [00:00<00:00, 436kB/s]
Downloading: 100% 232k/232k [00:00<00:00, 912kB/s]
Downloading: 100% 440M/440M [00:08<00:00, 51.4MB/s]
Selected optimization level O2: FP16 training with FP32 batchnorm and FP32
master weights.

Defaults for this optimization level are:
enabled   : True
opt_level   : O2
cast_model_type   : torch.float16
patch_torch_functions : False
keep_batchnorm_fp32   : True
master_weights   : True
loss_scale   : dynamic
Processing user overrides (additional kwargs that are not None)…
After processing overrides, optimization options are:
enabled   : True
opt_level   : O2
cast_model_type   : torch.float16
patch_torch_functions : False
keep_batchnorm_fp32   : True
master_weights   : True
loss_scale   : dynamic
Warning: multi_tensor_applier fused unscale kernel is unavailable, possibly
because apex was installed without --cuda_ext --cpp_ext. Using Python fallback.
Original ImportError was: ModuleNotFoundError("No module named 'amp_C'")
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 32768.0
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:136:
UserWarning: Detected call of `lr_scheduler.step()` before `optimizer.step()`.
In PyTorch 1.1.0 and later, you should call them in the opposite order:
`optimizer.step()` before `lr_scheduler.step()`. Failure to do this will result
in PyTorch skipping the first value of the learning rate schedule. See more
details at https://pytorch.org/docs/stable/optim.html#how-to-adjust-learning-

```
rate
  "https://pytorch.org/docs/stable/optim.html#how-to-adjust-learning-rate",
UserWarning)
step: 0, loss: 0.6328253746032715
step: 10, loss: 0.6103754639625549
step: 20, loss: 0.3856714963912964
step: 30, loss: 0.1393957883119583
step: 40, loss: 0.1765221804380417
step: 50, loss: 0.10389523208141327
step: 60, loss: 0.11553406715393066
step: 70, loss: 0.026385825127363205
step: 80, loss: 0.051389988511800766
step: 90, loss: 0.09927055984735489
step: 100, loss: 0.0071890028193593025
step: 110, loss: 0.26556140184402466
step: 120, loss: 0.023457661271095276
step: 130, loss: 0.01379297487437725
step: 140, loss: 0.00209015398286283
step: 150, loss: 0.02370705083012581
step: 160, loss: 0.004031467251479626
step: 170, loss: 0.20166727900505066
step: 180, loss: 0.03046809509396553
step: 190, loss: 0.052427250891923904
step: 200, loss: 0.06157727539539337
step: 210, loss: 0.007041556295007467
step: 220, loss: 0.020971043035387993
step: 230, loss: 0.011473480612039566
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9668141592920354, f1=0.9597315436241611,
best_f1=0.9597315436241611
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.01792488433420658
step: 10, loss: 0.00782647356390953
step: 20, loss: 0.16014230251312256
step: 30, loss: 0.1593785434961319
step: 40, loss: 0.07365930080413818
step: 50, loss: 0.005195092409849167
step: 60, loss: 0.013420111499726772
step: 70, loss: 0.0938880443572998
step: 80, loss: 0.008678065612912178
```

```
step: 90, loss: 0.0027068473864346743
step: 100, loss: 0.1329144686460495
step: 110, loss: 0.1314181387424469
step: 120, loss: 0.035921379923820496
step: 130, loss: 0.016445208340883255
step: 140, loss: 0.000814572034869343
step: 150, loss: 0.014821207150816917
step: 160, loss: 0.03506682068109512
step: 170, loss: 0.0012440549908205867
step: 180, loss: 0.07782574743032455
step: 190, loss: 0.0033408936578780413
step: 200, loss: 0.003287007100880146
step: 210, loss: 0.0009351181797683239
step: 220, loss: 0.09925171732902527
step: 230, loss: 0.008618604391813278
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9767441860465117, 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: 0.010199536569416523
step: 10, loss: 0.007129963953047991
step: 20, loss: 0.005501138512045145
step: 30, loss: 0.014088592492043972
step: 40, loss: 0.047920677810907364
step: 50, loss: 0.01934928074479103
step: 60, loss: 0.009367890655994415
step: 70, loss: 0.0019946449901908636
step: 80, loss: 0.0024828696623444557
step: 90, loss: 0.021068844944238663
step: 100, loss: 0.0029837689362466335
step: 110, loss: 0.0005984138115309179
step: 120, loss: 0.017453908920288086
step: 130, loss: 0.00031821487937122583
step: 140, loss: 0.0030783924739807844
step: 150, loss: 0.019322263076901436
step: 160, loss: 0.007799824699759483
step: 170, loss: 0.0035058713983744383
step: 180, loss: 0.01648951880633831
step: 190, loss: 0.0037169770803302526
step: 200, loss: 0.07820051908493042
```

```
step: 210, loss: 0.009687347337603569
step: 220, loss: 0.004310948774218559
step: 230, loss: 0.018410498276352882
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9788182831661093, f1=0.9696969696969697,
best_f1=0.9696969696969697
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0010201437398791313
step: 10, loss: 0.005887189880013466
step: 20, loss: 0.0056731509976089
step: 30, loss: 0.0009044689359143376
step: 40, loss: 0.002252852777019143
step: 50, loss: 0.0006214429158717394
step: 60, loss: 0.0009372188942506909
step: 70, loss: 0.0004170264001004398
step: 80, loss: 0.000997008290141821
step: 90, loss: 0.005551119334995747
step: 100, loss: 0.0012163883075118065
step: 110, loss: 0.0009063984616659582
step: 120, loss: 0.004430186469107866
step: 130, loss: 0.001011249958537519
step: 140, loss: 0.0004732070374302566
step: 150, loss: 0.2270107865333557
step: 160, loss: 0.020377403125166893
step: 170, loss: 0.0384901687502861
step: 180, loss: 0.0007493505254387856
step: 190, loss: 0.001524322316981852
step: 200, loss: 0.0036994870752096176
step: 210, loss: 0.14176134765148163
step: 220, loss: 0.0004991292953491211
step: 230, loss: 0.004736266098916531
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9843749999999999, f1=0.980963045912654,
best_f1=0.980963045912654
```

```
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0028591982554644346
step: 10, loss: 0.00064756587380543335
step: 20, loss: 0.0015521567547693849
step: 30, loss: 0.00019148980209138244
step: 40, loss: 0.0007148428703658283
step: 50, loss: 0.0025109443813562393
step: 60, loss: 0.076754130423069
step: 70, loss: 0.0007034274167381227
step: 80, loss: 0.0007096272311173379
step: 90, loss: 0.0023137000389397144
step: 100, loss: 0.0005060555413365364
step: 110, loss: 0.0003047799982596189
step: 120, loss: 0.00022395639098249376
step: 130, loss: 0.018499305471777916
step: 140, loss: 0.0017306567169725895
step: 150, loss: 0.0008509734179824591
step: 160, loss: 0.00044459925265982747
step: 170, loss: 0.0004343627952039242
step: 180, loss: 0.00022754189558327198
step: 190, loss: 0.05608052387833595
step: 200, loss: 0.009347678162157536
step: 210, loss: 0.0018621011404320598
step: 220, loss: 0.0009226193651556969
step: 230, loss: 0.0021627454552799463
/usr/local/lib/python3.7/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=0.9808773903262092, f1=0.9763779527559054,
best_f1=0.980963045912654
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0099926574304699898
step: 10, loss: 0.0037944887299090624
step: 20, loss: 0.0008569427882321179
step: 30, loss: 0.00018839156837202609
step: 40, loss: 0.0002023469569394365
step: 50, loss: 0.00014298844325821847
step: 60, loss: 0.00019478274043649435
step: 70, loss: 0.0005304620135575533
step: 80, loss: 0.003683892311528325
step: 90, loss: 0.0002481688861735165
step: 100, loss: 0.021367738023400307
step: 110, loss: 0.0009986136574298143
step: 120, loss: 0.014021213166415691
```

```
step: 130, loss: 0.0032282527536153793
step: 140, loss: 0.0006201882497407496
step: 150, loss: 0.0676233321428299
step: 160, loss: 0.0012558333110064268
step: 170, loss: 0.0006288660224527121
step: 180, loss: 0.007629347499459982
step: 190, loss: 0.00039851327892392874
step: 200, loss: 0.0002256120351376012
step: 210, loss: 0.0001955215266207233
Gradient overflow.  Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 220, loss: 0.00017003185348585248
step: 230, loss: 0.06634372472763062
/usr/local/lib/python3.7/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=0.9810055865921787, f1=0.9732142857142857,
best_f1=0.980963045912654
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0003179889463353902
step: 10, loss: 0.0035470614675432444
step: 20, loss: 0.0010406909277662635
step: 30, loss: 0.0015386033337563276
step: 40, loss: 0.00040914476267062133
step: 50, loss: 0.00016944303933996707
step: 60, loss: 0.00018353150517214099
step: 70, loss: 0.019650749862194006
step: 80, loss: 0.0017972884234040976
step: 90, loss: 0.0001796460710465908
step: 100, loss: 0.00017409160500392318
step: 110, loss: 0.033816803246736526
step: 120, loss: 0.00011286497465334833
step: 130, loss: 0.00017826406110543758
step: 140, loss: 0.0003363180148880929
step: 150, loss: 0.006154491566121578
step: 160, loss: 0.06445659697055817
step: 170, loss: 0.0009090415551327169
step: 180, loss: 0.0016777014825493097
step: 190, loss: 0.0003110935213044286
step: 200, loss: 0.05023506283760071
step: 210, loss: 8.921284461393952e-05
step: 220, loss: 0.00028760521672666073
step: 230, loss: 0.00110783486161381
/usr/local/lib/python3.7/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=0.9797752808988766, f1=0.9808342728297633,
```

```
best_f1=0.980963045912654
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00020961380505468696
step: 10, loss: 0.0007864487706683576
step: 20, loss: 0.00016995955957099795
step: 30, loss: 0.000157848815433681
step: 40, loss: 0.0009129546815529466
step: 50, loss: 0.0005225837812758982
step: 60, loss: 0.00013474187289830297
step: 70, loss: 0.00031297776149585843
step: 80, loss: 0.00016347142809536308
step: 90, loss: 0.00018078525317832828
step: 100, loss: 0.00039657403249293566
step: 110, loss: 0.001632506842724979
step: 120, loss: 0.014970698393881321
step: 130, loss: 0.000407403422286734
step: 140, loss: 0.00011030719906557351
step: 150, loss: 0.000184784788871184
step: 160, loss: 0.00030136900022625923
step: 170, loss: 0.00020341775962151587
step: 180, loss: 0.0010131543967872858
step: 190, loss: 0.0003966558142565191
step: 200, loss: 0.0003735109930858016
step: 210, loss: 0.0027247900143265724
step: 220, loss: 0.00017812916485127062
step: 230, loss: 0.00027660245541483164
/usr/local/lib/python3.7/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=0.9798657718120806, f1=0.9720670391061451,
best_f1=0.980963045912654
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 5.050940308137797e-05
step: 10, loss: 0.0006292357575148344
step: 20, loss: 0.0006247548153623939
step: 30, loss: 0.00010151943570235744
step: 40, loss: 0.015263205394148827
step: 50, loss: 0.00010198281233897433
step: 60, loss: 0.0001157305741799064
step: 70, loss: 0.0003742183034773916
step: 80, loss: 8.476443326799199e-05
step: 90, loss: 0.0003033253306057304
step: 100, loss: 8.836662163957953e-05
step: 110, loss: 4.854464714298956e-05
```

```
step: 120, loss: 5.6970686273416504e-05
step: 130, loss: 0.00012568301463034004
step: 140, loss: 6.758780364179984e-05
step: 150, loss: 0.0007888160180300474
step: 160, loss: 0.00010412138362880796
step: 170, loss: 8.265528595075011e-05
step: 180, loss: 0.0003145733498968184
step: 190, loss: 7.140123489079997e-05
step: 200, loss: 0.02908698283135891
step: 210, loss: 0.00031279344693757594
step: 220, loss: 0.00015698802599217743
step: 230, loss: 0.0007589261513203382
/usr/local/lib/python3.7/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=0.9693530079455165, f1=0.9670079635949943,
best_f1=0.980963045912654
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.004629859700798988
step: 10, loss: 0.00010740024299593642
step: 20, loss: 0.01873115636408329
step: 30, loss: 0.007914287969470024
step: 40, loss: 0.00011224202899029478
step: 50, loss: 0.00015288361464627087
step: 60, loss: 0.00028402928728610277
step: 70, loss: 0.00026881915982812643
step: 80, loss: 8.747303218115121e-05
step: 90, loss: 0.00023022257664706558
step: 100, loss: 0.00015536582213826478
step: 110, loss: 0.00014835991896688938
step: 120, loss: 0.00013886881060898304
step: 130, loss: 0.0001272500230697915
step: 140, loss: 0.029072681441903114
step: 150, loss: 0.0181154552847147
step: 160, loss: 0.00013514513557311147
step: 170, loss: 0.00012425243039615452
step: 180, loss: 0.00023261395108420402
step: 190, loss: 0.0036469323094934225
step: 200, loss: 7.271322829183191e-05
step: 210, loss: 0.0001451301359338686
Gradient overflow.  Skipping step, loss scaler 0 reducing loss scale to 8192.0
step: 220, loss: 0.00010202536213910207
step: 230, loss: 0.00027659436454996467
/usr/local/lib/python3.7/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=0.9777777777777777, f1=0.966740576496674,
best_f1=0.980963045912654
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 5.602574674412608e-05
step: 10, loss: 0.00018021016148850322
step: 20, loss: 0.00035723194014281034
step: 30, loss: 0.009828047826886177
step: 40, loss: 0.0002799040521495044
step: 50, loss: 0.0003114357532467693
step: 60, loss: 0.003737029619514942
step: 70, loss: 6.803627911722288e-05
step: 80, loss: 4.711929432232864e-05
step: 90, loss: 6.887591734994203e-05
step: 100, loss: 6.47140623186715e-05
step: 110, loss: 0.00028062055935151875
step: 120, loss: 0.00040006765630096197
step: 130, loss: 0.0006111276452429593
step: 140, loss: 0.00010312644735677168
step: 150, loss: 0.0005037516821175814
step: 160, loss: 0.0001942618255270645
step: 170, loss: 9.726430289447308e-05
step: 180, loss: 0.00014346730313263834
step: 190, loss: 7.044243830023333e-05
step: 200, loss: 0.001570221851579845
step: 210, loss: 6.169363769004121e-05
step: 220, loss: 6.233790918486193e-05
step: 230, loss: 0.00011906477448064834
/usr/local/lib/python3.7/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=0.9755555555555556, f1=0.972129319955407,
best_f1=0.980963045912654
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00012736314965877682
step: 10, loss: 5.9767611674033105e-05
step: 20, loss: 0.0003135038714390248
step: 30, loss: 0.009522058069705963
step: 40, loss: 0.00022044072102289647
step: 50, loss: 0.008664265275001526
step: 60, loss: 0.00030162956682033837
step: 70, loss: 0.05599142983555794
step: 80, loss: 0.000851520337164402
step: 90, loss: 0.0003932616382371634
step: 100, loss: 3.8677975680911914e-05
```

```
step: 110, loss: 0.0001750697410898283
step: 120, loss: 0.00011925496073672548
step: 130, loss: 7.831188122509047e-05
step: 140, loss: 0.00013540392683353275
step: 150, loss: 0.0011816996848210692
step: 160, loss: 0.00012989800598006696
step: 170, loss: 0.00011658173752948642
step: 180, loss: 0.010694917291402817
step: 190, loss: 0.0012424761662259698
step: 200, loss: 3.991057019447908e-05
step: 210, loss: 6.220045906957239e-05
step: 220, loss: 0.00018148228991776705
step: 230, loss: 0.037205133587121964
/usr/local/lib/python3.7/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=0.9810901001112348, f1=0.9721913236929923,
best_f1=0.980963045912654
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00021827539603691548
step: 10, loss: 0.00015335493662860245
step: 20, loss: 0.00026702918694354594
step: 30, loss: 0.00010478722106199712
step: 40, loss: 0.00015158338646870106
step: 50, loss: 0.00013600016245618463
step: 60, loss: 9.782104461919516e-05
step: 70, loss: 7.262768485816196e-05
step: 80, loss: 8.315870945807546e-05
step: 90, loss: 0.00013365919585339725
step: 100, loss: 7.647228630958125e-05
step: 110, loss: 0.007331752683967352
step: 120, loss: 0.0029432105366140604
step: 130, loss: 0.00013222689449321479
step: 140, loss: 7.683486910536885e-05
step: 150, loss: 7.725120667601004e-05
step: 160, loss: 0.0001354548439849168
step: 170, loss: 0.00012466609684452415
step: 180, loss: 0.00010163323167944327
step: 190, loss: 0.00010486526298336685
step: 200, loss: 0.00024791524629108611
step: 210, loss: 3.591406129999086e-05
step: 220, loss: 5.959095506113954e-05
step: 230, loss: 0.0004647393070627004
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
```

```
epoch 13: dev_f1=0.98, f1=0.9700332963374029, best_f1=0.980963045912654
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 6.924672925379127e-05
step: 10, loss: 0.00039316085167229176
step: 20, loss: 0.000180906688910909
step: 30, loss: 0.00010733016824815422
step: 40, loss: 0.036153942346572876
step: 50, loss: 0.00012708500435110182
step: 60, loss: 6.877809209981933e-05
step: 70, loss: 6.0997397667961195e-05
step: 80, loss: 6.352281343424693e-05
step: 90, loss: 8.134428935591131e-05
step: 100, loss: 5.231839168118313e-05
step: 110, loss: 0.0004044450761284679
step: 120, loss: 0.00013691715139430016
step: 130, loss: 7.008054672041908e-05
step: 140, loss: 9.600788325769827e-05
step: 150, loss: 5.1364120736252517e-05
step: 160, loss: 0.0048683746717870235
step: 170, loss: 5.644865086651407e-05
step: 180, loss: 7.7332035289146e-05
step: 190, loss: 0.0005010411259718239
step: 200, loss: 6.230977305676788e-05
step: 210, loss: 0.0012249964056536555
step: 220, loss: 8.422701648669317e-05
step: 230, loss: 0.0003696857311297208
/usr/local/lib/python3.7/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=0.98, f1=0.9732739420935412, best_f1=0.980963045912654
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 4.602186527336016e-05
step: 10, loss: 7.605290738865733e-05
step: 20, loss: 6.7739216319751e-05
step: 30, loss: 9.182745270663872e-05
step: 40, loss: 0.0002995871764142066
step: 50, loss: 0.02182852104306221
step: 60, loss: 9.539919847156852e-05
step: 70, loss: 6.849571946077049e-05
step: 80, loss: 0.00011994559463346377
step: 90, loss: 0.00011178627755725756
step: 100, loss: 6.249373109312728e-05
step: 110, loss: 0.0001728535135043785
step: 120, loss: 6.427001062547788e-05
```

```
step: 130, loss: 0.00012016308028250933
step: 140, loss: 5.535289164981805e-05
step: 150, loss: 0.00022514123702421784
step: 160, loss: 3.758602906600572e-05
step: 170, loss: 3.2021293009165674e-05
step: 180, loss: 0.0008234437555074692
step: 190, loss: 0.0007194034405983984
step: 200, loss: 0.00012264354154467583
step: 210, loss: 9.003075683722273e-05
step: 220, loss: 0.0002166960621252656
step: 230, loss: 6.168849358800799e-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=0.9777777777777777, f1=0.9733333333333333,
best_f1=0.980963045912654
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.")
2473it [00:13, 188.91it/s]
load_f1 = 0.9821428571428571
```

```
real_f1 = 0.9821428571428571
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 [00:19, 225.41it/s]
```

1.3.2 DBLP-GoogleScholar - Running the matcher

```python
# 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 bert-base-uncased \
  --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 bert-base-uncased \
  --use_gpu \
  --fp16 \
  --checkpoint_path checkpoints/
```

```
Downloading: 100% 433/433 [00:00<00:00, 504kB/s]
Downloading: 100% 232k/232k [00:00<00:00, 1.22MB/s]
Downloading: 100% 440M/440M [00:11<00:00, 37.0MB/s]
Selected optimization level O2:  FP16 training with FP32 batchnorm and FP32
master weights.

Defaults for this optimization level are:
```

```
enabled                 : True
opt_level               : O2
cast_model_type         : torch.float16
patch_torch_functions   : False
keep_batchnorm_fp32     : True
master_weights          : True
loss_scale              : dynamic
Processing user overrides (additional kwargs that are not None)…
After processing overrides, optimization options are:
enabled                 : True
opt_level               : O2
cast_model_type         : torch.float16
patch_torch_functions   : False
keep_batchnorm_fp32     : True
master_weights          : True
loss_scale              : dynamic
```

Warning: multi_tensor_applier fused unscale kernel is unavailable, possibly because apex was installed without --cuda_ext --cpp_ext. Using Python fallback. Original ImportError was: ModuleNotFoundError("No module named 'amp_C'")
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 32768.0
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:136: UserWarning: Detected call of `lr_scheduler.step()` before `optimizer.step()`. In PyTorch 1.1.0 and later, you should call them in the opposite order: `optimizer.step()` before `lr_scheduler.step()`. Failure to do this will result in PyTorch skipping the first value of the learning rate schedule. See more details at https://pytorch.org/docs/stable/optim.html#how-to-adjust-learning-rate
  "https://pytorch.org/docs/stable/optim.html#how-to-adjust-learning-rate", UserWarning)

```
step: 0, loss: 0.620211660861969
step: 10, loss: 0.5198891162872314
step: 20, loss: 0.45027828216552734
step: 30, loss: 0.047647327184677124
step: 40, loss: 0.1682967096567154
step: 50, loss: 0.09712085127830505
step: 60, loss: 0.05414538457989693
step: 70, loss: 0.09050697833299637
step: 80, loss: 0.07514294236898422
step: 90, loss: 0.35679978132247925
step: 100, loss: 0.013281605206429958
step: 110, loss: 0.03360165283083916
step: 120, loss: 0.13389691710472107
step: 130, loss: 0.03440094739198685
step: 140, loss: 0.06585576385259628
step: 150, loss: 0.04540632292628288
```

```
step: 160, loss: 0.10813866555690765
step: 170, loss: 0.1983003467321396
step: 180, loss: 0.05463750287890434
step: 190, loss: 0.01997894048690796
step: 200, loss: 0.29899823665618896
step: 210, loss: 0.0779263824224472
step: 220, loss: 0.15488319098949432
step: 230, loss: 0.15293848514556885
step: 240, loss: 0.07385851442813873
step: 250, loss: 0.01021682471036911
step: 260, loss: 0.024034777656197548
step: 270, loss: 0.011711493134498596
step: 280, loss: 0.043403640389442444
step: 290, loss: 0.1181909367442131
step: 300, loss: 0.021355193108320236
step: 310, loss: 0.2815183401107788
step: 320, loss: 0.11297283321619034
step: 330, loss: 0.032392971217632294
step: 340, loss: 0.034369587898254395
step: 350, loss: 0.011839590966701508
step: 360, loss: 0.024117687717080116
step: 370, loss: 0.09739982336759567
step: 380, loss: 0.03245289623737335
step: 390, loss: 0.1633678376674652
step: 400, loss: 0.22218602895736694
step: 410, loss: 0.029796451330184937
step: 420, loss: 0.029682034626603127
step: 430, loss: 0.15330614149570465
step: 440, loss: 0.03440922871232033
step: 450, loss: 0.015446494333446026
step: 460, loss: 0.004424413666129112
step: 470, loss: 0.1311386376619339
step: 480, loss: 0.14359542727470398
step: 490, loss: 0.16655930876731873
step: 500, loss: 0.04612942785024643
step: 510, loss: 0.08438608795404434
step: 520, loss: 0.14986580610275269
step: 530, loss: 0.01488204300403595
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9405255878284925, f1=0.933579335793358,
best_f1=0.933579335793358
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
```

```
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.05233093723654747
step: 10, loss: 0.03890660032629967
step: 20, loss: 0.14259470999240875
step: 30, loss: 0.07224797457456589
step: 40, loss: 0.15989993512630463
step: 50, loss: 0.14552603662014008
step: 60, loss: 0.0031328722834587097
step: 70, loss: 0.02255094237625599
step: 80, loss: 0.13471809029579163
step: 90, loss: 0.01982971280813217
step: 100, loss: 0.024042408913373947
step: 110, loss: 0.06167444586753845
step: 120, loss: 0.08233251422643661
step: 130, loss: 0.10822433978319168
step: 140, loss: 0.050933413207530975
step: 150, loss: 0.05634232982993126
step: 160, loss: 0.010077916085720062
step: 170, loss: 0.015591628849506378
step: 180, loss: 0.023225123062729836
step: 190, loss: 0.05839761346578598
step: 200, loss: 0.002501139650121331
step: 210, loss: 0.013156593777239323
step: 220, loss: 0.09003039449453354
step: 230, loss: 0.005798471625894308
step: 240, loss: 0.05563482642173767
step: 250, loss: 0.07933655381202698
step: 260, loss: 0.003773412201553583
step: 270, loss: 0.20738691091537476
step: 280, loss: 0.015832656994462013
step: 290, loss: 0.11436371505260468
step: 300, loss: 0.15776051580905914
step: 310, loss: 0.02166842855513096
step: 320, loss: 0.09662068635225296
step: 330, loss: 0.08461590111255646
step: 340, loss: 0.031808797270059586
step: 350, loss: 0.0019538283813744783
step: 360, loss: 0.08465400338172913
step: 370, loss: 0.0708480179309845
step: 380, loss: 0.04469585791230202
step: 390, loss: 0.07389258593320847
step: 400, loss: 0.01096771378070116
step: 410, loss: 0.030044084414839745
step: 420, loss: 0.010866274125874043
step: 430, loss: 0.028429223224520683
step: 440, loss: 0.19941022992134094
step: 450, loss: 0.046408236026763916
```

```
step: 460, loss: 0.051631860435009
step: 470, loss: 0.09856204688549042
step: 480, loss: 0.3055647015571594
step: 490, loss: 0.009990720078349113
step: 500, loss: 0.23805655539035797
step: 510, loss: 0.02696511335670948
step: 520, loss: 0.12696263194084167
step: 530, loss: 0.0097251953555713367
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9481000926784059, f1=0.9450853714813106,
best_f1=0.9450853714813106
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.11234118044376373
step: 10, loss: 0.04636390507221222
step: 20, loss: 0.12940388917922974
step: 30, loss: 0.0727221816778183
step: 40, loss: 0.08530095964670181
step: 50, loss: 0.15657086670398712
step: 60, loss: 0.08394121378660202
step: 70, loss: 0.009101087227463722
step: 80, loss: 0.006055274046957493
step: 90, loss: 0.00598114775493741
step: 100, loss: 0.11693280935287476
step: 110, loss: 0.0046592191793022454
step: 120, loss: 0.0252920500934124
step: 130, loss: 0.006570320576429367
step: 140, loss: 0.03538621962070465
step: 150, loss: 0.01615666225552559
step: 160, loss: 0.009945645928382874
step: 170, loss: 0.07527970522642136
step: 180, loss: 0.006470350082963705
step: 190, loss: 0.017423491925001144
step: 200, loss: 0.11826492846012115
step: 210, loss: 0.06324555724859238
step: 220, loss: 0.08644970506429672
step: 230, loss: 0.11784248054027557
step: 240, loss: 0.010364744812250137
step: 250, loss: 0.06098647788167
step: 260, loss: 0.012140237726271152
step: 270, loss: 0.04539098963141441
```

```
step: 280, loss: 0.12159612774848938
step: 290, loss: 0.005863488651812077
step: 300, loss: 0.0076057445257902145
step: 310, loss: 0.006593648809939623
step: 320, loss: 0.00770585285499692
step: 330, loss: 0.006333013065159321
step: 340, loss: 0.006145872641354799
step: 350, loss: 0.0010060049826279283
step: 360, loss: 0.04693809151649475
step: 370, loss: 0.04633662477135658
step: 380, loss: 0.05419714003801346
step: 390, loss: 0.13953787088394165
step: 400, loss: 0.022263377904891968
step: 410, loss: 0.005117344204336405
step: 420, loss: 0.14837293326854706
step: 430, loss: 0.03004787489771843
step: 440, loss: 0.0876847505569458
step: 450, loss: 0.08346634358167648
step: 460, loss: 0.06339360028505325
step: 470, loss: 0.03136089816689491
step: 480, loss: 0.010180804878473282
step: 490, loss: 0.016000838950276375
step: 500, loss: 0.010124248452484608
step: 510, loss: 0.002556161256507039
step: 520, loss: 0.2068503201007843
step: 530, loss: 0.035690803080797195
/usr/local/lib/python3.7/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=0.9480093676814988, f1=0.9351376574895007,
best_f1=0.9450853714813106
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.007651783525943756
step: 10, loss: 0.007747984491288662
step: 20, loss: 0.003798985853791237
step: 30, loss: 0.005484074354171753
step: 40, loss: 0.0006829002522863448
step: 50, loss: 0.0024147944059222937
step: 60, loss: 0.11154317855834961
step: 70, loss: 0.010034816339612007
step: 80, loss: 0.02163839153945446
step: 90, loss: 0.01650879718363285
step: 100, loss: 0.01368021685630083
step: 110, loss: 0.026085462421178818
step: 120, loss: 0.0002282242785440758
step: 130, loss: 0.011059500277042389
```

```
step: 140, loss: 0.0008595511317253113
step: 150, loss: 0.010360253043472767
step: 160, loss: 0.13086742162704468
step: 170, loss: 0.0028291363269090652
step: 180, loss: 0.002287530107423663
step: 190, loss: 0.00222400831989944
step: 200, loss: 0.04207802191376686
step: 210, loss: 0.00724708940833807
step: 220, loss: 0.0020380914211273193
step: 230, loss: 0.13252609968185425
step: 240, loss: 0.020757384598255157
step: 250, loss: 0.038554392755031586
step: 260, loss: 0.0910162478685379
step: 270, loss: 0.01984843611717224
step: 280, loss: 0.02860099822282791
Gradient overflow.  Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 290, loss: 0.15569733083248138
step: 300, loss: 0.000856010417919606
step: 310, loss: 0.00472077913582325
step: 320, loss: 0.04289701581001282
step: 330, loss: 0.03074340894818306
step: 340, loss: 0.1381680965423584
step: 350, loss: 0.024507543072104454
step: 360, loss: 0.017858272418379784
step: 370, loss: 0.0031684692949056625
step: 380, loss: 0.0039327326230704784
step: 390, loss: 0.0035052443854510784
step: 400, loss: 0.006329922936856747
step: 410, loss: 0.017280766740441322
step: 420, loss: 0.005191698670387268
step: 430, loss: 0.035024404525756836
step: 440, loss: 0.0018705211114138365
step: 450, loss: 0.002539150882512331
step: 460, loss: 0.000718166003935039
step: 470, loss: 0.016735989600419998
step: 480, loss: 0.18793798983097076
step: 490, loss: 0.006689141504466534
step: 500, loss: 0.03681725636124611
step: 510, loss: 0.040876604616642
step: 520, loss: 0.0719016045331955
step: 530, loss: 0.012250687927007675
/usr/local/lib/python3.7/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=0.9475638051044084, f1=0.9313543599257884,
best_f1=0.9450853714813106
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
```

```
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.15635819733142853
step: 10, loss: 0.06100168824195862
step: 20, loss: 0.021887755021452904
step: 30, loss: 0.004180803429335356
step: 40, loss: 0.10453283041715622
step: 50, loss: 0.0015420898562297225
step: 60, loss: 0.08831655234098434
step: 70, loss: 0.008313958533108234
step: 80, loss: 0.0028694516513496637
step: 90, loss: 0.007013029418885708
step: 100, loss: 0.024081232026219368
step: 110, loss: 0.0008686897926963866
step: 120, loss: 0.006341390777379274
step: 130, loss: 0.0006233368767425418
step: 140, loss: 0.002528228797018528
step: 150, loss: 0.011857670731842518
step: 160, loss: 0.0013863000785931945
step: 170, loss: 0.015474239364266396
step: 180, loss: 0.0007275070529431105
step: 190, loss: 0.004140516743063927
step: 200, loss: 0.0013800282031297684
step: 210, loss: 0.004275189712643623
step: 220, loss: 0.02035343274474144
step: 230, loss: 0.03688102215528488
step: 240, loss: 0.003670687787234783
step: 250, loss: 0.002661708975210786
step: 260, loss: 0.006824732292443514
step: 270, loss: 0.0006044851616024971
step: 280, loss: 0.003905388992279768
step: 290, loss: 0.1552412062883377
step: 300, loss: 0.030453016981482506
step: 310, loss: 0.0012260830262675881
step: 320, loss: 0.11858335137367249
step: 330, loss: 0.014666816219687462
step: 340, loss: 0.004011617973446846
step: 350, loss: 0.00295799458399415
step: 360, loss: 0.0035473997704684734
step: 370, loss: 0.016402091830968857
step: 380, loss: 0.0016930276760831475
step: 390, loss: 0.0002879010571632534
step: 400, loss: 0.006458886433392763
step: 410, loss: 0.02234538644552231
step: 420, loss: 0.0035137375816702843
step: 430, loss: 0.0037326260935515165
step: 440, loss: 0.013662921264767647
step: 450, loss: 0.01216923538595438
step: 460, loss: 0.0008224134799093008
```

```
step: 470, loss: 0.0009308111621066928
step: 480, loss: 0.006741562392562628
step: 490, loss: 0.001164212473668158
step: 500, loss: 0.013335928320884705
step: 510, loss: 0.16832800209522247
step: 520, loss: 0.0021908460184931755
step: 530, loss: 0.014894774183630943
/usr/local/lib/python3.7/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=0.9397477814105558, f1=0.9335810496980957,
best_f1=0.9450853714813106
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0020952897612005472
step: 10, loss: 0.0001165009816759266
step: 20, loss: 0.1167994812130928
step: 30, loss: 0.005653956905007362
step: 40, loss: 0.033457208424806595
step: 50, loss: 0.025357132777571678
step: 60, loss: 0.00088995473925024227
step: 70, loss: 0.004138217307627201
step: 80, loss: 0.008831402286887169
step: 90, loss: 0.010356949642300606
step: 100, loss: 0.06344641000032425
step: 110, loss: 0.0012862951261922717
step: 120, loss: 0.004321553744375706
step: 130, loss: 0.031562160700559616
step: 140, loss: 0.0075373174622654915
step: 150, loss: 0.0009218697668984532
step: 160, loss: 0.0014183638850226998
step: 170, loss: 0.00933071505278349
step: 180, loss: 0.0027152537368237972
step: 190, loss: 0.0023836204782128334
step: 200, loss: 0.028400126844644547
step: 210, loss: 0.029840189963579178
step: 220, loss: 0.024125395342707634
step: 230, loss: 0.0003351588675286621
step: 240, loss: 0.08970962464809418
step: 250, loss: 0.029142489656805992
step: 260, loss: 0.00037351626087911427
step: 270, loss: 0.08172225952148438
step: 280, loss: 0.025287680327892303
step: 290, loss: 0.001000813441351056
step: 300, loss: 0.0009330813190899789
step: 310, loss: 0.0010529081337153912
step: 320, loss: 0.001612629508599639
```

```
step: 330, loss: 0.011041772551834583
step: 340, loss: 0.02143397368490696
step: 350, loss: 0.12683618068695068
step: 360, loss: 0.02313319779932499
step: 370, loss: 0.0037051150575280197
step: 380, loss: 0.0011049644090235233
step: 390, loss: 0.037256572395563126
step: 400, loss: 5.75034755456727e-05
step: 410, loss: 0.00041484017856419086
step: 420, loss: 0.015536496415734291
step: 430, loss: 0.0006996538140811026
step: 440, loss: 0.0076814196072518826
step: 450, loss: 0.007169710472226143
step: 460, loss: 0.00036499349516816437
step: 470, loss: 0.007848951034247875
step: 480, loss: 0.002770141465589404
step: 490, loss: 0.011443265713751316
step: 500, loss: 0.009670421481132507
step: 510, loss: 0.0033660114277154207
step: 520, loss: 0.026807893067598343
step: 530, loss: 0.0049149454571306705
/usr/local/lib/python3.7/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=0.9412317818523741, f1=0.9330819981149858,
best_f1=0.9450853714813106
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.022482167929410934
step: 10, loss: 0.0024709461722522974
step: 20, loss: 0.00016023374337237328
step: 30, loss: 0.006644922774285078
step: 40, loss: 0.03064381517469883
step: 50, loss: 0.004434865899384022
step: 60, loss: 0.035442955791950226
step: 70, loss: 0.00022609240841120481
step: 80, loss: 0.001880955183878541
step: 90, loss: 0.0072547332383692265
step: 100, loss: 0.0003054914704989642
step: 110, loss: 0.001045950804837048
step: 120, loss: 0.0015737281646579504
step: 130, loss: 0.004815349821001291
step: 140, loss: 0.0003389579651411623
step: 150, loss: 0.0011819397332146764
step: 160, loss: 0.0007262652507051826
step: 170, loss: 0.0005231171380728483
step: 180, loss: 0.0019402956822887063
```

```
step: 190, loss: 0.001327711739577353
step: 200, loss: 0.00391562283039093
step: 210, loss: 0.0028987161349505186
step: 220, loss: 0.00014260713942348957
step: 230, loss: 0.01009590644389391
step: 240, loss: 0.00029521732358261943
step: 250, loss: 0.004386547952890396
step: 260, loss: 0.0013234559446573257
step: 270, loss: 0.04652772843837738
step: 280, loss: 0.044174496084451675
step: 290, loss: 0.0015796617371961474
step: 300, loss: 0.005700692068785429
step: 310, loss: 0.08769163489341736
step: 320, loss: 0.000547171977814287
step: 330, loss: 0.00010505871614441276
step: 340, loss: 0.01709018275141716
step: 350, loss: 0.003822882194072008
step: 360, loss: 0.002357610734179616
step: 370, loss: 0.003056602319702506
step: 380, loss: 0.005993092432618141
step: 390, loss: 0.0004042880900669843
step: 400, loss: 0.017517801374197006
step: 410, loss: 0.001983063528314233
step: 420, loss: 0.00020660743757616729
step: 430, loss: 0.0037167591508477926
step: 440, loss: 0.0003689056320581585
step: 450, loss: 0.000530458812136203
step: 460, loss: 0.0007567715365439653
step: 470, loss: 0.03367133066058159
step: 480, loss: 0.0005064456490799785
step: 490, loss: 0.000643103092443198
step: 500, loss: 0.0014547297032549977
step: 510, loss: 0.033362288028001785
step: 520, loss: 0.03194228187203407
step: 530, loss: 0.03723897784948349
/usr/local/lib/python3.7/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=0.9294062646096307, f1=0.9346314325452018,
best_f1=0.9450853714813106
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0944775715470314
step: 10, loss: 0.04301105812191963
step: 20, loss: 0.0002026785077759996
step: 30, loss: 0.029201284050941467
step: 40, loss: 0.0006061961757950485
```

```
step: 50, loss: 0.003000528085976839
step: 60, loss: 0.0008799658389762044
step: 70, loss: 0.0008312500431202352
step: 80, loss: 0.059782978147268295
step: 90, loss: 0.0007825592765584588
step: 100, loss: 0.00021206731616985053
step: 110, loss: 0.0003167421091347933
step: 120, loss: 0.0024671361315995455
step: 130, loss: 0.0002670924295671284
step: 140, loss: 0.005584940779954195
step: 150, loss: 0.009966854006052017
step: 160, loss: 0.010742926970124245
step: 170, loss: 0.05819888412952423
step: 180, loss: 0.00044966363930143416
step: 190, loss: 0.0023438928183168173
step: 200, loss: 0.0022510201670229435
step: 210, loss: 0.010976525023579597
step: 220, loss: 0.0018878328846767545
step: 230, loss: 0.002847790950909257
step: 240, loss: 0.0020915400236845016
step: 250, loss: 0.00011150935461046174
step: 260, loss: 0.00010897697211476043
step: 270, loss: 0.023271549493074417
step: 280, loss: 0.013420699164271355
step: 290, loss: 0.005202204920351505
step: 300, loss: 0.007139430847018957
step: 310, loss: 0.04702724143862724
step: 320, loss: 0.000394322705687955
step: 330, loss: 0.0031687230803072453
step: 340, loss: 0.00015725824050605297
step: 350, loss: 9.437783592147753e-05
step: 360, loss: 0.0014343311777338386
step: 370, loss: 0.001464056666009128
step: 380, loss: 0.0017697543371468782
step: 390, loss: 0.10871636122465134
step: 400, loss: 0.002992029767483473
step: 410, loss: 0.0002237840963061899
step: 420, loss: 0.0016152530442923307
step: 430, loss: 0.04361981898546219
step: 440, loss: 0.15147076547145844
Gradient overflow.  Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 450, loss: 0.00045853445772081614
step: 460, loss: 0.00012239234638400376
step: 470, loss: 0.0001554545306134969
step: 480, loss: 0.00012199379125377163
step: 490, loss: 0.0642336905002594
step: 500, loss: 0.007057225797325373
step: 510, loss: 0.0009908683132380247
```

```
step: 520, loss: 0.004507166333496571
step: 530, loss: 0.03152082487940788
/usr/local/lib/python3.7/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=0.9406264609630668, f1=0.931098696461825,
best_f1=0.9450853714813106
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0026218879502266645
step: 10, loss: 0.012796865776181221
step: 20, loss: 0.030078284442424774
step: 30, loss: 0.0026625508908182383
step: 40, loss: 0.04324067011475563
step: 50, loss: 0.0005632779793813825
step: 60, loss: 0.000123669917229563
step: 70, loss: 0.017914097756147385
step: 80, loss: 0.04358018562197685
step: 90, loss: 2.513361323508434e-05
step: 100, loss: 0.00108260044362396
step: 110, loss: 5.672048791893758e-05
step: 120, loss: 0.00040986062958836555
step: 130, loss: 0.011427802965044975
step: 140, loss: 0.0016254113288596272
step: 150, loss: 0.0011317436583340168
step: 160, loss: 0.003565968247130513
step: 170, loss: 0.00959751009941101
step: 180, loss: 0.00043484187335707247
step: 190, loss: 0.0008132922230288386
step: 200, loss: 0.002758611924946308
step: 210, loss: 0.0005767885013483465
step: 220, loss: 0.0002120180579368025
step: 230, loss: 0.00015060383884701878
step: 240, loss: 0.00017289008246734738
step: 250, loss: 0.0019873876590281725
step: 260, loss: 0.00676176929846406
step: 270, loss: 7.412239210680127e-05
step: 280, loss: 0.008826700039207935
step: 290, loss: 0.04201119765639305
step: 300, loss: 2.049993236141745e-05
step: 310, loss: 0.001967252930626273
step: 320, loss: 0.00014866302080918103
step: 330, loss: 0.00043477071449160576
step: 340, loss: 3.194507007719949e-05
step: 350, loss: 0.00012912835518363863
step: 360, loss: 2.232896804343909e-05
step: 370, loss: 0.00680931843817234
```

```
step: 380, loss: 0.00030128820799291134
step: 390, loss: 0.0009672027663327754
step: 400, loss: 0.2944793403148651
step: 410, loss: 0.01987561583518982
step: 420, loss: 0.00011578621342778206
step: 430, loss: 0.00019630344468168914
step: 440, loss: 0.0014861708041280508
step: 450, loss: 0.14540298283100128
step: 460, loss: 0.015920313075184822
step: 470, loss: 0.00034766006865538657
step: 480, loss: 0.0028751222416758537
step: 490, loss: 0.00701560964807868
step: 500, loss: 0.0029450496658682823
step: 510, loss: 0.002604010282084346
step: 520, loss: 0.0066939787939190865
step: 530, loss: 0.00013358409341890365
/usr/local/lib/python3.7/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=0.9427107591988821, f1=0.9355586462679647,
best_f1=0.9450853714813106
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0006357435486279428
step: 10, loss: 0.00016540945216547698
step: 20, loss: 0.0002170888619730249
step: 30, loss: 0.00017906824359670281
step: 40, loss: 0.0001081281661754474
step: 50, loss: 0.00027605213108472526
step: 60, loss: 0.00019353328389115632
step: 70, loss: 6.892043893458322e-05
step: 80, loss: 0.00040257067303173244
step: 90, loss: 0.0009665217949077487
step: 100, loss: 8.670681563671678e-05
step: 110, loss: 0.00024255047901533544
step: 120, loss: 9.669257997302338e-05
step: 130, loss: 0.0003095285501331091
step: 140, loss: 0.06731318682432175
step: 150, loss: 0.00042964491876773536
step: 160, loss: 0.005575270392000675
step: 170, loss: 0.00014516613737214357
step: 180, loss: 0.0005161995068192482
step: 190, loss: 0.0004933088202960789
step: 200, loss: 0.014576783403754234
step: 210, loss: 0.0003066952631343156
step: 220, loss: 0.0020359728951007128
step: 230, loss: 0.001323480624705553
```

```
step: 240, loss: 0.0006386907771229744
step: 250, loss: 0.005727978423237801
step: 260, loss: 0.0029564956203103065
step: 270, loss: 0.0007371801184490323
step: 280, loss: 0.1681644469499588
step: 290, loss: 0.00024308194406330585
step: 300, loss: 0.0004867469542659819
step: 310, loss: 0.00028484579524956644
step: 320, loss: 0.0012273021275177598
step: 330, loss: 0.0008046728326007724
step: 340, loss: 0.013100785203278065
step: 350, loss: 0.0006162743084132671
step: 360, loss: 0.00028617875068448484
step: 370, loss: 0.011348847299814224
step: 380, loss: 0.15549376606941223
step: 390, loss: 0.000759061542339623
step: 400, loss: 0.004417884163558483
step: 410, loss: 0.002858918160200119
step: 420, loss: 0.0019530714489519596
step: 430, loss: 0.0018727956339716911
step: 440, loss: 0.0037825305480509996
step: 450, loss: 0.0003153980360366404
step: 460, loss: 0.0004188393068034202
step: 470, loss: 0.0030633520800620317
step: 480, loss: 0.00015669876302126795
step: 490, loss: 0.0072942255064845085
step: 500, loss: 0.00624457560479641
step: 510, loss: 0.0129269789904356
step: 520, loss: 0.0009807520546019077
step: 530, loss: 0.0009609129629097879
/usr/local/lib/python3.7/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=0.9401947148817803, f1=0.9380203515263645,
best_f1=0.9450853714813106
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0008667180081829429
step: 10, loss: 0.0012947161449119449
step: 20, loss: 0.0003835085954051465
step: 30, loss: 0.0003316118963994086
step: 40, loss: 0.00026611299836076796
step: 50, loss: 0.0016156903002411127
step: 60, loss: 0.022742297500371933
step: 70, loss: 7.336202543228865e-05
step: 80, loss: 0.00023996559320949018
step: 90, loss: 9.9888929980807e-05
```

```
step: 100, loss: 0.011858879588544369
step: 110, loss: 0.0004087313136551529
step: 120, loss: 0.00047113242908380926
step: 130, loss: 0.00013162096729502082
step: 140, loss: 0.024390222504734993
step: 150, loss: 4.097199780517258e-05
step: 160, loss: 0.016682149842381477
step: 170, loss: 7.242238643812016e-05
step: 180, loss: 0.0023731058463454247
step: 190, loss: 0.005413692444562912
step: 200, loss: 0.004933669697493315
step: 210, loss: 0.00021267880219966173
step: 220, loss: 0.001396775827743113
step: 230, loss: 6.545337964780629e-05
step: 240, loss: 0.0005139120039530098
step: 250, loss: 0.0015297472709789872
step: 260, loss: 0.0001171140611404553
step: 270, loss: 0.00034439688897691667
step: 280, loss: 7.204111170722172e-05
step: 290, loss: 0.0044275084510445595
step: 300, loss: 7.456340244971216e-05
step: 310, loss: 0.0010948857525363564
step: 320, loss: 7.038993499008939e-05
step: 330, loss: 0.0001403132191626355
step: 340, loss: 0.0010208599269390106
step: 350, loss: 0.00012692922609858215
step: 360, loss: 5.600827716989443e-05
step: 370, loss: 0.0005702206981368363
step: 380, loss: 0.00013486365787684917
step: 390, loss: 0.00014565762830898166
step: 400, loss: 0.001218862715177238
step: 410, loss: 0.00022284412989392877
step: 420, loss: 5.767217226093635e-05
step: 430, loss: 0.0001646404853090644
step: 440, loss: 0.0003414677339605987
step: 450, loss: 0.000393520895158872
step: 460, loss: 0.016283096745610237
step: 470, loss: 0.0001964321854757145
step: 480, loss: 0.0012698742793872952
step: 490, loss: 0.0029962186235934496
step: 500, loss: 0.00032954997732304037
step: 510, loss: 0.00016539696662221104
step: 520, loss: 4.236009772284888e-05
step: 530, loss: 0.0004060036444570869
/usr/local/lib/python3.7/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=0.9404651162790698, f1=0.9296296296296297,
```

```
best_f1=0.9450853714813106
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 3.1562372896587476e-05
step: 10, loss: 2.1930114598944783e-05
step: 20, loss: 0.04181075468659401
step: 30, loss: 0.009841651655733585
step: 40, loss: 0.0006829541525803506
step: 50, loss: 0.023439079523086548
step: 60, loss: 0.010908818803727627
step: 70, loss: 0.0001693713420536369
step: 80, loss: 5.7533943361818209e-05
step: 90, loss: 3.4358337870799005e-05
step: 100, loss: 0.0001638026733417064
step: 110, loss: 6.393976946128532e-05
step: 120, loss: 2.3152051653596573e-05
step: 130, loss: 0.0012353014899417758
step: 140, loss: 0.051887284964323044
step: 150, loss: 0.00042665351065807045
step: 160, loss: 0.00025774064124561846
step: 170, loss: 2.342758307349868e-05
step: 180, loss: 0.05178562551736832
step: 190, loss: 0.003292984329164028
step: 200, loss: 0.0002815847401507199
step: 210, loss: 6.60197256365791e-05
step: 220, loss: 0.00039583133184351027
step: 230, loss: 6.885750917717814e-05
step: 240, loss: 4.422353231348097e-05
step: 250, loss: 4.1251612856285647e-05
step: 260, loss: 7.073012966429815e-05
step: 270, loss: 6.686306005576625e-05
step: 280, loss: 0.0022519463673233986
step: 290, loss: 0.000583428714890033
step: 300, loss: 0.0003850914363283664
step: 310, loss: 2.7633022909867577e-05
step: 320, loss: 2.273110112582799e-05
step: 330, loss: 0.00011123695730930194
step: 340, loss: 0.00016039705951698124
step: 350, loss: 0.0031758900731801987
step: 360, loss: 0.003205028595402837
step: 370, loss: 0.0002861428656615317
step: 380, loss: 0.0002762045478448272
step: 390, loss: 0.00010604471754049882
step: 400, loss: 0.1051715835928917
step: 410, loss: 0.0003236234770156443
step: 420, loss: 0.006853075698018074
step: 430, loss: 0.019093068316578865
```

```
step: 440, loss: 9.983897325582802e-05
step: 450, loss: 0.004374200478196144
step: 460, loss: 0.00023194242385216057
step: 470, loss: 0.00017421269149053842
step: 480, loss: 0.0011635645059868693
step: 490, loss: 0.0025915508158504963
step: 500, loss: 0.0005736117600463331
step: 510, loss: 0.0021265868563205004
step: 520, loss: 0.00013039108307566494
step: 530, loss: 0.00358349340967834
/usr/local/lib/python3.7/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=0.9428838951310862, f1=0.9343269678621332,
best_f1=0.9450853714813106
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.003883660538122058
step: 10, loss: 4.0838429413270205e-05
step: 20, loss: 0.0003980401379521936
step: 30, loss: 0.001544563565403223
step: 40, loss: 0.00025353257660754025
step: 50, loss: 0.020618535578250885
step: 60, loss: 4.8700880142860115e-05
step: 70, loss: 0.00016149351722560823
step: 80, loss: 0.0008802236989140511
step: 90, loss: 0.011593620292842388
step: 100, loss: 0.0018849273910745978
step: 110, loss: 0.0007159945671446621
step: 120, loss: 0.0019149092258885503
step: 130, loss: 0.0007572666509076953
step: 140, loss: 0.0003434709506109357
step: 150, loss: 0.00533656869083643
step: 160, loss: 7.203819404821843e-05
step: 170, loss: 0.00040937616722658277
step: 180, loss: 8.09029588708654e-05
step: 190, loss: 0.00010872448910959065
step: 200, loss: 6.549867975991219e-05
step: 210, loss: 0.00242374604567854
step: 220, loss: 8.456877549178898e-05
step: 230, loss: 9.619991760700941e-05
step: 240, loss: 0.00039743975503370166
step: 250, loss: 0.00022435748542193323
step: 260, loss: 0.0001204221262014471
step: 270, loss: 3.120871406281367e-05
step: 280, loss: 5.297297320794314e-05
step: 290, loss: 0.00034931523259729147
```

```
step: 300, loss: 3.804650987149216e-05
step: 310, loss: 0.0012479025172069669
step: 320, loss: 0.006236674729734659
step: 330, loss: 0.0016769074136391282
step: 340, loss: 9.069334191735834e-05
step: 350, loss: 5.476427759276703e-05
step: 360, loss: 0.00031816127011552453
step: 370, loss: 0.02207225188612938
step: 380, loss: 0.0005380421644076705
step: 390, loss: 0.004017170052975416
step: 400, loss: 0.0006062027532607317
step: 410, loss: 0.001053009764291346
step: 420, loss: 0.020877398550510406
step: 430, loss: 0.005308268126100302
step: 440, loss: 0.0014978524995967746
step: 450, loss: 0.001250597182661295
step: 460, loss: 0.0015575401484966278
step: 470, loss: 0.0003510827955324203
step: 480, loss: 0.0003688035940285772
step: 490, loss: 0.0025581626687198877
step: 500, loss: 0.0003656666085589677
step: 510, loss: 0.0005817149067297578
step: 520, loss: 0.001928870566189289
step: 530, loss: 0.0001385885989293456
/usr/local/lib/python3.7/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=0.9428172942817293, f1=0.9353647276084949,
best_f1=0.9450853714813106
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.003116944804787636
step: 10, loss: 3.347764868522063e-05
step: 20, loss: 0.00013143743854016066
step: 30, loss: 0.00010013456630986184
step: 40, loss: 0.0001315333356615156
step: 50, loss: 0.0023988233879208565
step: 60, loss: 1.7475133063271642e-05
step: 70, loss: 0.0009431135840713978
step: 80, loss: 7.404717325698584e-05
step: 90, loss: 0.00017437979113310575
step: 100, loss: 0.02803601324558258
step: 110, loss: 6.0227230278542265e-05
step: 120, loss: 0.0005969246267341077
step: 130, loss: 0.007077860180288553
step: 140, loss: 0.0002498703252058476
step: 150, loss: 3.970452962676063e-05
```

```
step: 160, loss: 0.0036424214486032724
step: 170, loss: 0.0012112708063796163
step: 180, loss: 0.00013984291581436992
step: 190, loss: 6.810403283452615e-05
step: 200, loss: 2.9689297662116587e-05
step: 210, loss: 3.712002217071131e-05
step: 220, loss: 4.935513061354868e-05
step: 230, loss: 6.34959724266082e-05
step: 240, loss: 3.20311555697117e-05
step: 250, loss: 0.00011431483289925382
step: 260, loss: 0.00019743455050047487
step: 270, loss: 0.00010628675227053463
step: 280, loss: 5.12501756020356e-05
step: 290, loss: 5.948275065748021e-05
step: 300, loss: 0.00044432309069670737
step: 310, loss: 0.0001701531291473657
step: 320, loss: 0.00014818414638284594
step: 330, loss: 0.0011308752000331879
step: 340, loss: 0.00024688965640962124
step: 350, loss: 0.0008299996261484921
step: 360, loss: 0.022853923961520195
step: 370, loss: 0.00011839645594591275
step: 380, loss: 0.0002821404195856303
step: 390, loss: 0.07555506378412247
step: 400, loss: 0.0003448709030635655
step: 410, loss: 4.044483648613095e-05
step: 420, loss: 8.83085303939879e-05
step: 430, loss: 6.985123036429286e-05
step: 440, loss: 0.00696060844433579445
step: 450, loss: 0.0004260904679540545
step: 460, loss: 0.004194257780909538
step: 470, loss: 0.0013691324274986982
step: 480, loss: 4.035875463159755e-05
step: 490, loss: 2.765917088254355e-05
step: 500, loss: 9.718078217701986e-05
step: 510, loss: 0.022004226222634315
step: 520, loss: 7.18409355613403e-05
step: 530, loss: 3.029476647498086e-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=0.9412852519648636, f1=0.9380449747590638,
best_f1=0.9450853714813106
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 2.3539567337138578e-05
step: 10, loss: 3.180038766004145e-05
```

```
step: 20, loss: 3.093664417974651e-05
step: 30, loss: 0.00026773876743391156
step: 40, loss: 0.08864908665418625
step: 50, loss: 9.162705100607127e-05
step: 60, loss: 3.32433519361075e-05
step: 70, loss: 0.00011819838255178183
step: 80, loss: 7.285436004167423e-05
step: 90, loss: 4.461194475879893e-05
step: 100, loss: 0.04313957318663597
step: 110, loss: 0.000885307730641216
step: 120, loss: 0.0028951705899089575
step: 130, loss: 0.07402981072664261
step: 140, loss: 6.143413338577375e-05
step: 150, loss: 2.9543678465415724e-05
step: 160, loss: 3.198599733877927e-05
step: 170, loss: 6.804140866734087e-05
step: 180, loss: 5.216763020143844e-05
step: 190, loss: 5.1855753554264084e-05
step: 200, loss: 4.838926179218106e-05
step: 210, loss: 0.0016104895621538162
step: 220, loss: 4.941943916492164e-05
step: 230, loss: 0.0015970563981682062
step: 240, loss: 0.0007981430971994996
step: 250, loss: 8.977087418315932e-05
step: 260, loss: 0.005037084221839905
step: 270, loss: 3.1504645448876545e-05
step: 280, loss: 1.644693656999152e-05
step: 290, loss: 2.1371552065829746e-05
step: 300, loss: 0.00021376731456257403
step: 310, loss: 0.000209154182812199
step: 320, loss: 0.0003479969745967537
step: 330, loss: 0.00011001239909091964
step: 340, loss: 8.194960537366569e-05
step: 350, loss: 3.0084087484283373e-05
step: 360, loss: 0.003200661391019821
step: 370, loss: 0.0004622801789082587
step: 380, loss: 7.332273526117206e-05
step: 390, loss: 0.0003475230769254267
step: 400, loss: 3.980670953751542e-05
step: 410, loss: 0.0003453663084656
step: 420, loss: 6.828294135630131e-05
step: 430, loss: 0.00028090013074688613
step: 440, loss: 0.00010203369311057031
step: 450, loss: 2.8136193577665836e-05
step: 460, loss: 5.409673030953854e-05
step: 470, loss: 0.001341169816441834
step: 480, loss: 2.2902575437910855e-05
step: 490, loss: 2.4586077415733598e-05
```

```
step: 500, loss: 0.00039677132735960186
step: 510, loss: 6.838350964244455e-05
step: 520, loss: 3.520882819429971e-05
step: 530, loss: 0.0011270437389612198
/usr/local/lib/python3.7/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=0.9419953596287702, f1=0.9349930843706776,
best_f1=0.9450853714813106
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 [00:26, 220.33it/s]
load_f1 = 0.9446768944676894
real_f1 = 0.9446768944676894
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 [00:19, 223.12it/s]
```

2 BASELINE

2.1 BASELINE STRUCTURED

2.1.1 Beer - Running the matcher

```python
# 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 bert-base-uncased \
  --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 bert-base-uncased \
  --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: 0.5951436161994934
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.4782608695652174, f1=0.45, best_f1=0.45
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.5059506297111511
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.5142857142857143, f1=0.5142857142857143,
best_f1=0.5142857142857143
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.

```
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.4737869203090668
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.5806451612903226, f1=0.5294117647058824,
best_f1=0.5294117647058824
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.2728976607322693
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.6956521739130435, f1=0.5517241379310344,
best_f1=0.5517241379310344
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.2917260229587555
/usr/local/lib/python3.7/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=0.6875000000000001, f1=0.5, best_f1=0.5517241379310344
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.19082775712013245
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.8148148148148148, f1=0.6206896551724138,
best_f1=0.6206896551724138
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
```

```
step: 0, loss: 0.10189282894134521
/usr/local/lib/python3.7/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=0.7777777777777778, f1=0.5789473684210527,
best_f1=0.6206896551724138
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.04210579767823219
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.8387096774193549, f1=0.6470588235294117,
best_f1=0.6470588235294117
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0026811426505446434
/usr/local/lib/python3.7/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=0.8387096774193549, f1=0.6470588235294117,
best_f1=0.6470588235294117
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.025330014526844025
/usr/local/lib/python3.7/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=0.6470588235294117,
best_f1=0.6470588235294117
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0026916740462183952
/usr/local/lib/python3.7/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=0.7692307692307692, f1=0.7586206896551724,
best_f1=0.6470588235294117
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
```

```
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.005531466566026211
/usr/local/lib/python3.7/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=0.7878787878787878, f1=0.6857142857142857,
best_f1=0.6470588235294117
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0036483248695731163
/usr/local/lib/python3.7/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=0.8, f1=0.6470588235294117, best_f1=0.6470588235294117
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0037169780116528273
/usr/local/lib/python3.7/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=0.7857142857142857, f1=0.7333333333333334,
best_f1=0.6470588235294117
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.005069883074611425
/usr/local/lib/python3.7/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=0.7857142857142857, f1=0.7333333333333334,
best_f1=0.6470588235294117
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.")
91it [00:00, 137098.30it/s]
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
load_f1 = 0.7999999999999999
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 [00:16, 271.11it/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 bert-base-uncased \
       --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 bert-base-uncased \
  --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: 0.6272245049476624
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 10, loss: 0.5736082792282104
step: 20, loss: 0.34831702709198

```
step: 30, loss: 0.18740178644657135
step: 40, loss: 0.22859103977680206
step: 50, loss: 0.04402647167444229
step: 60, loss: 0.11197713017463684
step: 70, loss: 0.0672559142112732
step: 80, loss: 0.06802605837583542
step: 90, loss: 0.14019127190113068
step: 100, loss: 0.0063547058962288552
step: 110, loss: 0.09664954245090485
step: 120, loss: 0.012116491794586182
step: 130, loss: 0.02486845664680004
step: 140, loss: 0.004789031110703945
step: 150, loss: 0.011197269894182682
step: 160, loss: 0.007236007135361433
step: 170, loss: 0.05827059596776962
step: 180, loss: 0.006281403824687004
step: 190, loss: 0.0023051362950354815
step: 200, loss: 0.032967709004879
step: 210, loss: 0.002137178787961602
step: 220, loss: 0.0029133346397429705
step: 230, loss: 0.020952044054865837
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9809203142536477, f1=0.9749430523917996,
best_f1=0.9749430523917996
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.022934794425964355
step: 10, loss: 0.002386761363595724
step: 20, loss: 0.05957091227173805
step: 30, loss: 0.20308151841163635
step: 40, loss: 0.04761108011007309
step: 50, loss: 0.007089200895279646
step: 60, loss: 0.0013774757971987128
step: 70, loss: 0.059876300394535065
step: 80, loss: 0.006100693251937628
step: 90, loss: 0.00436230655759573
step: 100, loss: 0.01658911444246769
step: 110, loss: 0.005991771817207336
step: 120, loss: 0.16805429756641388
step: 130, loss: 0.2539735436439514
step: 140, loss: 0.012648270465433598
```

```
step: 150, loss: 0.03386475145816803
step: 160, loss: 0.004350883886218071
step: 170, loss: 0.0021131907124072313
step: 180, loss: 0.006302378606051207
step: 190, loss: 0.043335214257240295
step: 200, loss: 0.0010457439348101616
step: 210, loss: 0.009859693236649036
step: 220, loss: 0.04741985350847244
step: 230, loss: 0.002648165449500084
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.987598647125141, f1=0.9819819819819819,
best_f1=0.9819819819819819
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.007300368510186672
step: 10, loss: 0.0015388474566861987
step: 20, loss: 0.0006799096590839326
step: 30, loss: 0.00906427577137947
step: 40, loss: 0.05107806995511055
step: 50, loss: 0.003651769831776619
step: 60, loss: 0.001797109143808484
step: 70, loss: 0.0015384610742330551
step: 80, loss: 0.001436686608940363
step: 90, loss: 0.011018961668014526
step: 100, loss: 0.003989658318459988
step: 110, loss: 0.0009442548616789281
step: 120, loss: 0.04479938745498657
step: 130, loss: 0.00047249949420802295
step: 140, loss: 0.008577162399888039
step: 150, loss: 0.0016996819758787751
step: 160, loss: 0.023966524749994278
step: 170, loss: 0.005133269354701042
step: 180, loss: 0.0008115096716210246
step: 190, loss: 0.0028037538286298513
step: 200, loss: 0.0011299968464300036
step: 210, loss: 0.0003834753588307649
step: 220, loss: 0.0014771951828151941
step: 230, loss: 0.00016554261674173176
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
```

```
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9887640449438202, f1=0.9864253393665158,
best_f1=0.9864253393665158
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00012173617142252624
step: 10, loss: 0.00017125913291238248
step: 20, loss: 0.0003510774113237858
step: 30, loss: 0.0026117628440260887
step: 40, loss: 0.0006974849384278059
step: 50, loss: 0.003439415944740176
step: 60, loss: 0.000793388404417783
step: 70, loss: 0.00031215333729051054
step: 80, loss: 0.01075664907693863
step: 90, loss: 0.0006510630482807755
step: 100, loss: 0.025065282359719276
step: 110, loss: 0.00036390358582139015
step: 120, loss: 0.02050042152404785
step: 130, loss: 0.0023783191572874784
step: 140, loss: 0.0009445723844692111
step: 150, loss: 0.061276815831661224
step: 160, loss: 0.0005314275622367859
step: 170, loss: 0.001378742977976799
step: 180, loss: 0.0010271752253174782
step: 190, loss: 0.0034784111194312572
Gradient overflow.  Skipping step, loss scaler 0 reducing loss scale to 8192.0
step: 200, loss: 0.00047951226588338614
step: 210, loss: 0.02487030252814293
step: 220, loss: 0.004676831420511007
step: 230, loss: 0.00216737762093544
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9830124575311437,
best_f1=0.9830124575311437
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0036192447878420353
step: 10, loss: 0.10303062200546265
```

```
step: 20, loss: 0.00607782369479537
step: 30, loss: 0.00019828815129585564
step: 40, loss: 0.0005128983757458627
step: 50, loss: 0.0002443357215706259
step: 60, loss: 0.0001999032247113064
step: 70, loss: 0.0003190765273757279
step: 80, loss: 0.0006472115055657923
step: 90, loss: 0.001629972248338163
step: 100, loss: 0.00018194458971265703
step: 110, loss: 0.00030124562908895314
step: 120, loss: 5.76565507799387e-05
step: 130, loss: 0.0004410024848766625
step: 140, loss: 0.0008389314753003418
step: 150, loss: 0.02168995328247547
step: 160, loss: 0.000578473205678165
step: 170, loss: 0.0070081851445138454
step: 180, loss: 0.0033213121350854635
step: 190, loss: 0.0038942359387874603
step: 200, loss: 0.0014632671372964978
step: 210, loss: 0.00031726292218081653
step: 220, loss: 0.000386374827940017
step: 230, loss: 0.00011905932478839532
/usr/local/lib/python3.7/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=0.983050847457627, f1=0.9728506787330317,
best_f1=0.9830124575311437
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.000905118475202471
step: 10, loss: 0.0003551135887391865
step: 20, loss: 0.003052356420084834
step: 30, loss: 0.09444867819547653
step: 40, loss: 0.1129281297326088
step: 50, loss: 0.0007556056370958686
step: 60, loss: 0.00029443236417137086
step: 70, loss: 0.003464545588940382
step: 80, loss: 0.011766097508370876
step: 90, loss: 0.007756147533655167
step: 100, loss: 0.038020841777324677
step: 110, loss: 0.0004567269061226398
step: 120, loss: 0.00028563174419105053
step: 130, loss: 0.013392725959420204
step: 140, loss: 0.0005155996768735349
step: 150, loss: 0.0019310866482555866
step: 160, loss: 0.0012806288432329893
step: 170, loss: 0.0002965754538308829
```

```
step: 180, loss: 0.0704735592007637
step: 190, loss: 0.00040641098166815937
step: 200, loss: 0.0022109285928308964
step: 210, loss: 0.0002983398735523224
step: 220, loss: 0.00012711016461253166
step: 230, loss: 0.07203879207372665
/usr/local/lib/python3.7/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=0.9841269841269841, f1=0.9761092150170648,
best_f1=0.9830124575311437
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.006690435577183962
step: 10, loss: 0.00013068558473605663
step: 20, loss: 0.00018792421906255186
step: 30, loss: 0.0013723153388127685
step: 40, loss: 0.0003614618326537311
step: 50, loss: 0.00032104685669764876
step: 60, loss: 0.040377549827098846
step: 70, loss: 0.013956270180642605
step: 80, loss: 0.00016588960716035217
step: 90, loss: 6.778870738344267e-05
step: 100, loss: 0.00012643684749491513
step: 110, loss: 0.0005365795805118978
step: 120, loss: 0.00012880739814136177
step: 130, loss: 0.00040149447158910334
step: 140, loss: 0.00016722534201107920
step: 150, loss: 0.01582835242152214
step: 160, loss: 0.043316859751939774
step: 170, loss: 0.0012720456579700112
step: 180, loss: 0.0007842901395633817
step: 190, loss: 0.00013048452092334628
step: 200, loss: 0.040573205798864365
step: 210, loss: 9.319237869931385e-05
step: 220, loss: 0.00018866221944335848
step: 230, loss: 0.01156504638493061
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
epoch 7: dev_f1=0.9898762654668166, f1=0.9773755656108598,
best_f1=0.9830124575311437
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 7.707301119808108e-05
step: 10, loss: 0.00017360415949951857
```

```
step: 20, loss: 0.004514187108725309
step: 30, loss: 0.010436397045850754
step: 40, loss: 0.001134455669671297
step: 50, loss: 0.0023668084759265184
step: 60, loss: 0.0004580229287967086
step: 70, loss: 0.00028303847648203373
step: 80, loss: 0.00013896847667638212
step: 90, loss: 8.205945050576702e-05
step: 100, loss: 0.00014291092520579696
step: 110, loss: 0.01877664402127266
step: 120, loss: 0.006745961960405111
step: 130, loss: 0.00030362021061591804
step: 140, loss: 8.999575220514089e-05
step: 150, loss: 7.180039392551407e-05
step: 160, loss: 0.012005187571048737
step: 170, loss: 7.175382052082568e-05
step: 180, loss: 0.00012391780910547823
step: 190, loss: 0.00016058134497143328
step: 200, loss: 0.00019854471611324698
step: 210, loss: 0.00020545651204884052
step: 220, loss: 0.001287960447371006
step: 230, loss: 0.00015468250785488635
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9909502262443439, f1=0.9794988610478361,
best_f1=0.9794988610478361
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00017439592920709401
step: 10, loss: 0.012102257460355759
step: 20, loss: 0.0011170804500579834
step: 30, loss: 0.0007559032528661191
step: 40, loss: 0.016679760068655014
step: 50, loss: 7.448132964782417e-05
step: 60, loss: 0.0006399195408448577
step: 70, loss: 0.0003628820995800197
step: 80, loss: 0.00010898290202021599
step: 90, loss: 0.00029861682560294867
step: 100, loss: 0.0001971609308384359
step: 110, loss: 0.00037497657467611134
step: 120, loss: 5.983533992548473e-05
step: 130, loss: 0.0004727381747215986
```

```
step: 140, loss: 0.03343388810753822
step: 150, loss: 0.00021377655502874404
step: 160, loss: 8.163407619576901e-05
step: 170, loss: 0.00012296847125981003
step: 180, loss: 0.0029728910885751247
step: 190, loss: 0.00029660252039320767
step: 200, loss: 0.00021137113799341023
step: 210, loss: 0.0005940190167166293
step: 220, loss: 8.459872333332896e-05
step: 230, loss: 0.0012158039025962353
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9920903954802259, f1=0.9819004524886877,
best_f1=0.9819004524886877
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00016336645057890564
step: 10, loss: 3.423415546421893e-05
step: 20, loss: 8.97437785170041e-05
step: 30, loss: 0.0002585502224974334
step: 40, loss: 6.740404205629602e-05
step: 50, loss: 0.00018259583157487214
step: 60, loss: 0.0019534290768206120
step: 70, loss: 0.00012776008225046098
step: 80, loss: 7.11213651811704e-05
step: 90, loss: 0.00010545110853854567
step: 100, loss: 6.786532321712002e-05
step: 110, loss: 7.400183676509187e-05
step: 120, loss: 0.00017203603056259453
step: 130, loss: 0.00010319970169803128
step: 140, loss: 0.033141326159238815
step: 150, loss: 0.016289198771119118
step: 160, loss: 3.132491838186979e-05
step: 170, loss: 9.749076707521453e-05
step: 180, loss: 7.247319445014e-05
step: 190, loss: 0.0253834780305624
step: 200, loss: 0.00015127507504075766
step: 210, loss: 3.0404697099584155e-05
step: 220, loss: 6.496383139165118e-05
step: 230, loss: 0.0025178883224725723
/usr/local/lib/python3.7/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=0.9909706546275394, f1=0.9819004524886877,
best_f1=0.9819004524886877
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 7.474932499462739e-05
step: 10, loss: 6.802104326197878e-05
step: 20, loss: 0.0027565991040319204
step: 30, loss: 0.017240170389413834
step: 40, loss: 0.00014015946362633258
step: 50, loss: 5.265326035441831e-05
step: 60, loss: 9.655233589001e-05
step: 70, loss: 0.00029499028460122645
step: 80, loss: 5.097561006550677e-05
step: 90, loss: 6.24005842837505e-05
step: 100, loss: 4.746089689433575e-05
step: 110, loss: 0.015073482878506184
step: 120, loss: 6.485152698587626e-05
step: 130, loss: 2.962142025353387e-05
step: 140, loss: 0.00022472890850622207
step: 150, loss: 0.030639519914984703
step: 160, loss: 0.00020257114374544472
step: 170, loss: 0.02406388893723488
step: 180, loss: 4.739651194540784e-05
step: 190, loss: 4.0383540181210265e-05
step: 200, loss: 0.00010029276745626703
step: 210, loss: 3.238292993046343e-05
step: 220, loss: 7.444697257597e-05
step: 230, loss: 4.943187377648428e-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 11: dev_f1=0.9853438556933484, f1=0.9784824462061155,
best_f1=0.9819004524886877
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 5.627406790154055e-05
step: 10, loss: 5.435072671389207e-05
step: 20, loss: 4.3956148147117347e-05
step: 30, loss: 0.00013628262968268245
step: 40, loss: 0.0008214561967179179
step: 50, loss: 0.00038283306639641523
step: 60, loss: 0.0003497059515211731
step: 70, loss: 6.241487426450476e-05
step: 80, loss: 5.4375002946471795e-05
step: 90, loss: 4.424635699251667e-05
```

```
step: 100, loss: 0.00011163135059177876
step: 110, loss: 6.22851075604558e-05
step: 120, loss: 6.0438622313085943e-05
step: 130, loss: 4.470538260648027e-05
step: 140, loss: 0.02541334554553032
step: 150, loss: 4.947617344441824e-05
step: 160, loss: 5.535728269023821e-05
step: 170, loss: 4.63887226942461e-05
step: 180, loss: 0.013739845715463161
step: 190, loss: 0.00014948398165870458
step: 200, loss: 2.619535916892346e-05
step: 210, loss: 3.835045936284587e-05
step: 220, loss: 3.9183923945529386e-05
step: 230, loss: 0.022650565952062607
/usr/local/lib/python3.7/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=0.9887387387387387, f1=0.9796380090497738,
best_f1=0.9819004524886877
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 5.150984725332819e-05
step: 10, loss: 0.006548650097101927
step: 20, loss: 0.00010180337267229334
step: 30, loss: 8.953845099313185e-05
step: 40, loss: 4.790223465533927e-05
step: 50, loss: 0.0004221196286380291
step: 60, loss: 4.6038061555009335e-05
step: 70, loss: 2.5960711354855448e-05
step: 80, loss: 4.626877125701867e-05
step: 90, loss: 5.100160342408344e-05
step: 100, loss: 2.5301244022557512e-05
step: 110, loss: 0.03277277201414108
step: 120, loss: 0.02241881750524044
step: 130, loss: 8.993552182801068e-05
step: 140, loss: 3.58916622644756e-05
step: 150, loss: 3.888693026965484e-05
step: 160, loss: 4.344801709521562e-05
step: 170, loss: 4.154527050559409e-05
step: 180, loss: 5.7652501709526405e-05
step: 190, loss: 5.777649857918732e-05
step: 200, loss: 4.641237683244981e-05
step: 210, loss: 1.989631709875539e-05
step: 220, loss: 4.4877338950755075e-05
step: 230, loss: 3.126889350824058e-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=0.9875706214689265, f1=0.9806598407281,
best_f1=0.9819004524886877
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 3.258006472606212e-05
step: 10, loss: 0.015618470497429371
step: 20, loss: 8.249521488323808e-05
step: 30, loss: 0.00018150685355067253
step: 40, loss: 0.0002479999966453761
step: 50, loss: 3.3802163670770824e-05
step: 60, loss: 4.5127544581191614e-05
step: 70, loss: 3.328055390738882e-05
step: 80, loss: 3.253192699048668e-05
step: 90, loss: 4.202994023216888e-05
step: 100, loss: 2.9596276363008656e-05
step: 110, loss: 0.00010510202264413238
step: 120, loss: 9.202285582432523e-05
step: 130, loss: 4.217155583319254e-05
step: 140, loss: 7.51671104808338e-05
step: 150, loss: 3.484421904431656e-05
step: 160, loss: 0.00620601668953896
step: 170, loss: 1.6975940525298938e-05
step: 180, loss: 0.0002587259514257312
step: 190, loss: 6.156806193757802e-05
step: 200, loss: 2.881036743929144e-05
step: 210, loss: 0.0017898408696055412
step: 220, loss: 0.00020805036183446646
step: 230, loss: 3.288960215286352e-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=0.9853438556933484, f1=0.9762174405436014,
best_f1=0.9819004524886877
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 3.257083517382853e-05
step: 10, loss: 1.8782489860313945e-05
step: 20, loss: 3.453216049820185e-05
step: 30, loss: 3.4907745430245996e-05
step: 40, loss: 4.555695704766549e-05
step: 50, loss: 0.046366482973098755
step: 60, loss: 7.640897092642263e-05
step: 70, loss: 0.0027907495386898518
step: 80, loss: 7.208502211142331e-05
step: 90, loss: 5.0090042350348085e-05
```

```
step: 100, loss: 6.695371848763898e-05
step: 110, loss: 2.514500738470815e-05
step: 120, loss: 0.00010725067113526165
step: 130, loss: 0.024307044222950935
step: 140, loss: 2.5644174456829205e-05
step: 150, loss: 0.00012179525219835341
step: 160, loss: 2.8087331884307787e-05
step: 170, loss: 2.8016818760079332e-05
step: 180, loss: 6.478762225015089e-05
step: 190, loss: 0.00029081100365146995
step: 200, loss: 5.568227425101213e-05
step: 210, loss: 4.245418313075788e-05
step: 220, loss: 4.243089279043488e-05
step: 230, loss: 5.48144853382837e-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=0.9864559819413092, f1=0.9762174405436014,
best_f1=0.9819004524886877
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.")
2473it [00:12, 202.86it/s]
load_f1 = 0.992108229988726
real_f1 = 0.990990990990991
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 [00:17, 247.61it/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 bert-base-uncased \
       --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 bert-base-uncased \
       --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: 0.623701810836792
step: 10, loss: 0.4945133924484253
step: 20, loss: 0.4843423068523407
step: 30, loss: 0.11872131377458572
step: 40, loss: 0.09801803529262543
step: 50, loss: 0.26610568165779114
step: 60, loss: 0.022304080426692963
step: 70, loss: 0.12701059877872467
step: 80, loss: 0.10472144186496735
step: 90, loss: 0.16097275912761688
step: 100, loss: 0.03193754330277443
step: 110, loss: 0.11913684010505676
step: 120, loss: 0.08952952921390533
step: 130, loss: 0.06178748980164528
step: 140, loss: 0.03615923970937729
step: 150, loss: 0.042633622884750366
step: 160, loss: 0.029341338202357292
step: 170, loss: 0.16557615995407104
```

```
step: 180, loss: 0.14337189495563507
step: 190, loss: 0.012324427254498005
step: 200, loss: 0.12273643165826797
step: 210, loss: 0.16687554121017456
step: 220, loss: 0.2613532841205597
step: 230, loss: 0.1444852650165558
step: 240, loss: 0.08140435814857483
step: 250, loss: 0.03508973866701126
step: 260, loss: 0.046839095652103424
step: 270, loss: 0.019742758944630623
step: 280, loss: 0.048276014626026154
step: 290, loss: 0.29139450192451477
step: 300, loss: 0.05600406229496002
step: 310, loss: 0.3354531526565552
step: 320, loss: 0.08252713084220886
step: 330, loss: 0.018460141494870186
step: 340, loss: 0.03173963353037834
step: 350, loss: 0.06917209923267365
step: 360, loss: 0.05163748189806938
step: 370, loss: 0.02933325432240963
step: 380, loss: 0.022103937342762947
step: 390, loss: 0.13489480316638947
step: 400, loss: 0.20461997389793396
step: 410, loss: 0.08663024753332138
step: 420, loss: 0.07537219673395157
step: 430, loss: 0.17122811079025269
step: 440, loss: 0.01581212319433689
step: 450, loss: 0.0031472190748900175
step: 460, loss: 0.0033160774037241936
step: 470, loss: 0.15659558773040771
step: 480, loss: 0.030907146632671356
step: 490, loss: 0.05060678347945213
step: 500, loss: 0.04918743297457695
step: 510, loss: 0.03545137122273445
step: 520, loss: 0.023964356631040573
step: 530, loss: 0.005591580644249916
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9252336448598131, f1=0.9255813953488372,
best_f1=0.9255813953488372
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
```

```
step: 0, loss: 0.2457345873117447
step: 10, loss: 0.09332336485385895
step: 20, loss: 0.04714571312069893
step: 30, loss: 0.009931259788572788
step: 40, loss: 0.07230591773986816
step: 50, loss: 0.09544889628887177
step: 60, loss: 0.009570027701556683
step: 70, loss: 0.02275906875729561
step: 80, loss: 0.08836831897497177
step: 90, loss: 0.014541668817400932
step: 100, loss: 0.007802006788551807
step: 110, loss: 0.004815710708498955
step: 120, loss: 0.1524265557527542
step: 130, loss: 0.1842881143093109
step: 140, loss: 0.01924448274075985
step: 150, loss: 0.018032241612672806
step: 160, loss: 0.006641145329922438
step: 170, loss: 0.03409732133150101
step: 180, loss: 0.09874418377876282
step: 190, loss: 0.14129221439361572
step: 200, loss: 0.006501054856926203
step: 210, loss: 0.08770856261253357
step: 220, loss: 0.11160922050476074
step: 230, loss: 0.01371404342353344
step: 240, loss: 0.10284619778394699
step: 250, loss: 0.1511460840702057
step: 260, loss: 0.008983870036900043
step: 270, loss: 0.11685728281736374
step: 280, loss: 0.10977207869291306
step: 290, loss: 0.012493791989982128
step: 300, loss: 0.15598392486572266
step: 310, loss: 0.013718869537115097
step: 320, loss: 0.10214202105998993
step: 330, loss: 0.04315662756562233
step: 340, loss: 0.012724166736006737
step: 350, loss: 0.0023688245564699173
step: 360, loss: 0.04496285319328308
step: 370, loss: 0.07883835583925247
step: 380, loss: 0.04705101251602173
step: 390, loss: 0.05528027191758156
step: 400, loss: 0.07638463377952576
step: 410, loss: 0.014742808416485786
step: 420, loss: 0.08849682658910751
step: 430, loss: 0.011346842162311077
step: 440, loss: 0.20701266825199127
step: 450, loss: 0.03397276625037193
step: 460, loss: 0.01170344278216362
step: 470, loss: 0.09013195335865021
```

```
step: 480, loss: 0.26276955008506775
step: 490, loss: 0.02012096904218197
step: 500, loss: 0.18753519654273987
step: 510, loss: 0.010199490934610367
step: 520, loss: 0.03160173445940018
step: 530, loss: 0.007118913810700178
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9384404924760602, f1=0.9347329986307622,
best_f1=0.9347329986307622
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.09993357211351395
step: 10, loss: 0.026116138324141502
step: 20, loss: 0.044791728258132935
step: 30, loss: 0.07612120360136032
step: 40, loss: 0.014292768202722073
step: 50, loss: 0.03884632885456085
step: 60, loss: 0.025354942306876183
step: 70, loss: 0.007498187944293022
step: 80, loss: 0.003285334212705493
step: 90, loss: 0.003967566881328821
step: 100, loss: 0.0633479580283165
step: 110, loss: 0.0033753076568245888
step: 120, loss: 0.0070038288831710815
step: 130, loss: 0.003102570306509733
step: 140, loss: 0.059780266135931015
step: 150, loss: 0.06304875016212463
step: 160, loss: 0.00854500848799944
step: 170, loss: 0.1195751428604126
step: 180, loss: 0.0062765381298959255
step: 190, loss: 0.00620250916108489
step: 200, loss: 0.005108142737299204
step: 210, loss: 0.17528173327445984
step: 220, loss: 0.06703665107488632
step: 230, loss: 0.16429346799850464
step: 240, loss: 0.0013828587252646685
step: 250, loss: 0.04333362355828285
step: 260, loss: 0.026347726583480835
step: 270, loss: 0.005160727072507143
step: 280, loss: 0.20104379951953888
step: 290, loss: 0.005163144785910845
```

```
step: 300, loss: 0.053770214319229126
step: 310, loss: 0.03344222530722618
step: 320, loss: 0.01789964735507965
step: 330, loss: 0.0047316597774624825
step: 340, loss: 0.007677934598177671
step: 350, loss: 0.004512555431574583
step: 360, loss: 0.049707964062690735
step: 370, loss: 0.005383518058806658
step: 380, loss: 0.04153832420706749
step: 390, loss: 0.009643038734793663
step: 400, loss: 0.09429330378770828
step: 410, loss: 0.007349151652306318
step: 420, loss: 0.15440435707569122
step: 430, loss: 0.00842659454792738
step: 440, loss: 0.008188154548406601
step: 450, loss: 0.06800371408462524
step: 460, loss: 0.013571567833423615
step: 470, loss: 0.047313932329416275
step: 480, loss: 0.00288943643681705
step: 490, loss: 0.010064059868454933
step: 500, loss: 0.008483384735882282
step: 510, loss: 0.011150562204420567
step: 520, loss: 0.05916542559862137
step: 530, loss: 0.07660632580518723
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.942951438000943, f1=0.9248937175247992,
best_f1=0.9248937175247992
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0022343031596392393
step: 10, loss: 0.004047555383294821
step: 20, loss: 0.00051093695219660616
step: 30, loss: 0.00030681132921017706
step: 40, loss: 0.0009334409842267632
step: 50, loss: 0.0015619012992829084
step: 60, loss: 0.013950572349131107
step: 70, loss: 0.01072100829333067
step: 80, loss: 0.0011664509074762464
step: 90, loss: 0.04720030725002289
step: 100, loss: 0.0023866414558142424
step: 110, loss: 0.006206636317074299
```

```
step: 120, loss: 0.0164936613291502
step: 130, loss: 0.0015860182465985417
step: 140, loss: 0.008011490106582642
step: 150, loss: 0.002635512501001358
step: 160, loss: 0.13738608360290527
step: 170, loss: 0.0036747646518051624
step: 180, loss: 0.001766571425832808
step: 190, loss: 0.002567718271166086
Gradient overflow.  Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 200, loss: 0.09597384929656982
step: 210, loss: 0.08075617998838425
step: 220, loss: 0.004137587267905474
step: 230, loss: 0.00956612266600132
step: 240, loss: 0.005390509497374296
step: 250, loss: 0.005603599362075329
step: 260, loss: 0.006070077884942293
step: 270, loss: 0.005399823654443026
step: 280, loss: 0.018240416422486305
step: 290, loss: 0.029959436506032944
step: 300, loss: 0.0002079273690469563
Gradient overflow.  Skipping step, loss scaler 0 reducing loss scale to 8192.0
step: 310, loss: 0.0020800030324608088
step: 320, loss: 0.035774461925029755
step: 330, loss: 0.10830123722553253
step: 340, loss: 0.26464971899986267
step: 350, loss: 0.02999197319149971
step: 360, loss: 0.04763977974653244
step: 370, loss: 0.0023350936826318502
step: 380, loss: 0.005143157206475735
step: 390, loss: 0.002394250128418207
step: 400, loss: 0.09177491813898087
step: 410, loss: 0.003183964639902115
step: 420, loss: 0.01260372344404459
step: 430, loss: 0.23093119263648987
step: 440, loss: 0.013958356343209743
step: 450, loss: 0.012871598824858665
step: 460, loss: 0.0012399308616295457
step: 470, loss: 0.0182722788304905
step: 480, loss: 0.031076300889253616
step: 490, loss: 0.000704517588019371
step: 500, loss: 0.004578254651278257
step: 510, loss: 0.0038408248219639063
step: 520, loss: 0.015036883763968945
step: 530, loss: 0.01787586882710457
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
```

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=0.9435707678075855, f1=0.9355432780847146,
best_f1=0.9355432780847146
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.006591624114662409
step: 10, loss: 0.0755748599767685
step: 20, loss: 0.10575177520513535
step: 30, loss: 0.01595289073884487
step: 40, loss: 0.001620826544240117
step: 50, loss: 0.011536487378180027
step: 60, loss: 0.14158578217029572
step: 70, loss: 0.011118919588625431
step: 80, loss: 0.018060285598039627
step: 90, loss: 0.013694501481950283
step: 100, loss: 0.013533495366573334
step: 110, loss: 0.0002581815351732075
step: 120, loss: 0.0022587471175938845
step: 130, loss: 0.016949405893683434
step: 140, loss: 0.0004510303260758519
step: 150, loss: 0.013489976525306702
step: 160, loss: 0.00026604149024933577
step: 170, loss: 0.014019312337040901
step: 180, loss: 0.0007533621392212808
step: 190, loss: 0.0011723223142325878
step: 200, loss: 0.0009676067857071757
step: 210, loss: 0.06788473576307297
step: 220, loss: 0.001299989060498774
step: 230, loss: 0.008429566398262978
step: 240, loss: 0.00439410237595439
step: 250, loss: 0.001619584858417511
step: 260, loss: 0.0018543086480349302
step: 270, loss: 0.001476938370615244
step: 280, loss: 0.008012509904801846
step: 290, loss: 0.08182721585035324
step: 300, loss: 0.007485808804631233
step: 310, loss: 0.009081757627427578
step: 320, loss: 0.07404407113790512
step: 330, loss: 0.0056714825332164764
step: 340, loss: 0.0005204027984291315
step: 350, loss: 0.0037911077961325645
step: 360, loss: 0.009461716748774052
step: 370, loss: 0.08359763771295547
step: 380, loss: 0.0008900928660295904
step: 390, loss: 0.00015484794857911766

```
step: 400, loss: 0.0008689332753419876
step: 410, loss: 0.00043587805703282356
step: 420, loss: 0.00477600796148181
step: 430, loss: 0.0019182127434760332
step: 440, loss: 0.06071166694164276
step: 450, loss: 0.01992277055978775
step: 460, loss: 0.0011824461398646235
step: 470, loss: 0.021472634747624397
step: 480, loss: 0.004641628358513117
step: 490, loss: 0.0006958594894967973
step: 500, loss: 0.002315310761332512
step: 510, loss: 0.1160380095243454
step: 520, loss: 0.004147810395807028
step: 530, loss: 0.11369283497333527
/usr/local/lib/python3.7/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=0.9319323892188214, f1=0.9262672811059909,
best_f1=0.9355432780847146
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0005046886508353055
step: 10, loss: 0.04343898221850395
step: 20, loss: 0.0715894028544426
step: 30, loss: 0.0006135411094874144
step: 40, loss: 0.00068857561529614031
step: 50, loss: 0.07358212023973465
step: 60, loss: 0.00021978377480991185
step: 70, loss: 0.0006023208261467516
step: 80, loss: 0.006626402493566275
step: 90, loss: 0.0002465713187120855
step: 100, loss: 0.0090911195650398731
step: 110, loss: 0.003815107047557831
step: 120, loss: 0.0006481820018962026
step: 130, loss: 0.017835542559623718
step: 140, loss: 0.007887715473771095
step: 150, loss: 0.0005993720842525363
step: 160, loss: 0.003993155434727669
step: 170, loss: 0.0009693157044239342
step: 180, loss: 0.000168906626640819
step: 190, loss: 0.0005757362232543528
step: 200, loss: 0.023878157138824463
step: 210, loss: 0.0017647066852077842
step: 220, loss: 0.0019719810225069523
step: 230, loss: 0.00067393347853794699
step: 240, loss: 0.006507054436951876
step: 250, loss: 0.001119795604608953
```

```
step: 260, loss: 0.0017466249410063028
step: 270, loss: 0.08607885986566544
step: 280, loss: 0.01240420900285244
step: 290, loss: 0.0005156475235708058
step: 300, loss: 0.0006546066724695265
step: 310, loss: 0.0007575028575956821
step: 320, loss: 0.00043696127249859273
step: 330, loss: 0.0007488135015591979
step: 340, loss: 0.05457005277276039
step: 350, loss: 0.004152423702180386
step: 360, loss: 0.008092576637864113
step: 370, loss: 0.0024012576323002577
step: 380, loss: 0.04405370354652405
step: 390, loss: 0.014109142124652863
step: 400, loss: 7.208426541183144e-05
step: 410, loss: 0.0003562263445928693
step: 420, loss: 0.004188107326626778
step: 430, loss: 0.00012165070074843243
step: 440, loss: 0.00015048369823489338
step: 450, loss: 0.043448206037282944
step: 460, loss: 0.0009914875263348222
step: 470, loss: 0.000982682453468442
step: 480, loss: 0.008275349624454975
step: 490, loss: 0.019931679591536522
step: 500, loss: 0.00493704853579402
step: 510, loss: 0.00034518196480348706
step: 520, loss: 0.020769720897078514
step: 530, loss: 0.01788395456969738
/usr/local/lib/python3.7/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=0.935831381733021, f1=0.9246511627906977,
best_f1=0.9355432780847146
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.000597335456404835
step: 10, loss: 0.0005497077945619822
step: 20, loss: 0.00089346665820561349
step: 30, loss: 0.012620136141777039
step: 40, loss: 8.565274038119242e-05
step: 50, loss: 0.0013460279442369938
step: 60, loss: 0.008171312510967255
step: 70, loss: 0.00030869783950038254
step: 80, loss: 0.0016263002762570977
step: 90, loss: 0.003173267701640725
step: 100, loss: 0.008326356299221516
step: 110, loss: 0.0033809104934334755
```

```
step: 120, loss: 0.0005970386555418372
step: 130, loss: 0.020606689155101776
step: 140, loss: 0.0005651454557664692
step: 150, loss: 0.001046106917783618
step: 160, loss: 8.863180846674368e-05
step: 170, loss: 0.00040038322913460433
step: 180, loss: 0.0023693435359746218
step: 190, loss: 0.12230940908193588
step: 200, loss: 0.0004249429621268064
step: 210, loss: 0.12698183953762054
step: 220, loss: 0.006071074400097132
step: 230, loss: 0.027336986735463142
step: 240, loss: 0.0018735831836238503
step: 250, loss: 0.003760288469493389
step: 260, loss: 0.00010671155177988112
step: 270, loss: 0.0009831914212554693
step: 280, loss: 0.00011506000009831041
step: 290, loss: 0.00013806014612782747
step: 300, loss: 0.0013565237168222666
step: 310, loss: 7.564183761132881e-05
step: 320, loss: 6.596014281967655e-05
step: 330, loss: 0.0007624577265232801
step: 340, loss: 0.02989668771624565
step: 350, loss: 0.0001535210758447647
step: 360, loss: 0.007193200755864382
step: 370, loss: 0.00019146222262120247
step: 380, loss: 0.0001365290954709053
step: 390, loss: 0.00043114711297675967
step: 400, loss: 0.0007166724535636604
step: 410, loss: 0.00014692039985675365
step: 420, loss: 0.0273460540920496
step: 430, loss: 0.00010106937406817451
step: 440, loss: 0.0006157496245577931
step: 450, loss: 0.001103741000406444
step: 460, loss: 0.00016547339328099042
step: 470, loss: 0.00270066992379725
step: 480, loss: 0.004776963498443365
step: 490, loss: 0.07019282877445221
step: 500, loss: 0.00020263949409127235
step: 510, loss: 0.00010529880819376558
step: 520, loss: 0.0312512032687664
step: 530, loss: 0.002824777038767934
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9463459759481962, f1=0.936111111111111,
best_f1=0.936111111111111
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00030552438693121076
step: 10, loss: 0.02917538583278656
step: 20, loss: 5.01362701470498e-05
step: 30, loss: 0.00026371105923317373
step: 40, loss: 0.001738260849379003
step: 50, loss: 0.005900297313928604
step: 60, loss: 0.004279551561921835
step: 70, loss: 7.643007847946137e-05
step: 80, loss: 0.0001579958334332332
step: 90, loss: 0.00014432729221880436
step: 100, loss: 8.238048030762002e-05
step: 110, loss: 0.000270967953838408
step: 120, loss: 0.066722072660923
step: 130, loss: 0.00047754429397173226
step: 140, loss: 0.0297384113073349
step: 150, loss: 8.921051630750299e-05
step: 160, loss: 0.01645777001976967
step: 170, loss: 0.001375530264340341
step: 180, loss: 0.0004301028384361416
step: 190, loss: 0.00025352791999466717
step: 200, loss: 0.00014421796367969364
step: 210, loss: 0.0018637500470504165
step: 220, loss: 0.0017514284700155258
step: 230, loss: 0.0011533888755366206
step: 240, loss: 0.000780132133513689
step: 250, loss: 0.00016217563825193793
step: 260, loss: 7.197131344582886e-05
step: 270, loss: 0.0056600552052259445
step: 280, loss: 0.06025013327598572
step: 290, loss: 0.002882894594222307
step: 300, loss: 0.0014669212978333235
step: 310, loss: 0.00032781969639472663
step: 320, loss: 0.0015362645499408245
step: 330, loss: 0.00014124519657343626
step: 340, loss: 0.026928558945655823
step: 350, loss: 0.00039271573768928647
step: 360, loss: 0.00015756157517898828
step: 370, loss: 0.0001609845639904961
step: 380, loss: 0.008608348667621613
step: 390, loss: 0.11245760321617126
step: 400, loss: 0.00025591018493287265
step: 410, loss: 6.57343480270356e-05
```

```
step: 420, loss: 0.00023768868413753808
step: 430, loss: 0.054770469665527344
step: 440, loss: 0.00205816188827157
step: 450, loss: 0.003124341368675232
step: 460, loss: 0.0016856420552358031
step: 470, loss: 7.800924504408613e-05
step: 480, loss: 6.860475696157664e-05
step: 490, loss: 0.000833055644761771
step: 500, loss: 0.0035217381082475185
step: 510, loss: 5.249390596873127e-05
step: 520, loss: 6.191492138896137e-05
step: 530, loss: 3.340310649946332e-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=0.9423347398030943, f1=0.9317004239284032,
best_f1=0.936111111111111
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00022621643438469619
step: 10, loss: 0.000896924058906734
step: 20, loss: 0.0003844252787530422
step: 30, loss: 0.0004838519962504506
step: 40, loss: 0.0010180543176829815
step: 50, loss: 3.1864317861618474e-05
step: 60, loss: 4.1171955672325566e-05
step: 70, loss: 0.00015092660032678396
step: 80, loss: 0.0181149672716856
step: 90, loss: 9.768502786755562e-05
step: 100, loss: 9.018395212478936e-05
step: 110, loss: 4.7312056267401204e-05
step: 120, loss: 6.948080408619717e-05
step: 130, loss: 0.00801053736358881
step: 140, loss: 0.0001583405683049932
step: 150, loss: 4.655354132410139e-05
step: 160, loss: 0.0008647011127322912
step: 170, loss: 0.003923120442777872
step: 180, loss: 0.0009081456228159368
step: 190, loss: 0.0002502574643585831
step: 200, loss: 0.0001471615832997486
step: 210, loss: 0.0001239240518771112
step: 220, loss: 0.00032156682573258877
step: 230, loss: 5.038714880356565e-05
step: 240, loss: 0.0001675678649917245
step: 250, loss: 8.496308873873204e-05
step: 260, loss: 0.0051398747600615025
step: 270, loss: 3.900654337485321e-05
```

```
step: 280, loss: 0.00044282973976805806
step: 290, loss: 0.03713442012667656
step: 300, loss: 0.00027403084095567465
step: 310, loss: 0.040152616798877716
step: 320, loss: 0.0005926844896748662
step: 330, loss: 0.0006021006847731769
step: 340, loss: 0.0005994337261654437
step: 350, loss: 0.001784753636457026
step: 360, loss: 0.00010028797260019928
step: 370, loss: 0.006189106032252312
step: 380, loss: 0.00013336492702364922
step: 390, loss: 8.875445928424597e-05
step: 400, loss: 0.001001322758384049
step: 410, loss: 5.7995297538582236e-05
step: 420, loss: 0.0009527879301458597
step: 430, loss: 7.119965448509902e-05
step: 440, loss: 0.007255688309669495
step: 450, loss: 0.0007649409235455096
step: 460, loss: 8.944931323640049e-05
step: 470, loss: 6.172717257868499e-05
step: 480, loss: 0.0001066049953806214
step: 490, loss: 0.0012126382207497954
step: 500, loss: 0.00014789614942856133
step: 510, loss: 0.0005021662800572813
step: 520, loss: 0.00012733486073557287
step: 530, loss: 0.00037053803680464625
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9488372093023255, f1=0.9356183418249191,
best_f1=0.9356183418249191
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0004827382508665323
step: 10, loss: 4.816809814656153e-05
step: 20, loss: 3.857330375467427e-05
step: 30, loss: 3.8244812458287925e-05
step: 40, loss: 0.00013440399197861552
step: 50, loss: 0.0001894456654554233
step: 60, loss: 9.143861097982153e-05
step: 70, loss: 4.948638525092974e-05
step: 80, loss: 3.005466533068102e-05
step: 90, loss: 9.05727210920304e-05
```

```
step: 100, loss: 4.7986264689825475e-05
step: 110, loss: 1.9076860553468578e-05
step: 120, loss: 4.6421238948823884e-05
step: 130, loss: 0.0003997432650066912
step: 140, loss: 0.002562482375651598
step: 150, loss: 6.391377246472985e-05
step: 160, loss: 0.0001542846584925428
step: 170, loss: 4.6134220610838383e-05
step: 180, loss: 3.629050843301229e-05
step: 190, loss: 5.2209383284207433e-05
step: 200, loss: 9.330963803222403e-05
step: 210, loss: 7.701084541622549e-05
step: 220, loss: 0.003957867156714201
step: 230, loss: 0.0020757820457220078
step: 240, loss: 4.4707114284392446e-05
step: 250, loss: 4.2182735342066735e-05
step: 260, loss: 5.192250682739541e-05
step: 270, loss: 5.8716119383461773e-05
step: 280, loss: 3.418556661927141e-05
step: 290, loss: 3.0877654353389516e-05
step: 300, loss: 2.4347929866053164e-05
step: 310, loss: 8.059187530307099e-05
step: 320, loss: 8.276756125269458e-05
step: 330, loss: 4.734910180559382e-05
step: 340, loss: 0.005160999950021505
step: 350, loss: 0.0006954938289709389
step: 360, loss: 4.328727300162427e-05
step: 370, loss: 0.00043943943455815315
step: 380, loss: 0.0031417496502399445
step: 390, loss: 9.065320045920089e-05
step: 400, loss: 6.016994302626699e-05
step: 410, loss: 0.00013748122728429735
step: 420, loss: 5.6598280934849754e-05
step: 430, loss: 0.007193821482360363
step: 440, loss: 0.01556831318885088
step: 450, loss: 0.00013936235336586833
step: 460, loss: 0.00014894199557602406
step: 470, loss: 0.00023645082546863705
step: 480, loss: 4.9425652832724154e-05
step: 490, loss: 0.011545627377927303
step: 500, loss: 0.003590164938941598
step: 510, loss: 0.00013562914682552218
step: 520, loss: 0.00047739126603119075
step: 530, loss: 0.000136528629809618
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
```

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=0.9494855004677268, f1=0.9364269141531322,
best_f1=0.9364269141531322
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0010647047311067581
step: 10, loss: 0.003976549953222275
step: 20, loss: 6.573754944838583e-05
step: 30, loss: 5.029698149883188e-05
step: 40, loss: 0.0014686079230159521
step: 50, loss: 9.361447882838547e-05
step: 60, loss: 0.016615290194749832
step: 70, loss: 7.953529711812735e-05
step: 80, loss: 0.0014585808385163546
step: 90, loss: 9.052921086549759e-05
step: 100, loss: 0.0003699812514241785
step: 110, loss: 7.804340566508472e-05
step: 120, loss: 7.625245052622631e-05
step: 130, loss: 2.558095548010897e-05
step: 140, loss: 0.0007713122759014368
step: 150, loss: 2.3412816517520696e-05
step: 160, loss: 2.735386078711599e-05
step: 170, loss: 3.560093682608567e-05
step: 180, loss: 3.365305019542575e-05
step: 190, loss: 0.00021954927069600672
step: 200, loss: 0.0017981103155761957
step: 210, loss: 4.4275417167227715e-05
step: 220, loss: 9.58348682615906e-05
step: 230, loss: 4.8657409934094176e-05
step: 240, loss: 2.9957174774608575e-05
step: 250, loss: 2.168812534364406e-05
step: 260, loss: 3.9917504182085395e-05
step: 270, loss: 3.173856021021493e-05
step: 280, loss: 0.01785924658179283
step: 290, loss: 0.00011738097964553162
step: 300, loss: 0.00010066694812849164
step: 310, loss: 3.0017052267794497e-05
step: 320, loss: 5.0597722292877734e-05
step: 330, loss: 5.258203964331187e-05
step: 340, loss: 7.179451495176181e-05
step: 350, loss: 0.03968669846653938
step: 360, loss: 4.344941407907754e-05
step: 370, loss: 0.00024267312255688012
step: 380, loss: 2.6761714252643287e-05
step: 390, loss: 3.9825939893489704e-05

```
step: 400, loss: 0.0042685153894126415
step: 410, loss: 3.253190516261384e-05
step: 420, loss: 0.00042324053356423974
step: 430, loss: 3.7436919228639454e-05
step: 440, loss: 0.0054803467355668545
step: 450, loss: 3.477007703622803e-05
step: 460, loss: 0.0031426637433469296
step: 470, loss: 9.93408466456458e-05
step: 480, loss: 5.2718249207828194e-05
step: 490, loss: 3.8963928091106936e-05
step: 500, loss: 4.4796368456445634e-05
step: 510, loss: 8.027190051507205e-05
step: 520, loss: 6.566941738128662e-05
step: 530, loss: 0.00035837641917169094
/usr/local/lib/python3.7/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=0.9428436466698158, f1=0.9317004239284032,
best_f1=0.9364269141531322
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 8.459534001303837e-05
step: 10, loss: 1.819408316805493e-05
step: 20, loss: 2.08872770595555522e-05
step: 30, loss: 0.0006088250665925443
step: 40, loss: 3.631127401604317e-05
step: 50, loss: 0.05328140780329704
step: 60, loss: 0.00050997355572002828
step: 70, loss: 0.00026594602968543777
step: 80, loss: 3.865825419779867e-05
step: 90, loss: 0.00010625830327626318
step: 100, loss: 2.5122242732322775e-05
step: 110, loss: 0.00012783434067387134
step: 120, loss: 3.624366581789218e-05
step: 130, loss: 3.582212229957804e-05
step: 140, loss: 2.4023685909924097e-05
step: 150, loss: 2.2574749891646206e-05
step: 160, loss: 1.931908991537057e-05
step: 170, loss: 3.036698035430163e-05
step: 180, loss: 4.6220895455917343e-05
step: 190, loss: 2.3662458261242136e-05
step: 200, loss: 5.6490876886527985e-05
step: 210, loss: 4.383763007353991e-05
step: 220, loss: 3.469974035397172e-05
step: 230, loss: 5.2473864343483e-05
step: 240, loss: 2.2790620278101414e-05
step: 250, loss: 1.7683676560409367e-05
```

```
step: 260, loss: 6.064464832888916e-05
step: 270, loss: 2.9134846045053564e-05
step: 280, loss: 0.00013019927428103983
step: 290, loss: 0.0001418616739101708
step: 300, loss: 2.6362966309534386e-05
step: 310, loss: 0.0001848804095061496
step: 320, loss: 1.5716814232291654e-05
step: 330, loss: 2.050735565717332e-05
step: 340, loss: 2.0175786630716175e-05
step: 350, loss: 7.72670391597785e-05
step: 360, loss: 0.0019275328377261758
step: 370, loss: 0.00025304959854111075
step: 380, loss: 3.852199006360024e-05
step: 390, loss: 0.000350680376868695
step: 400, loss: 1.8447357433615252e-05
step: 410, loss: 0.000121142111311201
step: 420, loss: 0.004787764512002468
step: 430, loss: 0.012519774958491325
step: 440, loss: 0.0026029450818896294
step: 450, loss: 0.0007737683481536806
step: 460, loss: 7.308842032216489e-05
step: 470, loss: 4.295016697142273e-05
step: 480, loss: 0.0008513072389177978
step: 490, loss: 6.115691940067336e-05
step: 500, loss: 0.09146160632371902
step: 510, loss: 0.002074414398521185
step: 520, loss: 0.18205951154232025
step: 530, loss: 2.5640156309236772e-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=0.9408450704225352, f1=0.9317757009345794,
best_f1=0.9364269141531322
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 3.201825893484056e-05
step: 10, loss: 1.5511917808908038e-05
step: 20, loss: 1.2446044820535462e-05
step: 30, loss: 2.4314425900229253e-05
step: 40, loss: 2.5978255507652648e-05
step: 50, loss: 8.003312541404739e-05
step: 60, loss: 0.0001803119812393561
step: 70, loss: 1.7333479263470508e-05
step: 80, loss: 1.3921233403380029e-05
step: 90, loss: 0.03282007575035095
step: 100, loss: 0.0027372322510927916
step: 110, loss: 3.6818160879192874e-05
```

```
step: 120, loss: 7.8066426794976e-05
step: 130, loss: 1.5638528566341847e-05
step: 140, loss: 1.9762217561947182e-05
step: 150, loss: 2.5326993636554107e-05
step: 160, loss: 1.7169568309327587e-05
step: 170, loss: 0.00019523125956766307
step: 180, loss: 2.3598893676535226e-05
step: 190, loss: 3.304578422103077e-05
step: 200, loss: 2.2279880795395002e-05
step: 210, loss: 4.676529351854697e-05
step: 220, loss: 2.287170718773268e-05
step: 230, loss: 1.3079367818136234e-05
step: 240, loss: 3.784167711273767e-05
step: 250, loss: 1.7579246559762396e-05
step: 260, loss: 7.715277752140537e-05
step: 270, loss: 1.3794588085147552e-05
step: 280, loss: 3.803240178967826e-05
step: 290, loss: 2.0451236196095124e-05
step: 300, loss: 3.614879460656084e-05
step: 310, loss: 0.00018852738139685243
step: 320, loss: 3.271270543336868e-05
step: 330, loss: 5.976256215944886e-05
step: 340, loss: 2.4057027985691093e-05
step: 350, loss: 3.7555426388280466e-05
step: 360, loss: 2.8964850571355782e-05
step: 370, loss: 6.5422151123772621e-05
step: 380, loss: 0.000237657455727458
step: 390, loss: 0.0038222612347453833
step: 400, loss: 1.58991551870713e-05
step: 410, loss: 3.452642704360187e-05
step: 420, loss: 1.748623253661208e-05
step: 430, loss: 0.003484277054667473
step: 440, loss: 3.175454185111448e-05
step: 450, loss: 2.655286698427517e-05
step: 460, loss: 5.479627725435421e-05
step: 470, loss: 1.1965511475864332e-05
step: 480, loss: 2.8227506845723838e-05
step: 490, loss: 2.7814890927402303e-05
step: 500, loss: 1.3362364370550495e-05
step: 510, loss: 8.30833159852773e-05
step: 520, loss: 0.0003628152480814606
step: 530, loss: 1.4200631994754076e-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=0.9392369288742346, f1=0.9324577861163227,
best_f1=0.9364269141531322
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
```

```
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 2.6541600163909607e-05
step: 10, loss: 1.2285887351026759e-05
step: 20, loss: 1.9710058040800504e-05
step: 30, loss: 2.8682692573056556e-05
step: 40, loss: 3.0380651878658682e-05
step: 50, loss: 0.0005022357800044119
step: 60, loss: 1.6368541764677502e-05
step: 70, loss: 1.7940576071850955e-05
step: 80, loss: 1.9624507331172936e-05
step: 90, loss: 2.940807280538138e-05
step: 100, loss: 1.923678610182833e-05
step: 110, loss: 2.1969872250338085e-05
step: 120, loss: 2.147570194210857e-05
step: 130, loss: 0.013411998748779297
step: 140, loss: 2.3917797079775482e-05
step: 150, loss: 1.3813196346745826e-05
step: 160, loss: 2.2793918105890043e-05
step: 170, loss: 2.4290990040753968e-05
step: 180, loss: 6.418691191356629e-05
step: 190, loss: 5.3929721616441384e-05
step: 200, loss: 2.408320506219752e-05
step: 210, loss: 2.1863015717826784e-05
step: 220, loss: 1.6223422790062614e-05
step: 230, loss: 6.910636147949845e-05
step: 240, loss: 1.1339691809553187e-05
step: 250, loss: 3.05238354485482e-05
step: 260, loss: 2.062636303890031e-05
step: 270, loss: 2.853445039363578e-05
step: 280, loss: 1.4033002116775606e-05
step: 290, loss: 3.1669089366914704e-05
step: 300, loss: 2.5211678803316317e-05
step: 310, loss: 6.554162973770872e-05
step: 320, loss: 2.8787160772481002e-05
step: 330, loss: 8.084272121777758e-05
step: 340, loss: 2.5181443561450578e-05
Gradient overflow.  Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 350, loss: 1.6800819139461964e-05
step: 360, loss: 0.00014425716653931886
step: 370, loss: 1.5437399270012975e-05
step: 380, loss: 2.2905951482243836e-05
step: 390, loss: 0.03165927901864052
step: 400, loss: 4.278244523447938e-05
step: 410, loss: 1.6409489035140723e-05
step: 420, loss: 1.5269675714080222e-05
step: 430, loss: 0.0002851001627277583
step: 440, loss: 3.907041173079051e-05
```

```
step: 450, loss: 2.8925525839440525e-05
step: 460, loss: 1.7545800801599398e-05
step: 470, loss: 1.645409793127328e-05
step: 480, loss: 1.949770921783056e-05
step: 490, loss: 2.669720743142534e-05
step: 500, loss: 1.2181561942270491e-05
step: 510, loss: 2.2846239517093636e-05
step: 520, loss: 1.418564988853177e-05
step: 530, loss: 2.631590177770704e-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=0.9406819243344232, f1=0.9326521133302369,
best_f1=0.9364269141531322
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 1.5813584468560293e-05
step: 10, loss: 2.0275943825254217e-05
step: 20, loss: 2.111365756718442e-05
step: 30, loss: 0.002570496406406164
step: 40, loss: 0.04804998263716698
step: 50, loss: 2.29470242629759e-05
step: 60, loss: 1.0445644875289872e-05
step: 70, loss: 4.975925912731327e-05
step: 80, loss: 1.7430227671866305e-05
step: 90, loss: 2.1609572286251932e-05
step: 100, loss: 0.00014972231292631477
step: 110, loss: 1.5321913451771252e-05
step: 120, loss: 0.006603142246603966
step: 130, loss: 0.052741099148988724
step: 140, loss: 1.785121094144415e-05
step: 150, loss: 1.778784098860342e-05
step: 160, loss: 2.249584395030979e-05
step: 170, loss: 1.6674099242663942e-05
step: 180, loss: 1.2825988960685208e-05
step: 190, loss: 3.353107604198158e-05
step: 200, loss: 2.2867652660352178e-05
step: 210, loss: 3.032918721146416e-05
step: 220, loss: 1.426765629730653e-05
step: 230, loss: 5.510833580046892e-05
step: 240, loss: 1.5910449292277917e-05
step: 250, loss: 8.344577508978546e-05
step: 260, loss: 1.855158552643843e-05
step: 270, loss: 2.9738970624748617e-05
step: 280, loss: 1.8662931324797682e-05
step: 290, loss: 2.4673145162523724e-05
step: 300, loss: 2.9229966457933187e-05
```

```
step: 310, loss: 4.2316580220358446e-05
step: 320, loss: 5.865887214895338e-05
step: 330, loss: 5.933361535426229e-05
step: 340, loss: 1.1425369848439004e-05
step: 350, loss: 1.0944801033474505e-05
step: 360, loss: 5.124739254824817e-05
step: 370, loss: 1.050523769663414e-05
step: 380, loss: 1.849574073276017e-05
step: 390, loss: 3.5428147384664044e-05
step: 400, loss: 1.5727971913293004e-05
step: 410, loss: 1.2945251910423394e-05
step: 420, loss: 2.5777664632187225e-05
step: 430, loss: 7.73898427723907e-05
step: 440, loss: 2.174004475818947e-05
step: 450, loss: 1.6740836144890636e-05
step: 460, loss: 1.3146413039066829e-05
step: 470, loss: 0.010207602754235268
step: 480, loss: 1.1403010830690619e-05
step: 490, loss: 1.259881173609756e-05
step: 500, loss: 6.125640356913209e-05
step: 510, loss: 1.6364625480491668e-05
step: 520, loss: 1.4446452041738667e-05
step: 530, loss: 3.222087980248034e-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=0.9404536862003782, f1=0.9317647058823528,
best_f1=0.9364269141531322
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 [00:21, 261.19it/s]
load_f1 = 0.9496738117427772
real_f1 = 0.9473684210526314
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 [00:17, 255.67it/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 bert-base-uncased \
       --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 bert-base-uncased \
       --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: 0.553176999092102
step: 10, loss: 0.37075871229171753
step: 20, loss: 0.3815804421901703
step: 30, loss: 0.32474467158317566
step: 40, loss: 0.18695521354675293
step: 50, loss: 0.47336795926094055
step: 60, loss: 0.34243133664131165
step: 70, loss: 0.19472730159759521
step: 80, loss: 0.1865738481283188

```
step: 90, loss: 0.32823246717453003
step: 100, loss: 0.4229927062988281
step: 110, loss: 0.28339892625808716
step: 120, loss: 0.30333632230758667
step: 130, loss: 0.2383698970079422
step: 140, loss: 0.209610253572464
step: 150, loss: 0.1403825879096985
step: 160, loss: 0.3175959885120392
step: 170, loss: 0.27820149064064026
step: 180, loss: 0.1085294559597969
step: 190, loss: 0.1558297872543335
step: 200, loss: 0.26177698373794556
step: 210, loss: 0.23595941066741943
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.6309278350515464, f1=0.6778242677824268,
best_f1=0.6778242677824268
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.13386915624141693
step: 10, loss: 0.1764293611049652
step: 20, loss: 0.1914094090461731
step: 30, loss: 0.1546858698129654
step: 40, loss: 0.1810772866010666
step: 50, loss: 0.17681853473186493
step: 60, loss: 0.3216722011566162
step: 70, loss: 0.12081484496593475
step: 80, loss: 0.14184829592704773
step: 90, loss: 0.05664047971367836
step: 100, loss: 0.014181986451148987
step: 110, loss: 0.15990062057971954
step: 120, loss: 0.16830426454544067
step: 130, loss: 0.010728895664215088
step: 140, loss: 0.26073896884918213
step: 150, loss: 0.20040559768676758
step: 160, loss: 0.13124924898147583
step: 170, loss: 0.10871139913797379
step: 180, loss: 0.3088846206665039
step: 190, loss: 0.13027262687683105
step: 200, loss: 0.045634277164936066
step: 210, loss: 0.1846799999475479
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
```

```
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.6719056974459725, f1=0.7309236947791166,
best_f1=0.7309236947791166
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.06004958227276802
step: 10, loss: 0.10635025799274445
step: 20, loss: 0.05147949978709221
step: 30, loss: 0.15782535076141357
step: 40, loss: 0.116518035531044
step: 50, loss: 0.04973814636468887
step: 60, loss: 0.1907472163438797
step: 70, loss: 0.045914407819509506
step: 80, loss: 0.13354326784610748
step: 90, loss: 0.023892411962151527
step: 100, loss: 0.08620160073041916
step: 110, loss: 0.09644295275211334
step: 120, loss: 0.08247297257184982
step: 130, loss: 0.08084753155708313
step: 140, loss: 0.07109463959932327
step: 150, loss: 0.19519008696079254
step: 160, loss: 0.022435860708355904
step: 170, loss: 0.07534468919038773
step: 180, loss: 0.05536482855677605
step: 190, loss: 0.22243273258209229
step: 200, loss: 0.04985963553190231
step: 210, loss: 0.08377277851104736
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.6730038022813688, f1=0.6974951830443159,
best_f1=0.6974951830443159
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.14760351181030273
step: 10, loss: 0.04263206198811531
step: 20, loss: 0.1812637895345688
```

```
step: 30, loss: 0.12882211804389954
step: 40, loss: 0.007867935113608837
step: 50, loss: 0.19991566240787506
step: 60, loss: 0.08126000314950943
step: 70, loss: 0.12810015678405762
step: 80, loss: 0.129172682762146
step: 90, loss: 0.008817647583782673
step: 100, loss: 0.25365862250328064
step: 110, loss: 0.08368035405874252
step: 120, loss: 0.0619543120265007
step: 130, loss: 0.11860432475805283
step: 140, loss: 0.06808052957057953
step: 150, loss: 0.07050155848264694
step: 160, loss: 0.03716902807354927
step: 170, loss: 0.119008868932724
step: 180, loss: 0.3855869174003601
step: 190, loss: 0.06437060236930847
step: 200, loss: 0.24660737812519073
step: 210, loss: 0.18397054076194763
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.6851063829787234, f1=0.673773987206823,
best_f1=0.673773987206823
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.10411219298839569
step: 10, loss: 0.12383384257555008
step: 20, loss: 0.045382242649793625
step: 30, loss: 0.049464307725429535
Gradient overflow.  Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 40, loss: 0.03241768106818199
step: 50, loss: 0.007785408291965723
step: 60, loss: 0.16804976761341095
step: 70, loss: 0.13361811637878418
step: 80, loss: 0.0050644297152757645
step: 90, loss: 0.16677452623844147
step: 100, loss: 0.0031725012231618166
step: 110, loss: 0.20864784717559814
step: 120, loss: 0.05428912118077278
step: 130, loss: 0.08937251567840576
step: 140, loss: 0.019596533849835396
step: 150, loss: 0.0494365431368351
```

```
step: 160, loss: 0.03839796036481857
step: 170, loss: 0.04224742203950882
step: 180, loss: 0.08357441425323486
step: 190, loss: 0.008291158825159073
step: 200, loss: 0.052284691482782364
step: 210, loss: 0.026739904657006264
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.6926229508196722, f1=0.6903765690376569,
best_f1=0.6903765690376569
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.021278141066432
step: 10, loss: 0.00664261681959033
step: 20, loss: 0.008158817887306213
step: 30, loss: 0.00164721591863778121
step: 40, loss: 0.010487882420420647
step: 50, loss: 0.03871608152985573
step: 60, loss: 0.06270872056484222
step: 70, loss: 0.012674887664616108
step: 80, loss: 0.25282052159309387
step: 90, loss: 0.2391124963760376
step: 100, loss: 0.017678843811154366
step: 110, loss: 0.00875455979257822
step: 120, loss: 0.09099459648132324
step: 130, loss: 0.08683932572603226
step: 140, loss: 0.028514433652162552
step: 150, loss: 0.02107793092727661
step: 160, loss: 0.009277050383388996
step: 170, loss: 0.040710095316171646
step: 180, loss: 0.0080043850466609
step: 190, loss: 0.06980721652507782
step: 200, loss: 0.002887702314183116
step: 210, loss: 0.0030486697796732187
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.696, f1=0.6854838709677419, best_f1=0.6854838709677419
```

```
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.09157610684633255
step: 10, loss: 0.03134679049253464
step: 20, loss: 0.0027936608530580997
step: 30, loss: 0.016835199669003487
step: 40, loss: 0.15010465681552887
step: 50, loss: 0.06943988800048828
step: 60, loss: 0.006099372170865536
step: 70, loss: 0.015037044882774353
step: 80, loss: 0.03919978067278862
step: 90, loss: 0.0553741455078125
step: 100, loss: 0.0008478343952447176
step: 110, loss: 0.07308578491210938
step: 120, loss: 0.043857499957084656
step: 130, loss: 0.005885037127882242
step: 140, loss: 0.016769269481301308
step: 150, loss: 0.0022808313369750977
step: 160, loss: 0.12504923343658447
step: 170, loss: 0.004621051251888275
step: 180, loss: 0.126916766166687
step: 190, loss: 0.05096321552991867
step: 200, loss: 0.010616830550134182
step: 210, loss: 0.020096996799111366
/usr/local/lib/python3.7/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=0.680161943319838, f1=0.6893787575150301,
best_f1=0.6854838709677419
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.02273685112595558
step: 10, loss: 0.050212275236845016
step: 20, loss: 0.0015391485067084432
step: 30, loss: 0.03419705480337143
step: 40, loss: 0.002895776415243745
step: 50, loss: 0.0005456604994833469
step: 60, loss: 0.11564520746469498
step: 70, loss: 0.002704182406887412
step: 80, loss: 0.01624414511024952
step: 90, loss: 0.016332285478711128
step: 100, loss: 0.009256136603653431
step: 110, loss: 0.13362886011600494
step: 120, loss: 0.0013739484129473567
step: 130, loss: 0.0020197227131575346
step: 140, loss: 0.09246761351823807
```

```
step: 150, loss: 0.0029901848174631596
step: 160, loss: 0.024425847455859184
step: 170, loss: 0.017518620938062668
step: 180, loss: 0.0644061267375946
step: 190, loss: 0.04004389047622681
step: 200, loss: 0.003771659452468157
step: 210, loss: 0.08771924674510956
/usr/local/lib/python3.7/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=0.6747967479674798, f1=0.7157464212678937,
best_f1=0.6854838709677419
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0032838950864970684
step: 10, loss: 0.0015765713760629296
step: 20, loss: 0.0008286090451292694
step: 30, loss: 0.0008437371579930186
step: 40, loss: 0.02185591496527195
step: 50, loss: 0.0018368617165833712
step: 60, loss: 0.1118161603808403
step: 70, loss: 0.0007330940570682287
step: 80, loss: 0.0004993480397388339
step: 90, loss: 0.0003309838066343218
step: 100, loss: 0.001675663166679442
step: 110, loss: 0.058996669948101044
step: 120, loss: 0.010174836963415146
step: 130, loss: 0.04012371972203255
step: 140, loss: 0.028211846947669983
step: 150, loss: 0.05904652178287506
step: 160, loss: 0.00806430820375681
step: 170, loss: 0.009545784443616867
step: 180, loss: 0.03997355327010155
step: 190, loss: 0.0014047850854694843
step: 200, loss: 0.13899162411689758
step: 210, loss: 0.0011989723425358534
/usr/local/lib/python3.7/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=0.6769230769230768, f1=0.7099567099567099,
best_f1=0.6854838709677419
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00253865122795105
step: 10, loss: 0.04497307538986206
step: 20, loss: 0.003476504236459732
```

```
step: 30, loss: 0.03537723049521446
step: 40, loss: 0.0008051018812693655
step: 50, loss: 0.04756921902298927
step: 60, loss: 0.02347647026181221
step: 70, loss: 0.05008142814040184
step: 80, loss: 0.02365628257393837
step: 90, loss: 0.022984851151704788
step: 100, loss: 0.01035258173942566
step: 110, loss: 0.0001803929771995172
step: 120, loss: 0.000629729067441076
step: 130, loss: 0.08082003146409988
step: 140, loss: 0.006179216783493757
step: 150, loss: 0.015862850472331047
step: 160, loss: 0.0018803656566888094
step: 170, loss: 0.00047567387809976935
step: 180, loss: 0.04508793354034424
step: 190, loss: 0.050447843968868256
step: 200, loss: 0.01525567565113306
step: 210, loss: 0.005360471550375223
/usr/local/lib/python3.7/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=0.672340425531915, f1=0.7031578947368422,
best_f1=0.6854838709677419
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.03647060692310333
step: 10, loss: 0.03065279871225357
step: 20, loss: 0.03142113238573074
step: 30, loss: 0.00028267010929994285
step: 40, loss: 0.019810069352388382
step: 50, loss: 0.006032058969140053
step: 60, loss: 0.010684352368116379
step: 70, loss: 0.015891581773757935
step: 80, loss: 0.0534202940762043
step: 90, loss: 0.003176626283675432
step: 100, loss: 0.00989631749689579
step: 110, loss: 0.07130631804466248
step: 120, loss: 0.04566318169236183
step: 130, loss: 0.006578630767762661
step: 140, loss: 0.0008521853014826775
step: 150, loss: 0.0008520475239492953
step: 160, loss: 0.020100431516766548
step: 170, loss: 0.04763733223080635
step: 180, loss: 0.00291487667709589
step: 190, loss: 0.013627405278384686
step: 200, loss: 0.0017555204685777426
```

```
step: 210, loss: 0.0004756607813760638
/usr/local/lib/python3.7/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=0.6802218114602587, f1=0.7030075187969925,
best_f1=0.6854838709677419
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.053718965500593185
step: 10, loss: 0.000927229761146009
step: 20, loss: 0.0023915558122098446
step: 30, loss: 0.0036654530558735132
step: 40, loss: 0.0038826230447739363
step: 50, loss: 0.008736418560147285
step: 60, loss: 0.0004439296026248485
step: 70, loss: 0.011049496941268444
step: 80, loss: 0.0005042661796323955
step: 90, loss: 0.018574655055999756
step: 100, loss: 0.0015944541664794087
step: 110, loss: 0.0005040388787165284
step: 120, loss: 0.00016897585010156035
step: 130, loss: 0.00011009447189280763
step: 140, loss: 0.0019345609471201897
step: 150, loss: 0.023536229506134987
step: 160, loss: 0.0021291887387633324
step: 170, loss: 0.0009395680390298367
step: 180, loss: 0.0032101725228130817
step: 190, loss: 0.0020732807461172342
step: 200, loss: 0.0006171476561576128
step: 210, loss: 0.03447621688246727
/usr/local/lib/python3.7/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=0.6561797752808989, f1=0.7015250544662309,
best_f1=0.6854838709677419
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.03292252868413925
step: 10, loss: 0.00010061921057058498
step: 20, loss: 0.015750795602798462
step: 30, loss: 0.010494865477085114
step: 40, loss: 0.01691197231411934
step: 50, loss: 0.0010730003705248237
step: 60, loss: 0.0010072838049381971
step: 70, loss: 0.18881277740001678
step: 80, loss: 0.0376126728951931
```

```
step: 90, loss: 0.0017010383307933807
step: 100, loss: 0.0002014428609982133
step: 110, loss: 0.00021740474039688706
step: 120, loss: 0.0009590124245733023
step: 130, loss: 0.018875740468502045
step: 140, loss: 0.06337208300828934
step: 150, loss: 0.002301980974152684
step: 160, loss: 0.11245302110910416
step: 170, loss: 0.00020272481197025627
step: 180, loss: 0.024968262761831284
step: 190, loss: 0.0003228765563108027
step: 200, loss: 0.0012089713709428906
step: 210, loss: 0.006855428218841553
/usr/local/lib/python3.7/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=0.6680851063829787, f1=0.7127882599580714,
best_f1=0.6854838709677419
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0013189768651500344
step: 10, loss: 0.0610395222902298
step: 20, loss: 0.0004100239893887192
step: 30, loss: 0.0001559251541038975
step: 40, loss: 0.0001438520848751068
step: 50, loss: 0.01454825047403574
step: 60, loss: 0.00013899730402044952
step: 70, loss: 9.943875193130225e-05
step: 80, loss: 0.00026766050072338134
step: 90, loss: 0.0008950581541284919
step: 100, loss: 0.011935665272176266
step: 110, loss: 0.00036319729406386614
step: 120, loss: 0.025531688705086708
step: 130, loss: 0.00698267575353384
step: 140, loss: 0.02979668788611889
step: 150, loss: 0.0019070249982178211
step: 160, loss: 0.0009958998998627067
step: 170, loss: 0.0001733565586619079
step: 180, loss: 0.00036767605342902243
step: 190, loss: 0.00047333957627415657
step: 200, loss: 0.00015000003622844815
step: 210, loss: 0.00048697448801249266
/usr/local/lib/python3.7/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=0.6550218340611355, f1=0.7041036717062635,
best_f1=0.6854838709677419
```

```
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0005863960250280797
step: 10, loss: 9.534873242955655e-05
step: 20, loss: 0.0002934423682745546
step: 30, loss: 0.011280297301709652
step: 40, loss: 0.00018523797916714102
step: 50, loss: 0.0001115706327254884
step: 60, loss: 0.020379232242703438
step: 70, loss: 0.00029730767710134387
step: 80, loss: 7.251750503201038e-05
step: 90, loss: 0.013674531131982803
step: 100, loss: 0.00012369938485790044
step: 110, loss: 0.00016775137919466943
step: 120, loss: 0.0001315134868491441
step: 130, loss: 0.0001801490579964593
step: 140, loss: 8.958158286986873e-05
step: 150, loss: 7.582223770441487e-05
step: 160, loss: 0.0012507893843576312
step: 170, loss: 0.00011196670675417408
step: 180, loss: 7.926589751150459e-05
step: 190, loss: 0.0003284939157310873
step: 200, loss: 0.009300082921981812
step: 210, loss: 0.002201198134571314
/usr/local/lib/python3.7/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=0.6738197424892702, f1=0.6993603411513859,
best_f1=0.6854838709677419
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.")
2293it [00:04, 477.74it/s]
load_f1 = 0.6975806451612904
real_f1 = 0.6958250497017893
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 [00:17, 251.15it/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 bert-base-uncased \
        --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 bert-base-uncased \
```

```
    --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: 0.578006386756897
step: 10, loss: 0.3667844235897064
step: 20, loss: 0.26713109016418457
step: 30, loss: 0.41870149970054626
step: 40, loss: 0.4327484667301178
step: 50, loss: 0.3580265939235687
step: 60, loss: 0.26098865270614624

```
step: 70, loss: 0.2171774059534073
step: 80, loss: 0.2960411012172699
step: 90, loss: 0.3011975586414337
step: 100, loss: 0.32247868180274963
step: 110, loss: 0.29177945852279663
step: 120, loss: 0.11316505819559097
step: 130, loss: 0.1109917014837265
step: 140, loss: 0.04615601524710655
step: 150, loss: 0.14227037131786346
step: 160, loss: 0.05160421133041382
step: 170, loss: 0.18749964237213135
step: 180, loss: 0.039927467703819275
step: 190, loss: 0.1832856833934784
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.5925925925925924, f1=0.6257309941520468,
best_f1=0.6257309941520468
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.21678146719932556
step: 10, loss: 0.10321661084890366
step: 20, loss: 0.12600238621234894
step: 30, loss: 0.19641466438770294
step: 40, loss: 0.05432778224349022
step: 50, loss: 0.0462467335164547
step: 60, loss: 0.1916395127773285
step: 70, loss: 0.15145288407802582
step: 80, loss: 0.21150481700897217
step: 90, loss: 0.22393518686294556
step: 100, loss: 0.013603896833956242
step: 110, loss: 0.09710870683193207
step: 120, loss: 0.12689363956451416
step: 130, loss: 0.08803948760032654
step: 140, loss: 0.09384607523679733
step: 150, loss: 0.06114745885133743
step: 160, loss: 0.024653751403093338
step: 170, loss: 0.2206433266401291
step: 180, loss: 0.1956881433725357
step: 190, loss: 0.04542078822851181
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
```

```
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.774025974025974, f1=0.8031088082901554,
best_f1=0.8031088082901554
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.06391490250825882
step: 10, loss: 0.18199267983436584
step: 20, loss: 0.020483477041125298
step: 30, loss: 0.05779064819216728
step: 40, loss: 0.07509326189756393
step: 50, loss: 0.11959967762231827
step: 60, loss: 0.03777550160884857
step: 70, loss: 0.12483709305524826
step: 80, loss: 0.059039406478405
step: 90, loss: 0.1019350066781044
step: 100, loss: 0.01243123784661293
step: 110, loss: 0.008483045734465122
step: 120, loss: 0.01395280472934246
step: 130, loss: 0.009786529466509819
step: 140, loss: 0.008119343779981136
step: 150, loss: 0.05310855060815811
step: 160, loss: 0.13815374672412872
step: 170, loss: 0.17282043397426605
step: 180, loss: 0.04282096400856972
step: 190, loss: 0.12485847622156143
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.8054054054054054, f1=0.8089887640449438,
best_f1=0.8089887640449438
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.019090067595243454
step: 10, loss: 0.04227316752076149
step: 20, loss: 0.02888135239481926
step: 30, loss: 0.0354495532810688
step: 40, loss: 0.05578320473432541
step: 50, loss: 0.08758358657360077
step: 60, loss: 0.12379251420497894
```

```
step: 70, loss: 0.0359012670814991
step: 80, loss: 0.08871569484472275
step: 90, loss: 0.064661405980587
step: 100, loss: 0.04364991560578346
step: 110, loss: 0.015475600957870483
step: 120, loss: 0.10447827726602554
step: 130, loss: 0.12453751266002655
step: 140, loss: 0.03618118166923523
step: 150, loss: 0.005623432341963053
step: 160, loss: 0.003703933209180832
step: 170, loss: 0.037177279591560364
step: 180, loss: 0.0016194452764466405
step: 190, loss: 0.10971371829509735
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.8106666666666666, f1=0.8108108108108106,
best_f1=0.8108108108108106
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.09507670998573303
step: 10, loss: 0.01801467128098011
step: 20, loss: 0.008178919553756714
step: 30, loss: 0.0034816451370716095
step: 40, loss: 0.06094171851873398
step: 50, loss: 0.042058542370796204
step: 60, loss: 0.00842760968953371
step: 70, loss: 0.0016833962872624397
step: 80, loss: 0.06271612644195557
step: 90, loss: 0.008846067823469639
step: 100, loss: 0.0007911864668130875
step: 110, loss: 0.0033616251312196255
step: 120, loss: 0.03545084223151207
step: 130, loss: 0.1720031052827835
step: 140, loss: 0.05420827865600586
step: 150, loss: 0.18296018242835999
step: 160, loss: 0.014125276356935501
step: 170, loss: 0.004101645201444626
Gradient overflow.  Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 180, loss: 0.14161746203899384
step: 190, loss: 0.04186953976750374
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
```

```
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.8247422680412371, f1=0.8041775456919059,
best_f1=0.8041775456919059
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.018299264833331108
step: 10, loss: 0.0706033781170845
step: 20, loss: 0.0021520708687603474
step: 30, loss: 0.0008413285249844193
step: 40, loss: 0.16637542843818665
step: 50, loss: 0.006495966576039791
step: 60, loss: 0.019954491406679153
step: 70, loss: 0.006856208201497793
step: 80, loss: 0.026293300092220306
step: 90, loss: 0.06077868118882179
step: 100, loss: 0.0010700454004108906
step: 110, loss: 0.09222665429115295
step: 120, loss: 0.06189001724123955
step: 130, loss: 0.014239431358873844
step: 140, loss: 0.0708874985575676
step: 150, loss: 0.024753360077738762
step: 160, loss: 0.029956139624118805
step: 170, loss: 0.014534193091094494
step: 180, loss: 0.024664150550961494
step: 190, loss: 0.014038165099918842
/usr/local/lib/python3.7/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=0.8011695906432748, f1=0.8, best_f1=0.8041775456919059
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.004283578135073185
step: 10, loss: 0.007403087802231312
step: 20, loss: 0.01948234625160694
step: 30, loss: 0.012381923384964466
step: 40, loss: 0.013128619641065598
step: 50, loss: 0.017140164971351624
step: 60, loss: 0.011041024699807167
step: 70, loss: 0.004229220096021891
step: 80, loss: 0.010619920678436756
step: 90, loss: 0.0006190272397361696
step: 100, loss: 0.0005374985048547387
```

```
step: 110, loss: 0.0009966654470190406
step: 120, loss: 0.0006659763748757541
step: 130, loss: 0.010254524648189545
step: 140, loss: 0.0003672422026284039
step: 150, loss: 0.0004902761429548264
step: 160, loss: 0.11569881439208984
step: 170, loss: 0.17090961337089539
step: 180, loss: 0.0011537846876308322
step: 190, loss: 0.0005499045946635306
/usr/local/lib/python3.7/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=0.8150134048257373, f1=0.8188976377952756,
best_f1=0.8041775456919059
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.007423918694257736
step: 10, loss: 0.024822020903229713
step: 20, loss: 0.0002931813651230186
step: 30, loss: 0.14170129597187042
step: 40, loss: 0.0003386471653357148
step: 50, loss: 0.0003747375449165702
step: 60, loss: 0.0009310602326877415
step: 70, loss: 0.002681027865037322
step: 80, loss: 0.00027779347146861255
step: 90, loss: 0.0015030629001557827
step: 100, loss: 0.004172705579549074
step: 110, loss: 0.0004177007940597832
step: 120, loss: 0.003522885264828801
step: 130, loss: 0.003809180110692978
step: 140, loss: 0.0014592584921047091
step: 150, loss: 0.01022462546825409
step: 160, loss: 0.0020405289251357317
step: 170, loss: 0.011486314237117767
step: 180, loss: 0.0050278021954
step: 190, loss: 0.09625375270843506
/usr/local/lib/python3.7/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=0.772972972972973, f1=0.7913279132791328,
best_f1=0.8041775456919059
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0010582326212897897
step: 10, loss: 0.007601955905556679
step: 20, loss: 0.004789138212800026
```

```
step: 30, loss: 0.0006382492138072848
step: 40, loss: 0.00034983843215741217
step: 50, loss: 0.0002501318813301623
step: 60, loss: 0.0005159824504517019
step: 70, loss: 0.0015989609528332949
step: 80, loss: 0.00038343656342476606
step: 90, loss: 0.05021959915757179
step: 100, loss: 0.03837137669324875
step: 110, loss: 0.0008063691784627736
step: 120, loss: 0.015553615987300873
step: 130, loss: 0.0009173551807180047
step: 140, loss: 0.0006379891419783235
step: 150, loss: 0.08374587446451187
step: 160, loss: 0.0025705844163894653
step: 170, loss: 0.001814075279980898
step: 180, loss: 0.0016611373284831643
step: 190, loss: 0.0007810536189936101
/usr/local/lib/python3.7/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=0.8043478260869565, f1=0.8201058201058201,
best_f1=0.80417754456919059
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0011651248205453157
step: 10, loss: 0.00033723286469466984
step: 20, loss: 0.004775306675583124
step: 30, loss: 0.00010994664626196027
step: 40, loss: 0.0006387754110619426
step: 50, loss: 0.00027531140949577093
step: 60, loss: 0.001358936307951808
step: 70, loss: 0.00046497874427586794
step: 80, loss: 0.012359707616269588
step: 90, loss: 0.00015734668704681098
step: 100, loss: 0.0006964753265492618
step: 110, loss: 0.09315750747919083
step: 120, loss: 0.1792207509279251
step: 130, loss: 0.000571977871004492
step: 140, loss: 0.00037235618219710886
step: 150, loss: 0.0006689861183986068
step: 160, loss: 0.002861651126295328
step: 170, loss: 0.007625051774084568
step: 180, loss: 0.000725769205018878
step: 190, loss: 0.0005957429530099034
/usr/local/lib/python3.7/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=0.7851002865329514, f1=0.8067226890756303,
best_f1=0.8041775456919059
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00017935731739271432
step: 10, loss: 0.00045868451707065105
step: 20, loss: 0.0030050575733184814
step: 30, loss: 0.004264747723937035
step: 40, loss: 0.07043381035327911
step: 50, loss: 0.020919272676110268
step: 60, loss: 0.00020723290799651295
step: 70, loss: 0.0003574301954358816
step: 80, loss: 0.0017390273278579116
step: 90, loss: 0.00016848974337335676
step: 100, loss: 0.0004678461409639567
step: 110, loss: 0.00017545087030157447
step: 120, loss: 0.00034113359288312495
step: 130, loss: 0.00014407407434191555
step: 140, loss: 0.010381980799138546
step: 150, loss: 0.00015106530918274075
step: 160, loss: 0.00044080879888497293
step: 170, loss: 0.0016314693493768573
step: 180, loss: 0.0005908140446990728
step: 190, loss: 0.0008336669998243451
/usr/local/lib/python3.7/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=0.8125, f1=0.8315789473684211, best_f1=0.8041775456919059
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00011976848327321932
step: 10, loss: 0.0007660356350243092
step: 20, loss: 0.00023896862694527954
step: 30, loss: 0.000508622732013464
step: 40, loss: 0.00101803510915488
step: 50, loss: 0.0006523474585264921
step: 60, loss: 0.002603694563731551
step: 70, loss: 0.0004581788962241262
step: 80, loss: 0.00017054707859642804
step: 90, loss: 0.00010797684808494523
step: 100, loss: 0.0007712286897003651
step: 110, loss: 0.00017910920723807067
step: 120, loss: 0.0003611682914197445
step: 130, loss: 0.00014176232798490673
step: 140, loss: 0.0007547543500550091
step: 150, loss: 0.00023328509996645153
```

```
step: 160, loss: 0.00010990414011757821
step: 170, loss: 0.0004457206232473254
step: 180, loss: 9.608842810848728e-05
step: 190, loss: 0.00011515753431012854
/usr/local/lib/python3.7/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=0.7958656330749353, f1=0.8177083333333333,
best_f1=0.8041775456919059
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0015532633988186717
step: 10, loss: 0.00015539838932454586
step: 20, loss: 0.05318048596382141
step: 30, loss: 0.0018836443778127432
step: 40, loss: 0.0005798023194074631
step: 50, loss: 0.000521624053362757
step: 60, loss: 0.00011401788651710376
step: 70, loss: 0.00032466652919538328
step: 80, loss: 0.0002836047497112304
step: 90, loss: 0.0012984948698431253
step: 100, loss: 0.002807339420542121
step: 110, loss: 9.750460594659671e-05
step: 120, loss: 0.004061332903802395
step: 130, loss: 9.816832607612014e-05
step: 140, loss: 0.0005891090258955956
step: 150, loss: 8.288770914077759e-05
step: 160, loss: 0.0031296603847295046
step: 170, loss: 0.0003983031492680311
step: 180, loss: 0.0005595030961558223
step: 190, loss: 0.0007030838169157505
/usr/local/lib/python3.7/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=0.8099173553719008, f1=0.825136612021858,
best_f1=0.8041775456919059
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00019721062562894076
step: 10, loss: 0.00016673516074661165
step: 20, loss: 0.00010716429096646607
step: 30, loss: 0.00014006956189405173
step: 40, loss: 0.00020183758169878274
step: 50, loss: 0.00014875202032271773
step: 60, loss: 0.0002771540603134781
step: 70, loss: 0.000484591320855543
```

```
step: 80, loss: 0.0002025998692261055
step: 90, loss: 0.0003695801424328238
step: 100, loss: 0.002650762675330043
step: 110, loss: 0.00033906748285517097
step: 120, loss: 0.0003152707649860531
step: 130, loss: 0.00031989518902264535
step: 140, loss: 0.0002297960309078917
step: 150, loss: 0.00019581365631893277
step: 160, loss: 0.00023171697102952749
step: 170, loss: 0.00013528557610698044
step: 180, loss: 0.00022068011458031833
step: 190, loss: 0.00012157580204075202
/usr/local/lib/python3.7/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=0.8216216216216217, f1=0.8355795148247979,
best_f1=0.8041775456919059
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0014226253842934966
step: 10, loss: 8.992398215923458e-05
step: 20, loss: 0.00024372483312617987
step: 30, loss: 0.00011004682164639235
step: 40, loss: 0.00011094957153545693
step: 50, loss: 0.00010259001282975078
step: 60, loss: 9.165731898974627e-05
step: 70, loss: 0.0006891295779496431
step: 80, loss: 0.0013188401935622096
step: 90, loss: 7.921276846900582e-05
step: 100, loss: 0.00020252274407539517
step: 110, loss: 9.605394734535366e-05
step: 120, loss: 0.00013164174742996693
step: 130, loss: 0.0001369585224892944
step: 140, loss: 0.0015692946035414934
step: 150, loss: 0.0001856963208410889
step: 160, loss: 0.00033063048613257706
step: 170, loss: 0.00022744634770788252
step: 180, loss: 0.00023460618103854358
step: 190, loss: 0.00011388608982088044
/usr/local/lib/python3.7/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=0.8174386920980926, f1=0.8273972602739726,
best_f1=0.8041775456919059
mv: cannot stat '*_dev.pt': No such file or directory
Selected optimization level O2:  FP16 training with FP32 batchnorm and FP32
master weights.
```

```
Defaults for this optimization level are:
enabled                : True
opt_level              : O2
cast_model_type        : torch.float16
patch_torch_functions  : False
keep_batchnorm_fp32    : True
master_weights         : True
loss_scale             : dynamic
Processing user overrides (additional kwargs that are not None)…
After processing overrides, optimization options are:
enabled                : True
opt_level              : O2
cast_model_type        : torch.float16
patch_torch_functions  : False
keep_batchnorm_fp32    : True
master_weights         : True
loss_scale             : dynamic
Warning:  multi_tensor_applier fused unscale kernel is unavailable, possibly
because apex was installed without --cuda_ext --cpp_ext. Using Python fallback.
Original ImportError was: ModuleNotFoundError("No module named 'amp_C'")
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
2049it [00:08, 232.06it/s]
load_f1 = 0.8293963254593176
real_f1 = 0.8177083333333333
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 [00:17, 256.72it/s]
```

2.2   BASELINE TEXTUAL

2.2.1   Abt-Buy - Running the matcher

```python
# 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 bert-base-uncased \
  --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 bert-base-uncased \
  --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: 0.625508189201355
step: 10, loss: 0.34955671429634094
step: 20, loss: 0.31388863921165466
step: 30, loss: 0.38484176993370056
step: 40, loss: 0.2767445743083954
step: 50, loss: 0.2614802420139313
step: 60, loss: 0.24148686230182648
step: 70, loss: 0.36785873770713806
step: 80, loss: 0.3441850244998932
step: 90, loss: 0.23476211726665497
step: 100, loss: 0.22136765718460083
step: 110, loss: 0.23919777572155
step: 120, loss: 0.1473180204629898
step: 130, loss: 0.06386634707450867
step: 140, loss: 0.18440231680870056
step: 150, loss: 0.17254899442195892
step: 160, loss: 0.14628566801548004
step: 170, loss: 0.14882813394069672
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.7178217821782178, f1=0.6964705882352942,
best_f1=0.6964705882352942
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.09744121134281158
step: 10, loss: 0.25015920400619507
step: 20, loss: 0.038159992545843124
step: 30, loss: 0.2876400053501129
step: 40, loss: 0.12036231905221939
step: 50, loss: 0.07875987887382507
step: 60, loss: 0.1438644528388977
step: 70, loss: 0.03565936163067818
```

```
step: 80, loss: 0.06718917191028595
step: 90, loss: 0.21696950495243073
step: 100, loss: 0.15898795425891876
step: 110, loss: 0.07766512036323547
step: 120, loss: 0.04942278936505318
step: 130, loss: 0.10154952108860016
step: 140, loss: 0.3824010193347931
step: 150, loss: 0.15958309173583984
step: 160, loss: 0.2405923455953598
step: 170, loss: 0.04539227485656738
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.7921225382932167, f1=0.7581699346405228,
best_f1=0.7581699346405228
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.1932191252708435
step: 10, loss: 0.05453795939683914
step: 20, loss: 0.11608632653951645
step: 30, loss: 0.057744283229112625
step: 40, loss: 0.06838446110486984
step: 50, loss: 0.09533453732728958
step: 60, loss: 0.11468178778886795
step: 70, loss: 0.06762492656707764
step: 80, loss: 0.06971783936023712
step: 90, loss: 0.07068713009357452
step: 100, loss: 0.004139591474086046
step: 110, loss: 0.13476869463920593
step: 120, loss: 0.07682481408119202
step: 130, loss: 0.14468732476234436
step: 140, loss: 0.061672549694776535
step: 150, loss: 0.00830790400505066
step: 160, loss: 0.01976865902543068
step: 170, loss: 0.04742838069796562
/usr/local/lib/python3.7/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=0.7909319899244334, f1=0.8019559902200489,
best_f1=0.7581699346405228
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
```

```
step: 0, loss: 0.005553450435400009
step: 10, loss: 0.03825712576508522
step: 20, loss: 0.001549420179799199
step: 30, loss: 0.008876792155206203
step: 40, loss: 0.00074394402327016
step: 50, loss: 0.09157276153564453
step: 60, loss: 0.12751369178295135
step: 70, loss: 0.009679645299911499
step: 80, loss: 0.02629970759153366
step: 90, loss: 0.02925812639296055
step: 100, loss: 0.058948859572410583
step: 110, loss: 0.13416925072669983
step: 120, loss: 0.009600217454135418
step: 130, loss: 0.018833819776773453
step: 140, loss: 0.058461710810661316
step: 150, loss: 0.1849406361579895
step: 160, loss: 0.08577606827020645
step: 170, loss: 0.01749419793486595
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.8087167070217919, f1=0.7904761904761904,
best_f1=0.7904761904761904
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.011569793336093426
step: 10, loss: 0.009656297042965889
step: 20, loss: 0.016318071633577347
step: 30, loss: 0.01812056452035904
step: 40, loss: 0.05209031701087952
step: 50, loss: 0.01728471741080284
step: 60, loss: 0.00829263310879469
step: 70, loss: 0.14010784029960632
step: 80, loss: 0.0584423765540123
step: 90, loss: 0.08920453488826752
step: 100, loss: 0.00977635569870472
step: 110, loss: 0.0705575942993164
step: 120, loss: 0.01946926675736904
step: 130, loss: 0.02154293470084672
step: 140, loss: 0.0033348542638123035
step: 150, loss: 0.10276451706886292
step: 160, loss: 0.030677273869514465
step: 170, loss: 0.01146057527512312
```

```
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.8229166666666666, f1=0.8088235294117646,
best_f1=0.8088235294117646
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0023193038068711758
step: 10, loss: 0.07406347244977951
step: 20, loss: 0.0026761882472783327
step: 30, loss: 0.0011138890404254198
step: 40, loss: 0.006983571220189333
step: 50, loss: 0.08653867989778519
step: 60, loss: 0.019583627581596375
step: 70, loss: 0.06431665271520615
step: 80, loss: 0.028258871287107468
step: 90, loss: 0.05712313577532768
step: 100, loss: 0.002050303388386965
step: 110, loss: 0.0014015790075063705
step: 120, loss: 0.04760381206870079
step: 130, loss: 0.015055825933814049
step: 140, loss: 0.01420763973146677
step: 150, loss: 0.09584487974643707
step: 160, loss: 0.041719987988471985
step: 170, loss: 0.002318022074177861
/usr/local/lib/python3.7/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=0.8078817733990147, f1=0.7799043062200957,
best_f1=0.8088235294117646
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0030371444299817085
step: 10, loss: 0.011282984167337418
step: 20, loss: 0.0018917940324172378
step: 30, loss: 0.016932141035795212
step: 40, loss: 0.0011639980366453528
step: 50, loss: 0.01670253649353981
step: 60, loss: 0.0003074818232562393
Gradient overflow.  Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 70, loss: 0.12101195752620697
step: 80, loss: 0.05288374796509743
```

```
step: 90, loss: 0.00264176819473505
step: 100, loss: 0.012951969169080257
step: 110, loss: 0.02074175886809826
step: 120, loss: 0.0008957340614870191
step: 130, loss: 0.1394144892692566
step: 140, loss: 0.000937768432777375
step: 150, loss: 0.0026695362757891417
step: 160, loss: 0.0004732813686132431
step: 170, loss: 0.019067712128162384
/usr/local/lib/python3.7/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=0.8174807197943444, f1=0.7989821882951653,
best_f1=0.8088235294117646
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.003477830905467272
step: 10, loss: 0.002674720250070095
step: 20, loss: 0.006285428535193205
step: 30, loss: 0.029276251792907715
step: 40, loss: 0.0058335065841674805
step: 50, loss: 0.0013440718175843358
step: 60, loss: 0.0004563866532407701
step: 70, loss: 0.01742713153362274
step: 80, loss: 0.009366853162646294
step: 90, loss: 0.001375709194689989
step: 100, loss: 0.024880848824977875
step: 110, loss: 0.06878218799829483
step: 120, loss: 0.0036669981200248003
step: 130, loss: 0.002973382594063878
step: 140, loss: 0.00694191362708807
step: 150, loss: 0.01603853888809681
step: 160, loss: 0.0018884656019508839
step: 170, loss: 0.0017194576794281602
/usr/local/lib/python3.7/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=0.7877237851662404, f1=0.7989821882951653,
best_f1=0.8088235294117646
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00168672448489815
step: 10, loss: 0.0075935134664177895
step: 20, loss: 0.0007842002087272704
step: 30, loss: 0.029799766838550568
step: 40, loss: 0.0006573935970664024
```

```
step: 50, loss: 0.0010624835267663002
step: 60, loss: 0.004898776300251484
step: 70, loss: 0.0782499834895134
step: 80, loss: 0.010694785043597221
step: 90, loss: 0.02438315935432911
step: 100, loss: 0.021596217527985573
step: 110, loss: 0.0007731050718575716
step: 120, loss: 0.010645737871527672
step: 130, loss: 0.0613638311624527
step: 140, loss: 0.0004412233247421682
step: 150, loss: 0.010739644058048725
step: 160, loss: 0.023575112223625183
step: 170, loss: 0.028755178675055504
/usr/local/lib/python3.7/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=0.7921760391198044, f1=0.7761904761904763,
best_f1=0.8088235294117646
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.05136848986148834
step: 10, loss: 0.00034752002102322876
step: 20, loss: 0.04098503664135933
step: 30, loss: 0.00913605373352766
step: 40, loss: 0.07818183302879333
step: 50, loss: 0.04481148719787598
step: 60, loss: 0.00108641083335226774
step: 70, loss: 0.002016047714278102
step: 80, loss: 0.004654434975236654
step: 90, loss: 0.007369736209511757
step: 100, loss: 0.00041014852467924356
step: 110, loss: 0.0021613040007650852
step: 120, loss: 0.0005581460427492857
step: 130, loss: 0.001263494836166501
step: 140, loss: 0.031724654138088226
step: 150, loss: 0.0013327470514923334
step: 160, loss: 0.0011712285922840238
step: 170, loss: 0.00019434298155582931
/usr/local/lib/python3.7/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=0.7946666666666666, f1=0.8072916666666666,
best_f1=0.8088235294117646
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.042212605476379395
```

```
step: 10, loss: 0.0008827677229419351
step: 20, loss: 0.00034025852801278234
step: 30, loss: 0.0004277122498024255
step: 40, loss: 0.0035820272751152515
step: 50, loss: 0.0020866170525550842
step: 60, loss: 0.017613209784030914
step: 70, loss: 0.0003097702865488827
step: 80, loss: 0.0003527569060679525
step: 90, loss: 0.0004112979513593018
step: 100, loss: 0.0016677769599482417
step: 110, loss: 0.08326825499534607
step: 120, loss: 0.0003020609437953681
step: 130, loss: 0.00021616708545479923
step: 140, loss: 0.0006234032334759831
step: 150, loss: 0.00545419380068779
step: 160, loss: 0.0011147307232022285
step: 170, loss: 0.00084348936798051
/usr/local/lib/python3.7/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=0.7967914438502675, f1=0.8, best_f1=0.8088235294117646
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.02157534658908844
step: 10, loss: 0.0004364682245068252
step: 20, loss: 0.0009225541725754738
step: 30, loss: 0.00022742759028915316
step: 40, loss: 0.0018768339650705457
step: 50, loss: 0.000201660324819386
step: 60, loss: 0.0007964922697283328
step: 70, loss: 0.035934142768383026
step: 80, loss: 0.00015903006671523437
step: 90, loss: 0.00042463047429919243
step: 100, loss: 0.0007558751385658979
step: 110, loss: 0.000620368286035955
step: 120, loss: 0.0053620897233486176
step: 130, loss: 0.009580422192811966
step: 140, loss: 0.01840750128030777
step: 150, loss: 0.001587473787367344
step: 160, loss: 0.0006227113190107048
step: 170, loss: 0.0004017505270894617
/usr/local/lib/python3.7/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=0.8010335917312662, f1=0.8020050125313284,
best_f1=0.8088235294117646
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
```

An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.03939204290509224
step: 10, loss: 0.00022182203247211874
step: 20, loss: 0.0010882418137043715
step: 30, loss: 0.0005988414050079882
step: 40, loss: 0.014990770258009434
step: 50, loss: 0.0033236548770219088
step: 60, loss: 0.012731666676700115
step: 70, loss: 0.001829917193390429
step: 80, loss: 0.0009580819751136005
step: 90, loss: 0.0010997417848557234
step: 100, loss: 0.00032260167063213885
step: 110, loss: 0.0007753566605970263
step: 120, loss: 0.020781615749001503
step: 130, loss: 0.0003042806638404727
step: 140, loss: 0.0009651294676586986
step: 150, loss: 0.0024414549116045237
step: 160, loss: 0.008525168523192406
step: 170, loss: 0.002108542248606682
/usr/local/lib/python3.7/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=0.7980535279805353, f1=0.7807228915662651,
best_f1=0.8088235294117646
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0004081880615558475
step: 10, loss: 0.00022592463938053697
step: 20, loss: 0.0005917787202633917
step: 30, loss: 0.00040671956958824921
step: 40, loss: 0.00018210783309768885
step: 50, loss: 0.00019688754400704056
step: 60, loss: 0.005386231932789087
step: 70, loss: 0.00037097735912539065
step: 80, loss: 0.2213425189256668
step: 90, loss: 0.0004177512601017952
step: 100, loss: 0.0002752718282863498
step: 110, loss: 0.0002888008311856538
step: 120, loss: 0.04332461208105087
step: 130, loss: 0.0008414093754254282
step: 140, loss: 0.00026725741918198764
step: 150, loss: 0.004742114804685116
step: 160, loss: 0.00030819629319012165
step: 170, loss: 0.00022990598517935723
/usr/local/lib/python3.7/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=0.7989276139410187, f1=0.796875, best_f1=0.8088235294117646
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.1149492934346199
step: 10, loss: 0.0006335630896501243
step: 20, loss: 0.05382486805319786
step: 30, loss: 0.0019371388480067253
step: 40, loss: 0.0004896469181403518
step: 50, loss: 0.00019698671530932188
step: 60, loss: 0.015597470104694366
step: 70, loss: 0.00024354133347515017
step: 80, loss: 0.014773000963032246
step: 90, loss: 0.0019365038024261594
step: 100, loss: 0.0004276104737073183
step: 110, loss: 0.0001682407601037994
step: 120, loss: 0.00035932360333390534
step: 130, loss: 0.0033603529445827007
step: 140, loss: 0.0002440542884869501
step: 150, loss: 0.00026147093740291893
step: 160, loss: 0.00018204544903710485
step: 170, loss: 0.00031304280855692923
/usr/local/lib/python3.7/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=0.8010610079575595, f1=0.7958656330749355,
best_f1=0.8088235294117646
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 [00:07, 246.61it/s]
load_f1 = 0.8179419525065964
real_f1 = 0.8157894736842105
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, 216.83it/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 bert-base-uncased \
       --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 bert-base-uncased \
```

```
  --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: 0.6271450519561768
step: 10, loss: 0.6153005957603455
step: 20, loss: 0.43128204345703125
step: 30, loss: 0.16884230077266693
step: 40, loss: 0.1917475163936615
step: 50, loss: 0.07869302481412888
step: 60, loss: 0.21761369705200195

```
step: 70, loss: 0.04846937954425812
step: 80, loss: 0.0776652842760086
step: 90, loss: 0.046428173780441284
step: 100, loss: 0.02929617650806904
step: 110, loss: 0.2844969630241394
step: 120, loss: 0.022593745961785316
step: 130, loss: 0.02304801158607006
step: 140, loss: 0.004168352112174034
step: 150, loss: 0.018481934443116188
step: 160, loss: 0.06406774371862411
step: 170, loss: 0.05344878509640694
step: 180, loss: 0.07674388587474823
step: 190, loss: 0.04305620864033699
step: 200, loss: 0.13743504881858826
step: 210, loss: 0.0036993296816945076
step: 220, loss: 0.03428885340690613
step: 230, loss: 0.01585022732615471
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9706546275395034, f1=0.9704545454545453,
best_f1=0.9704545454545453
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0088838255032897
step: 10, loss: 0.027011457830667496
step: 20, loss: 0.14829230308532715
step: 30, loss: 0.0148554015904665
step: 40, loss: 0.0316527858376503
step: 50, loss: 0.010610530152916908
step: 60, loss: 0.0034286254085600376
step: 70, loss: 0.07989975810050964
step: 80, loss: 0.04195771738886833
step: 90, loss: 0.028412576764822006
step: 100, loss: 0.08127564191818237
step: 110, loss: 0.09665244817733765
step: 120, loss: 0.14297246932983398
step: 130, loss: 0.012990279123187065
step: 140, loss: 0.002620941959321499
step: 150, loss: 0.022611021995544434
step: 160, loss: 0.05070710927248001
step: 170, loss: 0.02042328380048275
step: 180, loss: 0.0051484727300703526
```

```
step: 190, loss: 0.0017049964517354965
step: 200, loss: 0.0018102937610819936
step: 210, loss: 0.0011036432115361094
step: 220, loss: 0.05046715587377548
step: 230, loss: 0.02833561785519123
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9719416386083053, f1=0.9720670391061451,
best_f1=0.9720670391061451
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.009377866052091122
step: 10, loss: 0.000890303635969758
step: 20, loss: 0.00305143673904240413
step: 30, loss: 0.0022690966725349426
step: 40, loss: 0.028189022094011307
step: 50, loss: 0.011848445050418377
step: 60, loss: 0.010071715340018272
step: 70, loss: 0.004480781964957714
step: 80, loss: 0.0012737656943500042
step: 90, loss: 0.042407553642988205
step: 100, loss: 0.0013612427283078432
step: 110, loss: 0.0006357281235978007
step: 120, loss: 0.010805091820657253
step: 130, loss: 0.0005154961836524308
step: 140, loss: 0.00241091032512486
step: 150, loss: 0.02431389130651951
step: 160, loss: 0.012370679527521133
step: 170, loss: 0.012343217618763447
step: 180, loss: 0.005123824346810579
step: 190, loss: 0.01966150291264057
step: 200, loss: 0.0028034034185111523
step: 210, loss: 0.0049489825032651424
step: 220, loss: 0.0016615958884358406
step: 230, loss: 0.04266339913010597
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9787709497206705, f1=0.9731543624161074,
best_f1=0.9731543624161074
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0008424604311585426
step: 10, loss: 0.0021672130096703768
step: 20, loss: 0.00108085956890136
step: 30, loss: 0.0012089816154912114
step: 40, loss: 0.08310473710298538
step: 50, loss: 0.0019555757848918438
step: 60, loss: 0.0012368496973067522
step: 70, loss: 0.001350420294329524
step: 80, loss: 0.006833021529018879
step: 90, loss: 0.04013385623693466
step: 100, loss: 0.031000718474388123
step: 110, loss: 0.001808822969906032
step: 120, loss: 0.01570291817188263
step: 130, loss: 0.0019540726207196712
step: 140, loss: 0.0006584858638234437
step: 150, loss: 0.2094680368900299
step: 160, loss: 0.0031581574585288763
step: 170, loss: 0.0018967926735058427
step: 180, loss: 0.0006552713457494974
step: 190, loss: 0.0014854577602818608
step: 200, loss: 0.002490042243152857
step: 210, loss: 0.2015429586172104
step: 220, loss: 0.0004762661410495639
step: 230, loss: 0.0018056719563901424
/usr/local/lib/python3.7/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=0.9787234042553192, f1=0.9807909604519773,
best_f1=0.9731543624161074
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0008257052395492792
step: 10, loss: 0.0006347854505293071
step: 20, loss: 0.000865981332026422
step: 30, loss: 0.0004132275062147528
step: 40, loss: 0.001310772611759603
step: 50, loss: 0.0006805428420193493
step: 60, loss: 0.0028999524656683207
step: 70, loss: 0.0003550726978573948
step: 80, loss: 0.0004795589193236083
step: 90, loss: 0.007233936805278063
step: 100, loss: 0.0003679631045088172
```

```
step: 110, loss: 0.0002012737386394292
Gradient overflow.  Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 120, loss: 0.049383193254470825
step: 130, loss: 0.0017914199270308018
step: 140, loss: 0.0065753706730902195
step: 150, loss: 0.0007626080187037587
step: 160, loss: 0.00075200927676633
step: 170, loss: 0.005378887988626957
step: 180, loss: 0.0009184793452732265
step: 190, loss: 0.06941521912813187
step: 200, loss: 0.006601470522582531
step: 210, loss: 0.0005409940495155752
step: 220, loss: 0.0012296239146962762
step: 230, loss: 0.005494930315762758
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9809203142536477, 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: 0.004012235440313816
step: 10, loss: 0.0003324321296531707
step: 20, loss: 0.002303722547367215
step: 30, loss: 0.00019296881509944797
step: 40, loss: 0.00017057951481547207
step: 50, loss: 0.0002506877644918859
step: 60, loss: 0.00012973815319128335
step: 70, loss: 0.0005652345716953278
step: 80, loss: 0.0005483920685946941
step: 90, loss: 0.0004498197231441736
step: 100, loss: 0.0052136704325675964
step: 110, loss: 0.000591420044656843
step: 120, loss: 0.011302351951599121
step: 130, loss: 0.0008310950943268836
step: 140, loss: 0.04339362308382988
step: 150, loss: 0.17386634647846222
step: 160, loss: 0.002376652555540204
step: 170, loss: 0.0007145072449930012
step: 180, loss: 0.022567199543118477
step: 190, loss: 0.0015926639316603541
step: 200, loss: 0.00019354767573531717
step: 210, loss: 0.0007546328706666827
```

```
step: 220, loss: 0.0005504671717062593
step: 230, loss: 0.07597699016332626
/usr/local/lib/python3.7/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=0.9732739420935412, f1=0.9688195991091313,
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: 0.0010365251218900084
step: 10, loss: 0.00043732079211622477
step: 20, loss: 0.0033697725739330053
step: 30, loss: 0.005209349095821381
step: 40, loss: 0.0004337483842391521
step: 50, loss: 0.0005398810608312488
step: 60, loss: 0.001485001645050943
step: 70, loss: 0.0345061793923378
step: 80, loss: 0.0018170162802562118
step: 90, loss: 8.739143231650814e-05
step: 100, loss: 0.00016150882584042847
step: 110, loss: 0.0002713151043280959
step: 120, loss: 0.00010861544433282688
step: 130, loss: 9.250095899915323e-05
step: 140, loss: 0.00029670391813851893
step: 150, loss: 0.007512656506150961
step: 160, loss: 0.08532062917947769
step: 170, loss: 0.0006568029057234526
step: 180, loss: 0.006081215105950832
step: 190, loss: 0.0007209383766166866
step: 200, loss: 0.07312536239624023
step: 210, loss: 0.0003052636166103184
step: 220, loss: 0.18199506402015686
step: 230, loss: 0.004074115306138992
/usr/local/lib/python3.7/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=0.9777777777777777, f1=0.9732142857142857,
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: 0.00042994300019927323
step: 10, loss: 0.0016563043463975191
step: 20, loss: 0.0013335446128621697
step: 30, loss: 0.0003195655590388924
step: 40, loss: 0.0005175152909941971
step: 50, loss: 0.0004717349074780941
```

```
step: 60, loss: 0.00015820268890820444
step: 70, loss: 0.0009589222609065473
step: 80, loss: 0.0006130185211077332
step: 90, loss: 7.982734678080305e-05
step: 100, loss: 0.00045762636000290513
step: 110, loss: 0.1347406804561615
step: 120, loss: 0.0036429036408662796
step: 130, loss: 0.007606690749526024
step: 140, loss: 0.00022977097250986844
step: 150, loss: 0.00028265881701372564
step: 160, loss: 0.00040988618275150657
step: 170, loss: 0.002501159207895398
step: 180, loss: 0.003135404549539089
step: 190, loss: 0.002000238047912717
step: 200, loss: 0.0012482242891564965
step: 210, loss: 0.002806038362905383
step: 220, loss: 0.00037010753294453025
step: 230, loss: 0.00032251651282422245
/usr/local/lib/python3.7/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=0.9797297297297298, f1=0.9808342728297633,
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: 0.00013852721895091236
step: 10, loss: 0.0005453271442092955
step: 20, loss: 0.0028776288963854313
step: 30, loss: 0.0003472879179753363
step: 40, loss: 0.015499849803745747
step: 50, loss: 9.592338028596714e-05
step: 60, loss: 0.00020079025125596672
step: 70, loss: 0.0008052625344134867
step: 80, loss: 9.902416786644608e-05
step: 90, loss: 0.0004446821694727987
step: 100, loss: 0.00019602762768045068
step: 110, loss: 8.360071660717949e-05
step: 120, loss: 9.181489440379664e-05
step: 130, loss: 0.00010527660197112709
step: 140, loss: 0.0008703062776476145
step: 150, loss: 0.00027315036277286671
step: 160, loss: 0.00029561787960119545
step: 170, loss: 0.0007926073740236461
step: 180, loss: 0.0012960205785930157
step: 190, loss: 0.000516809057444334
step: 200, loss: 0.0005083896685391665
step: 210, loss: 0.0009832603391259909
```

```
step: 220, loss: 0.00028504765941761434
step: 230, loss: 0.00028536602621898055
/usr/local/lib/python3.7/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=0.9763779527559054, f1=0.9807909604519773,
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: 0.00036132882814854383
step: 10, loss: 0.0001810544345062226
step: 20, loss: 0.00031753876828588545
step: 30, loss: 0.0007364408229477704
step: 40, loss: 0.00015657990297768265
step: 50, loss: 0.00019116712792310864
step: 60, loss: 0.0003894547699019313
step: 70, loss: 0.0004039512714371085
step: 80, loss: 0.0002745245583355427
step: 90, loss: 0.0005845357081852853
step: 100, loss: 0.0002603383327368647
step: 110, loss: 0.0003304643032606691
step: 120, loss: 0.00027320446679368615
step: 130, loss: 0.0008421125239692628
step: 140, loss: 0.060741450637578964
step: 150, loss: 0.0166164543479681
step: 160, loss: 0.0006050083902664483
step: 170, loss: 0.00022141887166071683
step: 180, loss: 0.00029121656552888453
step: 190, loss: 0.003759015817195177
step: 200, loss: 0.00022870775137562305
step: 210, loss: 5.973963561700657e-05
step: 220, loss: 0.00012047526251990348
step: 230, loss: 0.0003662552044261247
/usr/local/lib/python3.7/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=0.9787709497206705, f1=0.9809203142536477,
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: 0.00012300476373638958
step: 10, loss: 0.0002754885936155915
step: 20, loss: 0.000827212177682668
step: 30, loss: 0.001417450257577002
step: 40, loss: 0.00039926738827489316
step: 50, loss: 0.0002466684381943196
```

```
step: 60, loss: 0.000761483795940876
step: 70, loss: 0.0002011418982874602
step: 80, loss: 0.00010115825716638938
step: 90, loss: 0.00012350598990451545
step: 100, loss: 0.00011933620407944545
step: 110, loss: 0.006277669221162796
step: 120, loss: 0.00015861308202147484
step: 130, loss: 8.263959171017632e-05
step: 140, loss: 8.723133214516565e-05
step: 150, loss: 0.0002639539889059961
step: 160, loss: 0.00015941268065944314
step: 170, loss: 0.0004873527796007693
step: 180, loss: 0.00017574333469383419
step: 190, loss: 0.0001234941155416891
step: 200, loss: 9.623437654227018e-05
step: 210, loss: 5.299536496750079e-05
step: 220, loss: 8.820331277092919e-05
step: 230, loss: 0.00017641477461438626
/usr/local/lib/python3.7/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=0.9798657718120806, f1=0.9820224719101124,
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: 9.552408300805837e-05
step: 10, loss: 6.946772191440687e-05
step: 20, loss: 0.00014468468725681305
step: 30, loss: 0.00011384129174984992
step: 40, loss: 0.00015367643209174275
step: 50, loss: 0.0013036163290962577
step: 60, loss: 0.00015826641174498945
step: 70, loss: 5.079769834992476e-05
step: 80, loss: 0.00011777428153436631
step: 90, loss: 7.854637078708038e-05
step: 100, loss: 6.97678406140767e-05
step: 110, loss: 8.465073187835515e-05
step: 120, loss: 8.642466855235398e-05
step: 130, loss: 7.061679207254201e-05
step: 140, loss: 0.0024976967833936214
step: 150, loss: 0.00013549112190958112
step: 160, loss: 0.00016607521683909
step: 170, loss: 0.00010258772817905992
step: 180, loss: 0.004322263412177563
step: 190, loss: 0.00015578877355437726
step: 200, loss: 8.034692291403189e-05
step: 210, loss: 0.00015409791376441717
```

```
step: 220, loss: 0.000274631951469928
step: 230, loss: 0.0014812322333455086
/usr/local/lib/python3.7/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=0.978865406006674, f1=0.9776286353467561,
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: 0.00048712201532907784
step: 10, loss: 0.0003158916952088475
step: 20, loss: 0.004023152869194746
step: 30, loss: 0.00018499261932447553
step: 40, loss: 0.0002518461551517248
step: 50, loss: 0.0002333280717721209
step: 60, loss: 0.00013949019194114953
step: 70, loss: 6.102844781707972e-05
step: 80, loss: 0.0001138713996624574
step: 90, loss: 0.00011826224363176152
step: 100, loss: 6.0127982578705996e-05
step: 110, loss: 0.00774032436311245
step: 120, loss: 0.0004144276026636362
step: 130, loss: 0.00032061644014902413
step: 140, loss: 0.00023962634440977126
step: 150, loss: 0.00020384743402246386
step: 160, loss: 0.00022291215496370569
step: 170, loss: 0.00019619902013801038
step: 180, loss: 0.00015636863827239722
step: 190, loss: 0.00011609386274358258
step: 200, loss: 0.000464424432720989
step: 210, loss: 6.962892803130671e-05
step: 220, loss: 0.00017312593990936875
step: 230, loss: 6.083375046728179e-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=0.9799554565701558, f1=0.9744160177975528,
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: 0.00015723789692856371
step: 10, loss: 9.421409049537033e-05
step: 20, loss: 0.000186983568710275
step: 30, loss: 0.00013942105579189956
step: 40, loss: 0.020689813420176506
step: 50, loss: 0.00018326667486689985
```

```
step: 60, loss: 7.473822188330814e-05
step: 70, loss: 0.0008653935510665178
step: 80, loss: 0.00013094609312247485
step: 90, loss: 0.00010649021714925766
step: 100, loss: 0.00016111622971948236
step: 110, loss: 0.00011684656055876985
step: 120, loss: 4.017371975351125e-05
step: 130, loss: 9.705795673653483e-05
step: 140, loss: 0.0001027550533763133
step: 150, loss: 7.207548333099112e-05
step: 160, loss: 0.0009882566519081593
step: 170, loss: 4.185995567240752e-05
step: 180, loss: 8.647425420349464e-05
step: 190, loss: 0.00017233644030056894
step: 200, loss: 6.323288107523695e-05
step: 210, loss: 8.326645183842629e-05
step: 220, loss: 0.0005851500318385661
step: 230, loss: 0.0007989166188053787
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9810901001112348, f1=0.9722530521642618,
best_f1=0.9722530521642618
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 5.5398239055648446e-05
step: 10, loss: 4.4971468014409766e-05
step: 20, loss: 0.00019314058590068036
step: 30, loss: 5.463447087095119e-05
step: 40, loss: 7.337704300880432e-05
step: 50, loss: 0.022162288427352905
step: 60, loss: 8.372803131351247e-05
step: 70, loss: 0.00020014468464069068
step: 80, loss: 0.0002531872596591711
step: 90, loss: 0.00011610258661676198
step: 100, loss: 9.698172652861103e-05
step: 110, loss: 8.22337024146691e-05
step: 120, loss: 0.0001365473581245169
step: 130, loss: 0.0027968697249889374
step: 140, loss: 7.751585508231074e-05
step: 150, loss: 0.00031504558864980936
step: 160, loss: 7.459281187038869e-05
step: 170, loss: 4.671805800171569e-05
```

```
step: 180, loss: 0.00015702626842539757
step: 190, loss: 0.00018995914433617145
step: 200, loss: 0.000308766815578565
step: 210, loss: 6.686674169031903e-05
step: 220, loss: 7.630258915014565e-05
step: 230, loss: 7.715090032434091e-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=0.9810901001112348, f1=0.9733333333333333,
best_f1=0.9722530521642618
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.")
2473it [00:13, 178.67it/s]
load_f1 = 0.9810901001112348
real_f1 = 0.98
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, 218.50it/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 bert-base-uncased \
  --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 bert-base-uncased \
  --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: 0.6352456212043762
step: 10, loss: 0.532656192779541
step: 20, loss: 0.4920925796031952
step: 30, loss: 0.06079695373773575
step: 40, loss: 0.12186449021100998
step: 50, loss: 0.17821063101291656
step: 60, loss: 0.2380864918231964
step: 70, loss: 0.1546720266342163
step: 80, loss: 0.13755249977111816
step: 90, loss: 0.1255534291267395
step: 100, loss: 0.01607920229434967
step: 110, loss: 0.10515327751636505
step: 120, loss: 0.14257046580314636
step: 130, loss: 0.054857488721609116
step: 140, loss: 0.08720608055591583
step: 150, loss: 0.03432334586977959
step: 160, loss: 0.0390322208404541
step: 170, loss: 0.19540369510650635
step: 180, loss: 0.04003281891345978
step: 190, loss: 0.00964298378676176
Gradient overflow.  Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 200, loss: 0.24825255572795868
step: 210, loss: 0.09779851883649826
step: 220, loss: 0.22028222680091858
step: 230, loss: 0.15740138292312622
step: 240, loss: 0.08587373793125153
step: 250, loss: 0.011150641366839409
step: 260, loss: 0.030726589262485504
step: 270, loss: 0.0185868777334690l
```

```
step: 280, loss: 0.037817247211933136
step: 290, loss: 0.07088350504636765
step: 300, loss: 0.020986733958125114
step: 310, loss: 0.30826064944267273
step: 320, loss: 0.13266515731811523
step: 330, loss: 0.013411153107881546
step: 340, loss: 0.04856182262301445
step: 350, loss: 0.032818373292684555
step: 360, loss: 0.032238882035017014
step: 370, loss: 0.10225626081228256
step: 380, loss: 0.023422878235578537
step: 390, loss: 0.09636527299880981
step: 400, loss: 0.23779603838920593
step: 410, loss: 0.058281686156988144
step: 420, loss: 0.016752738505601883
step: 430, loss: 0.14313413202762604
step: 440, loss: 0.045502543449401855
step: 450, loss: 0.025120466947555542
step: 460, loss: 0.005255893338471651
step: 470, loss: 0.1647808849811554
step: 480, loss: 0.053660567849874496
step: 490, loss: 0.08075788617134094
step: 500, loss: 0.07023809105157852
step: 510, loss: 0.059497442096471786
step: 520, loss: 0.059146683663129807
step: 530, loss: 0.008487471379339695
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.940959409594096, f1=0.9316081330868762,
best_f1=0.9316081330868762
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.05307995527982712
step: 10, loss: 0.027038387954235077
step: 20, loss: 0.03937145322561264
step: 30, loss: 0.034684669226408005
step: 40, loss: 0.10374635457992554
step: 50, loss: 0.08791260421276093
step: 60, loss: 0.00403040973469615
step: 70, loss: 0.01799757033586502
step: 80, loss: 0.10266252607107162
step: 90, loss: 0.011822097934782505
```

```
step: 100, loss: 0.015914836898446083
step: 110, loss: 0.11195503920316696
step: 120, loss: 0.07486983388662338
step: 130, loss: 0.08539516478776932
step: 140, loss: 0.04493777081370354
step: 150, loss: 0.03416001424193382
step: 160, loss: 0.0036681590136140585
step: 170, loss: 0.01093273889273405
step: 180, loss: 0.011515399441123009
step: 190, loss: 0.031393103301525116
step: 200, loss: 0.0042618680745363235
step: 210, loss: 0.04001837596297264
step: 220, loss: 0.055941034108400345
step: 230, loss: 0.007188259623944759
step: 240, loss: 0.023960178717970848
step: 250, loss: 0.06588230282068253
step: 260, loss: 0.006638604681938887
step: 270, loss: 0.24075284600257874
step: 280, loss: 0.015767240896821022
step: 290, loss: 0.012794791720807552
step: 300, loss: 0.17454369366168976
step: 310, loss: 0.013573684729635715
step: 320, loss: 0.10062653571367264
step: 330, loss: 0.03637361153960228
step: 340, loss: 0.057780925184488297
step: 350, loss: 0.002392341149970889
step: 360, loss: 0.0874578133225441
step: 370, loss: 0.15053382515907288
step: 380, loss: 0.11606507748365402
step: 390, loss: 0.04081861302256584
step: 400, loss: 0.005448398645967245
step: 410, loss: 0.021442698314785957
step: 420, loss: 0.01940445974469185
step: 430, loss: 0.02780776470899582
step: 440, loss: 0.22075439989566803
step: 450, loss: 0.031591128557920456
step: 460, loss: 0.042020369321107864
step: 470, loss: 0.0857573077082634
step: 480, loss: 0.1926824152469635
step: 490, loss: 0.021002596244215965
step: 500, loss: 0.16381320357322693
step: 510, loss: 0.014498014003038406
step: 520, loss: 0.042679768055677414
step: 530, loss: 0.07909803092479706
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
```

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=0.9438515081206497, f1=0.9354988399071927,
best_f1=0.9354988399071927
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.08881131559610367
step: 10, loss: 0.05082077160477638
step: 20, loss: 0.14803123474121094
step: 30, loss: 0.02156892605125904
step: 40, loss: 0.053945232182741165
step: 50, loss: 0.03322656452655792
step: 60, loss: 0.07750485092401505
step: 70, loss: 0.02244856022298336
step: 80, loss: 0.001689729979261756
step: 90, loss: 0.0052084531635046005
step: 100, loss: 0.09629654884338379
step: 110, loss: 0.002583834808319807
step: 120, loss: 0.013688302598893642
step: 130, loss: 0.02073977328836918
step: 140, loss: 0.025178194046020508
step: 150, loss: 0.010461008176207542
step: 160, loss: 0.006455943454056978
step: 170, loss: 0.05908512324094772
step: 180, loss: 0.0027213275898247957
step: 190, loss: 0.0029012414161115885
step: 200, loss: 0.06631483137607574
step: 210, loss: 0.08709229528903961
step: 220, loss: 0.03519751876592636
step: 230, loss: 0.08904144912958145
step: 240, loss: 0.007115130312740803
step: 250, loss: 0.043806906789541245
step: 260, loss: 0.035538025200366974
step: 270, loss: 0.0021400433033704758
step: 280, loss: 0.1290028691291809
step: 290, loss: 0.01291780173778534
step: 300, loss: 0.04254065454006195
step: 310, loss: 0.00626433826982975
step: 320, loss: 0.027869101613759995
step: 330, loss: 0.008855672553181648
step: 340, loss: 0.015904366970062256
step: 350, loss: 0.0033845435827970505
step: 360, loss: 0.04372011125087738
step: 370, loss: 0.1230047345161438
step: 380, loss: 0.02883102558553219
step: 390, loss: 0.05789574235677719

```
step: 400, loss: 0.0031391477677971125
step: 410, loss: 0.004923119675368071
step: 420, loss: 0.05904005467891693
step: 430, loss: 0.004538670182228088
step: 440, loss: 0.03969212621450424
step: 450, loss: 0.050042275339365005
step: 460, loss: 0.05100805684924126
step: 470, loss: 0.1640632003545761
step: 480, loss: 0.005017090123146772
step: 490, loss: 0.03212356194853783
step: 500, loss: 0.03612346202135086
step: 510, loss: 0.006496811751276255
step: 520, loss: 0.12753599882125854
step: 530, loss: 0.023568004369735718
/usr/local/lib/python3.7/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=0.9413936317489617, f1=0.933579335793358,
best_f1=0.9354988399071927
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0022208592854440212
step: 10, loss: 0.01092605758458376
step: 20, loss: 0.0009163780487142503
step: 30, loss: 0.0005403236136771739
step: 40, loss: 0.006002435460686684
step: 50, loss: 0.006505188532173634
step: 60, loss: 0.002018225844949484
step: 70, loss: 0.0027375505305826664
step: 80, loss: 0.013471471145749092
step: 90, loss: 0.019490810111165047
step: 100, loss: 0.0006568660028278828
step: 110, loss: 0.05766066163778305
step: 120, loss: 0.0005804208340123296
step: 130, loss: 0.00023348417016677558
step: 140, loss: 0.001046252204105258
step: 150, loss: 0.021371839568018913
step: 160, loss: 0.06006613373756409
step: 170, loss: 0.03196197375655174
step: 180, loss: 0.001566588063724339
step: 190, loss: 0.005957355257123709
step: 200, loss: 0.0007454715087078512
step: 210, loss: 0.008089900948107243
step: 220, loss: 0.0027332683093845844
step: 230, loss: 0.26554879546165466
step: 240, loss: 0.11600121855735779
step: 250, loss: 0.045750461518764496
```

```
step: 260, loss: 0.0034014026168733835
step: 270, loss: 0.008206683211028576
step: 280, loss: 0.07357639074325562
step: 290, loss: 0.061247654259204865
step: 300, loss: 0.00023671610688325018
step: 310, loss: 0.003964574076235294
step: 320, loss: 0.08019253611564636
step: 330, loss: 0.0027857949025928974
step: 340, loss: 0.0376526415348053
step: 350, loss: 0.0020336767192929983
step: 360, loss: 0.04809261113405228
step: 370, loss: 0.015599925071001053
step: 380, loss: 0.0068705035373568535
step: 390, loss: 0.0010084818350151181
step: 400, loss: 0.0014838753268122673
step: 410, loss: 0.004281720612198114
step: 420, loss: 0.0007736764964647591
step: 430, loss: 0.07263999432325363
step: 440, loss: 0.007088647224009037
step: 450, loss: 0.001683628186583519
step: 460, loss: 0.00013786710042040795
step: 470, loss: 0.01417747512459755
step: 480, loss: 0.1459924280643463
step: 490, loss: 0.00280265810340643
step: 500, loss: 0.007866870611906052
step: 510, loss: 0.05457304045557976
step: 520, loss: 0.04830816760659218
step: 530, loss: 0.021316329017281532
/usr/local/lib/python3.7/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=0.9427640763145649, f1=0.9211267605633803,
best_f1=0.9354988399071927
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0065442584455013275
step: 10, loss: 0.015401566401124
step: 20, loss: 0.0005035459762439132
step: 30, loss: 0.007809363771229982
step: 40, loss: 0.06171258166432381
step: 50, loss: 0.00033302936935797334
step: 60, loss: 0.021417973563075066
step: 70, loss: 0.00597331952303648
step: 80, loss: 0.00046365405432879925
step: 90, loss: 0.06677165627479553
step: 100, loss: 0.03670931234955788
step: 110, loss: 0.0003650574362836778
```

```
step: 120, loss: 0.28127509355545044
step: 130, loss: 0.0014671020908281207
step: 140, loss: 0.016812866553664207
step: 150, loss: 0.0006488834624178708
step: 160, loss: 0.0008260426111519337
step: 170, loss: 0.04333383962512016
step: 180, loss: 0.0008339261985383928
step: 190, loss: 0.0011994760716333985
step: 200, loss: 0.0005833275499753654
step: 210, loss: 0.000534181366674602
step: 220, loss: 0.0021475430112332106
step: 230, loss: 0.001652037026360631
step: 240, loss: 0.0020197622943669558
step: 250, loss: 0.0007197251543402672
step: 260, loss: 0.0010759307770058513
step: 270, loss: 0.00020407364354468882
step: 280, loss: 0.009686995297670364
step: 290, loss: 0.12256216257810593
step: 300, loss: 0.0033532017841935158
step: 310, loss: 0.00026903304387815297
step: 320, loss: 0.17594820261001587
step: 330, loss: 0.028917592018842697
step: 340, loss: 0.004591467324644327
step: 350, loss: 0.00312393088825047
step: 360, loss: 0.012666448950767517
step: 370, loss: 0.004586067982017994
step: 380, loss: 0.009680964052677155
step: 390, loss: 0.0003097923472523689
step: 400, loss: 0.018099164590239525
step: 410, loss: 0.00317774317227304
step: 420, loss: 0.00208106217905879
step: 430, loss: 0.0008747575338929892
step: 440, loss: 0.02083706296980381
step: 450, loss: 0.0005861615645699203
step: 460, loss: 0.00038406564272008836
step: 470, loss: 0.006254795007407665
step: 480, loss: 0.001845062244683504
step: 490, loss: 0.0005312507273629308
step: 500, loss: 0.008805819787085056
step: 510, loss: 0.1228424683213234
step: 520, loss: 0.002689980436116457
step: 530, loss: 0.003558252239599824
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9443665264142123, f1=0.927796130250118,
best_f1=0.927796130250118
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0014722750056535006
step: 10, loss: 0.000580055289901793
step: 20, loss: 0.03210533410310745
step: 30, loss: 0.014781951904296875
step: 40, loss: 0.00038857501931488514
step: 50, loss: 0.08755011856555939
step: 60, loss: 0.00015033849922474474
step: 70, loss: 0.0016178057994693518
step: 80, loss: 0.013393661007285118
step: 90, loss: 0.00021304235269781202
step: 100, loss: 0.0016620771493762732
step: 110, loss: 0.0009014095412567258
step: 120, loss: 0.00016047271492425352
step: 130, loss: 0.04183531925082207
step: 140, loss: 0.027408452704548836
step: 150, loss: 0.00257907179184258
step: 160, loss: 0.0017884689150378108
step: 170, loss: 0.0019395156996324658
step: 180, loss: 0.00017158810805995017
step: 190, loss: 0.00037509328103624284
step: 200, loss: 0.026874886825680733
step: 210, loss: 0.0003220711078029126
step: 220, loss: 0.002009822055697441
step: 230, loss: 0.0001576813228894025
step: 240, loss: 0.07514437288045883
step: 250, loss: 0.00031724601285532117
step: 260, loss: 0.0002520749403629452
step: 270, loss: 0.0355580635368824
step: 280, loss: 0.0009859175188466907
step: 290, loss: 0.001747244386933744
step: 300, loss: 0.00274930358864367
step: 310, loss: 0.0012286537094041705
step: 320, loss: 0.0005308396648615599
step: 330, loss: 0.0003123650676570833
step: 340, loss: 0.04458165168762207
step: 350, loss: 0.0014054151251912117
step: 360, loss: 0.01646062545478344
step: 370, loss: 0.0017760739428922534
step: 380, loss: 9.918784053297713e-05
step: 390, loss: 0.015925323590636253
step: 400, loss: 0.0046504465863108635
step: 410, loss: 0.0007588238804601133
```

```
step: 420, loss: 0.02249894104897976
step: 430, loss: 0.0032003656961023808
step: 440, loss: 0.0005326466052792966
step: 450, loss: 0.0014044526033103466
step: 460, loss: 0.00026330258697271347
step: 470, loss: 0.001081818132661283
step: 480, loss: 0.00121210771612823
step: 490, loss: 0.0060701146721839905
step: 500, loss: 0.002458169823512435
step: 510, loss: 0.0035434644669294357
step: 520, loss: 0.004271852318197489
step: 530, loss: 0.013587447814643383
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9445221445221446, f1=0.9327731092436975,
best_f1=0.9327731092436975
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00012078944564564154
step: 10, loss: 0.00014291731349658222
step: 20, loss: 0.009677836671471596
step: 30, loss: 0.00020126551680732518
step: 40, loss: 0.00048303065705113113
step: 50, loss: 0.00028624440892599523
Gradient overflow.  Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 60, loss: 0.0038012166041880846
step: 70, loss: 8.811733277980238e-05
step: 80, loss: 0.0014472572365775704
step: 90, loss: 0.014678142964839935
step: 100, loss: 9.394087828695774e-05
step: 110, loss: 0.0009082570904865861
step: 120, loss: 0.00029536665533669293
step: 130, loss: 0.001845344202592969
step: 140, loss: 0.000203657808015123
step: 150, loss: 0.00043896649731323123
step: 160, loss: 0.0008892358164303005
step: 170, loss: 0.0008139497367665172
step: 180, loss: 0.004433443769812584
step: 190, loss: 0.0030799710657447577
step: 200, loss: 0.00030224976944737136
step: 210, loss: 0.028197510167956352
step: 220, loss: 0.0003109747776761651
```

```
step: 230, loss: 0.002215280197560787
step: 240, loss: 0.0029697760473936796
step: 250, loss: 0.0002696537703741342
step: 260, loss: 0.0002178713766625151
step: 270, loss: 0.0002813331375364214
step: 280, loss: 0.017970560118556023
step: 290, loss: 9.982487972592935e-05
step: 300, loss: 0.002331648487597704
step: 310, loss: 0.0002233903214801103
step: 320, loss: 0.00033017329405993223
step: 330, loss: 0.011236673220992088
step: 340, loss: 0.005445412360131741
step: 350, loss: 0.0006032777018845081
step: 360, loss: 0.0008671144023537636
step: 370, loss: 0.009700414724647999
step: 380, loss: 0.0002848602889571339
step: 390, loss: 0.00037600655923597515
step: 400, loss: 0.01208296325057745
step: 410, loss: 0.001140745123848319
step: 420, loss: 0.0004999063676223159
step: 430, loss: 0.00026117145898751915
step: 440, loss: 8.459188393317163e-05
step: 450, loss: 9.718616638565436e-05
step: 460, loss: 0.0007450153934769332
step: 470, loss: 0.003453758778050542
step: 480, loss: 0.0011901456164196134
step: 490, loss: 0.000536040693987160 9
step: 500, loss: 0.0002722563222050667
step: 510, loss: 0.0020879856310784817
step: 520, loss: 0.08259718120098114
step: 530, loss: 0.016101332381367683
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9475638051044084, f1=0.9349442379182156,
best_f1=0.9349442379182156
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00033275847090408206
step: 10, loss: 0.0888376235961914
step: 20, loss: 0.00011313831782899797
step: 30, loss: 0.004674664232879877
step: 40, loss: 0.00019020926265511662
```

```
step: 50, loss: 0.004598146304488182
step: 60, loss: 0.000991497072391212
step: 70, loss: 0.00012934762344229966
step: 80, loss: 0.0003294001508038491
step: 90, loss: 0.0009120390750467777
step: 100, loss: 0.00018762817489914596
step: 110, loss: 0.00021395388466771692
step: 120, loss: 0.0005554570234380662
step: 130, loss: 0.0009164518560282886
step: 140, loss: 0.008454961702227592
step: 150, loss: 9.441168367629871e-05
step: 160, loss: 0.008967995643615723
step: 170, loss: 0.0017946702428162098
step: 180, loss: 0.00062397069996222285
step: 190, loss: 0.002800486981868744
step: 200, loss: 0.0005824528052471578
step: 210, loss: 0.0001416657614754513
step: 220, loss: 0.001734041958115995
step: 230, loss: 0.0010843495838344097
step: 240, loss: 0.00017280899919569492
step: 250, loss: 4.524394898908213e-05
step: 260, loss: 2.534641680540517e-05
step: 270, loss: 0.0009379719267599285
step: 280, loss: 0.05364244431257248
step: 290, loss: 3.4081323974533007e-05
step: 300, loss: 0.0001253883820027113
step: 310, loss: 0.00032816905877552927
step: 320, loss: 6.64361723465845e-05
step: 330, loss: 0.0007293806411325932
step: 340, loss: 6.607022078242153e-05
step: 350, loss: 2.8624466722249053e-05
step: 360, loss: 0.0001775790733518079
step: 370, loss: 0.0003979323082603514
step: 380, loss: 0.13780739903450012
step: 390, loss: 0.1443852037191391
step: 400, loss: 0.0003352757776156068
step: 410, loss: 0.00036524931783787906
step: 420, loss: 0.0033569294027984142
step: 430, loss: 0.025360478088259697
step: 440, loss: 0.0004312653618399054
step: 450, loss: 0.0003283280530013144
step: 460, loss: 0.00024406707962043583
step: 470, loss: 0.0003828303306363523
step: 480, loss: 0.002544017741456628
step: 490, loss: 0.005405545234680176
step: 500, loss: 0.005126782227307558
step: 510, loss: 0.0024588333908468485
step: 520, loss: 0.0014431508025154471
```

```
step: 530, loss: 9.423073788639158e-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=0.9346497414198403, f1=0.9262564584311883,
best_f1=0.9349442379182156
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.035900358110666275
step: 10, loss: 0.08569949120283127
step: 20, loss: 0.005672179628163576
step: 30, loss: 0.000983733800239861
step: 40, loss: 0.0005025133723393083
step: 50, loss: 0.00017161104187835008
step: 60, loss: 0.0003240134392399341
step: 70, loss: 0.004549202974885702
step: 80, loss: 0.006504269782453775
step: 90, loss: 0.0002531386853661388
step: 100, loss: 0.00014951755292713642
step: 110, loss: 0.0005245048669166863
step: 120, loss: 0.0002129328204318881
step: 130, loss: 0.008850546553730965
step: 140, loss: 0.03750121593475342
step: 150, loss: 0.0002702741476241499
step: 160, loss: 0.0023221548181027174
step: 170, loss: 0.012745882384479046
step: 180, loss: 3.3772215829230845e-05
step: 190, loss: 4.07075131079182e-05
step: 200, loss: 0.00010512029257370159
step: 210, loss: 0.00010857820598175749
step: 220, loss: 0.000540293229278177
step: 230, loss: 7.047659164527431e-05
step: 240, loss: 5.20244320796337e-05
step: 250, loss: 0.0004027351678814739
step: 260, loss: 0.050170186907052994
step: 270, loss: 3.324328281451017e-05
step: 280, loss: 0.00041969577432610095
step: 290, loss: 0.024570681154727936
step: 300, loss: 0.00022643906413577497
step: 310, loss: 0.027915531769394875
step: 320, loss: 0.00014720480248797685
step: 330, loss: 0.00037520338082686067
step: 340, loss: 2.6504883862799034e-05
step: 350, loss: 0.000402268604375422
step: 360, loss: 2.590149233583361e-05
step: 370, loss: 0.0012039928697049618
step: 380, loss: 4.8166883061639965e-05
```

```
step: 390, loss: 2.422889338049572e-05
step: 400, loss: 0.000761119881644845
step: 410, loss: 3.1077874155016616e-05
step: 420, loss: 3.7254198105074465e-05
step: 430, loss: 2.5372439267812297e-05
step: 440, loss: 8.324240479851142e-05
step: 450, loss: 6.196922186063603e-05
step: 460, loss: 3.4971515560755506e-05
step: 470, loss: 2.09358458960196e-05
step: 480, loss: 3.118707172689028e-05
step: 490, loss: 0.00019429599934240207
step: 500, loss: 0.0014911815524101257
step: 510, loss: 0.008601892739534378
step: 520, loss: 3.112759441137314e-05
step: 530, loss: 0.003350219689309597
/usr/local/lib/python3.7/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=0.9359742054352833, f1=0.9274563820018366,
best_f1=0.9349442379182156
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0019074585288763046
step: 10, loss: 0.00030217779567465186
step: 20, loss: 0.06172877177596092
step: 30, loss: 8.043101843213663e-05
step: 40, loss: 0.007077313028275967
step: 50, loss: 0.11527369916439056
step: 60, loss: 3.7698580854339525e-05
step: 70, loss: 0.0003129063989035785
step: 80, loss: 0.0003905271296389401
step: 90, loss: 0.00038872819277457893
step: 100, loss: 4.242406430421397e-05
step: 110, loss: 7.542043749708682e-05
step: 120, loss: 4.4429092667996883e-05
step: 130, loss: 5.744767258875072e-05
step: 140, loss: 0.0034593220334500074
step: 150, loss: 2.9678678401978686e-05
step: 160, loss: 3.157074024784379e-05
step: 170, loss: 0.004440210293978453
step: 180, loss: 0.014275982044637203
step: 190, loss: 0.0009144871146418154
step: 200, loss: 0.0005382047384046018
step: 210, loss: 0.006974448449909687
step: 220, loss: 0.0012798068346455693
step: 230, loss: 0.00044984364649280906
step: 240, loss: 0.00016055967716965824
```

```
step: 250, loss: 7.072094740578905e-05
step: 260, loss: 8.953754877438769e-05
step: 270, loss: 0.00019886373775079846
step: 280, loss: 0.00019089740817435086
step: 290, loss: 9.361794946016744e-05
step: 300, loss: 3.408826523809694e-05
step: 310, loss: 0.0005290798144415021
step: 320, loss: 0.0012029283680021763
step: 330, loss: 0.0004954914911650121
step: 340, loss: 0.0001736489066388458
step: 350, loss: 0.0006868992932140827
step: 360, loss: 5.698019958799705e-05
step: 370, loss: 0.0017844465328380466
step: 380, loss: 0.011674645356833935
step: 390, loss: 0.0005058327806182206
step: 400, loss: 0.00019093760005642116
step: 410, loss: 0.0016935961320996284
step: 420, loss: 0.00010964230023091659
step: 430, loss: 0.00011153284140164033
step: 440, loss: 0.005069886799901724
step: 450, loss: 3.154016667394899e-05
step: 460, loss: 6.978342571528628e-05
step: 470, loss: 0.06540845334529877
step: 480, loss: 0.0018054820830002427
step: 490, loss: 0.004261494614183903
step: 500, loss: 0.004970652516931295
step: 510, loss: 8.936840458773077e-05
step: 520, loss: 0.00012036192492814735
step: 530, loss: 0.0006637206533923745
/usr/local/lib/python3.7/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=0.9424426766495088, f1=0.9349670122525918,
best_f1=0.9349442379182156
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 5.4150030337041244e-05
step: 10, loss: 0.0005612321547232568
step: 20, loss: 0.00010091267176903784
step: 30, loss: 0.0008609667420387268
step: 40, loss: 9.940540621755645e-05
step: 50, loss: 0.002215864835307002
step: 60, loss: 0.00028721755370497704
step: 70, loss: 6.252850289456546e-05
step: 80, loss: 6.537089211633429e-05
step: 90, loss: 3.635667235357687e-05
step: 100, loss: 0.00020437703642528504
```

```
step: 110, loss: 0.0012185610830783844
step: 120, loss: 9.330708417110145e-05
step: 130, loss: 9.665815014159307e-05
step: 140, loss: 0.0007849445100873709
step: 150, loss: 2.1494533939403482e-05
step: 160, loss: 3.738162195077166e-05
step: 170, loss: 3.065380224143155e-05
step: 180, loss: 0.00012036920816171914
step: 190, loss: 0.00032386204111389816
step: 200, loss: 0.0025339967105537653
step: 210, loss: 0.0014430651208385825
step: 220, loss: 0.001346826902590692
step: 230, loss: 1.810812136682216e-05
step: 240, loss: 0.0007832312257960439
step: 250, loss: 1.834304930525832e-05
step: 260, loss: 9.615595627110451e-05
step: 270, loss: 2.7704281819751486e-05
step: 280, loss: 3.538517194101587e-05
step: 290, loss: 4.926165274810046e-05
step: 300, loss: 3.831532740150578e-05
step: 310, loss: 4.379485471872613e-05
step: 320, loss: 2.9275293854880147e-05
step: 330, loss: 2.8080368792871013e-05
step: 340, loss: 3.29681315633934e-05
step: 350, loss: 3.991400444647297e-05
step: 360, loss: 1.8413851648801938e-05
step: 370, loss: 0.000699808239005506
step: 380, loss: 2.5174864276777953e-05
step: 390, loss: 3.033691245946102e-05
step: 400, loss: 5.0011101848213e-05
step: 410, loss: 4.078200436197221e-05
step: 420, loss: 5.241105463937856e-05
step: 430, loss: 2.3740760298096575e-05
step: 440, loss: 0.00010826743528014049
step: 450, loss: 0.0003316767397336662
step: 460, loss: 0.0010340400040149689
step: 470, loss: 6.627482798648998e-05
step: 480, loss: 0.0021977233700454235
step: 490, loss: 0.005473866127431393
step: 500, loss: 5.261131809675135e-05
step: 510, loss: 7.009629189269617e-05
step: 520, loss: 8.390354923903942e-05
step: 530, loss: 2.5453733542235568e-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 11: dev_f1=0.9432029795158287, f1=0.9325842696629213,
best_f1=0.9349442379182156
```

```
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 2.684379433048889e-05
step: 10, loss: 2.0559560653055087e-05
step: 20, loss: 7.10170206730254e-05
step: 30, loss: 0.017044879496097565
step: 40, loss: 0.0008147166226990521
step: 50, loss: 0.03231536224484444
step: 60, loss: 0.001815488561987877
step: 70, loss: 2.3398082703351974e-05
step: 80, loss: 4.620687468559481e-05
step: 90, loss: 2.3789079932612367e-05
step: 100, loss: 4.731542139779776e-05
step: 110, loss: 2.658274934219662e-05
step: 120, loss: 5.2597872127080336e-05
step: 130, loss: 0.0011934642679989338
step: 140, loss: 3.467969145276584e-05
step: 150, loss: 5.3532672609435394e-05
step: 160, loss: 1.7165915778605267e-05
step: 170, loss: 0.0010301629081368446
step: 180, loss: 0.0004798089503310621
step: 190, loss: 0.0006501968600787222
step: 200, loss: 2.911157935159281e-05
Gradient overflow.  Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 210, loss: 3.0559491278836504e-05
step: 220, loss: 5.469157622428611e-05
step: 230, loss: 2.780738941510208e-05
step: 240, loss: 1.894240631372668e-05
step: 250, loss: 1.7694877897156402e-05
step: 260, loss: 1.9944875020883046e-05
step: 270, loss: 2.692179987207055e-05
step: 280, loss: 0.001557032112032175
step: 290, loss: 6.852135265944526e-05
step: 300, loss: 0.00062432431150227779
step: 310, loss: 0.0001528951688669622
step: 320, loss: 1.6066893294919282e-05
step: 330, loss: 3.5295375710120425e-05
step: 340, loss: 2.102134021697566e-05
step: 350, loss: 0.06537846475839615
step: 360, loss: 6.644369568675756e-05
step: 370, loss: 3.8473601307487115e-05
step: 380, loss: 0.0004073140735272318
step: 390, loss: 8.273391722468659e-05
step: 400, loss: 2.7476717150420882e-05
step: 410, loss: 0.00013294164091348648
step: 420, loss: 0.004273209720849991
step: 430, loss: 0.008469559252262115
```

```
step: 440, loss: 5.51508528587874e-05
step: 450, loss: 0.000255387945799157
step: 460, loss: 0.00014757426106370986
step: 470, loss: 1.9717597751878202e-05
step: 480, loss: 0.00038425609818659723
step: 490, loss: 6.563869828823954e-05
step: 500, loss: 3.27457491948735e-05
step: 510, loss: 0.0004365809727460146
step: 520, loss: 2.2623031327384524e-05
step: 530, loss: 3.9045375160640106e-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=0.9432723863103609, f1=0.9338959212376934,
best_f1=0.9349442379182156
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0021187704987823963
step: 10, loss: 1.6614578271401115e-05
step: 20, loss: 3.8112502807052806e-05
step: 30, loss: 2.4049908461165614e-05
step: 40, loss: 3.515990829328075e-05
step: 50, loss: 2.8497399398474954e-05
step: 60, loss: 2.8828082577092573e-05
step: 70, loss: 1.8596158042782918e-05
step: 80, loss: 1.3135260815033689e-05
step: 90, loss: 0.00607477780431509
step: 100, loss: 4.562237882055342e-05
step: 110, loss: 1.8775113858282566e-05
step: 120, loss: 6.44843967165798e-05
step: 130, loss: 2.3520051399827935e-05
step: 140, loss: 5.7684512285050005e-05
step: 150, loss: 0.0002632945543155074
step: 160, loss: 4.811092003365047e-05
step: 170, loss: 0.00022079811606090516
step: 180, loss: 3.191962241544388e-05
step: 190, loss: 3.0158334993757308e-05
step: 200, loss: 1.8801290934788994e-05
step: 210, loss: 3.293979534646496e-05
step: 220, loss: 1.8268556232214905e-05
step: 230, loss: 1.5288399857201092e-05
step: 240, loss: 2.9216182156233117e-05
step: 250, loss: 1.7024334738380276e-05
step: 260, loss: 0.00016152577882166952
step: 270, loss: 1.6905110896914266e-05
step: 280, loss: 0.0005121264839544892
step: 290, loss: 3.820767597062513e-05
```

```
step: 300, loss: 1.740063089528121e-05
step: 310, loss: 0.0005043118144385517
step: 320, loss: 0.0024921754375100136
step: 330, loss: 0.0013212822377681732
step: 340, loss: 2.0548113752738573e-05
step: 350, loss: 2.083496656268835e-05
step: 360, loss: 7.121159433154389e-05
step: 370, loss: 2.127837979060132e-05
step: 380, loss: 1.9441917174845003e-05
step: 390, loss: 0.0002532093785703182
step: 400, loss: 5.2248298743506894e-05
step: 410, loss: 0.00107634998857975
step: 420, loss: 2.8899290555273183e-05
step: 430, loss: 8.142222941387445e-05
step: 440, loss: 3.002036282850895e-05
step: 450, loss: 6.147746171336621e-05
step: 460, loss: 0.0016409845557063818
step: 470, loss: 1.2826050806324929e-05
step: 480, loss: 1.8302038370165974e-05
step: 490, loss: 1.981060631806031e-05
step: 500, loss: 0.00020601005235221237
step: 510, loss: 3.336286681587808e-05
step: 520, loss: 0.0003100034373346716
step: 530, loss: 0.00019265408627688885
/usr/local/lib/python3.7/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=0.9415067852129153, f1=0.930341280972417,
best_f1=0.9349442379182156
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.000868661911226809
step: 10, loss: 1.227846496476559e-05
step: 20, loss: 1.4278882190410513e-05
step: 30, loss: 1.5456047549378127e-05
step: 40, loss: 5.9080804930999875e-05
step: 50, loss: 0.00029316023574210703
step: 60, loss: 1.8834507500287145e-05
step: 70, loss: 5.071176929050125e-05
step: 80, loss: 3.094153726124205e-05
step: 90, loss: 1.6428275557700545e-05
step: 100, loss: 2.1978346921969205e-05
step: 110, loss: 4.7870282287476584e-05
step: 120, loss: 3.1450494134332985e-05
step: 130, loss: 0.028756888583302498
step: 140, loss: 5.5009510106174275e-05
step: 150, loss: 1.2233758752699941e-05
```

```
step: 160, loss: 0.0011182681191712618
step: 170, loss: 6.393908552126959e-05
step: 180, loss: 1.326564324699575e-05
step: 190, loss: 1.3988328646519221e-05
step: 200, loss: 2.3740514734527096e-05
step: 210, loss: 1.349660669802688e-05
step: 220, loss: 2.189666156482417e-05
step: 230, loss: 1.9546254407032393e-05
step: 240, loss: 1.3265635061543435e-05
step: 250, loss: 2.4202247004723176e-05
step: 260, loss: 0.0002520121051929891
step: 270, loss: 1.4144735359877814e-05
step: 280, loss: 2.0417646737769246e-05
step: 290, loss: 0.0006317399092949927
step: 300, loss: 3.9391739846905693e-05
step: 310, loss: 7.702823495492339e-05
step: 320, loss: 4.342442116467282e-05
step: 330, loss: 3.638501948444173e-05
step: 340, loss: 0.0002186959027312696
step: 350, loss: 2.982066507684067e-05
step: 360, loss: 0.0001576800859766081
step: 370, loss: 1.877510476333555e-05
step: 380, loss: 5.421416062745266e-05
step: 390, loss: 0.00196754839271307
step: 400, loss: 0.00014047326112631708
step: 410, loss: 1.3004852917219978e-05
step: 420, loss: 1.6193351257243194e-05
step: 430, loss: 2.2459260435425676e-05
step: 440, loss: 5.0391921831760556e-05
step: 450, loss: 0.00011274118151050061
step: 460, loss: 1.9694754882948473e-05
step: 470, loss: 5.077738023828715e-05
step: 480, loss: 1.8253618691232987e-05
step: 490, loss: 2.209755257354118e-05
step: 500, loss: 6.998073302172482e-05
step: 510, loss: 2.1717778508900665e-05
step: 520, loss: 2.375137592025567e-05
step: 530, loss: 2.117362237186171e-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=0.9446768944676894, f1=0.9369202226345084,
best_f1=0.9349442379182156
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 1.6055788364610635e-05
step: 10, loss: 1.56497389980359e-05
```

```
step: 20, loss: 1.619734575797338e-05
step: 30, loss: 2.5267048840760253e-05
step: 40, loss: 0.06811527907848358
step: 50, loss: 4.610866017173976e-05
step: 60, loss: 6.52066373731941e-05
step: 70, loss: 5.738526306231506e-05
step: 80, loss: 1.4141021893010475e-05
step: 90, loss: 1.6636933651170693e-05
step: 100, loss: 3.104678398813121e-05
step: 110, loss: 0.00035107709118165076
step: 120, loss: 1.848823558248114e-05
step: 130, loss: 0.09543679654598236
step: 140, loss: 1.5020188584458083e-05
step: 150, loss: 1.6610865714028478e-05
step: 160, loss: 1.4677460058010183e-05
step: 170, loss: 1.3247007700556424e-05
step: 180, loss: 1.4390592696145177e-05
step: 190, loss: 1.7389445929438807e-05
step: 200, loss: 1.5571500625810586e-05
step: 210, loss: 0.00045110666542313993
step: 220, loss: 1.3828144801664166e-05
step: 230, loss: 4.031208300148137e-05
step: 240, loss: 1.5411324056913145e-05
step: 250, loss: 2.2958149202167988e-05
step: 260, loss: 2.5434053895878606e-05
step: 270, loss: 1.5467214325326495e-05
step: 280, loss: 1.4167132576403674e-05
step: 290, loss: 1.1891031135746744e-05
step: 300, loss: 1.2811158740078099e-05
step: 310, loss: 7.318014104384929e-05
step: 320, loss: 5.340143616194837e-05
step: 330, loss: 2.658828452695161e-05
step: 340, loss: 1.1511061529745348e-05
step: 350, loss: 1.1350874956406187e-05
step: 360, loss: 0.0016202251426875591
step: 370, loss: 1.284839163417928e-05
step: 380, loss: 2.0801364371436648e-05
step: 390, loss: 0.00016117018822114915
step: 400, loss: 1.9393430193304084e-05
step: 410, loss: 2.340115315746516e-05
step: 420, loss: 3.463691973593086e-05
step: 430, loss: 2.4768720322754234e-05
step: 440, loss: 0.00020703398331534117
step: 450, loss: 1.1674964298435953e-05
step: 460, loss: 1.3403475350060035e-05
step: 470, loss: 0.0012729786103591323
step: 480, loss: 3.249774454161525e-05
step: 490, loss: 1.2174137737019919e-05
```

```
step: 500, loss: 1.4610395737690851e-05
step: 510, loss: 2.6403866286273114e-05
step: 520, loss: 1.4424172150029335e-05
step: 530, loss: 0.0002801791997626424
/usr/local/lib/python3.7/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=0.9463955637707948, f1=0.9371534195933457,
best_f1=0.9349442379182156
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 [00:26, 215.35it/s]
load_f1 = 0.9479553903345725
real_f1 = 0.9480519480519481
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 [00:19, 226.79it/s]
```

3 DK

```
[ ]: # Para executar 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 12.7 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: 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: 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: 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: 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: 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: 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: 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: 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: 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: 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: 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: 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: 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: 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: 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: 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: 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: 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=936c07b8b1446469558e86da218c38468de7f68a8438700f3b3ff9128ce1ea9f
  Stored in directory: /tmp/pip-ephem-wheel-cache-h5o83u1n/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 bert-base-uncased \
       --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 bert-base-uncased \
       --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: 0.5824135541915894
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.35294117647058826, f1=0.3333333333333333,
best_f1=0.3333333333333333
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.5115506649017334
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.4, f1=0.4, best_f1=0.4
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
```

```
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.5545016527175903
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.6000000000000001, f1=0.5000000000000001,
best_f1=0.5000000000000001
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.2693830728530884
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.6956521739130435, f1=0.5384615384615384,
best_f1=0.5384615384615384
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.3034687638282776
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.7586206896551724, 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: 0.19246216118335724
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
  warnings.warn(SAVE_STATE_WARNING, UserWarning)
```

```
epoch 6: dev_f1=0.8, f1=0.7200000000000001, best_f1=0.7200000000000001
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.028797170147299767
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
  warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 7: dev_f1=0.8666666666666666, f1=0.7333333333333334,
best_f1=0.7333333333333334
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.06198470667004585
/usr/local/lib/python3.7/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=0.8275862068965518, f1=0.689655172413793,
best_f1=0.7333333333333334
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.002338794758543372
/usr/local/lib/python3.7/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=0.8275862068965518, f1=0.689655172413793,
best_f1=0.7333333333333334
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.01109262928366661
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.8750000000000001, f1=0.7741935483870968,
best_f1=0.7741935483870968
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
```

```
step: 0, loss: 0.0022751064971089363
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
epoch 11: dev_f1=0.8666666666666666, f1=0.7333333333333334,
best_f1=0.7741935483870968
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0031614163890480995
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
epoch 12: dev_f1=0.8666666666666666, f1=0.8125000000000001,
best_f1=0.7741935483870968
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0029105134308338165
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
epoch 13: dev_f1=0.8666666666666666, f1=0.8125000000000001,
best_f1=0.7741935483870968
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0032919321674853563
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
epoch 14: dev_f1=0.8666666666666666, f1=0.8125000000000001,
best_f1=0.7741935483870968
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.002949628047645092
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
epoch 15: dev_f1=0.8666666666666666, f1=0.8125000000000001,
best_f1=0.7741935483870968
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.")
91it [00:00, 136950.72it/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 = 0.9032258064516129
real_f1 = 0.8666666666666666
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, 201.03it/s]
```

3.1.2 DBLP-ACM - Running the matcher

```python
# 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 bert-base-uncased \
```

```
    --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 bert-base-uncased \
    --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

255

/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: 0.5789301991462708
step: 10, loss: 0.6137872338294983
step: 20, loss: 0.3091577887535095
step: 30, loss: 0.1237759217619896
step: 40, loss: 0.24463814496994019
step: 50, loss: 0.2299698144197464
step: 60, loss: 0.11966574192047119
step: 70, loss: 0.0129395117983222
step: 80, loss: 0.05258003994822502
step: 90, loss: 0.17542080581188202
step: 100, loss: 0.005857539363205433
step: 110, loss: 0.17704127728939056
step: 120, loss: 0.009904581122100353
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 130, loss: 0.004264908377081156
step: 140, loss: 0.002968188840895891
step: 150, loss: 0.010766617953777313
step: 160, loss: 0.006788244470953941
step: 170, loss: 0.15663154423236847
step: 180, loss: 0.09027890861034393
step: 190, loss: 0.004888301249593496
step: 200, loss: 0.004113960079848766
step: 210, loss: 0.0030322372913360596
step: 220, loss: 0.02290383353829384
step: 230, loss: 0.04465249553322792
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9784335981838819, f1=0.9748283752860413,
best_f1=0.9748283752860413
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.014785799197852612
step: 10, loss: 0.002340613631531596

```
step: 20, loss: 0.06454721838235855
step: 30, loss: 0.21658356487751007
step: 40, loss: 0.05651390179991722
step: 50, loss: 0.021752888336777687
step: 60, loss: 0.008269977755844593
step: 70, loss: 0.16264723241329193
step: 80, loss: 0.02326761931180954
step: 90, loss: 0.013595158234238625
step: 100, loss: 0.1310613751411438
step: 110, loss: 0.010860726237297058
step: 120, loss: 0.0018513525137677789
step: 130, loss: 0.011559084989130497
step: 140, loss: 0.008608533069491386
step: 150, loss: 0.003684011520817876
step: 160, loss: 0.0014799832133576274
step: 170, loss: 0.0009657465852797031
step: 180, loss: 0.0032463683746755123
step: 190, loss: 0.0019865590147674084
step: 200, loss: 0.012713221833109856
step: 210, loss: 0.0013918100157752633
step: 220, loss: 0.02818098850548674
step: 230, loss: 0.00275380234234035
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9910112359550561, f1=0.9864864864864865,
best_f1=0.9864864864864865
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0075344455973654985
step: 10, loss: 0.0038136125076562166
step: 20, loss: 0.0008156837429851294
step: 30, loss: 0.02373586595058441
step: 40, loss: 0.011023142375051975
step: 50, loss: 0.016957266256213188
step: 60, loss: 0.001377667998895049
step: 70, loss: 0.0011758541222661734
step: 80, loss: 0.0272555872797966
step: 90, loss: 0.16801594197750092
step: 100, loss: 0.03303420543670654
step: 110, loss: 0.006063293199986219
step: 120, loss: 0.013111559674143791
step: 130, loss: 0.010423421859741211
```

```
step: 140, loss: 0.0088510103523373123
step: 150, loss: 0.08775696158409119
step: 160, loss: 0.008073955774307251
step: 170, loss: 0.00827016867697239
step: 180, loss: 0.004778970964252949
step: 190, loss: 0.0031471664551645517
step: 200, loss: 0.09852371364831924
step: 210, loss: 0.017431670799851418
step: 220, loss: 0.0011333677684888244
step: 230, loss: 0.0012397998943924904
/usr/local/lib/python3.7/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=0.9865168539325843, f1=0.9807474518686297,
best_f1=0.9864864864864865
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0025665683206170797
step: 10, loss: 0.00037208598223514855
step: 20, loss: 0.0007187819574028254
step: 30, loss: 0.0005324904923327267
step: 40, loss: 0.007887943647801876
step: 50, loss: 0.001898848102428019
step: 60, loss: 0.059793103486299515
step: 70, loss: 0.0009489909280091524
step: 80, loss: 0.0023602142464369535
step: 90, loss: 0.002676488598808646
step: 100, loss: 0.001125846989452839
step: 110, loss: 0.0018777906661853194
step: 120, loss: 0.10564666986465454
step: 130, loss: 0.03806113451719284
step: 140, loss: 0.001880354480817914
step: 150, loss: 0.15520089864730835
step: 160, loss: 0.07911291718482971
step: 170, loss: 0.03405435383319855
step: 180, loss: 0.0007204685243777931
step: 190, loss: 0.011924228630959988
step: 200, loss: 0.004841760266572237
step: 210, loss: 0.014650719240307808
step: 220, loss: 0.0005343310185708106
step: 230, loss: 0.0015393487410619855
/usr/local/lib/python3.7/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=0.9898762654668166, f1=0.9829738933030647,
best_f1=0.9864864864864865
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
```

```
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0031950471457093954
step: 10, loss: 0.002927830209955573
step: 20, loss: 0.029693299904465675
step: 30, loss: 0.00016181018145289272
step: 40, loss: 0.0026565983425825834
step: 50, loss: 0.00035297288559377193
step: 60, loss: 0.004111907910555601
step: 70, loss: 0.0004934178432449698
step: 80, loss: 0.023174215108156204
step: 90, loss: 0.0023681907914578915
step: 100, loss: 0.00033510272623971105
step: 110, loss: 0.00018988960073329508
step: 120, loss: 0.00010109437425853685
step: 130, loss: 0.005273755639791489
step: 140, loss: 0.003439879510551691
step: 150, loss: 0.002187910722568631
step: 160, loss: 0.0014069804456084967
step: 170, loss: 0.009851833805441856
step: 180, loss: 0.0012482665479183197
step: 190, loss: 0.01845129206776619
step: 200, loss: 0.010135342366993427
step: 210, loss: 0.000838595733512193
step: 220, loss: 0.00139195891097188
step: 230, loss: 0.0006822920404374599
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9921259842519685, f1=0.9841269841269841,
best_f1=0.9841269841269841
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00099791272077770944
step: 10, loss: 0.0004169714520685375
step: 20, loss: 0.0013955594040453434
step: 30, loss: 0.0005217864527367055
step: 40, loss: 0.0002951831847894937
step: 50, loss: 0.00048648074152879417
step: 60, loss: 0.00031185135594569147
step: 70, loss: 0.002967894310131669
step: 80, loss: 0.002492028521373868
step: 90, loss: 0.00036553386598825455
```

```
step: 100, loss: 0.004606527741998434
step: 110, loss: 0.0006851645885035396
step: 120, loss: 0.0008989517809823155
step: 130, loss: 0.004955099895596504
step: 140, loss: 0.0019232322229072452
step: 150, loss: 0.0003414804523345083
step: 160, loss: 0.012212550267577171
step: 170, loss: 0.000250812154263258
step: 180, loss: 0.015562781132757664
step: 190, loss: 0.006809313781559467
step: 200, loss: 0.0003291608882136643
step: 210, loss: 0.0014165174216032028
step: 220, loss: 0.0006656166515313089
step: 230, loss: 0.09115897864103317
/usr/local/lib/python3.7/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=0.9876819708846584, f1=0.9809203142536477,
best_f1=0.9841269841269841
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.007227765396237373
step: 10, loss: 0.0007093027816154063
step: 20, loss: 0.00019175377383362502
step: 30, loss: 0.00035861521610058844
step: 40, loss: 0.00015407246246468276
step: 50, loss: 0.00029042194364592433
step: 60, loss: 0.0002989218628499657
step: 70, loss: 0.001264981459826231
step: 80, loss: 0.00013484360533766448
step: 90, loss: 5.7578912674216554e-05
step: 100, loss: 0.00012075181439286098
step: 110, loss: 0.00010959240898955613
step: 120, loss: 7.558652578154579e-05
step: 130, loss: 0.00010028122778749093
step: 140, loss: 0.0001182976266136393
step: 150, loss: 0.004582305904477835
step: 160, loss: 0.07772865891456604
step: 170, loss: 0.014116867445409298
step: 180, loss: 0.003979107830673456
step: 190, loss: 0.0003603527438826859
step: 200, loss: 0.07276445627212524
step: 210, loss: 9.831254283199087e-05
step: 220, loss: 0.0006962300976738334
step: 230, loss: 0.0040547773241996765
/usr/local/lib/python3.7/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=0.9843400447427293, f1=0.9798206278026906,
best_f1=0.9841269841269841
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0004801071190740913
step: 10, loss: 0.00068961066642670929
step: 20, loss: 0.007008313201367855
step: 30, loss: 0.0028122293297201395
step: 40, loss: 0.00168038927949965
step: 50, loss: 0.005630503874272108
step: 60, loss: 0.006518939975649118
step: 70, loss: 0.00041659557609818876
step: 80, loss: 0.001426554867066443
step: 90, loss: 9.913367830449715e-05
step: 100, loss: 0.0024035212118178606
step: 110, loss: 0.007738585118204355
step: 120, loss: 0.00018537489813752472
step: 130, loss: 0.0004743838217109442
step: 140, loss: 0.0002847960568033159
step: 150, loss: 0.00046130624832585454
step: 160, loss: 0.0009200139320455492
step: 170, loss: 0.00010645444126566872
step: 180, loss: 0.00028737157117575407
step: 190, loss: 0.0002194373810198158
step: 200, loss: 0.0001372954429825768
step: 210, loss: 0.00010712930088629946
step: 220, loss: 0.00019100200734101236
step: 230, loss: 0.004157502204179764
/usr/local/lib/python3.7/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=0.9842696629213483, f1=0.9797297297297298,
best_f1=0.9841269841269841
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 8.574008825235069e-05
step: 10, loss: 0.0012375657679513097
step: 20, loss: 0.002960280515253544
step: 30, loss: 8.106107998173684e-05
step: 40, loss: 0.2015402913093567
step: 50, loss: 0.00012117742153350264
step: 60, loss: 0.00013762719754459531
step: 70, loss: 0.0004500467621255666
step: 80, loss: 0.0001467951515223831
step: 90, loss: 0.00021676765754818916
```

```
step: 100, loss: 0.0006885628099553287
step: 110, loss: 6.976993608986959e-05
step: 120, loss: 5.2992145356256515e-05
step: 130, loss: 0.00019848822557833046
step: 140, loss: 0.0002482416166458279
step: 150, loss: 0.00017151788051705807
step: 160, loss: 7.896271563367918e-05
step: 170, loss: 0.000569333671592176
step: 180, loss: 0.0034801666624844074
step: 190, loss: 0.00012177600001450628
step: 200, loss: 8.306517702294514e-05
step: 210, loss: 0.00014532654313370585
step: 220, loss: 7.319061842281371e-05
step: 230, loss: 0.0003772143099922687
/usr/local/lib/python3.7/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=0.9864864864864865, f1=0.9830890642615557,
best_f1=0.9841269841269841
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 7.941273361211643e-05
step: 10, loss: 5.1097169489366934e-05
step: 20, loss: 6.276130443438888e-05
step: 30, loss: 0.00029184223967604034
step: 40, loss: 0.00011186082701897249
step: 50, loss: 0.00029550740146078177
step: 60, loss: 0.02004363015294075
step: 70, loss: 0.00018578773597255355
step: 80, loss: 0.000137742783408612
step: 90, loss: 0.0010312896920368075
step: 100, loss: 0.0004580074746627360
step: 110, loss: 0.0006533077685162425
step: 120, loss: 9.450244397157803e-05
step: 130, loss: 0.0014871980529278517
step: 140, loss: 0.03895723819732666
step: 150, loss: 0.011745353229343891
step: 160, loss: 4.053274096804671e-05
step: 170, loss: 7.526126137236133e-05
step: 180, loss: 5.8591289416654035e-05
step: 190, loss: 0.0051427362486720085
step: 200, loss: 0.0001087260025087744
step: 210, loss: 0.00011465224088169634
step: 220, loss: 6.990075053181499e-05
step: 230, loss: 0.0003258753858972341
/usr/local/lib/python3.7/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=0.9864253393665158, f1=0.9818181818181818,
best_f1=0.9841269841269841
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 6.470645166700706e-05
step: 10, loss: 0.00023184741439763457
step: 20, loss: 0.0022831172682344913
step: 30, loss: 0.021799257025122643
step: 40, loss: 0.00013607714208774269
step: 50, loss: 8.351740689249709e-05
step: 60, loss: 0.00011450974852778018
step: 70, loss: 0.00035442746593616903
step: 80, loss: 6.911229866091162e-05
step: 90, loss: 3.972050399170257e-05
step: 100, loss: 9.25113563425839e-05
step: 110, loss: 0.04089193046092987
step: 120, loss: 4.515515320235863e-05
step: 130, loss: 4.072500087204389e-05
step: 140, loss: 0.00012911784870084375
step: 150, loss: 0.0408177375793457
step: 160, loss: 0.00011258802987867966
step: 170, loss: 0.03427638113498688
step: 180, loss: 8.276602602563798e-05
step: 190, loss: 8.068985334830359e-05
step: 200, loss: 0.0009587613749317825
step: 210, loss: 4.3113806896144524e-05
step: 220, loss: 0.00010809883679030463
step: 230, loss: 6.338296952890232e-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 11: dev_f1=0.9865168539325843, f1=0.9797297297297298,
best_f1=0.9841269841269841
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0001590008760103956
step: 10, loss: 5.1155711844330654e-05
step: 20, loss: 5.8312300097895786e-05
step: 30, loss: 0.0001252057554665953
step: 40, loss: 7.141969399526715e-05
step: 50, loss: 0.0004905465757474303
step: 60, loss: 0.0002488421159796417
step: 70, loss: 0.00010138455400010571
step: 80, loss: 0.00011954170622630045
step: 90, loss: 3.684512557811104e-05
```

```
step: 100, loss: 0.00038532985490746796
step: 110, loss: 5.324908488546498e-05
step: 120, loss: 0.00012355482613202184
step: 130, loss: 6.421072612283751e-05
step: 140, loss: 0.0059969741851091385
step: 150, loss: 0.00020262597536220252
step: 160, loss: 6.509012746391818e-05
step: 170, loss: 5.904031058889814e-05
step: 180, loss: 0.046905580908060074
step: 190, loss: 0.0003023470926564187
step: 200, loss: 4.963235551258549e-05
step: 210, loss: 0.0002131218498107046
step: 220, loss: 0.00020052219042554498
step: 230, loss: 0.05082706734538078
/usr/local/lib/python3.7/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=0.9853768278965129, f1=0.9774266365688488,
best_f1=0.9841269841269841
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00019610421441029757
step: 10, loss: 0.0001391398545820266
step: 20, loss: 8.048207382671535e-05
step: 30, loss: 0.00023955918732099235
step: 40, loss: 6.796608795411885e-05
step: 50, loss: 9.560440958011895e-05
step: 60, loss: 0.00011887424625456333
step: 70, loss: 3.302659388282336e-05
step: 80, loss: 6.321974069578573e-05
step: 90, loss: 4.187270678812638e-05
step: 100, loss: 3.3725620596669614e-05
step: 110, loss: 0.03273634985089302
step: 120, loss: 0.021573437377810478
step: 130, loss: 0.0001704225578578189
step: 140, loss: 4.834785431739874e-05
step: 150, loss: 8.15453749964945e-05
step: 160, loss: 0.00011445068957982585
step: 170, loss: 9.845731256064028e-05
step: 180, loss: 5.5840861023170874e-05
step: 190, loss: 4.7643981815781444e-05
step: 200, loss: 0.0001348501245956868
step: 210, loss: 6.577472959179431e-05
step: 220, loss: 4.259176421328448e-05
step: 230, loss: 4.697036638390273e-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=0.9865168539325843, f1=0.9797297297297298,
best_f1=0.9841269841269841
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 3.903513425029814e-05
step: 10, loss: 0.009226281195878983
step: 20, loss: 8.433168841293082e-05
step: 30, loss: 0.0001681896683294326
step: 40, loss: 0.0018067860510200262
step: 50, loss: 4.870898919762112e-05
step: 60, loss: 0.00022253005590755492
step: 70, loss: 5.4290092521114275e-05
step: 80, loss: 5.032485933043063e-05
step: 90, loss: 5.296550443745218e-05
step: 100, loss: 8.731109846848994e-05
step: 110, loss: 0.00014577059482689947
step: 120, loss: 2.6399991838843562e-05
step: 130, loss: 3.997820749646053e-05
step: 140, loss: 0.00021935621043667197
step: 150, loss: 5.154585960553959e-05
step: 160, loss: 0.0005315426387824118
step: 170, loss: 2.9793553039780818e-05
step: 180, loss: 0.00017173841479234397
step: 190, loss: 0.00010822412150446326
step: 200, loss: 3.91089815821033e-05
step: 210, loss: 5.824050822411664e-05
step: 220, loss: 0.0007905701640993357
step: 230, loss: 0.011172872968018055
/usr/local/lib/python3.7/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=0.9875706214689265, f1=0.9807037457434733,
best_f1=0.9841269841269841
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 7.856070442358032e-05
step: 10, loss: 2.6642073862603866e-05
step: 20, loss: 4.9006674089469016e-05
step: 30, loss: 3.500535603961907e-05
step: 40, loss: 9.050042717717588e-05
step: 50, loss: 0.05239781737327576
step: 60, loss: 6.498255243059248e-05
step: 70, loss: 4.769804945681244e-05
step: 80, loss: 8.847429126035422e-05
step: 90, loss: 4.853800783166662e-05
```

```
step: 100, loss: 3.925145210814662e-05
step: 110, loss: 0.00014071198529563844
step: 120, loss: 0.00014545275189448148
step: 130, loss: 0.013762429356575012
step: 140, loss: 3.168613329762593e-05
step: 150, loss: 7.199036190286279e-05
step: 160, loss: 2.902259075199254e-05
step: 170, loss: 3.523204577504657e-05
step: 180, loss: 0.00023342552594840527
step: 190, loss: 0.0005426799762062728
step: 200, loss: 4.94161868118681e-05
step: 210, loss: 4.851128323934972e-05
step: 220, loss: 0.00012573215644806623
step: 230, loss: 4.846827141591348e-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=0.9875706214689265, f1=0.9807037457434733,
best_f1=0.9841269841269841
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.")
2473it [00:14, 165.91it/s]
load_f1 = 0.9921259842519685
real_f1 = 0.9876819708846584
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, 191.44it/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 bert-base-uncased \
       --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 bert-base-uncased \
       --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: 0.5941812992095947
step: 10, loss: 0.5643873810768127
step: 20, loss: 0.5053332448005676
step: 30, loss: 0.1382715255022049
step: 40, loss: 0.06386411190032959
step: 50, loss: 0.2568763196468353
Gradient overflow.  Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 60, loss: 0.0614011324942112
step: 70, loss: 0.11283409595489502
step: 80, loss: 0.13586702942848206
step: 90, loss: 0.1993851512670517
step: 100, loss: 0.02375837042927742
step: 110, loss: 0.043974436819553375
step: 120, loss: 0.1050015389919281
step: 130, loss: 0.06462683528661728
```

```
step: 140, loss: 0.06457775086164474
step: 150, loss: 0.03822995349764824
step: 160, loss: 0.06171903759241104
step: 170, loss: 0.1826198697090149
step: 180, loss: 0.059611253440380096
step: 190, loss: 0.010811558924615383
step: 200, loss: 0.1327529102563858
step: 210, loss: 0.16594883799552917
step: 220, loss: 0.23945407569408417
step: 230, loss: 0.16991564631462097
step: 240, loss: 0.02302728407084942
step: 250, loss: 0.027178358286619186
step: 260, loss: 0.0667194202542305
step: 270, loss: 0.02007257379591465
step: 280, loss: 0.14153800904750824
step: 290, loss: 0.19589656591415405
step: 300, loss: 0.07285904884338379
step: 310, loss: 0.2925444543361664
step: 320, loss: 0.09666147828102112
step: 330, loss: 0.020689064636826515
step: 340, loss: 0.04177229851484299
step: 350, loss: 0.07293211668729782
step: 360, loss: 0.03356677293777466
step: 370, loss: 0.07338785380125046
step: 380, loss: 0.05319073796272278
step: 390, loss: 0.1228046715259552
step: 400, loss: 0.2532435655593872
step: 410, loss: 0.06499288976192474
step: 420, loss: 0.04461269453167915
step: 430, loss: 0.16231374442577362
step: 440, loss: 0.01692924089729786
step: 450, loss: 0.0049302284605801105
step: 460, loss: 0.009584357030689716
step: 470, loss: 0.1515865921974182
step: 480, loss: 0.05973268672823906
step: 490, loss: 0.0907684937119484
step: 500, loss: 0.19415059885707092
step: 510, loss: 0.060291845351457596
step: 520, loss: 0.13726806640625
step: 530, loss: 0.0023774015717208385
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9315448658649398, f1=0.9326568265682658,
```

```
best_f1=0.9326568265682658
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.2780977189540863
step: 10, loss: 0.08657890558242798
step: 20, loss: 0.015238344669342041
step: 30, loss: 0.006529510486871004
step: 40, loss: 0.05136450007557869
step: 50, loss: 0.10035192966461182
step: 60, loss: 0.012172658927738667
step: 70, loss: 0.06216040253639221
step: 80, loss: 0.06926295906305313
step: 90, loss: 0.014123805798590183
step: 100, loss: 0.007270304951816797
step: 110, loss: 0.047078683972358704
step: 120, loss: 0.10443152487277985
step: 130, loss: 0.19014109671115875
step: 140, loss: 0.014650438912212849
step: 150, loss: 0.03840923681855202
step: 160, loss: 0.007224952336400747
step: 170, loss: 0.04511386528611183
step: 180, loss: 0.044283196330070496
step: 190, loss: 0.07948414236307144
step: 200, loss: 0.006078329402953386
step: 210, loss: 0.11212272197008133
step: 220, loss: 0.08632946014404297
step: 230, loss: 0.011044835671782494
step: 240, loss: 0.04679754003882408
step: 250, loss: 0.10941175371408463
step: 260, loss: 0.008289448916912079
step: 270, loss: 0.06584442406892776
step: 280, loss: 0.03593801334500313
step: 290, loss: 0.015030864626169205
step: 300, loss: 0.14843736588954926
step: 310, loss: 0.010996242053806782
step: 320, loss: 0.05791084095835686
step: 330, loss: 0.07475490868091583
step: 340, loss: 0.016913369297981262
step: 350, loss: 0.0029076675418764353
step: 360, loss: 0.05816299840807915
step: 370, loss: 0.12896151840686798
step: 380, loss: 0.08881200850009918
step: 390, loss: 0.17374949157238007
step: 400, loss: 0.0611075684428215
step: 410, loss: 0.011395051144063473
step: 420, loss: 0.009089576080441475
step: 430, loss: 0.017355166375637054
```

```
step: 440, loss: 0.24719834327697754
step: 450, loss: 0.025613877922296524
step: 460, loss: 0.011727412231266499
step: 470, loss: 0.11246947944164276
step: 480, loss: 0.2906055152416229
step: 490, loss: 0.013998513109982014
step: 500, loss: 0.298051655292511
step: 510, loss: 0.015595877543091774
step: 520, loss: 0.03110715188086033
step: 530, loss: 0.007796839345246553
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9352245862884161, f1=0.9298578199052132,
best_f1=0.9298578199052132
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.07157765328884125
step: 10, loss: 0.06882112473249435
step: 20, loss: 0.06251675635576248
step: 30, loss: 0.14206433296203613
step: 40, loss: 0.004314890597015619
step: 50, loss: 0.07629803568124771
step: 60, loss: 0.05760900676250458
step: 70, loss: 0.012950544245541096
step: 80, loss: 0.0030708713456988335
step: 90, loss: 0.005837948061525822
step: 100, loss: 0.026986777782440186
step: 110, loss: 0.0008679360034875572
step: 120, loss: 0.0023735901340842247
step: 130, loss: 0.004432807210832834
step: 140, loss: 0.02385549433529377
step: 150, loss: 0.018410928547382355
step: 160, loss: 0.005132345482707024
step: 170, loss: 0.04999226704239845
step: 180, loss: 0.006798971444368362
step: 190, loss: 0.006134665105491877
step: 200, loss: 0.005482227075845003
step: 210, loss: 0.0429612472653389
step: 220, loss: 0.08177008479833603
step: 230, loss: 0.025430910289287567
step: 240, loss: 0.00464272778481245
step: 250, loss: 0.013050743378698826
```

```
step: 260, loss: 0.051194123923778534
step: 270, loss: 0.01048758253455162
step: 280, loss: 0.2109597772359848
step: 290, loss: 0.0034236120991408825
step: 300, loss: 0.11203441768884659
step: 310, loss: 0.006900348234921694
step: 320, loss: 0.026294007897377014
step: 330, loss: 0.003948966506868601
step: 340, loss: 0.016912322491407394
step: 350, loss: 0.0013569567818194628
step: 360, loss: 0.018942272290587425
step: 370, loss: 0.008030423894524574
step: 380, loss: 0.013602189719676971
step: 390, loss: 0.012945551425218582
step: 400, loss: 0.043862827122211456
step: 410, loss: 0.002217957517132163
step: 420, loss: 0.19910794496536255
step: 430, loss: 0.011816040612757206
step: 440, loss: 0.03656338155269623
step: 450, loss: 0.08612114191055298
step: 460, loss: 0.03437225520610809
step: 470, loss: 0.07602237164974213
step: 480, loss: 0.002456828486174345
step: 490, loss: 0.0038674494717270136
step: 500, loss: 0.017827309668064117
step: 510, loss: 0.0050650304183363914
step: 520, loss: 0.13993249833583832
step: 530, loss: 0.021520711481571198
/usr/local/lib/python3.7/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=0.9351251158480075, f1=0.9252206223873664,
best_f1=0.9298578199052132
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0041231997311115265
step: 10, loss: 0.0038674913812428713
step: 20, loss: 0.005117810796946287
step: 30, loss: 0.0015925210900604725
step: 40, loss: 0.0006788374157622457
step: 50, loss: 0.0027012068312615156
step: 60, loss: 0.00232420745305717
step: 70, loss: 0.00755506195127964
step: 80, loss: 0.009756827726960182
step: 90, loss: 0.008976851589977741
step: 100, loss: 0.03467829152941704
step: 110, loss: 0.061270296573638916
```

```
step: 120, loss: 0.0003092783153988421
step: 130, loss: 0.007448209915310144
step: 140, loss: 0.003762062406167388
step: 150, loss: 0.002179835457354784
step: 160, loss: 0.00325801782310009
step: 170, loss: 0.0009943465702235699
step: 180, loss: 0.008626949973404408
step: 190, loss: 0.017518023028969765
step: 200, loss: 0.000902339001186192
step: 210, loss: 0.031445298343896866
step: 220, loss: 0.007941445335745811
step: 230, loss: 0.037310682237148285
step: 240, loss: 0.0035359107423573732
step: 250, loss: 0.008290288969874382
step: 260, loss: 0.0012937899446114898
step: 270, loss: 0.008264867588877678
step: 280, loss: 0.014006241224706173
step: 290, loss: 0.048446204513311386
step: 300, loss: 0.0007556918426416814
step: 310, loss: 0.03901955857872963
step: 320, loss: 0.029797594994306564
step: 330, loss: 0.009532667696475983
step: 340, loss: 0.0023361516650766134
step: 350, loss: 0.073492631316185
step: 360, loss: 0.12254122644662857
step: 370, loss: 0.01855367049574852
step: 380, loss: 0.02060972899198532
step: 390, loss: 0.0006884007016196847
step: 400, loss: 0.0027361148968338966
step: 410, loss: 0.0027228707913309336
step: 420, loss: 0.004619908984750509
step: 430, loss: 0.1353176087141037
step: 440, loss: 0.005338752642273903
step: 450, loss: 0.0013638571836054325
step: 460, loss: 0.003381518181413412
step: 470, loss: 0.015982842072844505
step: 480, loss: 0.061667900532484055
step: 490, loss: 0.006762935314327478
step: 500, loss: 0.0039923666045069695
step: 510, loss: 0.005626735743135214
Gradient overflow.  Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 520, loss: 0.15956678986549377
step: 530, loss: 0.0168771892786026
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9393237610004631, f1=0.9346919870310328,
best_f1=0.9346919870310328
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.009611968882381916
step: 10, loss: 0.04045988991856575
step: 20, loss: 0.0726461261510849
step: 30, loss: 0.0010131775634363294
step: 40, loss: 0.0030772974714636803
step: 50, loss: 0.0024911765940487385
step: 60, loss: 0.1097845733165741
step: 70, loss: 0.003279578872025013
step: 80, loss: 0.005363688338547945
step: 90, loss: 0.008700139820575714
step: 100, loss: 0.026296114549040794
step: 110, loss: 0.00027598056476563215
step: 120, loss: 0.02463817410171032
step: 130, loss: 0.00042040928383357823
step: 140, loss: 0.0007551597664132714
step: 150, loss: 0.0003391029022168368
step: 160, loss: 0.00028497533639892936
step: 170, loss: 0.027159474790096283
step: 180, loss: 0.00039516930701211095
step: 190, loss: 0.0003902839380316436
step: 200, loss: 0.0004607538867276162
step: 210, loss: 0.013660683296620846
step: 220, loss: 0.001550542307086289
step: 230, loss: 0.0021565838251262903
step: 240, loss: 0.004057195968925953
step: 250, loss: 0.0017805217066779733
step: 260, loss: 0.03309642896056175
step: 270, loss: 0.0014873897889629006
step: 280, loss: 0.01941286399960518
step: 290, loss: 0.001888848259113729
step: 300, loss: 0.015164827927947044
step: 310, loss: 0.0030661076307296753
step: 320, loss: 0.12612655758857727
step: 330, loss: 0.05286338925361633
step: 340, loss: 0.0011349754640832543
step: 350, loss: 0.0011508592870086432
step: 360, loss: 0.025486718863248825
step: 370, loss: 0.006516643334180117
step: 380, loss: 0.0014617926208302379
step: 390, loss: 0.00020764171495102346
step: 400, loss: 0.0038392499554902315
```

```
step: 410, loss: 0.004825236275792122
step: 420, loss: 0.0036449439357966185
step: 430, loss: 0.006148270331323147
step: 440, loss: 0.004137112759053707
step: 450, loss: 0.0018380648689344525
step: 460, loss: 0.0003066406352445483
step: 470, loss: 0.0066450415179133415
step: 480, loss: 0.010101149789988995
step: 490, loss: 0.00016616724315099418
step: 500, loss: 0.0021800866816192865
step: 510, loss: 0.0036731434520334005
step: 520, loss: 0.07624439150094986
step: 530, loss: 0.00458444794639945
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9423697556477639, f1=0.9402573529411765,
best_f1=0.9402573529411765
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.014047208242118359
step: 10, loss: 0.003837019670754671
step: 20, loss: 0.05430829897522926
step: 30, loss: 0.00557180680334568
step: 40, loss: 0.0010023800423368812
step: 50, loss: 0.1482698768377304
step: 60, loss: 0.0001587003789609298
step: 70, loss: 0.006814118940383196
step: 80, loss: 0.0013915977906435728
step: 90, loss: 0.00016132215387187898
step: 100, loss: 0.028105566278100014
step: 110, loss: 0.011489455588161945
step: 120, loss: 0.000732011569198221
step: 130, loss: 0.001256960560567677
step: 140, loss: 0.005761244334280491
step: 150, loss: 0.0023682666942477226
step: 160, loss: 0.002510342514142394
step: 170, loss: 0.000766777026001364
step: 180, loss: 0.0004058222402818501
step: 190, loss: 0.00013468334509525448
step: 200, loss: 0.003750129835680127
step: 210, loss: 0.0017395186005160213
step: 220, loss: 0.0034353581722825766
```

```
step: 230, loss: 0.0002871159522328526
step: 240, loss: 0.018229762092232704
step: 250, loss: 0.0007393336854875088
step: 260, loss: 0.00013179791858419776
step: 270, loss: 0.09133229404687881
step: 280, loss: 0.0019574405159801245
step: 290, loss: 0.00036437882226891816
step: 300, loss: 0.00016136294289026416
step: 310, loss: 0.0009770046453922987
step: 320, loss: 0.0007078393246047199
step: 330, loss: 0.000421685486799106
step: 340, loss: 0.14815236628055573
step: 350, loss: 0.0015005464665591717
step: 360, loss: 0.0037887776270508766
step: 370, loss: 0.0009495826088823378
step: 380, loss: 0.00013805902563035488
step: 390, loss: 0.016013724729418755
step: 400, loss: 0.00011727940000128001
step: 410, loss: 0.0001994982158066705
step: 420, loss: 0.0016280082054436207
step: 430, loss: 0.000627703033387661
step: 440, loss: 0.00020124050206504762
step: 450, loss: 0.0003740485990419984
step: 460, loss: 0.00014622653543483466
step: 470, loss: 0.002947507193312049
step: 480, loss: 0.004941871855407953
step: 490, loss: 0.007671453524380922
step: 500, loss: 0.0011141039431095123
step: 510, loss: 0.002371882786974311
step: 520, loss: 0.0006616150494664907
step: 530, loss: 0.006080096587538719
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9434137291280148, f1=0.9382830626450116,
best_f1=0.9382830626450116
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0008506195736117661
step: 10, loss: 0.0003008809871971607
step: 20, loss: 0.0005879531381651759
step: 30, loss: 0.0008225663914345205
step: 40, loss: 8.762008656049147e-05
```

```
step: 50, loss: 0.00554334931075573
step: 60, loss: 0.0004035922174807638
step: 70, loss: 0.02476579137146473
step: 80, loss: 0.0007362254546023905
step: 90, loss: 0.00021581596229225397
step: 100, loss: 0.00010617219231789932
step: 110, loss: 0.0008416488999500871
step: 120, loss: 0.007578074932098389
step: 130, loss: 0.0016807059291750193
step: 140, loss: 0.001094402396120131
step: 150, loss: 0.0003009513020515442
step: 160, loss: 0.0005435283528640866
step: 170, loss: 0.00032041672966443
step: 180, loss: 0.0001019629489746876
step: 190, loss: 0.01004644576460123
step: 200, loss: 0.00051596894627437
step: 210, loss: 0.002079906640574336
step: 220, loss: 0.0001637922105146572
step: 230, loss: 0.17489507794380188
step: 240, loss: 0.0005520256818272173
step: 250, loss: 0.00026371682179160416
step: 260, loss: 4.5100474380888045e-05
step: 270, loss: 0.00018128432566300035
step: 280, loss: 0.0004016971797682345
step: 290, loss: 0.0004019980551674962
step: 300, loss: 0.003789068665355444
step: 310, loss: 0.00013748722267337143
step: 320, loss: 0.00013703620061278343
step: 330, loss: 0.00032082010875456035
step: 340, loss: 0.01677740551531315
step: 350, loss: 0.008199213072657585
step: 360, loss: 0.011951806955039501
step: 370, loss: 0.00011540594277903438
step: 380, loss: 0.0009770442266017199
step: 390, loss: 0.00016026897355914116
step: 400, loss: 0.00115681323222816
step: 410, loss: 0.00019969405548181385
step: 420, loss: 0.0002456358342897147
step: 430, loss: 9.120434697251767e-05
step: 440, loss: 0.0007782129687257111
step: 450, loss: 0.00036265328526496887
step: 460, loss: 0.0005901945987716317
step: 470, loss: 0.002805457217618823
step: 480, loss: 0.16708363592624664
step: 490, loss: 0.00013869033136870712
step: 500, loss: 0.0011876319767907262
step: 510, loss: 0.0007865465013310313
step: 520, loss: 0.0007563892286270857
```

```
step: 530, loss: 0.02952703647315502
/usr/local/lib/python3.7/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=0.9388707419505367, f1=0.9328984156570364,
best_f1=0.9382830626450116
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00019449747924227268
step: 10, loss: 0.028702594339847565
step: 20, loss: 0.00022207100118976086
step: 30, loss: 0.0001581020187586546
step: 40, loss: 0.0002351687871851027
step: 50, loss: 0.0009815540397539735
step: 60, loss: 0.0006588065298274159
step: 70, loss: 0.00010843015479622409
step: 80, loss: 0.004151594825088978
step: 90, loss: 0.00015647959662601352
step: 100, loss: 9.023993334267288e-05
step: 110, loss: 0.0010543884709477425
step: 120, loss: 0.0026573038194328547
step: 130, loss: 0.00012743133993353695
step: 140, loss: 0.0008106849272735417
step: 150, loss: 0.0002230617101304233
step: 160, loss: 0.016147928312420845
step: 170, loss: 0.016077116131782532
step: 180, loss: 0.0012262981617823243
step: 190, loss: 0.007712661288678646
step: 200, loss: 0.000547017902135849
step: 210, loss: 0.00572526641190052
step: 220, loss: 0.0003017602430190891
step: 230, loss: 5.2952189435018227e-05
step: 240, loss: 0.00017343691433779895
step: 250, loss: 6.333467172225937e-05
step: 260, loss: 6.765213038306683e-05
step: 270, loss: 0.0011983970180153847
step: 280, loss: 0.029498260468244553
step: 290, loss: 9.64468126767315e-05
step: 300, loss: 0.025754446163773537
step: 310, loss: 0.00047740156878717244
step: 320, loss: 0.003470911178737879
step: 330, loss: 0.00029126217123121023
step: 340, loss: 0.0003149854310322553
step: 350, loss: 0.01268818974494934
step: 360, loss: 0.00019978363707195967
step: 370, loss: 0.0005548756453208625
step: 380, loss: 0.002191924722865224
```

```
step: 390, loss: 0.14698758721351624
step: 400, loss: 7.194244972197339e-05
step: 410, loss: 0.0001094574254238978
step: 420, loss: 0.0016865466022863984
step: 430, loss: 0.05716206878423691
step: 440, loss: 0.0031389615032821894
step: 450, loss: 0.00014494560309685767
step: 460, loss: 0.0007080926443450153
step: 470, loss: 7.646766607649624e-05
step: 480, loss: 4.620238905772567e-05
step: 490, loss: 0.0021018204279243946
step: 500, loss: 0.0017818756168708205
step: 510, loss: 0.00010103568638442084
step: 520, loss: 0.0004557289066724479
step: 530, loss: 0.00017032174218911678
/usr/local/lib/python3.7/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=0.941340782122905, f1=0.9360902255639098,
best_f1=0.9382830626450116
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.000455373025033623
step: 10, loss: 8.95665361895226e-05
step: 20, loss: 0.012045766226947308
step: 30, loss: 0.001182005275040865
step: 40, loss: 0.00015408832405228168
step: 50, loss: 0.00011625871411524713
step: 60, loss: 4.7956906200852245e-05
step: 70, loss: 0.07553111016750336
step: 80, loss: 0.0029340507462620735
step: 90, loss: 0.00012468238128349185
step: 100, loss: 5.7779328926699236e-05
step: 110, loss: 6.0101716371718794e-05
step: 120, loss: 5.5395765230059624e-05
step: 130, loss: 0.013239657506346703
step: 140, loss: 0.010860650800168514
step: 150, loss: 4.0899922169046476e-05
step: 160, loss: 0.003669860539957881
step: 170, loss: 0.010667390190064907
step: 180, loss: 5.084967051516287e-05
step: 190, loss: 6.301762186922133e-05
step: 200, loss: 0.001803673687390983
step: 210, loss: 4.090848597115837e-05
step: 220, loss: 8.812213491182774e-05
step: 230, loss: 7.133622420951724e-05
step: 240, loss: 0.00010319965076632798
```

```
step: 250, loss: 7.385495700873435e-05
step: 260, loss: 0.0019582565873861313
step: 270, loss: 3.355508670210838e-05
step: 280, loss: 0.00032043427927419543
step: 290, loss: 0.03107600100338459
step: 300, loss: 7.081685907905921e-05
step: 310, loss: 0.02356523461639881
step: 320, loss: 7.716578693361953e-05
step: 330, loss: 0.00011655578418867663
step: 340, loss: 3.5875003959517926e-05
step: 350, loss: 0.007955399341881275
step: 360, loss: 2.9685968911508098e-05
step: 370, loss: 0.0032036560587584972
step: 380, loss: 5.200956002227031e-05
step: 390, loss: 3.530990579747595e-05
step: 400, loss: 0.00016505367239005864
step: 410, loss: 7.908752741059288e-05
step: 420, loss: 0.006729843094944954
step: 430, loss: 8.44625974423252e-05
step: 440, loss: 0.0009811274940147996
step: 450, loss: 0.00010374156408943236
step: 460, loss: 0.00041291210800409317
step: 470, loss: 2.40127610595664e-05
step: 480, loss: 4.429062028066255e-05
step: 490, loss: 0.0010694322409108281
step: 500, loss: 0.0005275965668261051
step: 510, loss: 0.0001608380989637226
step: 520, loss: 0.0004907906986773014
step: 530, loss: 0.0017802906222641468
/usr/local/lib/python3.7/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=0.9384902143522833, f1=0.9311475409836065,
best_f1=0.9382830626450116
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0013351424131542444
step: 10, loss: 2.6791092750499956e-05
step: 20, loss: 0.0007586884312331676
step: 30, loss: 9.601030615158379e-05
step: 40, loss: 0.0009322646073997021
step: 50, loss: 0.0002265843068016693
step: 60, loss: 0.0009182404610328376
step: 70, loss: 0.009744195267558098
step: 80, loss: 0.0011757529573515058
step: 90, loss: 0.00035360263427719474
step: 100, loss: 4.814880958292633e-05
```

```
step: 110, loss: 2.0764391592820175e-05
step: 120, loss: 0.0002061252889689058
step: 130, loss: 0.0016081412322819233
step: 140, loss: 0.007163052447140217
step: 150, loss: 9.66408260865137e-05
step: 160, loss: 2.813590799632948e-05
step: 170, loss: 6.731161556672305e-05
step: 180, loss: 9.320356912212446e-05
step: 190, loss: 0.0391877256333828
step: 200, loss: 0.007419873960316181
step: 210, loss: 0.0003570143599063158
step: 220, loss: 0.00046635387116111815
step: 230, loss: 0.00104476825799793
step: 240, loss: 0.00035368872340768576
step: 250, loss: 2.5916167942341417e-05
step: 260, loss: 0.00018000764248427004
step: 270, loss: 6.691200542263687e-05
step: 280, loss: 0.0003094544226769358
step: 290, loss: 3.795038355747238e-05
step: 300, loss: 3.0419603717746213e-05
step: 310, loss: 0.0022534485906362534
step: 320, loss: 0.0007959509966894984
step: 330, loss: 0.00043725661817006767
step: 340, loss: 0.00047804333735257387
step: 350, loss: 0.04786144196987152
step: 360, loss: 4.831310798181221e-05
step: 370, loss: 0.0016887669917196035
step: 380, loss: 0.00021856777311768383
step: 390, loss: 0.0005432229372672737
step: 400, loss: 0.0006937248399481177
step: 410, loss: 0.00042424455750733614
step: 420, loss: 0.00010015265670605
step: 430, loss: 0.00029115588404238224
step: 440, loss: 0.006133111659437418
step: 450, loss: 3.402869697310962e-05
step: 460, loss: 3.6413181078387424e-05
step: 470, loss: 5.233680712990463e-05
step: 480, loss: 0.000781570328399539
step: 490, loss: 0.0002310015115654096
step: 500, loss: 0.006461269687861204
step: 510, loss: 0.0009308846201747656
step: 520, loss: 4.441849887371063e-05
step: 530, loss: 4.4498301576822996e-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=0.9406264609630668, f1=0.9339578454332552,
best_f1=0.9382830626450116
```

```
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.000314731034450233
step: 10, loss: 0.00034799828426912427
step: 20, loss: 4.006343078799546e-05
step: 30, loss: 2.7983573090750724e-05
step: 40, loss: 0.0005905182915739715
step: 50, loss: 8.725409134058282e-05
step: 60, loss: 0.031125372275710106
step: 70, loss: 2.8795728212571703e-05
step: 80, loss: 7.234921213239431e-05
step: 90, loss: 3.050903251278214e-05
step: 100, loss: 0.00038512342143803835
step: 110, loss: 2.9857434128643945e-05
Gradient overflow.  Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 120, loss: 3.146062954328954e-05
step: 130, loss: 1.984409573196899e-05
step: 140, loss: 0.17119383811950684
step: 150, loss: 1.656610766076483e-05
step: 160, loss: 0.09015153348445892
step: 170, loss: 0.00020546669838950038
step: 180, loss: 0.0006367875612340868
step: 190, loss: 0.006387187168002129
step: 200, loss: 0.007533607073128223
step: 210, loss: 0.00045875422074459493
step: 220, loss: 0.003474452067166567
step: 230, loss: 0.00010994142212439328
step: 240, loss: 5.811066512251273e-05
step: 250, loss: 2.0816240066778846e-05
step: 260, loss: 8.770735439611599e-05
step: 270, loss: 8.974088996183127e-05
step: 280, loss: 0.0001896398316603154
step: 290, loss: 0.0036021603737026453
step: 300, loss: 0.002300660591572523
step: 310, loss: 0.00042168295476585627
step: 320, loss: 0.003634663298726082
step: 330, loss: 0.0017638663994148374
step: 340, loss: 0.0003379739064257592
step: 350, loss: 0.0006439135177060962
step: 360, loss: 0.00019127241102978587
step: 370, loss: 0.0001902558869915083
step: 380, loss: 5.248470188234933e-05
step: 390, loss: 0.0007128320867195725
step: 400, loss: 0.0008334564045071602
step: 410, loss: 6.16836769040674e-05
step: 420, loss: 0.004146751016378403
step: 430, loss: 9.314445924246684e-05
```

```
step: 440, loss: 0.00030462394352070987
step: 450, loss: 7.98014661995694e-05
step: 460, loss: 0.0009914366528391838
step: 470, loss: 0.0009585041552782059
step: 480, loss: 0.0001914559688884765
step: 490, loss: 0.00019035476725548506
step: 500, loss: 8.650852396385744e-05
step: 510, loss: 0.00017522698908578604
step: 520, loss: 0.00020599788695108145
step: 530, loss: 4.190931576886214e-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 11: dev_f1=0.9408424041646947, f1=0.9318723201524535,
best_f1=0.9382830626450116
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 6.418035627575591e-05
step: 10, loss: 0.0003571631386876106
step: 20, loss: 2.812388993334025e-05
step: 30, loss: 0.0008414312615059316
step: 40, loss: 0.00019293249351903796
step: 50, loss: 0.08740647882223129
step: 60, loss: 0.0017405145335942507
step: 70, loss: 0.00167789519764482983
step: 80, loss: 2.2902546334080398e-05
step: 90, loss: 6.790741463191807e-05
step: 100, loss: 0.0005022193654440343
step: 110, loss: 0.003400301793590188
step: 120, loss: 2.9934619306004606e-05
step: 130, loss: 0.00047900358913466334
step: 140, loss: 4.7019755584187806e-05
step: 150, loss: 3.318962626508437e-05
step: 160, loss: 3.207089321222156e-05
step: 170, loss: 6.056451456970535e-05
step: 180, loss: 2.495833905413747e-05
step: 190, loss: 0.000182046351255849
step: 200, loss: 3.0613529816037044e-05
step: 210, loss: 7.845721120247617e-05
step: 220, loss: 0.003796194214373827
step: 230, loss: 3.6402783734956756e-05
step: 240, loss: 0.0001221836282638833
step: 250, loss: 1.9263088688603602e-05
step: 260, loss: 3.63726940122433e-05
step: 270, loss: 5.173025783733465e-05
step: 280, loss: 0.0018445936730131507
step: 290, loss: 9.796566155273467e-05
```

```
step: 300, loss: 0.00011559268750716001
step: 310, loss: 3.870600266964175e-05
step: 320, loss: 2.3706876163487323e-05
step: 330, loss: 3.431574077694677e-05
step: 340, loss: 7.308595377253368e-05
step: 350, loss: 4.816855289391242e-05
step: 360, loss: 0.000871892727445811
step: 370, loss: 0.0003313029301352799
step: 380, loss: 2.055953154922463e-05
step: 390, loss: 9.237902850145474e-05
step: 400, loss: 0.003909861668944359
step: 410, loss: 0.00044074462493881583
step: 420, loss: 0.000650393427349627
step: 430, loss: 0.0055122836492955685
step: 440, loss: 2.8984552045585588e-05
step: 450, loss: 0.00014830034342594445
step: 460, loss: 3.479349106783047e-05
step: 470, loss: 6.014843165758066e-05
step: 480, loss: 0.0003691283636726439
step: 490, loss: 0.00135265092831105
step: 500, loss: 2.199720620410517e-05
step: 510, loss: 0.00041157047962769866
step: 520, loss: 0.00027459609555080533
step: 530, loss: 6.166471575852484e-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=0.9457943925233645, f1=0.9359513791491351,
best_f1=0.9359513791491351
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00012900344154331833
step: 10, loss: 4.346905916463584e-05
step: 20, loss: 1.339601385552669e-05
step: 30, loss: 2.765560020634439e-05
step: 40, loss: 2.1487005142262205e-05
step: 50, loss: 4.2102816223632544e-05
step: 60, loss: 3.024725810973905e-05
step: 70, loss: 4.0628365240991116e-05
step: 80, loss: 2.0026707716169767e-05
step: 90, loss: 0.023641474545001984
step: 100, loss: 0.000649342080578208
step: 110, loss: 0.004033796023577452
```

```
step: 120, loss: 0.000290151423541829
step: 130, loss: 5.6896558817243204e-05
step: 140, loss: 1.813446215237491e-05
step: 150, loss: 5.246836008154787e-05
step: 160, loss: 1.848085412348155e-05
step: 170, loss: 2.096929711115081e-05
step: 180, loss: 3.135798397124745e-05
step: 190, loss: 4.355091732577421e-05
step: 200, loss: 7.868644024711102e-05
step: 210, loss: 1.989634802157525e-05
step: 220, loss: 3.7755573430331424e-05
step: 230, loss: 1.883842924144119e-05
step: 240, loss: 2.3096148652257398e-05
step: 250, loss: 2.000812310143374e-05
step: 260, loss: 1.4271410691435449e-05
step: 270, loss: 1.3664233847521245e-05
step: 280, loss: 3.5132568882545456e-05
step: 290, loss: 0.00034402761957608163
step: 300, loss: 4.5749780838377774e-05
step: 310, loss: 0.00018922869639936835
step: 320, loss: 4.795884160557762e-05
step: 330, loss: 2.933900395873934e-05
step: 340, loss: 2.7606500225374475e-05
step: 350, loss: 2.0451176169444807e-05
step: 360, loss: 1.8614977307152003e-05
step: 370, loss: 6.23813975835219e-05
step: 380, loss: 0.00020750105613842607
step: 390, loss: 0.003016574075445533
step: 400, loss: 9.579947800375521e-05
step: 410, loss: 0.00030804661219008267
step: 420, loss: 2.0306139049353078e-05
step: 430, loss: 0.0011428085854277015
step: 440, loss: 4.801710383617319e-05
step: 450, loss: 6.050852607586421e-05
step: 460, loss: 9.797484381124377e-05
step: 470, loss: 4.230225022183731e-05
step: 480, loss: 2.1468305931193754e-05
step: 490, loss: 3.568506144802086e-05
step: 500, loss: 2.0685338313342072e-05
step: 510, loss: 0.0014836761401966214
step: 520, loss: 0.0004534600884653628
step: 530, loss: 1.5493298633373342e-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=0.9402843601895735, f1=0.9292161520190024,
best_f1=0.9359513791491351
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
```

```
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 9.398618567502126e-05
step: 10, loss: 1.64915290955219448e-05
step: 20, loss: 2.1855716113350354e-05
step: 30, loss: 2.8813146855100058e-05
step: 40, loss: 2.9692730095121078e-05
step: 50, loss: 0.03821853548288345
step: 60, loss: 1.4170833310345188e-05
step: 70, loss: 0.00010903707152465358
step: 80, loss: 3.103692506556399e-05
step: 90, loss: 0.0003631806466728449
step: 100, loss: 3.721453686011955e-05
step: 110, loss: 0.00016572137246839702
step: 120, loss: 0.00176671810913831
step: 130, loss: 0.026514139026403427
step: 140, loss: 4.313373574404977e-05
step: 150, loss: 2.620828126964625e-05
step: 160, loss: 1.6551224689465016e-05
step: 170, loss: 0.0006833224906586111
step: 180, loss: 1.7456381101510487e-05
step: 190, loss: 2.7448761102277786e-05
step: 200, loss: 0.002730715088546276
step: 210, loss: 6.107822991907597e-05
step: 220, loss: 2.122996374964714e-05
step: 230, loss: 2.8661203032243066e-05
step: 240, loss: 1.8529202861827798e-05
step: 250, loss: 5.1983570301672444e-05
step: 260, loss: 5.9004036302212626e-05
step: 270, loss: 2.2500020349980332e-05
step: 280, loss: 1.85628050530795e-05
step: 290, loss: 6.792823387069464e-05
step: 300, loss: 2.810204568959307e-05
step: 310, loss: 0.0009059629519470036
step: 320, loss: 0.00010739979916252196
step: 330, loss: 2.8530615963973105e-05
step: 340, loss: 2.2474145225714892e-05
step: 350, loss: 1.8536777133704163e-05
step: 360, loss: 8.554366650059819e-05
step: 370, loss: 1.973248799913563e-05
step: 380, loss: 3.9248232496902347e-05
step: 390, loss: 0.03034222312271595
step: 400, loss: 4.1998628148576245e-05
step: 410, loss: 1.716953920549713e-05
step: 420, loss: 8.574388630222529e-05
step: 430, loss: 3.9186328649520874e-05
step: 440, loss: 2.8865833883173764e-05
step: 450, loss: 0.0024288841523230076
```

```
step: 460, loss: 2.4053191737039015e-05
step: 470, loss: 2.320379644515924e-05
step: 480, loss: 3.0450877602561377e-05
step: 490, loss: 2.336040597583633e-05
step: 500, loss: 1.3057008800387848e-05
step: 510, loss: 0.0002679030003491789
step: 520, loss: 1.6674135622452013e-05
step: 530, loss: 4.5110955397831276e-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=0.9355900329102022, f1=0.9318288669487541,
best_f1=0.9359513791491351
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.11432890594005585
step: 10, loss: 2.2429350792663172e-05
step: 20, loss: 2.6171775971306488e-05
step: 30, loss: 0.00012500063166953623
step: 40, loss: 0.04502024129033089
step: 50, loss: 7.378914597211406e-05
step: 60, loss: 0.0014072456397116184
step: 70, loss: 8.950468327384442e-05
step: 80, loss: 4.435061782714911e-05
step: 90, loss: 2.015350037254393e-05
step: 100, loss: 3.486867717583664e-05
step: 110, loss: 8.684057684149593e-05
step: 120, loss: 0.0006673400639556348
step: 130, loss: 0.07446606457233429
step: 140, loss: 1.8406115486868657e-05
step: 150, loss: 2.0421717636054382e-05
step: 160, loss: 1.850312401074916e-05
step: 170, loss: 2.0388009943417273e-05
step: 180, loss: 1.6342604794772342e-05
step: 190, loss: 4.424117287271656e-05
step: 200, loss: 1.8879318304243498e-05
step: 210, loss: 8.077365055214614e-05
step: 220, loss: 1.3518964806280565e-05
step: 230, loss: 1.3924995073466562e-05
step: 240, loss: 2.639192098286003e-05
step: 250, loss: 1.668531695031561e-05
step: 260, loss: 1.72477539308602e-05
step: 270, loss: 1.748253453115467e-05
step: 280, loss: 2.0976671294192784e-05
step: 290, loss: 1.4580609786207788e-05
step: 300, loss: 5.52138990315143e-05
step: 310, loss: 0.00016301081632263958
```

```
Gradient overflow.  Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 320, loss: 9.275582124246284e-05
step: 330, loss: 6.919995212228969e-05
step: 340, loss: 1.614887332834769e-05
step: 350, loss: 1.2270999832253437e-05
step: 360, loss: 3.994162398157641e-05
step: 370, loss: 1.410006734658964e-05
step: 380, loss: 1.6417157894466072e-05
step: 390, loss: 3.386054595466703e-05
step: 400, loss: 0.00012319588859099895
step: 410, loss: 2.0928086087224074e-05
step: 420, loss: 3.128315438516438e-05
step: 430, loss: 4.725227699964307e-05
step: 440, loss: 3.4920954931294546e-05
step: 450, loss: 0.0024981950409710407
step: 460, loss: 1.4979186744312756e-05
step: 470, loss: 0.005371051840484142
step: 480, loss: 1.3228371244622394e-05
step: 490, loss: 1.6268129911622964e-05
step: 500, loss: 2.994173701154068e-05
step: 510, loss: 1.5578958482365124e-05
step: 520, loss: 2.0290912289055996e-05
step: 530, loss: 1.8030066712526605e-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=0.9380863039399624, f1=0.9360902255639098,
best_f1=0.9359513791491351
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 [00:28, 200.43it/s]
load_f1 = 0.9377901578458682
real_f1 = 0.935813953488372
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, 193.33it/s]
```

3.1.4 Amazon-Google - Running the matcher

```python
[ ]: # 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 bert-base-uncased \
       --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 bert-base-uncased \
    --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: 0.572301983833313
step: 10, loss: 0.3941054940223694
step: 20, loss: 0.3654656410217285
step: 30, loss: 0.299784779548645
step: 40, loss: 0.16337235271930695

```
step: 50, loss: 0.4199814796447754
step: 60, loss: 0.24419060349464417
step: 70, loss: 0.15371102094650269
step: 80, loss: 0.23000980913639069
step: 90, loss: 0.2820333242416382
step: 100, loss: 0.3874475955963135
step: 110, loss: 0.25426873564720154
step: 120, loss: 0.28519344329833984
step: 130, loss: 0.20220158994197845
step: 140, loss: 0.13728071749210358
step: 150, loss: 0.13825064897537231
step: 160, loss: 0.2566460072994232
step: 170, loss: 0.29734113812446594
step: 180, loss: 0.07367555797100067
step: 190, loss: 0.1728423684835434
step: 200, loss: 0.21142742037773132
step: 210, loss: 0.2368597835302353
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.6289308176100629, f1=0.6859504132231405,
best_f1=0.6859504132231405
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.07640405744314194
step: 10, loss: 0.1703793704509735
step: 20, loss: 0.1745409220457077
step: 30, loss: 0.1668172925710678
step: 40, loss: 0.22956176102161407
step: 50, loss: 0.16090832650661469
step: 60, loss: 0.3464376628398895
step: 70, loss: 0.09285777807235718
step: 80, loss: 0.1406853049993515
step: 90, loss: 0.062974050640105988
step: 100, loss: 0.011090509593486786
step: 110, loss: 0.17843160033226013
step: 120, loss: 0.16283495724201202
step: 130, loss: 0.013889086432754993
step: 140, loss: 0.18979990482330322
step: 150, loss: 0.16619355976581573
step: 160, loss: 0.14498940110206604
step: 170, loss: 0.1104385107755661
step: 180, loss: 0.23433561623096466
```

```
step: 190, loss: 0.2016455978155136
step: 200, loss: 0.07181383669376373
step: 210, loss: 0.1937071681022644
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.6561886051080551, f1=0.7054263565891473,
best_f1=0.7054263565891473
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0904419794678688
step: 10, loss: 0.08893804252147675
step: 20, loss: 0.04314545914530754
step: 30, loss: 0.14380747079849243
step: 40, loss: 0.058134738355875015
step: 50, loss: 0.04963115230202675
step: 60, loss: 0.10345923900604248
step: 70, loss: 0.0508284829556942
step: 80, loss: 0.1720404028892517
step: 90, loss: 0.06239734962582588
step: 100, loss: 0.18835270404815674
step: 110, loss: 0.11292258650064468
step: 120, loss: 0.09297993779182434
step: 130, loss: 0.127125546336174
step: 140, loss: 0.08184226602315903
step: 150, loss: 0.19515156745910645
step: 160, loss: 0.017924867570400238
step: 170, loss: 0.06026681140065193
step: 180, loss: 0.07014056295156479
step: 190, loss: 0.23061585426330566
step: 200, loss: 0.036090172827243805
step: 210, loss: 0.11473441869020462
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.6639175257731958, f1=0.6805845511482255,
best_f1=0.6805845511482255
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
```

```
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.09613039344549179
step: 10, loss: 0.04047086834907532
step: 20, loss: 0.264282763004303
step: 30, loss: 0.03992411494255066
step: 40, loss: 0.02423546090722084
step: 50, loss: 0.192145437002182
step: 60, loss: 0.366325706243515
step: 70, loss: 0.12497251480817795
step: 80, loss: 0.07381749153137207
step: 90, loss: 0.01807868853211403
step: 100, loss: 0.10294680297374725
step: 110, loss: 0.1210656389594078
step: 120, loss: 0.09058862924575806
step: 130, loss: 0.1163761243224144
step: 140, loss: 0.09339318424463272
step: 150, loss: 0.01910310797393322
step: 160, loss: 0.04744334891438484
step: 170, loss: 0.16692176461219788
step: 180, loss: 0.2286226749420166
step: 190, loss: 0.03397594019770622
step: 200, loss: 0.21535718441009521
step: 210, loss: 0.12660491466522217
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.6822033898305084, f1=0.6906779661016949,
best_f1=0.6906779661016949
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.2409161478281021
step: 10, loss: 0.09312979876995087
step: 20, loss: 0.12772473692893982
step: 30, loss: 0.06881605088710785
step: 40, loss: 0.03193971514701843
step: 50, loss: 0.010953844524919987
step: 60, loss: 0.07795514911413193
step: 70, loss: 0.029269492253661156
step: 80, loss: 0.03813057020306587
step: 90, loss: 0.036703743040561676
step: 100, loss: 0.00836554728448391
step: 110, loss: 0.1549123376607895
step: 120, loss: 0.12257692962884903
```

```
step: 130, loss: 0.06790617853403091
step: 140, loss: 0.04927223175764084
step: 150, loss: 0.028469009324908257
step: 160, loss: 0.05608278885483742
step: 170, loss: 0.0556715726852417
step: 180, loss: 0.09934604167938232
step: 190, loss: 0.002820629859343171
step: 200, loss: 0.08378953486680984
step: 210, loss: 0.01250764261931181
/usr/local/lib/python3.7/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=0.673773987206823, f1=0.6960167714884696,
best_f1=0.6906779661016949
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.028826449066400528
step: 10, loss: 0.027839146554470062
step: 20, loss: 0.01938534900546074
step: 30, loss: 0.0003175801830366254
step: 40, loss: 0.05317449942231178
step: 50, loss: 0.0036123956087976694
step: 60, loss: 0.07237084209918976
step: 70, loss: 0.051917415112257004
step: 80, loss: 0.2146013081073761
step: 90, loss: 0.08238024264574051
step: 100, loss: 0.006208932958543301
step: 110, loss: 0.023129906505346298
step: 120, loss: 0.05846504494547844
step: 130, loss: 0.05813491716980934
step: 140, loss: 0.047490011900663376
step: 150, loss: 0.008781298995018005
step: 160, loss: 0.007454437669366598
step: 170, loss: 0.043963391333818436
step: 180, loss: 0.019166886806488037
step: 190, loss: 0.062043558806180954
step: 200, loss: 0.0013538019265979528
step: 210, loss: 0.03331022337079048
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.7042253521126761, f1=0.6858237547892722,
best_f1=0.6858237547892722
```

```
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0048814266920089972
step: 10, loss: 0.01857932284474373
step: 20, loss: 0.031336553394794464
step: 30, loss: 0.023132100701332092
step: 40, loss: 0.0029299480374902487
step: 50, loss: 0.055837467312812805
step: 60, loss: 0.03590191528201103
step: 70, loss: 0.017007695510983467
step: 80, loss: 0.03929060325026512
step: 90, loss: 0.01602630689740181
step: 100, loss: 0.00038122956175357103
step: 110, loss: 0.09292154759168625
Gradient overflow.  Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 120, loss: 0.06733439117670059
step: 130, loss: 0.013500734232366085
step: 140, loss: 0.007962401024997234
step: 150, loss: 0.04461180046200752
step: 160, loss: 0.05218010023236275
step: 170, loss: 0.002356053562834859
step: 180, loss: 0.07687848061323166
step: 190, loss: 0.036777373403310776
step: 200, loss: 0.0027891923673450947
step: 210, loss: 0.04181869700551033
/usr/local/lib/python3.7/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=0.683206106870229, f1=0.6880907372400757,
best_f1=0.6858237547892722
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.052142709493637085
step: 10, loss: 0.046751800924539566
step: 20, loss: 0.04486225172877312
step: 30, loss: 0.07829297333955765
step: 40, loss: 0.003583156503736973
step: 50, loss: 0.0009121280163526535
step: 60, loss: 0.04178067296743393
step: 70, loss: 0.0030922635924071074
step: 80, loss: 0.029476746916770935
step: 90, loss: 0.058098189532756805
step: 100, loss: 0.0011958916438743472
step: 110, loss: 0.08829181641340256
step: 120, loss: 0.0005749156698584557
step: 130, loss: 0.005619400180876255
```

```
step: 140, loss: 0.05100309103727341
step: 150, loss: 0.006666963454335928
step: 160, loss: 0.0465724803507328
step: 170, loss: 0.0002868187439162284
step: 180, loss: 0.03494931384921074
step: 190, loss: 0.002419155091047287
step: 200, loss: 0.00030142979812808335
step: 210, loss: 0.11991662532091141
/usr/local/lib/python3.7/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=0.6929460580912863, f1=0.6981519507186859,
best_f1=0.6858237547892722
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.01671033725142479
step: 10, loss: 0.009868151508271694
step: 20, loss: 0.0005503348656930029
step: 30, loss: 0.004136851988732815
step: 40, loss: 0.0018176076700910926
step: 50, loss: 0.0012315837666392326
step: 60, loss: 0.18269602954387665
step: 70, loss: 0.022534994408488274
step: 80, loss: 0.0007177735678851604
step: 90, loss: 0.003470040624961257
step: 100, loss: 0.05619870126247406
step: 110, loss: 0.03247244656085968
step: 120, loss: 0.0023698643781244755
step: 130, loss: 0.014406121335923672
step: 140, loss: 0.010081594809889793
step: 150, loss: 0.07376833260059357
step: 160, loss: 0.0025611247401684523
step: 170, loss: 0.017614159733057022
step: 180, loss: 0.014488487504422665
step: 190, loss: 0.0026204430032521486
step: 200, loss: 0.1189228892326355
step: 210, loss: 0.019568676128983498
/usr/local/lib/python3.7/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=0.6933867735470941, f1=0.6851485148514852,
best_f1=0.6858237547892722
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.019200801849365234
step: 10, loss: 0.012185790576040745
```

```
step: 20, loss: 0.0003892574750352651
step: 30, loss: 0.020605523139238358
step: 40, loss: 0.0001629416219657287
step: 50, loss: 0.0008329750271514058
step: 60, loss: 0.004478591959923506
step: 70, loss: 0.031497735530138016
step: 80, loss: 0.0031908145174384117
Gradient overflow.  Skipping step, loss scaler 0 reducing loss scale to 8192.0
step: 90, loss: 0.008657888509333134
step: 100, loss: 0.030557088553905487
step: 110, loss: 0.00017885548004414886
step: 120, loss: 0.01699383743107319
step: 130, loss: 0.05571303889155388
step: 140, loss: 0.0013341326266527176
step: 150, loss: 0.05760533735156059
step: 160, loss: 0.015276359394192696
step: 170, loss: 0.0013533206656575203
step: 180, loss: 0.04480905830860138
step: 190, loss: 0.0855480283498764
step: 200, loss: 0.0015235567698255181
step: 210, loss: 0.01583554968237877
/usr/local/lib/python3.7/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=0.6899383983572897, f1=0.6940451745379876,
best_f1=0.6858237547892722
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.023856276646256447
step: 10, loss: 0.028520723804831505
step: 20, loss: 0.00935782678425312
step: 30, loss: 0.0026838176418095827
step: 40, loss: 0.021532660350203514
step: 50, loss: 0.017657985910773277
step: 60, loss: 0.001782726845704019
step: 70, loss: 0.0037113421130925417
step: 80, loss: 0.05483262240886688
step: 90, loss: 0.010741185396909714
step: 100, loss: 0.024613086134195328
step: 110, loss: 0.09358198940753937
step: 120, loss: 0.07742271572351456
step: 130, loss: 0.00029983060085214674
step: 140, loss: 8.539092232240364e-05
step: 150, loss: 0.0014003252144902945
step: 160, loss: 0.017459500581026077
step: 170, loss: 0.026072517037391663
step: 180, loss: 0.00034199556102976203
```

```
step: 190, loss: 0.013623356819152832
step: 200, loss: 0.0004784954944625497
step: 210, loss: 0.0001972979516722262
/usr/local/lib/python3.7/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=0.698989898989899, f1=0.6944444444444444,
best_f1=0.6858237547892722
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.06805618852376938
step: 10, loss: 0.006442139856517315
step: 20, loss: 0.024676073342561722
step: 30, loss: 0.01580211892724037
step: 40, loss: 0.00833108089864254
step: 50, loss: 0.002629659604281187
step: 60, loss: 0.0008792132721282542
step: 70, loss: 0.0014075948856770992
step: 80, loss: 0.0003743976994883269
step: 90, loss: 0.00715081300586462
step: 100, loss: 0.007385516073554754
step: 110, loss: 0.0008013063343241811
step: 120, loss: 0.002309567527845502
step: 130, loss: 0.0006861591245979071
step: 140, loss: 0.0023346957750618458
step: 150, loss: 0.010719646699726582
step: 160, loss: 0.012819251045584679
step: 170, loss: 0.00550058763474226
step: 180, loss: 0.0003755370562430471
step: 190, loss: 0.0016693283105269074
step: 200, loss: 0.0003042641328647733
step: 210, loss: 0.040483709424734116
/usr/local/lib/python3.7/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=0.6954732510288065, f1=0.7087576374745417,
best_f1=0.6858237547892722
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.031114855781197548
step: 10, loss: 0.0023251017555594444
step: 20, loss: 0.025805797427892685
step: 30, loss: 0.034925878047943115
step: 40, loss: 0.03486964479088783
step: 50, loss: 0.0005311189452186227
step: 60, loss: 0.00020569877233356237
```

```
step: 70, loss: 0.030446602031588554
step: 80, loss: 0.005530581343919039
step: 90, loss: 0.00019766183686442673
step: 100, loss: 0.0003711206081788987
step: 110, loss: 0.00040917444857768714
step: 120, loss: 8.805565448710695e-05
step: 130, loss: 0.00670149848610163
step: 140, loss: 0.04700175300240517
step: 150, loss: 0.00033407853334210813
step: 160, loss: 0.0017867237329483032
step: 170, loss: 0.0018989361124113202
step: 180, loss: 0.03316314518451691
step: 190, loss: 0.00022220493701752275
step: 200, loss: 0.0005900968099012971
step: 210, loss: 0.00063729169778552583
/usr/local/lib/python3.7/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=0.6712018140589568, f1=0.6832579185520362,
best_f1=0.6858237547892722
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0001778160221874714
step: 10, loss: 0.0014148764312267303
step: 20, loss: 0.015374080277979374
step: 30, loss: 0.00024111261882353574
step: 40, loss: 0.00013531585864257067
step: 50, loss: 0.007092924788594246
step: 60, loss: 0.017602767795324326
step: 70, loss: 0.00014024657139088959
step: 80, loss: 0.00015422569413203746
step: 90, loss: 0.00021224345255177468
step: 100, loss: 0.0007800977327860892
step: 110, loss: 0.00014162609295453876
step: 120, loss: 0.019378304481506348
step: 130, loss: 0.003513060975819826
step: 140, loss: 0.015972357243299484
step: 150, loss: 0.00010073567682411522
step: 160, loss: 9.707530261948705e-05
step: 170, loss: 0.000444816512754187
step: 180, loss: 0.0018295724876224995
step: 190, loss: 0.00011006985732819885
step: 200, loss: 0.0007553450996056199
step: 210, loss: 0.04123806580901146
/usr/local/lib/python3.7/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=0.6723768736616702, f1=0.7106382978723405,
best_f1=0.6858237547892722
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00016352182137779891
step: 10, loss: 0.0002077579847536981
step: 20, loss: 0.0009378936956636608
step: 30, loss: 0.034532710909843445
step: 40, loss: 0.0006323182606138289
step: 50, loss: 8.370322757400572e-05
step: 60, loss: 0.0003510572714731097
step: 70, loss: 0.00016738398699089885
step: 80, loss: 9.27329674595967e-05
step: 90, loss: 0.01892555132508278
step: 100, loss: 0.0001053515516105108
step: 110, loss: 8.462019468424842e-05
step: 120, loss: 8.915547368815169e-05
step: 130, loss: 0.00012842040450777858
step: 140, loss: 0.00011005345004377887
step: 150, loss: 0.00010367207141825929
step: 160, loss: 0.0007853468996472657
step: 170, loss: 0.00012120702012907714
step: 180, loss: 5.7925572036765516e-05
step: 190, loss: 0.0003026595513802022
step: 200, loss: 0.0006712895701639354
step: 210, loss: 0.0007433033315464854
/usr/local/lib/python3.7/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=0.676595744680851, f1=0.7130801687763714,
best_f1=0.6858237547892722
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.")
2293it [00:08, 283.24it/s]
load_f1 = 0.6836935166994106
real_f1 = 0.686046511627907
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, 190.79it/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 bert-base-uncased \
  --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 bert-base-uncased \
  --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: 0.5823693871498108

```
step: 10, loss: 0.36139798164367676
step: 20, loss: 0.2790636122226715
step: 30, loss: 0.4119768440723419
step: 40, loss: 0.44671520590782166
step: 50, loss: 0.30343276262283325
step: 60, loss: 0.26127833127975464
step: 70, loss: 0.22446975111961365
step: 80, loss: 0.28559038043022156
step: 90, loss: 0.2478112429380417
step: 100, loss: 0.35517439246177673
step: 110, loss: 0.27189111709594727
step: 120, loss: 0.1476224660873413
step: 130, loss: 0.17033906280994415
step: 140, loss: 0.04734409600496292
step: 150, loss: 0.12294584512710571
step: 160, loss: 0.05829283967614174
step: 170, loss: 0.22430208325386047
step: 180, loss: 0.05505599454045296
step: 190, loss: 0.16613318026065826
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.6539509536784741, f1=0.6574585635359116,
best_f1=0.6574585635359116
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.17967119812965393
step: 10, loss: 0.07112343609333038
step: 20, loss: 0.05950772389769554
step: 30, loss: 0.16536715626716614
step: 40, loss: 0.22316820919513702
step: 50, loss: 0.04124399274587631
step: 60, loss: 0.17963923513889313
step: 70, loss: 0.10944081097841263
step: 80, loss: 0.33517390489578247
step: 90, loss: 0.2218671590089798
step: 100, loss: 0.026005398482084274
step: 110, loss: 0.12230682373046875
step: 120, loss: 0.19994384050369263
step: 130, loss: 0.10001827031373978
step: 140, loss: 0.10683367401361465
step: 150, loss: 0.05890798941254616
step: 160, loss: 0.029864901676774025
```

```
step: 170, loss: 0.11519993096590042
step: 180, loss: 0.33510681986808777
step: 190, loss: 0.02481904998421669
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.7867036011080332, f1=0.7616438356164383,
best_f1=0.7616438356164383
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.022343872115015984
step: 10, loss: 0.02482236921787262
step: 20, loss: 0.011309939436614513
step: 30, loss: 0.0161169171333313
step: 40, loss: 0.05391566827893257
step: 50, loss: 0.12146729975938797
step: 60, loss: 0.01932242326438427
step: 70, loss: 0.0649748295545578
step: 80, loss: 0.05909239500761032
step: 90, loss: 0.08042450994253159
step: 100, loss: 0.061029501259326935
step: 110, loss: 0.007285099010914564
step: 120, loss: 0.009488922543823719
step: 130, loss: 0.02574612945318222
step: 140, loss: 0.016337376087903976
step: 150, loss: 0.09487242251634598
step: 160, loss: 0.08640403300523758
step: 170, loss: 0.06463710218667984
step: 180, loss: 0.03787438943982124
step: 190, loss: 0.09734451025724411
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.8021108179419525, f1=0.802030456852792,
best_f1=0.802030456852792
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.02283041551709175
```

```
step: 10, loss: 0.16990026831626892
step: 20, loss: 0.03812193498015404
step: 30, loss: 0.009410946629941463
step: 40, loss: 0.016425572335720062
Gradient overflow.  Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 50, loss: 0.07011391967535019
step: 60, loss: 0.08538215607404709
step: 70, loss: 0.05389038845896721
step: 80, loss: 0.12212590873241425
step: 90, loss: 0.02496788464486599
step: 100, loss: 0.017213841900229454
step: 110, loss: 0.0017357668839395046
step: 120, loss: 0.04424126818776131
step: 130, loss: 0.14886124432086945
step: 140, loss: 0.016344722360372543
step: 150, loss: 0.008121415972709656
step: 160, loss: 0.008815024979412556
step: 170, loss: 0.12436394393444061
step: 180, loss: 0.012891613878309727
step: 190, loss: 0.04867830127477646
/usr/local/lib/python3.7/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=0.7671957671957672, f1=0.7777777777777779,
best_f1=0.802030456852792
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0721939206123352
step: 10, loss: 0.015240266919136047
step: 20, loss: 0.10138542950153351
step: 30, loss: 0.010217140428721905
step: 40, loss: 0.10061194002628326
step: 50, loss: 0.15603618323802948
step: 60, loss: 0.08403822034597397
step: 70, loss: 0.03257272019982338
step: 80, loss: 0.024161657318472862
step: 90, loss: 0.011556667275726795
step: 100, loss: 0.0017372473375871778
step: 110, loss: 0.037494976073503494
step: 120, loss: 0.0059323860332369804
step: 130, loss: 0.10684336721897125
step: 140, loss: 0.004927216097712517
step: 150, loss: 0.16192373633384705
step: 160, loss: 0.00948312971740961
step: 170, loss: 0.0022482769563794136
step: 180, loss: 0.003131476230919361
step: 190, loss: 0.038464222103357315
```

```
/usr/local/lib/python3.7/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=0.7616438356164383, f1=0.7548209366391184,
best_f1=0.802030456852792
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.004416508600115776
step: 10, loss: 0.18004998564720154
step: 20, loss: 0.003540550824254751
step: 30, loss: 0.0013530862051993608
step: 40, loss: 0.004347852431237698
step: 50, loss: 0.0028399445582181215
step: 60, loss: 0.16819961369037628
step: 70, loss: 0.17688293755054474
step: 80, loss: 0.026408646255731583
step: 90, loss: 0.016260787844657898
step: 100, loss: 0.0005952690844424069
step: 110, loss: 0.04186815395951271
step: 120, loss: 0.0017931152833625674
step: 130, loss: 0.004089527763426304
step: 140, loss: 0.003595352405682206
step: 150, loss: 0.010654152370989323
step: 160, loss: 0.020657211542129517
step: 170, loss: 0.04332399740815163
step: 180, loss: 0.003158275969326496
step: 190, loss: 0.004267342854291201
/usr/local/lib/python3.7/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=0.7819148936170213, f1=0.7798408488063661,
best_f1=0.802030456852792
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.014026794582605362
step: 10, loss: 0.0012075980193912983
step: 20, loss: 0.012435481883585453
step: 30, loss: 0.005342384334653616
step: 40, loss: 0.005264537874609232
step: 50, loss: 0.013309570960700512
step: 60, loss: 0.028523048385977745
step: 70, loss: 0.02391928620636463
step: 80, loss: 0.006843126844614744
step: 90, loss: 0.005413732491433 36205
step: 100, loss: 0.0013278024271130562
step: 110, loss: 0.003363928757607937
```

```
step: 120, loss: 0.00267079402692616
step: 130, loss: 0.0032707606442272663
step: 140, loss: 0.0007370196981355548
step: 150, loss: 0.022953882813453674
step: 160, loss: 0.009629325941205025
step: 170, loss: 0.0013098041526973248
step: 180, loss: 0.01389120239764452
step: 190, loss: 0.0026498616207391024
/usr/local/lib/python3.7/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=0.7880434782608696, f1=0.7989417989417988,
best_f1=0.802030456852792
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0014815190806984901
step: 10, loss: 0.013223609887063503
step: 20, loss: 0.0004260240530129522
step: 30, loss: 0.0063969907350838184
step: 40, loss: 0.0004663862055167556
step: 50, loss: 0.15459348261356354
Gradient overflow.  Skipping step, loss scaler 0 reducing loss scale to 8192.0
step: 60, loss: 0.0005536035168915987
step: 70, loss: 0.001675630221143365
step: 80, loss: 0.00018038696725852787
step: 90, loss: 0.0002726721577346325
step: 100, loss: 0.0008367647533304989
step: 110, loss: 0.0004922044463455677
step: 120, loss: 0.0010283977026119828
step: 130, loss: 0.0005596933187916875
step: 140, loss: 0.0005626978818327188
step: 150, loss: 0.0027743473183363676
step: 160, loss: 0.0008936345693655312
step: 170, loss: 0.002606310648843646
step: 180, loss: 0.004496345296502113
step: 190, loss: 0.4852410852909088
/usr/local/lib/python3.7/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=0.7688311688311688, f1=0.7735368956743003,
best_f1=0.802030456852792
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.003125400049611926
step: 10, loss: 0.04939427971839905
step: 20, loss: 0.0015230823773890734
```

```
step: 30, loss: 0.0028543751686811447
step: 40, loss: 0.001208114204928279
step: 50, loss: 0.004229413811117411
step: 60, loss: 0.018002083525061607
step: 70, loss: 0.0006583057111129165
step: 80, loss: 0.0005042277043685317
step: 90, loss: 0.0009581436170265079
step: 100, loss: 0.0814618170261383
step: 110, loss: 0.0035992045886814594
step: 120, loss: 0.0026816665194928646
step: 130, loss: 0.0025244627613574266
step: 140, loss: 0.00228467071428895
step: 150, loss: 0.004849170334637165
step: 160, loss: 0.00027753645554184914
step: 170, loss: 0.0009322082041762769
step: 180, loss: 0.00434296578168869
step: 190, loss: 0.0002506980672478676
/usr/local/lib/python3.7/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=0.7873563218390804, f1=0.776536312849162,
best_f1=0.802030456852792
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0007533561438322067
step: 10, loss: 0.0005206839996390045
step: 20, loss: 0.20725691318511963
step: 30, loss: 0.0037379942368716
step: 40, loss: 0.04035133123397827
step: 50, loss: 0.0006451589870266616
step: 60, loss: 0.0006848065531812608
step: 70, loss: 0.0007278256816789508
step: 80, loss: 0.000760201015509665
step: 90, loss: 0.0006546287913806736
step: 100, loss: 0.0023698064032942057
step: 110, loss: 0.0017040452221408486
step: 120, loss: 0.15135179460048676
step: 130, loss: 0.00043905971688218415
step: 140, loss: 0.001573885208927095
step: 150, loss: 0.0006396325770765543
step: 160, loss: 0.018822096288204193
step: 170, loss: 0.0013020007172599435
step: 180, loss: 0.00029432872543111444
step: 190, loss: 0.0007162682595662773
/usr/local/lib/python3.7/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=0.7903225806451613, f1=0.8115183246073298,
best_f1=0.802030456852792
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0005538600962609053
step: 10, loss: 0.0005234460113570094
step: 20, loss: 0.004193445667624474
step: 30, loss: 0.006658586207777262
step: 40, loss: 0.0003849398344755173
step: 50, loss: 0.0003051572712138295
step: 60, loss: 0.0004162489785812795
step: 70, loss: 0.0004327136557549238
step: 80, loss: 0.00046319596003741026
step: 90, loss: 0.0005331168649718165
step: 100, loss: 0.0016615181230008602
step: 110, loss: 0.0005565047031268477
step: 120, loss: 0.0006035658880136907
step: 130, loss: 0.00512717617675662
step: 140, loss: 0.0005544417072087526
step: 150, loss: 0.00025543637457303703
step: 160, loss: 0.0011810294818133116
step: 170, loss: 0.0028845653869211674
step: 180, loss: 0.00040378450648859143
step: 190, loss: 0.0007576559437438846
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.8042895442359249, f1=0.7814207650273224,
best_f1=0.7814207650273224
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.006039765663444996
step: 10, loss: 0.0006932242540642619
step: 20, loss: 0.000359586498234421
step: 30, loss: 0.14162983000278473
step: 40, loss: 0.0006049525691196322
step: 50, loss: 0.0010551774175837636
step: 60, loss: 0.020154988393187523
step: 70, loss: 0.00045878274249844253
step: 80, loss: 0.0019173604669049382
step: 90, loss: 0.0003233580500818789
step: 100, loss: 0.0006293276092037559
```

```
step: 110, loss: 0.0006966934888623655
step: 120, loss: 0.00017628143541514874
step: 130, loss: 0.0006320445681922138
step: 140, loss: 0.001473252777941525
step: 150, loss: 0.00038687142659910023
step: 160, loss: 0.00021077644487377256
step: 170, loss: 0.0015882629668340087
step: 180, loss: 8.76850972417742e-05
step: 190, loss: 0.00017582850705366582
/usr/local/lib/python3.7/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=0.7966573816155988, f1=0.7912087912087912,
best_f1=0.7814207650273224
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0005166057962924242
step: 10, loss: 0.0002861553803086281
step: 20, loss: 0.001296639209613204
step: 30, loss: 0.0018618948524817824
step: 40, loss: 0.0009737786022014916
step: 50, loss: 0.0002940944686997682
step: 60, loss: 0.0001525125262560323
step: 70, loss: 0.0018098036525771022
step: 80, loss: 0.0007169791497290134
step: 90, loss: 0.0003508404770400375
step: 100, loss: 0.0011442331597208977
step: 110, loss: 0.0001864479563664645
step: 120, loss: 0.004817552398890257
step: 130, loss: 0.00012880380381830037
step: 140, loss: 0.00022456690203398466
step: 150, loss: 0.0001697426341706887
step: 160, loss: 0.0001360053865937516
step: 170, loss: 0.00018407318566460162
step: 180, loss: 0.00013748220226261765
step: 190, loss: 0.0010834136046469212
/usr/local/lib/python3.7/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=0.7947368421052631, f1=0.7908163265306122,
best_f1=0.7814207650273224
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00013998591748531908
step: 10, loss: 0.000955234922003001
step: 20, loss: 0.00019947507826145738
```

```
step: 30, loss: 8.278261520899832e-05
step: 40, loss: 0.0024064930621534586
step: 50, loss: 0.00022571507724933326
step: 60, loss: 0.019904294982255253
step: 70, loss: 0.00011534907389432192
step: 80, loss: 9.766790753928944e-05
step: 90, loss: 0.0027706639375537634
step: 100, loss: 0.002492742147296667
step: 110, loss: 0.0005418646614998579
step: 120, loss: 0.0002979286073241383
step: 130, loss: 0.00017668626969680019
step: 140, loss: 0.00020590041822288185
step: 150, loss: 0.00017978948017116636
step: 160, loss: 0.0006636455655097961
step: 170, loss: 0.00021460893913172185
step: 180, loss: 0.0002210819802712649
step: 190, loss: 0.00019875913858413696
/usr/local/lib/python3.7/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=0.7978723404255318, f1=0.7866323907455013,
best_f1=0.7814207650273224
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.005377138499170542
step: 10, loss: 0.0002377331693423912
step: 20, loss: 0.00019460574549157172
step: 30, loss: 0.00010719015699578449
step: 40, loss: 0.13149785995483398
step: 50, loss: 0.00013558450154960155
step: 60, loss: 0.00025498951436020434
step: 70, loss: 0.00059710297500071466
step: 80, loss: 0.00028105787350796163
step: 90, loss: 0.0001561603567097336
step: 100, loss: 0.00026766999950825423
step: 110, loss: 0.00016833478002808988
step: 120, loss: 0.0002828467986546457
step: 130, loss: 0.00016365162446163595
step: 140, loss: 0.00019201147370040417
step: 150, loss: 0.00025680565158836544
step: 160, loss: 0.0001658838737057522
step: 170, loss: 0.04905526340007782
step: 180, loss: 0.002641662722453475
step: 190, loss: 0.00022922085190657526
/usr/local/lib/python3.7/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=0.7958115183246073, f1=0.7869674185463658,
best_f1=0.7814207650273224
mv: cannot stat '*_dev.pt': No such file or directory
Selected optimization level O2:  FP16 training with FP32 batchnorm and FP32
master weights.

Defaults for this optimization level are:
enabled                : True
opt_level              : O2
cast_model_type        : torch.float16
patch_torch_functions  : False
keep_batchnorm_fp32    : True
master_weights         : True
loss_scale             : dynamic
Processing user overrides (additional kwargs that are not None)…
After processing overrides, optimization options are:
enabled                : True
opt_level              : O2
cast_model_type        : torch.float16
patch_torch_functions  : False
keep_batchnorm_fp32    : True
master_weights         : True
loss_scale             : dynamic
Warning:  multi_tensor_applier fused unscale kernel is unavailable, possibly
because apex was installed without --cuda_ext --cpp_ext. Using Python fallback.
Original ImportError was: ModuleNotFoundError("No module named 'amp_C'")
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
2049it [00:11, 174.22it/s]
load_f1 = 0.5156794425087108
real_f1 = 0.5094664371772805
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, 190.97it/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 bert-base-uncased \
       --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 bert-base-uncased \
       --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: 0.6247397661209106
step: 10, loss: 0.3557359278202057
step: 20, loss: 0.2834051549434662
step: 30, loss: 0.40132585167884827
step: 40, loss: 0.2753695845603943
step: 50, loss: 0.26979121565818787
step: 60, loss: 0.23669739067554474
step: 70, loss: 0.3674873411655426
step: 80, loss: 0.33642902970314026
step: 90, loss: 0.2037467509508133
step: 100, loss: 0.09567026793956757
step: 110, loss: 0.131685271859169
step: 120, loss: 0.12959757447242737
step: 130, loss: 0.04608611389994621
step: 140, loss: 0.13999848067760468
step: 150, loss: 0.39430058002471924
step: 160, loss: 0.14999598264694214
step: 170, loss: 0.23324286937713623
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.7563451776649747, f1=0.727710843373494,
best_f1=0.727710843373494
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.09803439676761627
step: 10, loss: 0.20545488595962524
step: 20, loss: 0.09366738796234131
step: 30, loss: 0.27811795473098755
step: 40, loss: 0.05070352181792259
step: 50, loss: 0.06934833526611328
step: 60, loss: 0.14549550414085388
step: 70, loss: 0.12068112939596176
step: 80, loss: 0.09786354005336761
step: 90, loss: 0.16819195449352264
step: 100, loss: 0.07156727463006973
step: 110, loss: 0.08533837646245956
step: 120, loss: 0.07879835367202759
step: 130, loss: 0.053931478410959244
step: 140, loss: 0.35383936762809753
step: 150, loss: 0.14788289368152618
step: 160, loss: 0.18028166890144348
step: 170, loss: 0.03680713102221489
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.8037383177570093, f1=0.7667436489607391,
best_f1=0.7667436489607391
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.1639876812696457
step: 10, loss: 0.06437782198190689
step: 20, loss: 0.09415853023529053
step: 30, loss: 0.1554536670446396
step: 40, loss: 0.0693693608045578
step: 50, loss: 0.022579507902264595
step: 60, loss: 0.023029319941997528
step: 70, loss: 0.07122821360826492
step: 80, loss: 0.05866663530468941
step: 90, loss: 0.08580131083726883
step: 100, loss: 0.03139529749751091

```
step: 110, loss: 0.05998484417796135
step: 120, loss: 0.07870571315288544
step: 130, loss: 0.17779803276062012
step: 140, loss: 0.024851318448781967
step: 150, loss: 0.019277743995189667
step: 160, loss: 0.017877349629998207
step: 170, loss: 0.06171262264251709
/usr/local/lib/python3.7/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=0.8019559902200489, f1=0.7865707434052758,
best_f1=0.7667436489607391
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.006530554033815861
step: 10, loss: 0.019609889015555382
step: 20, loss: 0.04399288445711136
step: 30, loss: 0.025226695463061333
step: 40, loss: 0.0030955791007727385
step: 50, loss: 0.10359357297420502
step: 60, loss: 0.11500443518161774
step: 70, loss: 0.006238303147256374
Gradient overflow.  Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 80, loss: 0.02726191282272339
step: 90, loss: 0.08451920002698898
step: 100, loss: 0.17299172282218933
step: 110, loss: 0.08107675611972809
step: 120, loss: 0.018166884779930115
step: 130, loss: 0.03380856662988663
step: 140, loss: 0.026949871331453323
step: 150, loss: 0.13063260912895203
step: 160, loss: 0.08970829099416733
step: 170, loss: 0.006755107548087835
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.8421052631578948, f1=0.8105263157894737,
best_f1=0.8105263157894737
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.007052374072372913
step: 10, loss: 0.08909985423088074
```

```
step: 20, loss: 0.06588388234376907
step: 30, loss: 0.013061020523309708
step: 40, loss: 0.013039574958384037
step: 50, loss: 0.004659783095121384
step: 60, loss: 0.03800704702734947
step: 70, loss: 0.23741455376148224
step: 80, loss: 0.0489974170923233
step: 90, loss: 0.06120522692799568
step: 100, loss: 0.01729172095656395
step: 110, loss: 0.03956054151058197
step: 120, loss: 0.025052864104509354
step: 130, loss: 0.030158866196870804
step: 140, loss: 0.03643541783094406
step: 150, loss: 0.02451668307185173
step: 160, loss: 0.09319374710321426
step: 170, loss: 0.0175810307264328
/usr/local/lib/python3.7/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=0.8407960199004973, f1=0.836272040302267,
best_f1=0.81052631157894737
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.007393453735858202
step: 10, loss: 0.011073608882725239
step: 20, loss: 0.01889253780245781
step: 30, loss: 0.04147773236036301
step: 40, loss: 0.05325925350189209
step: 50, loss: 0.07707115262746811
step: 60, loss: 0.03636804223060608
step: 70, loss: 0.09096931666135788
step: 80, loss: 0.024120764806866646
step: 90, loss: 0.010934561491012573
step: 100, loss: 0.007033841218799353
step: 110, loss: 0.0029913310427218676
step: 120, loss: 0.05412686988711357
step: 130, loss: 0.009501706808805466
step: 140, loss: 0.03084452450275421
step: 150, loss: 0.015896888449788094
step: 160, loss: 0.016999203711748123
step: 170, loss: 0.003082261187955737
/usr/local/lib/python3.7/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=0.8393285371702638, f1=0.8137254901960784,
best_f1=0.81052631157894737
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
```

An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0014485983410850167
step: 10, loss: 0.056258365511894226
step: 20, loss: 0.03008987382054329
step: 30, loss: 0.029297756031155586
step: 40, loss: 0.008847227320075035
step: 50, loss: 0.04356733337044716
step: 60, loss: 0.0013196842046454549
step: 70, loss: 0.00048074114602059126
step: 80, loss: 0.04309306666254997
step: 90, loss: 0.00025988678680732846
step: 100, loss: 0.0004626772424671799
step: 110, loss: 0.008421923033893108
step: 120, loss: 0.001648954814299941
step: 130, loss: 0.12773381173610687
step: 140, loss: 0.011524415574967861
step: 150, loss: 0.0017628296045586467
step: 160, loss: 0.00038232735823839903
step: 170, loss: 0.013440708629786968
/usr/local/lib/python3.7/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=0.8277511961722487, f1=0.8173076923076923,
best_f1=0.8105263157894737
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.01979573257267475
step: 10, loss: 0.0020578971598297358
step: 20, loss: 0.025443060323596
step: 30, loss: 0.04129326343536377
step: 40, loss: 0.0033373581245541573
step: 50, loss: 0.001474421238526702
step: 60, loss: 0.014353714883327484
step: 70, loss: 0.0007475012680515647
step: 80, loss: 0.015224956907331944
step: 90, loss: 0.001989985816180706
step: 100, loss: 0.018865926191210747
step: 110, loss: 0.05880993977189064
step: 120, loss: 0.0350002646446228
step: 130, loss: 0.0037435346748679876
step: 140, loss: 0.01149837113916874
step: 150, loss: 0.021116366609930992
step: 160, loss: 0.006164606660604477
step: 170, loss: 0.0017917925724759698
/usr/local/lib/python3.7/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=0.8383838383838383, f1=0.8223350253807107,
best_f1=0.81052631578947370
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0005018077208660543
step: 10, loss: 0.00056534580653533334
step: 20, loss: 0.0005109384073875844
step: 30, loss: 0.006676857825368643
step: 40, loss: 0.00029798093601122995
step: 50, loss: 0.00048322626389563084
step: 60, loss: 0.00779092637822032
step: 70, loss: 0.057401143014431
step: 80, loss: 0.0016724836314097047
step: 90, loss: 0.04991661384701729
step: 100, loss: 0.007119432091712952
step: 110, loss: 0.002348655601963401
step: 120, loss: 0.0056169223971664906
step: 130, loss: 0.07143401354551315
step: 140, loss: 0.0005462634726427495
step: 150, loss: 0.0005910326144658029
step: 160, loss: 0.021150846034288406
step: 170, loss: 0.002123346086591482
/usr/local/lib/python3.7/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=0.8368421052631579, f1=0.8358974358974359,
best_f1=0.81052631578947370
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.052415672689676285
step: 10, loss: 0.0013870454858988523
step: 20, loss: 0.0022158552892506123
step: 30, loss: 0.011646276339888573
step: 40, loss: 0.0002445212157908827
step: 50, loss: 0.05683981999754906
step: 60, loss: 0.00022050221741665155
step: 70, loss: 0.001052939798682928
step: 80, loss: 0.0032348872628062963
step: 90, loss: 0.015621309168636799
step: 100, loss: 0.00034587091067805886
step: 110, loss: 0.006925664842128754
step: 120, loss: 0.0009932677494361997
step: 130, loss: 0.006491872947663069
step: 140, loss: 0.002214438281953335
step: 150, loss: 0.0012581741902977228
```

```
step: 160, loss: 0.001544465310871601
step: 170, loss: 0.0002776343608275056
/usr/local/lib/python3.7/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=0.8324607329842932, f1=0.8372093023255814,
best_f1=0.81052631157894737
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.022469041869044304
step: 10, loss: 0.004128649830818176
step: 20, loss: 0.23532213270664215
step: 30, loss: 0.00014429680595640093
step: 40, loss: 0.00020066062279511243
step: 50, loss: 0.004929929506033659
step: 60, loss: 0.0035269649233669043
step: 70, loss: 0.0007842086488381028
step: 80, loss: 0.0007843874045647681
step: 90, loss: 0.0002897685917560011
step: 100, loss: 0.004590378142893314
step: 110, loss: 0.04852055385708809
step: 120, loss: 0.0002118435368174687
step: 130, loss: 0.00012230032007209957
step: 140, loss: 0.013733474537730217
step: 150, loss: 0.0041802567429840565
step: 160, loss: 0.004010302014648914
step: 170, loss: 0.0049280025996267796
/usr/local/lib/python3.7/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=0.8052631578947368,
best_f1=0.81052631157894737
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.028586555272340775
step: 10, loss: 0.000138975286972709
step: 20, loss: 0.0001560179516673088
step: 30, loss: 0.00022802987950854003
step: 40, loss: 0.00024837470846250653
step: 50, loss: 9.569436952006072e-05
step: 60, loss: 0.006319834850728512
step: 70, loss: 0.05224290490150452
step: 80, loss: 7.18730443622917e-05
step: 90, loss: 0.00018588286184240133
step: 100, loss: 0.0004863426147494465
step: 110, loss: 0.00022006934159435332
```

```
step: 120, loss: 0.04662118852138519
step: 130, loss: 0.002932666800916195
step: 140, loss: 0.005481495056301355
step: 150, loss: 0.00038962418329901993
step: 160, loss: 0.00018373466446064413
step: 170, loss: 0.00015723310934845358
/usr/local/lib/python3.7/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=0.8342245989304814, f1=0.806282722513089,
best_f1=0.81052631578947737
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.027101965621113777
step: 10, loss: 0.0010031445417553186
step: 20, loss: 8.422793325735256e-05
step: 30, loss: 0.00010993202886311337
step: 40, loss: 0.00016059441259130836
step: 50, loss: 0.0003915915731340647
step: 60, loss: 0.016320014372467995
step: 70, loss: 0.000235377621720545
step: 80, loss: 0.0001642543647903949
step: 90, loss: 7.128328434191644e-05
step: 100, loss: 0.00018172601994592696
step: 110, loss: 7.810909301042557e-05
step: 120, loss: 0.04238767549395561
step: 130, loss: 0.00066809740383178
step: 140, loss: 9.989050886360928e-05
step: 150, loss: 0.01725485548377037
step: 160, loss: 0.00041188867180608213
step: 170, loss: 0.00019367859933166268
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.8503937007874015, f1=0.8286445012787724,
best_f1=0.8286445012787724
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.000240803012275137
step: 10, loss: 0.00010501353244762868
step: 20, loss: 0.0009393573272973299
step: 30, loss: 0.00010471000132383779
```

```
step: 40, loss: 0.0001952246093424037
step: 50, loss: 0.0002549690834712237
step: 60, loss: 0.00010861168266274035
step: 70, loss: 0.00011427969002397731
step: 80, loss: 0.0024676385801285505
step: 90, loss: 0.05531714856624603
step: 100, loss: 0.0001428145042154938
step: 110, loss: 0.00032266779453493655
step: 120, loss: 0.00610985467210412
step: 130, loss: 0.00019488527323119342
step: 140, loss: 0.00011650851229205728
step: 150, loss: 0.00023297264124266803
step: 160, loss: 4.595372956828214e-05
step: 170, loss: 0.00019231399346608669
/usr/local/lib/python3.7/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=0.8488063660477452, f1=0.824742268041237,
best_f1=0.8286445012787724
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00039390684105455875
step: 10, loss: 0.0009647891856729984
step: 20, loss: 0.010717416182160378
step: 30, loss: 0.00012746674474328756
step: 40, loss: 0.00014511636982206255
step: 50, loss: 0.0013555031036958098
step: 60, loss: 0.0044702948071062565
step: 70, loss: 0.00013254614896140993
step: 80, loss: 0.016949187964200974
step: 90, loss: 0.00013278382539283484
step: 100, loss: 0.000492600433062762
step: 110, loss: 0.00018438750703353435
step: 120, loss: 0.00017108075553551316
step: 130, loss: 0.0027756975032389164
step: 140, loss: 0.0003277519135735929
step: 150, loss: 0.010654158890247345
step: 160, loss: 0.00018583291966933757
step: 170, loss: 0.003023434430360794
/usr/local/lib/python3.7/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=0.8488063660477452, f1=0.8385416666666666,
best_f1=0.8286445012787724
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 [00:08, 231.98it/s]
load_f1 = 0.5216
real_f1 = 0.5220125786163522
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, 192.34it/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 bert-base-uncased \
  --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 bert-base-uncased \
  --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: 0.6194750070571899
step: 10, loss: 0.6161248683929443
step: 20, loss: 0.3646714389324188
step: 30, loss: 0.07974057644605637
step: 40, loss: 0.23349368572235107
step: 50, loss: 0.058283086866140366
step: 60, loss: 0.10044883191585541
step: 70, loss: 0.07128609716892242
step: 80, loss: 0.06452182680368423
step: 90, loss: 0.02671315148472786
step: 100, loss: 0.012483140453696251
step: 110, loss: 0.308052122592926
step: 120, loss: 0.017459839582443237
step: 130, loss: 0.011220098473131657
step: 140, loss: 0.0038542302791029215
step: 150, loss: 0.008862766437232494
step: 160, loss: 0.0030408089514821768
step: 170, loss: 0.14826180040836334
step: 180, loss: 0.02860269695520401
step: 190, loss: 0.08088603615760803
step: 200, loss: 0.21614088118076324
step: 210, loss: 0.015486326068639755
step: 220, loss: 0.031737830489873886
step: 230, loss: 0.020072808489203453
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9705215419501134, f1=0.9704545454545453,
best_f1=0.9704545454545453
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
```

```
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0073662251234054565
step: 10, loss: 0.01342717744410038
step: 20, loss: 0.07732295989990234
step: 30, loss: 0.017877647653222084
step: 40, loss: 0.020342664793133736
step: 50, loss: 0.002430680440738797
step: 60, loss: 0.0030196174047887325
step: 70, loss: 0.09509935975074768
step: 80, loss: 0.008904285728931427
step: 90, loss: 0.0032832338474690914
step: 100, loss: 0.10952574014663696
step: 110, loss: 0.20117107033729553
step: 120, loss: 0.15603390336036682
step: 130, loss: 0.014829511754214764
step: 140, loss: 0.0017981319688260555
step: 150, loss: 0.01569884642958641
step: 160, loss: 0.16152271628379822
step: 170, loss: 0.0017253595869988203
step: 180, loss: 0.008099624887108803
step: 190, loss: 0.004323799163103104
step: 200, loss: 0.0020966229494661093
step: 210, loss: 0.005608984734863043
step: 220, loss: 0.15206070244312286
step: 230, loss: 0.022555917501449585
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.972129319955407, f1=0.9718785151856018,
best_f1=0.9718785151856018
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.012753073126077652
step: 10, loss: 0.003391615580767393
step: 20, loss: 0.0026529557071626186
step: 30, loss: 0.011045488528907299
step: 40, loss: 0.11922239512205124
step: 50, loss: 0.0251388531178236
step: 60, loss: 0.004208309110254049
step: 70, loss: 0.025801509618759155
step: 80, loss: 0.0012001912109553814
step: 90, loss: 0.009886585175991058
```

```
step: 100, loss: 0.0008118034456856549
step: 110, loss: 0.0005542384460568428
step: 120, loss: 0.007345037069171667
step: 130, loss: 0.0005014548078179359
step: 140, loss: 0.007338046096265316
step: 150, loss: 0.015074782073497772
step: 160, loss: 0.004116553347557783
step: 170, loss: 0.003682695794850588
step: 180, loss: 0.02389516495168209
step: 190, loss: 0.0019012369448319077
step: 200, loss: 0.02804599329829216
step: 210, loss: 0.01989644579589367
step: 220, loss: 0.0004446805687621236
step: 230, loss: 0.0037094682920724154
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9752252252252253, f1=0.9762174405436014,
best_f1=0.9762174405436014
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00012402908760122955
step: 10, loss: 0.00025415653362870216
step: 20, loss: 0.00025439599994570017
step: 30, loss: 0.0021591896656900644
step: 40, loss: 0.028383836150169373
step: 50, loss: 0.00036227976670488715
step: 60, loss: 0.0003771544143091887
step: 70, loss: 0.0009666305268183351
step: 80, loss: 0.00193393649533391
step: 90, loss: 0.001081919064745307
step: 100, loss: 0.00050795276183635
step: 110, loss: 0.001288932398892939
step: 120, loss: 0.02604277990758419
step: 130, loss: 0.0006667233537882566
step: 140, loss: 0.0012615375453606248
step: 150, loss: 0.10786391794681549
step: 160, loss: 0.10004390776157379
step: 170, loss: 0.010935881175100803
step: 180, loss: 0.001379733206704259
step: 190, loss: 0.0032453418243676424
step: 200, loss: 0.006291971076279879
step: 210, loss: 0.17183321714401245
```

```
step: 220, loss: 0.0001818019663915038
step: 230, loss: 0.009873474948108196
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
  warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 4: dev_f1=0.9797752808988766, f1=0.9730337078651685,
best_f1=0.9730337078651685
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0005451628239825368
step: 10, loss: 0.00020262147882021964
step: 20, loss: 0.015504357405006886
step: 30, loss: 0.0001868314138846472
step: 40, loss: 0.0012964140623807907
step: 50, loss: 0.00020481737738009542
step: 60, loss: 0.09952779859304428
step: 70, loss: 0.0004382619808893651
step: 80, loss: 0.00022264209110289812
step: 90, loss: 0.003093302482739091
step: 100, loss: 0.0009967224905267358
step: 110, loss: 0.0024319260846823454
step: 120, loss: 0.001637789187952876
step: 130, loss: 0.005023275502026081
step: 140, loss: 0.004977117292582989
step: 150, loss: 0.004216309636831284
step: 160, loss: 0.011044611223042011
step: 170, loss: 0.0036645757500082254
step: 180, loss: 0.0006603632355108857
step: 190, loss: 0.09698843210935593
step: 200, loss: 0.0003144099027849734
step: 210, loss: 0.001642611576244235
step: 220, loss: 0.0010985886910930276
step: 230, loss: 0.00027295740437693894
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9798657718120806, f1=0.9753914988814317,
best_f1=0.9753914988814317
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
```

An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.023281658068299294
step: 10, loss: 0.000890693801920861
step: 20, loss: 0.0033239468466490507
step: 30, loss: 0.00010917381587205455
step: 40, loss: 0.0006090556853450835
step: 50, loss: 0.00021360802929848433
step: 60, loss: 0.00020411642617546022
step: 70, loss: 0.0007053544977679849
step: 80, loss: 0.0052948566153645515
step: 90, loss: 0.002000825246796012
step: 100, loss: 0.03423747792840004
step: 110, loss: 0.0006507838843390346
step: 120, loss: 0.002986431121826172
step: 130, loss: 0.0056368205696344376
step: 140, loss: 0.0003512481926009059
step: 150, loss: 0.22213885188102722
step: 160, loss: 0.019485656172037125
step: 170, loss: 0.0007494613528251648
step: 180, loss: 0.037206877022981644
step: 190, loss: 0.004200577735900879
step: 200, loss: 0.000856558617670089
step: 210, loss: 0.0013216867810115218
step: 220, loss: 0.006780377589166164
step: 230, loss: 0.09361729770898819
/usr/local/lib/python3.7/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=0.9700996677740864, f1=0.9720044792833147,
best_f1=0.9753914988814317
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.010997564531862736
step: 10, loss: 0.0009401583229191601
step: 20, loss: 0.001669325865805149
step: 30, loss: 0.002643708372488618
step: 40, loss: 0.002806133823469281
step: 50, loss: 0.001995963742956519
step: 60, loss: 0.00023166494793258607
step: 70, loss: 0.01763947494328022
step: 80, loss: 0.0055902814492583275
step: 90, loss: 0.00018868160259444267
step: 100, loss: 0.0007501806830987334
step: 110, loss: 0.00036276219179853797
step: 120, loss: 0.0009431431535631418
step: 130, loss: 0.0200492013245821

```
step: 140, loss: 0.00032895541517063975
step: 150, loss: 0.007622320204973221
step: 160, loss: 0.060898054391145706
step: 170, loss: 0.0021532792598800911
step: 180, loss: 0.0006986769731156528
step: 190, loss: 0.00029442779486998916
step: 200, loss: 0.05338705703616142
step: 210, loss: 0.000209630306692655593
step: 220, loss: 0.0023365411907434464
step: 230, loss: 0.007831570692360401
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9819819819819819, f1=0.9718785151856018,
best_f1=0.9718785151856018
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0004193398344796151
step: 10, loss: 0.00025966137764044106
step: 20, loss: 0.0012117901351302862
step: 30, loss: 0.00033176448778249323
step: 40, loss: 0.006811528000980616
step: 50, loss: 0.0035609796177595854
step: 60, loss: 0.00013345239858608693
step: 70, loss: 0.00017868535360321403
step: 80, loss: 0.00015544844791293144
step: 90, loss: 0.00010994343002886948
step: 100, loss: 0.0002762439544312656
step: 110, loss: 0.013667821884155273
step: 120, loss: 0.00029984707362018526
step: 130, loss: 0.0002536377578508109
step: 140, loss: 6.970459071453661e-05
step: 150, loss: 0.00029449930298142135
step: 160, loss: 0.00028067186940461397
step: 170, loss: 0.0003326224396005273
step: 180, loss: 0.0004560007946565747
step: 190, loss: 0.0003020634758286178
step: 200, loss: 0.0004379774909466505
step: 210, loss: 0.0011177943088114262
step: 220, loss: 0.00012392719509080052
step: 230, loss: 0.0003272832836955786
/usr/local/lib/python3.7/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=0.9799107142857142, f1=0.9710467706013363,
best_f1=0.9718785151856018
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 4.29868778115565094e-05
step: 10, loss: 0.000912858871743083
step: 20, loss: 0.0023990122135728598
step: 30, loss: 8.318775508087128e-05
step: 40, loss: 0.011642088182270527
step: 50, loss: 4.95616186526604e-05
step: 60, loss: 0.00011786506365751848
step: 70, loss: 0.00024410116020590067
step: 80, loss: 9.167271491605788e-05
step: 90, loss: 0.00011274839926045388
step: 100, loss: 0.0006965339416638017
step: 110, loss: 3.833464870695025e-05
step: 120, loss: 4.1561037505744025e-05
step: 130, loss: 0.0013020883779972792
step: 140, loss: 4.627997986972332e-05
step: 150, loss: 0.0001229079207405448
step: 160, loss: 0.00032445212127640843
step: 170, loss: 8.313944999827072e-05
step: 180, loss: 0.007619715295732021
step: 190, loss: 9.758550731930882e-05
step: 200, loss: 9.196051541948691e-05
step: 210, loss: 0.00010489217675058171
step: 220, loss: 7.222348358482122e-05
step: 230, loss: 0.00023090248578228056
/usr/local/lib/python3.7/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=0.980963045912654, f1=0.968609865470852,
best_f1=0.9718785151856018
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0002449035528115928
step: 10, loss: 0.05585525184869766
step: 20, loss: 0.085670106112957
Gradient overflow.  Skipping step, loss scaler 0 reducing loss scale to 32768.0
step: 30, loss: 0.012515585869550705
step: 40, loss: 0.00039065987220965326
step: 50, loss: 0.0009422289440408349
step: 60, loss: 0.0012978101149201393
step: 70, loss: 0.0004530657606665045
step: 80, loss: 0.00020306259102653712
```

```
step: 90, loss: 0.00172189821023494
step: 100, loss: 0.0001300701405853033
step: 110, loss: 0.00025678728707134724
step: 120, loss: 8.373724267585203e-05
step: 130, loss: 9.662059164838865e-05
step: 140, loss: 0.032384295016527176
step: 150, loss: 0.017053600400686264
step: 160, loss: 9.638887422624975e-05
step: 170, loss: 0.00011405273107811809
step: 180, loss: 0.0010326550109311938
step: 190, loss: 0.0035033582244068384
step: 200, loss: 5.582606900134124e-05
step: 210, loss: 0.000733471882995218
step: 220, loss: 6.501511234091595e-05
step: 230, loss: 0.0004392405680846423
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9820224719101124, f1=0.9774266365688488,
best_f1=0.9774266365688488
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 6.710442539770156e-05
step: 10, loss: 0.00010335882689105347
step: 20, loss: 0.00020921658142469823
step: 30, loss: 0.0014432409079745412
step: 40, loss: 9.91829001577571e-05
step: 50, loss: 0.00024843381834216416
step: 60, loss: 0.000856187951285392
step: 70, loss: 0.00017394873430021107
step: 80, loss: 4.728895873995498e-05
step: 90, loss: 6.311019387794659e-05
step: 100, loss: 6.189555278979242e-05
step: 110, loss: 0.00020288307860028
step: 120, loss: 4.9591446440899745e-05
step: 130, loss: 3.548484164639376e-05
step: 140, loss: 6.027493145666085e-05
step: 150, loss: 0.0009968304075300694
step: 160, loss: 5.1803737733280286e-05
step: 170, loss: 8.135528332786635e-05
step: 180, loss: 0.000173702763277106
step: 190, loss: 5.61961714993231e-05
step: 200, loss: 0.00028400789597071707
```

```
step: 210, loss: 9.431868966203183e-05
step: 220, loss: 8.000041270861402e-05
step: 230, loss: 0.0014716187724843621
/usr/local/lib/python3.7/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=0.9763779527559054, f1=0.9685393258426966,
best_f1=0.9774266365688488
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 6.561485497513786e-05
step: 10, loss: 5.3364648920251057e-05
step: 20, loss: 9.353779751108959e-05
step: 30, loss: 0.0001190494149341248
step: 40, loss: 0.00033301967778243124
step: 50, loss: 0.0013871947303414345
step: 60, loss: 0.0006152118439786136
step: 70, loss: 7.034646114334464e-05
step: 80, loss: 0.00020031616440974176
step: 90, loss: 5.2619139751186594e-05
step: 100, loss: 0.0004880890191998333
step: 110, loss: 0.0002807346754707396
step: 120, loss: 5.970513302600011e-05
step: 130, loss: 5.640030576614663e-05
step: 140, loss: 0.00011293105490040034
step: 150, loss: 0.00015013596566859633
step: 160, loss: 5.908196544623934e-05
step: 170, loss: 7.243028085213155e-05
step: 180, loss: 0.005496774334460497
Gradient overflow.  Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 190, loss: 0.02734309248626232
step: 200, loss: 4.094756513950415e-05
step: 210, loss: 5.9830908867297694e-05
step: 220, loss: 0.00014583885786123574
step: 230, loss: 0.006487508304417133
/usr/local/lib/python3.7/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=0.9787709497206705, f1=0.96875, best_f1=0.9774266365688488
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 7.479346095351502e-05
step: 10, loss: 7.391707913484424e-05
step: 20, loss: 0.00023488550505135208
step: 30, loss: 6.527345249196514e-05
step: 40, loss: 6.080164894228801e-05
```

```
step: 50, loss: 5.9858302847715095e-05
step: 60, loss: 4.8172259994316846e-05
step: 70, loss: 3.0784740374656394e-05
step: 80, loss: 4.771642852574587e-05
step: 90, loss: 0.0005872056935913861
step: 100, loss: 3.336993540870026e-05
step: 110, loss: 8.669902308611199e-05
step: 120, loss: 0.0011662017786875367
step: 130, loss: 5.668696030625142e-05
step: 140, loss: 5.71615673834458e-05
step: 150, loss: 8.604550384916365e-05
step: 160, loss: 0.0002100130368489772
step: 170, loss: 0.00039263913640752435
step: 180, loss: 4.31853550253436e-05
step: 190, loss: 5.835989577462897e-05
step: 200, loss: 7.238052785396576e-05
step: 210, loss: 2.431424945825711e-05
step: 220, loss: 4.868466567131691e-05
step: 230, loss: 6.1016948166070506e-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=0.977728285077951, f1=0.9698996655518396,
best_f1=0.9774266365688488
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 5.488747774506919e-05
step: 10, loss: 5.10647842020262e-05
step: 20, loss: 3.84393533749997154e-05
step: 30, loss: 5.974433588562533e-05
step: 40, loss: 0.014307687059044838
step: 50, loss: 5.3744963224744424e-05
step: 60, loss: 4.833558341488242e-05
step: 70, loss: 5.3662326536141336e-05
step: 80, loss: 5.420409070211463e-05
step: 90, loss: 0.0009028479689732194
step: 100, loss: 4.5553362724604085e-05
step: 110, loss: 5.358768248697743e-05
step: 120, loss: 3.0273571610450745e-05
step: 130, loss: 4.140469536650926e-05
step: 140, loss: 4.4578162487596273e-05
step: 150, loss: 3.555994408088736e-05
step: 160, loss: 0.01109416875988245
step: 170, loss: 3.980353721999563e-05
step: 180, loss: 3.8297614082694054e-05
step: 190, loss: 0.00010308334458386526
step: 200, loss: 3.2799776818137616e-05
```

```
step: 210, loss: 0.00013119280629325658
step: 220, loss: 3.902861863025464e-05
step: 230, loss: 0.001969914184883237
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
epoch 14: dev_f1=0.9799107142857142, f1=0.9665178571428571,
best_f1=0.9774266365688488
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 3.156655657221563e-05
step: 10, loss: 2.0995266822865233e-05
step: 20, loss: 3.3466734748799354e-05
step: 30, loss: 3.526595173752867e-05
step: 40, loss: 0.0003110423276666552
step: 50, loss: 0.03520059213042259
step: 60, loss: 4.955410258844495e-05
step: 70, loss: 3.478911821730435e-05
step: 80, loss: 0.0002521382411941886
step: 90, loss: 6.2996317865327e-05
step: 100, loss: 4.366923894849606e-05
step: 110, loss: 3.701599416672252e-05
step: 120, loss: 5.376990884542465e-05
step: 130, loss: 6.336024671327323e-05
step: 140, loss: 3.3448057365603745e-05
step: 150, loss: 6.808878242736682e-05
step: 160, loss: 2.925343687820714e-05
step: 170, loss: 2.7256901375949383e-05
step: 180, loss: 4.8743211664259434e-05
step: 190, loss: 0.0003428220807109028
step: 200, loss: 5.140840221429244e-05
step: 210, loss: 7.39823590265587e-05
step: 220, loss: 0.0002078630350297317
step: 230, loss: 7.048348925309256e-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=0.9788182831661093, f1=0.9675977653631285,
best_f1=0.9774266365688488
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.")
2473it [00:14, 166.45it/s]
load_f1 = 0.9810055865921787
real_f1 = 0.9821826280623607
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, 190.94it/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 bert-base-uncased \
       --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 bert-base-uncased \
  --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: 0.6272004246711731
step: 10, loss: 0.5235728025436401
step: 20, loss: 0.5073021054267883
step: 30, loss: 0.10616632550954819
step: 40, loss: 0.21608470380306244
step: 50, loss: 0.1492369920015335
step: 60, loss: 0.07054004818201065
step: 70, loss: 0.08430299907922745
step: 80, loss: 0.02450070157647133
step: 90, loss: 0.2002764642238617
step: 100, loss: 0.028222206979990005
step: 110, loss: 0.04467798396945
step: 120, loss: 0.17538297176361084
step: 130, loss: 0.06589365005493164
step: 140, loss: 0.07256913930177689
step: 150, loss: 0.05100129544734955
step: 160, loss: 0.060991380363702774
step: 170, loss: 0.27634960412979126
step: 180, loss: 0.03698642551898956
step: 190, loss: 0.016247019171714783
step: 200, loss: 0.2553931176662445
step: 210, loss: 0.05721835047006607
step: 220, loss: 0.19579879939556122
step: 230, loss: 0.15862980484962463
step: 240, loss: 0.06677629053592682
step: 250, loss: 0.011067197658121586
step: 260, loss: 0.022216342389583588
step: 270, loss: 0.011600811965763569
step: 280, loss: 0.06013108417391777
step: 290, loss: 0.09856702387332916
step: 300, loss: 0.016628216952085495
step: 310, loss: 0.3863380551338196
step: 320, loss: 0.12281493842601776
step: 330, loss: 0.04300514981150627
step: 340, loss: 0.0727325826883316
step: 350, loss: 0.028976909816265106
step: 360, loss: 0.019710782915353775
step: 370, loss: 0.1554291844367981
step: 380, loss: 0.050133660435676575
step: 390, loss: 0.12444085627794266
step: 400, loss: 0.19818562269210815

```
step: 410, loss: 0.06177016720175743
step: 420, loss: 0.037731193006038666
step: 430, loss: 0.20412544906139374
step: 440, loss: 0.09866957366466522
step: 450, loss: 0.01208152249455452
step: 460, loss: 0.005931690335273743
step: 470, loss: 0.1063336506485939
step: 480, loss: 0.0716637447476387
step: 490, loss: 0.10688390582799911
step: 500, loss: 0.12221809476613998
step: 510, loss: 0.06144881993532181
step: 520, loss: 0.0702042207121849
step: 530, loss: 0.009836861863732338
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9405255878284925, f1=0.9377593360995851,
best_f1=0.9377593360995851
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.14288875460624695
step: 10, loss: 0.06131424754858017
step: 20, loss: 0.008292056620121002
step: 30, loss: 0.058203067630529404
step: 40, loss: 0.08619215339422226
step: 50, loss: 0.19534224271774292
step: 60, loss: 0.005707398988306522
step: 70, loss: 0.023439066484570503
step: 80, loss: 0.06768332421779633
step: 90, loss: 0.009037629701197147
step: 100, loss: 0.013542176224291325
step: 110, loss: 0.07860809564590454
step: 120, loss: 0.04427005723118782
step: 130, loss: 0.0987568348646164
step: 140, loss: 0.041056860238313675
step: 150, loss: 0.05708111450076103
step: 160, loss: 0.026686975732445717
step: 170, loss: 0.007395267952233553
step: 180, loss: 0.04555478319525719
step: 190, loss: 0.057669468224048615
step: 200, loss: 0.02169552631676197
step: 210, loss: 0.023508861660957336
step: 220, loss: 0.049335725605487823
```

```
step: 230, loss: 0.005089908372610807
step: 240, loss: 0.03694797307252884
step: 250, loss: 0.023923128843307495
step: 260, loss: 0.0044083306565880775
step: 270, loss: 0.1293390691280365
step: 280, loss: 0.033159613609313965
step: 290, loss: 0.014812936075031757
step: 300, loss: 0.16617673635482788
step: 310, loss: 0.016358425840735435
step: 320, loss: 0.053163930773735046
step: 330, loss: 0.042985063046216965
step: 340, loss: 0.01803339645266533
step: 350, loss: 0.004884444177150726
step: 360, loss: 0.01846652664244175
step: 370, loss: 0.16563540697097778
step: 380, loss: 0.07807080447673798
step: 390, loss: 0.054316818714141846
step: 400, loss: 0.04873327538371086
step: 410, loss: 0.004623436369001865
step: 420, loss: 0.017648285254836082
step: 430, loss: 0.007643613498657942
step: 440, loss: 0.20685136318206787
step: 450, loss: 0.032134462147951126
step: 460, loss: 0.02397949807345867
step: 470, loss: 0.12230385839939117
step: 480, loss: 0.19799016416072845
step: 490, loss: 0.01499460730701685
step: 500, loss: 0.10351144522428513
step: 510, loss: 0.010308869183063507
step: 520, loss: 0.05392074212431908
step: 530, loss: 0.005216307006776333
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9431345353675451, f1=0.9425287356321841,
best_f1=0.9425287356321841
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.11098097264766693
step: 10, loss: 0.07666566222906113
step: 20, loss: 0.043073903769254684
step: 30, loss: 0.089785635471344
step: 40, loss: 0.11462970823049545
```

```
step: 50, loss: 0.08958758413791656
step: 60, loss: 0.014724367298185825
step: 70, loss: 0.0014668825315311551
step: 80, loss: 0.0051310257986187935
step: 90, loss: 0.006068656221032143
step: 100, loss: 0.04995546489953995
step: 110, loss: 0.0013560012448579073
step: 120, loss: 0.014665033668279648
step: 130, loss: 0.0026652689557522535
step: 140, loss: 0.015101131983101368
step: 150, loss: 0.010269733145833015
step: 160, loss: 0.0058285389095544815
step: 170, loss: 0.0763477310538292
step: 180, loss: 0.007921409793198109
step: 190, loss: 0.0028929964173585176
step: 200, loss: 0.03562583029270172
step: 210, loss: 0.06871092319488525
step: 220, loss: 0.01750458963215351
step: 230, loss: 0.18137693405151367
step: 240, loss: 0.002310781739652157
step: 250, loss: 0.018983112648129463
step: 260, loss: 0.012777854688465595
step: 270, loss: 0.004828695207834244
step: 280, loss: 0.09714604914188385
step: 290, loss: 0.00921417772769928
step: 300, loss: 0.033964600414037704
step: 310, loss: 0.004470459651201963
step: 320, loss: 0.04122941568493843
step: 330, loss: 0.0018927904311567545
step: 340, loss: 0.019195031374692917
step: 350, loss: 0.0034629369620233774
step: 360, loss: 0.02764681726694107
step: 370, loss: 0.022948380559682846
step: 380, loss: 0.006561726331710815
step: 390, loss: 0.021461213007569313
step: 400, loss: 0.006984437815845013
step: 410, loss: 0.004919407423585653
step: 420, loss: 0.04898110032081604
step: 430, loss: 0.005915164016187191
step: 440, loss: 0.036522697657346725
step: 450, loss: 0.09842580556869507
step: 460, loss: 0.04361281543970108
step: 470, loss: 0.1199057474732399
step: 480, loss: 0.026036398485302925
step: 490, loss: 0.0216888934373856
step: 500, loss: 0.042767394334077835
step: 510, loss: 0.003363149706274271
step: 520, loss: 0.1067817285656929
```

```
step: 530, loss: 0.002842986024916172
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9439338235294118, f1=0.9352914180816887,
best_f1=0.9352914180816887
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.033740390092134476
step: 10, loss: 0.010532183572649956
step: 20, loss: 0.0018326519057154655
step: 30, loss: 0.001579827512614429
step: 40, loss: 0.0012204106897115707
step: 50, loss: 0.0036830792669206858
step: 60, loss: 0.0046413978561758995
step: 70, loss: 0.02352041006088257
step: 80, loss: 0.014415772631764412
step: 90, loss: 0.015117165632545948
step: 100, loss: 0.006948092486709356
step: 110, loss: 0.07793375104665756
step: 120, loss: 0.00047046825056895614
step: 130, loss: 0.00033412702032364905
step: 140, loss: 0.0011833348544314504
step: 150, loss: 0.01163561549037695
Gradient overflow.  Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 160, loss: 0.00919048860669136
step: 170, loss: 0.05164431780576706
step: 180, loss: 0.008004611358046532
step: 190, loss: 0.005251975264400244
step: 200, loss: 0.030644405633211136
step: 210, loss: 0.009546671994030476
step: 220, loss: 0.01827685907483101
step: 230, loss: 0.0385027751326561
step: 240, loss: 0.01522266399115324
step: 250, loss: 0.04569340869784355
step: 260, loss: 0.0013314869720488787
step: 270, loss: 0.006375585217028856
step: 280, loss: 0.01774822548031807
step: 290, loss: 0.037680529057979584
step: 300, loss: 0.00040790141792967916
step: 310, loss: 0.0021345065906643867
step: 320, loss: 0.09997234493494034
step: 330, loss: 0.002084638923406601
```

```
step: 340, loss: 0.0009889323264360428
step: 350, loss: 0.0074003166519105434
step: 360, loss: 0.009110280312597752
step: 370, loss: 0.0018353033810853958
step: 380, loss: 0.003679964691400528
step: 390, loss: 0.0013888422399759293
step: 400, loss: 0.01073186844587326
step: 410, loss: 0.0009988319361582398
step: 420, loss: 0.0009056529379449785
step: 430, loss: 0.2086169719696045
step: 440, loss: 0.0013713842490687966
step: 450, loss: 0.0027041276916861534
step: 460, loss: 0.0009840400889515877
step: 470, loss: 0.014690746553242207
step: 480, loss: 0.11242925375699997
step: 490, loss: 0.005162014625966549
step: 500, loss: 0.0017483358969911933
step: 510, loss: 0.02563529834151268
step: 520, loss: 0.06061875820159912
step: 530, loss: 0.016510339453816414
/usr/local/lib/python3.7/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=0.9418819188191883, f1=0.9310027598896043,
best_f1=0.9352914180816887
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.019639020785689354
step: 10, loss: 0.0419631190598011
step: 20, loss: 0.007081306539475918
step: 30, loss: 0.017749568447470665
step: 40, loss: 0.004715407267212868
step: 50, loss: 0.0008197634015232325
step: 60, loss: 0.1334879994392395
step: 70, loss: 0.003988267853856087
step: 80, loss: 0.0013969797873869538
step: 90, loss: 0.01665639691054821
step: 100, loss: 0.014468623325228691
step: 110, loss: 0.0006038848659954965
step: 120, loss: 0.025274403393268585
step: 130, loss: 0.0009008751367218792
step: 140, loss: 0.004323720466345549
step: 150, loss: 0.004058232996612787
step: 160, loss: 0.0005972103681415319
step: 170, loss: 0.02184934727847576
step: 180, loss: 0.001340566435828805
step: 190, loss: 0.0015326832653954625
```

```
step: 200, loss: 0.0005040615797042847
step: 210, loss: 0.0010265057208016515
step: 220, loss: 0.0012519779847934842
step: 230, loss: 0.003255135379731655
step: 240, loss: 0.006182704586535692
step: 250, loss: 0.0022237251978367567
step: 260, loss: 0.07742716372013092
step: 270, loss: 0.004707083571702242
step: 280, loss: 0.0012679058127105236
step: 290, loss: 0.1083468422293663
step: 300, loss: 0.001889124047011137
step: 310, loss: 0.001265066210180521
step: 320, loss: 0.20797783136367798
step: 330, loss: 0.007567860186100006
step: 340, loss: 0.0028523497749119997
step: 350, loss: 0.002007841132581234
step: 360, loss: 0.009524885565042496
step: 370, loss: 0.0038357821758836508
step: 380, loss: 0.002425567479804158
step: 390, loss: 0.00030943553429096937
step: 400, loss: 0.02158728614449501
step: 410, loss: 0.0005181105807423592
step: 420, loss: 0.0006942061008885503
step: 430, loss: 0.003348582424223423
step: 440, loss: 0.01251804269850254
step: 450, loss: 0.35068389773368835
step: 460, loss: 0.011592944152653217
step: 470, loss: 0.01923641934990883
step: 480, loss: 0.01267493236809969
step: 490, loss: 0.0031214230693876743
step: 500, loss: 0.005353225395083427
step: 510, loss: 0.05186744034290314
step: 520, loss: 0.002062127459794283
step: 530, loss: 0.00859035737812519
/usr/local/lib/python3.7/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=0.9429234338747099, f1=0.9294336118848654,
best_f1=0.9352914180816887
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0022042966447770596
step: 10, loss: 0.005314744543284178
step: 20, loss: 0.11751555651426315
step: 30, loss: 0.03947502747178078
step: 40, loss: 0.0048834169283509254
step: 50, loss: 0.03212522342801094
```

```
step: 60, loss: 0.00044238465488888323
step: 70, loss: 0.0010357487481087446
step: 80, loss: 0.033158741891384125
step: 90, loss: 0.021143700927495956
step: 100, loss: 0.0771137923002243
step: 110, loss: 0.0018515175906941295
step: 120, loss: 0.00031796353869140155
step: 130, loss: 0.0027306987904012203
step: 140, loss: 0.0017729797400534153
step: 150, loss: 0.0002591146621853113
step: 160, loss: 0.001160445623099804
step: 170, loss: 0.0008183013997040689
step: 180, loss: 0.00022569352586288005
step: 190, loss: 0.0006258194916881621
step: 200, loss: 0.00014513432688545436
step: 210, loss: 0.001925033750012517
step: 220, loss: 0.0025275058578699827
step: 230, loss: 0.00033298644120804965
step: 240, loss: 0.15827083587646484
step: 250, loss: 0.0037474865093827248
step: 260, loss: 0.00758483586832881
step: 270, loss: 0.10261627286672592
step: 280, loss: 0.0036953389644622803
step: 290, loss: 0.00047219530097208917
step: 300, loss: 0.0006791278719902039
step: 310, loss: 0.0010142945684492588
step: 320, loss: 0.001575126894749701
step: 330, loss: 0.002969358116388321
step: 340, loss: 0.053577642887830734
step: 350, loss: 0.0007964418618939817
step: 360, loss: 0.010601851157844067
step: 370, loss: 0.0018512862734496593
step: 380, loss: 0.000511490274220705
step: 390, loss: 0.01776484027504921
step: 400, loss: 0.0051491702906787395
step: 410, loss: 0.0008601297740824521
step: 420, loss: 0.001998472260311246
step: 430, loss: 0.002798121189698577
step: 440, loss: 0.0008934882353059947
step: 450, loss: 0.0008485745056532323
step: 460, loss: 0.0008721243357285857
step: 470, loss: 0.0072496505454182625
step: 480, loss: 0.0023898605722934008
step: 490, loss: 0.006004024296998978
step: 500, loss: 0.0028429124504327774
step: 510, loss: 0.012903902679681778
step: 520, loss: 0.0966711938381195
step: 530, loss: 0.009109114296734333
```

```
/usr/local/lib/python3.7/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=0.9337717238139972, f1=0.9251764705882354,
best_f1=0.9352914180816887
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.004848989192396402
step: 10, loss: 0.0025244993157684803
step: 20, loss: 0.0010460257763043046
step: 30, loss: 0.0009797875536605716
step: 40, loss: 0.07600057125091553
step: 50, loss: 0.011286946013569832
step: 60, loss: 0.0007554487092420459
step: 70, loss: 0.0005808662390336394
step: 80, loss: 0.005165865644812584
step: 90, loss: 9.056152339326218e-05
step: 100, loss: 0.0001916752225952223
step: 110, loss: 0.004827518947422504
step: 120, loss: 0.00037873885594308376
step: 130, loss: 0.0013969405554234982
step: 140, loss: 0.0006865523755550385
step: 150, loss: 0.005982842296361923
step: 160, loss: 0.000419168034568429
step: 170, loss: 0.0032943100668489933
step: 180, loss: 0.0024655915331095457
step: 190, loss: 0.0025333696976304054
step: 200, loss: 0.000658104894682765
step: 210, loss: 0.005331553053110838
step: 220, loss: 0.0007956284680403769
step: 230, loss: 0.04417821019887924
step: 240, loss: 0.0013141511008143425
step: 250, loss: 0.020853858441114426
step: 260, loss: 0.0001737868442432955
step: 270, loss: 0.001854151370935142
step: 280, loss: 0.0015514246188104153
step: 290, loss: 0.0005422562826424837
step: 300, loss: 0.0024811476469039917
step: 310, loss: 0.0007841947954148054
step: 320, loss: 0.00015179529145825654
step: 330, loss: 0.0007812652620486915
step: 340, loss: 0.09908942878246307
step: 350, loss: 0.0024240389466285706
step: 360, loss: 0.0012987595982849598
step: 370, loss: 0.0009199728956446052
step: 380, loss: 0.0002993193920701742
step: 390, loss: 0.004757903516292572
```

```
step: 400, loss: 0.007238794583827257
step: 410, loss: 0.002082107588648796
step: 420, loss: 0.0006316208746284246
step: 430, loss: 0.00011138594709336758
step: 440, loss: 0.00037582352524623275
step: 450, loss: 0.0010071111610159278
step: 460, loss: 0.0011608779896050692
step: 470, loss: 0.006261701695621014
step: 480, loss: 0.001074352883733809
step: 490, loss: 0.0005348783452063799
step: 500, loss: 0.0014875531196594238
step: 510, loss: 0.0014282800257205963
step: 520, loss: 0.013070868328213692
step: 530, loss: 0.001356185064651072
/usr/local/lib/python3.7/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=0.9436422915696321, f1=0.9327102803738317,
best_f1=0.9352914180816887
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0002564545429777354
step: 10, loss: 0.018846753984689713
step: 20, loss: 8.835546032059938e-05
step: 30, loss: 7.925854879431427e-05
step: 40, loss: 0.00013519621279556304
step: 50, loss: 0.0007700116257183254
step: 60, loss: 0.0017727125668898225
step: 70, loss: 0.0014928141608834267
step: 80, loss: 0.006341694854199886
step: 90, loss: 0.0008242556941695511
step: 100, loss: 0.00015310866001527756
step: 110, loss: 0.0001792669208953157
step: 120, loss: 0.0007706396281719208
step: 130, loss: 0.00020649044017773122
step: 140, loss: 0.0013928741682320833
step: 150, loss: 0.00042227667290717363
step: 160, loss: 0.004527080338448286
step: 170, loss: 0.12328045070171356
step: 180, loss: 0.00013499501801561564
step: 190, loss: 0.002457579830661416
step: 200, loss: 0.001471907366067171
step: 210, loss: 0.00032357353484258056
step: 220, loss: 0.0021448128391057253
step: 230, loss: 0.0014170485083013773
step: 240, loss: 0.00030919871642254293
step: 250, loss: 0.00016973938909359276
```

```
step: 260, loss: 6.286644929787144e-05
step: 270, loss: 0.0012946794740855694
step: 280, loss: 0.007322737947106361
step: 290, loss: 5.273654096527025e-05
step: 300, loss: 0.0005341256037354469
step: 310, loss: 0.0001607511076144874
step: 320, loss: 0.0012367385206744075
step: 330, loss: 0.000487250741571188
step: 340, loss: 0.00021796207875013351
step: 350, loss: 0.000287471164483577
step: 360, loss: 0.00018070425721816719
step: 370, loss: 4.584066118695773e-05
step: 380, loss: 0.0003336460213176906
step: 390, loss: 0.1210169792175293
step: 400, loss: 0.0006906163180246949
step: 410, loss: 9.41435937420465e-05
step: 420, loss: 0.0024384858552366495
step: 430, loss: 0.030811244621872902
step: 440, loss: 0.0009525472414679825
step: 450, loss: 0.018312877044081688
step: 460, loss: 5.6238794059026986e-05
step: 470, loss: 0.00017352370196022093
step: 480, loss: 0.00012679815699812025
step: 490, loss: 0.004097582306712866
step: 500, loss: 5.770352800027467e-05
step: 510, loss: 0.00010760006989585236
step: 520, loss: 0.002110548783093691
step: 530, loss: 0.001251010224223137
/usr/local/lib/python3.7/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=0.9407925407925408, f1=0.9304713019132057,
best_f1=0.9352914180816887
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0013861593324691057
step: 10, loss: 6.625542300753295e-05
step: 20, loss: 0.014251741580665112
step: 30, loss: 0.00014252049732021987
step: 40, loss: 8.392035670112818e-05
step: 50, loss: 3.6125635233474895e-05
step: 60, loss: 8.440203964710236e-05
step: 70, loss: 0.000940535101108253
step: 80, loss: 0.004113515838980675
step: 90, loss: 7.088040729286149e-05
step: 100, loss: 9.840691927820444e-05
step: 110, loss: 3.3392330806236714e-05
```

```
step: 120, loss: 0.00031365518225356936
step: 130, loss: 9.281449456466362e-05
step: 140, loss: 0.00021231210848782212
step: 150, loss: 0.0004812789265997708
step: 160, loss: 0.0010333404643461108
step: 170, loss: 0.006356037221848965
step: 180, loss: 0.00010095862671732903
step: 190, loss: 7.901651406427845e-05
step: 200, loss: 0.00030223626527003944
step: 210, loss: 6.974046846153215e-05
step: 220, loss: 0.00014513617497868836
step: 230, loss: 3.119832763331942e-05
step: 240, loss: 9.089826926356182e-05
step: 250, loss: 6.941175524843857e-05
step: 260, loss: 0.013012897223234177
step: 270, loss: 7.326139893848449e-05
step: 280, loss: 0.00038217316614463925
step: 290, loss: 0.025785524398088455
step: 300, loss: 6.000600478728302e-05
step: 310, loss: 0.01875491440296173
step: 320, loss: 0.022554395720362663
step: 330, loss: 0.0014916479121893644
step: 340, loss: 4.726499173557386e-05
step: 350, loss: 0.003247916465625167
step: 360, loss: 2.6545683795120567e-05
step: 370, loss: 7.541397644672543e-05
step: 380, loss: 9.575439617037773e-05
step: 390, loss: 3.7586938560707495e-05
step: 400, loss: 0.002654924290254712
step: 410, loss: 4.081217048224062e-05
step: 420, loss: 6.105812644818798e-05
step: 430, loss: 8.823150710668415e-05
step: 440, loss: 0.0006943597691133618
step: 450, loss: 0.0002344265376450494
step: 460, loss: 0.00012481538578867912
step: 470, loss: 2.367005799897015e-05
step: 480, loss: 6.403630686691031e-05
step: 490, loss: 0.0851978287100792
step: 500, loss: 0.0019950876012444496
step: 510, loss: 0.0004910128191113472
step: 520, loss: 4.268127304385416e-05
step: 530, loss: 0.002121128374710679
/usr/local/lib/python3.7/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=0.9415067852129153, f1=0.9288389513108615,
best_f1=0.9352914180816887
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
```

An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00032221493893302977
step: 10, loss: 5.262840204522945e-05
step: 20, loss: 0.00010356791608501226
step: 30, loss: 5.479238461703062e-05
step: 40, loss: 0.0018577948212623596
step: 50, loss: 2.79016021522269334e-05
step: 60, loss: 3.7571589928120375e-05
step: 70, loss: 9.830026829149574e-05
step: 80, loss: 3.307903534732759e-05
step: 90, loss: 0.0009805632289499044
step: 100, loss: 2.3025637347018346e-05
step: 110, loss: 2.2150092263473198e-05
step: 120, loss: 0.000659878016449511
step: 130, loss: 4.459768388187513e-05
step: 140, loss: 0.004928731359541416
step: 150, loss: 2.87691545963753e-05
step: 160, loss: 2.2414451450458728e-05
step: 170, loss: 0.00028872984694316983
step: 180, loss: 0.0001674913364695385
step: 190, loss: 0.00241421558894217
step: 200, loss: 0.00017656019190326333
step: 210, loss: 0.0030649900436401367
step: 220, loss: 0.00039186934009194374
step: 230, loss: 0.0004703010490629822
step: 240, loss: 9.56835356191732e-05
step: 250, loss: 6.925882189534605e-05
step: 260, loss: 0.00044181567500034541
step: 270, loss: 8.594502287451178e-05
step: 280, loss: 0.0001311350060859695
step: 290, loss: 0.00035311165265738964
step: 300, loss: 4.051040014019236e-05
step: 310, loss: 0.00017587428737897426
step: 320, loss: 0.001345415716059506
step: 330, loss: 0.0001090957157430239
step: 340, loss: 6.347798625938594e-05
step: 350, loss: 0.00041243768646381795
step: 360, loss: 9.8641132353805e-05
step: 370, loss: 0.00488103786483407
step: 380, loss: 0.0028964588418602943
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 390, loss: 0.022983413189649582
step: 400, loss: 0.00023434274771716446
step: 410, loss: 0.0029637408442795277
step: 420, loss: 9.168749238597229e-05
step: 430, loss: 0.00018935193656943738
step: 440, loss: 0.002612876007333398

```
step: 450, loss: 0.07264342159032822
step: 460, loss: 0.0001127942741732113
step: 470, loss: 0.00044436260941438377
step: 480, loss: 0.0001321997115155682
step: 490, loss: 0.01568804867565632
step: 500, loss: 0.0227727759629488
step: 510, loss: 0.0003702256944961846
step: 520, loss: 0.001465339446440339
step: 530, loss: 0.0015578370075672865
/usr/local/lib/python3.7/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=0.9396471680594244, f1=0.9327808471454879,
best_f1=0.9352914180816887
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0026948435697704554
step: 10, loss: 0.002280389191582799
step: 20, loss: 0.00015770200116094202
step: 30, loss: 7.079356873873621e-05
step: 40, loss: 0.0003314973146189004
step: 50, loss: 0.011087297461926937
step: 60, loss: 0.010978983715176582
step: 70, loss: 5.544964369619265e-05
step: 80, loss: 0.00013492839934770018
step: 90, loss: 0.0038221480790525675
step: 100, loss: 0.001370733487419784
step: 110, loss: 0.0006006488692946732
step: 120, loss: 7.2822644142434e-05
step: 130, loss: 0.00066634301389567554
step: 140, loss: 9.347878949483857e-05
step: 150, loss: 2.5942043066606857e-05
step: 160, loss: 3.4169839636888355e-05
step: 170, loss: 0.00012138004240114242
step: 180, loss: 0.0003504337801132351
step: 190, loss: 0.0016508070984855294
step: 200, loss: 0.0033031560014933348
step: 210, loss: 0.031755756586790085
step: 220, loss: 0.0029163160361349583
step: 230, loss: 3.0138888178044e-05
step: 240, loss: 0.00018724090477917343
step: 250, loss: 2.2347552658175118e-05
step: 260, loss: 0.0006202272488735616
step: 270, loss: 0.00010863073111977428
step: 280, loss: 0.001996374223381281
step: 290, loss: 0.0002615425328258425
step: 300, loss: 0.00010597814252832904
```

```
step: 310, loss: 0.00012911845988128334
step: 320, loss: 4.8217367293545976e-05
step: 330, loss: 0.0001250309287570417
step: 340, loss: 0.00015654378512408584
step: 350, loss: 9.481603774474934e-05
step: 360, loss: 0.0010961823863908648
step: 370, loss: 0.0005616115522570908
step: 380, loss: 5.23304843227379e-05
step: 390, loss: 4.084609463461675e-05
step: 400, loss: 4.684424857259728e-05
step: 410, loss: 0.00027635821606963873
step: 420, loss: 0.006102574989199638
step: 430, loss: 8.540256385458633e-05
step: 440, loss: 0.00016888520622160286
step: 450, loss: 0.00048502441495656967
step: 460, loss: 0.009083203040063381
step: 470, loss: 0.0003103544295299798
step: 480, loss: 0.0012636955361813307
step: 490, loss: 0.0014311660779640079
step: 500, loss: 0.0003022932796739042
step: 510, loss: 5.759291525464505e-05
step: 520, loss: 9.10982271307148e-05
step: 530, loss: 0.0009946102509275079
/usr/local/lib/python3.7/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=0.9432723863103609, f1=0.9316360207449317,
best_f1=0.9352914180816887
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 4.7746805648785084e-05
step: 10, loss: 2.827019852702506e-05
step: 20, loss: 0.00023561855778098106
step: 30, loss: 0.0003527898225001991
step: 40, loss: 0.0013562690000981092
step: 50, loss: 0.10336959362030029
step: 60, loss: 0.0014103177236393094
step: 70, loss: 0.00016773765673860908
step: 80, loss: 6.282466347329319e-05
step: 90, loss: 0.0001398472231812775
step: 100, loss: 4.7430352424271405e-05
step: 110, loss: 0.00015282671665772796
step: 120, loss: 2.0268833395675756e-05
step: 130, loss: 0.0017410427099093795
step: 140, loss: 0.00016864579811226577
step: 150, loss: 4.030410127597861e-05
step: 160, loss: 3.791056224144995e-05
```

```
step: 170, loss: 9.162104834103957e-05
step: 180, loss: 6.773199856979772e-05
step: 190, loss: 6.793341162847355e-05
step: 200, loss: 4.649524998967536e-05
step: 210, loss: 8.008073200471699e-05
step: 220, loss: 0.00010946614929707721
step: 230, loss: 5.689289537258446e-05
step: 240, loss: 0.000514787738211453
step: 250, loss: 3.068306250497699e-05
step: 260, loss: 5.726050221710466e-05
step: 270, loss: 4.116209311177954e-05
step: 280, loss: 0.0015138343442231417
step: 290, loss: 0.00020065391436219215
step: 300, loss: 0.00017940135148819536
step: 310, loss: 4.554582483251579e-05
step: 320, loss: 2.3152026187744923e-05
step: 330, loss: 5.184067049412988e-05
step: 340, loss: 2.8974225642741658e-05
step: 350, loss: 0.00011766857642214745
step: 360, loss: 7.684656156925485e-05
step: 370, loss: 6.869983189972118e-05
step: 380, loss: 0.0004006443778052926
step: 390, loss: 6.742229743395001e-05
step: 400, loss: 0.00011187895142938942
step: 410, loss: 0.001212661387398839
step: 420, loss: 0.005529988091439009
step: 430, loss: 0.01370376069098711
step: 440, loss: 1.8819901015376672e-05
step: 450, loss: 0.00042594512342475355
step: 460, loss: 2.264150680275634e-05
step: 470, loss: 3.22773921652697e-05
step: 480, loss: 0.008234171196818352
step: 490, loss: 0.002870891708880663
step: 500, loss: 8.609989890828729e-05
step: 510, loss: 0.0003095369029324502
step: 520, loss: 5.061526098870672e-05
step: 530, loss: 2.970814603031613e-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=0.9425393883225208, f1=0.9363086936308694,
best_f1=0.9352914180816887
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.004718996584415436
step: 10, loss: 2.0600469724740833e-05
step: 20, loss: 2.004092493734788e-05
```

```
step: 30, loss: 5.5333068303298205e-05
step: 40, loss: 2.1636058590956964e-05
step: 50, loss: 0.0004411092959344387
step: 60, loss: 2.4258342818939127e-05
step: 70, loss: 1.9806861018878408e-05
step: 80, loss: 4.252983126207255e-05
step: 90, loss: 0.015363655984401703
step: 100, loss: 0.00192410743329674
step: 110, loss: 5.475041689351201e-05
step: 120, loss: 2.7245663659414276e-05
step: 130, loss: 3.6243909562472254e-05
step: 140, loss: 3.812789873336442e-05
step: 150, loss: 0.0001527641579741612
step: 160, loss: 2.077175122394692e-05
step: 170, loss: 5.228129157330841e-05
step: 180, loss: 7.374207052635029e-05
step: 190, loss: 1.8592618289403617e-05
step: 200, loss: 2.9312290280358866e-05
step: 210, loss: 5.901730401092209e-05
step: 220, loss: 2.501413291611243e-05
step: 230, loss: 1.9494089428917505e-05
step: 240, loss: 0.0001279749267268926
step: 250, loss: 1.7925816791830584e-05
step: 260, loss: 2.5915647711372003e-05
step: 270, loss: 1.563483601785265e-05
step: 280, loss: 0.0006416037795133889
step: 290, loss: 3.272401590947993e-05
step: 300, loss: 0.00017067475710064173
step: 310, loss: 0.0003993926220573485
step: 320, loss: 0.0005518191028386354
step: 330, loss: 0.001543352147564292
step: 340, loss: 0.0001006650272756815
step: 350, loss: 2.572618905105628e-05
step: 360, loss: 2.4459368432871997e-05
step: 370, loss: 2.9518208975787275e-05
step: 380, loss: 3.7503272324102e-05
step: 390, loss: 0.0022898055613040924
step: 400, loss: 4.9019734433386475e-05
step: 410, loss: 0.0004343913751654327
step: 420, loss: 0.00012882117880508304
step: 430, loss: 2.3032802346278913e-05
step: 440, loss: 4.364532287581824e-05
step: 450, loss: 0.00037938301102258265
step: 460, loss: 0.001111416146159172
step: 470, loss: 1.4800353710597847e-05
step: 480, loss: 2.963112638099119e-05
step: 490, loss: 2.216134816990234e-05
step: 500, loss: 1.7638880308368243e-05
```

```
step: 510, loss: 3.482849933789112e-05
step: 520, loss: 0.0003348281024955213
step: 530, loss: 1.642085589992348e-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=0.9407337723424272, f1=0.926944971537002,
best_f1=0.9352914180816887
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0003245527623221278
step: 10, loss: 1.3459350157063454e-05
step: 20, loss: 1.813065682654269e-05
step: 30, loss: 2.2626742065767758e-05
step: 40, loss: 8.19938286440447e-05
step: 50, loss: 3.270694287493825e-05
step: 60, loss: 1.3708910955756437e-05
step: 70, loss: 2.033194868999999e-05
step: 80, loss: 1.8879358322010376e-05
step: 90, loss: 1.8682052541407757e-05
step: 100, loss: 9.438523557037115e-05
step: 110, loss: 3.261708843638189e-05
step: 120, loss: 0.00022101987269707024
step: 130, loss: 0.03098961152136326
step: 140, loss: 1.676686770224478e-05
step: 150, loss: 4.372306648292579e-05
step: 160, loss: 0.00012707366840913892
step: 170, loss: 0.0008588081691414118
step: 180, loss: 3.328493403387256e-05
step: 190, loss: 2.1255698811728507e-05
step: 200, loss: 7.798159640515223e-05
step: 210, loss: 1.7717236914904788e-05
step: 220, loss: 1.7624019164941274e-05
step: 230, loss: 2.7063346351496875e-05
Gradient overflow.  Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 240, loss: 1.5754036212456413e-05
step: 250, loss: 2.9231061489554122e-05
step: 260, loss: 2.6947122023557313e-05
step: 270, loss: 1.936742410180159e-05
step: 280, loss: 1.425280515832128e-05
step: 290, loss: 3.1294512155000012e-05
step: 300, loss: 3.915305933333002e-05
step: 310, loss: 3.333609492983669e-05
step: 320, loss: 7.241106504807249e-05
step: 330, loss: 0.00022635261120740324
step: 340, loss: 6.170944834593683e-05
step: 350, loss: 3.978548920713365e-05
```

```
step: 360, loss: 6.817837856942788e-05
step: 370, loss: 2.0853569367318414e-05
step: 380, loss: 8.418483776040375e-05
step: 390, loss: 0.0028893568087369204
step: 400, loss: 0.00017294203280471265
step: 410, loss: 5.3389518143376336e-05
step: 420, loss: 5.0564183766255155e-05
step: 430, loss: 0.00034710083855316043
step: 440, loss: 6.116409349488094e-05
step: 450, loss: 3.33464631694369e-05
step: 460, loss: 1.8194003132521175e-05
step: 470, loss: 2.4444336304441094e-05
step: 480, loss: 2.2217109290068038e-05
step: 490, loss: 6.085087079554796e-05
step: 500, loss: 1.3176226275390945e-05
step: 510, loss: 4.5487813622457907e-05
step: 520, loss: 1.965053343155887e-05
step: 530, loss: 5.42418347322382e-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=0.9414542020774315, f1=0.9248334919124643,
best_f1=0.9352914180816887
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 4.861511479248293e-05
step: 10, loss: 3.9954418753040954e-05
step: 20, loss: 4.09401509386953e-05
step: 30, loss: 0.0002514992083888501
step: 40, loss: 0.07536045461893082
step: 50, loss: 3.357307650730945e-05
step: 60, loss: 3.0232182325562462e-05
step: 70, loss: 8.880647510522977e-05
step: 80, loss: 1.8436052414472215e-05
step: 90, loss: 1.8696955521591008e-05
step: 100, loss: 0.00021628969989251345
step: 110, loss: 4.466665268409997e-05
step: 120, loss: 1.8167795133194886e-05
step: 130, loss: 0.06735361367464066
step: 140, loss: 0.00012068869546055794
step: 150, loss: 1.769490154401865e-05
step: 160, loss: 2.3743290512356907e-05
step: 170, loss: 1.9810686353594065e-05
step: 180, loss: 1.4636456398875453e-05
step: 190, loss: 2.2340080249705352e-05
step: 200, loss: 1.641707967792172e-05
step: 210, loss: 0.00019976745534222573
```

```
step: 220, loss: 9.434369712835178e-05
step: 230, loss: 3.167265822412446e-05
step: 240, loss: 2.1270867364364676e-05
step: 250, loss: 2.9555230867117643e-05
step: 260, loss: 7.939182978589088e-05
step: 270, loss: 3.222905797883868e-05
step: 280, loss: 1.706885996100027e-05
step: 290, loss: 1.3340111763682216e-05
step: 300, loss: 2.5623350666137412e-05
step: 310, loss: 2.659353776834905e-05
step: 320, loss: 0.00014113535871729255
step: 330, loss: 0.00014589977217838168
step: 340, loss: 2.698358366615139e-05
step: 350, loss: 1.3850398318027146e-05
step: 360, loss: 0.0014912665355950594
step: 370, loss: 1.4692363038193434e-05
step: 380, loss: 3.150598058709875e-05
step: 390, loss: 3.900527735822834e-05
step: 400, loss: 7.302843005163595e-05
step: 410, loss: 0.03693424165248871
step: 420, loss: 2.1661940991180018e-05
step: 430, loss: 0.000324700289638713
step: 440, loss: 3.181197098456323e-05
step: 450, loss: 1.6081798094091937e-05
step: 460, loss: 1.7389338609063998e-05
step: 470, loss: 0.004711776506155729
step: 480, loss: 1.2393936231092084e-05
step: 490, loss: 1.527723179606255e-05
step: 500, loss: 2.23696024477249e-05
step: 510, loss: 2.1043670130893588e-05
step: 520, loss: 3.156663660774939e-05
step: 530, loss: 0.000603840162511915
/usr/local/lib/python3.7/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=0.9411206077872746, f1=0.9241050119331742,
best_f1=0.9352914180816887
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 [00:29, 196.05it/s]
load_f1 = 0.943344081068632
real_f1 = 0.9415554532903819
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, 192.83it/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 bert-base-uncased \
```

```
  --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 bert-base-uncased \
  --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: 0.5723162293434143
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.4761904761904762, f1=0.3829787234042553,
best_f1=0.3829787234042553
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.5089961290359497
/usr/local/lib/python3.7/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=0.4761904761904762, f1=0.4761904761904762,
best_f1=0.3829787234042553
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.600621223449707
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.5714285714285714, f1=0.5333333333333333,
best_f1=0.5333333333333333
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.2770836651325226
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.

```
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.6250000000000001, f1=0.5454545454545454,
best_f1=0.5454545454545454
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.2899373471736908
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.6666666666666665, f1=0.5641025641025641,
best_f1=0.5641025641025641
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.22987358272075653
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.8387096774193549, f1=0.7058823529411764,
best_f1=0.7058823529411764
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.05309892073273659
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.8484848484848484, f1=0.7058823529411764,
best_f1=0.7058823529411764
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
```

```
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00888088345527649
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.8666666666666666, f1=0.7857142857142857,
best_f1=0.7857142857142857
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0068386271595954895
/usr/local/lib/python3.7/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=0.8666666666666666, f1=0.7586206896551724,
best_f1=0.7857142857142857
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.02832161635160446
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
epoch 10: dev_f1=0.8275862068965518, f1=0.7857142857142857,
best_f1=0.7857142857142857
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0015190683770924807
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.8750000000000001, f1=0.7058823529411764,
best_f1=0.7058823529411764
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.14676958322525024
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
```

```
   warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9032258064516129, f1=0.7741935483870968,
best_f1=0.7741935483870968
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
   warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0029221305157989264
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
   warnings.warn("An input tensor was not cuda.")
epoch 13: dev_f1=0.9032258064516129, f1=0.7741935483870968,
best_f1=0.7741935483870968
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
   warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0018053180538117886
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
   warnings.warn("An input tensor was not cuda.")
epoch 14: dev_f1=0.9032258064516129, f1=0.7741935483870968,
best_f1=0.7741935483870968
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
   warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0043836915865540504
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
   warnings.warn("An input tensor was not cuda.")
epoch 15: dev_f1=0.9032258064516129, f1=0.7741935483870968,
best_f1=0.7741935483870968
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.")
91it [00:00, 129427.49it/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 = 0.8235294117647058
real_f1 = 0.8
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
4398it [00:22, 196.88it/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 bert-base-uncased \
       --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 bert-base-uncased \
  --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: 0.5879495143890381
step: 10, loss: 0.6351197361946106
step: 20, loss: 0.34574073553085327
step: 30, loss: 0.11681168526411057
step: 40, loss: 0.31944727897644043
step: 50, loss: 0.05297078937292099
step: 60, loss: 0.029154961928725243
step: 70, loss: 0.036405228078365326
step: 80, loss: 0.13111251592636108
step: 90, loss: 0.1730131357908249
step: 100, loss: 0.018699346110224724
step: 110, loss: 0.15918830037117004
step: 120, loss: 0.013745211064815521
step: 130, loss: 0.006614878308027983
step: 140, loss: 0.0025431180838495493
step: 150, loss: 0.035667259246110916
step: 160, loss: 0.011252583004534245
step: 170, loss: 0.060955967754125595
step: 180, loss: 0.03254802152514458
step: 190, loss: 0.002762076910585165
step: 200, loss: 0.014281907118856907
step: 210, loss: 0.008263975381851196
step: 220, loss: 0.00136531179305166
step: 230, loss: 0.003931193146854639
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9819413092550789, 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: 0.05390576273202896
step: 10, loss: 0.0026842777151614428
step: 20, loss: 0.12898890674114227
step: 30, loss: 0.2168799340724945
step: 40, loss: 0.0233453456312418
step: 50, loss: 0.0036158678121864796
step: 60, loss: 0.0028858589939773083
step: 70, loss: 0.2202688455581665
step: 80, loss: 0.029705142602324486
```

```
step: 90, loss: 0.0036493621300905943
step: 100, loss: 0.014117994345724583
step: 110, loss: 0.08072357624769211
step: 120, loss: 0.002212034771218896
step: 130, loss: 0.009624618105590343
step: 140, loss: 0.0023933937773108482
step: 150, loss: 0.0066144089214503765
step: 160, loss: 0.024550406262278557
step: 170, loss: 0.00129384093452245
step: 180, loss: 0.01312487293034792
step: 190, loss: 0.12223276495933533
step: 200, loss: 0.00226397393271327
step: 210, loss: 0.0018056257395073771
step: 220, loss: 0.020421171560883522
step: 230, loss: 0.004712779074907303
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
  warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 2: dev_f1=0.9898762654668166, f1=0.9864559819413092,
best_f1=0.9864559819413092
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0030346377775272131
step: 10, loss: 0.02469855360686779
step: 20, loss: 0.0007814903510734439
step: 30, loss: 0.022559912875294685
step: 40, loss: 0.10054778307676315
step: 50, loss: 0.006515864748507738
step: 60, loss: 0.04003532603383064
step: 70, loss: 0.015405433252453804
step: 80, loss: 0.002343378961086273
step: 90, loss: 0.06467949599027634
step: 100, loss: 0.0008441125391982496
step: 110, loss: 0.0012339416425675154
step: 120, loss: 0.012915899977087975
step: 130, loss: 0.0009190545533783734
step: 140, loss: 0.028543950989842415
step: 150, loss: 0.0015570688992738724
step: 160, loss: 0.003814501455053687
step: 170, loss: 0.0052234698086977005
step: 180, loss: 0.009063372388482094
step: 190, loss: 0.008730482310056686
step: 200, loss: 0.003764956258237362
```

```
step: 210, loss: 0.0011029099114239216
step: 220, loss: 0.0010992359602823853
step: 230, loss: 0.0012637045001611114
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9921436588103255, 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: 0.0007234452059492469
step: 10, loss: 0.00039289050619117916
step: 20, loss: 0.00060007771007716656
step: 30, loss: 0.0013302351580932736
step: 40, loss: 0.008734439499676228
step: 50, loss: 0.0007849635439924896
step: 60, loss: 0.002537741092965007
step: 70, loss: 0.00034267493174411356
step: 80, loss: 0.0022585338447242975
step: 90, loss: 0.007754125166684389
step: 100, loss: 0.0007282645092345774
step: 110, loss: 0.0028620774392038584
step: 120, loss: 0.007108010351657867
step: 130, loss: 0.0053014010190963745
step: 140, loss: 0.0005013084155507386
step: 150, loss: 0.2019900530576706
step: 160, loss: 0.0009371594060212374
step: 170, loss: 0.14957131445407867
step: 180, loss: 0.0006661242805421352
step: 190, loss: 0.00564616871997714
step: 200, loss: 0.004561245441436768
step: 210, loss: 0.09485898166894913
step: 220, loss: 0.0006562020862475038
step: 230, loss: 0.007256180047988892
/usr/local/lib/python3.7/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=0.9853768278965129, f1=0.9783352337514253,
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: 0.0014444597763940692
```

```
step: 10, loss: 0.0010038189357146662
step: 20, loss: 0.0020126334857195616
step: 30, loss: 0.00045191129902377725
step: 40, loss: 0.00021836013183929026
step: 50, loss: 0.0001528878347016871
step: 60, loss: 0.0004475772730074823
step: 70, loss: 0.0001936507469508797
step: 80, loss: 0.009728369303047657
step: 90, loss: 0.0005792819429188967
step: 100, loss: 0.0002345589455217123
step: 110, loss: 0.0014998021069914103
step: 120, loss: 5.9447294916026294e-05
step: 130, loss: 0.0008990878704935312
step: 140, loss: 0.0015950131928548217
step: 150, loss: 0.009235890582203865
step: 160, loss: 0.0030216870363801718
step: 170, loss: 0.010244395583868027
step: 180, loss: 0.004695449955761433
step: 190, loss: 0.004693671129643917
step: 200, loss: 0.023004302754998207
step: 210, loss: 0.00028646443388424814
step: 220, loss: 0.0028451785910874605
step: 230, loss: 0.0004607686714734882
/usr/local/lib/python3.7/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=0.9876265466816648, f1=0.9852774631936579,
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: 0.00478289183229208
step: 10, loss: 0.00041232671355828643
step: 20, loss: 0.000691188732162118
step: 30, loss: 0.025773653760552406
step: 40, loss: 0.0001502961531514302
step: 50, loss: 0.00015251834702212363
step: 60, loss: 9.188438707496971e-05
step: 70, loss: 0.0036190368700772524
step: 80, loss: 0.005182295106351376
step: 90, loss: 0.0009801110718399286
step: 100, loss: 0.03518830984830856
step: 110, loss: 0.0018849615007638931
step: 120, loss: 0.00016500003403052688
step: 130, loss: 0.003767314599826932
step: 140, loss: 0.0001761463936418295
step: 150, loss: 0.0008286773227155209
step: 160, loss: 0.0078040361404418945
```

```
step: 170, loss: 0.0012574568390846252
step: 180, loss: 0.01594812609255314
step: 190, loss: 0.02565000392496586
step: 200, loss: 0.0012633062433451414
step: 210, loss: 0.003537634154781699
step: 220, loss: 0.00016266273451037705
step: 230, loss: 0.052543267607688904
/usr/local/lib/python3.7/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=0.9875706214689265, f1=0.9818181818181818,
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: 0.001972736092284322
step: 10, loss: 0.0001184842549264431
step: 20, loss: 9.681722440291196e-05
step: 30, loss: 0.0023350210394710302
step: 40, loss: 0.00013099003990646452
step: 50, loss: 0.00026476336643099785
step: 60, loss: 0.0008130475180223584
step: 70, loss: 0.004751952830702066
step: 80, loss: 8.734393486520275e-05
step: 90, loss: 7.204952999018133e-05
step: 100, loss: 7.263312727445737e-05
step: 110, loss: 0.00011355804599588737
step: 120, loss: 4.915283716400154e-05
step: 130, loss: 0.00017375661991536617
step: 140, loss: 0.00010109956201631576
step: 150, loss: 0.005143456626683474
step: 160, loss: 0.06225467100739479
step: 170, loss: 0.0008251820108853281
step: 180, loss: 0.0003327679296489805
step: 190, loss: 0.00019587329006753862
step: 200, loss: 0.02306114323437214
step: 210, loss: 0.00010224224388366565
step: 220, loss: 0.005968289915472269
step: 230, loss: 0.00015880742284934968
/usr/local/lib/python3.7/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=0.9852774631936579, f1=0.9829738933030647,
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: 0.00011551351053640246
```

```
step: 10, loss: 0.0001525240804767236
step: 20, loss: 0.00044083886314183474
step: 30, loss: 0.00013391162792686373
step: 40, loss: 0.0006051035597920418
step: 50, loss: 0.00017655329429544508
step: 60, loss: 0.0579773411154747
step: 70, loss: 0.0004746206686832011
step: 80, loss: 0.00010590430611046031
step: 90, loss: 6.993782881181687e-05
step: 100, loss: 0.0003514644631650299
step: 110, loss: 0.13076075911521912
step: 120, loss: 0.025742443278431892
step: 130, loss: 0.0056189014576375484
step: 140, loss: 0.00031807288178242743
step: 150, loss: 0.0006666792905889452
step: 160, loss: 0.00019269382755737752
step: 170, loss: 0.00024589995155110955
step: 180, loss: 0.005230267066508532
step: 190, loss: 0.00017291698895860463
step: 200, loss: 8.719740435481071e-05
step: 210, loss: 0.0001421451597707346
step: 220, loss: 0.00011356857430655509
step: 230, loss: 0.00014789898705203086
/usr/local/lib/python3.7/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=0.983277591973244, f1=0.9832402234636871,
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: 0.00011781267676269636
step: 10, loss: 0.0002889663737732917
step: 20, loss: 0.001804152736440301
step: 30, loss: 0.0002576246333774179
step: 40, loss: 0.0044439793564379215
step: 50, loss: 5.5939293815754354e-05
step: 60, loss: 6.605085945921019e-05
step: 70, loss: 0.00073156354483217
step: 80, loss: 7.002842176007107e-05
step: 90, loss: 6.305424903985113e-05
step: 100, loss: 0.00011311419075354934
step: 110, loss: 7.899488264229149e-05
step: 120, loss: 3.400357672944665e-05
step: 130, loss: 4.1892326407833025e-05
step: 140, loss: 3.7955465813865885e-05
step: 150, loss: 0.000581083295401858
step: 160, loss: 8.273313869722188e-05
```

```
step: 170, loss: 7.419603207381442e-05
step: 180, loss: 0.005658951587975025
step: 190, loss: 0.00022005170467309654
step: 200, loss: 0.04656320437788963
step: 210, loss: 0.002171424450352788
step: 220, loss: 0.00017101748380810022
step: 230, loss: 0.00017254478007089347
/usr/local/lib/python3.7/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=0.9799107142857142,
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: 0.00018628520774655044
step: 10, loss: 0.0002628426591400057
step: 20, loss: 0.00017298245802521706
step: 30, loss: 5.6910317653091624e-05
step: 40, loss: 0.00011106068996014073
step: 50, loss: 0.00013041192141827196
step: 60, loss: 0.00026726644136942923
Gradient overflow.  Skipping step, loss scaler 0 reducing loss scale to 32768.0
step: 70, loss: 0.00019825121853500605
step: 80, loss: 3.824257873930037e-05
step: 90, loss: 0.0005536496755667031
step: 100, loss: 6.830380880273879e-05
step: 110, loss: 0.0012217385228723288
step: 120, loss: 0.00433831661939621
step: 130, loss: 5.6154221965698525e-05
step: 140, loss: 0.04454032704234123
step: 150, loss: 0.011202300898730755
step: 160, loss: 3.275924609624781e-05
step: 170, loss: 9.294466144638136e-05
step: 180, loss: 0.0005862741963937879
step: 190, loss: 0.014473901130259037
step: 200, loss: 0.00017339135229121894
step: 210, loss: 3.3626936783548445e-05
step: 220, loss: 4.263439404894598e-05
step: 230, loss: 0.0005508980248123407
/usr/local/lib/python3.7/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=0.9866071428571428, f1=0.9821826280623607,
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: 3.778023165068589e-05
step: 10, loss: 0.00039337313501164317
step: 20, loss: 0.0035738544538617134
step: 30, loss: 0.020904377102851868
step: 40, loss: 0.00037730945041403174
step: 50, loss: 4.256975807948038e-05
step: 60, loss: 0.0009115960565395653
step: 70, loss: 0.0002498266112525016
step: 80, loss: 0.0001226866152137518
step: 90, loss: 0.0009565692162141204
step: 100, loss: 0.0006287145661190152
step: 110, loss: 0.04843176528811455
step: 120, loss: 6.134423892945051e-05
step: 130, loss: 0.0002642635372467339
step: 140, loss: 0.006916780024766922
step: 150, loss: 0.013471723534166813
step: 160, loss: 0.00012671838339883834
step: 170, loss: 0.019948724657297134
step: 180, loss: 0.0005352067528292537
step: 190, loss: 4.224723670631647e-05
step: 200, loss: 0.0004216574889142066
step: 210, loss: 4.9603931984165683e-05
step: 220, loss: 4.8471538320882246e-05
step: 230, loss: 0.0021268706768751144
/usr/local/lib/python3.7/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=0.9855072463768116, f1=0.978865406006674,
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: 0.00020585661695804447
step: 10, loss: 0.00023934507044032216
step: 20, loss: 5.889394014957361e-05
step: 30, loss: 5.8766003348864615e-05
step: 40, loss: 0.00037561325007118285
step: 50, loss: 0.0003129302931483835
step: 60, loss: 0.0002905502915382385
step: 70, loss: 7.248880865518004e-05
step: 80, loss: 0.00014954968355596066
step: 90, loss: 3.10198993247468e-05
step: 100, loss: 7.298479613382369e-05
step: 110, loss: 0.00013770622899755836
step: 120, loss: 5.515721932169981e-05
step: 130, loss: 4.6750035835430026e-05
step: 140, loss: 0.011258103884756565
step: 150, loss: 0.000302895117783919
```

```
step: 160, loss: 3.574703077902086e-05
step: 170, loss: 0.0001980039814952761
step: 180, loss: 0.0209537111222744
step: 190, loss: 4.795128552359529e-05
step: 200, loss: 4.379410529509187e-05
step: 210, loss: 0.00010554441541898996
step: 220, loss: 0.006266613956540823
step: 230, loss: 0.03709025681018829
/usr/local/lib/python3.7/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=0.9865771812080537, f1=0.9831649831649831,
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: 8.772046567173675e-05
step: 10, loss: 0.001109635573811829
step: 20, loss: 0.0020912804175168276
step: 30, loss: 6.580150511581451e-05
step: 40, loss: 9.25034619285725e-05
step: 50, loss: 0.00011046246800106019
step: 60, loss: 8.630973024992272e-05
step: 70, loss: 3.544343780959025e-05
step: 80, loss: 5.596923438133672e-05
step: 90, loss: 5.9567744756350294e-05
step: 100, loss: 2.638949081301689e-05
step: 110, loss: 0.002330099930986762
step: 120, loss: 0.050241775810718536
step: 130, loss: 7.576151983812451e-05
step: 140, loss: 3.4167438570875674e-05
step: 150, loss: 4.0007842471823096e-05
step: 160, loss: 0.00012075588892912492
step: 170, loss: 0.00010935167083516717
step: 180, loss: 0.00011847605492221192
step: 190, loss: 2.9589289624709636e-05
step: 200, loss: 0.0009296130738221109
step: 210, loss: 2.5212273612851277e-05
step: 220, loss: 0.00015661866927985102
step: 230, loss: 7.428721437463537e-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=0.987709497206704, f1=0.9820627802690582,
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: 3.351912528160028e-05
step: 10, loss: 0.005523880943655968
step: 20, loss: 0.0001470683782827109
step: 30, loss: 0.00935161393135786
step: 40, loss: 0.007014417555183172
step: 50, loss: 6.116246368037537e-05
step: 60, loss: 5.0615195505088195e-05
step: 70, loss: 3.358611138537526e-05
step: 80, loss: 8.215112757170573e-05
step: 90, loss: 4.36465852544643e-05
step: 100, loss: 4.032059950986877e-05
step: 110, loss: 6.38954879832454e-05
step: 120, loss: 2.0511153707047924e-05
step: 130, loss: 3.280014789197594e-05
step: 140, loss: 0.00027953696553595364
step: 150, loss: 3.978308814112097e-05
step: 160, loss: 0.0064516691491007805
step: 170, loss: 8.293409337056801e-05
step: 180, loss: 0.00010312608355889097
step: 190, loss: 3.660708898678422e-05
step: 200, loss: 0.00011502976121846586
step: 210, loss: 5.502076237462461e-05
step: 220, loss: 0.0008735916344448924
step: 230, loss: 0.002164522185921669
/usr/local/lib/python3.7/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=0.987709497206704, f1=0.9832026875699889,
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: 4.86369026475586e-05
step: 10, loss: 2.5402232495252974e-05
step: 20, loss: 3.88124244636856e-05
step: 30, loss: 3.7026722566224635e-05
step: 40, loss: 0.00014920382818672806
step: 50, loss: 0.03518601879477501
step: 60, loss: 0.00013637264783028513
step: 70, loss: 0.0002053125062957406
step: 80, loss: 7.307058695005253e-05
step: 90, loss: 0.00011112997162854299
step: 100, loss: 0.00026511639589443803
step: 110, loss: 4.054691089550033e-05
step: 120, loss: 0.0002516602980904281
step: 130, loss: 0.02910393476486206
step: 140, loss: 2.2686699594487436e-05
step: 150, loss: 0.00025340908905491233
```

```
step: 160, loss: 6.941657193237916e-05
step: 170, loss: 8.181881275959313e-05
step: 180, loss: 0.0002829797158483416
step: 190, loss: 0.0002808517892844975
step: 200, loss: 0.00013480654160957783
step: 210, loss: 7.520983490394428e-05
step: 220, loss: 0.0002244346251245588
step: 230, loss: 3.4207954740850255e-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=0.9866071428571428, f1=0.9820627802690582,
best_f1=0.984304932735426
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.")
2473it [00:14, 168.08it/s]
load_f1 = 0.9932432432432432
real_f1 = 0.9921259842519685
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, 190.60it/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 bert-base-uncased \
       --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 bert-base-uncased \
       --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: 0.6056830286979675
step: 10, loss: 0.48671266436576843
step: 20, loss: 0.4952046573162079
step: 30, loss: 0.1341300904750824
step: 40, loss: 0.05080261826515198
step: 50, loss: 0.30849355459213257
step: 60, loss: 0.06354303658008575
step: 70, loss: 0.2153526246547699
step: 80, loss: 0.1531357616186142
step: 90, loss: 0.2993803322315216
step: 100, loss: 0.029736565425992012
step: 110, loss: 0.07863210141658783
step: 120, loss: 0.08997026085853577
step: 130, loss: 0.024150926619768143
step: 140, loss: 0.033066023141145706
step: 150, loss: 0.04951201751828194
step: 160, loss: 0.030431905761361122
step: 170, loss: 0.17827516794204712
step: 180, loss: 0.08608818799257278
```

```
step: 190, loss: 0.012509978376328945
step: 200, loss: 0.136866495013237
step: 210, loss: 0.08008138090372086
step: 220, loss: 0.22274905443191528
step: 230, loss: 0.14563053846359253
step: 240, loss: 0.01800447888672352
step: 250, loss: 0.0491865836083889
step: 260, loss: 0.08452793210744858
step: 270, loss: 0.10682818293571472
Gradient overflow.  Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 280, loss: 0.08151289820671082
step: 290, loss: 0.20138494670391083
step: 300, loss: 0.035791803151369095
step: 310, loss: 0.24357402324676514
step: 320, loss: 0.04698223993182182
step: 330, loss: 0.016436053439974785
step: 340, loss: 0.0616389624774456
step: 350, loss: 0.1640530526638031
step: 360, loss: 0.0801292136311531
step: 370, loss: 0.06537435948848724
step: 380, loss: 0.0780545324087143
step: 390, loss: 0.19008587300777435
step: 400, loss: 0.1979454755783081
step: 410, loss: 0.04874361306428909
step: 420, loss: 0.03949306532740593
step: 430, loss: 0.21747484803199768
step: 440, loss: 0.04117783159017563
step: 450, loss: 0.006181924603879452
step: 460, loss: 0.02059437893331051
step: 470, loss: 0.10012863576412201
step: 480, loss: 0.059720829129219055
step: 490, loss: 0.16258442401885986
step: 500, loss: 0.11362351477146149
step: 510, loss: 0.04616708308458328
step: 520, loss: 0.06633046269416809
step: 530, loss: 0.0022200376261025667
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9311475409836065, f1=0.9335192933519293,
best_f1=0.9335192933519293
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
```

```
step: 0, loss: 0.32736778259277344
step: 10, loss: 0.09538831561803818
step: 20, loss: 0.011863930150866508
step: 30, loss: 0.018558889627456665
step: 40, loss: 0.057668011635541916
step: 50, loss: 0.06914232671260834
step: 60, loss: 0.08681558817625046
step: 70, loss: 0.1029915139079094
step: 80, loss: 0.04197295382618904
step: 90, loss: 0.010253914631903172
step: 100, loss: 0.020336942747235298
step: 110, loss: 0.015646900981664658
step: 120, loss: 0.15157704055309296
step: 130, loss: 0.11721634864807129
step: 140, loss: 0.019785122945904732
step: 150, loss: 0.1392667442560196
step: 160, loss: 0.007501333951950073
step: 170, loss: 0.00579499639570713
step: 180, loss: 0.025112604722380638
step: 190, loss: 0.07412412017583847
step: 200, loss: 0.009859699755907059
step: 210, loss: 0.07333612442016602
step: 220, loss: 0.14755812287330627
step: 230, loss: 0.09406350553035736
step: 240, loss: 0.1404649317264557
step: 250, loss: 0.14598813652992249
step: 260, loss: 0.0071792397648096085
step: 270, loss: 0.11122774332761765
step: 280, loss: 0.015152291394770145
step: 290, loss: 0.02878561057150364
step: 300, loss: 0.19770193099975586
step: 310, loss: 0.0129060298204422
step: 320, loss: 0.11531262844800949
step: 330, loss: 0.08798030763864517
step: 340, loss: 0.052827879786491394
step: 350, loss: 0.0013317351695150137
step: 360, loss: 0.011479396373033524
step: 370, loss: 0.15175282955169678
step: 380, loss: 0.0427955687046051
step: 390, loss: 0.055422160774469376
step: 400, loss: 0.06953930854797363
step: 410, loss: 0.007388134021311998
step: 420, loss: 0.03590783104300499
step: 430, loss: 0.01305961050093174
step: 440, loss: 0.15576881170272827
step: 450, loss: 0.058814775198698044
step: 460, loss: 0.016009507700800896
step: 470, loss: 0.131790891289711
```

```
step: 480, loss: 0.276548832654953
step: 490, loss: 0.01302926056087017
step: 500, loss: 0.2997358441352844
step: 510, loss: 0.016481023281812668
step: 520, loss: 0.05668371170759201
step: 530, loss: 0.014469383284449577
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9354243542435424, f1=0.9327188940092166,
best_f1=0.9327188940092166
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.24939970672130585
step: 10, loss: 0.04198877885937691
step: 20, loss: 0.03865250572562218
step: 30, loss: 0.09467437863349915
step: 40, loss: 0.06603863835334778
step: 50, loss: 0.051603931933641434
step: 60, loss: 0.0285358689725399
step: 70, loss: 0.00424532312899828
step: 80, loss: 0.00333752972073853
step: 90, loss: 0.020661812275648117
step: 100, loss: 0.03500435873866081
step: 110, loss: 0.0013908512191846967
step: 120, loss: 0.017931632697582245
step: 130, loss: 0.006471480708569288
step: 140, loss: 0.009537650272250175
step: 150, loss: 0.02164044976234436
step: 160, loss: 0.011805471032857895
step: 170, loss: 0.07213246077299118
step: 180, loss: 0.007356804795563221
step: 190, loss: 0.011235922574996948
step: 200, loss: 0.038233090192079544
step: 210, loss: 0.08814413100481033
step: 220, loss: 0.09466797858476639
step: 230, loss: 0.0759652853012085
step: 240, loss: 0.023069854825735092
step: 250, loss: 0.029205821454524994
step: 260, loss: 0.017059680074453354
step: 270, loss: 0.0043665338307619095
step: 280, loss: 0.15108883380889893
step: 290, loss: 0.0032393834553658962
```

```
step: 300, loss: 0.020099008455872536
step: 310, loss: 0.00988102052360773
step: 320, loss: 0.013665435835719109
step: 330, loss: 0.003982939291745424
step: 340, loss: 0.0034012198448181152
step: 350, loss: 0.006805736105889082
step: 360, loss: 0.045689214020967484
step: 370, loss: 0.003684328170493245
step: 380, loss: 0.012462570331990719
step: 390, loss: 0.025184009224176407
step: 400, loss: 0.008088530972599983
step: 410, loss: 0.004237615503370762
step: 420, loss: 0.1432313621044159
step: 430, loss: 0.03685395419597626
step: 440, loss: 0.005988053046166897
step: 450, loss: 0.06118442490696907
step: 460, loss: 0.028526952490210533
step: 470, loss: 0.04922507330775261
step: 480, loss: 0.02219744771718979
step: 490, loss: 0.02312634326517582
step: 500, loss: 0.00931494403630495
step: 510, loss: 0.016796208918094635
step: 520, loss: 0.12684382498264313
step: 530, loss: 0.024367153644561768
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9391955098222639, f1=0.926829268292683,
best_f1=0.926829268292683
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.009339220821857452
step: 10, loss: 0.0044335490092635155
step: 20, loss: 0.026616381481289864
step: 30, loss: 0.0007198153180070221
step: 40, loss: 0.038019414991140366
step: 50, loss: 0.004984982777386904
step: 60, loss: 0.018056239932775497
step: 70, loss: 0.0869821086525917
step: 80, loss: 0.0007868855609558523
step: 90, loss: 0.020868385210633278
step: 100, loss: 0.02394157275557518
step: 110, loss: 0.013768717646598816
```

```
step: 120, loss: 0.033447034657001495
step: 130, loss: 0.0031491422560065985
step: 140, loss: 0.0010262442519888282
step: 150, loss: 0.017861701548099518
step: 160, loss: 0.051754847168922424
step: 170, loss: 0.0031844917684793472
step: 180, loss: 0.004009251948446035
step: 190, loss: 0.006798724178224802
step: 200, loss: 0.014395911246538162
step: 210, loss: 0.0033675245940685272
step: 220, loss: 0.003952420316636562
step: 230, loss: 0.04662406072020531
step: 240, loss: 0.001200301107019186
step: 250, loss: 0.003226139582693577
step: 260, loss: 0.0007537920610047877
step: 270, loss: 0.005812426097691059
step: 280, loss: 0.09957852959632874
step: 290, loss: 0.035617027431726456
step: 300, loss: 0.0020054348278790712
step: 310, loss: 0.007574854884296656
step: 320, loss: 0.02399243600666523
step: 330, loss: 0.06971930712461472
step: 340, loss: 0.00099464925006032
step: 350, loss: 0.1418742686510086
step: 360, loss: 0.05192306637763977
step: 370, loss: 0.01131319161504507
step: 380, loss: 0.037814319133758545
step: 390, loss: 0.0632757917046547
step: 400, loss: 0.011770457029342651
step: 410, loss: 0.003290473949164152
step: 420, loss: 0.010031045414507389
step: 430, loss: 0.0875239148736
step: 440, loss: 0.02174782194197178
step: 450, loss: 0.050848253071308136
step: 460, loss: 0.0018623630749061704
step: 470, loss: 0.0013745836913585663
step: 480, loss: 0.09186716377735138
step: 490, loss: 0.010472786612808704
step: 500, loss: 0.017956597730517387
step: 510, loss: 0.005510191898792982
step: 520, loss: 0.02494591660797596
step: 530, loss: 0.047882165759801865
/usr/local/lib/python3.7/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=0.932666060054595, f1=0.9307201458523247,
best_f1=0.926829268292683
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
```

```
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.03794964775443077
step: 10, loss: 0.027283180505037308
step: 20, loss: 0.01942852884531021
step: 30, loss: 0.02383052371442318
step: 40, loss: 0.08990051597356796
step: 50, loss: 0.010679371654987335
step: 60, loss: 0.13901618123054504
step: 70, loss: 0.014845027588307858
step: 80, loss: 0.02747073769569397
step: 90, loss: 0.05422923341393471
step: 100, loss: 0.011852667666971684
step: 110, loss: 0.0013614208437502384
step: 120, loss: 0.006379642523825169
step: 130, loss: 0.0005686875665560365
step: 140, loss: 0.004598767962306738
step: 150, loss: 0.0019551070872694254
step: 160, loss: 0.0003935680433642119
step: 170, loss: 0.017736496403813362
step: 180, loss: 0.0007363775512203574
step: 190, loss: 0.000932585506234318
step: 200, loss: 0.0009043471654877067
step: 210, loss: 0.018693163990974426
step: 220, loss: 0.0018205762607976794
step: 230, loss: 0.005784440785646439
step: 240, loss: 0.0012880938593298197
step: 250, loss: 0.0044695232063531876
step: 260, loss: 0.002189900726079941
step: 270, loss: 0.00024354978813789785
step: 280, loss: 0.004171317908912897
step: 290, loss: 0.07628636062145233
step: 300, loss: 0.003184887347742915
step: 310, loss: 0.00403132289648056
step: 320, loss: 0.023923086002469063
step: 330, loss: 0.10124138742685318
step: 340, loss: 0.009599590674042702
step: 350, loss: 0.0027733941096812487
step: 360, loss: 0.01418119017034769
step: 370, loss: 0.028549090027809143
step: 380, loss: 0.0012097564758732915
step: 390, loss: 0.000647296488750726
step: 400, loss: 0.010681627318263054
step: 410, loss: 0.008631756529211998
step: 420, loss: 0.005814466159790754
step: 430, loss: 0.00045634451089426875
step: 440, loss: 0.024229271337389946
step: 450, loss: 0.0020521690603345633
```

```
step: 460, loss: 0.09543853998184204
step: 470, loss: 0.00687940651550889
step: 480, loss: 0.019907433539628983
step: 490, loss: 0.0019678995013237
step: 500, loss: 0.0037279590032994747
step: 510, loss: 0.06164151430130005
step: 520, loss: 0.03935135155916214
step: 530, loss: 0.0009094922570511699
/usr/local/lib/python3.7/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=0.9367441860465117, f1=0.9299303944315545,
best_f1=0.926829268292683
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.005554516799747944
step: 10, loss: 0.00227854442415624857
step: 20, loss: 0.12328599393367767
step: 30, loss: 0.06185663864016533
step: 40, loss: 0.0005983130540698767
step: 50, loss: 0.02599221281707287
step: 60, loss: 0.00013047936954535544
step: 70, loss: 0.001372210681438446
step: 80, loss: 0.03025827184319496
step: 90, loss: 0.006505743134766817
step: 100, loss: 0.005143646150827408
step: 110, loss: 0.04206559807062149
step: 120, loss: 0.002096751006320119
step: 130, loss: 0.007199229672551155
step: 140, loss: 0.06778635829687119
step: 150, loss: 0.0011361652286723256
step: 160, loss: 0.012381902895867825
step: 170, loss: 0.0019018049351871014
step: 180, loss: 0.0011807094560936093
step: 190, loss: 0.02156616933643818
step: 200, loss: 0.04314490780234337
step: 210, loss: 0.02092619799077511
step: 220, loss: 0.018607251346111298
step: 230, loss: 0.1001490131020546
step: 240, loss: 0.07113561034202576
step: 250, loss: 0.0034308135509490967
step: 260, loss: 0.0023680238518863916
step: 270, loss: 0.0781107023358345
step: 280, loss: 0.0012781384866684675
step: 290, loss: 0.00021697735064662993
step: 300, loss: 0.010161800310015678
step: 310, loss: 0.00044439564226195216
```

```
step: 320, loss: 0.00048000813694670796
step: 330, loss: 0.00014029645535629243
step: 340, loss: 0.0956144705414772
step: 350, loss: 0.0009134876308962703
step: 360, loss: 0.022119618952274323
step: 370, loss: 0.02253863774240017
step: 380, loss: 0.00020596652757376432
step: 390, loss: 0.04438285157084465
step: 400, loss: 0.0026242416352033615
step: 410, loss: 0.0006402177968993783
step: 420, loss: 0.0017397957853972912
step: 430, loss: 0.0012692735763266683
step: 440, loss: 0.0015166024677455425
step: 450, loss: 0.0010082758963108063
step: 460, loss: 0.000631962320767343
step: 470, loss: 0.002874253084883094
step: 480, loss: 0.0030456294771283865
step: 490, loss: 0.009398948401212692
step: 500, loss: 0.0007880909834057093
step: 510, loss: 0.0016222366830334067
step: 520, loss: 0.006323203910142183
step: 530, loss: 0.02933003380894661
/usr/local/lib/python3.7/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=0.9352112676056337, f1=0.9321561338289963,
best_f1=0.926829268292683
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0006959354504942894
step: 10, loss: 0.007610516157001257
step: 20, loss: 0.001968431519344449
step: 30, loss: 0.0029490762390196323
step: 40, loss: 0.00034719990799203515
step: 50, loss: 0.012215007096529007
step: 60, loss: 0.002982890233397484
step: 70, loss: 0.00021638930775225163
step: 80, loss: 0.0019849399104714394
step: 90, loss: 0.0052072107791900635
step: 100, loss: 0.006848773919045925
step: 110, loss: 0.0016440856270492077
step: 120, loss: 0.0078095123171806335
step: 130, loss: 0.003304634476080537
step: 140, loss: 0.0018163374625146639
step: 150, loss: 0.0005451479810290039
step: 160, loss: 0.00012470530054997653
step: 170, loss: 0.00034823984606191516
```

```
step: 180, loss: 0.0001621696283109486
step: 190, loss: 0.023840848356485367
step: 200, loss: 0.00045185245107859373
step: 210, loss: 0.0006272773025557399
step: 220, loss: 0.005370244849473238
step: 230, loss: 0.035916637629270554
step: 240, loss: 0.0009105698554776609
step: 250, loss: 0.006196146830916405
step: 260, loss: 0.00014676939463242888
step: 270, loss: 0.0008699442842043936
step: 280, loss: 0.0001353724510408938
step: 290, loss: 0.001302648219279945
step: 300, loss: 0.003075928892940283
step: 310, loss: 0.0003206967667210847
step: 320, loss: 0.0010070214048027992
step: 330, loss: 0.004830007441341877
step: 340, loss: 0.03550759702920914
step: 350, loss: 0.0009281316306442022
step: 360, loss: 0.0014218565775081515
step: 370, loss: 0.004361363127827644
step: 380, loss: 0.0003060217131860554
step: 390, loss: 0.0021163499914109707
step: 400, loss: 0.00869528204202652
step: 410, loss: 0.0006794519140385091
step: 420, loss: 0.0009551741532050073
step: 430, loss: 0.0004713804228231311
step: 440, loss: 0.019640401005744934
step: 450, loss: 0.001385033130645752
step: 460, loss: 0.000963413855060935
step: 470, loss: 0.006307045929133892
step: 480, loss: 0.0025783188175410032
step: 490, loss: 0.00039238433237187564
step: 500, loss: 0.00029780069598928094
step: 510, loss: 0.0021114563569426537
step: 520, loss: 0.04169868305325508
step: 530, loss: 0.13045839965343475
/usr/local/lib/python3.7/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=0.9324137931034482, f1=0.930909090909091,
best_f1=0.926829268292683
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.041670061647892
step: 10, loss: 0.024370282888412476
step: 20, loss: 0.0015955112176015973
step: 30, loss: 0.00016007987142074853
```

```
step: 40, loss: 0.005397573113441467
step: 50, loss: 0.006050364580005407
step: 60, loss: 0.002183083910495043
step: 70, loss: 0.00016071554273366928
step: 80, loss: 0.004260585177689791
step: 90, loss: 0.0012740340316668153
step: 100, loss: 0.00014489333261735737
step: 110, loss: 0.0063695297576487064
step: 120, loss: 0.03587467968463898
step: 130, loss: 0.003174409968778491
step: 140, loss: 0.011510745622217655
step: 150, loss: 0.00045720700290985405
step: 160, loss: 0.04115695506334305
step: 170, loss: 0.005289482418447733
step: 180, loss: 0.0014852641616016626
step: 190, loss: 0.0005209124065004289
step: 200, loss: 0.0004472793370950967
step: 210, loss: 0.009370808489620686
step: 220, loss: 0.0007408653036691248
step: 230, loss: 0.0006232518353499472
step: 240, loss: 0.004827433731406927
step: 250, loss: 0.002376715885475278
step: 260, loss: 6.332679186016321e-05
step: 270, loss: 0.004792178049683571
step: 280, loss: 0.02085241861641407
step: 290, loss: 0.0012903622118756175
step: 300, loss: 0.008436402305960655
step: 310, loss: 0.037063054740428925
step: 320, loss: 0.01869329810142517
step: 330, loss: 0.18807344138622284
step: 340, loss: 0.188223734498024
step: 350, loss: 0.0010119216749444604
step: 360, loss: 0.018057463690638542
step: 370, loss: 0.00042739370837807655
step: 380, loss: 0.0022502674255520105
step: 390, loss: 0.10012997686862946
step: 400, loss: 0.0004541263042483479
step: 410, loss: 0.021019993349909782
step: 420, loss: 0.003241005819290876
step: 430, loss: 0.07804747670888901
step: 440, loss: 0.014534824527800083
step: 450, loss: 0.00013980931544210762
step: 460, loss: 0.0004035850870423019
step: 470, loss: 0.0011897203512489796
step: 480, loss: 0.017784271389245987
step: 490, loss: 0.005054904147982597
step: 500, loss: 0.021266059949994087
step: 510, loss: 0.011286921799182892
```

```
Gradient overflow.  Skipping step, loss scaler 0 reducing loss scale to 32768.0
step: 520, loss: 0.006900493521243334
step: 530, loss: 0.0007926359539851546
/usr/local/lib/python3.7/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=0.929550827423168, f1=0.9286726499763817,
best_f1=0.926829268292683
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.007915862835943699
step: 10, loss: 0.0009961959440261126
step: 20, loss: 0.0026381940115243196
step: 30, loss: 0.0006350145558826625
step: 40, loss: 0.003929448779672384
step: 50, loss: 0.0002756406320258975
step: 60, loss: 0.0004148985317442566
step: 70, loss: 0.006808914244174957
step: 80, loss: 0.012327347882091999
step: 90, loss: 5.5249282013392076e-05
step: 100, loss: 4.616181831806898e-05
step: 110, loss: 0.0003137844614684582
step: 120, loss: 0.0006865108734928071
step: 130, loss: 0.010897754691541195
step: 140, loss: 0.009136009961366653
step: 150, loss: 0.00023914290068205446
step: 160, loss: 0.001908357604406774
step: 170, loss: 0.006786716636270285
step: 180, loss: 0.000269969052169472
step: 190, loss: 0.022734306752681732
step: 200, loss: 0.0034798167180269957
step: 210, loss: 0.0006051189266145229
step: 220, loss: 0.0018625186057761312
step: 230, loss: 7.19817981007509e-05
step: 240, loss: 0.007040295749902725
step: 250, loss: 0.0004813862906303257
step: 260, loss: 0.0405578576028347
step: 270, loss: 3.0326687920023687e-05
step: 280, loss: 0.037007737904787064
step: 290, loss: 0.03841273486614227
step: 300, loss: 0.00011258501763222739
step: 310, loss: 0.024701951071619987
step: 320, loss: 0.020824836567044258
step: 330, loss: 0.0003915833367500454
step: 340, loss: 0.005898700095713139
step: 350, loss: 0.003520638681948185
step: 360, loss: 0.00444834806680679
```

```
step: 370, loss: 0.012839782051742077
step: 380, loss: 0.0001455462770536542
step: 390, loss: 0.00020196553668938577
step: 400, loss: 0.0005841788370162249
step: 410, loss: 7.58079913794063e-05
step: 420, loss: 4.7822271881159395e-05
step: 430, loss: 0.00011907513544429094
step: 440, loss: 0.003221531631425023
step: 450, loss: 0.00017780107737053186
step: 460, loss: 0.0003522511979099363
step: 470, loss: 4.801083923666738e-05
step: 480, loss: 0.000520434114150703
step: 490, loss: 0.003524093423038721
step: 500, loss: 0.00648772157728672
step: 510, loss: 0.00024727245909161866
step: 520, loss: 0.010837799869477749
step: 530, loss: 0.00036568002542480826
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9394785847299814, f1=0.9340710004610421,
best_f1=0.9340710004610421
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 4.454058580449782e-05
step: 10, loss: 4.421252015163191e-05
step: 20, loss: 6.364652654156089e-05
step: 30, loss: 0.00013531475269701332
step: 40, loss: 0.0024958932772278786
step: 50, loss: 0.0005191202508285642
step: 60, loss: 0.0009859109995886683
step: 70, loss: 2.6601521312841214e-05
step: 80, loss: 3.400949208298698e-05
step: 90, loss: 0.0023721216712146997
step: 100, loss: 9.588438842911273e-05
step: 110, loss: 0.00010919304622802883
step: 120, loss: 0.00024070648942142725
step: 130, loss: 0.000521952984854579
step: 140, loss: 0.01036748941987753
step: 150, loss: 0.0015949145890772343
step: 160, loss: 2.8817916245316155e-05
step: 170, loss: 5.8981720940209925e-05
step: 180, loss: 0.00023081515973899513
```

```
step: 190, loss: 0.000544873415492475
step: 200, loss: 0.00017304430366493762
step: 210, loss: 0.0011860972736030817
step: 220, loss: 0.009278766810894012
step: 230, loss: 0.0001313849788857624
step: 240, loss: 0.01182801928371191
step: 250, loss: 3.005020334967412e-05
step: 260, loss: 0.0010305749019607902
step: 270, loss: 0.0002448589075356722
step: 280, loss: 0.00032830191776156425
step: 290, loss: 2.936202145065181e-05
step: 300, loss: 4.784147677128203e-05
step: 310, loss: 0.00010681013372959569
step: 320, loss: 0.0017713407287374139
step: 330, loss: 0.00010952635784633458
step: 340, loss: 0.17442962527275085
step: 350, loss: 0.0010495488531887531
step: 360, loss: 0.0002038375532720238
step: 370, loss: 0.007093470543622971
step: 380, loss: 0.0027806952130049467
step: 390, loss: 0.0009079980663955212
step: 400, loss: 0.0013860510662198067
step: 410, loss: 0.0011048997985199094
step: 420, loss: 0.008816234767436981
step: 430, loss: 0.0006162190111353993
step: 440, loss: 0.012708866968750954
step: 450, loss: 0.00011291760165477172
step: 460, loss: 0.0005043551209382713
step: 470, loss: 0.001075942418538034
step: 480, loss: 0.00021091099188197404
step: 490, loss: 0.0009229364222846925
step: 500, loss: 0.09813597798347473
step: 510, loss: 0.0002661156468093395
step: 520, loss: 0.004409634508192539
step: 530, loss: 0.0007620699470862746
/usr/local/lib/python3.7/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=0.9348335677449601, f1=0.9333333333333333,
best_f1=0.9340710004610421
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.007697161752730608
step: 10, loss: 0.06860776990652084
step: 20, loss: 0.000189512618817389
step: 30, loss: 0.0013449442340061069
step: 40, loss: 0.0013458606554195285
```

```
step: 50, loss: 0.0522882379591465
step: 60, loss: 0.027614498510956764
step: 70, loss: 0.00012275883636903018
step: 80, loss: 0.00011720779002644122
step: 90, loss: 0.0004550693556666374
step: 100, loss: 0.0006398643599823117
step: 110, loss: 0.002220471389591694
step: 120, loss: 7.472754805348814e-05
step: 130, loss: 7.898345938883722e-05
step: 140, loss: 0.07773122936487198
step: 150, loss: 0.0002799921203404665
step: 160, loss: 0.0040560816414654255
step: 170, loss: 0.0005378416972234845
step: 180, loss: 0.0004696075338870287
step: 190, loss: 0.0013175783678889275
step: 200, loss: 0.005740666761994362
step: 210, loss: 0.004590925760567188
step: 220, loss: 0.004136829171329737
step: 230, loss: 6.024260073900223e-05
step: 240, loss: 0.0001245315943378955
step: 250, loss: 3.889131039613858e-05
step: 260, loss: 0.0007925268146209419
step: 270, loss: 0.0002533411607146263
step: 280, loss: 0.0001901772920973599
step: 290, loss: 0.00018152742995880544
step: 300, loss: 0.0345175601541996
step: 310, loss: 0.00026636122493073344
step: 320, loss: 0.00020057686197105795
step: 330, loss: 0.0013436791487038136
step: 340, loss: 0.00012542912736535072
step: 350, loss: 0.0021855824161320925
step: 360, loss: 0.0004995348863303661
step: 370, loss: 0.009601721540093422
step: 380, loss: 4.183123746770434e-05
step: 390, loss: 0.00010121092782355845
step: 400, loss: 6.525596108986065e-05
step: 410, loss: 0.0003382716968189925
step: 420, loss: 0.0005807711859233677
step: 430, loss: 4.433613139553927e-05
step: 440, loss: 0.00013685862359125167
step: 450, loss: 6.406997999874875e-05
step: 460, loss: 0.0037528902757912874
step: 470, loss: 3.6857913073617965e-05
step: 480, loss: 0.0006023449823260307
step: 490, loss: 0.0005421192036010325
step: 500, loss: 5.544985833694227e-05
step: 510, loss: 0.00024409320030827075
step: 520, loss: 0.0002741527569014579
```

```
step: 530, loss: 4.3056763388449326e-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 11: dev_f1=0.9350163627863488, f1=0.926829268292683,
best_f1=0.9340710004610421
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 7.435472798533738e-05
step: 10, loss: 5.26384828845039e-05
step: 20, loss: 7.65760923968628e-05
step: 30, loss: 0.004512839950621128
step: 40, loss: 0.0002439456438878551
step: 50, loss: 0.021371006965637207
step: 60, loss: 0.011648290790617466
step: 70, loss: 0.0009974446147680283
Gradient overflow.  Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 80, loss: 5.675432475982234e-05
step: 90, loss: 0.000160598021466285
step: 100, loss: 5.742390567320399e-05
step: 110, loss: 0.00020124389266129583
step: 120, loss: 2.4679438865859993e-05
step: 130, loss: 5.195089761400595e-05
step: 140, loss: 2.0514620700851083e-05
step: 150, loss: 1.8868282495532185e-05
step: 160, loss: 3.989645483670756e-05
step: 170, loss: 0.00012409345072228462
step: 180, loss: 2.848939402610995e-05
step: 190, loss: 7.310553337447345e-05
step: 200, loss: 4.031500793644227e-05
step: 210, loss: 6.766356091247872e-05
step: 220, loss: 0.00010378483420936391
step: 230, loss: 2.5189863663399592e-05
step: 240, loss: 0.06353951245546341
step: 250, loss: 8.698959572939202e-05
step: 260, loss: 0.0002729384577833116
step: 270, loss: 0.05188968777656555
step: 280, loss: 0.0003495760611258447
step: 290, loss: 0.017027784138917923
step: 300, loss: 0.009394156746566296
step: 310, loss: 0.0002947361790575087
step: 320, loss: 0.0007806906360201538
step: 330, loss: 0.0006880760192871094
step: 340, loss: 5.739837069995701e-05
step: 350, loss: 0.005129473749548197
step: 360, loss: 0.004941971506923437
step: 370, loss: 0.0025847756769508123
```

```
step: 380, loss: 0.00041380972834303975
step: 390, loss: 0.00028513651341199875
step: 400, loss: 0.12568283081054688
step: 410, loss: 0.004356415942311287
step: 420, loss: 0.014074563048779964
step: 430, loss: 0.01524368766695261
step: 440, loss: 0.0013096253387629986
step: 450, loss: 0.0007511820294894278
step: 460, loss: 0.0031034129206091166
step: 470, loss: 0.0006226947880350053
step: 480, loss: 0.002003444591537118
step: 490, loss: 0.0007466258830390871
step: 500, loss: 0.000148570048622787
step: 510, loss: 0.017333460971713066
step: 520, loss: 0.0010805519996210933
step: 530, loss: 0.0004632060881704092
/usr/local/lib/python3.7/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=0.9330855018587362, f1=0.9276830953477659,
best_f1=0.9340710004610421
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00035129438037984073
step: 10, loss: 0.006358546670526266
step: 20, loss: 0.01981649361550808
step: 30, loss: 0.00016375964332837611
step: 40, loss: 0.0001306662888964638
step: 50, loss: 0.0016697155078873038
step: 60, loss: 0.0008342606597580016
step: 70, loss: 9.239177597919479e-05
step: 80, loss: 0.0004056835896335542
step: 90, loss: 0.03287338837981224
step: 100, loss: 0.008627488277852535
step: 110, loss: 0.0029804708901792765
step: 120, loss: 0.0006515515269711614
step: 130, loss: 0.00017377604672219604
step: 140, loss: 0.0002780131471809
step: 150, loss: 0.0005648695514537394
step: 160, loss: 0.004382043145596981
step: 170, loss: 0.00026393510052002966
step: 180, loss: 0.0003787569294217974
step: 190, loss: 0.00012455953401513398
step: 200, loss: 3.3467047614976764e-05
step: 210, loss: 5.910408799536526e-05
step: 220, loss: 0.00015867926413193345
step: 230, loss: 0.0004118185315746814
```

```
step: 240, loss: 0.03134601190686226
step: 250, loss: 3.0139999580569565e-05
step: 260, loss: 4.963453102391213e-05
step: 270, loss: 2.194517946918495e-05
step: 280, loss: 0.000356975884642452
step: 290, loss: 0.0003779747348744422
step: 300, loss: 2.9551845727837645e-05
step: 310, loss: 0.003247002372518182
step: 320, loss: 0.0016312624793499708
step: 330, loss: 0.0020500710234045982
step: 340, loss: 0.004532813560217619
step: 350, loss: 0.01653623767197132
step: 360, loss: 0.0009276469936594367
step: 370, loss: 0.00036204984644427896
step: 380, loss: 0.0013401854084804654
step: 390, loss: 0.041708238422870636
step: 400, loss: 0.00029365590307861567
step: 410, loss: 0.000869989744387567
step: 420, loss: 6.568984099430963e-05
step: 430, loss: 0.00879593100398779
step: 440, loss: 0.00021004687005188316
step: 450, loss: 0.0013452075654640794
step: 460, loss: 0.0026928421575576067
step: 470, loss: 5.6705743190832436e-05
step: 480, loss: 0.00018218497280031443
step: 490, loss: 0.001124494126997888
step: 500, loss: 0.0010373125551268458
step: 510, loss: 0.0004455716407392174
step: 520, loss: 0.005759123247116804
step: 530, loss: 5.523742220248096e-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=0.9345794392523366, f1=0.9312267657992565,
best_f1=0.9340710004610421
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00017421635857317597
step: 10, loss: 3.2352756534237415e-05
step: 20, loss: 0.0002985569299198687
step: 30, loss: 0.00012598819739650935
step: 40, loss: 0.00017096014926210046
step: 50, loss: 0.00023264587798621505
step: 60, loss: 6.622139335377142e-05
step: 70, loss: 0.00024054590903688222
step: 80, loss: 5.024009442422539e-05
step: 90, loss: 2.428444531688001e-05
```

```
step: 100, loss: 0.00037381291622295976
step: 110, loss: 5.6779084843583405e-05
step: 120, loss: 0.000296086713206023
step: 130, loss: 0.0404512956738472
step: 140, loss: 0.0009816192323341966
step: 150, loss: 0.0007836492732167244
step: 160, loss: 0.0005533905932679772
step: 170, loss: 0.0003464118344709277
step: 180, loss: 5.7293891586596146e-05
step: 190, loss: 0.000621766725089401
step: 200, loss: 0.000789186917245388
step: 210, loss: 3.5249879147158936e-05
step: 220, loss: 0.002676336793228984
step: 230, loss: 0.002437853952869773
step: 240, loss: 0.00012750955647788942
step: 250, loss: 8.299916225951165e-05
step: 260, loss: 0.0003866908955387771
step: 270, loss: 0.0001239673438249156
step: 280, loss: 4.5017288357485086e-05
step: 290, loss: 0.000180298593477346
step: 300, loss: 0.0007676705718040466
step: 310, loss: 0.054982613772153854
step: 320, loss: 0.024825716391205788
step: 330, loss: 0.000978928990662098
step: 340, loss: 0.0008342408109456301
step: 350, loss: 0.00021806989389006048
step: 360, loss: 0.08189328014850616
step: 370, loss: 2.3226642952067778e-05
step: 380, loss: 4.980304220225662e-05
step: 390, loss: 0.026337746530771255
step: 400, loss: 0.0007920251809991896
step: 410, loss: 4.129300941713154e-05
step: 420, loss: 0.00012182573118479922
step: 430, loss: 8.701248589204624e-05
step: 440, loss: 0.0026046892162412405
step: 450, loss: 0.0007194186910055578
step: 460, loss: 0.001530562760308385
step: 470, loss: 6.401701830327511e-05
step: 480, loss: 0.0003941963950637728
step: 490, loss: 0.0008849675068631768
step: 500, loss: 5.48542921023909e-05
step: 510, loss: 0.0010149864247068763
step: 520, loss: 0.0005220783059485257
step: 530, loss: 7.683315925532952e-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=0.937528921795465, f1=0.9306384933394579,
```

```
best_f1=0.9340710004610421
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 3.359236507094465e-05
step: 10, loss: 3.200231367372908e-05
step: 20, loss: 0.0010074661113321781
step: 30, loss: 0.002411931985989213
step: 40, loss: 0.05208873003721237
step: 50, loss: 0.0689588338136673
step: 60, loss: 5.935559966019355e-05
step: 70, loss: 0.0019506418611854315
step: 80, loss: 0.00037348348996602
step: 90, loss: 2.8102529540774412e-05
step: 100, loss: 4.9941252655511684e-05
step: 110, loss: 0.012945085763931274
step: 120, loss: 0.006390544585883617
step: 130, loss: 0.11271315813064575
step: 140, loss: 0.0002422169636702165
step: 150, loss: 4.550907760858536e-05
step: 160, loss: 0.00010635854414431378
step: 170, loss: 7.659800758119673e-05
step: 180, loss: 4.4533953769132495e-05
step: 190, loss: 0.00031806787592358887
step: 200, loss: 7.309420470846817e-05
step: 210, loss: 0.000795798609033227
step: 220, loss: 0.037366122007369995
step: 230, loss: 0.001104669296182692
step: 240, loss: 0.0034633981995284557
step: 250, loss: 0.00032319052843376994
step: 260, loss: 0.0002137532428605482
step: 270, loss: 2.9966895453981124e-05
step: 280, loss: 0.007704245392233133
step: 290, loss: 2.0037889044033363e-05
step: 300, loss: 2.000797030632384e-05
step: 310, loss: 0.00025842711329460144
step: 320, loss: 0.0008973596268333495
step: 330, loss: 0.00016521039651706815
step: 340, loss: 7.065061072353274e-05
step: 350, loss: 5.497105303220451e-05
step: 360, loss: 0.000584669120144099
step: 370, loss: 9.089898958336562e-05
step: 380, loss: 5.349602724891156e-05
step: 390, loss: 0.00011833666212623939
step: 400, loss: 4.6865508920745924e-05
step: 410, loss: 0.00013759502326138318
step: 420, loss: 0.0009303619153797626
step: 430, loss: 0.00025126218679361045
```

```
step: 440, loss: 0.0005013871123082936
step: 450, loss: 5.844051338499412e-05
step: 460, loss: 6.379274418577552e-05
step: 470, loss: 0.003371267346665263
step: 480, loss: 0.0002341497311135754
step: 490, loss: 2.3472270186175592e-05
step: 500, loss: 0.00038282445166260004
step: 510, loss: 0.001078105764463544
step: 520, loss: 4.2990741349058226e-05
step: 530, loss: 0.00031127355759963393
/usr/local/lib/python3.7/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=0.938134810710988, f1=0.9327231121281464,
best_f1=0.9340710004610421
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 [00:31, 182.09it/s]
load_f1 = 0.9342723004694835
real_f1 = 0.9341486359360301
```

```
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, 182.38it/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 bert-base-uncased \
  --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 bert-base-uncased \
  --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: 0.5652856826782227
step: 10, loss: 0.3907673954963684
step: 20, loss: 0.3769378960132599
step: 30, loss: 0.2921670079231262
step: 40, loss: 0.15990251302719116
step: 50, loss: 0.34373199939727783
step: 60, loss: 0.13302618265151978
step: 70, loss: 0.19179973006248474
step: 80, loss: 0.22902096807956696
step: 90, loss: 0.3585101366043091
step: 100, loss: 0.5478050708770752
step: 110, loss: 0.265102744102478
step: 120, loss: 0.38448017835617065
step: 130, loss: 0.20309247076511383
step: 140, loss: 0.1883535534143448
step: 150, loss: 0.16657309234142303
step: 160, loss: 0.25122445821762085
```

```
step: 170, loss: 0.2537182569503784
step: 180, loss: 0.11301770806312561
step: 190, loss: 0.17340192198753357
step: 200, loss: 0.19866417348384857
step: 210, loss: 0.2054128795862198
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.6484375, f1=0.6769825918762089, best_f1=0.6769825918762089
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.09086306393146515
step: 10, loss: 0.26500314474105835
step: 20, loss: 0.21525613963603973
step: 30, loss: 0.17670148611068726
step: 40, loss: 0.15084326267242432
step: 50, loss: 0.12640506029129028
step: 60, loss: 0.4048025906085968
step: 70, loss: 0.14677786827087402
step: 80, loss: 0.14890938997268677
step: 90, loss: 0.07325001806020737
step: 100, loss: 0.012609010562300682
step: 110, loss: 0.1326940804719925
step: 120, loss: 0.194444939494133
step: 130, loss: 0.04046957567334175
step: 140, loss: 0.1825299859046936
step: 150, loss: 0.24761606752872467
step: 160, loss: 0.16477146744728088
step: 170, loss: 0.11513444781303406
step: 180, loss: 0.3088395297527313
step: 190, loss: 0.17683729529380798
step: 200, loss: 0.05713687092065811
step: 210, loss: 0.198047935962677
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.6875, f1=0.6918238993710691, best_f1=0.6918238993710691
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
```

```
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.05261028930544853
step: 10, loss: 0.1586970090866089
step: 20, loss: 0.1735595166683197
step: 30, loss: 0.15817150473594666
step: 40, loss: 0.058478400111198425
step: 50, loss: 0.05877547711133957
step: 60, loss: 0.24301321804523468
step: 70, loss: 0.030456680804491043
step: 80, loss: 0.12822431325912476
step: 90, loss: 0.04621618986129761
step: 100, loss: 0.1479964554309845
step: 110, loss: 0.08152537047863007
step: 120, loss: 0.18420270085334778
step: 130, loss: 0.21009789407253265
step: 140, loss: 0.1012144461274147
step: 150, loss: 0.2573927342891693
step: 160, loss: 0.01280668843537569
step: 170, loss: 0.10003358125686646
step: 180, loss: 0.054406650364398956
step: 190, loss: 0.19407247006893158
step: 200, loss: 0.0963604524731636
step: 210, loss: 0.14001429080963135
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.6963249516441006, f1=0.7137254901960784,
best_f1=0.7137254901960784
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.09790296852588654
step: 10, loss: 0.08643648773431778
step: 20, loss: 0.21335488557815552
step: 30, loss: 0.10321828722953796
step: 40, loss: 0.026657624170184135
step: 50, loss: 0.14033056795597076
step: 60, loss: 0.28737935423851013
step: 70, loss: 0.17332793772220612
step: 80, loss: 0.045533012598752975
step: 90, loss: 0.03182852268218994
step: 100, loss: 0.19412310421466827
step: 110, loss: 0.09091690182685852
step: 120, loss: 0.14129887521266937
```

```
step: 130, loss: 0.25271955132484436
step: 140, loss: 0.09889674186706543
step: 150, loss: 0.06235896050930023
step: 160, loss: 0.01631062850356102
step: 170, loss: 0.22615444660186768
step: 180, loss: 0.5887908935546875
step: 190, loss: 0.11085732281208038
step: 200, loss: 0.1552671641111374
step: 210, loss: 0.299483060836792
/usr/local/lib/python3.7/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=0.6694915254237288, f1=0.6877637130801688,
best_f1=0.7137254901960784
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.13726630806922913
step: 10, loss: 0.09424791485071182
step: 20, loss: 0.15198804438114166
step: 30, loss: 0.022287892177700996
step: 40, loss: 0.2959010601043701
step: 50, loss: 0.0672507956624031
step: 60, loss: 0.08596711605787277
step: 70, loss: 0.09874825179576874
step: 80, loss: 0.017470940947532654
step: 90, loss: 0.20592668652534485
step: 100, loss: 0.003996573388576508
step: 110, loss: 0.11664550006389618
step: 120, loss: 0.09889265894889832
step: 130, loss: 0.07428630441427231
step: 140, loss: 0.04242611303925514
step: 150, loss: 0.03543508052825928
step: 160, loss: 0.22040468454360962
step: 170, loss: 0.055949967354536057
step: 180, loss: 0.11412177234888077
step: 190, loss: 0.06814837455749512
step: 200, loss: 0.083650141954422
step: 210, loss: 0.02408229373395443
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.696798493408663, f1=0.701492537313433,
best_f1=0.701492537313433
```

```
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.03935646638274193
step: 10, loss: 0.04559030756354332
step: 20, loss: 0.058024339377880096
step: 30, loss: 0.002080114558339119
step: 40, loss: 0.023844629526138306
step: 50, loss: 0.006660402752459049
step: 60, loss: 0.30838829278945923
step: 70, loss: 0.0500839538872242
step: 80, loss: 0.2286207377910614
step: 90, loss: 0.08929130434989929
step: 100, loss: 0.03042975254356861
step: 110, loss: 0.1036205068230629
step: 120, loss: 0.1324075311422348
step: 130, loss: 0.11004579812288284
step: 140, loss: 0.13713814318180084
step: 150, loss: 0.007469921838492155
step: 160, loss: 0.0035563965793699026
step: 170, loss: 0.14765632152557373
step: 180, loss: 0.02587031200528145
step: 190, loss: 0.19166824221611023
step: 200, loss: 0.004672127775847912
step: 210, loss: 0.042625393718481064
/usr/local/lib/python3.7/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=0.674698795180723, f1=0.6707818930041152,
best_f1=0.701492537313433
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.005826172884553671
step: 10, loss: 0.05333897843956947
step: 20, loss: 0.026358995586633682
step: 30, loss: 0.03513123095035553
step: 40, loss: 0.07039960473775864
step: 50, loss: 0.052981700748205185
step: 60, loss: 0.0398930124938488
step: 70, loss: 0.023088403046131134
step: 80, loss: 0.09701993316411972
step: 90, loss: 0.0462910495698452
step: 100, loss: 0.0062019783072173595
step: 110, loss: 0.15396836400032043
step: 120, loss: 0.26065900921821594
step: 130, loss: 0.03413809463381767
step: 140, loss: 0.004312231671065092
```

```
step: 150, loss: 0.0109795751005011463
step: 160, loss: 0.03449005261063576
step: 170, loss: 0.1275259256362915
step: 180, loss: 0.030738260596990585
step: 190, loss: 0.13738718628883362
step: 200, loss: 0.010047446936368942
step: 210, loss: 0.06365258246660233
/usr/local/lib/python3.7/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=0.6613545816733067, f1=0.6653306613226453,
best_f1=0.701492537313433
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.044730622321367264
step: 10, loss: 0.08082811534404755
step: 20, loss: 0.1056201383471489
step: 30, loss: 0.11674336344003677
step: 40, loss: 0.008841478265821934
step: 50, loss: 0.01088839489966631
step: 60, loss: 0.1199108436703682
step: 70, loss: 0.011371787637472153
step: 80, loss: 0.026220355182886124
step: 90, loss: 0.003658856963738799
step: 100, loss: 0.08172465115785599
step: 110, loss: 0.055743258446455
step: 120, loss: 0.006054713856428862
step: 130, loss: 0.012135591357946396
step: 140, loss: 0.06058010086417198
step: 150, loss: 0.023784618824720383
step: 160, loss: 0.038437824696302414
step: 170, loss: 0.029849810525774956
step: 180, loss: 0.05508897826075554
step: 190, loss: 0.010166436433792114
step: 200, loss: 0.002368057845160365
step: 210, loss: 0.22949719429016113
/usr/local/lib/python3.7/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=0.6680000000000001, f1=0.6787878787878788,
best_f1=0.701492537313433
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.012740237638354301
step: 10, loss: 0.09815485030412674
step: 20, loss: 0.006413740571588278
```

```
step: 30, loss: 0.023754987865686417
step: 40, loss: 0.06885318458080292
step: 50, loss: 0.009688016027212143
step: 60, loss: 0.26880231499671936
step: 70, loss: 0.010935651138424873
step: 80, loss: 0.006376605946570635
step: 90, loss: 0.004723094869405031
step: 100, loss: 0.02867664210498333
step: 110, loss: 0.01682164892554283
step: 120, loss: 0.010623996146023273
step: 130, loss: 0.007983023300766945
step: 140, loss: 0.07327984273433685
step: 150, loss: 0.0581275038421154
step: 160, loss: 0.016786864027380943
step: 170, loss: 0.011582969687879086
step: 180, loss: 0.043652232736349106
step: 190, loss: 0.00483017647638917
step: 200, loss: 0.08683383464813232
step: 210, loss: 0.03628423810005188
/usr/local/lib/python3.7/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=0.6651982378854626, f1=0.6546275395033859,
best_f1=0.701492537313433
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.03515547513961792
step: 10, loss: 0.07486793398857117
step: 20, loss: 0.0010720898862928152
step: 30, loss: 0.09961601346731186
step: 40, loss: 0.00861481111496687
step: 50, loss: 0.016121286898851395
step: 60, loss: 0.08905156701803207
step: 70, loss: 0.208916574716568
step: 80, loss: 0.026269666850566864
step: 90, loss: 0.20811666548252106
step: 100, loss: 0.12106245011091232
step: 110, loss: 0.01711735688149929
step: 120, loss: 0.009282526560127735
step: 130, loss: 0.026314876973628998
step: 140, loss: 0.03674235939979553
step: 150, loss: 0.04782969877123833
step: 160, loss: 0.010045405477285385
step: 170, loss: 0.006759746000170708
step: 180, loss: 0.13227058947086334
step: 190, loss: 0.07479646056890488
step: 200, loss: 0.0111766289919614
```

step: 210, loss: 0.06805715709924698
/usr/local/lib/python3.7/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=0.6826568265682657, f1=0.6593001841620626,
best_f1=0.701492537313433
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.018836595118045807
step: 10, loss: 0.14545203745365143
step: 20, loss: 0.1078914999961853
step: 30, loss: 0.001704772817902267
step: 40, loss: 0.028851358219981194
step: 50, loss: 0.00787155982106924
step: 60, loss: 0.0031019444577395916
step: 70, loss: 0.004058328922837973
step: 80, loss: 0.05911487713456154
step: 90, loss: 0.05999954417347908
step: 100, loss: 0.018253155052661896
step: 110, loss: 0.2092432975769043
step: 120, loss: 0.19366838037967682
step: 130, loss: 0.00787840411067009
step: 140, loss: 0.002275738399475813
step: 150, loss: 0.0026144967414438725
step: 160, loss: 0.01036666426807642
step: 170, loss: 0.004344361834228039
step: 180, loss: 0.0031032655388116837
step: 190, loss: 0.019318653270602226
step: 200, loss: 0.0010882215574383736
step: 210, loss: 0.007829640060663223
/usr/local/lib/python3.7/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=0.6710526315789473, f1=0.662280701754386,
best_f1=0.701492537313433
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.15441584587097168
step: 10, loss: 0.00041464975220151246
step: 20, loss: 0.042608827352523804
step: 30, loss: 0.003674035659059882
step: 40, loss: 0.06078360974788666
step: 50, loss: 0.0013467412209138274
step: 60, loss: 0.06792419403791428
step: 70, loss: 0.0007953917374834418
step: 80, loss: 0.02109525352716446

```
step: 90, loss: 0.044288598001003265
step: 100, loss: 0.019009245559573174
step: 110, loss: 0.00826225895434618
step: 120, loss: 0.006801344454288483
step: 130, loss: 0.013258401304483414
step: 140, loss: 0.04606947675347328
Gradient overflow.  Skipping step, loss scaler 0 reducing loss scale to 32768.0
step: 150, loss: 0.020366529002785683
step: 160, loss: 0.012261942960321903
step: 170, loss: 0.005278879776597023
step: 180, loss: 0.015064878389239311
step: 190, loss: 0.12608259916305542
step: 200, loss: 0.008177515119314194
step: 210, loss: 0.028264304623007774
/usr/local/lib/python3.7/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=0.6653992395437263, f1=0.6577437858508605,
best_f1=0.701492537313433
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.05513427406549454
step: 10, loss: 0.007641611620783806
step: 20, loss: 0.01275221910327673
step: 30, loss: 0.08130696415901184
step: 40, loss: 0.09095196425914764
step: 50, loss: 0.0017986692255362868
step: 60, loss: 0.047360360622406006
step: 70, loss: 0.04620056971907616
step: 80, loss: 0.02392730675637722
step: 90, loss: 0.002611400792375207
step: 100, loss: 0.011087877675890923
step: 110, loss: 0.014775658957660198
step: 120, loss: 0.003154212376102805
step: 130, loss: 0.0010525044053792953
step: 140, loss: 0.051223449409008026
step: 150, loss: 0.021846050396561623
step: 160, loss: 0.003623134922236204
step: 170, loss: 0.011365135200321674
step: 180, loss: 0.028846630826592445
step: 190, loss: 0.004128532949835062
step: 200, loss: 0.012836672365665436
step: 210, loss: 0.03666891157627106
/usr/local/lib/python3.7/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=0.679324894514768, f1=0.6767895878524947,
```

```
best_f1=0.701492537313433
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0008016815991140902
step: 10, loss: 0.02651875466108322
step: 20, loss: 0.001996008912101388
step: 30, loss: 0.0029966384172439575
step: 40, loss: 0.08377201110124588
step: 50, loss: 0.05492750555276871
step: 60, loss: 0.009741355665028095
step: 70, loss: 0.0030334703624248505
step: 80, loss: 0.0012430183123797178
step: 90, loss: 0.03278113529086113
step: 100, loss: 0.0035315747372806072
step: 110, loss: 0.011633730493485928
step: 120, loss: 0.04734884202480316
step: 130, loss: 0.052552852779626846
step: 140, loss: 0.01889384165406227
step: 150, loss: 0.00045076131937094033
step: 160, loss: 0.008613495156168938
step: 170, loss: 0.0011528534814715385
step: 180, loss: 0.020728429779410362
step: 190, loss: 0.003079934511333704
step: 200, loss: 0.020796209573745728
step: 210, loss: 0.027434833347797394
/usr/local/lib/python3.7/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=0.6779661016949152, f1=0.6724890829694323,
best_f1=0.701492537313433
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0018847546307370067
step: 10, loss: 0.0004362166509963572
step: 20, loss: 0.0022657299414277077
step: 30, loss: 0.16841654479503632
step: 40, loss: 0.0017538656247779727
step: 50, loss: 0.0014327747048810124
step: 60, loss: 0.03601112589240074
step: 70, loss: 0.0027955954428762197
step: 80, loss: 0.0022851969115436077
step: 90, loss: 0.0191498976200819
step: 100, loss: 0.004589024931192398
step: 110, loss: 0.00031852288520894945
step: 120, loss: 0.008587832562625408
step: 130, loss: 0.06624113023281097
```

```
step: 140, loss: 0.00039658925379626453
step: 150, loss: 0.0022045923396945
step: 160, loss: 0.008546956814825535
step: 170, loss: 0.0006160775083117187
step: 180, loss: 0.0007602993282489479
step: 190, loss: 0.051203321665525436
step: 200, loss: 0.005925463046878576
step: 210, loss: 0.010023328475654125
/usr/local/lib/python3.7/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=0.6720647773279351, f1=0.673469387755102,
best_f1=0.701492537313433
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.")
2293it [00:08, 278.96it/s]
load_f1 = 0.6946107784431138
real_f1 = 0.6956521739130436
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
```

```
    warnings.warn("An input tensor was not cuda.")
4398it [00:24, 178.58it/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 bert-base-uncased \
       --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 bert-base-uncased \
       --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: 0.5782621502876282
step: 10, loss: 0.354979544878006
step: 20, loss: 0.2731013000011444
step: 30, loss: 0.43319621682167053
step: 40, loss: 0.43185487389564514
step: 50, loss: 0.29107028245925903
step: 60, loss: 0.25781431794166565
step: 70, loss: 0.29815489053726196
step: 80, loss: 0.26085326075553894
step: 90, loss: 0.2669897973537445
step: 100, loss: 0.30330130457878113
step: 110, loss: 0.34197261929512024
step: 120, loss: 0.09649531543254852
step: 130, loss: 0.09927815943956375
step: 140, loss: 0.06911297142505646
step: 150, loss: 0.24752597510814667
step: 160, loss: 0.1067902147769928
step: 170, loss: 0.2119292914867401
step: 180, loss: 0.03482069820165634
```

```
step: 190, loss: 0.1881684958934784
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.6733668341708543, 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: 0.2998875379562378
step: 10, loss: 0.044975169003009796
step: 20, loss: 0.12249886989593506
step: 30, loss: 0.2880551815032959
step: 40, loss: 0.14970912039279938
step: 50, loss: 0.07896122336387634
step: 60, loss: 0.24194645881652832
step: 70, loss: 0.1769617795944214
step: 80, loss: 0.18135465681552887
step: 90, loss: 0.3552413582801819
step: 100, loss: 0.012356016784906387
step: 110, loss: 0.060385413467884064
step: 120, loss: 0.27190616726875305
step: 130, loss: 0.06648913025856018
step: 140, loss: 0.04197143763303757
step: 150, loss: 0.09708297252655029
step: 160, loss: 0.029514294117689133
step: 170, loss: 0.12260895222425461
step: 180, loss: 0.22709667682647705
step: 190, loss: 0.02003883197903633
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.8010610079575596, f1=0.7896103896103895,
best_f1=0.7896103896103895
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.03526846319437027
step: 10, loss: 0.0501880906522274
step: 20, loss: 0.09451814740896225
```

```
step: 30, loss: 0.017887573689222336
step: 40, loss: 0.07958859950304031
step: 50, loss: 0.24643279612064362
step: 60, loss: 0.003871069522574544
step: 70, loss: 0.06910131871700287
step: 80, loss: 0.015190882608294487
step: 90, loss: 0.10757546871900558
step: 100, loss: 0.05823846906423569
step: 110, loss: 0.1395512968301773
step: 120, loss: 0.07640766352415085
step: 130, loss: 0.008134079165756702
step: 140, loss: 0.013965423218905926
step: 150, loss: 0.10463521629571915
step: 160, loss: 0.03618304431438446
step: 170, loss: 0.07111161947250366
step: 180, loss: 0.009630517102777958
step: 190, loss: 0.12567846477031708
/usr/local/lib/python3.7/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=0.7934782608695652, f1=0.8169761273209548,
best_f1=0.7896103896103895
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.03068314865231514
step: 10, loss: 0.13601000607013702
step: 20, loss: 0.04649369791150093
step: 30, loss: 0.011709353886544704
step: 40, loss: 0.034868620336055756
step: 50, loss: 0.023539040237665176
step: 60, loss: 0.10473117977380753
step: 70, loss: 0.3529830574989319
step: 80, loss: 0.20885398983955383
step: 90, loss: 0.010671332478523254
step: 100, loss: 0.11645162105560303
step: 110, loss: 0.04253696650266647
step: 120, loss: 0.09452421963214874
step: 130, loss: 0.07309078425168991
step: 140, loss: 0.0878283903002739
step: 150, loss: 0.12723170220851898
step: 160, loss: 0.0066381958313286304
step: 170, loss: 0.0693470910191536
step: 180, loss: 0.01544719934463501
step: 190, loss: 0.17501147091388702
/usr/local/lib/python3.7/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=0.7704918032786885, f1=0.7786666666666667,
best_f1=0.7896103896103895
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.1806756854057312
step: 10, loss: 0.013895973563194275
step: 20, loss: 0.046953556264710426
step: 30, loss: 0.00511049572378397
step: 40, loss: 0.06158332899212837
step: 50, loss: 0.1047038659453392
step: 60, loss: 0.11234328150749207
step: 70, loss: 0.019676096737384796
step: 80, loss: 0.013144111260771751
step: 90, loss: 0.019158923998475075
step: 100, loss: 0.005917536094784737
step: 110, loss: 0.009909010492265224
step: 120, loss: 0.0037628142163157463
step: 130, loss: 0.11420723795890808
step: 140, loss: 0.05832366645336151
step: 150, loss: 0.04800385981798172
step: 160, loss: 0.011142592877149582
step: 170, loss: 0.0041717663407325745
Gradient overflow.  Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 180, loss: 0.12532973289489746
Gradient overflow.  Skipping step, loss scaler 0 reducing loss scale to 8192.0
step: 190, loss: 0.03594290092587471
/usr/local/lib/python3.7/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=0.7913279132791328, f1=0.8191489361702128,
best_f1=0.7896103896103895
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.026769591495394707
step: 10, loss: 0.15794119238853455
step: 20, loss: 0.03943765535950661
step: 30, loss: 0.0019378849538043141
step: 40, loss: 0.056069422513246536
step: 50, loss: 0.004031339660286903
step: 60, loss: 0.03272628411650658
step: 70, loss: 0.004784054588526487
step: 80, loss: 0.020683182403445244
step: 90, loss: 0.018508248031139374
step: 100, loss: 0.0014023496769368649
step: 110, loss: 0.1263478547334671
step: 120, loss: 0.014199376106262207
```

```
step: 130, loss: 0.008415773510932922
step: 140, loss: 0.0455123595893383
step: 150, loss: 0.002087977947667241
step: 160, loss: 0.035590920597314835
step: 170, loss: 0.0270572267472744
step: 180, loss: 0.02049342915415764
step: 190, loss: 0.013333415612578392
/usr/local/lib/python3.7/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=0.7780821917808218, f1=0.7862796833773087,
best_f1=0.7896103896103895
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0032642255537211895
step: 10, loss: 0.01004622969776392
step: 20, loss: 0.17812985181808472
step: 30, loss: 0.0057601542212069035
step: 40, loss: 0.006415836047381163
step: 50, loss: 0.011516694910824299
step: 60, loss: 0.0058692279271781445
step: 70, loss: 0.0029511938337236643
step: 80, loss: 0.02931848168373108
step: 90, loss: 0.008123956620693207
step: 100, loss: 0.0010464643128216267
step: 110, loss: 0.050682276487350464
step: 120, loss: 0.0023597029503434896
step: 130, loss: 0.001045609824359417
step: 140, loss: 0.0009542639018036425
step: 150, loss: 0.011891862377524376
step: 160, loss: 0.036210767924785614
step: 170, loss: 0.14013753831386566
step: 180, loss: 0.015934107825160027
step: 190, loss: 0.002941161161288619
/usr/local/lib/python3.7/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=0.7828418230563003, f1=0.7795698924731183,
best_f1=0.7896103896103895
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0026064130943268538
step: 10, loss: 0.049032747745513916
step: 20, loss: 0.004423834849148989
step: 30, loss: 0.09288257360458374
step: 40, loss: 0.01563139632344246
```

```
step: 50, loss: 0.10040686279535294
step: 60, loss: 0.06735201925039291
step: 70, loss: 0.005646057426929474
step: 80, loss: 0.0007263813749887049
step: 90, loss: 0.0006727800937369466
step: 100, loss: 0.027194317430257797
step: 110, loss: 0.008846556767821312
step: 120, loss: 0.004384610336273909
step: 130, loss: 0.0005325052770785987
step: 140, loss: 0.004664693959057331
step: 150, loss: 0.007151535712182522
step: 160, loss: 0.0005596785922534764
step: 170, loss: 0.0008721629856154323
step: 180, loss: 0.004303382709622383
step: 190, loss: 0.22234420478343964
/usr/local/lib/python3.7/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=0.7830687830687831, f1=0.7780678851174935,
best_f1=0.7896103896103895
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0010084420209750533
step: 10, loss: 0.07388052344322205
step: 20, loss: 0.0024663687217980623
step: 30, loss: 0.0007919908966869116
step: 40, loss: 0.005351308733224869
step: 50, loss: 0.006020638160407543
step: 60, loss: 0.004077460616827011
step: 70, loss: 0.008761354722082615
step: 80, loss: 0.010358858853578568
step: 90, loss: 0.06329087913036346
step: 100, loss: 0.029028726741671562
step: 110, loss: 0.006985506974160671
step: 120, loss: 0.008045039139688015
step: 130, loss: 0.1513843834400177
step: 140, loss: 0.0037268144078552723
step: 150, loss: 0.0020575327798724174
step: 160, loss: 0.004554847255349159
step: 170, loss: 0.011134592816233635
step: 180, loss: 0.09521184861660004
step: 190, loss: 0.05053999274969101
/usr/local/lib/python3.7/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=0.7712082262210797, f1=0.7830423940149626,
best_f1=0.7896103896103895
```

```
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.006185355596244335
step: 10, loss: 0.0038254272658377886
step: 20, loss: 0.0020371987484395504
step: 30, loss: 0.0005602336023002863
step: 40, loss: 0.08008576184511185
step: 50, loss: 0.026266030967235565
step: 60, loss: 0.003838102100417018
step: 70, loss: 0.003943517804145813
step: 80, loss: 0.007081673946231604
step: 90, loss: 0.0015132918488234282
step: 100, loss: 0.0012529162922874093
step: 110, loss: 0.0010903098154813051
step: 120, loss: 0.10992400348186493
step: 130, loss: 0.0035385009832680225
step: 140, loss: 0.002005018526688218
step: 150, loss: 0.003107621567323804
step: 160, loss: 0.0014022403629496694
step: 170, loss: 0.001118380925618112
step: 180, loss: 0.0010177608346566558
step: 190, loss: 0.00199604919180274
/usr/local/lib/python3.7/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=0.7851458885941645, f1=0.8134715025906736,
best_f1=0.7896103896103895
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0007011779234744608
step: 10, loss: 0.0006871689110994339
step: 20, loss: 0.0010631741024553776
step: 30, loss: 0.011987222358584404
step: 40, loss: 0.00036650634137913585
step: 50, loss: 0.05588288977742195
step: 60, loss: 0.0006934336852282286
step: 70, loss: 0.0006820374983362854
step: 80, loss: 0.000933525909204036
step: 90, loss: 0.0002784958924166858
step: 100, loss: 0.0007164230337366462
step: 110, loss: 0.001062346389517188
step: 120, loss: 0.05647636577486992
step: 130, loss: 0.0015150789404287934
step: 140, loss: 0.003627085825428367
step: 150, loss: 0.0011234792182222009
step: 160, loss: 0.0016489307163283229
```

```
step: 170, loss: 0.0014870526501908898
step: 180, loss: 0.01267226506024599
step: 190, loss: 0.09879404306411743
/usr/local/lib/python3.7/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=0.7769028871391076, f1=0.8, best_f1=0.7896103896103895
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0020980052649974823
step: 10, loss: 0.020248042419552803
step: 20, loss: 0.0009089100058190525
step: 30, loss: 0.001798369805328548
step: 40, loss: 0.005536659620702267
step: 50, loss: 0.03810668736696243
step: 60, loss: 0.006041122134774923
step: 70, loss: 0.0006406959146261215
step: 80, loss: 0.0007848537061363459
step: 90, loss: 0.0016175297787413
step: 100, loss: 0.001275237649679184
step: 110, loss: 0.004957225173711777
step: 120, loss: 0.0008066837908700109
step: 130, loss: 0.006963049061596394
step: 140, loss: 0.002390265231952071
step: 150, loss: 0.0020627903286367655
step: 160, loss: 0.0007126420387066901
step: 170, loss: 0.0014040714595466852
step: 180, loss: 0.0010640650289133191
step: 190, loss: 0.04478176310658455
/usr/local/lib/python3.7/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=0.7819148936170213, f1=0.8095238095238094,
best_f1=0.7896103896103895
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0021149124950170517
step: 10, loss: 0.0052154515869915485
step: 20, loss: 0.005207792390137911
step: 30, loss: 0.0029847491532564163
step: 40, loss: 0.0031132949516177177
step: 50, loss: 0.20890916883945465
step: 60, loss: 0.0012935335980728269
step: 70, loss: 0.008577670902013779
step: 80, loss: 0.0019142113160341978
step: 90, loss: 0.005318997427821159
```

```
step: 100, loss: 0.017637088894844055
step: 110, loss: 0.0006631235592067242
step: 120, loss: 0.0018084516050294042
step: 130, loss: 0.0005016233189962804
step: 140, loss: 0.0003719865344464779
step: 150, loss: 0.0006975685246288776
step: 160, loss: 0.0005469952593557537
step: 170, loss: 0.0004239374538883567
step: 180, loss: 0.000593235541600734
step: 190, loss: 0.0023057921789586544
/usr/local/lib/python3.7/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=0.7882037533512065, f1=0.8167539267015705,
best_f1=0.7896103896103895
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0008144961902871728
step: 10, loss: 0.0005201524472795427
step: 20, loss: 0.00045117794070392847
step: 30, loss: 0.0009808873292058706
step: 40, loss: 0.0038458104245364666
step: 50, loss: 0.0024207544047385454
step: 60, loss: 0.0007692394428886473
step: 70, loss: 0.0015639691846445203
step: 80, loss: 0.003589852713048458
step: 90, loss: 0.04106758162379265
step: 100, loss: 0.00646961759775877
step: 110, loss: 0.0014846139820292592
step: 120, loss: 0.10969969630241394
step: 130, loss: 0.0012404578737914562
step: 140, loss: 0.1679559350013733
step: 150, loss: 0.0016741689760237932
step: 160, loss: 0.00477292574942112
step: 170, loss: 0.0010086086113005877
step: 180, loss: 0.0030754278413951397
step: 190, loss: 0.0017442272510379553
/usr/local/lib/python3.7/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=0.7924528301886792, f1=0.8073878627968338,
best_f1=0.7896103896103895
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.004835131578147411
step: 10, loss: 0.001766012399457395
```

```
step: 20, loss: 0.0032281007152050734
step: 30, loss: 0.004968248773366213
step: 40, loss: 0.002064278582111001
step: 50, loss: 0.0013932292349636555
step: 60, loss: 0.001874046167358756
step: 70, loss: 0.010964402928948402
step: 80, loss: 0.03393801674246788
step: 90, loss: 0.00083813356468454
step: 100, loss: 0.0005751755088567734
step: 110, loss: 0.00040334323421120644
step: 120, loss: 0.0012721145758405328
step: 130, loss: 0.013888990506529808
step: 140, loss: 0.0038966492284089327
step: 150, loss: 0.0016545800026506186
step: 160, loss: 0.00044614425860345364
step: 170, loss: 0.01798618771135807
step: 180, loss: 0.0007217250531539321
step: 190, loss: 0.0005781762301921844
/usr/local/lib/python3.7/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=0.7936507936507936, f1=0.8041237113402062,
best_f1=0.7896103896103895
mv: cannot stat '*_dev.pt': No such file or directory
Selected optimization level O2:  FP16 training with FP32 batchnorm and FP32
master weights.

Defaults for this optimization level are:
enabled                : True
opt_level              : O2
cast_model_type        : torch.float16
patch_torch_functions  : False
keep_batchnorm_fp32    : True
master_weights         : True
loss_scale             : dynamic
Processing user overrides (additional kwargs that are not None)…
After processing overrides, optimization options are:
enabled                : True
opt_level              : O2
cast_model_type        : torch.float16
patch_torch_functions  : False
keep_batchnorm_fp32    : True
master_weights         : True
loss_scale             : dynamic
Warning:  multi_tensor_applier fused unscale kernel is unavailable, possibly
because apex was installed without --cuda_ext --cpp_ext. Using Python fallback.
Original ImportError was: ModuleNotFoundError("No module named 'amp_C'")
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
```

```
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
2049it [00:12, 166.12it/s]
load_f1 = 0.6074766355140186
real_f1 = 0.6037735849056604
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
4398it [00:24, 176.47it/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 bert-base-uncased \
       --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 bert-base-uncased \
--use_gpu \
--fp16 \
--checkpoint_path checkpoints/
```

Downloading: 100% 433/433 [00:00<00:00, 388kB/s]
Downloading: 100% 232k/232k [00:00<00:00, 3.27MB/s]
Downloading: 100% 440M/440M [00:06<00:00, 72.2MB/s]
Selected optimization level O2: FP16 training with FP32 batchnorm and FP32
master weights.

Defaults for this optimization level are:
enabled   : True
opt_level   : 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: 0.6373497247695923
step: 10, loss: 0.3721559941768646
step: 20, loss: 0.3066442310810089

```
step: 30, loss: 0.38014018535614014
step: 40, loss: 0.26081645488739014
step: 50, loss: 0.24878303706645966
step: 60, loss: 0.24997681379318237
step: 70, loss: 0.37006959319114685
step: 80, loss: 0.3759680688381195
step: 90, loss: 0.2548896074295044
step: 100, loss: 0.222920760512352
step: 110, loss: 0.1788758486509323
step: 120, loss: 0.12111561000347137
step: 130, loss: 0.06048804521560669
step: 140, loss: 0.22171001136302948
step: 150, loss: 0.15624015033245087
step: 160, loss: 0.11266078054904938
step: 170, loss: 0.2367572784423828
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.7008547008547009, f1=0.6709677419354838,
best_f1=0.6709677419354838
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.058229707181453705
step: 10, loss: 0.11479490995407104
step: 20, loss: 0.0347757488489151
step: 30, loss: 0.16419066488742828
step: 40, loss: 0.15731805562973022
step: 50, loss: 0.13692045211791992
step: 60, loss: 0.15995903313159943
step: 70, loss: 0.08281766623258591
step: 80, loss: 0.050005748867988586
step: 90, loss: 0.16822251677513123
step: 100, loss: 0.14461375772953033
step: 110, loss: 0.09603899717330933
step: 120, loss: 0.08204219490289688
step: 130, loss: 0.07510780543088913
step: 140, loss: 0.15176604688167572
step: 150, loss: 0.17673994600772858
step: 160, loss: 0.13191883265972137
step: 170, loss: 0.042249441146850586
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
```

```
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.8082901554404144, f1=0.7939698492462312,
best_f1=0.7939698492462312
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.09520711749792099
step: 10, loss: 0.16664718091487885
step: 20, loss: 0.03895142301917076
step: 30, loss: 0.0874994620680809
step: 40, loss: 0.13495080173015594
step: 50, loss: 0.08862970024347305
step: 60, loss: 0.1089923158288002
step: 70, loss: 0.0931791439652443
step: 80, loss: 0.0460478775203228
step: 90, loss: 0.02488229237496853
step: 100, loss: 0.10942574590444565
step: 110, loss: 0.08596231788396835
step: 120, loss: 0.12161177396774292
step: 130, loss: 0.031130878254771233
step: 140, loss: 0.07587211579084396
step: 150, loss: 0.026434509083628654
step: 160, loss: 0.04291718080639839
step: 170, loss: 0.08319764584302902
/usr/local/lib/python3.7/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=0.8040201005025125, f1=0.8019559902200489,
best_f1=0.7939698492462312
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.009364948607981205
step: 10, loss: 0.032714031636714935
step: 20, loss: 0.0060034762136638165
step: 30, loss: 0.2902781665325165
step: 40, loss: 0.004769536200910807
step: 50, loss: 0.15188340842723846
step: 60, loss: 0.22879542410373688
step: 70, loss: 0.012635506689548492
step: 80, loss: 0.12463007867336273
step: 90, loss: 0.061099834740161896
step: 100, loss: 0.20702725648880005
step: 110, loss: 0.09409885853528976
step: 120, loss: 0.006486514117568731
```

```
step: 130, loss: 0.0315246656537056
step: 140, loss: 0.025133166462183
step: 150, loss: 0.14704224467277527
step: 160, loss: 0.03316393122076988
step: 170, loss: 0.034432508051395416
/usr/local/lib/python3.7/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=0.786046511627907, f1=0.7963800904977376,
best_f1=0.7939698492462312
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.01592177525162697
step: 10, loss: 0.00814928486943245
step: 20, loss: 0.038551487028598785
step: 30, loss: 0.17474855482578278
step: 40, loss: 0.07688043266534805
step: 50, loss: 0.1355120837688446
step: 60, loss: 0.07339449226856232
step: 70, loss: 0.26180416345596313
step: 80, loss: 0.10814085602760315
step: 90, loss: 0.1241757720708847
step: 100, loss: 0.04321916401386261
step: 110, loss: 0.030607588589191437
step: 120, loss: 0.02683088928461075
step: 130, loss: 0.1979800909757614
step: 140, loss: 0.01187109388411045
step: 150, loss: 0.11060590296983719
step: 160, loss: 0.11030419170856476
step: 170, loss: 0.010309805162250996
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.8205128205128205, f1=0.8193384223918575,
best_f1=0.8193384223918575
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00797942467033863
step: 10, loss: 0.10712599754333496
step: 20, loss: 0.05964013189077377
step: 30, loss: 0.0737161785364151
step: 40, loss: 0.0497276671230793
```

```
step: 50, loss: 0.1344417780637741
step: 60, loss: 0.016885120421648026
step: 70, loss: 0.06302080303430557
step: 80, loss: 0.07030148059129715
step: 90, loss: 0.026351571083068848
step: 100, loss: 0.017087187618017197
step: 110, loss: 0.0018224731320515275
step: 120, loss: 0.05716472864151001
step: 130, loss: 0.016140643507242203
step: 140, loss: 0.1832413226366043
step: 150, loss: 0.15498459339141846
step: 160, loss: 0.0640069916844368
step: 170, loss: 0.0032473853789269924
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.8317307692307692, f1=0.794188861985472,
best_f1=0.794188861985472
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0010743483435362577
step: 10, loss: 0.011459735222160816
step: 20, loss: 0.002115854760631919
step: 30, loss: 0.04583398997783661
step: 40, loss: 0.0372355692088604
step: 50, loss: 0.022918419912457466
step: 60, loss: 0.0027004012372344732
step: 70, loss: 0.0003218232304789126
step: 80, loss: 0.006406849250197411
step: 90, loss: 0.00021757037029601634
step: 100, loss: 0.0002621258608996868
step: 110, loss: 0.04535377770662308
step: 120, loss: 0.005366533529013395
step: 130, loss: 0.14402399957180023
step: 140, loss: 0.0008195352857001126
step: 150, loss: 0.005334786139428616
step: 160, loss: 0.000272390287136659
step: 170, loss: 0.023099452257156372
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.836272040302267, f1=0.8109452736318408,
best_f1=0.8109452736318408
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.002890246454626322
step: 10, loss: 0.008053073659539223
step: 20, loss: 0.0042982823215425014
step: 30, loss: 0.013092819601297379
step: 40, loss: 0.0007725739269517362
step: 50, loss: 0.006032106000930071
step: 60, loss: 0.0029320931062102318
step: 70, loss: 0.0028405161574482918
step: 80, loss: 0.002388622146099806
step: 90, loss: 0.003929572645574808
step: 100, loss: 0.09033811837434769
step: 110, loss: 0.07008867710828781
step: 120, loss: 0.01919819973409176
step: 130, loss: 0.004686539992690086
step: 140, loss: 0.0329296737909317
step: 150, loss: 0.025357071310281754
step: 160, loss: 0.03611864522099495
step: 170, loss: 0.002767491154372692
/usr/local/lib/python3.7/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=0.8295165394402035, f1=0.8009950248756218,
best_f1=0.8109452736318408
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0012744634877890348
step: 10, loss: 0.006024669390171766
step: 20, loss: 0.006608565337955952
step: 30, loss: 0.05532778427004814
step: 40, loss: 0.18642717599868774
Gradient overflow.  Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 50, loss: 0.004359578248113394
step: 60, loss: 0.01309962384402752
step: 70, loss: 0.0007179719978012145
step: 80, loss: 0.0011101439595222473
step: 90, loss: 0.04283272475004196
step: 100, loss: 0.03368737921118736
step: 110, loss: 0.006751608569175005
step: 120, loss: 0.0012727768626064062
step: 130, loss: 0.052285924553871155
```

```
step: 140, loss: 0.010852070525288582
step: 150, loss: 0.0027690825518220663
step: 160, loss: 0.004226600285619497
step: 170, loss: 0.0014808328123763204
/usr/local/lib/python3.7/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=0.8312958435207823, f1=0.8019323671497586,
best_f1=0.8109452736318408
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.041142914444208145
step: 10, loss: 0.00019603797409217805
step: 20, loss: 0.06731490045785904
step: 30, loss: 0.012644757516682148
step: 40, loss: 0.025757968425750732
step: 50, loss: 0.12032090127468109
step: 60, loss: 0.0007471602293662727
step: 70, loss: 0.027741000056266785
step: 80, loss: 0.0051085264421999454
step: 90, loss: 0.0011001203674823046
step: 100, loss: 0.0005252755945548415
step: 110, loss: 0.003878816729411483
step: 120, loss: 0.01179398875683546
step: 130, loss: 0.024044658988714218
step: 140, loss: 0.009656500071287155
step: 150, loss: 0.005922697950154543
step: 160, loss: 0.06285074353218079
step: 170, loss: 0.006392533425241709
/usr/local/lib/python3.7/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=0.8316831683168316, f1=0.7780429594272077,
best_f1=0.8109452736318408
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.04881914332509041
step: 10, loss: 0.002229547593742609
step: 20, loss: 0.004674346186220646
step: 30, loss: 0.001004080637358129
step: 40, loss: 0.0014298309106379747
step: 50, loss: 0.0005381070077419281
step: 60, loss: 0.0034808439668267965
step: 70, loss: 0.025492584332823753
step: 80, loss: 0.0020868254359811544
step: 90, loss: 0.0023756222799420357
```

```
step: 100, loss: 0.0008373733726330101
step: 110, loss: 0.08927112817764282
step: 120, loss: 0.003371786791831255
step: 130, loss: 0.00020710799435619265
step: 140, loss: 0.015019599348306656
step: 150, loss: 0.005332986358553171
step: 160, loss: 0.00599583238363266
step: 170, loss: 0.0029512643814086914
/usr/local/lib/python3.7/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=0.8296296296296296, f1=0.8173076923076923,
best_f1=0.8109452736318408
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.014721553772687912
step: 10, loss: 0.0004920699284411967
step: 20, loss: 0.00036168392398394644
step: 30, loss: 0.0025096735917031765
step: 40, loss: 0.001244635903276503
step: 50, loss: 0.0004718563868664205
step: 60, loss: 0.005545745603740215
step: 70, loss: 0.020982157438993454
step: 80, loss: 0.0003536739095579833
step: 90, loss: 0.0013542763190343976
step: 100, loss: 0.0006190590211190283
step: 110, loss: 0.0007947086123749614
step: 120, loss: 0.04340243712067604
step: 130, loss: 0.04165467992424965
step: 140, loss: 0.02741178683936596
step: 150, loss: 0.21530768275260925
step: 160, loss: 0.0008311295532621443
step: 170, loss: 0.0003515627176966518
/usr/local/lib/python3.7/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=0.8329297820823245, f1=0.8133971291866029,
best_f1=0.8109452736318408
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.018315888941287994
step: 10, loss: 0.005507337860763073
step: 20, loss: 0.0003225411637686193
step: 30, loss: 0.000924637948628515
step: 40, loss: 0.019669005647301674
step: 50, loss: 0.003247430780902505
```

step: 60, loss: 0.021877655759453773
step: 70, loss: 0.0004380558093544096
step: 80, loss: 0.0010010653641074896
step: 90, loss: 0.00022744774469174445
step: 100, loss: 0.0037309606559574604
step: 110, loss: 0.00020571811182890087
step: 120, loss: 0.03522762656211853
step: 130, loss: 0.011388426646590233
step: 140, loss: 0.001194002223201096
step: 150, loss: 0.001674190629273653
step: 160, loss: 0.010333982296288013
step: 170, loss: 0.003582798643037677
/usr/local/lib/python3.7/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=0.8226221079691517, f1=0.8120300751879699,
best_f1=0.8109452736318408
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.004305231384932995
step: 10, loss: 0.00016584657714700219
step: 20, loss: 0.013088444247841835
step: 30, loss: 0.0017250578384846449
step: 40, loss: 0.001377870561555028
step: 50, loss: 0.00040238883229903877
step: 60, loss: 0.00551664037629962
step: 70, loss: 0.0006848580087535083
step: 80, loss: 0.005330360494554043
step: 90, loss: 0.00042493094224482775
step: 100, loss: 0.012090425938367844
step: 110, loss: 0.0017552459612488747
step: 120, loss: 0.041745588183403015
step: 130, loss: 0.001330366125330329
step: 140, loss: 0.0012845058226957917
step: 150, loss: 0.0003129805554635823
step: 160, loss: 0.0007059563649818301
step: 170, loss: 0.0007666859892196953
/usr/local/lib/python3.7/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=0.8211586901763224, f1=0.8088235294117646,
best_f1=0.8109452736318408
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0037938072346150875
step: 10, loss: 0.000609388982411474

```
step: 20, loss: 0.053263988345861435
step: 30, loss: 0.00017525660223327577
step: 40, loss: 0.0011958586983382702
step: 50, loss: 0.000658674631267786
step: 60, loss: 0.0026275671552866697
step: 70, loss: 0.005295877810567617
step: 80, loss: 0.016722390428185463
step: 90, loss: 0.0013595662312582135
step: 100, loss: 0.011196565814316273
step: 110, loss: 0.0002148147177649662
step: 120, loss: 0.0003521730250213295
step: 130, loss: 0.0011884482810273767
step: 140, loss: 0.001838593976572156
step: 150, loss: 0.013169157318770885
step: 160, loss: 0.0003221704682800919
step: 170, loss: 0.03447093814611435
/usr/local/lib/python3.7/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=0.8188585607940446, f1=0.8115942028985508,
best_f1=0.8109452736318408
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, 277.32it/s]
load_f1 = 0.4214876033057851
real_f1 = 0.41269841269841273
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 [00:18, 236.85it/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 bert-base-uncased \
       --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 bert-base-uncased \
       --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: 0.5984289646148682
step: 10, loss: 0.6309069991111755
step: 20, loss: 0.3546966016292572
step: 30, loss: 0.09655976295471191
step: 40, loss: 0.17539358139038086
step: 50, loss: 0.06712396442890167
step: 60, loss: 0.17435792088508606
step: 70, loss: 0.201813206076622

434
```

```
step: 80, loss: 0.028476113453507423
step: 90, loss: 0.12350612133741379
step: 100, loss: 0.0067200809717178345
step: 110, loss: 0.23084577918052673
step: 120, loss: 0.016604648903012276
step: 130, loss: 0.03743276745080948
step: 140, loss: 0.004087226931005716
step: 150, loss: 0.01537527609616518
step: 160, loss: 0.15068882703781128
step: 170, loss: 0.15937568247318268
step: 180, loss: 0.04287632927298546
step: 190, loss: 0.05039301514625549
step: 200, loss: 0.11427188664674759
step: 210, loss: 0.009329305961728096
step: 220, loss: 0.0061249504797160625
step: 230, loss: 0.02340780571103096
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9701657458563535, f1=0.9686800894854586,
best_f1=0.9686800894854586
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.013851052150130272
step: 10, loss: 0.00896868109703064
step: 20, loss: 0.15081098675727844
step: 30, loss: 0.017128974199295044
step: 40, loss: 0.09457256644964218
step: 50, loss: 0.005010006483644247
step: 60, loss: 0.021296672523021698
step: 70, loss: 0.10069683194160461
step: 80, loss: 0.009261484257876873
step: 90, loss: 0.013975290581583977
step: 100, loss: 0.13498049974441528
step: 110, loss: 0.09977183490991592
step: 120, loss: 0.09378498792648315
step: 130, loss: 0.014841698110103607
step: 140, loss: 0.004239553119987249
step: 150, loss: 0.011544683948159218
step: 160, loss: 0.024938123300671577
step: 170, loss: 0.0007957394700497389
step: 180, loss: 0.0014437763020396233
step: 190, loss: 0.005042870994657278
```

```
step: 200, loss: 0.0028269989416003227
step: 210, loss: 0.002628088230267167
step: 220, loss: 0.2800047993659973
step: 230, loss: 0.02612587995827198
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9787234042553192, f1=0.9731543624161074,
best_f1=0.9731543624161074
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.045350052416324615
step: 10, loss: 0.0036739949136972427
step: 20, loss: 0.04836370050907135
step: 30, loss: 0.02876836247742176
step: 40, loss: 0.09180104732513428
step: 50, loss: 0.0054868971928954124
step: 60, loss: 0.00502582686021924
step: 70, loss: 0.06869909912347794
step: 80, loss: 0.0013224952854216099
step: 90, loss: 0.06123630702495575
step: 100, loss: 0.0033728666603565216
step: 110, loss: 0.0010065010283142328
step: 120, loss: 0.009662720374763012
step: 130, loss: 0.00044324889313429594
step: 140, loss: 0.008170687593519688
step: 150, loss: 0.04428361728787422
step: 160, loss: 0.01705487072467804
step: 170, loss: 0.011886299587786198
step: 180, loss: 0.027774561196565628
step: 190, loss: 0.002250595949590206
step: 200, loss: 0.055200014263391495
step: 210, loss: 0.0037380452267825603
step: 220, loss: 0.0008323651854880154
step: 230, loss: 0.020257391035556793
/usr/local/lib/python3.7/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=0.9732739420935412, f1=0.9689578713968958,
best_f1=0.9731543624161074
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
```

```
step: 0, loss: 0.000995245180092752
step: 10, loss: 0.009164686314761639
step: 20, loss: 0.0013332006055861712
step: 30, loss: 0.02582743391394615
step: 40, loss: 0.003326427424326539
step: 50, loss: 0.0028233216144144535
step: 60, loss: 0.0006686746492050588
step: 70, loss: 0.004465918987989426
step: 80, loss: 0.005886731669306755
step: 90, loss: 0.0035246300976723433
step: 100, loss: 0.0006640676874667406
step: 110, loss: 0.006522298324853182
step: 120, loss: 0.010660193860530853
step: 130, loss: 0.0008751689456403255
step: 140, loss: 0.0004271193465683609
step: 150, loss: 0.2630937099456787
step: 160, loss: 0.1391354203224182
step: 170, loss: 0.008372090756893158
step: 180, loss: 0.0041752588003873825
step: 190, loss: 0.0073751420713961124
step: 200, loss: 0.0024316590279340744
step: 210, loss: 0.14490751922130585
step: 220, loss: 0.002840770408511162
step: 230, loss: 0.00581250200048089
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9832402234636871, f1=0.9752808988764046,
best_f1=0.9752808988764046
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0026955727953463793
step: 10, loss: 0.0007221636478789151
step: 20, loss: 0.0014410251751542091
step: 30, loss: 0.0004732509842142463
step: 40, loss: 0.0015311157330870628
step: 50, loss: 0.0007383530028164387
step: 60, loss: 0.07170861959457397
step: 70, loss: 0.000781480222940445
step: 80, loss: 0.0008264842326752841
step: 90, loss: 0.0030201307963579893
step: 100, loss: 0.0003705160634126514
step: 110, loss: 0.0008788537816144526
```

```
step: 120, loss: 0.00027132374816574156
step: 130, loss: 0.002779579721391201
step: 140, loss: 0.008617883548140526
step: 150, loss: 0.014534891583025455
step: 160, loss: 0.008775421418249607
step: 170, loss: 0.013560797087848186
step: 180, loss: 0.0006051842938177288
step: 190, loss: 0.09453287720680237
step: 200, loss: 0.0008401330560445786
step: 210, loss: 0.0011362485820427537
step: 220, loss: 0.0009760076645761728
step: 230, loss: 0.0007772105163894594
/usr/local/lib/python3.7/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=0.9798657718120806, f1=0.9732142857142857,
best_f1=0.9752808988764046
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0014684272464364767
step: 10, loss: 0.0015549955423921347
step: 20, loss: 0.00632155267521739
step: 30, loss: 0.001153071760199964
step: 40, loss: 0.002441217657178604
step: 50, loss: 0.0007600426324643195
step: 60, loss: 0.0004525560070760548
step: 70, loss: 0.0021152214612811804
step: 80, loss: 0.0009417766705155373
step: 90, loss: 0.00022821960737928748
step: 100, loss: 0.09424160420894623
step: 110, loss: 0.0015072317328304052
step: 120, loss: 0.00397541793063283
step: 130, loss: 0.0006031978409737349
step: 140, loss: 0.0004666615277528763
step: 150, loss: 0.00925297848880291
step: 160, loss: 0.00024032860528677702
step: 170, loss: 0.00023012679594103247
step: 180, loss: 0.0006472321110777557
step: 190, loss: 0.09762938320636749
step: 200, loss: 0.00036260270280763507
step: 210, loss: 0.004557713866233826
step: 220, loss: 0.01838451810181141
step: 230, loss: 0.05696210265159607
/usr/local/lib/python3.7/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=0.9787234042553192, f1=0.9752808988764046,
```

```
best_f1=0.9752808988764046
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0027445268351584673
step: 10, loss: 0.0006363533902913332
step: 20, loss: 0.01566896215081215
step: 30, loss: 0.0013519260101020336
step: 40, loss: 0.005884020123630762
step: 50, loss: 0.007424355484545231
step: 60, loss: 0.057914379984140396
step: 70, loss: 0.023513825610280037
step: 80, loss: 0.0005288023967295885
step: 90, loss: 0.0004240469424985349
step: 100, loss: 0.00023700515157543123
step: 110, loss: 0.00030378237715922296
step: 120, loss: 0.00018424075096845627
step: 130, loss: 0.00014873997133690864
step: 140, loss: 0.00029714446228201425
step: 150, loss: 0.003336466383188963
step: 160, loss: 0.06314016878604889
step: 170, loss: 0.06022866070270538
step: 180, loss: 0.0006608616677112877
step: 190, loss: 0.0011101956479251385
step: 200, loss: 0.052366990596055984
step: 210, loss: 0.0008043296402320266
step: 220, loss: 0.0100239934399724
step: 230, loss: 0.0020781387574970722
/usr/local/lib/python3.7/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=0.9809203142536477, f1=0.9786276715410572,
best_f1=0.9752808988764046
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0003972681879531592
step: 10, loss: 0.004967528395354748
step: 20, loss: 0.00042072052019648254
step: 30, loss: 0.00012007301847916096
step: 40, loss: 0.0018298751674592495
step: 50, loss: 0.0015085435006767511
step: 60, loss: 0.00019506212265696377
step: 70, loss: 0.0005417457432486117
step: 80, loss: 0.00015424579032696784
step: 90, loss: 0.00023310972028411925
step: 100, loss: 0.005309737287461758
step: 110, loss: 0.018334027379751205
```

```
step: 120, loss: 0.01320928055793047
step: 130, loss: 0.0010769636137410998
step: 140, loss: 0.0004878875915892422
step: 150, loss: 0.00021901819854974747
step: 160, loss: 0.0001888091501314193
step: 170, loss: 0.0024599346797913313
step: 180, loss: 0.0017938674427568913
step: 190, loss: 0.0006340989493764937
step: 200, loss: 0.0006916237180121243
step: 210, loss: 0.0015551565447822213
step: 220, loss: 0.0004981131642125547
step: 230, loss: 0.0006799310795031488
/usr/local/lib/python3.7/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=0.9776286353467561, f1=0.9743589743589743,
best_f1=0.9752808988764046
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 9.698908979771659e-05
step: 10, loss: 0.0002878982340916991
step: 20, loss: 0.001501289545558393
step: 30, loss: 0.00034001137828454375
step: 40, loss: 0.10902591049671173
step: 50, loss: 5.7779103372013196e-05
step: 60, loss: 0.00013286435569170862
step: 70, loss: 0.0017425973201170564
step: 80, loss: 0.00020849840075243264
step: 90, loss: 0.0024319372605532408
step: 100, loss: 0.00013228855095803738
step: 110, loss: 5.6173372286139056e-05
step: 120, loss: 6.935495184734464e-05
step: 130, loss: 5.115633030072786e-05
step: 140, loss: 4.8008241719799116e-05
step: 150, loss: 0.00030313999741338193
step: 160, loss: 5.7621440646471456e-05
step: 170, loss: 8.109051850624382e-05
step: 180, loss: 0.0018014473607763648
step: 190, loss: 0.0006952836993150413
step: 200, loss: 0.0008425044361501932
step: 210, loss: 0.0038159836549311876
step: 220, loss: 0.00010541867959545925
step: 230, loss: 0.00017670862143859267
/usr/local/lib/python3.7/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=0.983050847457627, f1=0.9773242630385486,
```

best_f1=0.9752808988764046
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 7.797450962243602e-05
step: 10, loss: 0.00022072743740864098
step: 20, loss: 0.00011302793427603319
step: 30, loss: 0.00028150880825705826
step: 40, loss: 0.000963641214184463
step: 50, loss: 0.0001192636409541592
step: 60, loss: 0.0003893494722433388
step: 70, loss: 0.0017613967647776008
step: 80, loss: 0.00016425202193204314
step: 90, loss: 0.098353311419487
step: 100, loss: 0.00014315196312963963
step: 110, loss: 0.00017966885934583843
step: 120, loss: 0.0011437444481998682
step: 130, loss: 0.0638740286231041
step: 140, loss: 0.04443597048521042
step: 150, loss: 0.014533407986164093
step: 160, loss: 0.0006833617226220667
step: 170, loss: 0.00014301619376055896
step: 180, loss: 0.008383904583752155
step: 190, loss: 0.0014926594449207187
step: 200, loss: 3.90694840461947e-05
step: 210, loss: 0.00011540992272784933
step: 220, loss: 6.72367968945764e-05
step: 230, loss: 0.0008465422433800995
/usr/local/lib/python3.7/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=0.9831271091113611, f1=0.9752252252252253,
best_f1=0.9752808988764046
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 5.288101237965748e-05
step: 10, loss: 0.00011028294102288783
step: 20, loss: 0.0009812969947233796
step: 30, loss: 0.0029399532359093428
step: 40, loss: 0.0004341907042544335
step: 50, loss: 0.000178035581484437
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 32768.0
step: 60, loss: 0.0028273717034608126
step: 70, loss: 0.00044263325980864465
step: 80, loss: 0.00011667661601677537
step: 90, loss: 0.00029668663046322763
step: 100, loss: 0.000277891056612134

```
step: 110, loss: 0.009352811612188816
step: 120, loss: 0.0001223678991664201
step: 130, loss: 0.00018085465126205236
step: 140, loss: 3.724443740793504e-05
step: 150, loss: 0.0032621752470731735
step: 160, loss: 0.0001697027764748782
step: 170, loss: 0.00076038867700845
step: 180, loss: 0.029632655903697014
step: 190, loss: 0.0001228434412041679
step: 200, loss: 0.0007520526414737105
step: 210, loss: 0.00014797880430705845
step: 220, loss: 8.925537986215204e-05
step: 230, loss: 0.00015483886818401515
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9854423292273236, f1=0.9710467706013363,
best_f1=0.9710467706013363
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0013890400296077132
step: 10, loss: 6.536470755236223e-05
step: 20, loss: 0.0006118747987784445
step: 30, loss: 0.00024306404520757496
step: 40, loss: 0.00045152968959882855
step: 50, loss: 0.005125865805894136
step: 60, loss: 0.0034589862916618586
step: 70, loss: 0.00014863707474432886
step: 80, loss: 0.00040368115878663957
step: 90, loss: 8.860675006872043e-05
step: 100, loss: 3.16455407300964e-05
step: 110, loss: 4.4369797251420096e-05
step: 120, loss: 3.753053169930354e-05
step: 130, loss: 7.370120874838904e-05
step: 140, loss: 0.0008733628201298416
step: 150, loss: 0.0011622001184150577
step: 160, loss: 0.00010064908565254882
step: 170, loss: 0.0005908144521526992
step: 180, loss: 0.006835748441517353
step: 190, loss: 9.005245374282822e-05
step: 200, loss: 3.5366738302400336e-05
step: 210, loss: 0.00010808841761900112
step: 220, loss: 0.0003660086658783257
```

step: 230, loss: 0.0018957251450046897
/usr/local/lib/python3.7/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=0.983277591973244, f1=0.9721913236929923,
best_f1=0.9710467706013363
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00019331910880282252
step: 10, loss: 0.0001370374084217474
step: 20, loss: 0.0001475857279729098
step: 30, loss: 0.00011251357500441372
step: 40, loss: 0.00017905501590576023
step: 50, loss: 0.0007865853840485215
step: 60, loss: 0.0001444533991161734
step: 70, loss: 5.123562004882842e-05
step: 80, loss: 4.695808456744999e-05
step: 90, loss: 0.0004777562862727791
step: 100, loss: 6.804862641729414e-05
step: 110, loss: 0.0038912754971534014
step: 120, loss: 7.550996087957174e-05
step: 130, loss: 4.6783385187154636e-05
step: 140, loss: 5.379128924687393e-05
step: 150, loss: 6.633213342865929e-05
step: 160, loss: 0.00017348950495943427
step: 170, loss: 4.5572665840154514e-05
step: 180, loss: 5.316307215252891e-05
step: 190, loss: 4.1903007513610646e-05
step: 200, loss: 0.0001752923271851614
step: 210, loss: 2.5700224796310067e-05
step: 220, loss: 0.00012698580394499004
step: 230, loss: 4.800660826731473e-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=0.984304932735426, f1=0.9808773903262092,
best_f1=0.9710467706013363
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00010330440272809938
step: 10, loss: 0.00026847634580917656
step: 20, loss: 0.0015346587169915438
step: 30, loss: 5.288760075927712e-05
step: 40, loss: 0.013415229506790638
step: 50, loss: 0.0001073743260349147
step: 60, loss: 8.750831329962239e-05

```
step: 70, loss: 4.074518074048683e-05
step: 80, loss: 3.356762317707762e-05
step: 90, loss: 8.64499161252752e-05
step: 100, loss: 3.6986319173593074e-05
step: 110, loss: 0.00044901767978444695
step: 120, loss: 0.00019247279851697385
step: 130, loss: 5.062769923824817e-05
step: 140, loss: 0.0013309925561770797
step: 150, loss: 3.888610808644444e-05
step: 160, loss: 0.0021860802080482244
step: 170, loss: 0.0002268755342811346
step: 180, loss: 7.126080890884623e-05
step: 190, loss: 2.8810638468712568e-05
step: 200, loss: 2.932835377578158e-05
step: 210, loss: 0.00010381893662270159
step: 220, loss: 0.000230448815273121
step: 230, loss: 0.0025051224511116743
/usr/local/lib/python3.7/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=0.984304932735426, f1=0.978675645342312,
best_f1=0.9710467706013363
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 3.096735599683598e-05
step: 10, loss: 2.1341853425838053e-05
step: 20, loss: 4.1330196836497635e-05
step: 30, loss: 0.00010872911661863327
step: 40, loss: 0.00010881891648750752
step: 50, loss: 0.030773457139730453
step: 60, loss: 3.4554854210000485e-05
step: 70, loss: 0.0001477933255955577
step: 80, loss: 0.0005969498888589442
step: 90, loss: 0.00010435389413032681
step: 100, loss: 4.963890023645945e-05
step: 110, loss: 5.408029755926691e-05
step: 120, loss: 4.1735638660611585e-05
step: 130, loss: 0.0035261265002191067
step: 140, loss: 0.000649604422505945
step: 150, loss: 0.00019504937517922372
step: 160, loss: 2.6583047656458803e-05
step: 170, loss: 2.6556910597719252e-05
step: 180, loss: 0.0004945112159475684
step: 190, loss: 0.00020583099103532732
step: 200, loss: 4.4454882299760357e-05
step: 210, loss: 9.21455430216156e-05
step: 220, loss: 0.001287955092266202
```

step: 230, loss: 3.140323678962886e-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=0.9842696629213483, f1=0.9797752808988766,
best_f1=0.9710467706013363
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.")
2473it [00:12, 201.62it/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 [00:18, 243.81it/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 bert-base-uncased \
  --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 bert-base-uncased \
  --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: 0.6151407361030579
step: 10, loss: 0.498611718416214
step: 20, loss: 0.4994441568851471
step: 30, loss: 0.08650235086679459
step: 40, loss: 0.19029004871845245
step: 50, loss: 0.059788137674331665
step: 60, loss: 0.37446287274360657
step: 70, loss: 0.16117902100086212
step: 80, loss: 0.1706964671611786
step: 90, loss: 0.0519549697637558
step: 100, loss: 0.04800776019692421
step: 110, loss: 0.061222366988658905
step: 120, loss: 0.14908680319786072
step: 130, loss: 0.05207423120737076
step: 140, loss: 0.05296257138252258
step: 150, loss: 0.08730778843164444
step: 160, loss: 0.029156316071748734
step: 170, loss: 0.158375084400177
step: 180, loss: 0.07641295343637466
step: 190, loss: 0.025662528350949287
step: 200, loss: 0.2834603786468506
step: 210, loss: 0.12235049158334732
step: 220, loss: 0.18399566411972046
```

```
step: 230, loss: 0.16160009801387787
step: 240, loss: 0.047593459486961365
step: 250, loss: 0.009007933549582958
step: 260, loss: 0.07953186333179474
step: 270, loss: 0.009064409881830215
step: 280, loss: 0.05208146199584007
step: 290, loss: 0.05063496530056
step: 300, loss: 0.014199203811585903
step: 310, loss: 0.2815638780593872
step: 320, loss: 0.07134076207876205
step: 330, loss: 0.04084891453385353
step: 340, loss: 0.053085315972566605
step: 350, loss: 0.05964807793498039
step: 360, loss: 0.05695812404155731
step: 370, loss: 0.16613955795764923
step: 380, loss: 0.04278506711125374
step: 390, loss: 0.09927941113710403
step: 400, loss: 0.19911642372608185
step: 410, loss: 0.040803212672472
step: 420, loss: 0.02254304476082325
step: 430, loss: 0.13155284523963928
step: 440, loss: 0.07054674625396729
step: 450, loss: 0.004385902546346188
step: 460, loss: 0.012911925092339516
step: 470, loss: 0.12330543249845505
step: 480, loss: 0.0586797259747982
step: 490, loss: 0.1763785183429718
step: 500, loss: 0.0716941580176353
step: 510, loss: 0.09550493210554123
step: 520, loss: 0.04052546247839928
step: 530, loss: 0.007444813381880522
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9375879868606289, f1=0.9285042333019755,
best_f1=0.9285042333019755
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.10558290779590607
step: 10, loss: 0.0455656535923481
step: 20, loss: 0.09935462474822998
step: 30, loss: 0.03832734003663063
step: 40, loss: 0.13085800409317017
```

```
step: 50, loss: 0.11892498284578323
step: 60, loss: 0.005241325590759516
step: 70, loss: 0.01885008066892624
step: 80, loss: 0.08842672407627106
step: 90, loss: 0.0075035616755485535
step: 100, loss: 0.03782433643937111
step: 110, loss: 0.04113822802901268
step: 120, loss: 0.14198914170265198
step: 130, loss: 0.10312672704458237
step: 140, loss: 0.04322122037410736
step: 150, loss: 0.08799736201763153
step: 160, loss: 0.011876348406076431
step: 170, loss: 0.038938459008932114
step: 180, loss: 0.017257677391171455
step: 190, loss: 0.12573271989822388
step: 200, loss: 0.01891661062836647
step: 210, loss: 0.03230121731758118
step: 220, loss: 0.07334835827350616
step: 230, loss: 0.01949165388941765
step: 240, loss: 0.03472708910703659
step: 250, loss: 0.032593291252851486
step: 260, loss: 0.00261638988740742
step: 270, loss: 0.23470713198184967
step: 280, loss: 0.017758939415216446
step: 290, loss: 0.022838864475488663
step: 300, loss: 0.06698677688837051
step: 310, loss: 0.011688657104969025
step: 320, loss: 0.18604399263858795
step: 330, loss: 0.0903654620051384
step: 340, loss: 0.020375076681375504
step: 350, loss: 0.001711913151666522
step: 360, loss: 0.094424307346344
step: 370, loss: 0.1508675217628479
step: 380, loss: 0.08141909539699554
step: 390, loss: 0.044105883687734604
step: 400, loss: 0.004266228526830673
step: 410, loss: 0.03803836181759834
step: 420, loss: 0.019923485815525055
step: 430, loss: 0.011435496620833874
step: 440, loss: 0.1251361221075058
step: 450, loss: 0.028346970677375793
step: 460, loss: 0.02177179977297783
step: 470, loss: 0.07596619427204132
step: 480, loss: 0.19151940941810608
step: 490, loss: 0.01228451356291771
step: 500, loss: 0.16835075616836548
step: 510, loss: 0.009236367419362068
step: 520, loss: 0.04895889014005661
```

```
step: 530, loss: 0.014841062016785145
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9428962996802193, f1=0.9340009103322713,
best_f1=0.9340009103322713
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.06524061411619186
step: 10, loss: 0.029602600261569023
step: 20, loss: 0.03448404744267464
step: 30, loss: 0.016370108351111412
step: 40, loss: 0.18149897456169128
step: 50, loss: 0.08365107327699661
step: 60, loss: 0.011132261715829372
step: 70, loss: 0.01861380785703659
step: 80, loss: 0.0038391663692891598
step: 90, loss: 0.0045702229253947735
step: 100, loss: 0.07371866703033447
step: 110, loss: 0.005107162520289421
step: 120, loss: 0.01789812371134758
step: 130, loss: 0.010905050672590733
step: 140, loss: 0.017450129613280296
step: 150, loss: 0.025260936468839645
step: 160, loss: 0.014916162006556988
step: 170, loss: 0.06078106909990311
step: 180, loss: 0.005396023392677307
step: 190, loss: 0.017885128036141396
step: 200, loss: 0.06370701640844345
step: 210, loss: 0.1559295803308487
step: 220, loss: 0.11713948100805283
step: 230, loss: 0.12331704795360565
step: 240, loss: 0.002143921097740531
step: 250, loss: 0.033940695226192474
step: 260, loss: 0.0183029156178236
step: 270, loss: 0.008891777135431767
step: 280, loss: 0.13353604078292847
step: 290, loss: 0.0018567681545391679
step: 300, loss: 0.02916271612048149
step: 310, loss: 0.011511236429214478
step: 320, loss: 0.037351094186306
step: 330, loss: 0.06715411692857742
step: 340, loss: 0.0033518995624035597
```

```
step: 350, loss: 0.015988094732165337
step: 360, loss: 0.039315998554229736
step: 370, loss: 0.04222545400261879
step: 380, loss: 0.03060954436659813
step: 390, loss: 0.02878705784678459
step: 400, loss: 0.02330838516354561
step: 410, loss: 0.015322616323828697
step: 420, loss: 0.19424153864383698
step: 430, loss: 0.003091546706855297
step: 440, loss: 0.0032645552419126034
step: 450, loss: 0.054036471992731094
step: 460, loss: 0.06515383720397949
step: 470, loss: 0.11395596712827682
step: 480, loss: 0.004905366338789463
step: 490, loss: 0.006538409274071455
step: 500, loss: 0.014749010093510151
step: 510, loss: 0.010048478841781616
step: 520, loss: 0.14688417315483093
step: 530, loss: 0.06721779704093933
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9432753888380604, f1=0.9370437956204379,
best_f1=0.9370437956204379
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.008904380723834038
step: 10, loss: 0.003949452191591263
step: 20, loss: 0.004837804473936558
step: 30, loss: 0.0017339453333988786
step: 40, loss: 0.0029417197220027447
step: 50, loss: 0.00964549370110035
step: 60, loss: 0.009431998245418072
step: 70, loss: 0.0030820576939731836
step: 80, loss: 0.003360083093866706
step: 90, loss: 0.09238351136445999
step: 100, loss: 0.004862761590629816
step: 110, loss: 0.008742421865463257
step: 120, loss: 0.0011908684391528368
step: 130, loss: 0.003645041724666953
step: 140, loss: 0.004497702233493328
step: 150, loss: 0.004230888094753027
step: 160, loss: 0.004655010532587767
```

```
step: 170, loss: 0.0016838803421705961
step: 180, loss: 0.005443362053483725
step: 190, loss: 0.013484693132340908
step: 200, loss: 0.008702441118657589
step: 210, loss: 0.027105871587991714
step: 220, loss: 0.0038557497318834066
step: 230, loss: 0.10772474110126495
step: 240, loss: 0.0042862980626523495
step: 250, loss: 0.03561095893383026
step: 260, loss: 0.0019520529313012958
step: 270, loss: 0.012215596623718739
step: 280, loss: 0.018566159531474113
step: 290, loss: 0.02200276404619217
step: 300, loss: 0.00029906368581578135
step: 310, loss: 0.002358893398195505
step: 320, loss: 0.05648963898420334
step: 330, loss: 0.001366603304632008
step: 340, loss: 0.060478899627923965
step: 350, loss: 0.009563934989273548
step: 360, loss: 0.03953031823039055
step: 370, loss: 0.006730073131620884
step: 380, loss: 0.019876591861248016
step: 390, loss: 0.0046530296094715595
step: 400, loss: 0.011304252780973911
step: 410, loss: 0.00978096667677164
step: 420, loss: 0.008133253082633018
step: 430, loss: 0.11054114997386932
Gradient overflow.  Skipping step, loss scaler 0 reducing loss scale to 32768.0
step: 440, loss: 0.025807706639170647
step: 450, loss: 0.01099482737481594
step: 460, loss: 0.01945132575929165
step: 470, loss: 0.05960230156779289
step: 480, loss: 0.10183247178792953
step: 490, loss: 0.004796828608959913
step: 500, loss: 0.007065026555210352
step: 510, loss: 0.06541437655687332
step: 520, loss: 0.03248481824994087
step: 530, loss: 0.022880196571350098
/usr/local/lib/python3.7/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=0.9361305361305362, f1=0.9333333333333333,
best_f1=0.9370437956204379
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.08912589401006699
step: 10, loss: 0.03862324729561806
```

```
step: 20, loss: 0.006208847276866436
step: 30, loss: 0.003753799246624112
step: 40, loss: 0.02384381927549839
step: 50, loss: 0.004383810795843601
step: 60, loss: 0.12086358666419983
step: 70, loss: 0.19680537283420563
step: 80, loss: 0.005156306549906731
step: 90, loss: 0.006354981102049351
step: 100, loss: 0.06642155349254608
step: 110, loss: 0.0005409918958321214
step: 120, loss: 0.005366073921322823
step: 130, loss: 0.0004383940249681473
step: 140, loss: 0.002132181078195572
step: 150, loss: 0.020698845386505127
step: 160, loss: 0.00026473868638277054
step: 170, loss: 0.0624498575925827
step: 180, loss: 0.001005880651064217
step: 190, loss: 0.0011539848055690527
step: 200, loss: 0.0028791711665689945
step: 210, loss: 0.0006793620996177197
step: 220, loss: 0.010668179020285606
step: 230, loss: 0.008349776268005371
step: 240, loss: 0.0032138654496520758
step: 250, loss: 0.0017709607491269708
step: 260, loss: 0.010667594149708748
step: 270, loss: 0.002001209184527397
step: 280, loss: 0.002103407634422183
step: 290, loss: 0.12426038831472397
step: 300, loss: 0.000418245792388916
step: 310, loss: 0.000402203033445403
step: 320, loss: 0.11806940287351608
step: 330, loss: 0.11429297178983688
step: 340, loss: 0.002484875964000821
step: 350, loss: 0.0026825706008821726
step: 360, loss: 0.005277905613183975
step: 370, loss: 0.055437564849853516
step: 380, loss: 0.0015057047130540013
step: 390, loss: 0.000498096109367907
step: 400, loss: 0.005715100094676018
step: 410, loss: 0.0009538005106151104
step: 420, loss: 0.007912706583738327
step: 430, loss: 0.006325682625174522
step: 440, loss: 0.02003752999007702
step: 450, loss: 0.019340885803103447
step: 460, loss: 0.00104944605845958
step: 470, loss: 0.005369126331061125
step: 480, loss: 0.10024020820856094
step: 490, loss: 0.0012233038432896137
```

```
step: 500, loss: 0.018038801848888397
step: 510, loss: 0.02371320128440857
step: 520, loss: 0.020172329619526863
step: 530, loss: 0.005414656363427639
/usr/local/lib/python3.7/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=0.940566037735849, f1=0.9307116104868914,
best_f1=0.9370437956204379
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0009600095800124109
step: 10, loss: 0.0005235603312030435
step: 20, loss: 0.10374331474304199
step: 30, loss: 0.0021930246148258448
step: 40, loss: 0.00031540298368781805
step: 50, loss: 0.049377523362636566
step: 60, loss: 0.00014161382569000125
step: 70, loss: 0.00023862929083406925
step: 80, loss: 0.00018618536705616862
step: 90, loss: 0.00017717316222842783
step: 100, loss: 0.061445996165275574
step: 110, loss: 0.0005773320444859564
step: 120, loss: 0.0004289918579161167
step: 130, loss: 0.005656825844198465
step: 140, loss: 0.0027461564168334007
step: 150, loss: 0.002017123159021139
step: 160, loss: 0.000824645918328315
step: 170, loss: 0.0033330018632113934
step: 180, loss: 0.0015987918013706803
step: 190, loss: 0.0014492650516331196
step: 200, loss: 0.005687995348125696
step: 210, loss: 0.01502626109868288
step: 220, loss: 0.010803074575960636
step: 230, loss: 0.002216694178059697
step: 240, loss: 0.03789214789867401
step: 250, loss: 0.06622784584760666
step: 260, loss: 0.0009032702655531466
step: 270, loss: 0.08040814846754074
step: 280, loss: 0.002779910573735833
step: 290, loss: 0.0005645607598125935
step: 300, loss: 0.0009104465716518462
step: 310, loss: 0.0005275861476548016
step: 320, loss: 0.0004045942041557282
step: 330, loss: 0.0008874892955645919
step: 340, loss: 0.12026768177747726
step: 350, loss: 0.0036504080053418875
```

```
step: 360, loss: 0.01670900732278824
step: 370, loss: 0.006203827448189259
step: 380, loss: 0.0010106973350048065
step: 390, loss: 0.008107256144285202
step: 400, loss: 0.00018949908553622663
step: 410, loss: 0.00024157826555892825
step: 420, loss: 0.05080390349030495
step: 430, loss: 0.0006488944054581225
step: 440, loss: 0.00013730583305004984
step: 450, loss: 0.0002755539317149669
step: 460, loss: 0.000630476395599544
step: 470, loss: 0.003791116876527667
step: 480, loss: 0.05642838403582573
step: 490, loss: 0.007940880954265594
step: 500, loss: 0.000495652318932116
step: 510, loss: 0.011270133778452873
step: 520, loss: 0.007844940759241581
step: 530, loss: 0.015047477558255196
/usr/local/lib/python3.7/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=0.937037037037037, f1=0.9345622119815669,
best_f1=0.9370437956204379
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.001762198400683701
step: 10, loss: 0.04614872485399246
step: 20, loss: 0.08495283126831055
step: 30, loss: 0.003821341088041663
step: 40, loss: 0.017215527594089508
step: 50, loss: 0.0032850028946995735
step: 60, loss: 0.0031086658127605915
step: 70, loss: 0.0012383689172565937
step: 80, loss: 0.0016527457628399134
step: 90, loss: 0.0002535604580771178
step: 100, loss: 0.00026045070262625813
step: 110, loss: 0.029041487723588943
step: 120, loss: 0.0006995477015152574
step: 130, loss: 0.012135009281337261
step: 140, loss: 0.0007879044278524816
step: 150, loss: 0.00040070584509521723
step: 160, loss: 0.00011132559302495793
step: 170, loss: 0.0007918429910205305
step: 180, loss: 0.0012603342765942216
step: 190, loss: 0.003316395916044712
step: 200, loss: 0.00012022574810544029
step: 210, loss: 0.04373618960380554
```

```
step: 220, loss: 0.0003124377108179033
step: 230, loss: 0.014501760713756084
step: 240, loss: 0.0021739862859249115
step: 250, loss: 0.0036295459140092134
step: 260, loss: 7.824500789865851e-05
step: 270, loss: 0.003978237509727478
step: 280, loss: 0.00014809585991315544
step: 290, loss: 0.00016840647731442004
step: 300, loss: 0.0012778755044564605
step: 310, loss: 4.060022911289707e-05
step: 320, loss: 0.0002453448250889778
step: 330, loss: 0.062389880418777466
step: 340, loss: 0.029132938012480736
step: 350, loss: 0.0010760592995211482
step: 360, loss: 0.0026682044845074415
step: 370, loss: 0.0006179802003316581
step: 380, loss: 0.0006062220199964941
step: 390, loss: 0.011426093056797981
step: 400, loss: 0.015713118016719818
step: 410, loss: 0.001998194260522723
step: 420, loss: 0.0014009373262524605
step: 430, loss: 0.010004639625549316
step: 440, loss: 0.0022888407111167908
step: 450, loss: 0.0020676949061453342
step: 460, loss: 0.001231490634381771
step: 470, loss: 0.007058532442897558
step: 480, loss: 0.001145065645687282
step: 490, loss: 0.0003064429620280862
step: 500, loss: 0.0005392829771153629
step: 510, loss: 0.005440752487629652
step: 520, loss: 0.05500676855444908
step: 530, loss: 0.013700399547815323
/usr/local/lib/python3.7/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=0.9424964936886395, f1=0.9350770667912189,
best_f1=0.9370437956204379
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0032278711441904306
step: 10, loss: 0.011908264830708504
step: 20, loss: 0.0007992186001501977
step: 30, loss: 0.11784708499908447
step: 40, loss: 0.005061198491603136
step: 50, loss: 0.0024486894253641367
step: 60, loss: 0.005445966962724924
step: 70, loss: 0.0007303571328520775
```

```
step: 80, loss: 0.0013811865355819464
step: 90, loss: 0.0018923073075711727
step: 100, loss: 0.0003188431728631258
step: 110, loss: 0.0014963001012802124
step: 120, loss: 0.0007198599632829428
step: 130, loss: 0.003525699721649289
step: 140, loss: 0.001964612863957882
step: 150, loss: 0.0008087062160484493
step: 160, loss: 0.003417676780372858
step: 170, loss: 0.011292004957795143
step: 180, loss: 0.0005249793757684529
step: 190, loss: 0.00306318630464375
step: 200, loss: 0.0014358183834701777
step: 210, loss: 0.010200487449765205
step: 220, loss: 0.0021623969078063965
step: 230, loss: 0.0011309412075206637
step: 240, loss: 0.020101821050047874
step: 250, loss: 0.002484625671058893
step: 260, loss: 0.00016601852257736027
step: 270, loss: 0.013582170009613037
step: 280, loss: 0.07201102375984192
step: 290, loss: 0.0003168561961501837
step: 300, loss: 0.0041598966345191
step: 310, loss: 0.0031282398849725723
step: 320, loss: 0.0008234693668782711
step: 330, loss: 0.0006481494638137519
step: 340, loss: 0.0006094048731029034
step: 350, loss: 0.0008307248936034739
step: 360, loss: 0.012036353349685669
step: 370, loss: 0.0004903123481199145
step: 380, loss: 0.004625431727617979
step: 390, loss: 0.1348719298839569
step: 400, loss: 0.0003047392820008099
step: 410, loss: 0.0009299339144490659
step: 420, loss: 0.004049584269523621
step: 430, loss: 0.025896424427628517
step: 440, loss: 0.01230963971465826
step: 450, loss: 0.00011028658627765253
step: 460, loss: 0.0005138297565281391
step: 470, loss: 0.00035739768645726144
step: 480, loss: 0.0004346403875388205
Gradient overflow.  Skipping step, loss scaler 0 reducing loss scale to 32768.0
step: 490, loss: 0.003082094481214881
step: 500, loss: 0.002112076384946704
step: 510, loss: 0.0006430052453652024
step: 520, loss: 0.002551455283537507
step: 530, loss: 0.0003768884635064751
/usr/local/lib/python3.7/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=0.9304267161410019, f1=0.9330889092575618,
best_f1=0.9370437956204379
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.005385462660342455
step: 10, loss: 0.015247256495058537
step: 20, loss: 0.01507467683404684
step: 30, loss: 0.0005995542160235345
step: 40, loss: 0.0003961297916248441
step: 50, loss: 0.0019224494462832808
step: 60, loss: 0.00015256600454449654
step: 70, loss: 0.0008623882313258946
step: 80, loss: 0.0035148952156305313
step: 90, loss: 7.394226850010455e-05
step: 100, loss: 0.00016357410640921444
step: 110, loss: 0.0001564693229738623
step: 120, loss: 0.00041850589332170784
step: 130, loss: 0.0029187400359660387
step: 140, loss: 0.02672170288860798
step: 150, loss: 0.005857937969267368
step: 160, loss: 0.000993768684566021
step: 170, loss: 0.014497078955173492
step: 180, loss: 0.00019159112707711756
step: 190, loss: 0.00041046395199373364
step: 200, loss: 0.0009450875222682953
step: 210, loss: 0.00024366623256355524
step: 220, loss: 0.03561023250222206
step: 230, loss: 0.00019716887618415058
step: 240, loss: 0.0002027001464739442
step: 250, loss: 0.019835907965898514
step: 260, loss: 0.029962463304400444
step: 270, loss: 0.00010657548409653828
step: 280, loss: 0.0025710423942655325
step: 290, loss: 0.03187975659966469
step: 300, loss: 0.00030262547079473734
step: 310, loss: 0.017620563507080078
step: 320, loss: 0.0002608778013382107
step: 330, loss: 0.0008856051717884839
step: 340, loss: 0.0008817656198516488
step: 350, loss: 0.04030049964785576
step: 360, loss: 0.00010748279601102695
step: 370, loss: 0.004427964333444834
step: 380, loss: 6.406501779565588e-05
step: 390, loss: 6.844118615845218e-05
step: 400, loss: 0.013918484561145306
```

```
step: 410, loss: 0.0003684677940327674
step: 420, loss: 0.021275997161865234
step: 430, loss: 7.153901970013976e-05
step: 440, loss: 0.0012883482268080115
step: 450, loss: 7.067326077958569e-05
step: 460, loss: 6.823390140198171e-05
step: 470, loss: 3.6274032026994973e-05
step: 480, loss: 5.947234603809193e-05
step: 490, loss: 0.0018937716959044337
step: 500, loss: 0.0008308116812258959
step: 510, loss: 0.00022044249635655433
step: 520, loss: 7.815063872840255e-05
step: 530, loss: 6.146530358819291e-05
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
epoch 9: dev_f1=0.9364107395195479, f1=0.9243065350258581,
best_f1=0.9370437956204379
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.006051960866898298
step: 10, loss: 0.00013817936996929348
step: 20, loss: 3.358930553076789e-05
step: 30, loss: 0.00018263564561493695
step: 40, loss: 0.0018219614867120981
step: 50, loss: 0.00010461773490533233
step: 60, loss: 0.00010282814037054777
step: 70, loss: 0.0001090989971999079
step: 80, loss: 4.778481161338277e-05
step: 90, loss: 0.01678111031651497
step: 100, loss: 0.00013860267063137144
step: 110, loss: 3.404427116038278e-05
step: 120, loss: 0.0014134857337921858
step: 130, loss: 0.00012495573901105672
step: 140, loss: 0.014825576916337013
step: 150, loss: 0.0007151653990149498
step: 160, loss: 8.083043212536722e-05
step: 170, loss: 0.0005307603860273957
step: 180, loss: 0.0008418934303335845
step: 190, loss: 0.0004950478905811906
step: 200, loss: 0.0008719688630662858
step: 210, loss: 9.26822394831106e-05
step: 220, loss: 0.006014689803123474
step: 230, loss: 0.003741063410416245
step: 240, loss: 0.007587210275232792
step: 250, loss: 0.0004492154985200614
step: 260, loss: 0.0012634817976504564
```

```
step: 270, loss: 0.0005070297047495842
step: 280, loss: 0.0005063664866611362
step: 290, loss: 0.00020957400556653738
step: 300, loss: 0.0037484089843928814
step: 310, loss: 0.0003292392357252538
step: 320, loss: 0.0012908661738038063
step: 330, loss: 0.0017994447844102979
step: 340, loss: 0.009158669970929623
step: 350, loss: 0.00013048553955741227
step: 360, loss: 0.0007071148720569909
step: 370, loss: 0.003760724328458309
step: 380, loss: 0.3022940754890442
step: 390, loss: 0.0005947727477177978
step: 400, loss: 0.013501091860234737
step: 410, loss: 0.016940820962190628
step: 420, loss: 0.0007826121291145682
step: 430, loss: 0.00018365599680691957
step: 440, loss: 0.0018448163755238056
step: 450, loss: 5.6601118558319286e-05
step: 460, loss: 0.0007708513294346631
step: 470, loss: 0.004652512725442648
step: 480, loss: 0.00027325566043145955
step: 490, loss: 0.0001858882897067815
step: 500, loss: 0.005451507866382599
step: 510, loss: 0.013291459530591965
step: 520, loss: 0.0001456875033909455
step: 530, loss: 0.0016932319849729538
/usr/local/lib/python3.7/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=0.941230911614993, f1=0.9344413665743306,
best_f1=0.9370437956204379
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0001160383180831559
step: 10, loss: 0.0007091814768500626
step: 20, loss: 0.0004310276999603957
step: 30, loss: 0.0006009223288856447
step: 40, loss: 0.002942407503724098
step: 50, loss: 0.00040150509448722005
step: 60, loss: 0.050766997039318085
step: 70, loss: 6.283899710979313e-05
step: 80, loss: 0.001326612662523985
step: 90, loss: 3.403505979804322e-05
step: 100, loss: 0.00148426101077348
step: 110, loss: 0.0026008840650320053
step: 120, loss: 0.00030794565100222826
```

```
step: 130, loss: 0.00015328630979638547
step: 140, loss: 0.00035586333251558244
step: 150, loss: 3.177203325321898e-05
step: 160, loss: 4.626892041414976e-05
step: 170, loss: 0.00010454103903612122
step: 180, loss: 0.0031863427720963955
step: 190, loss: 0.0004815188003703952
step: 200, loss: 0.005612696520984173
step: 210, loss: 0.0003301371762063354
step: 220, loss: 0.0021217993926256895
step: 230, loss: 0.000613864918705076
step: 240, loss: 0.0005558126140385866
step: 250, loss: 2.5334826204925776e-05
step: 260, loss: 0.0005404465482570231
step: 270, loss: 0.0013835816644132137
step: 280, loss: 0.0002080425329040736
step: 290, loss: 0.000659829645883292
step: 300, loss: 0.002165385289117694
step: 310, loss: 0.0015665048267692327
step: 320, loss: 0.0003279412048868835
step: 330, loss: 0.000719727948307991
step: 340, loss: 0.013687783852219582
step: 350, loss: 0.00015227605763357133
step: 360, loss: 0.00015920065925456583
step: 370, loss: 0.010535934008657932
step: 380, loss: 0.0006844506715424359
step: 390, loss: 3.6636778531828895e-05
step: 400, loss: 0.00010961174848489463
step: 410, loss: 0.0010880357585847378
step: 420, loss: 0.0020641740411520004
step: 430, loss: 0.00010380493040429428
step: 440, loss: 0.0012729560257866979
step: 450, loss: 0.0007002807105891407
step: 460, loss: 0.04299754649400711
step: 470, loss: 0.0010000338079407811
step: 480, loss: 0.0014090609038248658
step: 490, loss: 0.0020889148581773043
step: 500, loss: 0.0005430825985968113
step: 510, loss: 0.01497254054993391
step: 520, loss: 3.415363607928157e-05
step: 530, loss: 0.0013677716488018632
/usr/local/lib/python3.7/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=0.9362511893434825, f1=0.9275225011842728,
best_f1=0.9370437956204379
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
```

```
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 8.152195368893445e-05
step: 10, loss: 0.000594529090449214
step: 20, loss: 0.00014212523819878697
step: 30, loss: 0.0029545719735324383
step: 40, loss: 0.0012513544643297791
step: 50, loss: 0.014616010710597038
step: 60, loss: 0.002761795883998275
step: 70, loss: 0.0004371034156065434
step: 80, loss: 0.0008627185598015785
step: 90, loss: 0.001926093245856464
step: 100, loss: 0.004140346776694059
step: 110, loss: 0.000623348809313029
step: 120, loss: 7.005884981481358e-05
step: 130, loss: 0.0012913875980302691
step: 140, loss: 2.618433973111678e-05
step: 150, loss: 8.234747656388208e-05
step: 160, loss: 0.0002777794434223324
step: 170, loss: 0.00013949688582215458
step: 180, loss: 0.00013304651656653732
step: 190, loss: 7.831682887626812e-05
step: 200, loss: 0.0009099471499212086
step: 210, loss: 0.10664243251085281
step: 220, loss: 0.06698650866746902
step: 230, loss: 2.748431325017009e-05
step: 240, loss: 0.029580306261777878
step: 250, loss: 0.00020138385298196226
step: 260, loss: 0.000402259174734354
step: 270, loss: 0.04797749221324921
step: 280, loss: 0.0020235555712133646
step: 290, loss: 0.0010640803957358003
step: 300, loss: 0.008988097310066223
step: 310, loss: 0.0017925251740962267
step: 320, loss: 0.0008576316758990288
step: 330, loss: 0.0006665520486421883
step: 340, loss: 0.0005333898006938398
step: 350, loss: 0.005688670556992292
step: 360, loss: 0.0008250526152551174
step: 370, loss: 0.007181132212281227
step: 380, loss: 0.00036131354863755405
step: 390, loss: 0.0002527305914554745
step: 400, loss: 9.957607835531235e-05
step: 410, loss: 0.0008033594349399209
step: 420, loss: 0.00275238836184144
Gradient overflow.  Skipping step, loss scaler 0 reducing loss scale to 32768.0
step: 430, loss: 0.016737356781959534
step: 440, loss: 0.0005517117097042501
step: 450, loss: 0.054982807487249374
```

```
step: 460, loss: 3.206576730008237e-05
step: 470, loss: 5.954884545644745e-05
step: 480, loss: 0.0006215705070644617
step: 490, loss: 0.0004638737300410867
step: 500, loss: 9.271187445847318e-05
step: 510, loss: 0.001555637107230723
step: 520, loss: 3.0359529773704708e-05
step: 530, loss: 0.00069133365214131773
/usr/local/lib/python3.7/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=0.9358490566037737, f1=0.9256820319849483,
best_f1=0.9370437956204379
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00042412642505951226
step: 10, loss: 6.520071474369615e-05
step: 20, loss: 0.00055231514852494
step: 30, loss: 0.0001893620501505211
step: 40, loss: 3.247551649110392e-05
step: 50, loss: 0.00194718805141747
step: 60, loss: 0.0004938740748912096
step: 70, loss: 6.836315151304007e-05
step: 80, loss: 0.0002690812689252198
step: 90, loss: 0.024553533643484116
step: 100, loss: 0.0005218240548856556
step: 110, loss: 0.0002920805709436536
step: 120, loss: 0.00018999757594428957
step: 130, loss: 0.00015150191029533744
step: 140, loss: 0.004965293221175671
step: 150, loss: 0.0013258750550448895
step: 160, loss: 5.111169957672246e-05
step: 170, loss: 5.337366383173503e-05
step: 180, loss: 5.363774471334182e-05
step: 190, loss: 0.00015978625742718577
step: 200, loss: 6.482560274889693e-05
step: 210, loss: 0.00025843645562417805
step: 220, loss: 5.9895464801229537e-05
step: 230, loss: 6.072139149182476e-05
step: 240, loss: 0.00031323067378252745
step: 250, loss: 0.00014778853801544756
step: 260, loss: 0.007190557196736336
step: 270, loss: 4.2638937884476036e-05
step: 280, loss: 2.437768125673756e-05
step: 290, loss: 9.527287329547107e-05
step: 300, loss: 7.53647182136774e-05
step: 310, loss: 0.005774525459855795
```

```
step: 320, loss: 0.00256898021325469
step: 330, loss: 0.0014097450766712427
step: 340, loss: 2.826956369972322e-05
step: 350, loss: 3.246363849029876e-05
step: 360, loss: 0.000107012310763821
step: 370, loss: 0.0004137996002100408
step: 380, loss: 0.00011010811431333423
step: 390, loss: 0.0024486996699124575
step: 400, loss: 0.0002931435010395944
step: 410, loss: 0.006553260609507561
step: 420, loss: 0.03586304932832718
step: 430, loss: 0.0006816595559939742
step: 440, loss: 0.0001949752913787961
step: 450, loss: 0.00514135230332613
step: 460, loss: 0.013235070742666721
step: 470, loss: 0.00020710896933451295
step: 480, loss: 0.00017352664144709706
step: 490, loss: 3.4614022297319025e-05
step: 500, loss: 0.0006529099773615599
step: 510, loss: 0.05326385796070099
step: 520, loss: 0.0008923950372263789
step: 530, loss: 0.00018046863260678947
/usr/local/lib/python3.7/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=0.9345266132830901, f1=0.9275634995296332,
best_f1=0.9370437956204379
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0019172033062204719
step: 10, loss: 3.61174825229682e-05
step: 20, loss: 9.232215961674228e-05
step: 30, loss: 4.0086139051709324e-05
step: 40, loss: 0.00232674740254879
step: 50, loss: 0.00021519461006391793
step: 60, loss: 2.7644400688586757e-05
step: 70, loss: 0.0004288137424737215
step: 80, loss: 3.101552647422068e-05
step: 90, loss: 6.129900430096313e-05
step: 100, loss: 0.02506747469305992
step: 110, loss: 0.00010501797805773094
step: 120, loss: 0.0005138287669979036
step: 130, loss: 0.008153699338436127
step: 140, loss: 0.004576202016323805
step: 150, loss: 0.029725363478064537
step: 160, loss: 0.0009786138543859124
step: 170, loss: 0.0013019187608733773
```

```
step: 180, loss: 4.111107045901008e-05
step: 190, loss: 0.0003756332735065371
step: 200, loss: 0.005776016507297754
step: 210, loss: 3.960657340940088e-05
step: 220, loss: 0.00011576130782486871
step: 230, loss: 0.001357392524369061
step: 240, loss: 0.00438730837777257
step: 250, loss: 9.847761248238385e-05
step: 260, loss: 0.0008594162063673139
step: 270, loss: 0.00038569417665712535
step: 280, loss: 0.00042038559331558645
step: 290, loss: 0.00027351052267476916
step: 300, loss: 0.0010623699054121971
step: 310, loss: 0.00015276740305125713
step: 320, loss: 0.001136264530941844
step: 330, loss: 0.0028000620659440756
step: 340, loss: 0.0007712118094787002
step: 350, loss: 0.002578641753643751
step: 360, loss: 0.0010320640867576003
step: 370, loss: 6.289142766036093e-05
step: 380, loss: 0.0003593284636735916
step: 390, loss: 0.05029962211847305
step: 400, loss: 9.075069101527333e-05
step: 410, loss: 4.59542716271244e-05
step: 420, loss: 0.0006217111949808896
step: 430, loss: 0.000752398103941232
step: 440, loss: 0.00033534131944179535
step: 450, loss: 0.00016270054038614035
step: 460, loss: 0.00035480171209201217
step: 470, loss: 0.0002093836519634351
step: 480, loss: 5.6690321798669174e-05
step: 490, loss: 7.974875188665465e-05
step: 500, loss: 4.58138165413402e-05
step: 510, loss: 0.04684567078948021
step: 520, loss: 0.0002846584829967469
step: 530, loss: 0.0010326525662094355
/usr/local/lib/python3.7/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=0.9368372521899492, f1=0.9364426154549611,
best_f1=0.9370437956204379
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 4.440081465872936e-05
step: 10, loss: 5.664639320457354e-05
step: 20, loss: 8.526291640009731e-05
step: 30, loss: 0.00013054319424554706
```

```
step: 40, loss: 0.08032801002264023
step: 50, loss: 5.145139220985584e-05
step: 60, loss: 4.88908844999969e-05
step: 70, loss: 0.0008211205713450909
step: 80, loss: 4.557389911497012e-05
step: 90, loss: 4.3948082748102024e-05
step: 100, loss: 0.00045650440733879805
step: 110, loss: 0.0012246210826560855
step: 120, loss: 0.00019105669343844056
step: 130, loss: 0.10403910279273987
step: 140, loss: 4.557523425319232e-05
step: 150, loss: 8.243363117799163e-05
step: 160, loss: 5.547218461288139e-05
step: 170, loss: 0.00016803100879769772
step: 180, loss: 9.58741147769615e-05
step: 190, loss: 0.0001658374967519194
step: 200, loss: 0.00011551322677405551
step: 210, loss: 0.0015526128700003028
step: 220, loss: 6.634386227233335e-05
step: 230, loss: 0.0010198801755905151
step: 240, loss: 4.60703122371342e-05
step: 250, loss: 0.0001810524263419211
step: 260, loss: 0.0055924407206475735
step: 270, loss: 1.7426591512048617e-05
step: 280, loss: 3.0436527595156804e-05
step: 290, loss: 2.416899769741576e-05
step: 300, loss: 4.34864305134397e-05
step: 310, loss: 0.0005451447796076536
step: 320, loss: 0.0007119920337572694
step: 330, loss: 0.00034998764749616385
step: 340, loss: 0.00019947152759414166
step: 350, loss: 0.002161626936867833
step: 360, loss: 0.0018804160645231605
step: 370, loss: 0.00016275219968520105
step: 380, loss: 0.0006496002897620201
step: 390, loss: 0.005795896053314209
step: 400, loss: 0.0011329872068017721
step: 410, loss: 0.0007592998445034027
step: 420, loss: 7.051281863823533e-05
step: 430, loss: 0.0003246846026740968
step: 440, loss: 8.699990576133132e-05
step: 450, loss: 5.291198976919986e-05
step: 460, loss: 2.8795404432457872e-05
step: 470, loss: 0.0014030815800651908
step: 480, loss: 3.7615784094668925e-05
step: 490, loss: 1.537407297291793e-05
step: 500, loss: 0.0002840383385773748
step: 510, loss: 7.055851892801002e-05
```

```
step: 520, loss: 0.00020830114954151213
step: 530, loss: 0.0005928856553509831
/usr/local/lib/python3.7/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=0.9370892018779343, f1=0.9314685314685315,
best_f1=0.9370437956204379
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 [00:22, 252.71it/s]
load_f1 = 0.9381584974805315
real_f1 = 0.9388686131386861
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 [00:17, 255.71it/s]
```