GLow_30_1_2_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 | 17.02 MiB/s, done.
Resolving deltas: 100% (142/142), done.
/content/ditto
Looking in indexes: https://pypi.org/simple, https://us-python.pkg.dev/colab-
wheels/public/simple/
Collecting gensim==3.8.1
  Downloading gensim-3.8.1-cp37-cp37m-manylinux1_x86_64.whl (24.2 MB)
     |                      | 24.2 MB 1.4 MB/s
Collecting numpy==1.19.2
  Downloading numpy-1.19.2-cp37-cp37m-manylinux2010_x86_64.whl (14.5 MB)
     |                      | 14.5 MB 60.9 MB/s
Collecting regex==2019.12.20
  Downloading regex-2019.12.20-cp37-cp37m-manylinux2010_x86_64.whl (689 kB)
     |                      | 689 kB 70.5 MB/s
Collecting scipy==1.3.2
  Downloading scipy-1.3.2-cp37-cp37m-manylinux1_x86_64.whl (25.2 MB)
     |                      | 25.2 MB 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 58.6 MB/s
Requirement already satisfied: sklearn==0.0 in
/usr/local/lib/python3.7/dist-packages (from -r requirements.txt (line 6)) (0.0)
```

1

```
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 63.8 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 | 9.02 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 30.2 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 76.6 MB/s
Collecting torch==1.4.0
  Downloading torch-1.4.0-cp37-cp37m-manylinux1_x86_64.whl (753.4 MB)
     |                        | 753.4 MB 7.0 kB/s
Collecting tqdm==4.41.0
  Downloading tqdm-4.41.0-py2.py3-none-any.whl (56 kB)
     |                        | 56 kB 6.1 MB/s
Collecting transformers==3.1.0
  Downloading transformers-3.1.0-py3-none-any.whl (884 kB)
     |                        | 884 kB 75.3 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 71.4 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: 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: 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)
Collecting plac<1.2.0,>=0.9.6
  Downloading plac-1.1.3-py2.py3-none-any.whl (20 kB)
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)
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)
Collecting srsly<1.1.0,>=0.1.0
  Downloading srsly-1.0.5-cp37-cp37m-manylinux2014_x86_64.whl (184 kB)
     |                        | 184 kB 63.8 MB/s
Collecting catalogue<1.1.0,>=0.0.7
  Downloading catalogue-1.0.0-py2.py3-none-any.whl (7.7 kB)
Collecting thinc<7.4.0,>=7.3.0
  Downloading thinc-7.3.1-cp37-cp37m-manylinux1_x86_64.whl (2.2 MB)
     |                        | 2.2 MB 66.0 MB/s
Requirement already satisfied: setuptools in
/usr/local/lib/python3.7/dist-packages (from spacy==2.2.3->-r requirements.txt
(line 4)) (57.4.0)
Requirement already satisfied: 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 blis<0.5.0,>=0.4.0
  Downloading blis-0.4.1-cp37-cp37m-manylinux1_x86_64.whl (3.7 MB)
     |                        | 3.7 MB 56.2 MB/s
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: 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: scikit-learn in /usr/local/lib/python3.7/dist-
packages (from sklearn==0.0->-r requirements.txt (line 6)) (1.0.2)
Requirement already satisfied: protobuf>=3.8.0 in /usr/local/lib/python3.7/dist-
packages (from tensorboardX==2.0->-r requirements.txt (line 8)) (3.17.3)
Collecting tokenizers==0.8.1.rc2
  Downloading tokenizers-0.8.1rc2-cp37-cp37m-manylinux1_x86_64.whl (3.0 MB)
     |                        | 3.0 MB 53.6 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)
Collecting sacremoses
  Downloading sacremoses-0.0.53.tar.gz (880 kB)
     | | 880 kB 65.6 MB/s
Requirement already satisfied: packaging in /usr/local/lib/python3.7/dist-packages (from transformers==3.1.0->-r requirements.txt (line 11)) (21.3)
Requirement already satisfied: 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: 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: 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: 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=1449923 sha256=5f7900927dba9d2075808d61dbee130f5a11a21360e3c545d4ed6edf9b51c2f4
  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=143113a2d18da688f20d4b3bd5d73ea7655626dc18864177b88a1df023fd933d
  Stored in directory: /root/.cache/pip/wheels/87/39/dd/a83eeef36d0bf98e7a4d1933
a4ad2d660295a40613079bafc9
Successfully built nltk sacremoses
Installing collected packages: numpy, tqdm, srsly, regex, plac, blis,
tokenizers, thinc, sentencepiece, sacremoses, catalogue, transformers, torch,
tensorboardX, spacy, nltk, jsonlines, gensim
  Attempting uninstall: numpy
    Found existing installation: numpy 1.21.6
    Uninstalling numpy-1.21.6:
      Successfully uninstalled numpy-1.21.6
  Attempting uninstall: tqdm
    Found existing installation: tqdm 4.64.0
    Uninstalling tqdm-4.64.0:
      Successfully uninstalled tqdm-4.64.0
  Attempting uninstall: srsly
    Found existing installation: srsly 2.4.4
    Uninstalling srsly-2.4.4:
      Successfully uninstalled srsly-2.4.4
  Attempting uninstall: regex
    Found existing installation: regex 2022.6.2
    Uninstalling regex-2022.6.2:
      Successfully uninstalled regex-2022.6.2
  Attempting uninstall: blis
    Found existing installation: blis 0.7.8
    Uninstalling blis-0.7.8:
      Successfully uninstalled blis-0.7.8
  Attempting uninstall: thinc
    Found existing installation: thinc 8.1.0
    Uninstalling thinc-8.1.0:
      Successfully uninstalled thinc-8.1.0
  Attempting uninstall: catalogue
    Found existing installation: catalogue 2.0.8
    Uninstalling catalogue-2.0.8:
      Successfully uninstalled catalogue-2.0.8
  Attempting uninstall: torch
    Found existing installation: torch 1.12.0+cu113
    Uninstalling torch-1.12.0+cu113:
      Successfully uninstalled torch-1.12.0+cu113
  Attempting uninstall: spacy
    Found existing installation: spacy 3.4.1
    Uninstalling spacy-3.4.1:
      Successfully uninstalled spacy-3.4.1
  Attempting uninstall: nltk
    Found existing installation: nltk 3.7
    Uninstalling nltk-3.7:
      Successfully uninstalled nltk-3.7
```

```
  Attempting uninstall: gensim
    Found existing installation: gensim 3.6.0
    Uninstalling gensim-3.6.0:
      Successfully uninstalled gensim-3.6.0
ERROR: pip's dependency resolver does not currently take into account all
the packages that are installed. This behaviour is the source of the following
dependency conflicts.
xarray-einstats 0.2.2 requires numpy>=1.21, but you have numpy 1.19.2 which is
incompatible.
torchvision 0.13.0+cu113 requires torch==1.12.0, but you have torch 1.4.0 which
is incompatible.
torchtext 0.13.0 requires torch==1.12.0, but you have torch 1.4.0 which is
incompatible.
torchaudio 0.12.0+cu113 requires torch==1.12.0, but you have torch 1.4.0 which
is incompatible.
tensorflow 2.8.2+zzzcolab20220719082949 requires numpy>=1.20, but you have numpy
1.19.2 which is incompatible.
panel 0.12.1 requires tqdm>=4.48.0, but you have tqdm 4.41.0 which is
incompatible.
fastai 2.7.7 requires torch<1.13,>=1.7, but you have torch 1.4.0 which is
incompatible.
en-core-web-sm 3.4.0 requires spacy<3.5.0,>=3.4.0, but you have spacy 2.2.3
which is incompatible.
cmdstanpy 1.0.4 requires numpy>=1.21, but you have numpy 1.19.2 which is
incompatible.
Successfully installed blis-0.4.1 catalogue-1.0.0 gensim-3.8.1 jsonlines-1.2.0
nltk-3.4.5 numpy-1.19.2 plac-1.1.3 regex-2019.12.20 sacremoses-0.0.53
sentencepiece-0.1.85 spacy-2.2.3 srsly-1.0.5 tensorboardX-2.0 thinc-7.3.1
tokenizers-0.8.1rc2 torch-1.4.0 tqdm-4.41.0 transformers-3.1.0
/content/ditto
[nltk_data] Downloading package stopwords to /root/nltk_data…
[nltk_data]    Unzipping corpora/stopwords.zip.
```

[ ]: True

0.2 Install fp16 optimization

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

```
Cloning into 'apex'…
remote: Enumerating objects: 10085, done.
remote: Counting objects: 100% (201/201), done.
remote: Compressing objects: 100% (138/138), done.
remote: Total 10085 (delta 98), reused 131 (delta 59), pack-reused 9884
Receiving objects: 100% (10085/10085), 14.95 MiB | 13.49 MiB/s, done.
Resolving deltas: 100% (6902/6902), done.
/content/ditto/apex
Using pip 21.1.3 from /usr/local/lib/python3.7/dist-packages/pip (python 3.7)
Value for scheme.platlib does not match. Please report this to
<https://github.com/pypa/pip/issues/9617>
distutils: /usr/local/lib/python3.7/dist-packages
sysconfig: /usr/lib/python3.7/site-packages
Value for scheme.purelib does not match. Please report this to
<https://github.com/pypa/pip/issues/9617>
distutils: /usr/local/lib/python3.7/dist-packages
sysconfig: /usr/lib/python3.7/site-packages
Value for scheme.headers does not match. Please report this to
<https://github.com/pypa/pip/issues/9617>
distutils: /usr/local/include/python3.7/UNKNOWN
sysconfig: /usr/include/python3.7m/UNKNOWN
Value for scheme.scripts does not match. Please report this to
<https://github.com/pypa/pip/issues/9617>
distutils: /usr/local/bin
sysconfig: /usr/bin
Value for scheme.data does not match. Please report this to
<https://github.com/pypa/pip/issues/9617>
distutils: /usr/local
sysconfig: /usr
Additional context:
user = False
home = None
root = None
prefix = None
Non-user install because site-packages writeable
Created temporary directory: /tmp/pip-ephem-wheel-cache-ym70gx0y
Created temporary directory: /tmp/pip-req-tracker-m20tyguz
Initialized build tracking at /tmp/pip-req-tracker-m20tyguz
Created build tracker: /tmp/pip-req-tracker-m20tyguz
Entered build tracker: /tmp/pip-req-tracker-m20tyguz
```

```
Created temporary directory: /tmp/pip-install-s6qfvm15
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-1eeoeqsu
  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-m20tyguz'
    Running setup.py (path:/tmp/pip-req-build-1eeoeqsu/setup.py) egg_info for
package from file:///content/ditto/apex
    Created temporary directory: /tmp/pip-pip-egg-info-ro3w1yn1
    Running command python setup.py egg_info


    torch.__version__  = 1.4.0


    running egg_info
    creating /tmp/pip-pip-egg-info-ro3w1yn1/apex.egg-info
    writing /tmp/pip-pip-egg-info-ro3w1yn1/apex.egg-info/PKG-INFO
    writing dependency_links to /tmp/pip-pip-egg-info-ro3w1yn1/apex.egg-
info/dependency_links.txt
    writing top-level names to /tmp/pip-pip-egg-info-ro3w1yn1/apex.egg-
info/top_level.txt
    writing manifest file '/tmp/pip-pip-egg-info-ro3w1yn1/apex.egg-
info/SOURCES.txt'
    adding license file 'LICENSE'
    writing manifest file '/tmp/pip-pip-egg-info-ro3w1yn1/apex.egg-
info/SOURCES.txt'
  Source in /tmp/pip-req-build-1eeoeqsu 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-m20tyguz'
Created temporary directory: /tmp/pip-unpack-bj7ib6kq
Building wheels for collected packages: apex
  Created temporary directory: /tmp/pip-wheel-kh1loa2k
  Building wheel for apex (setup.py) …   Destination directory:
/tmp/pip-wheel-kh1loa2k
  Running command /usr/bin/python3 -u -c 'import io, os, sys, setuptools,
tokenize; sys.argv[0] = '"'"'/tmp/pip-req-build-1eeoeqsu/setup.py'"'"';
```

```
__file__='"'"'/tmp/pip-req-build-1eeoeqsu/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-kh1loa2k


  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-kh1loa2k/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=0a88c20167858771da39d82a19025da35f5dc99eabdb71f388d80fa8ce10bde9
  Stored in directory: /tmp/pip-ephem-wheel-cache-
ym70gx0y/wheels/6b/22/a2/1665526ee3c3061243260fd522525eca31398e04bfa5ad7e1f
Successfully built apex
Installing collected packages: apex
  Value for scheme.platlib does not match. Please report this to
<https://github.com/pypa/pip/issues/9617>
  distutils: /usr/local/lib/python3.7/dist-packages
  sysconfig: /usr/lib/python3.7/site-packages
  Value for scheme.purelib does not match. Please report this to
<https://github.com/pypa/pip/issues/9617>
  distutils: /usr/local/lib/python3.7/dist-packages
  sysconfig: /usr/lib/python3.7/site-packages
  Value for scheme.headers does not match. Please report this to
<https://github.com/pypa/pip/issues/9617>
  distutils: /usr/local/include/python3.7/apex
  sysconfig: /usr/include/python3.7m/apex
  Value for scheme.scripts does not match. Please report this to
<https://github.com/pypa/pip/issues/9617>
  distutils: /usr/local/bin
  sysconfig: /usr/bin
  Value for scheme.data does not match. Please report this to
<https://github.com/pypa/pip/issues/9617>
  distutils: /usr/local
  sysconfig: /usr
  Additional context:
  user = False
  home = None
  root = None
  prefix = None

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

Value for scheme.headers does not match. Please report this to
<https://github.com/pypa/pip/issues/9617>
distutils: /usr/local/include/python3.7/UNKNOWN
sysconfig: /usr/include/python3.7m/UNKNOWN
Value for scheme.scripts does not match. Please report this to
<https://github.com/pypa/pip/issues/9617>
distutils: /usr/local/bin
sysconfig: /usr/bin
Value for scheme.data does not match. Please report this to
<https://github.com/pypa/pip/issues/9617>
distutils: /usr/local
sysconfig: /usr
Additional context:
user = False
home = None
root = None
prefix = None
Successfully installed apex-0.1
Removed build tracker: '/tmp/pip-req-tracker-m20tyguz'
/content/ditto

```
# some issue with colab
!pip install --upgrade "urllib3==1.25.4" awscli
```

Looking in indexes: https://pypi.org/simple, https://us-python.pkg.dev/colab-wheels/public/simple/
Collecting urllib3==1.25.4
  Downloading urllib3-1.25.4-py2.py3-none-any.whl (125 kB)
     | | 125 kB 30.5 MB/s
Collecting awscli
  Downloading awscli-1.25.47-py3-none-any.whl (3.9 MB)
     | | 3.9 MB 66.5 MB/s
Collecting botocore==1.27.47
  Downloading botocore-1.27.47-py3-none-any.whl (9.0 MB)
     | | 9.0 MB 59.9 MB/s
Collecting s3transfer<0.7.0,>=0.6.0
  Downloading s3transfer-0.6.0-py3-none-any.whl (79 kB)
     | | 79 kB 10.5 MB/s
Collecting rsa<4.8,>=3.1.2
  Downloading rsa-4.7.2-py3-none-any.whl (34 kB)
Collecting colorama<0.4.5,>=0.2.5
  Downloading colorama-0.4.4-py2.py3-none-any.whl (16 kB)
Collecting docutils<0.17,>=0.10
  Downloading docutils-0.16-py2.py3-none-any.whl (548 kB)
     | | 548 kB 70.3 MB/s
Requirement already satisfied: PyYAML<5.5,>=3.10 in
/usr/local/lib/python3.7/dist-packages (from awscli) (3.13)
Collecting jmespath<2.0.0,>=0.7.1

```
  Downloading jmespath-1.0.1-py3-none-any.whl (20 kB)
Requirement already satisfied: python-dateutil<3.0.0,>=2.1 in
/usr/local/lib/python3.7/dist-packages (from botocore==1.27.47->awscli) (2.8.2)
Requirement already satisfied: six>=1.5 in /usr/local/lib/python3.7/dist-
packages (from python-dateutil<3.0.0,>=2.1->botocore==1.27.47->awscli) (1.15.0)
Requirement already satisfied: pyasn1>=0.1.3 in /usr/local/lib/python3.7/dist-
packages (from rsa<4.8,>=3.1.2->awscli) (0.4.8)
Installing collected packages: urllib3, jmespath, botocore, s3transfer, rsa,
docutils, colorama, awscli
  Attempting uninstall: urllib3
    Found existing installation: urllib3 1.24.3
    Uninstalling urllib3-1.24.3:
      Successfully uninstalled urllib3-1.24.3
  Attempting uninstall: rsa
    Found existing installation: rsa 4.9
    Uninstalling rsa-4.9:
      Successfully uninstalled rsa-4.9
  Attempting uninstall: docutils
    Found existing installation: docutils 0.17.1
    Uninstalling docutils-0.17.1:
      Successfully uninstalled docutils-0.17.1
ERROR: pip's dependency resolver does not currently take into account all

the packages that are installed. This behaviour is the source of the following

dependency conflicts.

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

incompatible.

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

incompatible.
Successfully installed awscli-1.25.47 botocore-1.27.47 colorama-0.4.4
docutils-0.16 jmespath-1.0.1 rsa-4.7.2 s3transfer-0.6.0 urllib3-1.25.4
```

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

```
Looking in indexes: https://pypi.org/simple, https://us-python.pkg.dev/colab-
wheels/public/simple/
Collecting torch==1.7.1
  Downloading torch-1.7.1-cp37-cp37m-manylinux1_x86_64.whl (776.8 MB)
     |                        | 776.8 MB 17 kB/s
Requirement already satisfied: 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 | 16.15 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/GLow_30_1_2/configs.json /content/ditto/
```

1 DA

1.1 DA STRUCTURED

1.1.1 Beer - Running the matcher

```
[ ]:  # Step 1:
      # train the wdc model with the save_model flag on
      # it will produce two files *_dev.pt and *_test.pt which are the
      # best checkpoints on the dev set and the test set.
      !CUDA_VISIBLE_DEVICES=0 python train_ditto.py \
        --task Structured/Beer \
        --batch_size 32 \
        --max_len 256 \
        --lr 3e-5 \
        --n_epochs 15 \
        --finetuning \
        --lm 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, 352kB/s]
Downloading: 100% 232k/232k [00:00<00:00, 352kB/s]
Downloading: 100% 440M/440M [00:08<00:00, 52.0MB/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: 0.5509892106056213
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.4571428571428571, f1=0.32432432432432434,
best_f1=0.32432432432432434
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.5032119750976562
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:

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=0.4761904761904762, f1=0.42857142857142855,
best_f1=0.42857142857142855
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.611190676689148
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.5263157894736842, f1=0.3939393939393939,
best_f1=0.3939393939393939
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.30005261301994324
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.64, f1=0.5625000000000001, best_f1=0.5625000000000001
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.2698204219341278
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.6956521739130435, 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: 0.4255441427230835
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:

```
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.7333333333333334,
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: 0.05421486869454384
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.6875000000000001,
best_f1=0.6875000000000001
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.017690740525722504
/usr/local/lib/python3.7/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=0.8125000000000001, f1=0.7333333333333334,
best_f1=0.6875000000000001
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0034423889592289925
/usr/local/lib/python3.7/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=0.689655172413793,
best_f1=0.6875000000000001
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.011498814448714256
/usr/local/lib/python3.7/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=0.8666666666666666, f1=0.7096774193548386,
best_f1=0.6875000000000001
```

```
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.005369340535253286
/usr/local/lib/python3.7/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=0.7096774193548386,
best_f1=0.6875000000000001
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00391433946788311
/usr/local/lib/python3.7/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=0.8387096774193549, f1=0.75, best_f1=0.6875000000000001
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00423455098643899
/usr/local/lib/python3.7/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=0.8387096774193549, f1=0.75, best_f1=0.6875000000000001
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0027639735490083694
/usr/local/lib/python3.7/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=0.8387096774193549, f1=0.75, best_f1=0.6875000000000001
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.004610847216099501
/usr/local/lib/python3.7/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=0.8387096774193549, f1=0.75, best_f1=0.6875000000000001
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, 134917.52it/s]
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
load_f1 = 0.7647058823529412
real_f1 = 0.742857142857143
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
4398it [00:22, 197.01it/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

# 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: 0.61360764503479
step: 10, loss: 0.6060507893562317
step: 20, loss: 0.29252126812934875
step: 30, loss: 0.0836595743894577
step: 40, loss: 0.265250563621521
step: 50, loss: 0.031064342707395554
step: 60, loss: 0.1298033595085144
step: 70, loss: 0.02207045815885067
step: 80, loss: 0.1375972032546997
step: 90, loss: 0.13806623220443726
step: 100, loss: 0.006056226324290037
step: 110, loss: 0.0412595197558403
step: 120, loss: 0.003394680330529809
step: 130, loss: 0.007184511981904507
step: 140, loss: 0.0015734051121398807
step: 150, loss: 0.009520241059362888
step: 160, loss: 0.008436452597379684
step: 170, loss: 0.024883117526769638
step: 180, loss: 0.008980232290923595
step: 190, loss: 0.003581976518034935
step: 200, loss: 0.17186276614665985
step: 210, loss: 0.00949359592050314
step: 220, loss: 0.015583659522235394
step: 230, loss: 0.06920652836561203
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9830124575311437, f1=0.9748858447488584,
best_f1=0.9748858447488584
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.001778988866135478
step: 10, loss: 0.0010173831833526492
step: 20, loss: 0.149968221783638
step: 30, loss: 0.2044285088777542
step: 40, loss: 0.07572926580905914

```
step: 50, loss: 0.0038105833809822798
step: 60, loss: 0.005344164092093706
step: 70, loss: 0.11141209304332733
step: 80, loss: 0.002032521413639188
step: 90, loss: 0.030820343643426895
step: 100, loss: 0.026891076937317848
step: 110, loss: 0.020916398614645004
step: 120, loss: 0.13848377764225006
step: 130, loss: 0.08541326224803925
step: 140, loss: 0.00373519747518003
step: 150, loss: 0.012934011407196522
step: 160, loss: 0.003929352853447199
step: 170, loss: 0.0024371235631406307
step: 180, loss: 0.0065992423333227634
step: 190, loss: 0.004592441953718662
step: 200, loss: 0.0014284143690019846
step: 210, loss: 0.00048077941755764186
step: 220, loss: 0.07583026587963104
step: 230, loss: 0.052222855389118195
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9899216125419933, f1=0.9887387387387387,
best_f1=0.9887387387387387
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0145880663767457
step: 10, loss: 0.0008764764061197639
step: 20, loss: 0.04063619300723076
step: 30, loss: 0.07550598680973053
step: 40, loss: 0.1502288579940796
step: 50, loss: 0.00995203573256731
step: 60, loss: 0.0029036090709269047
step: 70, loss: 0.016418475657701492
step: 80, loss: 0.0009566176449880004
step: 90, loss: 0.011895876377820969
step: 100, loss: 0.0023834598250687122
step: 110, loss: 0.0009680389193817973
step: 120, loss: 0.005158033687621355
step: 130, loss: 0.0015312707982957363
step: 140, loss: 0.0016029230318963528
step: 150, loss: 0.007489652372896671
step: 160, loss: 0.004613358993083239
```

```
step: 170, loss: 0.0036733457818627357
step: 180, loss: 0.004752924665808678
step: 190, loss: 0.003860162803903222
step: 200, loss: 0.0041845873929560184
step: 210, loss: 0.0013404831988736987
step: 220, loss: 0.0006585965165868402
step: 230, loss: 0.0004888717085123062
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.987709497206704,
best_f1=0.987709497206704
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0009260995429940522
step: 10, loss: 0.0005836343043483794
step: 20, loss: 0.0007531323935836554
step: 30, loss: 0.0007050394196994603
step: 40, loss: 0.0019735614769160748
step: 50, loss: 0.0015966343926265836
step: 60, loss: 0.0007179290987551212
step: 70, loss: 0.0012957908911630511
step: 80, loss: 0.002054451499134302
step: 90, loss: 0.0010225195437669754
step: 100, loss: 0.00034843917819671333
step: 110, loss: 0.0006075240671634674
step: 120, loss: 0.041891634464263916
step: 130, loss: 0.010944724082946777
step: 140, loss: 0.0009645503014326096
step: 150, loss: 0.153598263859744884
step: 160, loss: 0.03613527491688728
step: 170, loss: 0.008423724211752415
step: 180, loss: 0.0013857270823791623
step: 190, loss: 0.003168809926137328
step: 200, loss: 0.0009153328719548881
step: 210, loss: 0.1021098643541336
step: 220, loss: 0.0005398401990532875
step: 230, loss: 0.09273406118154526
/usr/local/lib/python3.7/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=0.9876819708846584, f1=0.984304932735426,
best_f1=0.987709497206704
```

```
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0018040071008726954
step: 10, loss: 0.0011056715156883001
step: 20, loss: 0.0027762586250901222
step: 30, loss: 0.011055098846554756
step: 40, loss: 0.0010542746167629957
step: 50, loss: 0.00040891391108743846
step: 60, loss: 0.0032910299487411976
step: 70, loss: 0.0003126460360363126
step: 80, loss: 0.00030490089557133615
step: 90, loss: 0.0013095347676426172
step: 100, loss: 0.0014885059790685773
step: 110, loss: 0.0015356160001829267
step: 120, loss: 0.00014682784967590123
step: 130, loss: 0.004203615710139275
step: 140, loss: 0.0007379847229458392
step: 150, loss: 0.042046431452035904
step: 160, loss: 0.000754447013605386
step: 170, loss: 0.015302612446248531
step: 180, loss: 0.00771752092987299
step: 190, loss: 0.003425250295549631
step: 200, loss: 0.001573648420162499
step: 210, loss: 0.0012247200356796384
step: 220, loss: 0.004613311495631933
step: 230, loss: 0.0007262814324349165
/usr/local/lib/python3.7/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=0.9875141884222476, f1=0.976054732041049,
best_f1=0.987709497206704
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0005588412168435752
step: 10, loss: 0.0007142314570955932
step: 20, loss: 0.0009316775831393898
step: 30, loss: 0.0011912856716662645
step: 40, loss: 0.0004058245103806257
step: 50, loss: 0.00024239761114586145
step: 60, loss: 0.00022450262622442096
step: 70, loss: 0.020429521799087524
step: 80, loss: 0.01049130316823721
step: 90, loss: 0.015370482578873634
step: 100, loss: 0.015816735103726387
step: 110, loss: 0.002717433264479041
step: 120, loss: 0.0008530538179911673
```

```
step: 130, loss: 0.0013184596318751574
step: 140, loss: 0.0002630821836646646
step: 150, loss: 0.0004346657660789788
step: 160, loss: 0.0007227725582197309
step: 170, loss: 0.0005477203521877527
step: 180, loss: 0.018213365226984024
step: 190, loss: 0.01418870035558939
step: 200, loss: 0.0006809272454120219
step: 210, loss: 0.0006750098546035588
step: 220, loss: 0.000954474788159132
step: 230, loss: 0.024846738204360008
/usr/local/lib/python3.7/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=0.9888392857142857, f1=0.9822616407982262,
best_f1=0.987709497206704
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0006073236581869423
step: 10, loss: 0.00034739854163490236
step: 20, loss: 0.0019870048854500055
step: 30, loss: 0.0011379417264834046
step: 40, loss: 0.00022489936964120716
step: 50, loss: 0.0004991954192519188
step: 60, loss: 0.0008439971134066582
step: 70, loss: 0.018087120726704597
step: 80, loss: 0.003673804458230734
step: 90, loss: 0.000179390495759435
step: 100, loss: 0.0002536542306188494
step: 110, loss: 0.00045365767437033355
step: 120, loss: 0.00018104715854860842
step: 130, loss: 0.0007800596067681909
step: 140, loss: 0.0003777756937779486
step: 150, loss: 0.002387842396274209
step: 160, loss: 0.09835189580917358
step: 170, loss: 0.004696089308708906
step: 180, loss: 0.002858423627912998
step: 190, loss: 0.0011876702774316072
step: 200, loss: 0.07623787969350815
step: 210, loss: 0.00021621248743031174
step: 220, loss: 0.0015257832128554583
step: 230, loss: 0.010907434858381748
/usr/local/lib/python3.7/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=0.9887133182844244, f1=0.9807037457434733,
best_f1=0.987709497206704
```

/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0006454878021031618
step: 10, loss: 0.0017904323758557439
step: 20, loss: 0.00023976605734787881
step: 30, loss: 0.0010267908219248056
step: 40, loss: 0.0004767374775838107
step: 50, loss: 0.0008320307242684066
step: 60, loss: 0.00018863870354834944
step: 70, loss: 0.0012320661917328835
step: 80, loss: 0.0009726324351504445
step: 90, loss: 0.0002097706455970183
step: 100, loss: 0.001443590852431953
step: 110, loss: 0.0030390319880098104
step: 120, loss: 0.0005823900573886931
step: 130, loss: 0.0009897882118821144
step: 140, loss: 0.0007755103870294988
step: 150, loss: 0.0019238763488829136
step: 160, loss: 0.0014514618087559938
step: 170, loss: 0.0006012141820974648
step: 180, loss: 0.0005461935070343316
step: 190, loss: 0.0002510317717678845
step: 200, loss: 0.00024350860621780157
step: 210, loss: 0.003192767035216093
step: 220, loss: 0.00029919121880084276
step: 230, loss: 0.0002080962440231815
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
epoch 8: dev_f1=0.9899216125419933, f1=0.9844097995545658,
best_f1=0.987709497206704
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00016832497203722596
step: 10, loss: 0.0005323499208316207
step: 20, loss: 0.0006506656063720584
step: 30, loss: 8.814300963422284e-05
step: 40, loss: 0.04501849785447121
step: 50, loss: 0.0027115014381706715
step: 60, loss: 0.0068714916706085205
step: 70, loss: 0.011982230469584465
step: 80, loss: 0.0028041619807481766
step: 90, loss: 0.000801150803454203
step: 100, loss: 0.0006667404668405652
step: 110, loss: 0.00045190146192908287
step: 120, loss: 8.201618038583547e-05

```
step: 130, loss: 9.108885569730774e-05
step: 140, loss: 0.00011377887858543545
step: 150, loss: 0.00031616969499737024
step: 160, loss: 0.00020624836906790733
step: 170, loss: 0.0009808932663872838
step: 180, loss: 0.0005565992323681712
step: 190, loss: 0.0004446184611879289
step: 200, loss: 0.0001305760524701327
step: 210, loss: 0.00023295250139199197
step: 220, loss: 8.090186747722328e-05
step: 230, loss: 0.0006242808303795755
/usr/local/lib/python3.7/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=0.9909706546275394, f1=0.9876265466816648,
best_f1=0.987709497206704
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00040160404751077294
step: 10, loss: 4.560279558063485e-05
step: 20, loss: 0.00016653905913699418
step: 30, loss: 0.01359280850738287
step: 40, loss: 9.769391908776015e-05
step: 50, loss: 0.00021185669174883515
step: 60, loss: 0.0018472623778507113
step: 70, loss: 0.00021024569286964834
step: 80, loss: 0.00019273464567959309
step: 90, loss: 0.0003737047372851521
step: 100, loss: 0.000681539298966527
step: 110, loss: 0.009923738427460194
step: 120, loss: 0.0002991129586007446
step: 130, loss: 0.00011337329488014802
step: 140, loss: 0.02940143086016178
step: 150, loss: 0.014705050736665726
step: 160, loss: 4.3562642531469464e-05
step: 170, loss: 0.00022615341003984213
step: 180, loss: 0.00011954231740674004
step: 190, loss: 0.012239170260727406
step: 200, loss: 6.141424819361418e-05
step: 210, loss: 6.663975364062935e-05
step: 220, loss: 0.00012248785060364753
step: 230, loss: 0.00011796412582043558
/usr/local/lib/python3.7/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=0.9898534385569334, f1=0.9853438556933484,
best_f1=0.987709497206704
```

```
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 8.412796159973368e-05
step: 10, loss: 7.121812814148143e-05
step: 20, loss: 8.706131484359503e-05
step: 30, loss: 0.026016762480139732
step: 40, loss: 0.00018203201761934906
step: 50, loss: 6.749082240276039e-05
step: 60, loss: 0.00033063662704080343
step: 70, loss: 0.00011610580259002745
step: 80, loss: 7.820593600627035e-05
step: 90, loss: 4.158243245910853e-05
step: 100, loss: 0.0001048026024363935
step: 110, loss: 0.0008053880883380771
step: 120, loss: 4.033874211017974e-05
step: 130, loss: 3.654247848317027e-05
step: 140, loss: 0.0004939441569149494
step: 150, loss: 0.023761603981256485
step: 160, loss: 7.268826448125765e-05
step: 170, loss: 0.03758939355611801
step: 180, loss: 6.862762529635802e-05
step: 190, loss: 0.00010068721167044714
step: 200, loss: 0.0002997994306497276
step: 210, loss: 4.7609893954358995e-05
step: 220, loss: 7.22430195310153e-05
step: 230, loss: 0.00010661759006325155
/usr/local/lib/python3.7/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=0.9898534385569334, f1=0.9830124575311437,
best_f1=0.987709497206704
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 7.49828977859579e-05
step: 10, loss: 0.00012243054516147822
step: 20, loss: 5.8164321671938524e-05
step: 30, loss: 7.776536222081631e-05
step: 40, loss: 8.5350991867017e-05
step: 50, loss: 0.0001070585785782896
step: 60, loss: 0.00012211456487420946
step: 70, loss: 5.818101271870546e-05
step: 80, loss: 7.474248559447005e-05
step: 90, loss: 6.11531431786716e-05
step: 100, loss: 9.668692655395716e-05
step: 110, loss: 6.830893835285679e-05
step: 120, loss: 0.000195507105672732
```

```
step: 130, loss: 7.637035741936415e-05
step: 140, loss: 6.812602805439383e-05
step: 150, loss: 6.367081368807703e-05
step: 160, loss: 5.899298412259668e-05
step: 170, loss: 0.00010684924200177193
step: 180, loss: 7.854123396100476e-05
step: 190, loss: 4.9154481530422345e-05
step: 200, loss: 3.371465209056623e-05
step: 210, loss: 8.840776717988774e-05
step: 220, loss: 6.350656622089446e-05
step: 230, loss: 0.04616161435842514
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
epoch 12: dev_f1=0.9876543209876544, f1=0.9843400447427293,
best_f1=0.987709497206704
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00011964143050136045
step: 10, loss: 0.0007645345758646727
step: 20, loss: 9.88761312328279e-05
step: 30, loss: 8.201796299545094e-05
step: 40, loss: 0.00010119413491338491
step: 50, loss: 0.00011763598013203591
step: 60, loss: 4.7196121158776805e-05
step: 70, loss: 2.90669959213119e-05
step: 80, loss: 5.2563500503310934e-05
step: 90, loss: 7.946918049128726e-05
step: 100, loss: 2.435517126286868e-05
step: 110, loss: 0.00020708564261440188
step: 120, loss: 0.03305528312921524
step: 130, loss: 6.890839722473174e-05
step: 140, loss: 8.192471432266757e-05
step: 150, loss: 6.55801486573182e-05
step: 160, loss: 0.0001195739969261922
step: 170, loss: 0.000105190061731264
step: 180, loss: 8.088609320111573e-05
step: 190, loss: 9.434366802452132e-05
step: 200, loss: 8.468616579193622e-05
step: 210, loss: 0.000323501939419657
step: 220, loss: 6.201273936312646e-05
step: 230, loss: 9.061618766281754e-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=0.9876543209876544, f1=0.9843400447427293,
best_f1=0.987709497206704
```

```
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 3.721736720763147e-05
step: 10, loss: 0.012067052535712719
step: 20, loss: 4.755270492751151e-05
step: 30, loss: 7.953868043841794e-05
step: 40, loss: 0.01075134240090847
step: 50, loss: 4.983349208487198e-05
step: 60, loss: 5.878967931494117e-05
step: 70, loss: 4.867156167165376e-05
step: 80, loss: 5.6865334045141935e-05
step: 90, loss: 0.0001545220147818327
step: 100, loss: 4.5626442442880943e-05
step: 110, loss: 7.885901868576184e-05
step: 120, loss: 2.332703661522828e-05
step: 130, loss: 4.644016735255718e-05
step: 140, loss: 0.00011988452752120793
step: 150, loss: 7.067836850183085e-05
step: 160, loss: 0.0003830104833468795
step: 170, loss: 4.13642410421744e-05
step: 180, loss: 4.296370025258511e-05
step: 190, loss: 5.5205888202181086e-05
step: 200, loss: 3.707929135998711e-05
step: 210, loss: 9.444751776754856e-05
step: 220, loss: 4.276497566024773e-05
step: 230, loss: 5.5286698625423014e-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=0.9898305084745763, f1=0.983050847457627,
best_f1=0.987709497206704
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 2.621380190248601e-05
step: 10, loss: 2.5874544007820077e-05
step: 20, loss: 0.00024226454843301326
step: 30, loss: 3.7664467527065426e-05
step: 40, loss: 7.766697672195733e-05
step: 50, loss: 0.03977149724960327
step: 60, loss: 5.1355404139030725e-05
step: 70, loss: 5.286749728838913e-05
step: 80, loss: 6.447064515668899e-05
step: 90, loss: 6.021445369697176e-05
step: 100, loss: 5.204162880545482e-05
step: 110, loss: 3.5067834687652066e-05
step: 120, loss: 5.388964927988127e-05
```

```
step: 130, loss: 7.498609193135053e-05
step: 140, loss: 4.42821801698301e-05
step: 150, loss: 8.088039612630382e-05
step: 160, loss: 2.746899554040283e-05
step: 170, loss: 2.7402136765886098e-05
step: 180, loss: 4.384453859529458e-05
step: 190, loss: 0.0001607047743164003
step: 200, loss: 9.855616372078657e-05
step: 210, loss: 0.00011933585483347997
step: 220, loss: 4.761904710903764e-05
step: 230, loss: 4.723175516119227e-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=0.9887133182844244, f1=0.9853438556933484,
best_f1=0.987709497206704
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 [00:15, 159.44it/s]
load_f1 = 0.9921612541993281
```

```
real_f1 = 0.9899216125419933
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
4398it [00:24, 181.98it/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, 486kB/s]
Downloading: 100% 232k/232k [00:00<00:00, 3.90MB/s]
Downloading: 100% 440M/440M [00:09<00:00, 44.3MB/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: 0.613934338092804
step: 10, loss: 0.5082128643989563
step: 20, loss: 0.4452504813671112
step: 30, loss: 0.06997070461511612
step: 40, loss: 0.11014629155397415
step: 50, loss: 0.18721503019332886
step: 60, loss: 0.06895910203456879
step: 70, loss: 0.09611670672893524
step: 80, loss: 0.0457177609205246
step: 90, loss: 0.2640192210674286
step: 100, loss: 0.07924122363328934
step: 110, loss: 0.08022493869066238
step: 120, loss: 0.12507089972496033
step: 130, loss: 0.07762271910905838
step: 140, loss: 0.10442501306533813
step: 150, loss: 0.09120186418294907
```

```
step: 160, loss: 0.0383390411734581
step: 170, loss: 0.18734470009803772
step: 180, loss: 0.12905855476856232
step: 190, loss: 0.023265380412340164
step: 200, loss: 0.13216833770275116
step: 210, loss: 0.14655452966690063
step: 220, loss: 0.24144193530082703
step: 230, loss: 0.18649506568908691
step: 240, loss: 0.07451999932527542
step: 250, loss: 0.04067869856953621
step: 260, loss: 0.07409151643514633
step: 270, loss: 0.01939357817173004
step: 280, loss: 0.04048396646976471
step: 290, loss: 0.042394667863845825
step: 300, loss: 0.027263274416327477
step: 310, loss: 0.0906909853219986
step: 320, loss: 0.06287535279989243
step: 330, loss: 0.057860054075717926
step: 340, loss: 0.03125019744038582
step: 350, loss: 0.04064049571752548
step: 360, loss: 0.03070683591067791
step: 370, loss: 0.10907158255577087
step: 380, loss: 0.11332546919584274
step: 390, loss: 0.1589994579553604
step: 400, loss: 0.3631631135940552
step: 410, loss: 0.05283457413315773
step: 420, loss: 0.053227271884679794
step: 430, loss: 0.20440471172332764
step: 440, loss: 0.031467255204916
step: 450, loss: 0.007331761531531811
step: 460, loss: 0.010424582287669182
step: 470, loss: 0.15816155076026917
step: 480, loss: 0.04498807713389397
step: 490, loss: 0.05979424715042114
step: 500, loss: 0.0762535110116005
step: 510, loss: 0.08564624190330505
step: 520, loss: 0.04503433406352997
step: 530, loss: 0.0028658669907599688
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.93202062822316, f1=0.927536231884058,
best_f1=0.927536231884058
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
```

```
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.1956147700548172
step: 10, loss: 0.14010795950889587
step: 20, loss: 0.007230376824736595
step: 30, loss: 0.02317636087536812
step: 40, loss: 0.07165294140577316
step: 50, loss: 0.08454151451587677
step: 60, loss: 0.014049187302589417
step: 70, loss: 0.03037022426724434
step: 80, loss: 0.07194548100233078
step: 90, loss: 0.041661448776721954
step: 100, loss: 0.020001620054244995
step: 110, loss: 0.0987481102347374
step: 120, loss: 0.03953232988715172
step: 130, loss: 0.1718534529209137
step: 140, loss: 0.04623088985681534
step: 150, loss: 0.06081540510058403
step: 160, loss: 0.012563962489366531
step: 170, loss: 0.03166091442108154
step: 180, loss: 0.03495490178465843
step: 190, loss: 0.054093074053525925
step: 200, loss: 0.007727175951004028
step: 210, loss: 0.0951017290353775
step: 220, loss: 0.13484549522399902
step: 230, loss: 0.005941242910921574
step: 240, loss: 0.049584146589040756
step: 250, loss: 0.09885513037443161
step: 260, loss: 0.002990849083289504
step: 270, loss: 0.21787047386169434
step: 280, loss: 0.0030217573512345552
step: 290, loss: 0.03377397730946541
step: 300, loss: 0.1130533367395401
step: 310, loss: 0.01932665891945362
step: 320, loss: 0.05203250050544739
step: 330, loss: 0.0413800030946731
step: 340, loss: 0.02669086493551731
step: 350, loss: 0.0010877667227759957
step: 360, loss: 0.09356673061847687
step: 370, loss: 0.1608896255493164
step: 380, loss: 0.08814701437950134
step: 390, loss: 0.04781879484653473
step: 400, loss: 0.046957846730947495
step: 410, loss: 0.01615333743393421
step: 420, loss: 0.01300432812422514
step: 430, loss: 0.018483096733689308
step: 440, loss: 0.1576240211725235
step: 450, loss: 0.020416073501110077
```

```
step: 460, loss: 0.015114519745111465
step: 470, loss: 0.009904108941555023
step: 480, loss: 0.21572451293468475
step: 490, loss: 0.019455745816230774
step: 500, loss: 0.19059783220291138
step: 510, loss: 0.012401782907545567
step: 520, loss: 0.06397898495197296
step: 530, loss: 0.05924451723694801
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9356505401596994, f1=0.9373246024321795,
best_f1=0.9373246024321795
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.03346934914588928
step: 10, loss: 0.0541888065636158
step: 20, loss: 0.12902846932411194
step: 30, loss: 0.12529242038726807
step: 40, loss: 0.029717106372117996
step: 50, loss: 0.13159780204296112
step: 60, loss: 0.02439424768090248
step: 70, loss: 0.010078513994812965
step: 80, loss: 0.002970106201246381
step: 90, loss: 0.005487777292728424
step: 100, loss: 0.08866236358880997
step: 110, loss: 0.0010315888794139028
step: 120, loss: 0.004780990071594715
step: 130, loss: 0.00698839919641614
step: 140, loss: 0.025573445484042168
step: 150, loss: 0.014395667240023613
step: 160, loss: 0.0064708092249929905
step: 170, loss: 0.060806598514318466
step: 180, loss: 0.007326286286115646
step: 190, loss: 0.013348439708352089
step: 200, loss: 0.05384363234043121
step: 210, loss: 0.039264075458049774
step: 220, loss: 0.11485884338617325
step: 230, loss: 0.06627652049064636
step: 240, loss: 0.0031007127836346626
step: 250, loss: 0.025571776553988457
step: 260, loss: 0.008420799858868122
step: 270, loss: 0.02386505715548992
```

```
step: 280, loss: 0.11756488680839539
step: 290, loss: 0.0018904379103332758
step: 300, loss: 0.008340919390320778
step: 310, loss: 0.05358602851629257
step: 320, loss: 0.032012827694416046
step: 330, loss: 0.0023750776890665293
step: 340, loss: 0.014050230383872986
step: 350, loss: 0.013914233073592186
step: 360, loss: 0.05564792454242706
step: 370, loss: 0.010071555152535439
step: 380, loss: 0.013834935612976551
step: 390, loss: 0.03584376350045204
step: 400, loss: 0.012088505551218987
step: 410, loss: 0.006354578770697117
step: 420, loss: 0.14969882369041443
step: 430, loss: 0.011920103803277016
step: 440, loss: 0.002132501918822527
step: 450, loss: 0.08592624217271805
step: 460, loss: 0.006167762912809849
step: 470, loss: 0.005790104623883963
step: 480, loss: 0.0033562059979885817
step: 490, loss: 0.008619189262390137
step: 500, loss: 0.020189370959997177
step: 510, loss: 0.008360831066966057
step: 520, loss: 0.12516364455223083
step: 530, loss: 0.04122161120176315
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9375866851595007, f1=0.9364858599907279,
best_f1=0.9364858599907279
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.010036027990281582
step: 10, loss: 0.0065558189526200294
step: 20, loss: 0.004528503865003586
step: 30, loss: 0.0010049828561022878
step: 40, loss: 0.004993733484297991
step: 50, loss: 0.0005125848692841828
step: 60, loss: 0.0007800692692399025
step: 70, loss: 0.00853730272501707
step: 80, loss: 0.0029322232585400343
step: 90, loss: 0.03672856092453003
```

```
step: 100, loss: 0.06479310989379883
step: 110, loss: 0.01806642860174179
step: 120, loss: 0.00018564745550975204
step: 130, loss: 0.0006582285859622061
step: 140, loss: 0.0010411322582513094
step: 150, loss: 0.0011261026374995708
step: 160, loss: 0.04512978345155716
step: 170, loss: 0.043578751385211945
step: 180, loss: 0.0062391310930252075
step: 190, loss: 0.0326557420194149
step: 200, loss: 0.061354849487543106
step: 210, loss: 0.024330290034413338
step: 220, loss: 0.00154464493971318
step: 230, loss: 0.02029651403427124
step: 240, loss: 0.0073782941326498985
step: 250, loss: 0.010680004954338074
step: 260, loss: 0.008982696570456028
step: 270, loss: 0.02917974255979061
step: 280, loss: 0.015648556873202324
step: 290, loss: 0.014416366815567017
step: 300, loss: 0.00019843342306558043
step: 310, loss: 0.07886991649866104
step: 320, loss: 0.019532978534698486
step: 330, loss: 0.010830240324139595
step: 340, loss: 0.07274630665779114
step: 350, loss: 0.03630658611655235
step: 360, loss: 0.039976220577955246
step: 370, loss: 0.007580047473311424
step: 380, loss: 0.000982298399321735
step: 390, loss: 0.01696426421403885
step: 400, loss: 0.05888251215219498
step: 410, loss: 0.003723912639543414
Gradient overflow.  Skipping step, loss scaler 0 reducing loss scale to 32768.0
step: 420, loss: 0.048240967094898224
step: 430, loss: 0.10430088639259338
step: 440, loss: 0.00722967553883791
step: 450, loss: 0.0019129014108330011
step: 460, loss: 0.0020404160022735596
step: 470, loss: 0.007343672215938568
step: 480, loss: 0.0038516575004905462
step: 490, loss: 0.0010678413091227412
step: 500, loss: 0.010023959912359715
step: 510, loss: 0.003828303422778845
step: 520, loss: 0.04377691075205803
step: 530, loss: 0.09692911058664322
/usr/local/lib/python3.7/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=0.9317972350230416, f1=0.9349220898258479,
best_f1=0.9364858599907279
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.012839573435485363
step: 10, loss: 0.0493207685649395
step: 20, loss: 0.04077386111021042
step: 30, loss: 0.0013693490764126182
step: 40, loss: 0.053036294877529144
step: 50, loss: 0.0037944503128528595
step: 60, loss: 0.0046800789423286915
step: 70, loss: 0.02637835405766964
step: 80, loss: 0.00043273603660054505
step: 90, loss: 0.0004761040036100894
step: 100, loss: 0.007375899702310562
step: 110, loss: 0.00038152484921738505
step: 120, loss: 0.0004714365932159126
Gradient overflow.  Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 130, loss: 0.0005595898837782443
step: 140, loss: 0.0004121988022234291
step: 150, loss: 0.002262201625853777
step: 160, loss: 0.00039949960773810744
step: 170, loss: 0.044757865369319916
step: 180, loss: 0.001883669407106936
step: 190, loss: 0.0034446311183273792
step: 200, loss: 0.003538762219250202
step: 210, loss: 0.007229969836771488
step: 220, loss: 0.0025058311875909567
step: 230, loss: 0.004095837473869324
step: 240, loss: 0.0023783687502145767
step: 250, loss: 0.0019853352569043636
step: 260, loss: 0.0010527988197281957
step: 270, loss: 0.00016674063226673752
step: 280, loss: 0.0038085924461483955
step: 290, loss: 0.12465646117925644
step: 300, loss: 0.003673934144899249
step: 310, loss: 0.0011199722066521645
step: 320, loss: 0.0745108351111412
step: 330, loss: 0.02580823190510273
step: 340, loss: 0.002064164262264967
step: 350, loss: 0.006912263575941324
step: 360, loss: 0.007277794647961855
step: 370, loss: 0.00909241009503603
step: 380, loss: 0.0013737418921664357
step: 390, loss: 0.0002650823735166341
step: 400, loss: 0.0011744149960577488
step: 410, loss: 0.0002386741543887183
```

```
step: 420, loss: 0.0013529341667890549
step: 430, loss: 0.00039975406252779067
step: 440, loss: 0.003147589275613427
step: 450, loss: 0.004347576294094324
step: 460, loss: 0.10527914762496948
step: 470, loss: 0.02410237118601799
step: 480, loss: 0.007835080847144127
step: 490, loss: 0.0008592365775257349
step: 500, loss: 0.0026379134505987167
step: 510, loss: 0.009951031766831875
step: 520, loss: 0.07097494602203369
step: 530, loss: 0.003615850117057562
/usr/local/lib/python3.7/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=0.9366391184573002, f1=0.9402099497946144,
best_f1=0.9364858599907279
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.010396366007626057
step: 10, loss: 0.002793720457702875
step: 20, loss: 0.11026394367218018
step: 30, loss: 0.0008188716601580381
step: 40, loss: 0.0017879604129120708
step: 50, loss: 0.014004328288137913
step: 60, loss: 0.000283078319625929
step: 70, loss: 0.0004212549247313291
step: 80, loss: 0.007956408895552158
step: 90, loss: 0.000473715306725353
step: 100, loss: 0.0013079320779070258
step: 110, loss: 0.01610811986029148
step: 120, loss: 0.0032987461891025305
step: 130, loss: 0.00027509153005667033
step: 140, loss: 0.003113705664873123
step: 150, loss: 0.0008387485868297517
step: 160, loss: 0.0003085996431764215
step: 170, loss: 0.001013087690807879
step: 180, loss: 0.0013439857866615057
step: 190, loss: 0.00033698894549161196
step: 200, loss: 0.00024111953098326921
step: 210, loss: 0.0011466031428426504
step: 220, loss: 0.046398088335990906
step: 230, loss: 0.027985168620944023
step: 240, loss: 0.10077822208404541
step: 250, loss: 0.0009960670722648501
step: 260, loss: 0.01877768337726593
step: 270, loss: 0.11677844077348709
```

```
step: 280, loss: 0.004582786001265049
step: 290, loss: 0.0004875946033280343
step: 300, loss: 0.004140487406402826
step: 310, loss: 0.0006182852666825056
step: 320, loss: 0.0004775654524564743
step: 330, loss: 0.005474055651575327
step: 340, loss: 0.050843559205532074
step: 350, loss: 0.011586022563278675
step: 360, loss: 0.018565841019153595
step: 370, loss: 0.0016359372530132532
step: 380, loss: 0.0005920235416851938
step: 390, loss: 0.029160238802433014
step: 400, loss: 0.00029319405439309776
step: 410, loss: 0.0004007990937680006
step: 420, loss: 0.006374744698405266
step: 430, loss: 0.0009057938004843891
step: 440, loss: 0.00026845268439501524
step: 450, loss: 0.000431251828558743
step: 460, loss: 0.00016507912368979305
step: 470, loss: 0.0003501226019579917
step: 480, loss: 0.0020065505523234606
step: 490, loss: 0.011343443766236305
step: 500, loss: 0.0009669616119936109
step: 510, loss: 0.03878592327237129
step: 520, loss: 0.0011211598757654428
step: 530, loss: 0.05348493158817291
/usr/local/lib/python3.7/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=0.9266697804764129, f1=0.9357541899441341,
best_f1=0.9364858599907279
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0008534913649782538
step: 10, loss: 0.003850545035675168
step: 20, loss: 0.0006376230739988387
step: 30, loss: 0.0015533401165157557
step: 40, loss: 0.00020776537712663412
step: 50, loss: 0.006746172439306974
step: 60, loss: 0.00433209678158164
step: 70, loss: 0.0002775152097456157
step: 80, loss: 0.0053441585041582584
step: 90, loss: 0.00015637799515388906
step: 100, loss: 0.00013909919653087854
step: 110, loss: 0.0042875539511442184
step: 120, loss: 0.0040756696835160255
step: 130, loss: 0.00037950687692500065
```

```
step: 140, loss: 0.00017285517242271453
step: 150, loss: 0.0006774748908355832
step: 160, loss: 0.0009464994654990733
step: 170, loss: 0.0010614024940878153
step: 180, loss: 0.0022003338672220707
step: 190, loss: 0.0028458419255912304
step: 200, loss: 0.00015491642989218235
step: 210, loss: 0.0007967435522004962
step: 220, loss: 0.00021377517259679735
step: 230, loss: 0.02284861169755459
step: 240, loss: 0.0002659448073245585
step: 250, loss: 0.00031806380138732493
step: 260, loss: 6.0087440942879766e-05
step: 270, loss: 9.366855374537408e-05
step: 280, loss: 0.0003123840142507106
step: 290, loss: 9.294680785387754e-05
step: 300, loss: 0.0024208456743508577
step: 310, loss: 0.07721994817256927
step: 320, loss: 0.001036832807585597
step: 330, loss: 0.004226783290505409
step: 340, loss: 0.013594734482467175
step: 350, loss: 0.000301238673273474
step: 360, loss: 0.03400915488600731
step: 370, loss: 0.0009042947785928845
step: 380, loss: 0.0020022511016577482
step: 390, loss: 9.418661647941917e-05
step: 400, loss: 0.008458890952169895
step: 410, loss: 0.0018016245448961854
step: 420, loss: 0.0003077961737290025
step: 430, loss: 5.532994327950291e-05
step: 440, loss: 0.00014639945584349334
step: 450, loss: 0.00012910277291666716
step: 460, loss: 0.0011904799612239003
step: 470, loss: 0.004262649919837713
step: 480, loss: 0.2016754150390625
step: 490, loss: 0.00019067231914959848
step: 500, loss: 0.00035242928424850106
step: 510, loss: 0.005133882164955139
step: 520, loss: 0.01761259324848652
step: 530, loss: 0.0053643835708498955
/usr/local/lib/python3.7/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=0.9347623485554521, f1=0.9377323420074348,
best_f1=0.9364858599907279
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
```

```
step: 0, loss: 0.00018206539971288294
step: 10, loss: 0.012868838384747505
step: 20, loss: 0.00015348580200225115
step: 30, loss: 0.008971890434622765
step: 40, loss: 0.0005534648080356419
step: 50, loss: 0.0035205124877393246
step: 60, loss: 0.0020258035510778427
step: 70, loss: 0.00011267873924225569
step: 80, loss: 0.01128922589123249
step: 90, loss: 0.0005116018000990152
step: 100, loss: 4.889712727162987e-05
step: 110, loss: 0.0003761308907996863
step: 120, loss: 0.054790809750556946
step: 130, loss: 0.00012732576578855515
step: 140, loss: 0.013537583872675896
step: 150, loss: 0.000723730365280062
step: 160, loss: 0.10025345534086227
step: 170, loss: 0.19775685667991638
step: 180, loss: 0.0002681297482922673
step: 190, loss: 0.0034166653640568256
step: 200, loss: 0.00210868163779378
step: 210, loss: 0.0031014918349683285
step: 220, loss: 0.002344634849578142
step: 230, loss: 0.00012638873886317015
step: 240, loss: 0.00046304488205350935
step: 250, loss: 0.0007252034265547991
step: 260, loss: 0.00011047320003854111
step: 270, loss: 0.008527781814336777
step: 280, loss: 0.03619414567947388
step: 290, loss: 0.00013441122428048402
step: 300, loss: 0.0010376116260886192
step: 310, loss: 0.00019447928934823722
step: 320, loss: 0.00043510159594006836
step: 330, loss: 6.423712329706177e-05
step: 340, loss: 8.58246658115308e-05
step: 350, loss: 0.00016133183089550585
step: 360, loss: 0.0031917900778353214
step: 370, loss: 0.0001997996005229652
step: 380, loss: 0.0005214763223193586
step: 390, loss: 0.12901125848293304
step: 400, loss: 0.0008633381803520024
step: 410, loss: 0.00013105219113640487
step: 420, loss: 0.002133836504071951
step: 430, loss: 0.05257078632712364
step: 440, loss: 0.0002277519815834239
step: 450, loss: 0.00017868027498479933
step: 460, loss: 0.000316916179144755
step: 470, loss: 8.89099610503763e-05
```

```
step: 480, loss: 8.756417082622647e-05
step: 490, loss: 0.0011698753805831075
step: 500, loss: 0.00014033864135853946
step: 510, loss: 0.00015008196351118386
step: 520, loss: 0.0017946739681065083
step: 530, loss: 0.007545233704149723
/usr/local/lib/python3.7/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=0.9348946135831382, f1=0.935813953488372,
best_f1=0.9364858599907279
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.009197330102324486
step: 10, loss: 7.613015623064712e-05
step: 20, loss: 0.0003250973823014647
step: 30, loss: 0.00093937071505934
step: 40, loss: 0.0005452515324577689
step: 50, loss: 2.780376235023141e-05
step: 60, loss: 5.155291364644654e-05
step: 70, loss: 0.003646361641585827
step: 80, loss: 0.0005817731725983322
step: 90, loss: 8.040168177103624e-05
step: 100, loss: 7.257080869749188e-05
step: 110, loss: 5.589974898612127e-05
step: 120, loss: 0.0003953043487854302
step: 130, loss: 0.00026634699931576401
step: 140, loss: 8.906058792490512e-05
step: 150, loss: 5.550196874537505e-05
step: 160, loss: 0.001206117682158947
step: 170, loss: 0.0036416472867131233
step: 180, loss: 6.126991502242163e-05
step: 190, loss: 0.00047457675100304186
step: 200, loss: 5.084211807115935e-05
step: 210, loss: 0.00018732687749434263
step: 220, loss: 8.670843089930713e-05
step: 230, loss: 6.12403528066352e-05
step: 240, loss: 8.408571011386812e-05
step: 250, loss: 5.9709731431212276e-05
step: 260, loss: 0.004378249403089285
step: 270, loss: 5.7752298744162545e-05
step: 280, loss: 0.007220068946480751
step: 290, loss: 0.05945190042257309
step: 300, loss: 0.00013673539797309786
step: 310, loss: 0.02529202401638031
step: 320, loss: 0.00011448845179984346
step: 330, loss: 0.0001895786845125258
```

```
step: 340, loss: 0.00010102095257025212
step: 350, loss: 0.0021284939721226692
step: 360, loss: 4.936772893415764e-05
step: 370, loss: 0.0022157924249768257
step: 380, loss: 0.0008394194883294404
step: 390, loss: 0.05942973122000694
step: 400, loss: 0.022158833220601082
step: 410, loss: 0.009444921277463436
step: 420, loss: 5.38441636308562e-05
step: 430, loss: 8.380404324270785e-05
step: 440, loss: 0.0033312987070530653
step: 450, loss: 0.00041719546425156295
step: 460, loss: 0.00016389560187235475
step: 470, loss: 0.003516304073855281
step: 480, loss: 0.0002570610085967928
step: 490, loss: 0.0007272197399288416
step: 500, loss: 0.0023559383116662502
step: 510, loss: 0.00030071751098148525
step: 520, loss: 0.0014173304662108421
step: 530, loss: 0.00019261091074440628
/usr/local/lib/python3.7/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=0.9266445811642213, f1=0.9337717238139972,
best_f1=0.9364858599907279
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0003528505621943623
step: 10, loss: 6.845030293334275e-05
step: 20, loss: 0.0002752391155809164
step: 30, loss: 8.178374264389277e-05
step: 40, loss: 0.00020149818737991154
step: 50, loss: 0.0009968988597393036
step: 60, loss: 0.00017280399333685637
step: 70, loss: 9.30982205318287e-05
step: 80, loss: 9.974028216674924e-05
step: 90, loss: 0.0005101111019030213
step: 100, loss: 0.00012255701585672796
step: 110, loss: 5.3365325584309176e-05
step: 120, loss: 0.00017148282495327294
step: 130, loss: 0.00019300385611131787
step: 140, loss: 0.01163062360137701
step: 150, loss: 0.0014187500346451998
step: 160, loss: 0.0002591057273093611
step: 170, loss: 0.00016914808657020233
step: 180, loss: 0.00014271019608713686
step: 190, loss: 8.936748781707138e-05
```

```
step: 200, loss: 0.0004736543050967157
step: 210, loss: 0.003344177268445492
step: 220, loss: 0.18883664906024933
step: 230, loss: 8.705165237188339e-05
step: 240, loss: 0.0001432945573469624
step: 250, loss: 0.00011299422476440668
step: 260, loss: 0.0011543100699782372
step: 270, loss: 0.000959704746492207
step: 280, loss: 0.00014640422887168825
step: 290, loss: 0.00010216025111731142
step: 300, loss: 8.180550503311679e-05
step: 310, loss: 7.582149555673823e-05
step: 320, loss: 0.0006477255374193192
step: 330, loss: 5.49009746464435e-05
step: 340, loss: 0.00021476979600265622
step: 350, loss: 0.00531064672395587
step: 360, loss: 0.00023330387193709612
step: 370, loss: 0.0068403747864067554
step: 380, loss: 0.00023970133042894304
step: 390, loss: 0.00014352967264130712
step: 400, loss: 0.0016609993763267994
step: 410, loss: 0.029473386704921722
step: 420, loss: 0.0001349167723674327
step: 430, loss: 6.647852569585666e-05
step: 440, loss: 0.019313538447022438
step: 450, loss: 0.00022129577700980008
step: 460, loss: 0.00019608855654951185
step: 470, loss: 0.0001392869744449854
step: 480, loss: 7.348298095166683e-05
step: 490, loss: 0.00011934716167161241
step: 500, loss: 0.004385172389447689
step: 510, loss: 5.7223904150305316e-05
step: 520, loss: 0.00019064053776673973
step: 530, loss: 0.0016413796693086624
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9398601398601399, f1=0.9453703703703704,
best_f1=0.9453703703703704
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.026088280603289604
step: 10, loss: 0.00024582637706771493
```

```
step: 20, loss: 0.00017898080113809556
step: 30, loss: 0.00013054636656306684
step: 40, loss: 0.0010877910535782576
step: 50, loss: 0.001215091673657298
step: 60, loss: 0.06667208671569824
step: 70, loss: 0.0007215086370706558
step: 80, loss: 0.00011206440103705972
step: 90, loss: 0.00027385420980863273
step: 100, loss: 0.0012538303853943944
step: 110, loss: 0.00023444740509148687
step: 120, loss: 4.1340139432577416e-05
step: 130, loss: 0.00011579989222809672
step: 140, loss: 0.03395124524831772
step: 150, loss: 8.24354647193104e-05
step: 160, loss: 0.009449001401662827
step: 170, loss: 0.00019655746291391551
step: 180, loss: 0.00013310502981767058
step: 190, loss: 0.009019912220537663
step: 200, loss: 0.010314377024769783
step: 210, loss: 0.0003691008605528623
step: 220, loss: 0.0029862921219319105
step: 230, loss: 0.0004108032153453678
step: 240, loss: 0.0005727591342292726
step: 250, loss: 4.1389226680621505e-05
step: 260, loss: 7.832699338905513e-05
step: 270, loss: 0.00027681211940944195
step: 280, loss: 0.0001315215340582654
step: 290, loss: 9.896483243210241e-05
step: 300, loss: 0.00016050244448706508
step: 310, loss: 0.0007531174924224615
step: 320, loss: 8.646656351629645e-05
step: 330, loss: 0.00031528074759989977
step: 340, loss: 0.0003487810608930886
step: 350, loss: 0.0014570645289495587
step: 360, loss: 0.0007904174271970987
step: 370, loss: 0.00020526461594272405
step: 380, loss: 6.328208110062405e-05
step: 390, loss: 4.817174340132624e-05
step: 400, loss: 2.7037331165047362e-05
step: 410, loss: 5.737476021749899e-05
step: 420, loss: 0.00011016060307156295
step: 430, loss: 3.711373210535385e-05
step: 440, loss: 5.1300463383086026e-05
step: 450, loss: 0.000689059728756547
step: 460, loss: 4.808003723155707e-05
step: 470, loss: 5.652193067362532e-05
step: 480, loss: 0.0012906717602163553
step: 490, loss: 9.744527051225305e-05
```

```
step: 500, loss: 0.0008215019479393959
step: 510, loss: 5.413613689597696e-05
step: 520, loss: 0.00019548155250959098
step: 530, loss: 0.00011572824587346986
/usr/local/lib/python3.7/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=0.9332079021636878, f1=0.9361305361305362,
best_f1=0.9453703703703704
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 5.28559903614223e-05
step: 10, loss: 2.6210103897028603e-05
step: 20, loss: 8.33989106467925e-05
step: 30, loss: 0.0007514553726650774
step: 40, loss: 0.00016407007933594286
step: 50, loss: 0.05101954564452171
step: 60, loss: 0.002887825248762965
step: 70, loss: 9.197460894938558e-05
step: 80, loss: 3.526731234160252e-05
step: 90, loss: 0.0005101598799228668
step: 100, loss: 0.0005079185939393938
step: 110, loss: 0.00011025385174434632
step: 120, loss: 2.005648821068462e-05
step: 130, loss: 0.00048135468387044966
step: 140, loss: 9.852150833467022e-05
step: 150, loss: 4.383974737720564e-05
step: 160, loss: 2.6947887818096206e-05
step: 170, loss: 3.348099562572315e-05
step: 180, loss: 3.004677637363784e-05
step: 190, loss: 9.155164298135787e-05
step: 200, loss: 0.00037530012195929885
step: 210, loss: 3.529547757352702e-05
step: 220, loss: 0.00015286101552192122
step: 230, loss: 2.1378888050094247e-05
step: 240, loss: 0.0001695889513939619
step: 250, loss: 7.36885194783099e-05
step: 260, loss: 4.9978072638623416e-05
step: 270, loss: 8.219020674005151e-05
step: 280, loss: 0.003779578721150756
step: 290, loss: 5.2900955779477954e-05
step: 300, loss: 0.0004583392583299428
step: 310, loss: 0.00018898621783591807
step: 320, loss: 1.8920396541943774e-05
step: 330, loss: 4.152918700128794e-05
step: 340, loss: 2.5059380277525634e-05
step: 350, loss: 0.00020433194003999233
```

```
step: 360, loss: 0.0006684482796117663
step: 370, loss: 9.57736701820977e-05
step: 380, loss: 0.0001544225960969925
step: 390, loss: 0.016340386122465134
step: 400, loss: 0.00017343703075312078
step: 410, loss: 5.351311483536847e-05
step: 420, loss: 0.007027212530374527
step: 430, loss: 0.023030713200569153
step: 440, loss: 2.812850652844645e-05
step: 450, loss: 8.541863644495606e-05
step: 460, loss: 0.00012725281703751534
step: 470, loss: 6.020966247888282e-05
step: 480, loss: 8.627366332802922e-05
step: 490, loss: 0.0003156780148856342
step: 500, loss: 3.995738006778993e-05
step: 510, loss: 0.0064966860227286816
step: 520, loss: 0.0002938823599833995
step: 530, loss: 4.2006922740256414e-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=0.9354078264969354, f1=0.9400749063670412,
best_f1=0.9453703703703704
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0069867572747170925
step: 10, loss: 2.1978550648782402e-05
step: 20, loss: 2.2591719243791886e-05
step: 30, loss: 0.001010425272397697
step: 40, loss: 3.852403096971102e-05
step: 50, loss: 0.00010705285967560485
step: 60, loss: 2.8865848435088992e-05
step: 70, loss: 3.230285437894054e-05
step: 80, loss: 5.4828553402330726e-05
step: 90, loss: 0.025303808972239494
step: 100, loss: 0.00020779561600647867
step: 110, loss: 2.9943350455141626e-05
step: 120, loss: 2.9748742235824466e-05
step: 130, loss: 3.252423266530968e-05
step: 140, loss: 3.142815694445744e-05
step: 150, loss: 0.0003869189531542361
step: 160, loss: 2.9255428671604022e-05
step: 170, loss: 0.00015462585724890232
step: 180, loss: 6.0776052123401314e-05
step: 190, loss: 1.872288885351736e-05
step: 200, loss: 3.8271617086138576e-05
step: 210, loss: 3.039821058337111e-05
```

```
step: 220, loss: 3.004091195180081e-05
step: 230, loss: 1.557892574055586e-05
step: 240, loss: 0.00015038893616292626
step: 250, loss: 3.379500412847847e-05
step: 260, loss: 9.962396143237129e-05
step: 270, loss: 1.588801205798518e-05
step: 280, loss: 2.9610702767968178e-05
step: 290, loss: 3.340472903801128e-05
step: 300, loss: 1.9542545487638563e-05
step: 310, loss: 0.001608490594662726
step: 320, loss: 0.0021312707103788853
step: 330, loss: 0.001338658039458096
step: 340, loss: 2.405669874860905e-05
step: 350, loss: 2.831097663147375e-05
Gradient overflow.  Skipping step, loss scaler 0 reducing loss scale to 32768.0
step: 360, loss: 5.537905963137746e-05
step: 370, loss: 0.00020022249373141676
step: 380, loss: 0.00022561097284778953
step: 390, loss: 0.00019773848180193454
step: 400, loss: 2.7643805879051797e-05
step: 410, loss: 0.0012808385072275996
step: 420, loss: 2.229130404884927e-05
step: 430, loss: 0.023649537935853004
step: 440, loss: 2.0625988327083178e-05
step: 450, loss: 3.288373045506887e-05
step: 460, loss: 0.001330300117842853
step: 470, loss: 2.371645314269699e-05
step: 480, loss: 5.4468244343297556e-05
Gradient overflow.  Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 490, loss: 0.000681342207826674
step: 500, loss: 0.00011310596892144531
step: 510, loss: 0.015189441852271557
step: 520, loss: 0.001005065510980785
step: 530, loss: 4.191395782981999e-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=0.9310670443814919, f1=0.9400749063670412,
best_f1=0.9453703703703704
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00011343284859322011
step: 10, loss: 1.7962176571018063e-05
step: 20, loss: 1.6081754438346252e-05
step: 30, loss: 1.7944372302736156e-05
step: 40, loss: 6.974130519665778e-05
step: 50, loss: 0.005358532071113586
```

```
step: 60, loss: 3.874915637425147e-05
step: 70, loss: 1.902065923786722e-05
step: 80, loss: 0.00011833675671368837
step: 90, loss: 2.351675539102871e-05
step: 100, loss: 0.0003983063797932118
step: 110, loss: 5.0462960643926635e-05
step: 120, loss: 2.591531847428996e-05
step: 130, loss: 0.012313400395214558
step: 140, loss: 0.0009043316822499037
step: 150, loss: 1.5213552615023218e-05
step: 160, loss: 0.0006573151331394911
step: 170, loss: 0.003227067878469825
step: 180, loss: 3.45355729223229e-05
step: 190, loss: 0.0012694899924099445
step: 200, loss: 8.133373194141313e-05
step: 210, loss: 2.1423413272714242e-05
step: 220, loss: 2.1520292648347095e-05
step: 230, loss: 9.740392124513164e-05
step: 240, loss: 1.983288530027494e-05
step: 250, loss: 3.687810749397613e-05
step: 260, loss: 0.0006777202943339944
step: 270, loss: 1.5031213479232974e-05
step: 280, loss: 1.4092564924794715e-05
step: 290, loss: 0.020047668367624283
step: 300, loss: 2.9670536605408415e-05
step: 310, loss: 0.018427779898047447
step: 320, loss: 0.0002042718551820144
step: 330, loss: 2.980045428557787e-05
step: 340, loss: 0.0001878845359897241
step: 350, loss: 7.30787796783261e-05
step: 360, loss: 0.00032677041599527
step: 370, loss: 1.9516277461661957e-05
step: 380, loss: 2.5420460588065907e-05
step: 390, loss: 0.02098827250301838
step: 400, loss: 0.00021952114184387028
step: 410, loss: 2.3613338271388784e-05
step: 420, loss: 0.00780859449878335
step: 430, loss: 0.13558706641197205
step: 440, loss: 0.00017801232752390206
step: 450, loss: 5.048049570177682e-05
step: 460, loss: 9.958689770428464e-05
step: 470, loss: 0.00016270442574750632
step: 480, loss: 3.077154906350188e-05
step: 490, loss: 3.969116005464457e-05
step: 500, loss: 1.794352829165291e-05
step: 510, loss: 0.00684317946434021
step: 520, loss: 0.0002774000167846 6797
step: 530, loss: 0.0009888290660455823
```

```
/usr/local/lib/python3.7/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=0.9324515824279641, f1=0.9360902255639098,
best_f1=0.9453703703703704
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 4.509543578024022e-05
step: 10, loss: 3.971498517785221e-05
step: 20, loss: 1.9147395505569875e-05
step: 30, loss: 8.107349276542664e-05
step: 40, loss: 0.07351942360401154
step: 50, loss: 0.00013193758786655962
step: 60, loss: 9.879401659418363e-06
step: 70, loss: 0.00040566551161548406
step: 80, loss: 2.3393578885588795e-05
step: 90, loss: 2.153152490791399e-05
step: 100, loss: 2.112494985340163e-05
step: 110, loss: 3.181569627486169e-05
step: 120, loss: 6.512519030366093e-05
step: 130, loss: 0.10380297154188156
step: 140, loss: 3.409739292692393e-05
step: 150, loss: 4.901258216705173e-05
step: 160, loss: 5.1674760470632464e-05
step: 170, loss: 2.060010410787072e-05
step: 180, loss: 1.038228401739616e-05
step: 190, loss: 4.782074029208161e-05
step: 200, loss: 4.593242556438781e-05
step: 210, loss: 0.0037274484056979418
step: 220, loss: 3.0492637961287983e-05
step: 230, loss: 0.0006253741448745131
step: 240, loss: 9.913330723065883e-05
step: 250, loss: 0.0005415150662884116
step: 260, loss: 4.08561245421879e-05
step: 270, loss: 2.261546796944458e-05
step: 280, loss: 2.044003122136928e-05
step: 290, loss: 1.4423962966247927e-05
step: 300, loss: 3.7067777157062665e-05
step: 310, loss: 3.9569429645780474e-05
step: 320, loss: 4.177675145911053e-05
step: 330, loss: 5.249411697150208e-05
step: 340, loss: 2.4158796804840676e-05
step: 350, loss: 1.4047667718841694e-05
step: 360, loss: 0.0020915570203214884
step: 370, loss: 1.5880519640631974e-05
step: 380, loss: 0.0003338168316986412
step: 390, loss: 4.997497671865858e-05
```

```
step: 400, loss: 2.629195478220936e-05
step: 410, loss: 2.1203306459938176e-05
step: 420, loss: 4.439058102434501e-05
step: 430, loss: 0.0003072820545639843
step: 440, loss: 5.726010931539349e-05
step: 450, loss: 1.2878126653959043e-05
step: 460, loss: 5.086036981083453e-05
step: 470, loss: 0.00315316254273057
step: 480, loss: 1.0736186595750041e-05
step: 490, loss: 1.3179879715607967e-05
step: 500, loss: 3.19811915687751e-05
step: 510, loss: 5.5342839914374053e-05
step: 520, loss: 1.949744728335645e-05
step: 530, loss: 0.0001341097813565284
/usr/local/lib/python3.7/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=0.935285781766651, f1=0.9407894736842105,
best_f1=0.9453703703703704
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 [00:20, 280.33it/s]
load_f1 = 0.9376181474480151
real_f1 = 0.9346497414198403
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, 277.52it/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: 0.5434144139289856
step: 10, loss: 0.3823399543762207
step: 20, loss: 0.37477490305900574
step: 30, loss: 0.32258114218711853
step: 40, loss: 0.17696405947208405
step: 50, loss: 0.44034701585769653
step: 60, loss: 0.18311454355716705
step: 70, loss: 0.12624257802963257
step: 80, loss: 0.21569672226905823
step: 90, loss: 0.363359659910202
step: 100, loss: 0.394257128238678
step: 110, loss: 0.2372412234544754
step: 120, loss: 0.19035498797893524
step: 130, loss: 0.19953617453575134
step: 140, loss: 0.21994824707508087
step: 150, loss: 0.19273599982261658
```

```
step: 160, loss: 0.22796499729156494
step: 170, loss: 0.3086649775505066
step: 180, loss: 0.08711553364992142
step: 190, loss: 0.13265776634216309
step: 200, loss: 0.2739883065223694
step: 210, loss: 0.1839381605386734
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.6404293381037568, f1=0.6323268206039075,
best_f1=0.6323268206039075
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.1075240895152092
step: 10, loss: 0.16614452004432678
step: 20, loss: 0.21451720595359802
step: 30, loss: 0.2866430878639221
step: 40, loss: 0.2459515780210495
step: 50, loss: 0.14319941401481628
step: 60, loss: 0.4001561403274536
step: 70, loss: 0.11650519073009491
step: 80, loss: 0.1990671157836914
step: 90, loss: 0.09421484172344208
step: 100, loss: 0.008277940563857555
step: 110, loss: 0.07497961074113846
step: 120, loss: 0.18063023686408997
step: 130, loss: 0.034392595291137695
step: 140, loss: 0.18351875245571136
step: 150, loss: 0.23681522905826569
step: 160, loss: 0.13535195589065552
step: 170, loss: 0.1488286852836609
step: 180, loss: 0.15940889716148376
step: 190, loss: 0.182070791721344
step: 200, loss: 0.04856667295098305
step: 210, loss: 0.12433665990829468
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.6830265848670757, f1=0.6708333333333333,
```

```
best_f1=0.6708333333333333
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.05691620334982872
step: 10, loss: 0.1504661589860916
step: 20, loss: 0.16778451204299927
step: 30, loss: 0.25222429633140564
step: 40, loss: 0.11161648482084274
step: 50, loss: 0.06341027468442917
step: 60, loss: 0.24596813321113586
step: 70, loss: 0.07870879024267197
step: 80, loss: 0.21909897029399872
step: 90, loss: 0.06598716229200363
step: 100, loss: 0.2469601333141327
step: 110, loss: 0.23451107740402222
step: 120, loss: 0.09911496192216873
step: 130, loss: 0.1133035197854042
step: 140, loss: 0.1439017653465271
step: 150, loss: 0.2673822343349457
step: 160, loss: 0.03618275374174118
step: 170, loss: 0.15383918583393097
step: 180, loss: 0.13727037608623505
step: 190, loss: 0.1931876838207245
step: 200, loss: 0.0668308287858963
step: 210, loss: 0.11872367560863495
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.7090558766859344, f1=0.6602687140115163,
best_f1=0.6602687140115163
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.13032038509845734
step: 10, loss: 0.08147811144590378
step: 20, loss: 0.13474252820014954
step: 30, loss: 0.15347692370414734
step: 40, loss: 0.020334392786026
step: 50, loss: 0.17192529141902924
step: 60, loss: 0.10283730179071426
step: 70, loss: 0.3344650864601135
step: 80, loss: 0.18116846680641174
step: 90, loss: 0.027297867462038994
```

```
step: 100, loss: 0.34588733315467834
step: 110, loss: 0.2002459168434143
step: 120, loss: 0.1327800154685974
step: 130, loss: 0.4090738296508789
step: 140, loss: 0.12034253031015396
step: 150, loss: 0.048721540719270706
step: 160, loss: 0.038944046944379807
step: 170, loss: 0.09557562321424484
step: 180, loss: 0.3871647119522095
step: 190, loss: 0.06750144064426422
step: 200, loss: 0.25029847025871277
step: 210, loss: 0.15257437527179718
/usr/local/lib/python3.7/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=0.6725978647686832, f1=0.6323809523809524,
best_f1=0.6602687140115163
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.29246899485588074
step: 10, loss: 0.0663321241736412
step: 20, loss: 0.2531580626964569
step: 30, loss: 0.046676263213157654
step: 40, loss: 0.11544736474752426
step: 50, loss: 0.0685892403125763
step: 60, loss: 0.0877435952425003
step: 70, loss: 0.19317007064819336
step: 80, loss: 0.08835513889789581
Gradient overflow.  Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 90, loss: 0.0792083740234375
step: 100, loss: 0.006859172601252794
step: 110, loss: 0.25867781043052673
step: 120, loss: 0.16783161461353302
step: 130, loss: 0.11209680885076523
step: 140, loss: 0.0689258873462677
step: 150, loss: 0.14476346969604492
step: 160, loss: 0.18658819794654846
step: 170, loss: 0.08254426717758179
step: 180, loss: 0.04942438751459122
step: 190, loss: 0.022585127502679825
step: 200, loss: 0.05238315090537071
step: 210, loss: 0.03189175948500633
/usr/local/lib/python3.7/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=0.6845124282982792, f1=0.661567877629063,
best_f1=0.6602687140115163
```

```
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.036258671432733536
step: 10, loss: 0.044209130108356476
step: 20, loss: 0.08524291217327118
step: 30, loss: 0.01522210892289877
step: 40, loss: 0.03478199243545532
step: 50, loss: 0.03269393742084503
step: 60, loss: 0.039633214473724365
step: 70, loss: 0.012402401305735111
step: 80, loss: 0.08546929061412811
step: 90, loss: 0.1155882403254509
step: 100, loss: 0.017045538872480392
step: 110, loss: 0.016512887552380562
step: 120, loss: 0.018426386639475822
step: 130, loss: 0.23243509232997894
step: 140, loss: 0.12982366979122162
step: 150, loss: 0.012962986715137959
step: 160, loss: 0.028802894055843353
step: 170, loss: 0.20357458293437958
step: 180, loss: 0.108008474111557
step: 190, loss: 0.03348458930850029
step: 200, loss: 0.03314198926091194
step: 210, loss: 0.1333669126033783
/usr/local/lib/python3.7/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=0.6782945736434107, f1=0.6305220883534137,
best_f1=0.6602687140115163
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.008201025426387787
step: 10, loss: 0.012597551569342613
step: 20, loss: 0.01755892112851143
step: 30, loss: 0.028729122132062912
step: 40, loss: 0.06333138793706894
step: 50, loss: 0.09282928705215454
step: 60, loss: 0.030311260372400284
step: 70, loss: 0.030941685661673546
step: 80, loss: 0.04151032865047455
step: 90, loss: 0.11013346165418625
step: 100, loss: 0.005585697013884783
step: 110, loss: 0.22471588850021362
step: 120, loss: 0.07909556478261948
step: 130, loss: 0.09482309222221375
step: 140, loss: 0.03022785671055317
```

```
step: 150, loss: 0.005826036911457777
step: 160, loss: 0.05504532903432846
step: 170, loss: 0.02035543881356716
step: 180, loss: 0.022567622363567352
step: 190, loss: 0.06585598737001419
step: 200, loss: 0.12270057201385498
step: 210, loss: 0.020176513120532036
/usr/local/lib/python3.7/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=0.6777777777777778, f1=0.6298342541436465,
best_f1=0.6602687140115163
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.03813336044549942
step: 10, loss: 0.051950909197330475
step: 20, loss: 0.030180664733052254
step: 30, loss: 0.03145000338554382
step: 40, loss: 0.01441612932831049
step: 50, loss: 0.09283449500799179
step: 60, loss: 0.22982068359851837
step: 70, loss: 0.04564806446433067
step: 80, loss: 0.0275464728474617
step: 90, loss: 0.028527090325951576
step: 100, loss: 0.029840078204870224
step: 110, loss: 0.07626857608556747
step: 120, loss: 0.11674680560827255
step: 130, loss: 0.0034818865824490786
step: 140, loss: 0.09638125449419022
step: 150, loss: 0.013650788925588131
step: 160, loss: 0.02280491404235363
step: 170, loss: 0.005541782360523939
step: 180, loss: 0.10646980255842209
step: 190, loss: 0.024318719282746315
step: 200, loss: 0.006971339229494333
step: 210, loss: 0.12148264050483704
/usr/local/lib/python3.7/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=0.6749116607773852, f1=0.6443661971830986,
best_f1=0.6602687140115163
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.01086603756994009
step: 10, loss: 0.044388722628355026
step: 20, loss: 0.0015463503077626228
```

```
step: 30, loss: 0.017910419031977654
step: 40, loss: 0.0657251626253128
step: 50, loss: 0.20473405718803406
step: 60, loss: 0.09129220992326736
step: 70, loss: 0.19022522866725922
step: 80, loss: 0.036637865006923676
step: 90, loss: 0.007282128091901541
step: 100, loss: 0.021125206723809242
step: 110, loss: 0.008844166062772274
step: 120, loss: 0.007906586863100529
step: 130, loss: 0.011861161328852177
step: 140, loss: 0.03238208219408989
step: 150, loss: 0.07858510315418243
step: 160, loss: 0.001864359830506146
step: 170, loss: 0.08528245985507965
step: 180, loss: 0.026723338291049004
step: 190, loss: 0.032848674803972244
step: 200, loss: 0.11070495843887329
step: 210, loss: 0.020563047379255295
/usr/local/lib/python3.7/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=0.6704761904761904, f1=0.6538461538461539,
best_f1=0.6602687140115163
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00634409487247467
step: 10, loss: 0.05429684743285179
step: 20, loss: 0.0009528380469419062
step: 30, loss: 0.0269283726811409
step: 40, loss: 0.004339852835983038
step: 50, loss: 0.062080591917037964
step: 60, loss: 0.02531030960381031
step: 70, loss: 0.027336139231920242
step: 80, loss: 0.0117233507335186
step: 90, loss: 0.047442540526390076
step: 100, loss: 0.046145785599946976
step: 110, loss: 0.024516146630048752
step: 120, loss: 0.037365105003118515
step: 130, loss: 0.0371159128844738
step: 140, loss: 0.01666734553873539
step: 150, loss: 0.1322038322687149
step: 160, loss: 0.02982565574347973
step: 170, loss: 0.030719708651304245
step: 180, loss: 0.012354628182947636
step: 190, loss: 0.060155972838401794
step: 200, loss: 0.008012364618480206
```

step: 210, loss: 0.046780772507190704
/usr/local/lib/python3.7/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=0.6679920477137177, f1=0.626984126984127,
best_f1=0.6602687140115163
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.08442655205726624
step: 10, loss: 0.01795981638133526
step: 20, loss: 0.025394072756171227
step: 30, loss: 0.0015009924536570907
step: 40, loss: 0.034022342413663864
step: 50, loss: 0.03386262431740761
step: 60, loss: 0.1297164261341095
step: 70, loss: 0.007500261999666691
step: 80, loss: 0.18360671401023865
step: 90, loss: 0.05574227496981621
step: 100, loss: 0.21540257334709167
step: 110, loss: 0.034185633063316345
step: 120, loss: 0.028818070888519287
step: 130, loss: 0.04166779715228081
step: 140, loss: 0.018840594217181206
step: 150, loss: 0.0023004808463156223
step: 160, loss: 0.04319607838988304
step: 170, loss: 0.12787030637264252
step: 180, loss: 0.00843667984008789
step: 190, loss: 0.011640525422990322
step: 200, loss: 0.0015326141146942973
step: 210, loss: 0.004533256869763136
/usr/local/lib/python3.7/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=0.6796875, f1=0.6455445544554456, best_f1=0.6602687140115163
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.04774558171629906
step: 10, loss: 0.0011328987311571836
step: 20, loss: 0.03329800069332123
step: 30, loss: 0.02573475055396557
step: 40, loss: 0.029503632336854935
step: 50, loss: 0.0102681964635849
step: 60, loss: 0.024892613291740417
step: 70, loss: 0.017998769879341125
step: 80, loss: 0.019424743950366974
step: 90, loss: 0.04704062268137932

```
step: 100, loss: 0.10649888217449188
step: 110, loss: 0.010094301775097847
step: 120, loss: 0.004402150399982929
step: 130, loss: 0.014233139343559742
step: 140, loss: 0.0034054103307425976
step: 150, loss: 0.030449680984020233
step: 160, loss: 0.0013055920135229826
step: 170, loss: 0.012780854478478432
step: 180, loss: 0.009127200581133366
step: 190, loss: 0.014651505276560783
step: 200, loss: 0.002840791130438447
step: 210, loss: 0.015225684270262718
/usr/local/lib/python3.7/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=0.6745562130177514, f1=0.6217821782178218,
best_f1=0.6602687140115163
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.06664949655532837
step: 10, loss: 0.0005417956272140145
step: 20, loss: 0.0815153419971466
step: 30, loss: 0.22164411842823029
step: 40, loss: 0.03726443648338318
step: 50, loss: 0.009639261290431023
step: 60, loss: 0.0021261433139443398
step: 70, loss: 0.031758058816194534
step: 80, loss: 0.004853283520787954
step: 90, loss: 0.0008378481725230813
Gradient overflow.  Skipping step, loss scaler 0 reducing loss scale to 8192.0
step: 100, loss: 0.001558643183670938
step: 110, loss: 0.003870993619784713
step: 120, loss: 0.06911928951740265
step: 130, loss: 0.004262561909854412
step: 140, loss: 0.0018739630468189716
step: 150, loss: 0.005796435289084911
step: 160, loss: 0.006482153665274382
step: 170, loss: 0.002456907881423831
step: 180, loss: 0.08066383004188538
step: 190, loss: 0.028921842575073242
step: 200, loss: 0.0020825748797506094
step: 210, loss: 0.0077072130516171455
/usr/local/lib/python3.7/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=0.6703910614525139, f1=0.632768361581921,
best_f1=0.6602687140115163
```

```
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.003595168935135007
step: 10, loss: 0.03299511969089508
step: 20, loss: 0.045291561633348465
step: 30, loss: 0.003825356252491474
step: 40, loss: 0.003134847618639469
step: 50, loss: 0.04149796813726425
step: 60, loss: 0.0012334410566836596
step: 70, loss: 0.003942342475056648
step: 80, loss: 0.0482274629175663
step: 90, loss: 0.09953691810369492
step: 100, loss: 0.008038888685405254
step: 110, loss: 0.00548273790627718
step: 120, loss: 0.03656699135899544
step: 130, loss: 0.003206951543688774
step: 140, loss: 0.04963789880275726
step: 150, loss: 0.006694663781672716
step: 160, loss: 0.011363727040588856
step: 170, loss: 0.013950797729194164
step: 180, loss: 0.008942659944295883
step: 190, loss: 0.0012861958239227533
step: 200, loss: 0.001840227167122066
step: 210, loss: 0.011081070639193058
/usr/local/lib/python3.7/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=0.6741573033707865, f1=0.6259541984732825,
best_f1=0.6602687140115163
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00051690055988729
step: 10, loss: 0.0005308325053192675
step: 20, loss: 0.003378119319677353
step: 30, loss: 0.030674301087856293
step: 40, loss: 0.0063453419134020805
step: 50, loss: 0.002676219679415226
step: 60, loss: 0.04937714338302612
step: 70, loss: 0.001819196273572743
step: 80, loss: 0.0010219716932624578
step: 90, loss: 0.013871180824935436
step: 100, loss: 0.006441762670874596
step: 110, loss: 0.001110854558646679
step: 120, loss: 0.006191359832882881
step: 130, loss: 0.009376094676554203
step: 140, loss: 0.0009132529376074672
```

```
step: 150, loss: 0.002283437643200159
step: 160, loss: 0.004335750360041857
step: 170, loss: 0.0006887535564601421
step: 180, loss: 0.0002992021618410945
step: 190, loss: 0.010078425519168377
step: 200, loss: 0.004206703510135412
step: 210, loss: 0.020142706111073494
/usr/local/lib/python3.7/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=0.6770428015564203, f1=0.6237816764132554,
best_f1=0.6602687140115163
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, 281.03it/s]
load_f1 = 0.7017543859649122
real_f1 = 0.69921875
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, 185.03it/s]
```

1.1.5 Walmart-Amazon - 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/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
```

81

```
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: 0.5713953971862793
step: 10, loss: 0.36597007513046265
step: 20, loss: 0.29090356826782227
step: 30, loss: 0.4396257996559143
step: 40, loss: 0.41501766443252563
step: 50, loss: 0.28450602293014526
step: 60, loss: 0.23495253920555115
step: 70, loss: 0.2580818831920624
step: 80, loss: 0.3696858584880829
step: 90, loss: 0.2715051770210266
step: 100, loss: 0.2673864960670471
step: 110, loss: 0.28615814447402954
step: 120, loss: 0.04424536973237991
Gradient overflow.  Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 130, loss: 0.1435668021440506
step: 140, loss: 0.04594852030277252
step: 150, loss: 0.14782533049583435
step: 160, loss: 0.034631531685590744
step: 170, loss: 0.2529412508010864
step: 180, loss: 0.04719490557909012
step: 190, loss: 0.18079321086406708
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
```

```
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.6386554621848739, f1=0.6551724137931034,
best_f1=0.6551724137931034
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.15462927520275116
step: 10, loss: 0.10671914368867874
step: 20, loss: 0.155114084482193
step: 30, loss: 0.08402692526578903
step: 40, loss: 0.10692373663187027
step: 50, loss: 0.09987255185842514
step: 60, loss: 0.3403746783733368
step: 70, loss: 0.2158690243959427
step: 80, loss: 0.22331339120864868
step: 90, loss: 0.15635773539543152
step: 100, loss: 0.011298095807433128
step: 110, loss: 0.07993175834417343
step: 120, loss: 0.22605013847351074
step: 130, loss: 0.038847699761390686
step: 140, loss: 0.05829255282878876
step: 150, loss: 0.11398463696241379
step: 160, loss: 0.018169106915593147
step: 170, loss: 0.16388072073459625
step: 180, loss: 0.11644705384969711
step: 190, loss: 0.08721638470888138
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.7863247863247862, f1=0.7877094972067039,
best_f1=0.7877094972067039
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.027608459815382957
step: 10, loss: 0.2068047672510147
step: 20, loss: 0.18215614557266235
step: 30, loss: 0.1409163475036621
step: 40, loss: 0.12309923768043518
step: 50, loss: 0.19497454166412354
```

```
step: 60, loss: 0.032188307493925095
step: 70, loss: 0.10452820360660553
step: 80, loss: 0.17867770791053772
step: 90, loss: 0.09748116135597229
step: 100, loss: 0.07328921556472778
step: 110, loss: 0.00998812634497881
step: 120, loss: 0.021200435236096382
step: 130, loss: 0.028454937040805817
step: 140, loss: 0.031100083142518997
step: 150, loss: 0.10852992534637451
step: 160, loss: 0.07934604585170746
step: 170, loss: 0.13658437132835388
step: 180, loss: 0.04974547028541565
step: 190, loss: 0.11717117577791214
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.7922437673130193, f1=0.7896253602305476,
best_f1=0.7896253602305476
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.04117247089743614
step: 10, loss: 0.1356584131717682
step: 20, loss: 0.07406715303659439
step: 30, loss: 0.08482196927070618
step: 40, loss: 0.12145064771175385
step: 50, loss: 0.017745008692145348
step: 60, loss: 0.1667526662349701
step: 70, loss: 0.11556793749332428
step: 80, loss: 0.10841846466064453
step: 90, loss: 0.017374811694025993
step: 100, loss: 0.035746391862630844
step: 110, loss: 0.0034292396157979965
step: 120, loss: 0.0650574341416359
step: 130, loss: 0.27608177065849304
step: 140, loss: 0.019772082567214966
step: 150, loss: 0.1074310839176178
step: 160, loss: 0.017524605616927147
step: 170, loss: 0.1388629972934723
step: 180, loss: 0.12089361250400543
step: 190, loss: 0.162501260638237
/usr/local/lib/python3.7/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=0.7821522309711286, f1=0.7609254498714654,
best_f1=0.7896253602305476
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.08281728625297546
step: 10, loss: 0.03849559649825096
step: 20, loss: 0.041657961905002594
step: 30, loss: 0.01568567007780075
step: 40, loss: 0.06963715702295303
step: 50, loss: 0.04461590573191643
step: 60, loss: 0.1246107742190361
step: 70, loss: 0.030137578025460243
step: 80, loss: 0.009864908643066883
step: 90, loss: 0.0258504506200552
Gradient overflow.  Skipping step, loss scaler 0 reducing loss scale to 8192.0
step: 100, loss: 0.00662204110994935
step: 110, loss: 0.00583443371579051
step: 120, loss: 0.07067828625440598
step: 130, loss: 0.0551731176674366
step: 140, loss: 0.02511274442076683
step: 150, loss: 0.01940545253455639
step: 160, loss: 0.31233111023902893
step: 170, loss: 0.1139197051525116
step: 180, loss: 0.0370701365172863
step: 190, loss: 0.07917950302362442
/usr/local/lib/python3.7/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=0.7903225806451613, f1=0.7830687830687831,
best_f1=0.7896253602305476
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.10054081678390503
step: 10, loss: 0.044361673295497894
step: 20, loss: 0.010160389356315136
step: 30, loss: 0.00109615339897573
step: 40, loss: 0.03498867154121399
step: 50, loss: 0.0972534567117691
step: 60, loss: 0.00943212304264307
step: 70, loss: 0.003285443875938654
step: 80, loss: 0.004601198248565197
step: 90, loss: 0.032069165259599686
step: 100, loss: 0.0022907163947820663
step: 110, loss: 0.0382375605404377
step: 120, loss: 0.0227949358522892
```

```
step: 130, loss: 0.004230060614645481
step: 140, loss: 0.027694690972566605
step: 150, loss: 0.0021362740080803633
step: 160, loss: 0.07327012717723846
step: 170, loss: 0.04482700303196907
step: 180, loss: 0.07534673064947128
step: 190, loss: 0.021200649440288544
/usr/local/lib/python3.7/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=0.772117962466488, f1=0.7659574468085106,
best_f1=0.7896253602305476
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.003288496285676956
step: 10, loss: 0.011541948653757572
step: 20, loss: 0.031035451218485832
step: 30, loss: 0.10196801275014877
step: 40, loss: 0.006145255174487829
step: 50, loss: 0.05757276341319084
step: 60, loss: 0.06846005469560623
step: 70, loss: 0.01766134984791279
step: 80, loss: 0.0024279432836920023
step: 90, loss: 0.017067966982722282
step: 100, loss: 0.002226448617875576
step: 110, loss: 0.008048581890761852
step: 120, loss: 0.0024557386059314013
step: 130, loss: 0.15148381888866425
step: 140, loss: 0.004480068106204271
step: 150, loss: 0.0010218415409326553
step: 160, loss: 0.16791176795959473
step: 170, loss: 0.0909385085105896
step: 180, loss: 0.00238889642059803
step: 190, loss: 0.001685967086814344
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.8062827225130891, f1=0.7946666666666666,
best_f1=0.7946666666666666
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0063565275631845
```

```
step: 10, loss: 0.02035815455019474
step: 20, loss: 0.005952455569058657
step: 30, loss: 0.014762070029973984
step: 40, loss: 0.0021533512044698
step: 50, loss: 0.00901661068201065
step: 60, loss: 0.0017091460758820176
step: 70, loss: 0.004638256039470434
step: 80, loss: 0.0008145183674059808
step: 90, loss: 0.0031912634149193764
step: 100, loss: 0.011429854668676853
step: 110, loss: 0.001698903040960431
step: 120, loss: 0.005310780368745327
step: 130, loss: 0.01751876249909401
step: 140, loss: 0.024650221690535545
step: 150, loss: 0.010251184925436974
step: 160, loss: 0.006353522185236216
step: 170, loss: 0.02789929322898388
step: 180, loss: 0.04883226752281189
step: 190, loss: 0.0022266595624387264
/usr/local/lib/python3.7/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=0.7774798927613942, f1=0.765498652291105,
best_f1=0.7946666666666666
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0060684275813400745
step: 10, loss: 0.03338444232940674
step: 20, loss: 0.0031447343062609434
step: 30, loss: 0.013660400174558163
step: 40, loss: 0.03500019386410713
step: 50, loss: 0.0045131198130548
step: 60, loss: 0.002446766709908843
step: 70, loss: 0.0026075770147144794
step: 80, loss: 0.001843073288910091
step: 90, loss: 0.01734914816915989
step: 100, loss: 0.043989844620227814
step: 110, loss: 0.001254464965313673
step: 120, loss: 0.011176369152963161
step: 130, loss: 0.0023328782990574837
step: 140, loss: 0.005780401639640331
step: 150, loss: 0.0046197520568966866
step: 160, loss: 0.001015214715152979
step: 170, loss: 0.0014972813660278916
step: 180, loss: 0.07434837520122528
step: 190, loss: 0.002931856317445636
/usr/local/lib/python3.7/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=0.7897435897435896, f1=0.7819548872180452,
best_f1=0.7946666666666666
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0019807424396276474
step: 10, loss: 0.05824992060661316
step: 20, loss: 0.031865015625953674
step: 30, loss: 0.0018833972280845046
step: 40, loss: 0.015832556411623955
step: 50, loss: 0.0005754731246270239
step: 60, loss: 0.000816097657661885
step: 70, loss: 0.0008115021628327668
step: 80, loss: 0.03136935457587242
step: 90, loss: 0.018660662695765495
step: 100, loss: 0.007297839038074017
step: 110, loss: 0.01653478294610977
step: 120, loss: 0.11227086186408997
step: 130, loss: 0.006080503109842539
step: 140, loss: 0.00301970262080431
step: 150, loss: 0.041819978505373
step: 160, loss: 0.01620326191186905
step: 170, loss: 0.0022895177826285362
step: 180, loss: 0.2327873408794403
step: 190, loss: 0.0014106258749961853
/usr/local/lib/python3.7/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=0.783289817232376, f1=0.783289817232376,
best_f1=0.7946666666666666
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0014418390346691012
step: 10, loss: 0.011065970174968243
step: 20, loss: 0.02498861774802208
step: 30, loss: 0.05318028852343559
step: 40, loss: 0.001423319918103516
step: 50, loss: 0.020419050008058548
step: 60, loss: 0.00048713828437030315
step: 70, loss: 0.014509234577417374
step: 80, loss: 0.0010689669288694859
step: 90, loss: 0.0037615953478962183
step: 100, loss: 0.0009547404479235411
step: 110, loss: 0.004425038117915392
step: 120, loss: 0.0006417073309421539
```

```
step: 130, loss: 0.0016537487972527742
step: 140, loss: 0.0046671610325574875
step: 150, loss: 0.0009090639650821686
step: 160, loss: 0.0007742232410237193
step: 170, loss: 0.0006685625994578004
step: 180, loss: 0.0016024879878386855
step: 190, loss: 0.0046195476315915585
/usr/local/lib/python3.7/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=0.773067331670823, f1=0.7450980392156864,
best_f1=0.7946666666666666
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.001564758480526507
step: 10, loss: 0.0052679977379739285
step: 20, loss: 0.0010680381674319506
step: 30, loss: 0.0005065581644885242
step: 40, loss: 0.0007652103086002171
step: 50, loss: 0.004546689800918102
step: 60, loss: 0.0015438739210367203
step: 70, loss: 0.0011170974466949701
step: 80, loss: 0.0017852156888693571
step: 90, loss: 0.0008792081498540938
step: 100, loss: 0.0009127809898927808
step: 110, loss: 0.0007510036812163889
step: 120, loss: 0.0004347946378402412
step: 130, loss: 0.0005159590509720147
step: 140, loss: 0.0031053603161126375
step: 150, loss: 0.0009550201939418912
step: 160, loss: 0.0008113458752632141
step: 170, loss: 0.001563416444696486
step: 180, loss: 0.0005220940802246332
step: 190, loss: 0.0008887051953934133
/usr/local/lib/python3.7/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=0.7989690721649483, f1=0.7788944723618091,
best_f1=0.7946666666666666
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.13375568389892578
step: 10, loss: 0.0803937241435051
step: 20, loss: 0.00524055166170001
step: 30, loss: 0.0006165573140606284
step: 40, loss: 0.0026473840698599815
```

```
step: 50, loss: 0.0023588803596794605
step: 60, loss: 0.0009515451965853572
step: 70, loss: 0.001736028236337006
step: 80, loss: 0.010810238309204578
step: 90, loss: 0.0011118088150396943
step: 100, loss: 0.0019483695505186915
step: 110, loss: 0.01506292074918747
step: 120, loss: 0.0020903206896036863
step: 130, loss: 0.00037787959445267916
step: 140, loss: 0.0003591335844248533
step: 150, loss: 0.0005096212844364345
step: 160, loss: 0.0007002914207987487
step: 170, loss: 0.0007201214320957661
step: 180, loss: 0.0021141599863767624
step: 190, loss: 0.04032883048057556
/usr/local/lib/python3.7/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=0.7928388746803069, f1=0.775, best_f1=0.7946666666666666
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.001620415598154068
step: 10, loss: 0.00040764687582850456
step: 20, loss: 0.0003274352930020541
step: 30, loss: 0.0005048796301707625
step: 40, loss: 0.0020006406120955944
step: 50, loss: 0.0010631382465362549
step: 60, loss: 0.004840065259486437
step: 70, loss: 0.00037488600355573
step: 80, loss: 0.0003542563645169139
step: 90, loss: 0.06722576916217804
step: 100, loss: 0.0021851882338523865
step: 110, loss: 0.0034830558579415083
step: 120, loss: 0.0031060767360031605
step: 130, loss: 0.0017864821711555123
step: 140, loss: 0.08998211473226547
step: 150, loss: 0.0008153589442372322
step: 160, loss: 0.0031459617894142866
step: 170, loss: 0.0011529983021318913
step: 180, loss: 0.0012179192854091525
step: 190, loss: 0.0008590673096477985
/usr/local/lib/python3.7/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=0.7969151670951158, f1=0.789873417721519,
best_f1=0.7946666666666666
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
```

```
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.031081942841410637
step: 10, loss: 0.000503536022733897
step: 20, loss: 0.0007295070681720972
step: 30, loss: 0.0008648857474327087
step: 40, loss: 0.001694891368970275
step: 50, loss: 0.010414054617285728
step: 60, loss: 0.006975201424211264
step: 70, loss: 0.0017490516183897853
step: 80, loss: 0.000776509812567383
step: 90, loss: 0.000833774683997035
step: 100, loss: 0.0022931070998311043
step: 110, loss: 0.0018416928360238671
step: 120, loss: 0.001659812987782061
step: 130, loss: 0.0008834523614495993
step: 140, loss: 0.12071750313043594
step: 150, loss: 0.0011006994172930717
step: 160, loss: 0.0005740750930272043
step: 170, loss: 0.012705820612609386
step: 180, loss: 0.0013890624977648258
step: 190, loss: 0.004342812113463879
/usr/local/lib/python3.7/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=0.7969151670951158, f1=0.7858942065491183,
best_f1=0.7946666666666666
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, 164.32it/s]
load_f1 = 0.7374301675977653
real_f1 = 0.7061994609164419
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, 183.05it/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: 0.6508017778396606
step: 10, loss: 0.37014082074165344
step: 20, loss: 0.30151501297950745
step: 30, loss: 0.3774060606956482
step: 40, loss: 0.26833775639533997

```
step: 50, loss: 0.2627197206020355
step: 60, loss: 0.2626081109046936
step: 70, loss: 0.37545379996299744
step: 80, loss: 0.4021260142326355
step: 90, loss: 0.23043376207351685
step: 100, loss: 0.22389978170394897
step: 110, loss: 0.26520395278930664
step: 120, loss: 0.18283258378505707
step: 130, loss: 0.0155501002445817
step: 140, loss: 0.26080521941185
step: 150, loss: 0.26983213424682617
step: 160, loss: 0.131405770778656
step: 170, loss: 0.2603968679904938
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.7428571428571429, f1=0.6924829157175398,
best_f1=0.6924829157175398
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.09318426996469498
step: 10, loss: 0.27214834094047546
step: 20, loss: 0.10133615881204605
step: 30, loss: 0.1970413476228714
step: 40, loss: 0.03626849874854088
step: 50, loss: 0.22509920597076416
step: 60, loss: 0.18079440295696259
step: 70, loss: 0.030362965539097786
Gradient overflow.  Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 80, loss: 0.1652430146932602
step: 90, loss: 0.09869427233934402
step: 100, loss: 0.12594100832939148
step: 110, loss: 0.08240988105535507
step: 120, loss: 0.28438976407051086
step: 130, loss: 0.10478556156158447
step: 140, loss: 0.29702168703079224
step: 150, loss: 0.12379471957683563
step: 160, loss: 0.16390804946422577
step: 170, loss: 0.11179737001657486
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
```

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=0.7713004484304932, f1=0.7428571428571429,
best_f1=0.7428571428571429
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.09332569688558578
step: 10, loss: 0.06343653798103333
step: 20, loss: 0.02359546534717083
step: 30, loss: 0.3062361180782318
step: 40, loss: 0.11452141404151917
step: 50, loss: 0.03373045101761818
step: 60, loss: 0.11460627615451813
step: 70, loss: 0.05287321284413338
step: 80, loss: 0.036168407648801804
step: 90, loss: 0.08440648019313812
step: 100, loss: 0.009607430547475815
step: 110, loss: 0.0750541090965271
step: 120, loss: 0.05009841173887253
step: 130, loss: 0.07029586285352707
step: 140, loss: 0.035264864563941956
step: 150, loss: 0.22849905490875244
step: 160, loss: 0.017936866730451584
step: 170, loss: 0.20577335357666016
/usr/local/lib/python3.7/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=0.7659574468085106, f1=0.7808219178082191,
best_f1=0.7428571428571429
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.012919265776872635
step: 10, loss: 0.02094007097184658
step: 20, loss: 0.08955798298120499
step: 30, loss: 0.05750463530421257
step: 40, loss: 0.006980641279369593
step: 50, loss: 0.05468466505408287
step: 60, loss: 0.09683454036712646
step: 70, loss: 0.016501083970069885
step: 80, loss: 0.08534888178110123
step: 90, loss: 0.0661911740899086
step: 100, loss: 0.24565835297107697
step: 110, loss: 0.0313933826982975
step: 120, loss: 0.03181761875748634
step: 130, loss: 0.06261614710092545

```
step: 140, loss: 0.031215060502290726
step: 150, loss: 0.14348874986171722
step: 160, loss: 0.05414145439863205
step: 170, loss: 0.16945752501487732
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.8128078817733989, f1=0.7971698113207547,
best_f1=0.7971698113207547
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.059104762971401215
step: 10, loss: 0.013607158325612545
step: 20, loss: 0.05217916890978813
step: 30, loss: 0.0034521969500929117
step: 40, loss: 0.015408781357109547
step: 50, loss: 0.01403993833810091
step: 60, loss: 0.0927409753203392
step: 70, loss: 0.10347811877727509
step: 80, loss: 0.042719561606645584
step: 90, loss: 0.0627584233880043
step: 100, loss: 0.016848847270011902
step: 110, loss: 0.032945193350315094
step: 120, loss: 0.0036046770401299
step: 130, loss: 0.026608996093273163
step: 140, loss: 0.024196309968829155
step: 150, loss: 0.09739633649587631
step: 160, loss: 0.04293755814433098
step: 170, loss: 0.004460733849555254
/usr/local/lib/python3.7/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=0.7990867579908676, f1=0.7955555555555555,
best_f1=0.7971698113207547
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.009692475199699402
step: 10, loss: 0.025939766317605972
step: 20, loss: 0.001979382010176778
step: 30, loss: 0.019940027967095375
step: 40, loss: 0.005318109877407551
step: 50, loss: 0.2319859266281128
```

```
step: 60, loss: 0.07336561381816864
step: 70, loss: 0.12984582781791687
step: 80, loss: 0.047625817358493805
step: 90, loss: 0.05133051797747612
step: 100, loss: 0.0032184591982513666
step: 110, loss: 0.008888894692063332
step: 120, loss: 0.014039186760783195
step: 130, loss: 0.005366511642932892
step: 140, loss: 0.01639285311102867
step: 150, loss: 0.00965900532901287
step: 160, loss: 0.07310846447944641
step: 170, loss: 0.01470182929188013
/usr/local/lib/python3.7/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=0.7967914438502675, f1=0.7969924812030076,
best_f1=0.7971698113207547
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0025829484220594168
step: 10, loss: 0.0023685910273343325
step: 20, loss: 0.011820957064628601
step: 30, loss: 0.007684449665248394
step: 40, loss: 0.012572857551276684
step: 50, loss: 0.003681881818920374
step: 60, loss: 0.0022688889876008034
step: 70, loss: 0.0011295879958197474
step: 80, loss: 0.0428287535905838
step: 90, loss: 0.0011422333773225546
step: 100, loss: 0.0005271583213470876
step: 110, loss: 0.01395932026207447
step: 120, loss: 0.01908322423696518
step: 130, loss: 0.10477763414382935
step: 140, loss: 0.02140052616596222
step: 150, loss: 0.0005705048679374158
step: 160, loss: 0.03465760871767998
step: 170, loss: 0.023062974214553833
/usr/local/lib/python3.7/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=0.7716535433070866, f1=0.7960199004975125,
best_f1=0.7971698113207547
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0007844371721148491
step: 10, loss: 0.0010299376444891095
```

step: 20, loss: 0.0023941195104271173
step: 30, loss: 0.007397486362606287
step: 40, loss: 0.00015137023001443595
step: 50, loss: 0.0022010228130966425
step: 60, loss: 0.002297842875123024
step: 70, loss: 0.005618109833449125
step: 80, loss: 0.007044227793812752
step: 90, loss: 0.004286285489797592
step: 100, loss: 0.05562727898359299
step: 110, loss: 0.06023945286870003
step: 120, loss: 0.00831915345042944
step: 130, loss: 0.0010879267938435078
step: 140, loss: 0.003700357163324952
step: 150, loss: 0.006168432999402285
step: 160, loss: 0.07342823594808578
step: 170, loss: 0.002280409215018153
/usr/local/lib/python3.7/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=0.781725888324873, f1=0.7897196261682243,
best_f1=0.7971698113207547
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.025070548057556152
step: 10, loss: 0.0022365818731486797
step: 20, loss: 0.00022406893549486995
step: 30, loss: 0.033337660133838654
step: 40, loss: 0.00482949847355485
step: 50, loss: 0.00026294219424016774
step: 60, loss: 0.0036465059965848923
step: 70, loss: 0.0006190433050505817
step: 80, loss: 0.020706601440906525
step: 90, loss: 0.044089339673519135
step: 100, loss: 0.003031775588169694
step: 110, loss: 0.005943763069808483
step: 120, loss: 0.0036545065231621265
step: 130, loss: 0.07264706492424011
step: 140, loss: 0.009714096784591675
step: 150, loss: 0.00015023231389932334
step: 160, loss: 0.04975646361708641
step: 170, loss: 0.003158397739753127
/usr/local/lib/python3.7/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=0.787128712871287, f1=0.8179669030732861,
best_f1=0.7971698113207547
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:

```
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.047648847103118896
step: 10, loss: 0.0045571839436888695
step: 20, loss: 0.012623528949916363
step: 30, loss: 0.00032272134558297694
step: 40, loss: 0.017326388508081436
step: 50, loss: 0.031150899827480316
step: 60, loss: 0.00011645457561826333
step: 70, loss: 0.011510123498737812
step: 80, loss: 0.00011729726975318044
step: 90, loss: 0.0008276844164356589
step: 100, loss: 0.00013283042062539607
step: 110, loss: 0.08267048746347427
step: 120, loss: 0.002229271689429879
step: 130, loss: 0.0004798964655492455
step: 140, loss: 0.0217727217823267
step: 150, loss: 0.053509317338466644
step: 160, loss: 0.0068598780781030655
step: 170, loss: 0.0027969495858997107
/usr/local/lib/python3.7/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=0.7642276422764228, f1=0.8061224489795918,
best_f1=0.7971698113207547
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.03261655941605568
step: 10, loss: 0.00010773850226541981
step: 20, loss: 0.0003137100429739803
step: 30, loss: 0.00013690217747353017
step: 40, loss: 7.789250958012417e-05
step: 50, loss: 0.026243047788739204
step: 60, loss: 0.002407355699688196
step: 70, loss: 0.008456428535282612
step: 80, loss: 0.0001612323831068352
step: 90, loss: 0.00013716451940126717
step: 100, loss: 0.0001871747663244605
step: 110, loss: 0.01634659618139267
step: 120, loss: 0.00030147621873766184
step: 130, loss: 0.0001119593289331533
step: 140, loss: 0.0035226482432335615
step: 150, loss: 0.0023427370470017195
step: 160, loss: 0.0022377038840204477
step: 170, loss: 0.00015707571583334357
/usr/local/lib/python3.7/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=0.7990314769975787, f1=0.7935779816513762,
best_f1=0.7971698113207547
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00012950404197908938
step: 10, loss: 0.00045638447045348585
step: 20, loss: 7.668208854738623e-05
step: 30, loss: 8.364645327674225e-05
step: 40, loss: 0.001190792303532362
step: 50, loss: 0.00028673664201050997
step: 60, loss: 0.000435427064076066
step: 70, loss: 0.05068981647491455
step: 80, loss: 0.00010477196337888017
step: 90, loss: 0.00030334541224874556
step: 100, loss: 0.00125010940246284
step: 110, loss: 0.014512897469103336
step: 120, loss: 0.008462676778435707
step: 130, loss: 0.0010573847685009241
step: 140, loss: 0.0015685658436268568
step: 150, loss: 0.0007948518032208085
step: 160, loss: 0.00011829325376311317
step: 170, loss: 0.0001178843667730689
/usr/local/lib/python3.7/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=0.7792207792207791, f1=0.7990430622009569,
best_f1=0.7971698113207547
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.011540760286152363
step: 10, loss: 9.197105100611225e-05
step: 20, loss: 0.0007342586177401245
step: 30, loss: 4.5442484406521544e-05
step: 40, loss: 7.498218474211171e-05
step: 50, loss: 0.000324094493407561
step: 60, loss: 0.0019679879769682884
step: 70, loss: 0.03485048934817314
step: 80, loss: 0.06671269237995148
step: 90, loss: 8.609583892393857e-05
step: 100, loss: 0.12202487885951996
step: 110, loss: 0.00030886256718076766
step: 120, loss: 0.05452214181423187
step: 130, loss: 0.00016749992209952325
step: 140, loss: 0.0005413363687694073
Gradient overflow.  Skipping step, loss scaler 0 reducing loss scale to 16384.0
```

```
step: 150, loss: 0.0007723423768766224
step: 160, loss: 0.0013912295689806342
step: 170, loss: 0.0003089234814979136
/usr/local/lib/python3.7/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=0.7823834196891193, f1=0.8076923076923078,
best_f1=0.7971698113207547
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0002157351264031604
step: 10, loss: 8.829720172798261e-05
step: 20, loss: 0.00012963431072421372
step: 30, loss: 0.0001308847713517025
step: 40, loss: 8.805794641375542e-05
step: 50, loss: 9.61672849371098e-05
step: 60, loss: 8.243740012403578e-05
step: 70, loss: 0.00032079964876174927
step: 80, loss: 0.001254648668691516
step: 90, loss: 0.00015864182205405086
step: 100, loss: 5.9782560128951445e-05
step: 110, loss: 0.00011992458166787401
step: 120, loss: 0.001849773689173162
step: 130, loss: 0.001533731585368514
step: 140, loss: 0.00027497828705236316
step: 150, loss: 5.870247332495637e-05
step: 160, loss: 4.395346695673652e-05
step: 170, loss: 0.00018369543249718845
/usr/local/lib/python3.7/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=0.7890818858560793, f1=0.810304449648712,
best_f1=0.7971698113207547
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00012659290223382413
step: 10, loss: 0.021372444927692413
step: 20, loss: 0.025938719511032104
step: 30, loss: 0.00041906186379492283
step: 40, loss: 0.0007824511267244816
step: 50, loss: 6.63423488731496e-05
step: 60, loss: 0.003790121292695403
step: 70, loss: 4.52674794360064e-05
step: 80, loss: 0.03561288118362427
step: 90, loss: 0.014140470884740353
step: 100, loss: 0.00033336502383463085
```

```
step: 110, loss: 7.006667146924883e-05
step: 120, loss: 0.00011436192289693281
step: 130, loss: 0.0005004284903407097
step: 140, loss: 9.403364674653858e-05
step: 150, loss: 0.0017260332824662328
step: 160, loss: 7.680700946366414e-05
step: 170, loss: 0.0011093338252976537
/usr/local/lib/python3.7/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=0.7704485488126649, f1=0.8148148148148149,
best_f1=0.7971698113207547
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, 226.91it/s]
load_f1 = 0.40476190476190477
real_f1 = 0.3724696356275304
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
```

```
4398it [00:23, 183.98it/s]
```

1.3   DA DIRTY

1.3.1   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 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, 522kB/s]
Downloading: 100% 232k/232k [00:00<00:00, 2.65MB/s]
Downloading: 100% 440M/440M [00:12<00:00, 36.4MB/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: 0.6374112367630005
step: 10, loss: 0.6111491322517395
step: 20, loss: 0.3884900212287903
step: 30, loss: 0.10773272812366486
step: 40, loss: 0.20705385506153107
step: 50, loss: 0.0583997517824173
step: 60, loss: 0.10728656500577927
step: 70, loss: 0.059156373143196106
step: 80, loss: 0.05844314768910408
step: 90, loss: 0.10975848138332367
step: 100, loss: 0.010257083922624588
step: 110, loss: 0.2951129078865051
step: 120, loss: 0.006196639034897089
step: 130, loss: 0.006480564828962088
step: 140, loss: 0.004214535001665354
step: 150, loss: 0.04568469524383545
step: 160, loss: 0.009131337516009808
step: 170, loss: 0.14555776119232178
```

```
step: 180, loss: 0.15463386476039886
step: 190, loss: 0.04021920636296272
step: 200, loss: 0.07479310780763626
step: 210, loss: 0.004508759826421738
step: 220, loss: 0.014682442881166935
step: 230, loss: 0.022360071539878845
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9642058165548099, f1=0.9592760180995475,
best_f1=0.9592760180995475
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0096937194466590088
step: 10, loss: 0.004716751631349325
step: 20, loss: 0.20243461430072784
step: 30, loss: 0.13871745765209198
step: 40, loss: 0.03426488861441612
step: 50, loss: 0.0035780263133347034
step: 60, loss: 0.002102046739310026
step: 70, loss: 0.05676734447479248
step: 80, loss: 0.027119185775518417
step: 90, loss: 0.0753648653626442
step: 100, loss: 0.06941071152687073
step: 110, loss: 0.05138611048460007
step: 120, loss: 0.12378107011318207
step: 130, loss: 0.005966637283563614
step: 140, loss: 0.0015214952873066068
step: 150, loss: 0.0713614895939827
step: 160, loss: 0.01445830799639225
step: 170, loss: 0.0007477056933566928
step: 180, loss: 0.0030019923578947783
step: 190, loss: 0.003011075546965003
step: 200, loss: 0.0008431681781075895
step: 210, loss: 0.0004772826796397567
step: 220, loss: 0.13583604991436005
step: 230, loss: 0.0025938530452549458
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.971815107102593, f1=0.972972972972973,
best_f1=0.972972972972973
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.003151038195937872
step: 10, loss: 0.0037254798226058483
step: 20, loss: 0.048445042222738266
step: 30, loss: 0.003410705365240574
step: 40, loss: 0.11386196315288544
step: 50, loss: 0.005112338345497847
step: 60, loss: 0.002403060207143426
step: 70, loss: 0.006269186269491911
step: 80, loss: 0.0016555151669308543
step: 90, loss: 0.025329889729619026
step: 100, loss: 0.0008446025894954801
step: 110, loss: 0.002279765671119094
step: 120, loss: 0.0017453193431720138
step: 130, loss: 0.0004660518025048077
step: 140, loss: 0.0032061166130006313
step: 150, loss: 0.0081625422462821
step: 160, loss: 0.02509242855012417
step: 170, loss: 0.030703090131282806
step: 180, loss: 0.007178278639912605
step: 190, loss: 0.009972725063562393
step: 200, loss: 0.18378527462482452
step: 210, loss: 0.006621772423386574
step: 220, loss: 0.00827522948384285
step: 230, loss: 0.09310071170330048
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9733924611973392, f1=0.9698996655518396,
best_f1=0.9698996655518396
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0026881496887654066
step: 10, loss: 0.0016907872632145882
step: 20, loss: 0.0008372999727725983
step: 30, loss: 0.0019373359391465783
step: 40, loss: 0.003889305517077446
step: 50, loss: 0.004832930862903595
```

```
step: 60, loss: 0.008800826966762543
step: 70, loss: 0.000494733452796936
step: 80, loss: 0.0014376427279785275
step: 90, loss: 0.0019094857852905989
step: 100, loss: 0.0011645450722426176
step: 110, loss: 0.0007096805493347347
step: 120, loss: 0.002709808060899377
step: 130, loss: 0.00239715538918972
step: 140, loss: 0.0005125892348587513
step: 150, loss: 0.1796960085630417
step: 160, loss: 0.14334706962108612
step: 170, loss: 0.00496085686609149
step: 180, loss: 0.0029577675741165876
step: 190, loss: 0.001152673503383994
step: 200, loss: 0.0016908959951251745
step: 210, loss: 0.00538281723856926
step: 220, loss: 0.0007190273609012365
step: 230, loss: 0.018810363486409187
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.977728285077951, f1=0.9732739420935412,
best_f1=0.9732739420935412
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.026441436260938644
step: 10, loss: 0.0008627603529021144
step: 20, loss: 0.0004928801208734512
step: 30, loss: 0.00020959442190360278
step: 40, loss: 0.02752051316201687
step: 50, loss: 0.00026576058007776737
step: 60, loss: 0.1845388561487198
step: 70, loss: 0.0011233786353841424
step: 80, loss: 0.0032041803933680058
step: 90, loss: 0.0018621855415403843
step: 100, loss: 0.0007277477998286486
step: 110, loss: 0.011228874325752258
step: 120, loss: 0.006968355271965265
step: 130, loss: 0.002676063682883978
step: 140, loss: 0.006733624264597893
step: 150, loss: 0.003188252681866288
step: 160, loss: 0.00041303259786218405
step: 170, loss: 0.0148519491776824
```

```
step: 180, loss: 0.004187657963484526
step: 190, loss: 0.08725697547197342
step: 200, loss: 0.005746576935052872
step: 210, loss: 0.005236005410552025
step: 220, loss: 0.0008532306528650224
step: 230, loss: 0.00032049359288066626
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9798206278026906, f1=0.970917225950783,
best_f1=0.970917225950783
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0022573198657482862
step: 10, loss: 0.00111444415475428104
step: 20, loss: 0.003650335595011711
step: 30, loss: 0.00024297831987496465
step: 40, loss: 0.00031760375713929534
step: 50, loss: 0.0005847482243552804
step: 60, loss: 0.00030291921575553715
step: 70, loss: 0.002102380385622382
step: 80, loss: 0.003713826183229685
step: 90, loss: 0.00043499653111211956
step: 100, loss: 0.005750415381044149
step: 110, loss: 0.0007868606480769813
step: 120, loss: 0.00047462419024668634
step: 130, loss: 0.0006862377631478012
step: 140, loss: 0.00041592857451178133
step: 150, loss: 0.002255016006529331
step: 160, loss: 0.002863791771233082
step: 170, loss: 0.001316435169428587
step: 180, loss: 0.008835085667669773
step: 190, loss: 0.08578633517026901
step: 200, loss: 0.0007113097817637026
step: 210, loss: 0.0013745344476774335
step: 220, loss: 0.0009466800838708878
step: 230, loss: 0.023656100034713745
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.980963045912654, f1=0.9763779527559054,
best_f1=0.9763779527559054
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0008249772363342345
step: 10, loss: 0.00012726857676170766
step: 20, loss: 0.00015605363296344876
step: 30, loss: 0.00017532167839817703
step: 40, loss: 0.00030041977879591286
step: 50, loss: 0.00013618904631584883
step: 60, loss: 0.00013988606224302202
step: 70, loss: 0.0002760797506198287
step: 80, loss: 0.0060876342467963696
step: 90, loss: 0.00038940884405747056
step: 100, loss: 0.00043649101280607283
step: 110, loss: 0.0034898535814136267
step: 120, loss: 0.003555585164576769
step: 130, loss: 0.0009791539050638676
step: 140, loss: 0.00100797577776111364
step: 150, loss: 0.000350981397787109
step: 160, loss: 0.06372598558664322
step: 170, loss: 0.0003570764674805105
step: 180, loss: 0.0008492153719998896
step: 190, loss: 0.00026281457394361496
step: 200, loss: 0.04901079460978508
step: 210, loss: 0.0001206037777592428
step: 220, loss: 0.0004461899516172707
step: 230, loss: 0.008297998458147049
/usr/local/lib/python3.7/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=0.9764309764309763,
best_f1=0.9763779527559054
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00023238058201996867
step: 10, loss: 0.001665815245360136
step: 20, loss: 0.00026459331274963915
step: 30, loss: 0.00032591939088888466
step: 40, loss: 0.00044514387263916433
step: 50, loss: 0.000549676304217428
step: 60, loss: 0.0022223670966923237
step: 70, loss: 0.0003440325381234288
step: 80, loss: 0.00026991937193088233
step: 90, loss: 7.222520071081817e-05
```

```
step: 100, loss: 0.0009156595333479345
step: 110, loss: 0.0030719900969415903
step: 120, loss: 0.03811157867312431
step: 130, loss: 8.653799159219489e-05
step: 140, loss: 9.464777394896373e-05
step: 150, loss: 0.0001756761921569705
step: 160, loss: 0.00014438756625168025
step: 170, loss: 0.000312784657580778
step: 180, loss: 0.012109913863241673
step: 190, loss: 0.0001434758014511317
step: 200, loss: 0.0001503960374975577
step: 210, loss: 0.0007158393273130059
step: 220, loss: 0.0005704754730686545
step: 230, loss: 0.00026445931871421635
/usr/local/lib/python3.7/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=0.9753363228699552,
best_f1=0.9763779527559054
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0008362606167793274
step: 10, loss: 0.0016649169847369194
step: 20, loss: 0.0005966824246570468
step: 30, loss: 0.0006528440862894058
step: 40, loss: 0.024640318006277084
step: 50, loss: 0.00020092409977223724
step: 60, loss: 0.00017221368034370244
step: 70, loss: 0.0001670478959567845
step: 80, loss: 0.00014663836918771267
step: 90, loss: 0.0003378555120434612
step: 100, loss: 0.0013914250303059816
step: 110, loss: 7.511823787353933e-05
step: 120, loss: 0.00010895268496824428
step: 130, loss: 7.2386501415167e-05
step: 140, loss: 5.552552829612978e-05
step: 150, loss: 0.00017667155771050602
step: 160, loss: 8.861959940986708e-05
step: 170, loss: 0.00026918540243059397
step: 180, loss: 0.00044155993964523077
step: 190, loss: 9.431586659047753e-05
Gradient overflow.  Skipping step, loss scaler 0 reducing loss scale to 32768.0
step: 200, loss: 0.00016058293113019317
step: 210, loss: 0.00010335926344851032
step: 220, loss: 0.00010536093759583309
step: 230, loss: 0.013789921067655087
/usr/local/lib/python3.7/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=0.978675645342312, f1=0.9742441209406495,
best_f1=0.9763779527559054
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0007878906326368451
step: 10, loss: 8.83090469869785e-05
step: 20, loss: 6.234750617295504e-05
step: 30, loss: 0.0011958021204918623
step: 40, loss: 7.362979522440583e-05
step: 50, loss: 6.489466613857076e-05
step: 60, loss: 7.23882403690368e-05
step: 70, loss: 0.00019106731633655727
step: 80, loss: 7.137422653613612e-05
step: 90, loss: 0.00015148910460993648
step: 100, loss: 0.014074564911425114
step: 110, loss: 0.0007045866805128753
step: 120, loss: 7.764371548546478e-05
step: 130, loss: 6.162939826026559e-05
step: 140, loss: 0.027337322011590004
step: 150, loss: 0.025783823803067207
step: 160, loss: 3.9142782043199986e-05
step: 170, loss: 6.842023867648095e-05
step: 180, loss: 0.00013015870354138315
step: 190, loss: 0.0001288753846893087
step: 200, loss: 5.312233042786829e-05
step: 210, loss: 0.0006089732050895691
step: 220, loss: 5.390239675762132e-05
step: 230, loss: 6.980499892961234e-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=0.9776286353467561, f1=0.9753914988814317,
best_f1=0.9763779527559054
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 5.5398577387677506e-05
step: 10, loss: 7.869541877880692e-05
step: 20, loss: 0.00015877805708441883
step: 30, loss: 0.00033459000405855477
step: 40, loss: 0.0017475336790084839
step: 50, loss: 0.00015894253738224506
step: 60, loss: 0.005205168854445219
step: 70, loss: 0.0002504417789168656
step: 80, loss: 0.00011664431804092601
```

```
step: 90, loss: 0.0002905378642026335
step: 100, loss: 6.943649350432679e-05
step: 110, loss: 0.00020801587379537523
step: 120, loss: 5.42692796443589e-05
step: 130, loss: 4.364908818388358e-05
step: 140, loss: 5.655962013406679e-05
step: 150, loss: 0.00043658987851813436
step: 160, loss: 6.683578249067068e-05
step: 170, loss: 9.093889821087942e-05
step: 180, loss: 0.0033208385575562716
step: 190, loss: 9.073258843272924e-05
step: 200, loss: 0.0005582260782830417
step: 210, loss: 4.552717655315064e-05
step: 220, loss: 0.0006226625991985202
step: 230, loss: 0.00910127442330122
/usr/local/lib/python3.7/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=0.9717514124293786, f1=0.9751693002257337,
best_f1=0.9763779527559054
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0002457438095007092
step: 10, loss: 5.17776352353394e-05
step: 20, loss: 7.084505341481417e-05
step: 30, loss: 0.004830458667129278
step: 40, loss: 0.0002801943337544799
step: 50, loss: 0.004522052593529224
step: 60, loss: 0.00012321160465944558
step: 70, loss: 0.005525513086467981
step: 80, loss: 0.0048625958152115345
step: 90, loss: 0.00036152268876321614
step: 100, loss: 3.8789203244959936e-05
step: 110, loss: 0.00040249587618745863
step: 120, loss: 0.0001275773101951927
step: 130, loss: 8.824106771498919e-05
step: 140, loss: 8.47189876367338e-05
step: 150, loss: 0.00010984705295413733
step: 160, loss: 0.00017382926307618618
step: 170, loss: 8.66501868586056e-05
step: 180, loss: 8.019598317332566e-05
step: 190, loss: 5.572689406108111e-05
step: 200, loss: 0.00010446326632518321
step: 210, loss: 0.00011751570855267346
step: 220, loss: 9.123125346377492e-05
step: 230, loss: 0.02634175308048725
/usr/local/lib/python3.7/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=0.9775280898876404, f1=0.9742441209406495,
best_f1=0.9763779527559054
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00012910740042570978
step: 10, loss: 0.0011707209050655365
step: 20, loss: 0.0004266592441126704
step: 30, loss: 0.00013416624278761446
step: 40, loss: 4.853735663346015e-05
step: 50, loss: 0.000996937626041472
step: 60, loss: 0.0005185264744795859
step: 70, loss: 0.00021283510432112962
step: 80, loss: 6.173613655846566e-05
step: 90, loss: 6.270000449148938e-05
step: 100, loss: 3.729177115019411e-05
step: 110, loss: 0.02927279658615589
step: 120, loss: 6.459283758886158e-05
step: 130, loss: 0.00018653178995009512
step: 140, loss: 0.0013927725376561284
step: 150, loss: 5.995689207338728e-05
step: 160, loss: 7.478212501155213e-05
step: 170, loss: 7.054676825646311e-05
step: 180, loss: 5.7773111620917916e-05
Gradient overflow.  Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 190, loss: 0.00023598611005581915
step: 200, loss: 0.00041462347144261
step: 210, loss: 4.5136424887459725e-05
step: 220, loss: 4.2343290260760114e-05
step: 230, loss: 5.2622464863816276e-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=0.978675645342312, f1=0.9754464285714286,
best_f1=0.9763779527559054
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0007401106413453817
step: 10, loss: 0.00040281793917529285
step: 20, loss: 0.00036387238651514053
step: 30, loss: 8.227104262914509e-05
step: 40, loss: 0.0005731173441745341
step: 50, loss: 0.00011780055501731113
step: 60, loss: 8.158345735864714e-05
step: 70, loss: 0.00010787024075398222
```

```
step: 80, loss: 0.0003807793837040663
step: 90, loss: 6.259629299165681e-05
step: 100, loss: 6.89519292791374e-05
step: 110, loss: 0.00016716925892978907
step: 120, loss: 3.920759991160594e-05
step: 130, loss: 4.344979242887348e-05
step: 140, loss: 0.0002902352425735444
step: 150, loss: 3.878925053868443e-05
step: 160, loss: 0.00010355365520808846
step: 170, loss: 2.0291332475608215e-05
step: 180, loss: 6.412035872926936e-05
step: 190, loss: 0.00011716688459273428
step: 200, loss: 3.468032446107827e-05
step: 210, loss: 0.0005392095190472901
step: 220, loss: 5.420640081865713e-05
step: 230, loss: 0.012526961974799633
/usr/local/lib/python3.7/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=0.9741863075196409, f1=0.9731543624161074,
best_f1=0.9763779527559054
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 2.499582478776574e-05
step: 10, loss: 2.8815929908887483e-05
step: 20, loss: 4.291670848033391e-05
step: 30, loss: 3.9769558497937396e-05
step: 40, loss: 5.100760972709395e-05
step: 50, loss: 0.01933087408542633
step: 60, loss: 8.253670966951177e-05
step: 70, loss: 6.868722994113341e-05
step: 80, loss: 4.4192867790116e-05
step: 90, loss: 0.0002602083550300449
step: 100, loss: 5.0481336074881256e-05
step: 110, loss: 3.2650859793648124e-05
step: 120, loss: 0.004527286160737276
step: 130, loss: 9.907638741424307e-05
step: 140, loss: 4.9842365115182474e-05
step: 150, loss: 0.00013426512305159122
step: 160, loss: 3.511955583235249e-05
step: 170, loss: 3.1354251404991373e-05
step: 180, loss: 0.0001659712434047833
step: 190, loss: 6.982959894230589e-05
step: 200, loss: 6.207804835867137e-05
step: 210, loss: 5.424936898634769e-05
step: 220, loss: 0.00020539491379167885
step: 230, loss: 3.9657203160459176e-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=0.9764837625979844, f1=0.9720670391061451,
best_f1=0.9763779527559054
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, 206.96it/s]
load_f1 = 0.9787234042553192
real_f1 = 0.9765363128491621
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, 237.96it/s]
```

1.3.2 DBLP-GoogleScholar - Running the matcher

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

```
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: 0.622779130935669
step: 10, loss: 0.5275973081588745
step: 20, loss: 0.48310959339141846
step: 30, loss: 0.11289907991886139
step: 40, loss: 0.19237300753593445
step: 50, loss: 0.1051773726940155
step: 60, loss: 0.038916174322366714
step: 70, loss: 0.1332879513502121
step: 80, loss: 0.0595669187605381
step: 90, loss: 0.16507588326931
step: 100, loss: 0.008498692885041237
step: 110, loss: 0.08463716506958008
step: 120, loss: 0.11474382132291794
step: 130, loss: 0.0652799978852272
step: 140, loss: 0.1260744035243988
step: 150, loss: 0.03715766966342926
step: 160, loss: 0.011899762786924839
step: 170, loss: 0.20283089578151703
step: 180, loss: 0.055387552827596664
step: 190, loss: 0.003894668072462082
step: 200, loss: 0.15793925523757935
step: 210, loss: 0.0726628303527832
step: 220, loss: 0.22228150069713593
step: 230, loss: 0.17135848104953766
step: 240, loss: 0.043372441083192825
step: 250, loss: 0.023197872564196587
```

```
step: 260, loss: 0.052153635770082474
step: 270, loss: 0.009614781476557255
step: 280, loss: 0.03135520592331886
step: 290, loss: 0.031743843108415604
step: 300, loss: 0.029772324487566948
step: 310, loss: 0.224617600440979
step: 320, loss: 0.12615352869033813
step: 330, loss: 0.0880129486322403
step: 340, loss: 0.06111243739724159
step: 350, loss: 0.03864394128322601
step: 360, loss: 0.09316810220479965
step: 370, loss: 0.06655983626842499
step: 380, loss: 0.003490251023322344
step: 390, loss: 0.16825971007347107
step: 400, loss: 0.2993960976600647
step: 410, loss: 0.07628637552261353
step: 420, loss: 0.08534898608922958
step: 430, loss: 0.15191909670829773
step: 440, loss: 0.039742544293403625
step: 450, loss: 0.0031632045283913612
step: 460, loss: 0.004211868159472942
step: 470, loss: 0.15681633353233337
step: 480, loss: 0.08089405298233032
step: 490, loss: 0.10740760713815689
step: 500, loss: 0.07589790970087051
step: 510, loss: 0.02808593586087227
step: 520, loss: 0.22000066936016083
step: 530, loss: 0.005513809621334076
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9378214118747078, f1=0.9333333333333335,
best_f1=0.9333333333333335
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.18094293773174286
step: 10, loss: 0.14935418963432312
step: 20, loss: 0.00870467722415924
step: 30, loss: 0.008358423598110676
step: 40, loss: 0.09689725190401077
step: 50, loss: 0.06440630555152893
step: 60, loss: 0.016148552298545837
step: 70, loss: 0.027182072401046753
```

```
step: 80, loss: 0.03061521053314209
step: 90, loss: 0.011298046447336674
step: 100, loss: 0.025286532938480377
step: 110, loss: 0.00429791584610939
step: 120, loss: 0.07362674921751022
step: 130, loss: 0.08713658154010773
step: 140, loss: 0.0268063023686409
step: 150, loss: 0.058482490479946136
step: 160, loss: 0.05283161997795105
step: 170, loss: 0.011720342561602592
step: 180, loss: 0.02963690087199211
step: 190, loss: 0.0834307074546814
step: 200, loss: 0.03853008523583412
step: 210, loss: 0.05145689845085144
step: 220, loss: 0.06053999438881874
step: 230, loss: 0.0098965959623456
step: 240, loss: 0.054112184792757034
step: 250, loss: 0.009877145290374756
step: 260, loss: 0.0021062230225652456
step: 270, loss: 0.11531636118888855
step: 280, loss: 0.02808830887079239
step: 290, loss: 0.025152157992124557
step: 300, loss: 0.18964961171150208
step: 310, loss: 0.007027726620435715
step: 320, loss: 0.09362034499645233
step: 330, loss: 0.07972729206085205
step: 340, loss: 0.02515587955713272
step: 350, loss: 0.0014690498355776072
step: 360, loss: 0.03319256007671356
step: 370, loss: 0.06648119539022446
step: 380, loss: 0.029195932671427727
step: 390, loss: 0.09204292297363281
step: 400, loss: 0.03838859498500824
step: 410, loss: 0.09918512403964996
step: 420, loss: 0.02702920511364937
step: 430, loss: 0.0912967249751091
step: 440, loss: 0.2084704339504242
step: 450, loss: 0.052516911178827286
step: 460, loss: 0.1318807750940323
step: 470, loss: 0.12025365233421326
step: 480, loss: 0.25759702920913696
step: 490, loss: 0.01609579659998417
step: 500, loss: 0.28244417905807495
step: 510, loss: 0.026727793738245964
step: 520, loss: 0.0678730383515358
step: 530, loss: 0.016048062592744827
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
```

```
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9387379087977891, f1=0.9426751592356688,
best_f1=0.9426751592356688
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0361579991877079
step: 10, loss: 0.03225671872496605
step: 20, loss: 0.027444962412118912
step: 30, loss: 0.08749599754810333
step: 40, loss: 0.009678586386144161
step: 50, loss: 0.09865362197160721
step: 60, loss: 0.0243461262434721
step: 70, loss: 0.009901993907988071
step: 80, loss: 0.021649440750479698
step: 90, loss: 0.008754130452871323
step: 100, loss: 0.08516976982355118
step: 110, loss: 0.023243263363838196
step: 120, loss: 0.01179206557571888
step: 130, loss: 0.006464681588113308
step: 140, loss: 0.024137670174241066
step: 150, loss: 0.02248375304043293
step: 160, loss: 0.007413546089082956
step: 170, loss: 0.1295415163040161
step: 180, loss: 0.05063636600971222
step: 190, loss: 0.003133955877274275
step: 200, loss: 0.06698116660118103
step: 210, loss: 0.04477537423372269
step: 220, loss: 0.10455337166786194
step: 230, loss: 0.10763264447450638
step: 240, loss: 0.0034286899026483297
step: 250, loss: 0.028674285858869553
step: 260, loss: 0.05083981156349182
step: 270, loss: 0.005372663494199514
step: 280, loss: 0.16593343019485474
step: 290, loss: 0.0031755685340613127
step: 300, loss: 0.04466870427131653
step: 310, loss: 0.01977984793484211
step: 320, loss: 0.061392322182655334
step: 330, loss: 0.004433652851730585
step: 340, loss: 0.042727500200271606
step: 350, loss: 0.010430168360471725
step: 360, loss: 0.042591679841279984
step: 370, loss: 0.033035825937986374
```

```
step: 380, loss: 0.11584734916687012
step: 390, loss: 0.011988498270511627
step: 400, loss: 0.023807061836123466
step: 410, loss: 0.00391494482755661
step: 420, loss: 0.1092425063252449
step: 430, loss: 0.04715125635266304
step: 440, loss: 0.03739115968346596
step: 450, loss: 0.06263525784015656
step: 460, loss: 0.04055683687329292
step: 470, loss: 0.09662601351737976
step: 480, loss: 0.005424105562269688
step: 490, loss: 0.004446505568921566
step: 500, loss: 0.0037889047525823116
step: 510, loss: 0.0052258591167628765
step: 520, loss: 0.04991353303194046
step: 530, loss: 0.010040398687124252
/usr/local/lib/python3.7/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=0.9385474860335196, f1=0.9419354838709678,
best_f1=0.9426751592356688
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.005723079200834036
step: 10, loss: 0.006814571563154459
step: 20, loss: 0.0014205977786332369
step: 30, loss: 0.006356928031891584
step: 40, loss: 0.016937127336859703
step: 50, loss: 0.07905756682157516
step: 60, loss: 0.0009415718959644437
step: 70, loss: 0.025147365406155586
step: 80, loss: 0.0044614337384700775
step: 90, loss: 0.054045334458351135
step: 100, loss: 0.06338650733232498
step: 110, loss: 0.03396662697196007
step: 120, loss: 0.00046988832764327526
step: 130, loss: 0.0033641948830336332
step: 140, loss: 0.010851985774934292
step: 150, loss: 0.03241880983114242<br>426
step: 160, loss: 0.006532840896397829
step: 170, loss: 0.0013073206646367908
step: 180, loss: 0.0021528161596506834
step: 190, loss: 0.04716425761580467
step: 200, loss: 0.016883935779333115
step: 210, loss: 0.03039388172328472
step: 220, loss: 0.012114390730857849
step: 230, loss: 0.04679577425122261
```

```
step: 240, loss: 0.003152471501380205
step: 250, loss: 0.03663688898086548
step: 260, loss: 0.008227696642279625
step: 270, loss: 0.006793361157178879
step: 280, loss: 0.10927943140268326
step: 290, loss: 0.09157305955886841
step: 300, loss: 0.00039416473009623587
step: 310, loss: 0.026689549908041954
step: 320, loss: 0.004467883612960577
step: 330, loss: 0.002528916345909238
step: 340, loss: 0.14500118792057037
step: 350, loss: 0.016260778531432152
step: 360, loss: 0.053769320249557495
step: 370, loss: 0.003101135604083538
step: 380, loss: 0.024801691994071007
step: 390, loss: 0.0007888007094152272
step: 400, loss: 0.09362640231847763
step: 410, loss: 0.020333178341388702
step: 420, loss: 0.003843069775030017
step: 430, loss: 0.0346059650182724
step: 440, loss: 0.0022273308131843805
step: 450, loss: 0.0026484401896595955
step: 460, loss: 0.0044701281003654
step: 470, loss: 0.014701714739203453
Gradient overflow.  Skipping step, loss scaler 0 reducing loss scale to 32768.0
step: 480, loss: 0.24645915627479553
step: 490, loss: 0.004608319140970707
step: 500, loss: 0.025165900588035583
step: 510, loss: 0.026782991364598274
step: 520, loss: 0.0914570540189743
step: 530, loss: 0.018058402463793755
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9410125406409661, f1=0.934622467771639,
best_f1=0.934622467771639
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.020834550261497498
step: 10, loss: 0.04787953570485115
step: 20, loss: 0.013070871122181416
step: 30, loss: 0.0011224907357245684
step: 40, loss: 0.0056644948199391365
```

```
step: 50, loss: 0.002371265785768628
step: 60, loss: 0.13789907097816467
step: 70, loss: 0.00698635820299387
step: 80, loss: 0.0011242344044148922
step: 90, loss: 0.07658139616250992
step: 100, loss: 0.017582625150680542
step: 110, loss: 0.0016375145642086864
step: 120, loss: 0.0038645442109555006
step: 130, loss: 0.0009484371403232217
step: 140, loss: 0.043578993529081345
step: 150, loss: 0.01558883860707283
step: 160, loss: 0.006042839027941227
step: 170, loss: 0.030664701014757156
step: 180, loss: 0.0023609264753758907
step: 190, loss: 0.0016433793352916837
step: 200, loss: 0.000805108982603997
step: 210, loss: 0.004941342864185572
step: 220, loss: 0.143914133310318
step: 230, loss: 0.019991092383861542
step: 240, loss: 0.0023381374776363373
step: 250, loss: 0.0037048368249088526
step: 260, loss: 0.0025966581888496876
step: 270, loss: 0.00011946802260354161
step: 280, loss: 0.001077710185199976
step: 290, loss: 0.14553874731063843
step: 300, loss: 0.011412191204726696
step: 310, loss: 0.002838225569576025
step: 320, loss: 0.05450898036360741
step: 330, loss: 0.04246392101049423
step: 340, loss: 0.007528673857450485
step: 350, loss: 0.029150232672691345
step: 360, loss: 0.0008545289165340364
step: 370, loss: 0.04179603233933449
step: 380, loss: 7.417218876071274e-05
step: 390, loss: 0.00038696586852893233
step: 400, loss: 0.02549278922379017
step: 410, loss: 0.01135576143860817
step: 420, loss: 0.0012358188396319747
step: 430, loss: 0.00297615141607821
step: 440, loss: 0.0055243829265236855
step: 450, loss: 0.1571044921875
step: 460, loss: 0.007279423996806145
step: 470, loss: 0.010462114587426186
step: 480, loss: 0.0030253021977841854
step: 490, loss: 0.0010076418984681368
step: 500, loss: 0.03595242649316788
step: 510, loss: 0.006472621578723192
step: 520, loss: 0.02957121655344963
```

```
step: 530, loss: 0.005758689250797033
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9411764705882353, f1=0.9410150891632373,
best_f1=0.9410150891632373
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00832322333008051
step: 10, loss: 0.0002250179386464879
step: 20, loss: 0.11102385073900223
step: 30, loss: 0.0011013925541192293
step: 40, loss: 0.0007326924824155867
step: 50, loss: 0.013825629837810993
step: 60, loss: 0.0005042552947998047
step: 70, loss: 0.00089591683335422873
step: 80, loss: 0.027659688144922256
step: 90, loss: 0.002166209276765585
step: 100, loss: 0.014434569515287876
step: 110, loss: 0.011797169223427773
step: 120, loss: 0.028645873069763184
step: 130, loss: 0.01226021721959114
step: 140, loss: 0.009101735427975655
step: 150, loss: 0.022244714200496674
step: 160, loss: 0.0018502939492464066
step: 170, loss: 0.0008387524867430329
step: 180, loss: 0.028505006805062294
step: 190, loss: 0.007003356236964464
Gradient overflow.  Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 200, loss: 0.0010064449161291122
step: 210, loss: 0.005690797232091427
step: 220, loss: 0.0024439250119030476
step: 230, loss: 0.00036989463842473924
step: 240, loss: 0.009634379297494888
step: 250, loss: 0.011921407654881477
step: 260, loss: 0.004600135143846273
step: 270, loss: 0.11845269799232483
step: 280, loss: 0.005335036665201187
step: 290, loss: 0.0007967686397023499
step: 300, loss: 0.002665080362930894
step: 310, loss: 0.009424257092177868
step: 320, loss: 0.00033285870449617505
step: 330, loss: 0.01204647496342659
```

```
step: 340, loss: 0.04133171588182449
step: 350, loss: 0.0053005702793598175
step: 360, loss: 0.006976163014769554
step: 370, loss: 0.0018751011230051517
step: 380, loss: 0.0020251476671546698
step: 390, loss: 0.015171248465776443
step: 400, loss: 0.00037978967884555546
step: 410, loss: 0.0005826774286106229
step: 420, loss: 0.012433603405952454
step: 430, loss: 0.0008864053525030613
step: 440, loss: 0.0009138880996033549
step: 450, loss: 0.0016529789427295327
step: 460, loss: 0.001130920136347413
step: 470, loss: 0.0034428881481289864
step: 480, loss: 0.0115025220438838
step: 490, loss: 0.01174985896795988
step: 500, loss: 0.01838747411966324
step: 510, loss: 0.01129350159317255
step: 520, loss: 0.01109018549323082
step: 530, loss: 0.11421897262334824
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9433255269320844, f1=0.9392675011590171,
best_f1=0.9392675011590171
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0001342212490271777
step: 10, loss: 0.00031746807508170605
step: 20, loss: 0.0011354265734553337
step: 30, loss: 0.011728874407708645
step: 40, loss: 0.005379855632781982
step: 50, loss: 0.009842222556471825
step: 60, loss: 0.04410979524254799
step: 70, loss: 0.00044789540697820485
step: 80, loss: 0.0015502457972615957
step: 90, loss: 0.0009636356262490153
step: 100, loss: 0.00012041611626045778
step: 110, loss: 0.0002761341747827828
step: 120, loss: 0.0004584275302477181
step: 130, loss: 0.01499202474951744
step: 140, loss: 0.00014824143727310002
step: 150, loss: 0.08500062674283981
```

```
step: 160, loss: 9.179919288726524e-05
step: 170, loss: 0.0004885879461653531
step: 180, loss: 0.004475985653698444
step: 190, loss: 0.00046471721725538373
step: 200, loss: 0.0002741771168075502
step: 210, loss: 0.0005286402883939445
step: 220, loss: 0.00041993040940724313
step: 230, loss: 0.0040346127934753895
step: 240, loss: 0.0005395518965087831
step: 250, loss: 0.0004765736812260002
step: 260, loss: 0.06524863839149475
step: 270, loss: 0.0005171767552383244
step: 280, loss: 0.0003921081079170108
step: 290, loss: 0.0012355224462226033
step: 300, loss: 0.003240796271711588
step: 310, loss: 0.0007008447428233922
step: 320, loss: 0.004242428578436375
step: 330, loss: 0.0005130689824000001
step: 340, loss: 0.009178945794701576
step: 350, loss: 0.05532969906926155
step: 360, loss: 0.0031671328470110893
step: 370, loss: 0.0018380949040874839
step: 380, loss: 0.0005007098661735654
step: 390, loss: 0.001222051098011434
step: 400, loss: 0.00829986110329628
step: 410, loss: 0.01696852408349514
step: 420, loss: 0.03321333974599838
step: 430, loss: 0.00048268030514009297
step: 440, loss: 0.002802795497700572
step: 450, loss: 0.0010657947277650237
step: 460, loss: 0.0011331530986353755
step: 470, loss: 0.005340824835002422
step: 480, loss: 0.09535203874111176
step: 490, loss: 0.00789873767644167
step: 500, loss: 0.06803953647613525
step: 510, loss: 0.043735720217227936
step: 520, loss: 0.08010351657867432
step: 530, loss: 0.0026987381279468536
/usr/local/lib/python3.7/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=0.9400187441424556, f1=0.9331476323119777,
best_f1=0.9392675011590171
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0008563774754293263
step: 10, loss: 0.0625843033194542
```

```
step: 20, loss: 0.002793934429064393
step: 30, loss: 0.017295174300670624
step: 40, loss: 0.0001372547703795135
step: 50, loss: 0.005169880110770464
step: 60, loss: 0.0013807631330564618
step: 70, loss: 0.0002995500690303743
step: 80, loss: 0.008668330498039722
step: 90, loss: 0.0007276290562003851
step: 100, loss: 0.0015268104616552591
step: 110, loss: 0.000639788806438446
step: 120, loss: 0.03314316272735596
step: 130, loss: 0.0015030507929623127
step: 140, loss: 0.004256189800798893
step: 150, loss: 0.008202278055250645
step: 160, loss: 0.028438687324523926
step: 170, loss: 0.013502940535545349
step: 180, loss: 0.002867850475013256
step: 190, loss: 0.0045578982681035995
step: 200, loss: 0.0009899245342239738
step: 210, loss: 0.0054556201212108135
step: 220, loss: 0.0027009134646505117
step: 230, loss: 0.0011586463078856468
step: 240, loss: 0.0014303330099210143
step: 250, loss: 7.496252510463819e-05
step: 260, loss: 9.480041626375169e-05
step: 270, loss: 0.010689113289117813
step: 280, loss: 0.0023824807722121477
step: 290, loss: 0.0002955622912850231
step: 300, loss: 0.0020285933278501034
step: 310, loss: 0.0003316628572065383
step: 320, loss: 0.015260877087712288
step: 330, loss: 0.00021366179862525314
step: 340, loss: 6.502788892248645e-05
step: 350, loss: 0.0011512167984619737
step: 360, loss: 0.00038860339554958045
step: 370, loss: 0.0013624678831547499
step: 380, loss: 0.022146038711071014
step: 390, loss: 0.15942390263080597
step: 400, loss: 0.0010818380396813154
step: 410, loss: 0.00041573235648684204
step: 420, loss: 0.0018302724929526448
step: 430, loss: 0.022652270272374153
step: 440, loss: 0.0024638460017740726
step: 450, loss: 0.0006740656681358814
step: 460, loss: 0.00029693098622374237
step: 470, loss: 0.00013094696623738855
step: 480, loss: 7.174410711741075e-05
step: 490, loss: 0.0038866540417075157
```

```
step: 500, loss: 0.0012696085032075644
step: 510, loss: 0.0002260289911646396
step: 520, loss: 0.01862211897969246
step: 530, loss: 0.0013195137726143003
/usr/local/lib/python3.7/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=0.9406307977736549, f1=0.9335776454420522,
best_f1=0.9392675011590171
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0015997110167518258
step: 10, loss: 0.000868145958520472
step: 20, loss: 0.0018190309638157487
step: 30, loss: 0.00033038476249203086
step: 40, loss: 0.0006976814474910498
step: 50, loss: 0.00020595520618371665
step: 60, loss: 0.0001481154322391376
step: 70, loss: 0.000216293876292184
step: 80, loss: 0.023171598091721535
step: 90, loss: 6.062351167201996e-05
step: 100, loss: 0.00013977254275232553
step: 110, loss: 4.5754793973173946e-05
step: 120, loss: 0.0010923986556008458
step: 130, loss: 0.0338941253721714
step: 140, loss: 0.00039232574636116624
step: 150, loss: 8.010405872482806e-05
step: 160, loss: 0.003921481315046549
step: 170, loss: 0.016700655221939087
step: 180, loss: 0.0005737433093599975
step: 190, loss: 0.00014710715913679451
step: 200, loss: 0.0011603519087657332
step: 210, loss: 0.0001817496377043426
step: 220, loss: 0.0004664068401325494
step: 230, loss: 0.0009061549790203571
step: 240, loss: 4.928962152916938e-05
step: 250, loss: 0.000927539193071425
step: 260, loss: 0.0028650495223701
step: 270, loss: 0.0009960492607206106
step: 280, loss: 0.013495701365172863
step: 290, loss: 0.04100016877055168
step: 300, loss: 4.7243531298590824e-05
step: 310, loss: 0.013239904306828976
step: 320, loss: 0.003991695586591959
step: 330, loss: 0.006420547142624855
step: 340, loss: 0.010385410860180855
step: 350, loss: 0.008068466559052467
```

```
step: 360, loss: 5.410626545199193e-05
step: 370, loss: 0.0010716163087636232
step: 380, loss: 0.0005719356122426689
step: 390, loss: 0.00011323072976665571
step: 400, loss: 0.0002403412654530257
step: 410, loss: 0.00365914567373693
step: 420, loss: 9.129913814831525e-05
step: 430, loss: 5.5018477723933756e-05
step: 440, loss: 0.0008348796400241554
step: 450, loss: 0.00056644924916327
step: 460, loss: 0.010445189662277699
step: 470, loss: 0.016230162233114243
step: 480, loss: 8.706992957741022e-05
step: 490, loss: 0.001867602113634348
step: 500, loss: 0.0026549948379397392
step: 510, loss: 0.004787901416420937
step: 520, loss: 0.0001503763924119994
step: 530, loss: 0.00013657924137078226
/usr/local/lib/python3.7/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=0.9412317818523741, f1=0.9389277389277391,
best_f1=0.9392675011590171
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 4.927481495542452e-05
step: 10, loss: 9.109747770708054e-05
step: 20, loss: 4.1959912778111175e-05
step: 30, loss: 2.5822575480560772e-05
step: 40, loss: 0.0002267227100674063
step: 50, loss: 0.0021396772935986652
step: 60, loss: 0.00017761881463229656
step: 70, loss: 9.024712926475331e-05
step: 80, loss: 0.00010889465920627117
step: 90, loss: 0.00032467208802700043
step: 100, loss: 0.0008203868637792766
step: 110, loss: 0.0005806826520711184
step: 120, loss: 0.0002882083354052156
step: 130, loss: 0.007844218984246254
step: 140, loss: 0.0034575669560581446
step: 150, loss: 0.0029785365331918
step: 160, loss: 0.0009911153465509415
step: 170, loss: 0.0002965344174299389
step: 180, loss: 0.0007771109230816364
step: 190, loss: 0.004488188307732344
step: 200, loss: 0.0010071403812617064
step: 210, loss: 0.00929993949830532
```

```
step: 220, loss: 0.013491489924490452
step: 230, loss: 0.0012525059282779694
step: 240, loss: 6.734677299391478e-05
step: 250, loss: 3.0743532988708466e-05
step: 260, loss: 0.0013644042192026973
step: 270, loss: 0.001009313273243606
step: 280, loss: 0.03626009076833725
step: 290, loss: 0.00016382870671804994
step: 300, loss: 0.0005436898791231215
step: 310, loss: 0.0005632793181575835
step: 320, loss: 0.0010812385007739067
step: 330, loss: 0.0005079521215520799
step: 340, loss: 0.0026809812989085913
step: 350, loss: 0.0016545187681913376
step: 360, loss: 9.436633990844712e-05
step: 370, loss: 0.007953505031764507
step: 380, loss: 0.003949563018977642
step: 390, loss: 5.9560949011938646e-05
step: 400, loss: 0.0015630333218723536
step: 410, loss: 0.0016870005056262016
step: 420, loss: 9.596301970304921e-05
step: 430, loss: 0.0001365634088870138
step: 440, loss: 0.006304907146841288
step: 450, loss: 7.703074516030028e-05
step: 460, loss: 0.0005286912783049047
step: 470, loss: 9.029858483700082e-05
step: 480, loss: 9.106930519919842e-05
step: 490, loss: 0.004508693236857653
step: 500, loss: 0.0133954593911767
step: 510, loss: 0.0003142130735795945
step: 520, loss: 0.002024455927312374
step: 530, loss: 0.002631843788549304
/usr/local/lib/python3.7/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=0.9406858202038925, f1=0.9368709972552608,
best_f1=0.9392675011590171
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00012154696014476947694
step: 10, loss: 0.029168393462896347
step: 20, loss: 0.0007585682906210423
step: 30, loss: 0.0016405177302658558
step: 40, loss: 0.00027458323165774345
step: 50, loss: 0.0031648033764213324
step: 60, loss: 0.018830761313438416
step: 70, loss: 0.005470042582601309
```

```
step: 80, loss: 3.20675790135283e-05
step: 90, loss: 0.00033465068554505706
step: 100, loss: 0.0006467649363912642
step: 110, loss: 0.00035670623765327036
step: 120, loss: 0.00018244137754663825
step: 130, loss: 0.0008016828796826303
step: 140, loss: 0.18271946907043457
step: 150, loss: 6.781150295864791e-05
step: 160, loss: 0.0007932519656606019
step: 170, loss: 0.0008174965041689575
step: 180, loss: 0.00024337437935173512
step: 190, loss: 0.002976955147460103
step: 200, loss: 0.0017875289777293801
step: 210, loss: 0.0005620368174277246
step: 220, loss: 0.003344695083796978
step: 230, loss: 0.0006880986620672047
step: 240, loss: 0.001100849243812263
step: 250, loss: 0.00021837638632860035
step: 260, loss: 0.0003866591432597488
step: 270, loss: 0.046796105802059174
step: 280, loss: 0.00044028557022102177
step: 290, loss: 0.0010756237898021936
step: 300, loss: 0.00034439796581864357
step: 310, loss: 0.00022764176537748426
step: 320, loss: 0.004883880261331797
step: 330, loss: 7.042507786536589e-05
step: 340, loss: 0.0013342826860025525
step: 350, loss: 0.0018748160218819976
step: 360, loss: 7.980683585628867e-05
step: 370, loss: 0.00167552859056741
step: 380, loss: 8.489465108141303e-05
step: 390, loss: 0.0001303347380599007
step: 400, loss: 0.0016080030472949147
step: 410, loss: 0.002557202475145459
step: 420, loss: 0.0005566520849242806
step: 430, loss: 0.001225977553986013
step: 440, loss: 0.0035787425003945827
step: 450, loss: 0.0068101584911346436
step: 460, loss: 0.0009266782435588539
step: 470, loss: 5.2533290727296844e-05
step: 480, loss: 0.0015076438430696726
step: 490, loss: 0.0030888670589774847
step: 500, loss: 0.022154338657855988
step: 510, loss: 0.0015220558270812035
step: 520, loss: 4.823501149076037e-05
step: 530, loss: 0.0004016751772724092
/usr/local/lib/python3.7/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=0.941342092914125, f1=0.9435185185185185,
best_f1=0.9392675011590171
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0018651513382792473
step: 10, loss: 0.002561309840530157
step: 20, loss: 0.002228444442152977
step: 30, loss: 0.0020766945090144873
step: 40, loss: 0.0007200376130640507
Gradient overflow.  Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 50, loss: 0.06811836361885071
step: 60, loss: 0.0015930033987388015
step: 70, loss: 0.00025940409977920353
step: 80, loss: 0.0030882172286510468
step: 90, loss: 0.010181318037211895
step: 100, loss: 0.008751255460083485
step: 110, loss: 6.350456533255056e-05
step: 120, loss: 4.775252455146983e-05
step: 130, loss: 0.0007037102477625012
step: 140, loss: 3.293683039373718e-05
step: 150, loss: 0.00013384883641265333
step: 160, loss: 0.00014484536950476468
step: 170, loss: 0.00019241841800976545
step: 180, loss: 0.00011892079783137888
step: 190, loss: 0.0003741414111573249
step: 200, loss: 8.839865768095478e-05
step: 210, loss: 4.194476787233725e-05
step: 220, loss: 0.000989115796983242
step: 230, loss: 8.287120726890862e-05
step: 240, loss: 0.0006556081352755427
step: 250, loss: 0.059531498700380325
step: 260, loss: 0.00027602355112321675
step: 270, loss: 0.0001556927163619548
step: 280, loss: 0.0009453105740249157
step: 290, loss: 0.0003642928204499185
step: 300, loss: 0.0009771184995770454
step: 310, loss: 0.00020755089644808322
step: 320, loss: 0.0009433940867893398
step: 330, loss: 0.0024071051739156246
step: 340, loss: 0.00016813245019875467
step: 350, loss: 0.001709298579953611
step: 360, loss: 0.0016937711043283343
step: 370, loss: 0.0029140703845769167
step: 380, loss: 0.06329704076051712
step: 390, loss: 0.0008988708723336458
step: 400, loss: 0.0014907743316143751
```

```
step: 410, loss: 0.0726521760225296
step: 420, loss: 0.004923081956803799
step: 430, loss: 0.003046793630346656
step: 440, loss: 0.0009939811425283551
step: 450, loss: 0.005933655425906181
step: 460, loss: 0.00010019337787525728
step: 470, loss: 6.983901403145865e-05
step: 480, loss: 0.0012111928081139922
step: 490, loss: 0.003750522155314684
step: 500, loss: 0.004867436829954386
step: 510, loss: 0.0006861689034849405
step: 520, loss: 1.5537878425675444e-05
step: 530, loss: 0.0013195336796343327
/usr/local/lib/python3.7/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=0.9396350023397285, f1=0.9405255878284925,
best_f1=0.9392675011590171
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.003838823875412345
step: 10, loss: 7.176691724453121e-05
step: 20, loss: 0.0008526674937456846
step: 30, loss: 0.0015145298093557358
step: 40, loss: 0.00026033478206954896
step: 50, loss: 0.004323444329202175
step: 60, loss: 2.9018587156315334e-05
step: 70, loss: 9.625737584428862e-05
step: 80, loss: 0.0012570281978696585
step: 90, loss: 0.010454618372023106
step: 100, loss: 0.0022058538161218166
step: 110, loss: 0.0007901926292106509
step: 120, loss: 0.0006491644307971001
step: 130, loss: 6.45088657620363e-05
step: 140, loss: 0.00021269066201057285
step: 150, loss: 0.015918513759970665
step: 160, loss: 4.58962022094056e-05
step: 170, loss: 0.0017010237788781524
step: 180, loss: 5.5436656111851335e-05
step: 190, loss: 0.00018924707546830177
step: 200, loss: 2.5167057174257934e-05
step: 210, loss: 5.871818939340301e-05
step: 220, loss: 0.005149298347532749
step: 230, loss: 0.00011543519940460101
step: 240, loss: 0.00020691142708528787
step: 250, loss: 3.309111707494594e-05
step: 260, loss: 0.00020381588547024876
```

```
step: 270, loss: 3.564605503925122e-05
step: 280, loss: 0.0001205094886245206
step: 290, loss: 0.0003108508826699108
step: 300, loss: 0.0006191869615577161
step: 310, loss: 0.001213090610690415
step: 320, loss: 0.002493141684681177
step: 330, loss: 0.0012553863925859332
step: 340, loss: 3.387543256394565e-05
step: 350, loss: 6.274668703554198e-05
step: 360, loss: 0.00013384345220401883
step: 370, loss: 0.0003714935446623713
step: 380, loss: 0.0007322326418943703
step: 390, loss: 0.0019354979740455747
step: 400, loss: 0.0007266262546181679
step: 410, loss: 0.0007851922418922186
step: 420, loss: 2.1117339201737195e-05
step: 430, loss: 0.022443123161792755
step: 440, loss: 3.0518811399815604e-05
step: 450, loss: 6.369245966197923e-05
step: 460, loss: 0.0011984427692368627
step: 470, loss: 0.0006780810072086751
step: 480, loss: 8.46593757160008e-05
step: 490, loss: 0.0017250519013032317
step: 500, loss: 0.0001926084514707327
step: 510, loss: 0.0002704447542782873
step: 520, loss: 0.00031163598760031164
step: 530, loss: 0.00017574468802195042
/usr/local/lib/python3.7/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=0.9408450704225352, f1=0.937471051412691,
best_f1=0.9392675011590171
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0010421063052490354
step: 10, loss: 4.4772201363230124e-05
step: 20, loss: 2.967572618217673e-05
step: 30, loss: 6.058184226276353e-05
step: 40, loss: 0.004144955426454544
step: 50, loss: 0.040789343416690826
step: 60, loss: 1.09783559310001715e-05
step: 70, loss: 0.0013647370506078005
step: 80, loss: 1.766486093401909e-05
step: 90, loss: 0.00036535211256705225
step: 100, loss: 7.902994548203424e-05
step: 110, loss: 0.0005997378029860556
step: 120, loss: 8.052241173572838e-05
```

```
step: 130, loss: 0.052251484245061874
step: 140, loss: 0.000640196492895484
step: 150, loss: 0.00021749800362158567
step: 160, loss: 0.0006160784396342933
step: 170, loss: 0.0008998675039038062
step: 180, loss: 0.03009360283613205
step: 190, loss: 6.108395609771833e-05
step: 200, loss: 7.960617949720472e-05
step: 210, loss: 0.0009546884102746844
step: 220, loss: 2.8647044018725865e-05
step: 230, loss: 0.018918389454483986
step: 240, loss: 0.001386389252729714
step: 250, loss: 2.2115973479230888e-05
step: 260, loss: 0.0007237168028950691
step: 270, loss: 0.00016302673611789942
step: 280, loss: 1.4595488210034091e-05
step: 290, loss: 0.03996060788631439
step: 300, loss: 0.00017860262596514076
step: 310, loss: 0.0002834591141436249
step: 320, loss: 0.0008629246149212122
step: 330, loss: 2.724100158957299e-05
step: 340, loss: 0.0008674767450429499
step: 350, loss: 0.0006546132499352098
step: 360, loss: 0.0014091384364292026
step: 370, loss: 0.0009226099355146289
step: 380, loss: 0.0018964458722621202
step: 390, loss: 0.004960492718964815
step: 400, loss: 0.00025596650084480643
step: 410, loss: 0.04681520164012909
step: 420, loss: 0.000247212708927691
step: 430, loss: 2.3613363737240434e-05
step: 440, loss: 0.02239587903022766
step: 450, loss: 0.0003202205407433212
step: 460, loss: 0.000408169929869473
step: 470, loss: 0.003463394707068801
step: 480, loss: 0.00011280316539341584
step: 490, loss: 0.00017435076006222516
step: 500, loss: 0.004999030847102404
step: 510, loss: 0.0031225045677274466
step: 520, loss: 0.0009488591458648443
step: 530, loss: 0.01908832974731922
/usr/local/lib/python3.7/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=0.941340782122905, f1=0.9395604395604396,
best_f1=0.9392675011590171
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
```

```
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 2.2112088117864914e-05
step: 10, loss: 0.0004121359088458121
step: 20, loss: 9.569540998199955e-05
step: 30, loss: 0.00011042547703254968
step: 40, loss: 0.15181785821914673
step: 50, loss: 0.00023662220337428153
step: 60, loss: 1.7355798263452016e-05
step: 70, loss: 0.008270013146102428
step: 80, loss: 3.0166747819748707e-05
step: 90, loss: 7.827100489521399e-05
step: 100, loss: 1.7240277884411626e-05
step: 110, loss: 0.0010906566167250276
step: 120, loss: 0.0017048262525349855
step: 130, loss: 0.1091959998011589
step: 140, loss: 0.0021015533711761236
step: 150, loss: 0.00011521351552801207
step: 160, loss: 7.76361339376308e-05
step: 170, loss: 0.0006015924154780805
step: 180, loss: 0.00011590989015530795
step: 190, loss: 2.935660995717626e-05
step: 200, loss: 1.9743278244277462e-05
step: 210, loss: 0.000628934009000659
step: 220, loss: 4.122447353438474e-05
step: 230, loss: 0.00019705195154529065
step: 240, loss: 8.335785241797566e-05
step: 250, loss: 2.4124225092236884e-05
step: 260, loss: 1.4949244359740987e-05
step: 270, loss: 0.00023426106781698763
step: 280, loss: 2.0361354472697712e-05
step: 290, loss: 1.0337615094613284e-05
step: 300, loss: 1.1391835869289935e-05
step: 310, loss: 0.00040664258995093405
step: 320, loss: 0.00013583623513113707
step: 330, loss: 0.0003584832011256367
step: 340, loss: 5.1624159823404625e-05
step: 350, loss: 4.455959060578607e-05
step: 360, loss: 0.0012899921275675297
step: 370, loss: 0.01417544949799776
step: 380, loss: 2.33302271226421e-05
step: 390, loss: 0.0001632719358894974
step: 400, loss: 8.378003258258104e-05
step: 410, loss: 8.31072247819975e-05
step: 420, loss: 4.535147309070453e-05
step: 430, loss: 7.582529360661283e-05
step: 440, loss: 0.05618361011147499
step: 450, loss: 0.0005280339391902089
step: 460, loss: 0.002630854258313775
```

```
step: 470, loss: 0.0015149656683206558
step: 480, loss: 1.1280097169219516e-05
step: 490, loss: 0.0013682199642062187
step: 500, loss: 0.0020652993116527796
step: 510, loss: 0.0002326446119695902
step: 520, loss: 1.5169138350756839e-05
step: 530, loss: 0.000658186327200383
/usr/local/lib/python3.7/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=0.9404096834264433, f1=0.9386446886446886,
best_f1=0.9392675011590171
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, 253.88it/s]
load_f1 = 0.9421028253821214
real_f1 = 0.9418874941887495
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, 256.22it/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: 0.5628355145454407
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.41379310344827586, f1=0.3373493975903615,
best_f1=0.3373493975903615
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.4936589002609253
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
epoch 2: dev_f1=0.2666666666666667, f1=0.2666666666666667,
best_f1=0.3373493975903615
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.

```

```
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.5533716678619385
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.5128205128205129, f1=0.3508771929824562,
best_f1=0.3508771929824562
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.3035537898540497
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.6086956521739131, f1=0.47058823529411764,
best_f1=0.47058823529411764
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.26104632019996643
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.6923076923076924, f1=0.5625000000000001,
best_f1=0.5625000000000001
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.175650492310524
/usr/local/lib/python3.7/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=0.6666666666666666, f1=0.64, best_f1=0.5625000000000001
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
```

```
step: 0, loss: 0.05452254042029381
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.7333333333333334, f1=0.5806451612903226,
best_f1=0.5806451612903226
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.008291578851640224
Gradient overflow.  Skipping step, loss scaler 0 reducing loss scale to 16384.0
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
  warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 8: dev_f1=0.742857142857143, f1=0.6153846153846153,
best_f1=0.6153846153846153
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.006321432068943977
/usr/local/lib/python3.7/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=0.6956521739130435, f1=0.5833333333333334,
best_f1=0.6153846153846153
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.004147296771407127
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.7692307692307692, f1=0.6875000000000001,
best_f1=0.6875000000000001
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
```

```
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.004414299037307501
/usr/local/lib/python3.7/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=0.6875000000000001,
best_f1=0.6875000000000001
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.003699528519064188
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.7857142857142857, f1=0.6875000000000001,
best_f1=0.6875000000000001
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.003985329996794462
/usr/local/lib/python3.7/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=0.7857142857142857, f1=0.6875000000000001,
best_f1=0.6875000000000001
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0020675500854849815
/usr/local/lib/python3.7/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=0.6875000000000001,
best_f1=0.6875000000000001
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.003553619608283043
/usr/local/lib/python3.7/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=0.6875000000000001,
best_f1=0.6875000000000001
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, 89240.51it/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 = 0.7586206896551724
real_f1 = 0.7586206896551724
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, 196.98it/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: 0.6287398338317871
step: 10, loss: 0.6170055866241455
step: 20, loss: 0.3245846927165985
step: 30, loss: 0.07589269429445267
step: 40, loss: 0.1162479966878891
step: 50, loss: 0.011368243023753166
step: 60, loss: 0.10020691156387329
step: 70, loss: 0.025660181418061256
step: 80, loss: 0.07038165628910065
step: 90, loss: 0.0632835403084755
step: 100, loss: 0.019798623397946358
step: 110, loss: 0.04488567262887955
step: 120, loss: 0.008560716174542904
step: 130, loss: 0.0161663256585598
step: 140, loss: 0.00861994456499815
step: 150, loss: 0.016185395419597626
step: 160, loss: 0.00800022017210722
step: 170, loss: 0.0464814156293869
step: 180, loss: 0.003243489423766732
step: 190, loss: 0.04911535233259201
step: 200, loss: 0.014451426453888416
Gradient overflow.  Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 210, loss: 0.004651886411011219
step: 220, loss: 0.0020774726290255785
step: 230, loss: 0.04399741813540459
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9865470852017937, f1=0.9841986455981941,
best_f1=0.9841986455981941
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.004976768046617508
```

```
step: 10, loss: 0.0036004697903990746
step: 20, loss: 0.21437394618988037
step: 30, loss: 0.07785701751708984
step: 40, loss: 0.0890127569437027
step: 50, loss: 0.0028668339364230633
step: 60, loss: 0.0015551658580079675
step: 70, loss: 0.1238347515463829
step: 80, loss: 0.0029887952841818333
step: 90, loss: 0.009945257566869259
step: 100, loss: 0.006151179783046246
step: 110, loss: 0.046558160334825516
step: 120, loss: 0.06838656216859818
step: 130, loss: 0.09017083048820496
step: 140, loss: 0.03170182555913925
step: 150, loss: 0.006148990243673325
step: 160, loss: 0.08399397134780884
step: 170, loss: 0.0021697520278394222
step: 180, loss: 0.0056805978529155254
step: 190, loss: 0.017505940049886703
step: 200, loss: 0.0010932416189461946
step: 210, loss: 0.00083641690434888
step: 220, loss: 0.10053575783967972
step: 230, loss: 0.005139052867889404
/usr/local/lib/python3.7/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=0.9831649831649831, f1=0.9852774631936579,
best_f1=0.9841986455981941
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0060157207772135735
step: 10, loss: 0.0013044532388448715
step: 20, loss: 0.003027952741831541
step: 30, loss: 0.08999254554510117
step: 40, loss: 0.07831455022096634
step: 50, loss: 0.013850501738488674
step: 60, loss: 0.008035908453166485
step: 70, loss: 0.007840143516659737
step: 80, loss: 0.0008278308669105172
step: 90, loss: 0.013708259910345078
step: 100, loss: 0.0019228693563491106
step: 110, loss: 0.0007019576733000576
step: 120, loss: 0.0032853700686246157
step: 130, loss: 0.0003793183423113078
step: 140, loss: 0.004102237522602081
step: 150, loss: 0.09882917255163193
step: 160, loss: 0.0015930233057588339
```

```
step: 170, loss: 0.009857628494501114
step: 180, loss: 0.008142880164086819
step: 190, loss: 0.0011559498962014914
step: 200, loss: 0.004751351661980152
step: 210, loss: 0.003968722186982632
step: 220, loss: 0.0010248887119814754
step: 230, loss: 0.0004605548456311226
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9898989898989898, f1=0.9842342342342343,
best_f1=0.9842342342342343
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00027793904882855713
step: 10, loss: 0.00023180263815447688
step: 20, loss: 0.0005135654355399311
step: 30, loss: 0.00017145216406788677
step: 40, loss: 0.09357640892267227
step: 50, loss: 0.010353020392358303
step: 60, loss: 0.007035946473479271
step: 70, loss: 0.0011747792595997453
step: 80, loss: 0.0009744453709572554
step: 90, loss: 0.0017496077343821526
step: 100, loss: 0.0002995044633280486
step: 110, loss: 0.00043126812670379877
step: 120, loss: 0.016616234555840492
step: 130, loss: 0.0020567025057971478
step: 140, loss: 0.00029266325873322785
Gradient overflow.  Skipping step, loss scaler 0 reducing loss scale to 8192.0
step: 150, loss: 0.046635936945676804
step: 160, loss: 0.00785272940993309
step: 170, loss: 0.1170072928071022
step: 180, loss: 0.0004029073170386255
step: 190, loss: 0.007480394095182419
step: 200, loss: 0.002417305950075388
step: 210, loss: 0.01250892598181963
step: 220, loss: 0.00020811246940866113
step: 230, loss: 0.009228890761733055
/usr/local/lib/python3.7/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=0.9807474518686297,
```

```
best_f1=0.9842342342342343
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0007268866756930947
step: 10, loss: 0.0007530837319791317
step: 20, loss: 0.0006472731474786997
step: 30, loss: 0.0004614882927853614
step: 40, loss: 0.0003841333382297307
step: 50, loss: 0.0647120401263237
step: 60, loss: 0.00017193725216202438
step: 70, loss: 0.03653615340590477
step: 80, loss: 0.00032528044539503753
step: 90, loss: 0.00048192418762482703
step: 100, loss: 0.02485293708741665
step: 110, loss: 0.0005810587317682803
step: 120, loss: 0.00019713971414603293
step: 130, loss: 0.0008560754940845072
step: 140, loss: 0.000842004141304642
step: 150, loss: 0.0004971435409970582
step: 160, loss: 0.09385579824447632
step: 170, loss: 0.012301388196647167
step: 180, loss: 0.0009307120926678181
step: 190, loss: 0.004864403512328863
step: 200, loss: 0.0002498562098480761
step: 210, loss: 0.00023424421669915318
step: 220, loss: 0.0002525036979932338
step: 230, loss: 0.00019514943414833397
/usr/local/lib/python3.7/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=0.9887640449438202, f1=0.9831271091113611,
best_f1=0.9842342342342343
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0002540623245295137
step: 10, loss: 0.00038269179640337825
step: 20, loss: 0.00024246214888989925
step: 30, loss: 0.0003222591185476631
step: 40, loss: 0.000184161719516851
step: 50, loss: 0.00012892948871012777
step: 60, loss: 0.00011485537106636912
step: 70, loss: 0.00012383621651679277
step: 80, loss: 0.00047578688827343285
step: 90, loss: 0.0004688352928496897
step: 100, loss: 0.020366452634334564
step: 110, loss: 0.0002126629406120628
```

```
step: 120, loss: 0.00011176343832630664
step: 130, loss: 9.03260734048672e-05
step: 140, loss: 7.9751291195862e-05
step: 150, loss: 0.09700535237789154
step: 160, loss: 0.00028571143047884107
step: 170, loss: 0.009959857910871506
step: 180, loss: 0.041766855865716934
step: 190, loss: 0.0021041659638285637
step: 200, loss: 0.00022617937065660954
step: 210, loss: 0.00020348653197288513
step: 220, loss: 7.771200034767389e-05
step: 230, loss: 0.03207036480307579
/usr/local/lib/python3.7/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=0.984304932735426, f1=0.9797297297297298,
best_f1=0.9842342342342343
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0006778785609640181
step: 10, loss: 0.00020313362528214231
step: 20, loss: 0.00011597051343414932
step: 30, loss: 0.0028276594821363688
step: 40, loss: 0.00013672430941369385
step: 50, loss: 7.39100287319161e-05
step: 60, loss: 0.00029758643358945847
step: 70, loss: 0.01473754271864891
step: 80, loss: 0.00012831110507249832
step: 90, loss: 4.2555468098726124e-05
step: 100, loss: 5.957140092505142e-05
step: 110, loss: 0.0007393266423605382
step: 120, loss: 0.00013469260011333972
step: 130, loss: 0.00015789766621310264
step: 140, loss: 0.00014263093180488795
step: 150, loss: 0.0006617932231165469
step: 160, loss: 0.053634706884622574
step: 170, loss: 0.03884851932525635
step: 180, loss: 0.0005268572713248432
step: 190, loss: 0.00048138879355974495
step: 200, loss: 0.04993297904729843
step: 210, loss: 8.958952821558341e-05
step: 220, loss: 0.0017694317502900958
step: 230, loss: 0.0001784090418368578
/usr/local/lib/python3.7/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=0.9809203142536477,
```

```
best_f1=0.9842342342342343
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00016301145660690963
step: 10, loss: 0.028104795143008232
step: 20, loss: 0.00013237232633400708
step: 30, loss: 0.00017842378292698413
step: 40, loss: 0.00030309276189655066
step: 50, loss: 0.0016597185749560595
step: 60, loss: 0.0007652775966562331
step: 70, loss: 0.018299970775842667
step: 80, loss: 0.00016522830992471427
step: 90, loss: 6.65901752654463e-05
step: 100, loss: 0.001293055946007371
step: 110, loss: 0.0005824011168442667
step: 120, loss: 0.0022679297253489494
step: 130, loss: 0.00011209851072635502
step: 140, loss: 6.716712232446298e-05
step: 150, loss: 6.490988016594201e-05
step: 160, loss: 0.000139689160278067
step: 170, loss: 9.4929440820124e-05
step: 180, loss: 8.602347952546552e-05
step: 190, loss: 0.00012810496264137328
step: 200, loss: 0.002335190074518323
step: 210, loss: 0.00020737761224154383
step: 220, loss: 0.006781162694096565
step: 230, loss: 0.001352639519609511
/usr/local/lib/python3.7/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=0.9887640449438202, f1=0.9731543624161074,
best_f1=0.9842342342342343
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0005625911289826035
step: 10, loss: 0.002048413036391139
step: 20, loss: 0.019039738923311234
step: 30, loss: 0.0013957041082903743
step: 40, loss: 0.011436576955020428
step: 50, loss: 0.005139630287885666
step: 60, loss: 0.00041890353895723820
step: 70, loss: 0.0021480401046574116
step: 80, loss: 0.07313874363899231
step: 90, loss: 0.0016744454624131322
step: 100, loss: 0.0013194566126912832
step: 110, loss: 0.002491159364581108
```

```
step: 120, loss: 0.0005371205043047667
step: 130, loss: 0.00023937804508022964
step: 140, loss: 0.00031308241887018085
step: 150, loss: 0.017006302252411842
step: 160, loss: 0.0017808381235226989
step: 170, loss: 0.000497817003633827
step: 180, loss: 0.11254528164863586
step: 190, loss: 0.00042249110992997885
step: 200, loss: 0.0013124685501679778
step: 210, loss: 0.0005225093918852508
step: 220, loss: 0.0005029041203670204
step: 230, loss: 0.0010997481876984239
/usr/local/lib/python3.7/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=0.9864253393665158, f1=0.9796839729119639,
best_f1=0.9842342342342343
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00027670388226397336
step: 10, loss: 0.0003649517602752894
step: 20, loss: 0.0012668658746406436
step: 30, loss: 0.027336524799466133
step: 40, loss: 0.0002416484639979899
step: 50, loss: 0.0002682362392079085
step: 60, loss: 0.00024955644039437175
step: 70, loss: 0.0004589023010339588
step: 80, loss: 0.0002726046950556338
step: 90, loss: 0.002595443045720458
step: 100, loss: 0.0007615590584464371
step: 110, loss: 0.00088715628953650359
step: 120, loss: 0.00043446669587586547
step: 130, loss: 0.0005722732748836279
step: 140, loss: 0.029476296156664482
step: 150, loss: 0.025189999490976334
step: 160, loss: 8.624249312561005e-05
step: 170, loss: 0.0005181374726817012
step: 180, loss: 0.0029053385369479656
step: 190, loss: 0.0009017442935146391
step: 200, loss: 0.0001269693748326972
step: 210, loss: 0.0017826813273131847
step: 220, loss: 0.00010783578909467906
step: 230, loss: 0.0014702263288199902
/usr/local/lib/python3.7/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=0.9852440408626559, f1=0.9786276715410572,
```

```
best_f1=0.9842342342342343
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0013213992351666093
step: 10, loss: 0.00012290573795326054
step: 20, loss: 0.0440414622426033
step: 30, loss: 0.029000360518693924
step: 40, loss: 0.0005324496887624264
step: 50, loss: 0.006254593376070261
step: 60, loss: 0.0003503359039314091
step: 70, loss: 0.006486298982053995
step: 80, loss: 0.00012735207565128803
step: 90, loss: 0.0002431704051559791
step: 100, loss: 0.00019366161723155528
step: 110, loss: 0.0028442975599318743
step: 120, loss: 0.00010441905760671943
step: 130, loss: 0.00011996964894933626
step: 140, loss: 0.0002332036820007488
step: 150, loss: 0.03108302131295204
step: 160, loss: 0.00010920778731815517
step: 170, loss: 0.018422279506921768
step: 180, loss: 0.00017074489733204246
step: 190, loss: 0.04522450268268585
step: 200, loss: 0.0003096498257946223
step: 210, loss: 0.0002040144318016246
step: 220, loss: 0.00010233134526060894
step: 230, loss: 0.00034873143886215985
/usr/local/lib/python3.7/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=0.9798206278026906,
best_f1=0.9842342342342343
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 9.95289083220996e-05
step: 10, loss: 0.00049233203753829
step: 20, loss: 9.797926031751558e-05
step: 30, loss: 0.00016467797104269266
step: 40, loss: 0.0003930600068997592
step: 50, loss: 0.0003104038769379258
step: 60, loss: 0.000537690008059144
step: 70, loss: 0.0050271134823560715
step: 80, loss: 0.0010425058426335454
step: 90, loss: 8.31836587167345e-05
step: 100, loss: 0.00011904873827006668
step: 110, loss: 0.0008104121661745012
```

```
step: 120, loss: 7.467936666216701e-05
step: 130, loss: 7.316585106309503e-05
step: 140, loss: 0.00010692621435737237
step: 150, loss: 0.00040067333611659706
step: 160, loss: 0.0002325134410057217
step: 170, loss: 0.00014761646161787212
step: 180, loss: 0.0002637183351907879
step: 190, loss: 0.00131084187887609
step: 200, loss: 4.426861050887965e-05
step: 210, loss: 0.00030098308343440294
step: 220, loss: 7.451918645529076e-05
step: 230, loss: 0.027549976482987404
/usr/local/lib/python3.7/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=0.9841986455981941, f1=0.9798657718120806,
best_f1=0.9842342342342343
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00018843045108951628
step: 10, loss: 0.00079743517562747
step: 20, loss: 0.00012810455518774688
step: 30, loss: 8.807966514723375e-05
step: 40, loss: 0.0001115100021706894
step: 50, loss: 0.00011959655967075378
step: 60, loss: 7.170601020334288e-05
step: 70, loss: 0.00035616260720416903
step: 80, loss: 0.00018921254377346486
step: 90, loss: 0.00015147638623602688
step: 100, loss: 4.397440352477133e-05
step: 110, loss: 0.0028856617864221334
step: 120, loss: 0.02672000601887703
step: 130, loss: 0.0002210071252193302
step: 140, loss: 0.00011574963718885556
step: 150, loss: 9.029218927025795e-05
step: 160, loss: 6.65154802845791e-05
step: 170, loss: 6.615507300011814e-05
step: 180, loss: 0.00012994790449738503
step: 190, loss: 5.248394882073626e-05
step: 200, loss: 0.0002034520439337939
step: 210, loss: 5.085672819404863e-05
step: 220, loss: 0.00013070525892544538
step: 230, loss: 7.189460302470252e-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=0.9852440408626559, f1=0.9796380090497738,
```

best_f1=0.9842342342342343
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 8.17993059172295e-05
step: 10, loss: 0.015811767429113388
step: 20, loss: 5.853818584000692e-05
step: 30, loss: 5.521334242075682e-05
step: 40, loss: 0.024112462997436523
step: 50, loss: 9.452247468288988e-05
step: 60, loss: 8.230312232626602e-05
step: 70, loss: 5.510869232239202e-05
step: 80, loss: 6.434617534978315e-05
step: 90, loss: 6.507922807941213e-05
step: 100, loss: 5.2892246458213776e-05
step: 110, loss: 0.005251242779195309
step: 120, loss: 4.0309212636202574e-05
step: 130, loss: 5.554458766710013e-05
step: 140, loss: 7.345463382080197e-05
step: 150, loss: 5.0727790949167684e-05
step: 160, loss: 0.00024357317306566983
step: 170, loss: 9.631291322875768e-05
step: 180, loss: 4.197122325422242e-05
step: 190, loss: 0.00020417780615389347
step: 200, loss: 6.554660649271682e-05
step: 210, loss: 6.112117989687249e-05
step: 220, loss: 3.856646435451694e-05
step: 230, loss: 0.0005535394884645939
/usr/local/lib/python3.7/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=0.9775784753363228,
best_f1=0.9842342342342343
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 5.495546429301612e-05
step: 10, loss: 3.414861566852778e-05
step: 20, loss: 7.955086039146408e-05
step: 30, loss: 8.730233093956485e-05
step: 40, loss: 9.404934826306999e-05
step: 50, loss: 0.03799119591712952
step: 60, loss: 6.342565029626712e-05
step: 70, loss: 9.075671550817788e-05
step: 80, loss: 9.851194772636518e-05
step: 90, loss: 0.0001315141562372446
step: 100, loss: 8.053216151893139e-05
step: 110, loss: 6.432658119592816e-05

```
step: 120, loss: 8.120117854559794e-05
step: 130, loss: 4.846740193897858e-05
step: 140, loss: 5.542780490941368e-05
step: 150, loss: 0.00020255890558473766
step: 160, loss: 6.386195309460163e-05
step: 170, loss: 3.507982910377905e-05
step: 180, loss: 0.00016891348059289157
step: 190, loss: 0.00015525927301496267
step: 200, loss: 6.165200466057286e-05
step: 210, loss: 9.058507566805929e-05
step: 220, loss: 0.0002461929398123175
step: 230, loss: 4.64630575152114e-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=0.9853107344632768, f1=0.9774774774774775,
best_f1=0.9842342342342343
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 [00:15, 160.17it/s]

```
load_f1 = 0.9898989898989898
real_f1 = 0.9898989898989898
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
4398it [00:23, 183.51it/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
```

156

```
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: 0.6264797449111938
step: 10, loss: 0.5796636939048767
step: 20, loss: 0.5105891227722168
step: 30, loss: 0.04149889573454857
step: 40, loss: 0.18932755291461945
step: 50, loss: 0.25733035802841187
step: 60, loss: 0.06434642523527145
step: 70, loss: 0.15270158648490906
step: 80, loss: 0.02858687750995159
step: 90, loss: 0.2736411988735199
step: 100, loss: 0.09923329204320908
step: 110, loss: 0.15016210079193115
step: 120, loss: 0.14926376938819885
step: 130, loss: 0.12824930250644684
step: 140, loss: 0.08006158471107483
step: 150, loss: 0.10085911303758621
step: 160, loss: 0.021000036969780922
step: 170, loss: 0.20480872690677643
step: 180, loss: 0.09711119532585144
step: 190, loss: 0.026989564299583435
```

```
step: 200, loss: 0.16218023002147675
step: 210, loss: 0.0882437527179718
step: 220, loss: 0.14681333303451538
step: 230, loss: 0.16993378102779388
step: 240, loss: 0.05940072238445282
step: 250, loss: 0.05520491674542427
step: 260, loss: 0.1410529613494873
step: 270, loss: 0.012315092608332634
step: 280, loss: 0.04700390249490738
step: 290, loss: 0.04785050079226494
step: 300, loss: 0.05958467721939087
step: 310, loss: 0.09795168787240982
step: 320, loss: 0.09017205983400345
step: 330, loss: 0.14253640174865723
step: 340, loss: 0.02440459653735161
step: 350, loss: 0.04501941055059433
step: 360, loss: 0.015924930572509766
step: 370, loss: 0.04987553134560585
step: 380, loss: 0.028937209397554398
step: 390, loss: 0.09460382163524628
step: 400, loss: 0.4052278399467468
step: 410, loss: 0.0835861936211586
step: 420, loss: 0.09184549748897552
step: 430, loss: 0.20632974803447723
step: 440, loss: 0.05220947787165642
step: 450, loss: 0.007643949240446091
step: 460, loss: 0.0045568812638521194
step: 470, loss: 0.12635540962219238
step: 480, loss: 0.03533947467803955
step: 490, loss: 0.05013275146484375
step: 500, loss: 0.14723052084445953
step: 510, loss: 0.12740114331245422
step: 520, loss: 0.0183267742395401
step: 530, loss: 0.0026379062328487635
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9313047487321346, f1=0.928801102434543,
best_f1=0.928801102434543
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.4282241761684418
step: 10, loss: 0.14382141828536987
```

```
step: 20, loss: 0.010002495720982552
step: 30, loss: 0.027746351435780525
step: 40, loss: 0.08094671368598938
step: 50, loss: 0.10785053670406342
step: 60, loss: 0.007966406643390656
step: 70, loss: 0.006841112859547138
step: 80, loss: 0.06009824201464653
step: 90, loss: 0.023692838847637177
step: 100, loss: 0.010499331168830395
step: 110, loss: 0.04325389117002487
step: 120, loss: 0.02892235666513443
step: 130, loss: 0.08882779628038406
step: 140, loss: 0.01987864263355732
step: 150, loss: 0.05871433764696121
step: 160, loss: 0.036236248910427094
step: 170, loss: 0.039043132215738297
step: 180, loss: 0.019424809142947197
step: 190, loss: 0.06440184265375137
step: 200, loss: 0.008500497788190842
step: 210, loss: 0.04987800121307373
step: 220, loss: 0.06631910055875778
step: 230, loss: 0.005671640858054161
step: 240, loss: 0.11282604187726974
step: 250, loss: 0.13542701303958893
step: 260, loss: 0.00848821084946394
step: 270, loss: 0.2919505834579468
step: 280, loss: 0.04637765511870384
step: 290, loss: 0.04090355709195137
step: 300, loss: 0.13528423011302948
step: 310, loss: 0.015960611402988434
step: 320, loss: 0.06502464413642883
step: 330, loss: 0.00960633996874094
step: 340, loss: 0.009605173021554947
step: 350, loss: 0.0028264345601201057
step: 360, loss: 0.07856465131044388
step: 370, loss: 0.14473913609981537
step: 380, loss: 0.08355088531970978
step: 390, loss: 0.1286901980638504
step: 400, loss: 0.08981747925281525
step: 410, loss: 0.046405769884586334
step: 420, loss: 0.024850983172655106
step: 430, loss: 0.023894062265753746
step: 440, loss: 0.17610956728458405
step: 450, loss: 0.021763669326901436
step: 460, loss: 0.09260276705026627
step: 470, loss: 0.05711446329951286
step: 480, loss: 0.2606547474861145
step: 490, loss: 0.021673955023288727
```

```
step: 500, loss: 0.2507379353046417
step: 510, loss: 0.022866982966661453
step: 520, loss: 0.023683272302150726
step: 530, loss: 0.05701705068349838
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9359513791491351, f1=0.9408476944573824,
best_f1=0.9408476944573824
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.02378014102578163
step: 10, loss: 0.046503495424985886
step: 20, loss: 0.08083922415971756
step: 30, loss: 0.034228891134262085
step: 40, loss: 0.02027580700814724
step: 50, loss: 0.13076484203338623
step: 60, loss: 0.010594965890049934
step: 70, loss: 0.0045076641254127026
step: 80, loss: 0.001964293885976076
step: 90, loss: 0.008689631707966328
step: 100, loss: 0.07555422931909561
step: 110, loss: 0.010058310814201832
step: 120, loss: 0.002464551944285631
step: 130, loss: 0.0032214648090302944
step: 140, loss: 0.01893952488899231
step: 150, loss: 0.009173804894089699
step: 160, loss: 0.00424785353243351
step: 170, loss: 0.08977067470550537
step: 180, loss: 0.0061074523255229
step: 190, loss: 0.010221621952950954
step: 200, loss: 0.09062105417251587
step: 210, loss: 0.04276618734002113
step: 220, loss: 0.16883179545402527
step: 230, loss: 0.04446353390812874
step: 240, loss: 0.0021331205498427153
step: 250, loss: 0.07015679776668549
step: 260, loss: 0.013573980890214443
step: 270, loss: 0.031260911375284195
step: 280, loss: 0.14855830371379852
step: 290, loss: 0.0040300870314240456
step: 300, loss: 0.019531184807419777
step: 310, loss: 0.016894502565264702
```

```
step: 320, loss: 0.013825053349137306
step: 330, loss: 0.0005851167370565236
step: 340, loss: 0.033035408705472946
step: 350, loss: 0.05793813616037369
step: 360, loss: 0.06407755613327026
step: 370, loss: 0.012588584795594215
step: 380, loss: 0.006076066754758358
step: 390, loss: 0.02180168405175209
step: 400, loss: 0.024262098595499992
step: 410, loss: 0.0047386265359818935
step: 420, loss: 0.16090525686740875
step: 430, loss: 0.03688650205731392
step: 440, loss: 0.0004332697717472911
step: 450, loss: 0.08810176700353622
step: 460, loss: 0.02592308819293976
step: 470, loss: 0.014568261802196503
step: 480, loss: 0.008333717472851276
step: 490, loss: 0.02139556035399437
step: 500, loss: 0.04548261687159538
step: 510, loss: 0.061884474009275436
step: 520, loss: 0.07810647785663605
step: 530, loss: 0.028316980227828026
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.937528921795465, f1=0.9397590361445785,
best_f1=0.9397590361445785
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00614946149289608
step: 10, loss: 0.0017209402285516262
step: 20, loss: 0.00041377736488357186
step: 30, loss: 0.0012282916577532887
step: 40, loss: 0.002112924586981535
step: 50, loss: 0.012344255112111568
step: 60, loss: 0.001160252490080893
step: 70, loss: 0.0012835027882829309
step: 80, loss: 0.0016844646306708455
step: 90, loss: 0.031342510133981705
step: 100, loss: 0.03053983859717846
step: 110, loss: 0.08120924979448318
step: 120, loss: 0.0023935867939144373
step: 130, loss: 0.0295831561088562
```

```
step: 140, loss: 0.0009356781374663115
step: 150, loss: 0.0035690967924892902
step: 160, loss: 0.09106329828500748
step: 170, loss: 0.004344150889664888
step: 180, loss: 0.001020793686620891
step: 190, loss: 0.11776486039161682
step: 200, loss: 0.03515485301613808
step: 210, loss: 0.04543352499604225
step: 220, loss: 0.06910441070795059
step: 230, loss: 0.09622053802013397
step: 240, loss: 0.006698907818645239
step: 250, loss: 0.07311729341745377
step: 260, loss: 0.0017759207403287292
step: 270, loss: 0.008103259839117527
step: 280, loss: 0.014604317024350166
step: 290, loss: 0.022977864369750023
step: 300, loss: 0.002766540041193366
step: 310, loss: 0.008854048326611519
step: 320, loss: 0.014939923770725727
step: 330, loss: 0.0038089780136942863
step: 340, loss: 0.05705561488866806
step: 350, loss: 0.10726699978113174
step: 360, loss: 0.01814541593194008
step: 370, loss: 0.012864870950579643
step: 380, loss: 0.008292542770504951
step: 390, loss: 0.002932235598564148
step: 400, loss: 0.01629350334405899
step: 410, loss: 0.008692096918821335
step: 420, loss: 0.003908325918018818
step: 430, loss: 0.23547011613845825
step: 440, loss: 0.007704154122620821
step: 450, loss: 0.0035476358607411385
step: 460, loss: 0.0010420707985758781
step: 470, loss: 0.02036360278725624
step: 480, loss: 0.12194480001926422
step: 490, loss: 0.001477784593589604
step: 500, loss: 0.0015963183250278234
step: 510, loss: 0.01896330714225769
Gradient overflow.  Skipping step, loss scaler 0 reducing loss scale to 32768.0
step: 520, loss: 0.1109549030661583
step: 530, loss: 0.02305159904062748
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9397260273972603, f1=0.9410689812699863,
best_f1=0.9410689812699863
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.019907278597354889
step: 10, loss: 0.04253368824720383
step: 20, loss: 0.09553717821836472
step: 30, loss: 0.003820653771981597
step: 40, loss: 0.00491533987224102
step: 50, loss: 0.0028913491405546665
step: 60, loss: 0.009862327948212624
step: 70, loss: 0.004498199559748173
step: 80, loss: 0.000882316380739212
step: 90, loss: 0.0006767820450477302
step: 100, loss: 0.0022152455057948828
step: 110, loss: 0.00041573005728423595
step: 120, loss: 0.0008791537838988006
step: 130, loss: 0.0004181790864095092
step: 140, loss: 0.0021528967190533876
step: 150, loss: 0.008331084623932838
step: 160, loss: 0.0006459890864789486
step: 170, loss: 0.009071671403944492
step: 180, loss: 0.000791162543464452
step: 190, loss: 0.013969296589493752
step: 200, loss: 0.0019563643727451563
step: 210, loss: 0.0032917477656155825
step: 220, loss: 0.006920026149600744
step: 230, loss: 0.027726629748940468
step: 240, loss: 0.01274082437157631
step: 250, loss: 0.0008840272785164416
step: 260, loss: 0.00062654464272778221
step: 270, loss: 0.0005174357793293893
step: 280, loss: 0.0006103711202740669
step: 290, loss: 0.1965404450893402
step: 300, loss: 0.0041307322680950165
step: 310, loss: 0.0007117776549421251
step: 320, loss: 0.01594020053744316
step: 330, loss: 0.018343940377235413
step: 340, loss: 0.003085664939135313
step: 350, loss: 0.005248062312602997
step: 360, loss: 0.0051747700199484825
step: 370, loss: 0.0019047128735110164
step: 380, loss: 0.0002966096217278391
step: 390, loss: 0.0003929976955987513
step: 400, loss: 0.004959827288985252
step: 410, loss: 0.0009485890041105449
step: 420, loss: 0.001862755510956049
```

```
step: 430, loss: 0.0029569033067673445
step: 440, loss: 0.03864328935742378
step: 450, loss: 0.016541199758648872
step: 460, loss: 0.000713221961632371
step: 470, loss: 0.0014660024316981435
step: 480, loss: 0.0016137680504471064
step: 490, loss: 0.0003380277776159346
step: 500, loss: 0.00046172604197636247
step: 510, loss: 0.1424761414527893
step: 520, loss: 0.0021376044023782015
step: 530, loss: 0.0030798083171248436
/usr/local/lib/python3.7/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=0.9287749287749287, f1=0.9359095193213949,
best_f1=0.9410689812699863
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00214552553370595
step: 10, loss: 0.0048523335717618465
step: 20, loss: 0.13947059214115143
step: 30, loss: 0.006380051840096712
step: 40, loss: 0.0005257982993498445
step: 50, loss: 0.016026217490434647
step: 60, loss: 0.00018302095122635365
step: 70, loss: 0.006065624300390482
step: 80, loss: 0.002164103789255023
step: 90, loss: 0.003387796925380826
step: 100, loss: 0.006712055299431086
step: 110, loss: 0.012701311148703098
step: 120, loss: 0.016804374754428864
step: 130, loss: 0.0026263538748025894
step: 140, loss: 0.001533582923002541
step: 150, loss: 0.04747790843248367
step: 160, loss: 0.009344718419015408
step: 170, loss: 0.00033521177829243243
step: 180, loss: 0.00040663316030986607
step: 190, loss: 0.00024688037228770554
step: 200, loss: 0.0012907484779134393
step: 210, loss: 0.0025356044061481953
step: 220, loss: 0.00992236752063036
step: 230, loss: 0.004517430905252695
step: 240, loss: 0.059908222407102585
step: 250, loss: 0.0034580484498292208
step: 260, loss: 0.0001600418472662568
step: 270, loss: 0.10763762146234512
step: 280, loss: 0.0648399218916893
```

```
step: 290, loss: 0.0014793338486924767
step: 300, loss: 0.0039058434776961803
step: 310, loss: 0.0086691658943891153
step: 320, loss: 0.0025271649938076735
step: 330, loss: 0.0016444443026557565
step: 340, loss: 0.0778818279504776
step: 350, loss: 0.003396779065951705
step: 360, loss: 0.00896213948726654
step: 370, loss: 0.003362886141985655
step: 380, loss: 0.00037587888073176146
step: 390, loss: 0.004522319417446852
step: 400, loss: 0.0022676601074635983
step: 410, loss: 0.002865451155230403
step: 420, loss: 0.01624351181089878
step: 430, loss: 0.0007310957298614085
step: 440, loss: 0.0007669473416171968
step: 450, loss: 0.0030823221895843744
step: 460, loss: 0.002088285516947508
step: 470, loss: 0.0024005777668207884
step: 480, loss: 0.012545347213745117
step: 490, loss: 0.013417490758001804
step: 500, loss: 0.0004900619387626648
step: 510, loss: 0.0011298370081931353
step: 520, loss: 0.0002888762974180281
step: 530, loss: 0.012786309234797955
/usr/local/lib/python3.7/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=0.9365225390984362, f1=0.942035600182565,
best_f1=0.9410689812699863
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00010592280159471557
step: 10, loss: 0.0003942172625102103
step: 20, loss: 0.00017595798999536783
step: 30, loss: 0.0019488041289150715
step: 40, loss: 0.03136494755744934
step: 50, loss: 0.0007972888997755945
step: 60, loss: 0.0002916343801189214
step: 70, loss: 0.011471526697278023
step: 80, loss: 0.003397201879086494
step: 90, loss: 0.02118891291320324
step: 100, loss: 0.07018017768859863
step: 110, loss: 0.0009830074850469828
step: 120, loss: 0.0006048679351806641
step: 130, loss: 0.0006809807382524014
step: 140, loss: 0.002399051794782281
```

```
step: 150, loss: 0.00953192263841629
step: 160, loss: 0.004716468974947929
step: 170, loss: 0.007118814159184694
step: 180, loss: 0.008843835443258286
step: 190, loss: 0.001547761494293809
step: 200, loss: 0.00022133854508865625
step: 210, loss: 0.000692821922712028
step: 220, loss: 0.0002506775490473956
step: 230, loss: 0.007374669425189495
step: 240, loss: 0.0009982057381421328
step: 250, loss: 0.00040318904211744666
step: 260, loss: 0.0009453139500692487
step: 270, loss: 0.0025809959042817354
step: 280, loss: 0.0003245171101298183
step: 290, loss: 0.0001663397706579417
step: 300, loss: 0.0006600231863558292
step: 310, loss: 0.0001073975145118311
step: 320, loss: 0.0004972700262442231
step: 330, loss: 0.002019620267674327
step: 340, loss: 0.02532227151095867
step: 350, loss: 0.022606132552027702
step: 360, loss: 0.001743991975672543
step: 370, loss: 0.002322485903277993
step: 380, loss: 0.002574289683252573
step: 390, loss: 0.0011184157337993383
step: 400, loss: 0.01023595966398716
step: 410, loss: 0.005021970719099045
step: 420, loss: 0.0004558605724014342
step: 430, loss: 9.53617854975164e-05
step: 440, loss: 0.00016743148444220424
step: 450, loss: 0.00376032292842865
step: 460, loss: 0.0001346164644928649
step: 480, loss: 0.09236464649438858
step: 490, loss: 0.002606153255328536
step: 500, loss: 0.00022337239352054894
step: 510, loss: 0.0021445313468575478
step: 520, loss: 0.02200537733733654
step: 530, loss: 0.002923469291999936
/usr/local/lib/python3.7/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=0.9379374708352777, f1=0.9413936317489617,
best_f1=0.9410689812699863
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0006544138304889202
step: 10, loss: 0.03570239618420601
```

```
step: 20, loss: 0.011199328117072582
step: 30, loss: 0.000630350667051971
step: 40, loss: 0.001198118319734931
step: 50, loss: 0.0020419645588845015
step: 60, loss: 0.0011268719099462032
step: 70, loss: 0.00033631210681051016
step: 80, loss: 0.001407264149747789
step: 90, loss: 0.0005983322625979781
step: 100, loss: 0.00018652663857210428
step: 110, loss: 0.00015885315951891243
step: 120, loss: 0.0002535762614570558
step: 130, loss: 0.0003223646490368992
step: 140, loss: 0.002309644129127264
step: 150, loss: 0.00014131856733001769
step: 160, loss: 0.05816630274057388
step: 170, loss: 0.0009059247095137835
step: 180, loss: 0.0003275318304076791
Gradient overflow.  Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 190, loss: 0.002807553159072995
step: 200, loss: 0.001613837550394237
step: 210, loss: 0.0001613945933058858
step: 220, loss: 0.0027942396700382233
step: 230, loss: 0.00028579641366377473
step: 240, loss: 8.408905705437064e-05
step: 250, loss: 6.832656072219834e-05
step: 260, loss: 7.598514639539644e-05
step: 270, loss: 0.0005232898984104395
step: 280, loss: 0.02706129476428032
step: 290, loss: 0.00021219828340690583
step: 300, loss: 0.0007233505020849407
step: 310, loss: 0.00032776768784970045
step: 320, loss: 0.023361265659332275
step: 330, loss: 0.00026126028387800731
step: 340, loss: 0.00041798703023232252
step: 350, loss: 0.00021445209858939052
step: 360, loss: 0.00498993881046772
step: 370, loss: 0.00075528584420681
step: 380, loss: 0.0009459700086154044
step: 390, loss: 0.06990984827280045
step: 400, loss: 0.028905479237437248
step: 410, loss: 0.00022102822549641132
step: 420, loss: 0.0009520823950879276
step: 430, loss: 0.05106959864497185
step: 440, loss: 0.0003161929489579052
step: 450, loss: 0.00031809776555746794
step: 460, loss: 0.00023858096392359585
step: 470, loss: 0.00015901026199571788
step: 480, loss: 0.00012786251318175346
```

```
step: 490, loss: 0.060140907764434814
step: 500, loss: 0.004495264496654272
step: 510, loss: 0.002480795606970787
step: 520, loss: 0.001965508097782731
step: 530, loss: 0.0008595475228503346
/usr/local/lib/python3.7/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=0.931892907468295, f1=0.9392111368909513,
best_f1=0.9410689812699863
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.002621621359139681
step: 10, loss: 0.00034448778023943305
step: 20, loss: 0.0013008207315579057
step: 30, loss: 0.0005037090741097927
step: 40, loss: 0.0003825408057309687
step: 50, loss: 0.00020575021335389465
step: 60, loss: 8.228506339946762e-05
step: 70, loss: 0.006074386648833752
step: 80, loss: 0.0003010049695149064
step: 90, loss: 0.0014812470180913806
step: 100, loss: 0.0001330364029854536
step: 110, loss: 0.00012709795555565506
step: 120, loss: 0.00040220029768534005
step: 130, loss: 0.0003073179686907679
step: 140, loss: 0.003208366222679615
step: 150, loss: 0.00012111983960494399
step: 160, loss: 0.0011198900174349546
step: 170, loss: 0.15147292613983154
step: 180, loss: 8.067591988947242e-05
step: 190, loss: 0.0014342796057462692
step: 200, loss: 0.0001023646182147786
step: 210, loss: 8.264795906143263e-05
step: 220, loss: 0.0001750914379954338
step: 230, loss: 8.742330101085827e-05
step: 240, loss: 0.06059933453798294
step: 250, loss: 0.0010072910226881504
step: 260, loss: 0.015440357849001884
step: 270, loss: 8.565836469642818e-05
step: 280, loss: 0.010838417336344719
step: 290, loss: 0.04183483496308327
step: 300, loss: 0.0003360615810379386
step: 310, loss: 0.026542847976088524
step: 320, loss: 0.000332935422193259
step: 330, loss: 0.0001863948709797114
step: 340, loss: 7.432234269799665e-05
```

```
step: 350, loss: 0.0005768080009147525
step: 360, loss: 0.00010703269799705595
step: 370, loss: 0.03989681601524353
step: 380, loss: 0.002762393793091178
step: 390, loss: 0.00023019984655547887
step: 400, loss: 0.0010311175137758255
step: 410, loss: 0.00036334936157800257
step: 420, loss: 5.990951831336133e-05
step: 430, loss: 0.00075879180803895
step: 440, loss: 0.0018122639739885926
step: 450, loss: 0.0003798298130277544
step: 460, loss: 0.00011830039147753268
step: 470, loss: 5.137227344675921e-05
step: 480, loss: 0.00015292508760467172
step: 490, loss: 0.00047658407129347324
step: 500, loss: 0.0020326185040175915
step: 510, loss: 0.0004310854710638523
step: 520, loss: 0.000301951018627733
step: 530, loss: 0.0001993140031117946
/usr/local/lib/python3.7/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=0.9379441023211748, f1=0.9418386491557222,
best_f1=0.9410689812699863
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0032473502214998007
step: 10, loss: 0.00014485811698250473
step: 20, loss: 0.0001415980514138937
step: 30, loss: 0.0001185989603982307
step: 40, loss: 0.0012327530421316624
step: 50, loss: 0.0007192183984443545
step: 60, loss: 0.0008512124186381698
step: 70, loss: 0.00039599102456122637
step: 80, loss: 0.00018608941172715276
step: 90, loss: 0.00024975574342533946
step: 100, loss: 0.00026034272741526365
step: 110, loss: 7.487408583983779e-05
step: 120, loss: 0.00022992298181634396
step: 130, loss: 0.00021792470943182707
step: 140, loss: 0.006843366660177708
step: 150, loss: 0.0005162103334441781
step: 160, loss: 0.0001935295294970274
step: 170, loss: 0.00034934416180476546
step: 180, loss: 0.00029760599136352554
step: 190, loss: 0.0002785835531540215
step: 200, loss: 0.00021653337171301246
```

```
step: 210, loss: 0.000312607066007331
step: 220, loss: 0.0037895238492637873
step: 230, loss: 0.0018189733382314444
step: 240, loss: 0.0011559005361050367
step: 250, loss: 0.001450743991881609
step: 260, loss: 0.00027515151305124164
step: 270, loss: 9.190602577291429e-05
step: 280, loss: 7.284348248504102e-05
step: 290, loss: 6.440307333832607e-05
step: 300, loss: 3.564564758562483e-05
step: 310, loss: 8.963185973698273e-05
step: 320, loss: 0.0009282948449254036
step: 330, loss: 0.00012139086902607232
step: 340, loss: 0.06529013812541962
step: 350, loss: 0.0001456641184631735
step: 360, loss: 0.0022608910221606493
step: 370, loss: 0.0036493302322924137
step: 380, loss: 0.00013203790877014399
step: 390, loss: 0.0005289770197123289
step: 400, loss: 0.00018573971465229988
step: 410, loss: 0.0020394327584654093
step: 420, loss: 0.0001427287788828835
step: 430, loss: 9.759997192304581e-05
step: 440, loss: 0.002281413646414876
step: 450, loss: 0.00022541036014445126
step: 460, loss: 5.874583075637929e-05
step: 470, loss: 0.00039489357732236385
step: 480, loss: 0.002415654482319951
step: 490, loss: 0.00036814436316490173
step: 500, loss: 0.0014431847957894206
step: 510, loss: 0.00011544768494786695
step: 520, loss: 9.207518451148644e-05
step: 530, loss: 0.0009262213134206831
/usr/local/lib/python3.7/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=0.9394785847299814, f1=0.9414498141263941,
best_f1=0.9410689812699863
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0018910960061475635
step: 10, loss: 0.0003156732418574393
step: 20, loss: 0.00014389576972462237
step: 30, loss: 0.0002286114322487265
step: 40, loss: 8.464180427836254e-05
step: 50, loss: 0.00039221704355441034
step: 60, loss: 0.04141397774219513
```

```
step: 70, loss: 5.900092583033256e-05
step: 80, loss: 0.0003111511468887329
step: 90, loss: 0.00020307950035203248
step: 100, loss: 0.009157117456197739
step: 110, loss: 0.00011330701818224043
step: 120, loss: 0.0020323689095675945
step: 130, loss: 0.0002132020308636129
step: 140, loss: 5.785674147773534e-05
step: 150, loss: 4.770676969201304e-05
step: 160, loss: 5.8411835198057815e-05
step: 170, loss: 0.00011827093840111047
step: 180, loss: 6.235302134882659e-05
step: 190, loss: 0.0013423513155430555
step: 200, loss: 0.011949460953474045
step: 210, loss: 0.00010002048657042906
step: 220, loss: 0.00117143162060529
step: 230, loss: 2.9834320230293088e-05
step: 240, loss: 0.00014997024845797569
step: 250, loss: 3.911961903213523e-05
step: 260, loss: 0.0002173220709664747
step: 270, loss: 6.0748217947548255e-05
step: 280, loss: 7.697546971030533e-05
step: 290, loss: 0.0004456517635844648
step: 300, loss: 0.0003244552353862673
step: 310, loss: 9.787515591597185e-05
step: 320, loss: 0.00020449032308533788
step: 330, loss: 0.0004955333424732089
step: 340, loss: 0.00017262162873521447
step: 350, loss: 0.00020422872330527753
step: 360, loss: 5.900464384467341e-05
step: 370, loss: 0.00022661659750156105
step: 380, loss: 0.0007324364269152284
step: 390, loss: 7.481126522179693e-05
step: 400, loss: 3.168187686242163e-05
step: 410, loss: 0.00010167995060328394
step: 420, loss: 0.0001257256808457896
step: 430, loss: 5.199280712986365e-05
step: 440, loss: 8.173583773896098e-05
step: 450, loss: 5.6416360166622326e-05
step: 460, loss: 9.309010056313127e-05
step: 470, loss: 0.00010676413512555882
step: 480, loss: 0.0002705891092773527
step: 490, loss: 6.688167195534334e-05
step: 500, loss: 0.00014865271805319935
step: 510, loss: 0.00013280462007969618
step: 520, loss: 7.150095188990235e-05
step: 530, loss: 0.0004098078643437475
/usr/local/lib/python3.7/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=0.9278937381404174, f1=0.9312617702448212,
best_f1=0.9410689812699863
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 8.715287549421191e-05
step: 10, loss: 9.434478124603629e-05
step: 20, loss: 0.09378344565629959
step: 30, loss: 0.001240090117789805
step: 40, loss: 0.00022810790687799454
step: 50, loss: 0.04052349179983139
step: 60, loss: 0.0006504969205707312
step: 70, loss: 0.0011054264614358544
step: 80, loss: 6.21793296886608e-05
step: 90, loss: 8.900385000742972e-05
step: 100, loss: 0.0020414297468960285
step: 110, loss: 9.400849376106635e-05
Gradient overflow.  Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 120, loss: 3.076145731029101e-05
step: 130, loss: 0.0011663156328722835
step: 140, loss: 0.0015459296992048621
step: 150, loss: 5.841030724695884e-05
step: 160, loss: 3.74474584532436e-05
step: 170, loss: 3.513900810503401e-05
step: 180, loss: 5.392428647610359e-05
step: 190, loss: 0.00017825327813625336
step: 200, loss: 9.552339179208502e-05
step: 210, loss: 4.0376635297434404e-05
step: 220, loss: 0.00011826519039459527
step: 230, loss: 5.377246634452604e-05
step: 240, loss: 0.0009292103932239115
step: 250, loss: 0.00014564207231160253
step: 260, loss: 0.00017955036310013384
step: 270, loss: 0.00034404045436531305
step: 280, loss: 0.001932558254338801
step: 290, loss: 0.003300898941233754
step: 300, loss: 0.0006401835125871003
step: 310, loss: 9.682479139883071e-05
step: 320, loss: 4.1883620724547654e-05
step: 330, loss: 0.00249828421510756
step: 340, loss: 3.94671515095979e-05
step: 350, loss: 0.00038728577783331275
step: 360, loss: 0.0005422018002718687
step: 370, loss: 0.0007345196790993214
step: 380, loss: 0.00014298023597802967
step: 390, loss: 4.03165613533929e-05
```

```
step: 400, loss: 5.6421828048769385e-05
step: 410, loss: 0.021121995523571968
step: 420, loss: 0.0031543681398034096
step: 430, loss: 0.016106072813272476
step: 440, loss: 0.0003172872238792479
step: 450, loss: 0.0016204708954319358
step: 460, loss: 0.00024159395252354443
step: 470, loss: 5.733309080824256e-05
step: 480, loss: 0.0015614326111972332
step: 490, loss: 0.023828906938433647
step: 500, loss: 0.000101591031125281
step: 510, loss: 0.00015912618255242705
step: 520, loss: 5.20667563250754e-05
step: 530, loss: 0.00030906108440831304
/usr/local/lib/python3.7/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=0.9373246024321795, f1=0.9408476944573824,
best_f1=0.9410689812699863
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.003078039037063718
step: 10, loss: 3.407307303859852e-05
step: 20, loss: 0.00017728732200339437
step: 30, loss: 0.0009323506383225322
step: 40, loss: 8.444044942734763e-05
step: 50, loss: 0.0008697164594195783
step: 60, loss: 8.397047349717468e-05
step: 70, loss: 3.005436883540824e-05
step: 80, loss: 7.207145972643048e-05
step: 90, loss: 0.022530220448970795
step: 100, loss: 0.003009763080626726
step: 110, loss: 5.350053106667474e-05
step: 120, loss: 0.00023589988995809108
step: 130, loss: 2.6378123948234133e-05
step: 140, loss: 0.00016709214833099395
step: 150, loss: 0.0012213286245241761
step: 160, loss: 2.4727851268835366e-05
step: 170, loss: 0.0001020770869217813
step: 180, loss: 0.0043305400758981705
step: 190, loss: 7.227167225209996e-05
step: 200, loss: 5.4147316404851153e-05
step: 210, loss: 0.00020873192988801748
step: 220, loss: 0.005083456635475159
step: 230, loss: 0.005153352860361338
step: 240, loss: 0.039471469819545746
step: 250, loss: 6.836559623479843e-05
```

```
step: 260, loss: 5.5396514653693885e-05
step: 270, loss: 9.463949390919879e-05
step: 280, loss: 5.116142347105779e-05
step: 290, loss: 2.3077423975337297e-05
step: 300, loss: 5.28731252416037e-05
step: 310, loss: 0.0010482431389391422
step: 320, loss: 0.0007026380626484752
step: 330, loss: 0.0012632766738533974
step: 340, loss: 2.4805816792650148e-05
step: 350, loss: 4.843399074161425e-05
step: 360, loss: 2.826304807967972e-05
step: 370, loss: 0.00015595200238749385
step: 380, loss: 0.012285543605685234
step: 390, loss: 0.00020630517974495888
step: 400, loss: 2.1512874809559435e-05
step: 410, loss: 0.00026635723770596087
step: 420, loss: 2.582292836450506e-05
step: 430, loss: 0.002717449329793453
step: 440, loss: 9.85519218374975e-05
step: 450, loss: 9.240825602319092e-05
step: 460, loss: 0.0005987995536997914
step: 470, loss: 2.691622285055928e-05
step: 480, loss: 0.010129563510417938
step: 490, loss: 2.1747742721345276e-05
step: 500, loss: 2.2905449441168457e-05
step: 510, loss: 7.289491622941568e-05
step: 520, loss: 0.0014766944805160165
step: 530, loss: 2.4917353584896773e-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=0.9397031539888682, f1=0.9402023919043238,
best_f1=0.9410689812699863
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00010338669380871579
step: 10, loss: 2.694545401027426e-05
step: 20, loss: 1.785129279596731e-05
step: 30, loss: 6.066533751436509e-05
step: 40, loss: 6.770540494471788e-05
step: 50, loss: 0.0001593145716469735
step: 60, loss: 6.450906221289188e-05
step: 70, loss: 8.423496183240786e-05
step: 80, loss: 0.0053588226437568665
step: 90, loss: 2.543929440435022e-05
step: 100, loss: 0.000984475016593933
step: 110, loss: 7.792904216330498e-05
```

```
step: 120, loss: 6.933875556569546e-05
step: 130, loss: 0.026273835450410843
step: 140, loss: 4.8128244088729843e-05
step: 150, loss: 0.00010321208537789062
step: 160, loss: 9.808482718653977e-05
step: 170, loss: 0.000292964861728251
step: 180, loss: 3.975301297032274e-05
step: 190, loss: 8.363943197764456e-05
step: 200, loss: 0.00016214251809287816
step: 210, loss: 3.003520760103129e-05
step: 220, loss: 3.9170554373413324e-05
step: 230, loss: 9.090275125345215e-05
step: 240, loss: 4.515424006967805e-05
step: 250, loss: 3.202108928235248e-05
step: 260, loss: 3.8929068978177384e-05
step: 270, loss: 1.8045009710476734e-05
step: 280, loss: 2.0566771127050743e-05
step: 290, loss: 0.0013858515303581953
step: 300, loss: 4.397188240545802e-05
step: 310, loss: 0.00010562088573351502
step: 320, loss: 8.67858761921525e-05
step: 330, loss: 7.161818939493969e-05
step: 340, loss: 3.5813183785649016e-05
step: 350, loss: 0.00017018176731653512
step: 360, loss: 5.9659098042175174e-05
step: 370, loss: 3.645984543254599e-05
step: 380, loss: 8.986005559563637e-05
step: 390, loss: 0.022940412163734436
step: 400, loss: 0.0002015558275161311
step: 410, loss: 2.3230155420606025e-05
step: 420, loss: 2.2570859073312022e-05
step: 430, loss: 0.0005448806914500892
step: 440, loss: 3.828256740234792e-05
step: 450, loss: 4.355188139015809e-05
step: 460, loss: 2.545786446717102e-05
step: 470, loss: 2.7994121410301886e-05
step: 480, loss: 1.819769386202097e-05
step: 490, loss: 3.431227742112242e-05
step: 500, loss: 1.639469337533228e-05
step: 510, loss: 0.00039359761285595596
step: 520, loss: 3.0430312108364888e-05
step: 530, loss: 4.371789691504091e-05
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
```

```
  warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 14: dev_f1=0.9398704902867714, f1=0.9372423270728356,
best_f1=0.9372423270728356
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 2.3725495339022018e-05
step: 10, loss: 2.5211797037627548e-05
step: 20, loss: 9.683192183729261e-05
step: 30, loss: 4.4411252019926906e-05
step: 40, loss: 0.03558916971087456
step: 50, loss: 0.00024175795260816813
step: 60, loss: 1.1928250387427397e-05
step: 70, loss: 0.0008486738079227507
step: 80, loss: 2.3293659978662618e-05
step: 90, loss: 2.4075985493254848e-05
step: 100, loss: 1.9132494344376028e-05
step: 110, loss: 3.196854959242046e-05
step: 120, loss: 9.033957758219913e-05
step: 130, loss: 0.04347885400056839
step: 140, loss: 4.221136623527855e-05
step: 150, loss: 5.08632801938802e-05
step: 160, loss: 2.294344267284032e-05
step: 170, loss: 3.1206593121169135e-05
step: 180, loss: 1.722912202239968e-05
step: 190, loss: 3.62341343134176e-05
step: 200, loss: 1.5858322512940504e-05
step: 210, loss: 9.645606041885912e-05
step: 220, loss: 2.7467971449368633e-05
step: 230, loss: 6.133898568805307e-05
step: 240, loss: 2.5252578780055046e-05
step: 250, loss: 0.00015347017324529588
step: 260, loss: 4.7264657041523606e-05
step: 270, loss: 3.979129542130977e-05
step: 280, loss: 3.071616811212152e-05
step: 290, loss: 1.4319820365926716e-05
step: 300, loss: 2.4082883101073094e-05
step: 310, loss: 3.632597508840263e-05
step: 320, loss: 9.934571426128969e-05
step: 330, loss: 6.669331924058497e-05
step: 340, loss: 2.533440783736296e-05
step: 350, loss: 2.2461932530859485e-05
step: 360, loss: 0.0025964719243347645
step: 370, loss: 1.5347988664871082e-05
step: 380, loss: 4.595115751726553e-05
step: 390, loss: 0.000650746573228389
step: 400, loss: 2.1267231204546988e-05
step: 410, loss: 1.8428656403557397e-05
```

```
step: 420, loss: 3.71324822481256e-05
step: 430, loss: 0.0034011222887784243
step: 440, loss: 3.238242425140925e-05
step: 450, loss: 1.536660965939518e-05
step: 460, loss: 1.992991383303888e-05
step: 470, loss: 0.011226708069443703
step: 480, loss: 1.4006892342877109e-05
step: 490, loss: 1.709136995486915e-05
step: 500, loss: 2.9607175747514702e-05
step: 510, loss: 2.9300808819243684e-05
step: 520, loss: 2.2168389477883466e-05
step: 530, loss: 3.3960939617827535e-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=0.9396035039188566, f1=0.9342465753424657,
best_f1=0.9372423270728356
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 [00:30, 189.07it/s]
```

```
load_f1 = 0.9390414146114472
real_f1 = 0.9384902143522833
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
4398it [00:24, 181.83it/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: 0.5498040914535522
step: 10, loss: 0.37779876589775085
step: 20, loss: 0.38229817152023315
step: 30, loss: 0.3414197266101837
step: 40, loss: 0.18779604136943817
step: 50, loss: 0.41652584075927734
step: 60, loss: 0.25878581404685974
step: 70, loss: 0.1763940453529358
step: 80, loss: 0.20959551632404327
step: 90, loss: 0.35307425260543823
step: 100, loss: 0.38716554641723633
step: 110, loss: 0.23845350742340088
step: 120, loss: 0.16744066774845123
step: 130, loss: 0.2685664892196655
step: 140, loss: 0.20945361256599426
step: 150, loss: 0.145716592669487
step: 160, loss: 0.26183387637138367
step: 170, loss: 0.28116127848625183
step: 180, loss: 0.12734976410865784
step: 190, loss: 0.17719824612140656
step: 200, loss: 0.313344806432724
```

```
step: 210, loss: 0.26320314407348633
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.6591760299625468, f1=0.6483300589390962,
best_f1=0.6483300589390962
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.10982842743396759
step: 10, loss: 0.13342329859733582
step: 20, loss: 0.14701780676841736
step: 30, loss: 0.13193215429782867
step: 40, loss: 0.11845387518405914
step: 50, loss: 0.181062251329422
step: 60, loss: 0.3943815231323242
step: 70, loss: 0.10819822549819946
step: 80, loss: 0.1971900910139084
step: 90, loss: 0.10545416921377182
step: 100, loss: 0.01892581395804882
step: 110, loss: 0.05569609999656677
step: 120, loss: 0.12285459041595459
step: 130, loss: 0.014573303051292896
step: 140, loss: 0.1706826090812683
step: 150, loss: 0.16561293601989746
step: 160, loss: 0.21735629439353943
step: 170, loss: 0.11218240857124329
step: 180, loss: 0.1317686289548874
step: 190, loss: 0.13911177217960358
step: 200, loss: 0.03660142049193382
step: 210, loss: 0.10887462645769119
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.7011952191235059, f1=0.6680672268907564,
best_f1=0.6680672268907564
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.051715247333049774
```

```
step: 10, loss: 0.16549478471279144
step: 20, loss: 0.16910479962825775
step: 30, loss: 0.31983110308647156
step: 40, loss: 0.10510765761137009
step: 50, loss: 0.08895600587129593
step: 60, loss: 0.17321735620498657
step: 70, loss: 0.04715650901198387
step: 80, loss: 0.10729236900806427
step: 90, loss: 0.04088016599416733
step: 100, loss: 0.17520888149738312
step: 110, loss: 0.15838564932346344
step: 120, loss: 0.07819822430610657
step: 130, loss: 0.135456845164299
step: 140, loss: 0.10626200586557388
step: 150, loss: 0.24412310123443604
step: 160, loss: 0.009628402069211006
step: 170, loss: 0.15552474558353424
step: 180, loss: 0.09027610719203949
step: 190, loss: 0.23999488353729248
step: 200, loss: 0.036798544228076935
step: 210, loss: 0.11245505511760712
/usr/local/lib/python3.7/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=0.6772908366533865, f1=0.654690618762475,
best_f1=0.6680672268907564
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.06713583320379257
step: 10, loss: 0.075599804520607
step: 20, loss: 0.021679142490029335
step: 30, loss: 0.1861591339111328
step: 40, loss: 0.024266330525279045
step: 50, loss: 0.06476384401321411
step: 60, loss: 0.05434846132993698
step: 70, loss: 0.2342473566532135
step: 80, loss: 0.15543851256370544
step: 90, loss: 0.03498319536447525
step: 100, loss: 0.2507036328315735
step: 110, loss: 0.05115671455860138
step: 120, loss: 0.1289844959974289
step: 130, loss: 0.20006297528743744
step: 140, loss: 0.06593553721904755
step: 150, loss: 0.026915282011032104
step: 160, loss: 0.023576881736516953
step: 170, loss: 0.0850742906332016
step: 180, loss: 0.3225179612636566
```

```
step: 190, loss: 0.04941257834434509
step: 200, loss: 0.2638801038265228
step: 210, loss: 0.07387332618236542
/usr/local/lib/python3.7/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=0.6846153846153845, f1=0.6614481409001958,
best_f1=0.6680672268907564
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.12477241456508636
step: 10, loss: 0.013700373470783234
step: 20, loss: 0.02192738465964794
step: 30, loss: 0.026220425963401794
step: 40, loss: 0.07253073155879974
step: 50, loss: 0.022711487486958504
step: 60, loss: 0.06492961943149567
step: 70, loss: 0.03664252534508705
step: 80, loss: 0.037848763167858124
step: 90, loss: 0.04787055402994156
step: 100, loss: 0.03652197867631912
step: 110, loss: 0.23017889261245728
step: 120, loss: 0.1306942105293274
step: 130, loss: 0.015473827719688416
step: 140, loss: 0.011062911711633205
step: 150, loss: 0.03377193957567215
step: 160, loss: 0.03261227533221245
step: 170, loss: 0.049793973565101624
step: 180, loss: 0.06183294951915741
step: 190, loss: 0.02959929034113884
step: 200, loss: 0.061041805893182755
step: 210, loss: 0.016397293657064438
/usr/local/lib/python3.7/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=0.6772908366533865, f1=0.6129032258064516,
best_f1=0.6680672268907564
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.02899480238556862
step: 10, loss: 0.0462799146771431
step: 20, loss: 0.041634686291217804
step: 30, loss: 0.01650536246597767
step: 40, loss: 0.0050338576547801495
step: 50, loss: 0.005760878790169954
step: 60, loss: 0.09755370020866394
```

```
step: 70, loss: 0.005023146979510784
step: 80, loss: 0.07714322954416275
step: 90, loss: 0.1979682594537735
step: 100, loss: 0.0052132271230220795
step: 110, loss: 0.005486433859914541
step: 120, loss: 0.006620378699153662
step: 130, loss: 0.02381177991628647
step: 140, loss: 0.04324512928724289
step: 150, loss: 0.0040918258018791676
step: 160, loss: 0.01001745369285345
step: 170, loss: 0.14789998531341553
step: 180, loss: 0.06096641346812248
step: 190, loss: 0.05871545895934105
step: 200, loss: 0.003648963291198015
step: 210, loss: 0.030315808951854706
/usr/local/lib/python3.7/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=0.6800804828973844, f1=0.6252587991718427,
best_f1=0.6680672268907564
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.007322218734771013
step: 10, loss: 0.014323972165584564
step: 20, loss: 0.0006835148087702692
step: 30, loss: 0.03851873055100441
step: 40, loss: 0.026293210685253143
step: 50, loss: 0.09103643149137497
step: 60, loss: 0.0027006445452570915
step: 70, loss: 0.007081540767103434
step: 80, loss: 0.04168139770627022
step: 90, loss: 0.05753052234649658
step: 100, loss: 0.001297361683100462
step: 110, loss: 0.11308598518371582
step: 120, loss: 0.027504513040184975
step: 130, loss: 0.010346801951527596
step: 140, loss: 0.004820907022804022
step: 150, loss: 0.014408527873456478
step: 160, loss: 0.0029216280672699213
step: 170, loss: 0.0004466455429792404
step: 180, loss: 0.005313217639923096
step: 190, loss: 0.05610351264476776
step: 200, loss: 0.002115415409207344
step: 210, loss: 0.06633972376585007
/usr/local/lib/python3.7/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=0.6680584551148224, f1=0.6033755274261604,
best_f1=0.6680672268907564
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0035432183649390936
step: 10, loss: 0.018533386290073395
step: 20, loss: 0.003791989991441369
step: 30, loss: 0.013059343211352825
Gradient overflow.  Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 40, loss: 0.19765834510326385
step: 50, loss: 0.0027853474020957947
step: 60, loss: 0.05391194671392441
step: 70, loss: 0.02606750652194023
step: 80, loss: 0.027433903887867928
step: 90, loss: 0.03242642804980278
step: 100, loss: 0.18535497784614563
step: 110, loss: 0.07539872080087662
step: 120, loss: 0.0015699905343353748
step: 130, loss: 0.001093888422474265
step: 140, loss: 0.11796796321868896
step: 150, loss: 0.0043112123385071754
step: 160, loss: 0.006786564365029335
step: 170, loss: 0.0022955809254199266
step: 180, loss: 0.021662726998329163
step: 190, loss: 0.006911387667059898
step: 200, loss: 0.001327352598309517
step: 210, loss: 0.16484519839286804
/usr/local/lib/python3.7/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=0.6968503937007874, f1=0.6153846153846154,
best_f1=0.6680672268907564
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.011962044052779675
step: 10, loss: 0.004652380011975765
step: 20, loss: 0.005024394486099482
step: 30, loss: 0.03373245894908905
step: 40, loss: 0.04232005402445793
step: 50, loss: 0.09055998176336288
step: 60, loss: 0.053642094135284424
step: 70, loss: 0.0006138205644674599
step: 80, loss: 0.002378253499045968
step: 90, loss: 0.00047785439528524876
step: 100, loss: 0.0004901181673631072
step: 110, loss: 0.0051302118226885796
```

```
step: 120, loss: 0.002896262798458338
step: 130, loss: 0.009339659474790096
step: 140, loss: 0.10493410378694534
step: 150, loss: 0.05260748788714409
step: 160, loss: 0.0021151998080313206
step: 170, loss: 0.038595590740442276
step: 180, loss: 0.0069406661204993725
step: 190, loss: 0.000325158704072237
step: 200, loss: 0.019184213131666183
step: 210, loss: 0.0017459879163652658
/usr/local/lib/python3.7/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=0.6796116504854368, f1=0.6144814090019569,
best_f1=0.6680672268907564
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.03388218581676483
step: 10, loss: 0.009261123836040497
step: 20, loss: 0.00037815660471096635
step: 30, loss: 0.016760891303420067
step: 40, loss: 0.008384027518332005
step: 50, loss: 0.01934649422764778
step: 60, loss: 0.017432112246751785
step: 70, loss: 0.031465794891119
step: 80, loss: 0.07543405890464783
step: 90, loss: 0.08959346264600754
step: 100, loss: 0.01589546911418438
step: 110, loss: 0.0004918609629385173
step: 120, loss: 0.05518413707613945
step: 130, loss: 0.0304089467972517
step: 140, loss: 0.02626798301935196
step: 150, loss: 0.03358534723520279
step: 160, loss: 0.0024544766638427973
step: 170, loss: 0.2520368695259094
step: 180, loss: 0.09644661098718643
step: 190, loss: 0.04739939793944359
step: 200, loss: 0.008076646365225315
step: 210, loss: 0.056075628846883774
/usr/local/lib/python3.7/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=0.6774847870182557, f1=0.6074380165289256,
best_f1=0.6680672268907564
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
```

```
step: 0, loss: 0.030870752409100533
step: 10, loss: 0.027170099318027496
step: 20, loss: 0.043015483766794205
step: 30, loss: 0.03936830535531044
step: 40, loss: 0.0038599062245339155
step: 50, loss: 0.048672426491975784
step: 60, loss: 0.0253412164747715
step: 70, loss: 0.0004627077141776681
step: 80, loss: 0.0449187345802784
step: 90, loss: 0.010423469357192516
step: 100, loss: 0.025668328627943993
step: 110, loss: 0.024129338562488556
step: 120, loss: 0.03211308643221855
step: 130, loss: 0.0059243240393698215
step: 140, loss: 0.0015507632633671165
step: 150, loss: 0.0004797623842023313
step: 160, loss: 0.0015229217242449522
step: 170, loss: 0.012950914911925793
step: 180, loss: 0.006541495211422443
step: 190, loss: 0.0005177112761884928
step: 200, loss: 0.0006364524597302079
step: 210, loss: 0.0004995209746994078
/usr/local/lib/python3.7/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=0.6734279918864097, f1=0.6050420168067228,
best_f1=0.6680672268907564
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.01787453331053257
step: 10, loss: 0.000297430349746719
step: 20, loss: 0.014543828554451466
step: 30, loss: 0.0006560804322361946
step: 40, loss: 0.010154885239899158
step: 50, loss: 0.0003584440564736724
step: 60, loss: 0.0005361092044040561
step: 70, loss: 0.009763163514435291
step: 80, loss: 0.08859609812498093
step: 90, loss: 0.001710523501969874
step: 100, loss: 0.0032056679483503103
step: 110, loss: 0.0016232574125751853
step: 120, loss: 0.0006485140766017139
step: 130, loss: 0.000695940398145467
step: 140, loss: 0.0002820062218233943
step: 150, loss: 0.012451490387320518
step: 160, loss: 0.0007124710828065872
step: 170, loss: 0.002036537742242217
```

```
step: 180, loss: 0.00021134468261152506
step: 190, loss: 0.00033743135281838477
step: 200, loss: 0.0005207629292272031
step: 210, loss: 0.02219071239233017
/usr/local/lib/python3.7/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=0.6553911205073996, f1=0.6004319654427647,
best_f1=0.6680672268907564
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.027350934222340584
step: 10, loss: 0.00015084302867762744
step: 20, loss: 0.009778988547623158
step: 30, loss: 0.009768380783498287
step: 40, loss: 0.005846688523888588
step: 50, loss: 0.012270929291844368
step: 60, loss: 0.0002772006264422089
step: 70, loss: 0.062045011669397354
step: 80, loss: 0.0011698288144543767
step: 90, loss: 0.0002144354220945388
step: 100, loss: 0.000143104451126419
step: 110, loss: 0.0003680475929286331
step: 120, loss: 0.0009081756579689682
step: 130, loss: 0.0014401610242202878
step: 140, loss: 0.0013333751121535897
step: 150, loss: 0.00026402398361824453
step: 160, loss: 0.013614602386951447
step: 170, loss: 0.0002712707791943103
step: 180, loss: 0.04009220376610756
step: 190, loss: 0.0002727325481828302
step: 200, loss: 0.0002865654823835939
step: 210, loss: 0.020165981724858284
/usr/local/lib/python3.7/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=0.6595289079229123, f1=0.6167400881057269,
best_f1=0.6680672268907564
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 9.801443957258016e-05
step: 10, loss: 0.021000105887651443
step: 20, loss: 0.0002562413574196398
step: 30, loss: 0.00011183074093423784
step: 40, loss: 0.00024280213983729482
step: 50, loss: 0.010557850822806358
```

```
step: 60, loss: 0.00020792809664271772
step: 70, loss: 0.00025280346744693816
step: 80, loss: 0.0001090197401936166
step: 90, loss: 0.00023681654420215636
step: 100, loss: 0.006776096299290657
step: 110, loss: 0.0003539700119290501
step: 120, loss: 0.03638570010662079
step: 130, loss: 0.00038108666194602847
step: 140, loss: 0.0059177023358643055
step: 150, loss: 0.00020290643442422152
step: 160, loss: 0.0033830299507826567
step: 170, loss: 0.00022869330132380128
step: 180, loss: 0.0013140486553311348
step: 190, loss: 0.00027867176686413586
step: 200, loss: 0.0007061630021780729
step: 210, loss: 0.00022713928774464875
/usr/local/lib/python3.7/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=0.6518847006651884, f1=0.6139954853273137,
best_f1=0.6680672268907564
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.000126939412439242
step: 10, loss: 0.00036834401544183491
step: 20, loss: 0.0005817097262479365
step: 30, loss: 0.02152130752801895
step: 40, loss: 0.0007277707918547094
step: 50, loss: 9.797293751034886e-05
step: 60, loss: 0.001901643117889762
step: 70, loss: 0.00031225773273035884
step: 80, loss: 9.062259050551802e-05
step: 90, loss: 0.0006278220098465681
step: 100, loss: 0.00018454385281074792
step: 110, loss: 0.00010881990601774305
step: 120, loss: 0.0001273860689252615
step: 130, loss: 0.00020170514471828938
step: 140, loss: 0.00013003252388443798
step: 150, loss: 7.917888433439657e-05
step: 160, loss: 0.0039586457423865795
step: 170, loss: 0.00011863602412631735
step: 180, loss: 0.0001021783536998555
step: 190, loss: 0.0001640539412619546
step: 200, loss: 0.008513140492141247
step: 210, loss: 0.0004170868778601289
/usr/local/lib/python3.7/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=0.6578366445916115, f1=0.609865470852018,
best_f1=0.6680672268907564
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, 282.15it/s]
load_f1 = 0.7033398821218074
real_f1 = 0.7000000000000001
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, 184.72it/s]
```

2.1.5 Walmart-Amazon - 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/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: 0.5859736204147339
step: 10, loss: 0.3428954482078552
step: 20, loss: 0.2946934401988983
step: 30, loss: 0.433014720678329475
step: 40, loss: 0.41834166646003723
step: 50, loss: 0.33564040064811707
step: 60, loss: 0.2624850869178772
step: 70, loss: 0.25834280252456665
step: 80, loss: 0.2958296239376068
step: 90, loss: 0.2785518765449524
step: 100, loss: 0.3322627544403076
step: 110, loss: 0.5067391991615295
step: 120, loss: 0.09253686666488647
step: 130, loss: 0.11667226999998093
step: 140, loss: 0.06084996089339256
step: 150, loss: 0.10934719443321228
step: 160, loss: 0.08438940346240997
step: 170, loss: 0.20361453294754028
step: 180, loss: 0.02084360457956791
step: 190, loss: 0.23858104646205902
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.6854219948849106, f1=0.7180851063829788,
best_f1=0.7180851063829788
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.28745779395103455
step: 10, loss: 0.05884002894163132
```

```
step: 20, loss: 0.08883176743984222
step: 30, loss: 0.17984971404075623
step: 40, loss: 0.15311773121356964
step: 50, loss: 0.019676411524415016
step: 60, loss: 0.18516451120376587
step: 70, loss: 0.14417676627635956
step: 80, loss: 0.16189737617969513
step: 90, loss: 0.1483948826789856
step: 100, loss: 0.015505686402320862
step: 110, loss: 0.1715383231639862
step: 120, loss: 0.25803011655807495
step: 130, loss: 0.12125241756439209
step: 140, loss: 0.08330117166042328
step: 150, loss: 0.10313200205564499
step: 160, loss: 0.037156011909246445
step: 170, loss: 0.15044814348220825
step: 180, loss: 0.15463685989379883
step: 190, loss: 0.01814812235534191
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.7922437673130193, f1=0.7734806629834254,
best_f1=0.7734806629834254
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.018573680892586708
step: 10, loss: 0.212227463722229
step: 20, loss: 0.06114271283149719
step: 30, loss: 0.13110890984535217
step: 40, loss: 0.12761494517326355
step: 50, loss: 0.10991301387548447
step: 60, loss: 0.012872488237917423
step: 70, loss: 0.0736360028386116
step: 80, loss: 0.09504639357328415
step: 90, loss: 0.0630415603518486
step: 100, loss: 0.0307109784334898
step: 110, loss: 0.014094065874814987
step: 120, loss: 0.15367792546749115
step: 130, loss: 0.013880091719329357
Gradient overflow.  Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 140, loss: 0.05014268681406975
step: 150, loss: 0.1377774327993393
step: 160, loss: 0.026717010885477066
```

```
step: 170, loss: 0.12461995333433151
step: 180, loss: 0.15582408010959625
step: 190, loss: 0.11308372765779495
/usr/local/lib/python3.7/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=0.7909604519774011, f1=0.8, best_f1=0.7734806629834254
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.03597605973482132
step: 10, loss: 0.13167685270309448
step: 20, loss: 0.038504213094711304
step: 30, loss: 0.0111116087064147
step: 40, loss: 0.09904304891824722
step: 50, loss: 0.00945945642888546
step: 60, loss: 0.1637646108865738
step: 70, loss: 0.03114270605146885
step: 80, loss: 0.07868309319019318
step: 90, loss: 0.0094047412276268
step: 100, loss: 0.03383363038301468
step: 110, loss: 0.0015194148290902376
step: 120, loss: 0.008659016340970993
step: 130, loss: 0.08792567998170853
step: 140, loss: 0.056013260036706924
step: 150, loss: 0.008355511352419853
step: 160, loss: 0.01768551766872406
step: 170, loss: 0.10280552506446838
step: 180, loss: 0.01142455730587244
step: 190, loss: 0.17399822175502777
/usr/local/lib/python3.7/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=0.7835616438356164, f1=0.7966101694915253,
best_f1=0.7734806629834254
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.19867181777954102
step: 10, loss: 0.012175059877336025
step: 20, loss: 0.0032583295833319426
step: 30, loss: 0.002666308544576168
step: 40, loss: 0.18350592255592346
step: 50, loss: 0.1428181678056717
step: 60, loss: 0.04832838848233223
step: 70, loss: 0.004698409233242273
step: 80, loss: 0.01396638061851263
step: 90, loss: 0.02447126805782318
```

```
step: 100, loss: 0.0009770295582711697
step: 110, loss: 0.0048856413923203945
step: 120, loss: 0.013841946609318256
step: 130, loss: 0.07971442490816116
step: 140, loss: 0.0554366298019886
step: 150, loss: 0.10102344304323196
step: 160, loss: 0.02849692478775978
step: 170, loss: 0.019994059577584267
step: 180, loss: 0.009950585663318634
step: 190, loss: 0.005170060321688652
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.7957559681697614, 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: 0.006370760500431061
step: 10, loss: 0.005274791270494461
step: 20, loss: 0.002459293929859996
step: 30, loss: 0.00062022494259923697
step: 40, loss: 0.028384145349264145
step: 50, loss: 0.004288894589990377
step: 60, loss: 0.03805924579501152
step: 70, loss: 0.0008120008278638124
step: 80, loss: 0.058616120368242264
step: 90, loss: 0.009955989196896553
step: 100, loss: 0.04781657084822655
step: 110, loss: 0.011715746484696865
step: 120, loss: 0.0033869221806526184
step: 130, loss: 0.002813513157889247
step: 140, loss: 0.035764213651418686
step: 150, loss: 0.0009845277527347207
step: 160, loss: 0.016507573425769806
step: 170, loss: 0.010427790693938732
step: 180, loss: 0.009960522875189781
step: 190, loss: 0.003520769765600562
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.8021390374331551, f1=0.8021680216802168,
best_f1=0.8021680216802168
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.002225299831479788
step: 10, loss: 0.05167379230260849
step: 20, loss: 0.04793612286448479
step: 30, loss: 0.025715947151184082
step: 40, loss: 0.0031139464117586613
step: 50, loss: 0.0657677873969078
step: 60, loss: 0.003557598451152444
step: 70, loss: 0.008377487771213055
step: 80, loss: 0.0030536726117134094
step: 90, loss: 0.0007288334309123456
step: 100, loss: 0.001139584113843739
step: 110, loss: 0.003807490691542625
step: 120, loss: 0.0072202133014798164
step: 130, loss: 0.02816847339272499
step: 140, loss: 0.0026234795805066824
step: 150, loss: 0.021073324605822563
step: 160, loss: 0.1251450926065445
step: 170, loss: 0.028834061697125435
step: 180, loss: 0.004939843434840441
step: 190, loss: 0.0066359941847622395
/usr/local/lib/python3.7/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=0.7842105263157894, f1=0.7924528301886792,
best_f1=0.8021680216802168
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00533713772892952
step: 10, loss: 0.06272028386592865
step: 20, loss: 0.0007554551702924073
step: 30, loss: 0.002356771845370531
step: 40, loss: 0.0011543519794940948
step: 50, loss: 0.0016208733431994915
step: 60, loss: 0.00039274521986953914
step: 70, loss: 0.04078863188624382
step: 80, loss: 9.443463932257146e-05
step: 90, loss: 0.012221350334584713
step: 100, loss: 0.04724647477269173
step: 110, loss: 0.0002531094360165298
step: 120, loss: 0.00023701558529864997
step: 130, loss: 0.0023038447834551334
```

```
step: 140, loss: 0.0016952389851212502
step: 150, loss: 0.0024232324212789536
step: 160, loss: 0.027927376329898834
step: 170, loss: 0.020700234919786453
step: 180, loss: 0.014240289106965065
step: 190, loss: 0.02152181603014469
/usr/local/lib/python3.7/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=0.8010752688172043, f1=0.7967032967032966,
best_f1=0.8021680216802168
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00402098661288619
step: 10, loss: 0.04860740154981613
step: 20, loss: 0.0015509355580434203
step: 30, loss: 0.0014875751221552491
step: 40, loss: 0.020751172676682472
step: 50, loss: 0.00068357627606601938
step: 60, loss: 0.0003846239997074008
step: 70, loss: 0.00034624221734702587
step: 80, loss: 0.000678381125908345
step: 90, loss: 0.024576453492045403
step: 100, loss: 0.014834805391728878
step: 110, loss: 0.00062365282792559801
step: 120, loss: 0.003241828177124262
step: 130, loss: 0.0030643485952168703
step: 140, loss: 0.08578767627477646
step: 150, loss: 0.007794082164764404
step: 160, loss: 0.00046220814692787826
step: 170, loss: 0.0007771681412123144
step: 180, loss: 0.0016957083716988564
step: 190, loss: 0.0005402928218245506
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.8074866310160429, f1=0.7956403269754768,
best_f1=0.7956403269754768
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0006374032236635685
step: 10, loss: 0.0005183526081964374
```

```
step: 20, loss: 0.018912658095359802
step: 30, loss: 0.000138548479299061
step: 40, loss: 0.0023765156511217356
step: 50, loss: 0.0002606858324725181
step: 60, loss: 0.0007253789808601141
step: 70, loss: 0.017807183787226677
step: 80, loss: 0.0002637276193127036
step: 90, loss: 0.00036066595930606127
step: 100, loss: 0.00020556421077344567
step: 110, loss: 0.02251235954463482
step: 120, loss: 0.10617945343255997
step: 130, loss: 0.001159132458269596
step: 140, loss: 0.0010145248379558325
step: 150, loss: 0.005065500736236572
step: 160, loss: 0.0026203615125268698
step: 170, loss: 0.0003117255982942879
step: 180, loss: 0.00020492062321864069
step: 190, loss: 0.00046443974133580923
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.8099173553719008, f1=0.8022922636103152,
best_f1=0.8022922636103152
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00027461853460408747
step: 10, loss: 0.0005490813055075705
step: 20, loss: 0.019681427627801895
step: 30, loss: 0.008612344972789288
step: 40, loss: 0.00019716337556019425
step: 50, loss: 0.029302814975380898
step: 60, loss: 0.00015984497440513223
step: 70, loss: 0.00020770668925251812
step: 80, loss: 0.00045287629473023117
step: 90, loss: 0.0003771288029383868
step: 100, loss: 0.0002389194123679772
step: 110, loss: 0.00017561452114023268
step: 120, loss: 0.0002940602716989815
step: 130, loss: 0.00024868943728506565
step: 140, loss: 0.00023980015248525888
step: 150, loss: 0.00013402572949416935
step: 160, loss: 0.0002629086666274816
step: 170, loss: 0.0004309725482016802
```

```
step: 180, loss: 0.000199835718376562
step: 190, loss: 0.000520171714015305
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.8209366391184573, f1=0.7899159663865546,
best_f1=0.7899159663865546
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00013041688362136483
step: 10, loss: 0.0010635816724970937
step: 20, loss: 0.00014599162386730313
step: 30, loss: 0.0022714927326887846
step: 40, loss: 0.0012977526057511568
step: 50, loss: 0.000675910385325551
step: 60, loss: 0.00015921150043141097
step: 70, loss: 0.00028983468655496836
step: 80, loss: 0.0007728604832664132
step: 90, loss: 0.0003312437329441309
step: 100, loss: 0.00016875023720785975
step: 110, loss: 0.00019528280245140195
step: 120, loss: 0.0002402004029136151
step: 130, loss: 0.00020311030675657094
step: 140, loss: 0.0016669981414452195
step: 150, loss: 0.00040724710561335087
step: 160, loss: 8.47812116262503e-05
step: 170, loss: 0.00035625617601908743
step: 180, loss: 8.376971527468413e-05
step: 190, loss: 9.884404425974935e-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=0.8123249299719888, f1=0.792022792022792,
best_f1=0.7899159663865546
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0007329173968173563
step: 10, loss: 0.00015976946451701224
step: 20, loss: 0.0004768257786054164
step: 30, loss: 0.12744872272014618
step: 40, loss: 0.00032068986911326647
step: 50, loss: 0.0005077358218841255
```

```
step: 60, loss: 0.0021129180677235126
step: 70, loss: 0.004465129692107439
step: 80, loss: 0.0003363546566106379
step: 90, loss: 0.0008894117781892419
step: 100, loss: 0.0031816211994737387
step: 110, loss: 0.00015154611901380122
step: 120, loss: 0.003956225235015154
step: 130, loss: 0.00044945275294594467
step: 140, loss: 0.00018602071213535964
step: 150, loss: 0.00014852467575110495
step: 160, loss: 0.0002255547879030928
step: 170, loss: 0.00019994605099782348
step: 180, loss: 0.00013045116793364286
step: 190, loss: 0.0027023274451494217
/usr/local/lib/python3.7/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=0.8140161725067385, f1=0.8213333333333334,
best_f1=0.7899159663865546
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.03490752726793289
step: 10, loss: 0.0001667511824052781
step: 20, loss: 0.00015734988846816123
step: 30, loss: 0.0006013228558003902
step: 40, loss: 0.00020705924544017762
step: 50, loss: 0.00016099157801363617
step: 60, loss: 0.0004992168978787959
step: 70, loss: 0.00012868507474195212
step: 80, loss: 9.169010445475578e-05
step: 90, loss: 0.0003658454224932939
step: 100, loss: 0.0025857435539364815
step: 110, loss: 0.00022971593716647476
step: 120, loss: 0.0005168535863049328
step: 130, loss: 0.00016175827477127314
step: 140, loss: 0.0003660922811832279
step: 150, loss: 9.622486686566845e-05
step: 160, loss: 0.0002807800192385912
step: 170, loss: 0.000169265505974181
step: 180, loss: 0.00021740405645687133
step: 190, loss: 0.00011435618216637522
/usr/local/lib/python3.7/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=0.8138297872340426, f1=0.8085106382978724,
best_f1=0.7899159663865546
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
```

An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0019315803656354547
step: 10, loss: 8.756924944464117e-05
step: 20, loss: 0.0002174217370338738
step: 30, loss: 9.26963402889669e-05
step: 40, loss: 0.0001130435848608613
step: 50, loss: 0.00011768618423957378
step: 60, loss: 9.38364610192366e-05
step: 70, loss: 0.005617887247353792
step: 80, loss: 0.0003678637440316379
step: 90, loss: 8.479381358483806e-05
step: 100, loss: 0.00020203863095957786
step: 110, loss: 0.00018982034816872329
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 120, loss: 0.0001628799655009061
step: 130, loss: 0.010208900086581707
step: 140, loss: 0.0014865404227748513
step: 150, loss: 0.00015308607544284314
step: 160, loss: 0.00014818692579865456
step: 170, loss: 0.00013035429583396763
step: 180, loss: 0.0034798337146639824
step: 190, loss: 0.00012367511226329952
/usr/local/lib/python3.7/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=0.8121546961325967, f1=0.8055555555555556,
best_f1=0.7899159663865546
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, 165.45it/s]
load_f1 = 0.8186813186813187
real_f1 = 0.8209366391184573
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, 185.29it/s]
```

2.2 BASELINE TEXTUAL

2.2.1 Abt-Buy - Running the matcher

```
[ ]: # Step 1:
     # train the wdc model with the save_model flag on
     # it will produce two files *_dev.pt and *_test.pt which are the
     # best checkpoints on the dev set and the test set.
     !CUDA_VISIBLE_DEVICES=0 python train_ditto.py \
       --task Textual/Abt-Buy \
       --batch_size 32 \
       --max_len 256 \
       --lr 3e-5 \
       --n_epochs 15 \
       --finetuning \
       --lm 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: 0.608964741230011
step: 10, loss: 0.37176182866096497
step: 20, loss: 0.3199981451034546
step: 30, loss: 0.4164694845676422
step: 40, loss: 0.28772231936454773
step: 50, loss: 0.2867978513240814
step: 60, loss: 0.22136126458644867

```
step: 70, loss: 0.37581419944763184
step: 80, loss: 0.33087587356567383
step: 90, loss: 0.278876930475235
Gradient overflow.  Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 100, loss: 0.18068687617778778
step: 110, loss: 0.27723193168640137
step: 120, loss: 0.22905020415782928
step: 130, loss: 0.07581739127635956
step: 140, loss: 0.2111777514219284
step: 150, loss: 0.25234454870224
step: 160, loss: 0.14972996711730957
step: 170, loss: 0.17359328269958496
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.7277227722772277, f1=0.7548076923076923,
best_f1=0.7548076923076923
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.14562007784843445
step: 10, loss: 0.1286705583333969
step: 20, loss: 0.0883573591709137
step: 30, loss: 0.29232487082481384
step: 40, loss: 0.06372775882482529
step: 50, loss: 0.1989271491765976
step: 60, loss: 0.10812437534332275
step: 70, loss: 0.1277197003364563
step: 80, loss: 0.02560921385884285
step: 90, loss: 0.04065953940153122
step: 100, loss: 0.11381419003009796
step: 110, loss: 0.0788901224732399
step: 120, loss: 0.300593763589859
step: 130, loss: 0.10390488803386688
step: 140, loss: 0.2110726684331894
step: 150, loss: 0.14764107763767242
step: 160, loss: 0.08441758900880814
step: 170, loss: 0.050237830728292465
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.787128712871287, f1=0.7838479809976248,
best_f1=0.7838479809976248
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0796331912279129
step: 10, loss: 0.023827265948057175
step: 20, loss: 0.07452094554901123
step: 30, loss: 0.24123752117156982
step: 40, loss: 0.07801631838083267
step: 50, loss: 0.06345465779304504
step: 60, loss: 0.10414409637451172
step: 70, loss: 0.08429516851902008
step: 80, loss: 0.057208701968193054
step: 90, loss: 0.15717074275016785
step: 100, loss: 0.007543105632066727
step: 110, loss: 0.034657176584005356
step: 120, loss: 0.0595102459192276
step: 130, loss: 0.16340163350105286
step: 140, loss: 0.09864652901887894
step: 150, loss: 0.1057158038020134
step: 160, loss: 0.07387940585613251
step: 170, loss: 0.22709529101848602
/usr/local/lib/python3.7/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=0.759493670886076, f1=0.7707317073170732,
best_f1=0.7838479809976248
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.018930504098534584
step: 10, loss: 0.06934985518455505
step: 20, loss: 0.02617844007909298
step: 30, loss: 0.12421999871730804
step: 40, loss: 0.0014624915784224868
step: 50, loss: 0.08205651491880417
step: 60, loss: 0.137105330824852
step: 70, loss: 0.002888160292059183
step: 80, loss: 0.10138147324323654
step: 90, loss: 0.1536918580532074
step: 100, loss: 0.16663600504398346
step: 110, loss: 0.09017772972583771
step: 120, loss: 0.0063294279389083385
step: 130, loss: 0.017151163890957832
step: 140, loss: 0.1006346195936203
step: 150, loss: 0.11568029224872589
```

```
step: 160, loss: 0.1979387253522873
step: 170, loss: 0.08775162696838379
/usr/local/lib/python3.7/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=0.7727272727272727, f1=0.8057553956834533,
best_f1=0.7838479809976248
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.04783746227622032
step: 10, loss: 0.005890728905797005
step: 20, loss: 0.01631157286465168
step: 30, loss: 0.06524590402841568
step: 40, loss: 0.0626714676618576
step: 50, loss: 0.012689689174294472
step: 60, loss: 0.0017868013819679618
step: 70, loss: 0.08367787301540375
step: 80, loss: 0.10623511672019958
step: 90, loss: 0.17485609650611877
step: 100, loss: 0.05752899870276451
step: 110, loss: 0.15500618517398834
step: 120, loss: 0.0127715440467
step: 130, loss: 0.043304286897182465
step: 140, loss: 0.02446885034441948
step: 150, loss: 0.13533082604408264
step: 160, loss: 0.05049283802509308
step: 170, loss: 0.01912762224674225
/usr/local/lib/python3.7/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=0.7655502392344498, f1=0.7916666666666666,
best_f1=0.7838479809976248
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.007921732030808926
step: 10, loss: 0.04073882848024368
step: 20, loss: 0.003280221950262785
step: 30, loss: 0.02998259849846363
step: 40, loss: 0.00848080962896347
step: 50, loss: 0.11534682661294937
step: 60, loss: 0.006901330780237913
step: 70, loss: 0.05241763964295387
step: 80, loss: 0.02519953064620495
step: 90, loss: 0.012700876221060753
step: 100, loss: 0.0013951101573184133
step: 110, loss: 0.004749370738863945
```

```
step: 120, loss: 0.01476444024592638
step: 130, loss: 0.03268996998667717
step: 140, loss: 0.03321652486920357
step: 150, loss: 0.07916546612977982
step: 160, loss: 0.04364090785384178
step: 170, loss: 0.00098451878875494
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.7893462469733656, f1=0.7731481481481483,
best_f1=0.7731481481481483
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0004011881537735462
step: 10, loss: 0.0006643960950896144
step: 20, loss: 0.004081712104380131
step: 30, loss: 0.01852068305015564
step: 40, loss: 0.13936731219291687
step: 50, loss: 0.034923139959573746
step: 60, loss: 0.0017194735119119287
step: 70, loss: 0.0005021622055210173
step: 80, loss: 0.01133015938103199
step: 90, loss: 0.0003994764410890639
step: 100, loss: 0.002827031072229147
step: 110, loss: 0.004791948478668928
step: 120, loss: 0.0039197830483317375
step: 130, loss: 0.11319585889577866
step: 140, loss: 0.056114841252565384
step: 150, loss: 0.01219034381210804
step: 160, loss: 0.0003668120480142534
step: 170, loss: 0.0655570849776268
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.7922705314009661, f1=0.7915690866510539,
best_f1=0.7915690866510539
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
```

```
step: 0, loss: 0.0009519907762296498
step: 10, loss: 0.0015598530881106853
step: 20, loss: 0.0002720303018577397
step: 30, loss: 0.052945103496313095
step: 40, loss: 0.00025824218755587935
step: 50, loss: 0.006406262516975403
step: 60, loss: 0.0010288541670888662
step: 70, loss: 0.0006217499612830579
step: 80, loss: 0.008118638768792152
step: 90, loss: 0.0005842589307576418
step: 100, loss: 0.0517183393239975
step: 110, loss: 0.05034938082098961
step: 120, loss: 0.0016124387038871646
step: 130, loss: 0.00637862179428339
step: 140, loss: 0.0008440644014626741
step: 150, loss: 0.021130423992872238
step: 160, loss: 0.003330261679366231
step: 170, loss: 0.007644359953701496
/usr/local/lib/python3.7/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=0.7777777777777778, f1=0.771551724137931,
best_f1=0.7915690866510539
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.005383950192481279
step: 10, loss: 0.00498559046536684
step: 20, loss: 0.0003322017437312752
step: 30, loss: 0.014244929887354374
step: 40, loss: 0.006013170350342989
step: 50, loss: 0.00029076822102069855
step: 60, loss: 0.04947999492287636
step: 70, loss: 0.018313266336917877
step: 80, loss: 0.00035656432737596333
step: 90, loss: 0.03964582458138466
step: 100, loss: 0.026357725262641907
step: 110, loss: 0.0005103240837343037
step: 120, loss: 0.001684972201474011
step: 130, loss: 0.08439728617668152
step: 140, loss: 0.0006876651314087212
step: 150, loss: 0.06488781422376633
step: 160, loss: 0.02588062174618244
step: 170, loss: 0.006827720906585455
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
```

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=0.8058252427184465, f1=0.8, best_f1=0.8
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.055033762007951736
step: 10, loss: 0.0001908992271637544
step: 20, loss: 0.043274443596601486
step: 30, loss: 0.009050547145307064
step: 40, loss: 0.003417401108890772
step: 50, loss: 0.06796982139348984
step: 60, loss: 0.00025074472068808973
step: 70, loss: 0.009357758797705173
step: 80, loss: 0.001794614945538342
step: 90, loss: 0.008849983103573322
step: 100, loss: 0.000156313632032834
step: 110, loss: 0.0014616578118875623
step: 120, loss: 0.0001701415458228439
step: 130, loss: 0.001973760314285755
step: 140, loss: 0.24192912876605988
step: 150, loss: 0.001482749474234879
step: 160, loss: 0.0010487547842785716
step: 170, loss: 0.00002758832124527544
/usr/local/lib/python3.7/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=0.7862407862407862, f1=0.7842227378190254, best_f1=0.8
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.05938039720058441
step: 10, loss: 0.02079075574874878
step: 20, loss: 0.00015949048975016922
step: 30, loss: 0.00014010693121235818
step: 40, loss: 0.0056085046380758286
step: 50, loss: 0.000360076199285686
step: 60, loss: 0.008629819378256798
step: 70, loss: 9.478355059400201e-05
step: 80, loss: 0.00010693569493014365
step: 90, loss: 0.0001338091678917408
step: 100, loss: 0.0005566537147387862
step: 110, loss: 0.2556253969669342
step: 120, loss: 0.00016643029812257737
step: 130, loss: 0.0001386226067552343
step: 140, loss: 0.0002039999671978876
step: 150, loss: 0.002087774919345975

```
step: 160, loss: 0.005788686219602823
step: 170, loss: 0.00045702431816607714
/usr/local/lib/python3.7/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=0.7480106100795756, f1=0.7892156862745098, best_f1=0.8
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.061189230531454086
step: 10, loss: 0.0005175832775421441
step: 20, loss: 0.00024766632122918963
step: 30, loss: 0.00024171810946427286
step: 40, loss: 0.001167385489679873
step: 50, loss: 0.0001378507586196065
step: 60, loss: 0.0030700028873980045
step: 70, loss: 0.06890933960676193
step: 80, loss: 0.0001368656085105613
step: 90, loss: 0.0003460514999460429
step: 100, loss: 0.0012906816555187106
Gradient overflow.  Skipping step, loss scaler 0 reducing loss scale to 8192.0
step: 110, loss: 0.00019505656382534653
step: 120, loss: 0.0012233239831402898
step: 130, loss: 0.0069166929461061954
step: 140, loss: 0.01622786559164524
step: 150, loss: 0.004151287488639355
step: 160, loss: 0.00024656258756294847
step: 170, loss: 0.00016504191444255412
/usr/local/lib/python3.7/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=0.7851851851851852, f1=0.7972027972027972, best_f1=0.8
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.018391720950603485
step: 10, loss: 0.0002558293635956943
step: 20, loss: 0.0015270088333636522
step: 30, loss: 0.00020974595099687576
step: 40, loss: 0.00027680417406372726
step: 50, loss: 0.00014268673839978874
step: 60, loss: 0.020870542153716087
step: 70, loss: 0.00027614986174739897
step: 80, loss: 0.003796263597905636
step: 90, loss: 0.0002120194403687492
step: 100, loss: 0.004123339429497719
step: 110, loss: 0.00012425618479028344
step: 120, loss: 0.030836785212159157
```

```
step: 130, loss: 0.00018400026601739228
step: 140, loss: 0.0006229413556866348
step: 150, loss: 0.00036146712955087423
step: 160, loss: 0.0033518278505653143
step: 170, loss: 0.02959415130317211
/usr/local/lib/python3.7/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=0.766839378238342, f1=0.784503631961259, best_f1=0.8
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0002853111072909087
step: 10, loss: 0.00013426702935248613
step: 20, loss: 0.00026332237757742405
step: 30, loss: 0.00020904367556795478
step: 40, loss: 0.000625958142336458
step: 50, loss: 0.0002688475651666522
step: 60, loss: 0.00012440900900401175
step: 70, loss: 0.0003121493791695684
step: 80, loss: 0.00035766520886681974
step: 90, loss: 0.00011240359890507534
step: 100, loss: 0.00011405926488805562
step: 110, loss: 0.00018545257626101375
step: 120, loss: 0.02293545939028263
step: 130, loss: 0.003294400405138731
step: 140, loss: 0.00030346852145157754
step: 150, loss: 0.00012068993237335235
step: 160, loss: 6.799866241635755e-05
step: 170, loss: 0.00013449552352540195
/usr/local/lib/python3.7/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=0.7688172043010754, f1=0.7786259541984734, best_f1=0.8
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00015947918291203678
step: 10, loss: 0.003735373029485345
step: 20, loss: 0.010113372467458248
step: 30, loss: 0.0003142955538351089
step: 40, loss: 0.000754607142880559
step: 50, loss: 0.00033520793658681214
step: 60, loss: 0.018829345703125
step: 70, loss: 0.00012034034443786368
step: 80, loss: 0.047195374965667725
step: 90, loss: 0.00017117975221481174
step: 100, loss: 0.001282449346035719
```

```
step: 110, loss: 0.00023159824195317924
step: 120, loss: 0.00027653115103021264
step: 130, loss: 0.005107981152832508
step: 140, loss: 0.0001711601798888296
step: 150, loss: 0.0012963417684659362
step: 160, loss: 0.009221234358847141
step: 170, loss: 0.0032733932603150606
/usr/local/lib/python3.7/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=0.7774936061381074, f1=0.7865707434052758, best_f1=0.8
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, 270.25it/s]
load_f1 = 0.8038740920096852
real_f1 = 0.8048780487804877
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, 248.58it/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: 0.6260179877281189
step: 10, loss: 0.6272473335266113
step: 20, loss: 0.445242315530777
step: 30, loss: 0.170008584856987
step: 40, loss: 0.35500437021255493
step: 50, loss: 0.07746218889951706
step: 60, loss: 0.11993733048439026
step: 70, loss: 0.07241245359182358
step: 80, loss: 0.063335959583520889
step: 90, loss: 0.05211516469717026
step: 100, loss: 0.005536961834877729
step: 110, loss: 0.28452956676483154
step: 120, loss: 0.027704857289791107
step: 130, loss: 0.0071223219856619835
step: 140, loss: 0.002566908486187458
step: 150, loss: 0.016703780740499496
step: 160, loss: 0.006554034072905779
step: 170, loss: 0.08134578913450241
step: 180, loss: 0.10202956944704056
step: 190, loss: 0.0717902258085208
step: 200, loss: 0.08535026758909225
step: 210, loss: 0.005083436146378517
step: 220, loss: 0.008666153065860271
step: 230, loss: 0.03337440267205238
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
```

```
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9678848283499446,
best_f1=0.9678848283499446
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.007230783347040415
step: 10, loss: 0.004978269804269075
step: 20, loss: 0.18062807619571686
step: 30, loss: 0.17280998826026917
step: 40, loss: 0.035905566066503525
step: 50, loss: 0.004077078774571419
step: 60, loss: 0.016993612051010132
step: 70, loss: 0.10806027799844742
step: 80, loss: 0.01940910518169403
step: 90, loss: 0.011914247646927834
step: 100, loss: 0.05414203926920891
step: 110, loss: 0.03771830350160599
step: 120, loss: 0.18400701880455017
step: 130, loss: 0.01298193447291851
step: 140, loss: 0.002112565329298377
step: 150, loss: 0.04937394708395004
step: 160, loss: 0.06402704119682312
step: 170, loss: 0.00164191541261971
step: 180, loss: 0.010511542670428753
step: 190, loss: 0.0072831157594919205
step: 200, loss: 0.00651921471580863
step: 210, loss: 0.0007690017228014767
step: 220, loss: 0.0840553417801857
step: 230, loss: 0.009204384870827198
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9704545454545453, f1=0.9679633867276889,
best_f1=0.9679633867276889
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.012899014167487621
step: 10, loss: 0.006694504991173744
step: 20, loss: 0.003770081792026758
```

```
step: 30, loss: 0.004864119924604893
step: 40, loss: 0.021142439916729927
step: 50, loss: 0.011000009253621101
step: 60, loss: 0.002406420186161995
step: 70, loss: 0.005648389924317598
step: 80, loss: 0.0007329056970775127
step: 90, loss: 0.07157924771308899
step: 100, loss: 0.002768379868939519
step: 110, loss: 0.006050541531294584
step: 120, loss: 0.025850793346762657
step: 130, loss: 0.0007437097374349833
step: 140, loss: 0.0024998243898153305
step: 150, loss: 0.03811869025230408
step: 160, loss: 0.02770012617111206
step: 170, loss: 0.012889519333839417
step: 180, loss: 0.007776603568345308
step: 190, loss: 0.012026198208332062
step: 200, loss: 0.06482347846031189
step: 210, loss: 0.01595143787562847
step: 220, loss: 0.04604296386241913
step: 230, loss: 0.14179947972297668
/usr/local/lib/python3.7/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=0.96875, f1=0.9720670391061451, best_f1=0.9679633867276889
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0005860367673449218
step: 10, loss: 0.00163787673227489
step: 20, loss: 0.0032016332261264324
step: 30, loss: 0.0026261257007718086
step: 40, loss: 0.0030815950594842434
step: 50, loss: 0.0024590229149907827
step: 60, loss: 0.002680039033293724
step: 70, loss: 0.0006926798960193992
step: 80, loss: 0.002603384433314204
step: 90, loss: 0.0018391460180282593
step: 100, loss: 0.0481957271695137
step: 110, loss: 0.00020756623416673392
step: 120, loss: 0.0015934703405946493
step: 130, loss: 0.03922612592577934
step: 140, loss: 0.001076463726349175
step: 150, loss: 0.19978973269462585
step: 160, loss: 0.004385118838399649
Gradient overflow.  Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 170, loss: 0.013684453442692757
step: 180, loss: 0.000817402615211904
```

```
step: 190, loss: 0.026035849004983902
step: 200, loss: 0.0008642776519991457
step: 210, loss: 0.0762990191578865
step: 220, loss: 0.0015179385663941503
step: 230, loss: 0.009694868698716164
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9752252252252253, f1=0.9728506787330317,
best_f1=0.9728506787330317
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.007991170510649681
step: 10, loss: 0.001594020170159638
step: 20, loss: 0.006718166172504425
step: 30, loss: 0.00025230227038264275
step: 40, loss: 0.0006233867025002837
step: 50, loss: 0.0009749357705004513
step: 60, loss: 0.0005166685441508889
step: 70, loss: 0.00024538743309676647
step: 80, loss: 0.00021962950995657593
step: 90, loss: 0.0004161326796747744
step: 100, loss: 0.001772893825545907
step: 110, loss: 0.019300969317555428
step: 120, loss: 0.00031146014225669205
step: 130, loss: 0.0007190472679212689
step: 140, loss: 0.0008780306670814753
step: 150, loss: 0.032319940626621246
step: 160, loss: 0.0005585881299339235
step: 170, loss: 0.04558049142360687
step: 180, loss: 0.006686880718916655
step: 190, loss: 0.10468197613954544
step: 200, loss: 0.00712910620495677
step: 210, loss: 0.006558018736541271
step: 220, loss: 0.0010279357666149735
step: 230, loss: 0.00040875139529816806
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9773242630385486, f1=0.9772727272727272,
best_f1=0.9772727272727272
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0014298806199803948
step: 10, loss: 0.02173979952931404
step: 20, loss: 0.006778893060982227
step: 30, loss: 0.002280393848195672
step: 40, loss: 0.00043857592390850186
step: 50, loss: 0.0008281274931505322
step: 60, loss: 0.00032135151559486985
step: 70, loss: 0.011406105011701584
step: 80, loss: 0.003572834189981222
step: 90, loss: 0.006148310843855143
step: 100, loss: 0.008729402907192707
step: 110, loss: 0.000654039322398603
step: 120, loss: 0.0007131243473850191
step: 130, loss: 0.0005553795490413904
step: 140, loss: 0.00027569313533604145
step: 150, loss: 0.004642079584300518
step: 160, loss: 0.0010424255160614848
step: 170, loss: 0.00023608254559803754
step: 180, loss: 0.000924255873542279
step: 190, loss: 0.0011449273442849517
step: 200, loss: 0.0001617792440811172
step: 210, loss: 0.0002572905214037746
step: 220, loss: 0.00020052920444868505
step: 230, loss: 0.06125007942318916
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9808773903262092, f1=0.9773755656108598,
best_f1=0.9773755656108598
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0005605787155218422
step: 10, loss: 0.00011290294060017914
step: 20, loss: 0.0002305156522197649
step: 30, loss: 0.00010948735143756494
step: 40, loss: 0.0001618025853531435
step: 50, loss: 0.00010472418216522783
step: 60, loss: 0.0010376794962212443
```

```
step: 70, loss: 0.00016857351874932647
step: 80, loss: 0.0008151212241500616
step: 90, loss: 0.00019050497212447226
step: 100, loss: 0.0002461467229295522
step: 110, loss: 0.0003145984373986721
step: 120, loss: 8.467448787996545e-05
step: 130, loss: 0.00013046774256508797
step: 140, loss: 0.00010830452083609998
step: 150, loss: 0.00013587603461928666
step: 160, loss: 0.0720972865819931
step: 170, loss: 0.0008293346618302166
step: 180, loss: 0.0005135972751304507
step: 190, loss: 0.00034832351957447827
step: 200, loss: 0.029582219198346138
step: 210, loss: 0.00010522630327614024
step: 220, loss: 0.03398378565907478
step: 230, loss: 0.00993336457759142
/usr/local/lib/python3.7/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=0.976271186440678, f1=0.9740698985343857,
best_f1=0.9773755656108598
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0004645257431548089
step: 10, loss: 0.0005145782488398254
step: 20, loss: 0.0003231401205994189
step: 30, loss: 0.00045738863991573453
step: 40, loss: 0.0014931682962924242
step: 50, loss: 0.00026952908956445754
step: 60, loss: 0.00029256631387397647
step: 70, loss: 0.00011704386270139366
step: 80, loss: 0.00024259774363599718
step: 90, loss: 8.26406103442423e-05
step: 100, loss: 0.04649874567985535
step: 110, loss: 0.0003689236764330417
step: 120, loss: 0.0023228360805660486
step: 130, loss: 0.00035351485712453723
step: 140, loss: 0.00015902916493359953
step: 150, loss: 0.0007039786432869732
step: 160, loss: 0.0015667147235944867
step: 170, loss: 0.0005800462095066905
step: 180, loss: 0.021722419187426567
step: 190, loss: 0.008524147793650627
step: 200, loss: 0.0007490087300539017
step: 210, loss: 0.0004636706435121596
step: 220, loss: 0.0016452680574730039
```

```
step: 230, loss: 0.0006974394782446325
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9819819819819819, f1=0.9753363228699552,
best_f1=0.9753363228699552
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00020017437054775655
step: 10, loss: 0.0018234089948236942
step: 20, loss: 0.0008050092146731913
step: 30, loss: 0.0001232438225997612
step: 40, loss: 0.04125375673174858
step: 50, loss: 0.0001320892624789849
step: 60, loss: 0.00022652457118965685
step: 70, loss: 0.00011978963448200375
step: 80, loss: 0.00023840602079872042
step: 90, loss: 0.0002966919564642012
step: 100, loss: 0.0158619973808527
step: 110, loss: 8.408509893342853e-05
step: 120, loss: 5.540937127079815e-05
step: 130, loss: 4.2629668314475566e-05
step: 140, loss: 0.0001269933272851631
step: 150, loss: 0.0007726395851932466
step: 160, loss: 0.0003452706150710583
step: 170, loss: 0.0002352118754060939
step: 180, loss: 0.0037090876139700413
step: 190, loss: 8.861265087034553e-05
step: 200, loss: 0.0049045211635529995
step: 210, loss: 0.001640578731894493
step: 220, loss: 7.952863234095275e-05
step: 230, loss: 0.0007458134787157178
/usr/local/lib/python3.7/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=0.9819004524886877, f1=0.9717514124293786,
best_f1=0.9753363228699552
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0005764744128100574
step: 10, loss: 4.939762220601551e-05
step: 20, loss: 0.0002370249858358875
```

```
step: 30, loss: 9.328324813395739e-05
step: 40, loss: 4.616055593942292e-05
step: 50, loss: 4.4451095163822174e-05
step: 60, loss: 0.001659413450397551
step: 70, loss: 6.830145139247179e-05
step: 80, loss: 0.00018145734793506563
step: 90, loss: 0.00010830048995558172
step: 100, loss: 5.2821134886471555e-05
step: 110, loss: 0.00012510477972682565
step: 120, loss: 6.798082904424518e-05
step: 130, loss: 6.0195907281013206e-05
step: 140, loss: 0.012676191516222057
step: 150, loss: 0.06137249246239662
step: 160, loss: 6.66164924041368e-05
step: 170, loss: 6.551742262672633e-05
step: 180, loss: 0.003973895218223333
step: 190, loss: 0.0020095547661185265
step: 200, loss: 0.00017487183504272252
step: 210, loss: 0.00011377512419130653
step: 220, loss: 0.00010993259638780728
step: 230, loss: 0.000629861606284976
/usr/local/lib/python3.7/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=0.977728285077951, f1=0.9678848283499446,
best_f1=0.9753363228699552
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 6.938086153240874e-05
step: 10, loss: 0.000779980793595314
step: 20, loss: 0.0014163509476929903
step: 30, loss: 0.00021555778221227229
step: 40, loss: 0.0002530465426389128
step: 50, loss: 0.00023453460016753525
step: 60, loss: 0.00017332645074930042
step: 70, loss: 9.204285743180662e-05
step: 80, loss: 3.866270344587974e-05
step: 90, loss: 0.000352701754309237
step: 100, loss: 7.791675307089463e-05
step: 110, loss: 6.625587411690503e-05
step: 120, loss: 4.231974889989942e-05
step: 130, loss: 3.334694338263944e-05
step: 140, loss: 9.423254959983751e-05
step: 150, loss: 0.0003229262074455619
step: 160, loss: 6.414132076315582e-05
step: 170, loss: 3.715220373123884e-05
step: 180, loss: 0.004312175326049328
```

```
step: 190, loss: 6.282705726334825e-05
step: 200, loss: 5.6193552154581994e-05
step: 210, loss: 3.869584543281235e-05
step: 220, loss: 5.3478368499781936e-05
step: 230, loss: 5.681228867615573e-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=0.9788182831661093, f1=0.9698324022346367,
best_f1=0.9753363228699552
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 6.708333239657804e-05
step: 10, loss: 3.3273005101364106e-05
step: 20, loss: 6.347700400510803e-05
step: 30, loss: 0.00010202266275882721
step: 40, loss: 0.00024990085512399673
step: 50, loss: 0.00012722676910925657
step: 60, loss: 9.446468175156042e-05
step: 70, loss: 7.279578858288005e-05
step: 80, loss: 0.0004214034997858107
step: 90, loss: 0.00015342766710091382
step: 100, loss: 2.7402280466048978e-05
step: 110, loss: 4.612117481883615e-05
step: 120, loss: 6.314613710856065e-05
step: 130, loss: 4.531907688942738e-05
step: 140, loss: 4.09012682212051e-05
step: 150, loss: 8.83951288415119e-05
step: 160, loss: 0.00015851495845708996
step: 170, loss: 5.4200503655010834e-05
step: 180, loss: 5.8970665122615173e-05
step: 190, loss: 4.664212974603288e-05
step: 200, loss: 2.2604461264563724e-05
step: 210, loss: 3.656590706668794e-05
step: 220, loss: 6.012328958604485e-05
step: 230, loss: 0.03690812364220619
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9820224719101124, f1=0.9741282339707535,
best_f1=0.9741282339707535
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
```

```
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 4.6679073420818895e-05
step: 10, loss: 8.180539589375257e-05
step: 20, loss: 0.00011925414582947269
step: 30, loss: 0.0001661035930737853
step: 40, loss: 3.3801687095547095e-05
step: 50, loss: 0.0006034228135831654
step: 60, loss: 6.047405622666702e-05
step: 70, loss: 3.1707164453109726e-05
step: 80, loss: 3.527679655235261e-05
step: 90, loss: 6.949828093638644e-05
step: 100, loss: 3.374911466380581e-05
step: 110, loss: 0.027720982208848
step: 120, loss: 0.0001759931183187291
step: 130, loss: 5.163591413293034e-05
step: 140, loss: 0.00011616632400546223
step: 150, loss: 4.011571945738979e-05
step: 160, loss: 0.00010366956848884001
step: 170, loss: 5.812329618493095e-05
step: 180, loss: 3.4182012313976884e-05
step: 190, loss: 3.34215656039305e-05
step: 200, loss: 3.545189611031674e-05
step: 210, loss: 2.1185140212764964e-05
step: 220, loss: 4.160456228419207e-05
step: 230, loss: 2.83747667708667e-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=0.9820224719101124, f1=0.9763779527559054,
best_f1=0.9741282339707535
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 7.805911445757374e-05
step: 10, loss: 8.032302139326693e-05
step: 20, loss: 3.8694324757670984e-05
step: 30, loss: 4.296541737858206e-05
step: 40, loss: 4.6222055971156806e-05
step: 50, loss: 7.290715439012274e-05
step: 60, loss: 4.625307337846607e-05
step: 70, loss: 4.6998764446470886e-05
step: 80, loss: 5.813484312966466e-05
step: 90, loss: 8.106313907774165e-05
step: 100, loss: 3.078836380154826e-05
step: 110, loss: 3.0702554795425385e-05
step: 120, loss: 2.208652949775569e-05
step: 130, loss: 3.435330290812999e-05
step: 140, loss: 5.5412827350664884e-05
```

```
step: 150, loss: 2.650804162840359e-05
step: 160, loss: 0.00026512646581977606
step: 170, loss: 2.571791446825955e-05
step: 180, loss: 3.4732569474726915e-05
step: 190, loss: 6.67126732878387e-05
step: 200, loss: 2.8907048545079306e-05
step: 210, loss: 3.724763519130647e-05
step: 220, loss: 0.0005024602869525552
step: 230, loss: 0.002683211350813508
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9820627802690582, 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: 2.2336203983286396e-05
step: 10, loss: 1.725521906337235e-05
step: 20, loss: 5.835137199028395e-05
step: 30, loss: 4.068771522725001e-05
step: 40, loss: 3.505358108668588e-05
step: 50, loss: 0.019087743014097214
step: 60, loss: 4.170200554654002e-05
step: 70, loss: 4.8389443691121414e-05
step: 80, loss: 3.9266084058908746e-05
step: 90, loss: 7.143418042687699e-05
step: 100, loss: 4.7877110773697495e-05
step: 110, loss: 2.4165330614778213e-05
step: 120, loss: 8.132861694321036e-05
step: 130, loss: 3.336227382533252e-05
step: 140, loss: 2.938030411314685e-05
step: 150, loss: 0.0003452726814430207
step: 160, loss: 2.1248530174489133e-05
step: 170, loss: 3.1659019441576675e-05
step: 180, loss: 3.940869646612555e-05
step: 190, loss: 3.971316255046986e-05
step: 200, loss: 5.121338108438067e-05
step: 210, loss: 0.00014873757027089596
step: 220, loss: 4.623709537554532e-05
step: 230, loss: 3.242350430809893e-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=0.9820627802690582, f1=0.9753363228699552,
best_f1=0.9752808988764046
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, 199.48it/s]
load_f1 = 0.9820224719101124
real_f1 = 0.9820224719101124
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, 247.36it/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: 0.6353820562362671
step: 10, loss: 0.5073322653770447
step: 20, loss: 0.5153066515922546
step: 30, loss: 0.15039213001728058
step: 40, loss: 0.16093450784683228
step: 50, loss: 0.12545731663703918
step: 60, loss: 0.03644094988703728
step: 70, loss: 0.13088616728782654
step: 80, loss: 0.06656893342733383
step: 90, loss: 0.2688131332397461
step: 100, loss: 0.019345665350556374
step: 110, loss: 0.09453622251749039
step: 120, loss: 0.058634184300899506
step: 130, loss: 0.07535288482904434
step: 140, loss: 0.19131241738796234
step: 150, loss: 0.056428734213113785
step: 160, loss: 0.018291281536221504
step: 170, loss: 0.1848556399345398
step: 180, loss: 0.08468259871006012
step: 190, loss: 0.009034608490765095
step: 200, loss: 0.11882679909467697
step: 210, loss: 0.07449843734502792
step: 220, loss: 0.24364151060581207
step: 230, loss: 0.14176180958747864
step: 240, loss: 0.07434066385030746
step: 250, loss: 0.01401505246758461
step: 260, loss: 0.07387635111808777
step: 270, loss: 0.043988361954689026
step: 280, loss: 0.06754173338413239
step: 290, loss: 0.04291051998734474
step: 300, loss: 0.021390268579125404
step: 310, loss: 0.1712079644203186
step: 320, loss: 0.08941511064767838
step: 330, loss: 0.010247890837490559
step: 340, loss: 0.04027310013771057
```

```
step: 350, loss: 0.11168801039457321
step: 360, loss: 0.08799289911985397
step: 370, loss: 0.07886168360710144
step: 380, loss: 0.008352170698344707
step: 390, loss: 0.17562159895896912
step: 400, loss: 0.15892194211483002
step: 410, loss: 0.06706181168556213
step: 420, loss: 0.09830500930547714
step: 430, loss: 0.11455438286066055
step: 440, loss: 0.03454729914665222
step: 450, loss: 0.0060407440178096294
step: 460, loss: 0.022377699613571167
step: 470, loss: 0.15100014209747314
step: 480, loss: 0.08393893390893936
step: 490, loss: 0.11040318012237549
step: 500, loss: 0.09225517511367798
step: 510, loss: 0.08995599299669266
step: 520, loss: 0.18596521019935608
step: 530, loss: 0.004845513496547937
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9426038264115725, f1=0.931098696461825,
best_f1=0.931098696461825
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.1382461041212082
step: 10, loss: 0.04282284528017044
step: 20, loss: 0.00938963983207941
step: 30, loss: 0.00649282755330205
step: 40, loss: 0.08714873343706131
step: 50, loss: 0.13163097202777863
step: 60, loss: 0.0028901593759655952
step: 70, loss: 0.01797778159379959
step: 80, loss: 0.04919851943850517
step: 90, loss: 0.01585349254310131
step: 100, loss: 0.010473194532096386
step: 110, loss: 0.02514643222093582
step: 120, loss: 0.056115422397851944
step: 130, loss: 0.052064236253499985
step: 140, loss: 0.015111136250197887
step: 150, loss: 0.04238656908273697
step: 160, loss: 0.03451288491487503
```

```
step: 170, loss: 0.0118088768795513264
step: 180, loss: 0.09729896485805511
step: 190, loss: 0.0488998144865036
step: 200, loss: 0.013832363300025463
step: 210, loss: 0.042775120586156845
step: 220, loss: 0.02667832002043724
step: 230, loss: 0.03642230108380318
step: 240, loss: 0.029201103374361992
step: 250, loss: 0.014055142179131508
step: 260, loss: 0.013570050708949566
step: 270, loss: 0.13080130517482758
step: 280, loss: 0.041927237063646317
step: 290, loss: 0.03934074193239212
step: 300, loss: 0.19770373404026031
step: 310, loss: 0.008122870698571205
step: 320, loss: 0.051686957478523254
step: 330, loss: 0.07382664084434509
step: 340, loss: 0.011689122766256332
step: 350, loss: 0.0037838867865502834
step: 360, loss: 0.04693153500556946
step: 370, loss: 0.1093669906258583
step: 380, loss: 0.04576779901981354
step: 390, loss: 0.07448209077119827
step: 400, loss: 0.051840029656887054
step: 410, loss: 0.27044054865837097
step: 420, loss: 0.01757739670574665
step: 430, loss: 0.12501665949821472
step: 440, loss: 0.17302566766738892
step: 450, loss: 0.08718055486679077
step: 460, loss: 0.022783761844038963
step: 470, loss: 0.09229721873998642
step: 480, loss: 0.2050756812095642
step: 490, loss: 0.008941804990172386
step: 500, loss: 0.2092979997396469
step: 510, loss: 0.029386015608906746
step: 520, loss: 0.11341958492994308
step: 530, loss: 0.010024160146713257
/usr/local/lib/python3.7/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=0.9424083769633508, f1=0.934844192634561,
best_f1=0.931098696461825
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.026973038911819458
step: 10, loss: 0.01353347860276699
step: 20, loss: 0.07133316993713379
```

```
step: 30, loss: 0.08935505151748657
step: 40, loss: 0.0022523519583046436
step: 50, loss: 0.1043737605214119
step: 60, loss: 0.01703302189707756
step: 70, loss: 0.003382388036698103
step: 80, loss: 0.009248235262930393
step: 90, loss: 0.007325091864913702
step: 100, loss: 0.033457931131124496
step: 110, loss: 0.005368509329855442
step: 120, loss: 0.0011279876343905926
step: 130, loss: 0.0010335531551390886
step: 140, loss: 0.04606708139181137
step: 150, loss: 0.031628020107746124
step: 160, loss: 0.016788102686405182
step: 170, loss: 0.03070908412337303
step: 180, loss: 0.005069802049547434
step: 190, loss: 0.001613510656170547
step: 200, loss: 0.06937608122825623
step: 210, loss: 0.025055408477783203
step: 220, loss: 0.043246544897556305
step: 230, loss: 0.09389092028141022
step: 240, loss: 0.0030385071877390146
step: 250, loss: 0.04488210752606392
step: 260, loss: 0.04156753420829773
step: 270, loss: 0.004339539911597967
step: 280, loss: 0.11728541553020477
step: 290, loss: 0.0010181983234360814
step: 300, loss: 0.040120091289281845
step: 310, loss: 0.0019953777082264423
step: 320, loss: 0.034853361546993256
step: 330, loss: 0.022413186728954315
step: 340, loss: 0.0018643955700099468
step: 350, loss: 0.005579223856329918
step: 360, loss: 0.01886582002043724
step: 370, loss: 0.005924299359321594
step: 380, loss: 0.01682928390800953
step: 390, loss: 0.025322940200567245
step: 400, loss: 0.004004334565252066
step: 410, loss: 0.005948617588728666
step: 420, loss: 0.16766159236431122
step: 430, loss: 0.0207191351801157
step: 440, loss: 0.009951671585440636
step: 450, loss: 0.05534409359097481
step: 460, loss: 0.015997495502233505
step: 470, loss: 0.06437166780233383
step: 480, loss: 0.0018873566295951605
step: 490, loss: 0.009788495488464832
step: 500, loss: 0.003755688900128007
```

```
step: 510, loss: 0.01047028973698616
Gradient overflow.  Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 520, loss: 0.07852314412593842
step: 530, loss: 0.10372145473957062
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9486348912540491, f1=0.9401947148817803,
best_f1=0.9401947148817803
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.005274541210383177
step: 10, loss: 0.005834496114403009
step: 20, loss: 0.0024789157323539257
step: 30, loss: 0.0005424150731414557
step: 40, loss: 0.002455309731885791
step: 50, loss: 0.014696570113301277
step: 60, loss: 0.0008020129753276706
step: 70, loss: 0.002359675941988826
step: 80, loss: 0.0004153094196226448
step: 90, loss: 0.0968398004770279
step: 100, loss: 0.006376064382493496
step: 110, loss: 0.014803450554609299
step: 120, loss: 0.00038057309575378895
step: 130, loss: 0.0006806051242165267
step: 140, loss: 0.0038640943821519613
step: 150, loss: 0.12534455955028534
step: 160, loss: 0.018504032865166664
step: 170, loss: 0.005357031244784594
step: 180, loss: 0.020235951989889145
step: 190, loss: 0.0358971431851387
step: 200, loss: 0.04170818626880646
step: 210, loss: 0.01401173323392868
step: 220, loss: 0.0008343904628418386
step: 230, loss: 0.0027725910767912865
step: 240, loss: 0.0070181540213525295
step: 250, loss: 0.0018852397333830595
step: 260, loss: 0.0006217756308615208
step: 270, loss: 0.010626215487718582
step: 280, loss: 0.0147246690467
step: 290, loss: 0.012536974623799324
step: 300, loss: 0.0001705121248960495
step: 310, loss: 0.002860068343579769
```

```
step: 320, loss: 0.05913637951016426
step: 330, loss: 0.061753977090120316
step: 340, loss: 0.016199981793761253
step: 350, loss: 0.0035879190545529127
step: 360, loss: 0.014963899739086628
step: 370, loss: 0.0020500035025179386
step: 380, loss: 0.01039524283260107
step: 390, loss: 0.0029197654221206903
step: 400, loss: 0.0010237162932753563
step: 410, loss: 0.0012848018668591976
step: 420, loss: 0.04210466518998146
step: 430, loss: 0.0020362737122923136
step: 440, loss: 0.0010059298947453499
step: 450, loss: 0.003738450352102518
step: 460, loss: 0.00017922790721058846
step: 470, loss: 0.0002772084844764322
step: 480, loss: 0.008122620172798634
step: 490, loss: 0.003852258436381817
step: 500, loss: 0.00922660157084465
step: 510, loss: 0.01657610386610031
step: 520, loss: 0.11054232716560364
step: 530, loss: 0.01507687009871006
/usr/local/lib/python3.7/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=0.9448529411764707, f1=0.932904411764706,
best_f1=0.9401947148817803
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.02690829150378704
step: 10, loss: 0.005277940537780523
step: 20, loss: 0.0029886991251260042
step: 30, loss: 0.0007077534100972116
step: 40, loss: 0.007650560699403286
step: 50, loss: 0.002951163798570633
step: 60, loss: 0.025746984407305717
step: 70, loss: 0.0008522020652890205
step: 80, loss: 0.00047073618043214083
step: 90, loss: 0.08391214907169342
step: 100, loss: 0.0016375056002289057
step: 110, loss: 0.0008560010464861989
step: 120, loss: 0.010891750454902649
step: 130, loss: 0.00027434006915427744
step: 140, loss: 0.003428451484069228
step: 150, loss: 0.05398145690560341
step: 160, loss: 0.0005891474429517984
step: 170, loss: 0.05842139199376106
```

```
step: 180, loss: 0.000568204908631742
step: 190, loss: 0.01885947398841381
step: 200, loss: 0.0005700523033738136
step: 210, loss: 0.024888312444090843
step: 220, loss: 0.0002916379307862371
step: 230, loss: 0.006371904630213976
step: 240, loss: 0.0016624287236481905
step: 250, loss: 0.0035772621631622314
step: 260, loss: 0.00088829844025895
step: 270, loss: 0.00013153011968825012
step: 280, loss: 0.008024878799915314
step: 290, loss: 0.12232472002506256
step: 300, loss: 0.0004592206096276641
step: 310, loss: 0.0005260832840576768
step: 320, loss: 0.17919808626174927
step: 330, loss: 0.00777842104434967
step: 340, loss: 0.007004708517342806
step: 350, loss: 0.0006538442685268819
step: 360, loss: 0.0014238671865314245
step: 370, loss: 0.0028367529157549143
step: 380, loss: 0.00015006140165496618
step: 390, loss: 0.00023163309379015118
step: 400, loss: 0.004644307307898998
step: 410, loss: 0.005987347569316626
step: 420, loss: 0.0029949755407869816
step: 430, loss: 0.005963711533695459
step: 440, loss: 0.08908254653215408
step: 450, loss: 0.00029912774334666166
step: 460, loss: 0.012435010634362698
step: 470, loss: 0.15635254979133606
step: 480, loss: 0.0014599986607208848
step: 490, loss: 0.00021994076087139547
step: 500, loss: 0.002943366067484021
step: 510, loss: 0.22917890548706055
step: 520, loss: 0.0009126780205406249
step: 530, loss: 0.0009423285955563188
/usr/local/lib/python3.7/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=0.9435559736594544, f1=0.9339578454332552,
best_f1=0.9401947148817803
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.03293495252728462
step: 10, loss: 0.00016668772150296718
step: 20, loss: 0.045897312462329865
step: 30, loss: 0.0013401622418314219
```

```
step: 40, loss: 0.007644464261829853
step: 50, loss: 0.003346678800880909
step: 60, loss: 8.529049955541268e-05
step: 70, loss: 0.00016514380695298314
step: 80, loss: 0.0052136084996163845
step: 90, loss: 0.000201104543521069
step: 100, loss: 0.035836197435855865
step: 110, loss: 0.00038756339927203953
step: 120, loss: 0.0001316186971962452
step: 130, loss: 0.00019732116197701544
step: 140, loss: 0.004121164791285992
step: 150, loss: 0.09646855294704437
step: 160, loss: 0.0022789195645600557
step: 170, loss: 0.0008673936245031655
step: 180, loss: 0.0005508067552000284
step: 190, loss: 0.023555072024464607
step: 200, loss: 0.00038024623063392937
step: 210, loss: 0.0024668346159160137
step: 220, loss: 0.0018651463324204087
step: 230, loss: 0.01827620342373848
step: 240, loss: 0.017849434167146683
step: 250, loss: 0.0009262178791686893
step: 260, loss: 0.0003236375341657549
step: 270, loss: 0.015644496306777
step: 280, loss: 0.000316145655233413
step: 290, loss: 0.0004789275408256799
step: 300, loss: 0.0011101554846391082
step: 310, loss: 0.0014864258700981736
step: 320, loss: 0.0001013902947306633
step: 330, loss: 0.00014678901061415672
step: 340, loss: 0.09548979252576828
step: 350, loss: 0.020293626934289932
step: 360, loss: 0.002118331380188465
step: 370, loss: 0.0321071557700634
step: 380, loss: 0.0015992731787264347
step: 390, loss: 0.027896979823708534
step: 400, loss: 0.00024721273803152144
step: 410, loss: 0.0011940234107896686
step: 420, loss: 0.0035162654239684343
step: 430, loss: 0.00021738688519690186
step: 440, loss: 4.677666584029794e-05
step: 450, loss: 0.0030987542122602463
step: 460, loss: 0.00020355143351480365
step: 470, loss: 0.006947772111743689
step: 480, loss: 0.0008791904547251761
step: 490, loss: 0.0065558538772165775
step: 500, loss: 0.00021911396470386535
step: 510, loss: 0.0013459998881444335
```

```
step: 520, loss: 0.0020270319655537605
step: 530, loss: 0.028129318729043007
/usr/local/lib/python3.7/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=0.9341603053435114, f1=0.9190225203641591,
best_f1=0.9401947148817803
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.006018589250743389
step: 10, loss: 0.001219673315063119
step: 20, loss: 0.0012357874074950814
step: 30, loss: 0.0007995617925189435
step: 40, loss: 0.0002510058111511171
step: 50, loss: 0.0008937724051065743
step: 60, loss: 0.002238925313577056
step: 70, loss: 0.00361591181717813
step: 80, loss: 0.006260947789996862
step: 90, loss: 0.001266620820388198
step: 100, loss: 8.773045556154102e-05
step: 110, loss: 0.00038897997001186013
step: 120, loss: 0.00013710049097426236
step: 130, loss: 0.011114098131656647
step: 140, loss: 4.2537514673313126e-05
step: 150, loss: 0.1036568284034729
step: 160, loss: 0.0004183700948487967
step: 170, loss: 0.04061039909720421
step: 180, loss: 0.0007281215512193739
step: 190, loss: 0.001299054827541113
step: 200, loss: 0.00012522809265647084
step: 210, loss: 0.0013548467541113496
step: 220, loss: 0.1000390574336052
step: 230, loss: 0.00023382318613585085
step: 240, loss: 0.000261968991253525
step: 250, loss: 0.00405695429071784
step: 260, loss: 0.0003027759084943682
step: 270, loss: 0.0002757013135123998
step: 280, loss: 0.007824975065886974
step: 290, loss: 0.0007776169804856181
step: 300, loss: 0.0013496385654434562
step: 310, loss: 7.480932981707156e-05
step: 320, loss: 8.888135198503733e-05
step: 330, loss: 0.004776468500494957
step: 340, loss: 0.01853233389556408
step: 350, loss: 0.0014049278106540442
step: 360, loss: 0.000641087710391301
step: 370, loss: 5.010610038880259e-05
```

```
step: 380, loss: 0.0022375828120857477
step: 390, loss: 6.158179166959599e-05
step: 400, loss: 0.005376606713980436
step: 410, loss: 0.0016596984351053834
step: 420, loss: 0.0013332524104043841
step: 430, loss: 0.00019065593369305134
step: 440, loss: 0.00017855939222499728
step: 450, loss: 0.004294349811971188
step: 460, loss: 0.001194715267047286
step: 470, loss: 0.0009935699636116624
step: 480, loss: 0.003625516314059496
step: 490, loss: 0.0002545539173297584
step: 500, loss: 0.0006960559985600412
step: 510, loss: 0.0028226228896528482
step: 520, loss: 0.059824489057064056
step: 530, loss: 0.003643846372142434
/usr/local/lib/python3.7/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=0.9408424041646947, f1=0.9248826291079812,
best_f1=0.9401947148817803
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0012469510547816753
step: 10, loss: 0.017039325088262558
step: 20, loss: 0.03591729328036308
step: 30, loss: 0.0006198594346642494
step: 40, loss: 4.391962647787295e-05
step: 50, loss: 0.003858183976262808
step: 60, loss: 0.0002707894891500473
step: 70, loss: 4.259733032085933e-05
step: 80, loss: 0.0005706443334929645
step: 90, loss: 0.0003449819050729275
step: 100, loss: 6.920587475178763e-05
step: 110, loss: 0.00010078279592562467
step: 120, loss: 0.0008436771458946168
step: 130, loss: 0.0001361678441753611
step: 140, loss: 0.007891563698649406
step: 150, loss: 4.2253912397427484e-05
step: 160, loss: 0.001836365438066423
step: 170, loss: 0.011963286437094212
step: 180, loss: 3.6427954910323024e-05
step: 190, loss: 0.001317257178016007
step: 200, loss: 0.0003531394468154758
step: 210, loss: 0.00031465551001019776
step: 220, loss: 0.00028976122848689556
step: 230, loss: 0.004298623185604811
```

```
step: 240, loss: 0.0003271752502769232
step: 250, loss: 2.4523018510080874e-05
step: 260, loss: 7.289511995622888e-05
step: 270, loss: 9.458241402171552e-05
step: 280, loss: 0.0031380439177155495
step: 290, loss: 0.015512421727180481
step: 300, loss: 0.0037424112670123577
step: 310, loss: 0.0007136098574846983
step: 320, loss: 0.013454732485115528
step: 330, loss: 0.0010756156407296658
step: 340, loss: 0.0002160482108592987
step: 350, loss: 6.554000719916075e-05
step: 360, loss: 0.0034772194921970367
step: 370, loss: 7.36572255846113e-05
step: 380, loss: 0.0002656869764905423
step: 390, loss: 0.13230346143245697
step: 400, loss: 6.732217298122123e-05
step: 410, loss: 0.00021098082652315497
step: 420, loss: 0.0011408058926463127
step: 430, loss: 0.015530903823673725
step: 440, loss: 0.00026790291303768754
step: 450, loss: 0.00018966877541970462
step: 460, loss: 8.511007763445377e-05
step: 470, loss: 0.00011938872921746224
step: 480, loss: 0.05506008118391037
step: 490, loss: 0.018917711451649666
step: 500, loss: 0.13796184957027435
step: 510, loss: 0.00018012127839028835
step: 520, loss: 0.0017553012585267425
step: 530, loss: 0.0005390428123064339
/usr/local/lib/python3.7/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=0.9422098936662043, f1=0.933456561922366,
best_f1=0.9401947148817803
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0054887086153030396
step: 10, loss: 0.0005412689642980695
step: 20, loss: 0.0006372471107169986
step: 30, loss: 0.0025481784250587225
step: 40, loss: 0.0002704368671402335
step: 50, loss: 0.0001093065511668101
step: 60, loss: 0.00010316498810425401
step: 70, loss: 0.03951750323176384
step: 80, loss: 0.00032251968514174223
step: 90, loss: 0.0025940993800759315
```

```
step: 100, loss: 5.8648016420193017e-05
step: 110, loss: 4.571415047394112e-05
step: 120, loss: 0.00021889664640184492
step: 130, loss: 0.02426040917634964
step: 140, loss: 0.006439877673983574
step: 150, loss: 0.004265526309609413
step: 160, loss: 0.0018289678264409304
step: 170, loss: 0.0029739034362137318
step: 180, loss: 0.00012874904496129602
step: 190, loss: 0.0002014017227338627
step: 200, loss: 0.0001069519275915809
step: 210, loss: 5.47153758816421e-05
step: 220, loss: 0.006548506207764149
step: 230, loss: 0.0005538287223316729
step: 240, loss: 0.00023727634106762707
step: 250, loss: 0.00026774313300848007
step: 260, loss: 0.027461666613817215
step: 270, loss: 0.00062328897183761
step: 280, loss: 0.0017716671572998166
step: 290, loss: 0.015231437981128693
step: 300, loss: 0.00010885237134061754
step: 310, loss: 0.013233541510999203
step: 320, loss: 0.013690668158233166
step: 330, loss: 0.0010528757702559233
step: 340, loss: 0.00010038320760941133
step: 350, loss: 0.003679159563034773
step: 360, loss: 2.6493225959711708e-05
step: 370, loss: 4.1373667045263574e-05
step: 380, loss: 7.623912097187713e-05
step: 390, loss: 8.115961099974811e-05
step: 400, loss: 0.016181768849492073
step: 410, loss: 7.52375417505391e-05
step: 420, loss: 4.232585706631653e-05
step: 430, loss: 5.6295386457350105e-05
step: 440, loss: 0.07160913944244385
step: 450, loss: 6.363869033521041e-05
step: 460, loss: 0.00013075330934952945
step: 470, loss: 2.8940294214407913e-05
step: 480, loss: 0.0002064796135528013
step: 490, loss: 0.00011401755182305351
step: 500, loss: 0.00043488069786690176
step: 510, loss: 0.00021001407003495842
step: 520, loss: 4.645061198971234e-05
step: 530, loss: 8.318448817590252e-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=0.9472234970169803, f1=0.9386446886446886,
```

```
best_f1=0.9401947148817803
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 3.449840369285084e-05
step: 10, loss: 0.00016442347259726375
step: 20, loss: 2.8384971301420592e-05
step: 30, loss: 3.296367867733352e-05
step: 40, loss: 8.423109102295712e-05
step: 50, loss: 5.858329313923605e-05
step: 60, loss: 3.86440078727901e-05
step: 70, loss: 4.026140959467739e-05
step: 80, loss: 2.429202686471399e-05
step: 90, loss: 3.660992661025375e-05
step: 100, loss: 2.5676939912955277e-05
step: 110, loss: 1.67709822562756e-05
step: 120, loss: 4.91226019221358e-05
step: 130, loss: 2.8880962418043055e-05
step: 140, loss: 0.0008548669866286218
step: 150, loss: 3.9940394344739616e-05
step: 160, loss: 2.8072663553757593e-05
step: 170, loss: 0.00038763682823628187
step: 180, loss: 7.927114347694442e-05
step: 190, loss: 5.096866880194284e-05
step: 200, loss: 0.00035672783269546926
step: 210, loss: 0.00037670941674150527
step: 220, loss: 8.356005855603144e-05
step: 230, loss: 0.00034978415351361036
step: 240, loss: 8.903434354579076e-05
step: 250, loss: 2.2328815248329192e-05
step: 260, loss: 0.003391255158931017
step: 270, loss: 0.007168347015976906
step: 280, loss: 4.0166782127926126e-05
step: 290, loss: 2.393063550698571e-05
step: 300, loss: 0.0006760630640201271
step: 310, loss: 3.909685619873926e-05
step: 320, loss: 0.0005643125041387975
step: 330, loss: 2.8425791242625564e-05
step: 340, loss: 4.9359066906617954e-05
step: 350, loss: 5.7686262152856216e-05
step: 360, loss: 2.682130980247166e-05
step: 370, loss: 0.0019289783667773008
step: 380, loss: 7.284917955985293e-05
step: 390, loss: 8.142585284076631e-05
step: 400, loss: 0.0004800416063517332
step: 410, loss: 0.0023058729711920023
step: 420, loss: 9.701358794700354e-05
step: 430, loss: 1.9520015484886244e-05
```

```
step: 440, loss: 0.0027873008511960506
step: 450, loss: 3.909412043867633e-05
step: 460, loss: 2.7234094886807725e-05
step: 470, loss: 0.00033011811319738626
step: 480, loss: 3.067274155910127e-05
step: 490, loss: 0.000648816698230803
step: 500, loss: 0.0044991374015808105
step: 510, loss: 3.0166127544362098e-05
step: 520, loss: 0.00021119943994563073
step: 530, loss: 8.578428969485685e-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=0.9473684210526316, f1=0.9348025711662074,
best_f1=0.9401947148817803
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00010249847400700673
step: 10, loss: 4.494798122323118e-05
step: 20, loss: 2.619148654048331e-05
step: 30, loss: 0.00011854506738018245
step: 40, loss: 5.948047328274697e-05
step: 50, loss: 9.574840078130364e-05
Gradient overflow.  Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 60, loss: 0.016935577616095543
step: 70, loss: 2.947289431176614e-05
step: 80, loss: 2.6459427317604423e-05
step: 90, loss: 3.725697024492547e-05
step: 100, loss: 3.1516105082118884e-05
step: 110, loss: 3.7495585274882615e-05
step: 120, loss: 3.8400350604206324e-05
step: 130, loss: 2.825785122695379e-05
step: 140, loss: 0.00011430826270952821
step: 150, loss: 1.3489059710991569e-05
step: 160, loss: 6.544915959239006e-05
step: 170, loss: 3.130220648017712e-05
step: 180, loss: 0.00010030931298388168
step: 190, loss: 8.411310409428552e-05
step: 200, loss: 0.0006753912894055247
step: 210, loss: 0.018447691574692726
step: 220, loss: 0.0012234754394739866
step: 230, loss: 1.3422049050859641e-05
step: 240, loss: 0.00022787631314713508
step: 250, loss: 1.7798873159335926e-05
step: 260, loss: 2.6682824682211503e-05
step: 270, loss: 4.521777736954391e-05
step: 280, loss: 4.636115409084596e-05
```

```
step: 290, loss: 0.0007126291748136282
step: 300, loss: 8.57700506458059e-05
step: 310, loss: 4.9883958126883954e-05
step: 320, loss: 0.0002644156338647008
step: 330, loss: 0.0004947728593833745
Gradient overflow.  Skipping step, loss scaler 0 reducing loss scale to 8192.0
step: 340, loss: 0.00026775497826747596
step: 350, loss: 5.8861947763944045e-05
step: 360, loss: 1.3630636203743052e-05
step: 370, loss: 0.0007160392706282437
step: 380, loss: 0.00015631093992851675
step: 390, loss: 1.6275575035251677e-05
step: 400, loss: 0.0018005073070526123
step: 410, loss: 5.332216824172065e-05
step: 420, loss: 7.520632789237425e-05
step: 430, loss: 2.4883811420295388e-05
step: 440, loss: 0.20726759731769562
step: 450, loss: 0.0006856822874397039
step: 460, loss: 0.026437869295477867
step: 470, loss: 0.00010378185106674209
step: 480, loss: 0.0015688029816374183
step: 490, loss: 0.00021011399803683162
step: 500, loss: 0.00022380526934284717
step: 510, loss: 8.874233026290312e-05
step: 520, loss: 7.566285785287619e-05
step: 530, loss: 6.233822205103934e-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=0.9323809523809523, f1=0.9241706161137442,
best_f1=0.9401947148817803
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00012748801964335144
step: 10, loss: 0.0003306334256194532
step: 20, loss: 0.00015423331933561713
step: 30, loss: 0.005069484002888203
step: 40, loss: 0.0004074301104992628
step: 50, loss: 0.03755149617791176
step: 60, loss: 0.0024532934185117483
step: 70, loss: 0.00022802315652370453
step: 80, loss: 5.5877160775708035e-05
step: 90, loss: 3.351490886416286e-05
step: 100, loss: 5.580268407356925e-05
step: 110, loss: 4.5789696741849184e-05
step: 120, loss: 3.7590114516206086e-05
step: 130, loss: 0.0007638622191734612
```

```
step: 140, loss: 5.702471753465943e-05
step: 150, loss: 2.4683020455995575e-05
step: 160, loss: 4.239724512444809e-05
step: 170, loss: 0.00012495627743192017
step: 180, loss: 0.00017727568047121167
step: 190, loss: 4.177571463515051e-05
step: 200, loss: 0.0010929268319159746
step: 210, loss: 5.963205694570206e-05
step: 220, loss: 0.0001239460543729365
step: 230, loss: 1.9844019334414043e-05
step: 240, loss: 0.00014038715744391084
step: 250, loss: 2.8332908186712302e-05
step: 260, loss: 3.917489084415138e-05
step: 270, loss: 0.0003986618830822408
step: 280, loss: 0.0006863302551209927
step: 290, loss: 3.961529000662267e-05
step: 300, loss: 0.00018224258383270353
step: 310, loss: 3.509996895445511e-05
step: 320, loss: 0.01807383820414543
step: 330, loss: 3.468065187917091e-05
step: 340, loss: 2.7043104637414217e-05
step: 350, loss: 0.00013249267067294568
step: 360, loss: 0.0001757764839567244
step: 370, loss: 0.00030327768763527274
step: 380, loss: 0.0004107895656488836
step: 390, loss: 0.0011632759124040604
step: 400, loss: 2.9968630769872107e-05
step: 410, loss: 2.4478009436279535e-05
step: 420, loss: 0.0033440415281802416
step: 430, loss: 0.0004159002855885774
step: 440, loss: 2.2187006834428757e-05
step: 450, loss: 0.0005256679141893983
step: 460, loss: 2.688414133444894e-05
step: 470, loss: 2.118138763762545e-05
step: 480, loss: 0.0033593943808227777
step: 490, loss: 0.00015152814739849418
step: 500, loss: 3.100975663983263e-05
step: 510, loss: 0.00012776031508110464
step: 520, loss: 1.9464254364720546e-05
step: 530, loss: 3.171392745571211e-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=0.9440820130475303, f1=0.9346314325452018,
best_f1=0.9401947148817803
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
```

```
step: 0, loss: 0.0004860812332481146
step: 10, loss: 2.57026022154605e-05
step: 20, loss: 2.1478934286278673e-05
step: 30, loss: 3.690328958327882e-05
step: 40, loss: 1.6152629541466013e-05
step: 50, loss: 0.0015133762499317527
step: 60, loss: 2.0484436390688643e-05
step: 70, loss: 3.3895103115355596e-05
step: 80, loss: 1.5884352251305245e-05
step: 90, loss: 0.00029343675123527646
step: 100, loss: 0.00012904488539788872
step: 110, loss: 3.529953028191812e-05
step: 120, loss: 4.2891395423794165e-05
step: 130, loss: 0.007245530374348164
step: 140, loss: 6.681526429019868e-05
step: 150, loss: 0.0007763320463709533
step: 160, loss: 1.799279561964795e-05
step: 170, loss: 4.758935756399296e-05
step: 180, loss: 1.965046612895094e-05
step: 190, loss: 3.278647636761889e-05
step: 200, loss: 1.4088885109231342e-05
step: 210, loss: 5.199732186156325e-05
step: 220, loss: 1.3191067409934476e-05
step: 230, loss: 1.6711046555428766e-05
step: 240, loss: 2.5327361072413623e-05
step: 250, loss: 1.2274712389626075e-05
step: 260, loss: 1.273289217351703e-05
step: 270, loss: 1.6859952665981837e-05
step: 280, loss: 1.4472572729573585e-05
step: 290, loss: 1.7474641936132684e-05
step: 300, loss: 2.4667875550221652e-05
step: 310, loss: 0.0003268296713940799
step: 320, loss: 0.002734582172706723
step: 330, loss: 0.00013265541929285973
step: 340, loss: 0.0007047271938063204
step: 350, loss: 1.6893896827241406e-05
step: 360, loss: 3.153066791128367e-05
step: 370, loss: 2.0287414372432977e-05
step: 380, loss: 1.846193481469527e-05
step: 390, loss: 0.00015700125368312
step: 400, loss: 1.3887690329283942e-05
step: 410, loss: 3.9807400753488764e-05
step: 420, loss: 4.2745414248201996e-05
step: 430, loss: 1.4155933058646042e-05
step: 440, loss: 3.9223239582497627e-05
step: 450, loss: 2.1508931240532547e-05
step: 460, loss: 0.0005592386587522924
step: 470, loss: 2.0606801626854576e-05
```

```
step: 480, loss: 1.3742407645622734e-05
step: 490, loss: 1.4461425053013954e-05
step: 500, loss: 1.8259677744936198e-05
step: 510, loss: 1.2889349818578921e-05
step: 520, loss: 0.0006576563464477658
step: 530, loss: 1.3343795217224397e-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=0.946146703806871, f1=0.937471051412691,
best_f1=0.9401947148817803
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0002313100703759119
step: 10, loss: 1.0415838005428668e-05
step: 20, loss: 1.138811967393849e-05
step: 30, loss: 1.5385239748866297e-05
step: 40, loss: 7.112858293112367e-05
step: 50, loss: 2.462343945808243e-05
step: 60, loss: 5.899511597817764e-05
step: 70, loss: 2.214920641563367e-05
step: 80, loss: 2.2123087546788156e-05
step: 90, loss: 1.3410874089458957e-05
step: 100, loss: 0.00013831595424562693
step: 110, loss: 1.1976701898674946e-05
step: 120, loss: 4.7129447921179235e-05
step: 130, loss: 0.015332682058215141
step: 140, loss: 0.000361207639798522
step: 150, loss: 1.033387343341019e-05
step: 160, loss: 0.0016584087861701846
step: 170, loss: 2.3833063096390106e-05
step: 180, loss: 1.659208101045806e-05
step: 190, loss: 1.4994007869972847e-05
step: 200, loss: 0.0002659619494806975
step: 210, loss: 1.637603600102011e-05
step: 220, loss: 1.6495267118443735e-05
step: 230, loss: 1.980308115889784e-05
step: 240, loss: 1.401432109560119e-05
step: 250, loss: 1.6286705431411974e-05
step: 260, loss: 2.000786116695963e-05
step: 270, loss: 1.173453620140208e-05
step: 280, loss: 1.6770696674939245e-05
step: 290, loss: 4.055539466207847e-05
step: 300, loss: 1.7806582036428154e-05
step: 310, loss: 4.587781586451456e-05
step: 320, loss: 0.0011446921853348613
step: 330, loss: 0.00024714009487070143
```

```
step: 340, loss: 3.924695556634106e-05
step: 350, loss: 2.757657966867555e-05
step: 360, loss: 7.288731285370886e-05
step: 370, loss: 1.5374080248875543e-05
step: 380, loss: 1.571305256220512e-05
step: 390, loss: 0.0028396009001880884
step: 400, loss: 0.0008237042347900569
step: 410, loss: 1.446877922717249e-05
step: 420, loss: 1.2408824659360107e-05
step: 430, loss: 1.579130912432447e-05
step: 440, loss: 0.0006547419470734894
step: 450, loss: 3.681727685034275e-05
step: 460, loss: 2.38809225265868e-05
step: 470, loss: 1.442780103388941e-05
step: 480, loss: 2.4011773348320276e-05
step: 490, loss: 3.51540838892106e-05
step: 500, loss: 1.3358500837057363e-05
step: 510, loss: 0.0006918307044543326
step: 520, loss: 3.850773282465525e-05
step: 530, loss: 0.0008538544061593711
/usr/local/lib/python3.7/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=0.9430970149253732, f1=0.9360518999073215,
best_f1=0.9401947148817803
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 1.4342174836201593e-05
step: 10, loss: 1.6077954569482245e-05
step: 20, loss: 1.1332217582094017e-05
step: 30, loss: 5.6958397181006 15e-05
step: 40, loss: 0.013987665995955467
step: 50, loss: 4.106436972506344e-05
step: 60, loss: 1.2181543752376456e-05
step: 70, loss: 1.681537469266914e-05
step: 80, loss: 0.0002659653837326914
step: 90, loss: 1.3414629393082578e-05
step: 100, loss: 9.609190601622686e-05
step: 110, loss: 5.09257115481887e-05
step: 120, loss: 1.892366708489135e-05
step: 130, loss: 0.07298208028078079
step: 140, loss: 0.009320687502622604
step: 150, loss: 1.747137139318511e-05
step: 160, loss: 1.3071902685624082e-05
step: 170, loss: 1.3302837032824755e-05
step: 180, loss: 5.558510747505352e-05
step: 190, loss: 1.661824899201747e-05
```

```
step: 200, loss: 1.4137307516648434e-05
step: 210, loss: 4.534135587164201e-05
step: 220, loss: 1.2822295502701309e-05
step: 230, loss: 1.651758429943584e-05
step: 240, loss: 1.3407170627033338e-05
step: 250, loss: 3.527300577843562e-05
step: 260, loss: 1.4811521396040916e-05
step: 270, loss: 1.5731537132523954e-05
step: 280, loss: 1.3816941645927727e-05
step: 290, loss: 1.0047031537396833e-05
step: 300, loss: 1.277014052902814e-05
step: 310, loss: 2.159854193450883e-05
step: 320, loss: 6.468484934885055e-05
step: 330, loss: 1.3667930943483952e-05
step: 340, loss: 2.9943905246909708e-05
step: 350, loss: 9.868223969533574e-06
step: 360, loss: 0.0035024345852434635
step: 370, loss: 1.3626899090013467e-05
step: 380, loss: 1.8238690245198086e-05
step: 390, loss: 2.0272429537726566e-05
step: 400, loss: 2.8353566449368373e-05
step: 410, loss: 1.6267937098746188e-05
step: 420, loss: 1.7888301954371855e-05
step: 430, loss: 3.1655767088523135e-05
step: 440, loss: 1.4766820640943479e-05
step: 450, loss: 1.4129746887192596e-05
step: 460, loss: 3.532904156600125e-05
step: 470, loss: 6.47989145363681e-05
step: 480, loss: 1.0233295142825227e-05
step: 490, loss: 1.1078918760176748e-05
step: 500, loss: 1.4256439499149565e-05
step: 510, loss: 1.2159222023910843e-05
step: 520, loss: 1.2282160241738893e-05
step: 530, loss: 2.9168428227421828e-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=0.94362292051756, f1=0.9358560221504384,
best_f1=0.9401947148817803
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 [00:23, 248.56it/s]
load_f1 = 0.9481961147086033
real_f1 = 0.9456221198156683
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, 256.60it/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 61.4 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: 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: 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: 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: 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: 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: 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: catalogue<1.1.0,>=0.0.7 in
/usr/local/lib/python3.7/dist-packages (from
spacy>=2.2.2->en_core_web_lg==2.2.5) (1.0.0)
Requirement already satisfied: cymem<2.1.0,>=2.0.2 in
/usr/local/lib/python3.7/dist-packages (from
spacy>=2.2.2->en_core_web_lg==2.2.5) (2.0.6)
Requirement already satisfied: 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: 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: 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: importlib-metadata>=0.20 in
/usr/local/lib/python3.7/dist-packages (from
catalogue<1.1.0,>=0.0.7->spacy>=2.2.2->en_core_web_lg==2.2.5) (4.12.0)
Requirement already satisfied: zipp>=0.5 in /usr/local/lib/python3.7/dist-
packages (from importlib-
metadata>=0.20->catalogue<1.1.0,>=0.0.7->spacy>=2.2.2->en_core_web_lg==2.2.5)
(3.8.1)
Requirement already satisfied: typing-extensions>=3.6.4 in
/usr/local/lib/python3.7/dist-packages (from importlib-
metadata>=0.20->catalogue<1.1.0,>=0.0.7->spacy>=2.2.2->en_core_web_lg==2.2.5)
(4.1.1)
Requirement already satisfied: 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: 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: 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=e3830efc7152fb48af70adb16e5d5ae180f6a29bd00bd6d925968164f4b35bf2
  Stored in directory: /tmp/pip-ephem-wheel-cache-d1wk6cdx/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",

249
```

```
UserWarning)
step: 0, loss: 0.5572822690010071
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.38095238095238093, f1=0.4137931034482759,
best_f1=0.4137931034482759
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.4956880211830139
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.5, f1=0.3880597014925373, best_f1=0.3880597014925373
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.5130605101585388
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.5217391304347826, f1=0.42857142857142855,
best_f1=0.42857142857142855
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.2886781692504883
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.6111111111111112, f1=0.4761904761904762,
```

```
best_f1=0.4761904761904762
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.23767723143100739
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.6956521739130435, f1=0.56, best_f1=0.56
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.2428884506225586
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.7692307692307692, f1=0.5833333333333334,
best_f1=0.5833333333333334
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.022770801559090614
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.7058823529411764,
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: 0.011586287058889866
/usr/local/lib/python3.7/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=0.8666666666666666, f1=0.7096774193548386,
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: 0.002045997651293874
/usr/local/lib/python3.7/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=0.7142857142857143,
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: 0.0043383436277508736
/usr/local/lib/python3.7/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=0.6666666666666666,
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: 0.0028888797387480736
/usr/local/lib/python3.7/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=0.8275862068965518, f1=0.6153846153846153,
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: 0.001958603970706463
/usr/local/lib/python3.7/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=0.8387096774193549, f1=0.7096774193548386,
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: 0.0023249892983585596
/usr/local/lib/python3.7/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=0.8387096774193549, f1=0.7096774193548386,
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: 0.0020695291459560394
/usr/local/lib/python3.7/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=0.8387096774193549, f1=0.7096774193548386,
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: 0.0031346718315035105
/usr/local/lib/python3.7/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=0.8387096774193549, f1=0.7096774193548386,
best_f1=0.7058823529411764
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, 102768.35it/s]
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
load_f1 = 0.8148148148148148
real_f1 = 0.8275862068965518
```

```
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, 245.95it/s]
```

3.1.2 DBLP-ACM - Running the matcher

```
[ ]:  # Step 1:
      # train the wdc model with the save_model flag on
      # it will produce two files *_dev.pt and *_test.pt which are the
      # best checkpoints on the dev set and the test set.
      !CUDA_VISIBLE_DEVICES=0 python train_ditto.py \
        --task Structured/DBLP-ACM \
        --batch_size 32 \
        --max_len 256 \
        --lr 3e-5 \
        --n_epochs 15 \
        --finetuning \
        --lm 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
/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: 0.6193021535873413
step: 10, loss: 0.6028382182121277
step: 20, loss: 0.336209237575531
step: 30, loss: 0.23590384423732758
step: 40, loss: 0.05278477072715759
step: 50, loss: 0.043239179998636246
step: 60, loss: 0.04994328320026398
step: 70, loss: 0.011044876649975777
step: 80, loss: 0.11049789190292358
step: 90, loss: 0.13780522346496582
step: 100, loss: 0.10830534249544144
step: 110, loss: 0.050908807665109634
step: 120, loss: 0.0035712027456611395
step: 130, loss: 0.009702678769826889
step: 140, loss: 0.0020606571342796087
step: 150, loss: 0.022425539791584015
step: 160, loss: 0.01586167886853218
step: 170, loss: 0.07907328754663467
step: 180, loss: 0.004419975448399782
```

```
step: 190, loss: 0.06707999855279922
step: 200, loss: 0.038237232714891434
step: 210, loss: 0.002631847048178315
step: 220, loss: 0.05569899454712868
step: 230, loss: 0.07024960964918137
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.990990990990991, f1=0.9819413092550789,
best_f1=0.9819413092550789
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0015683112433180213
step: 10, loss: 0.0020029633615165949
step: 20, loss: 0.093981072306633
Gradient overflow.  Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 30, loss: 0.12385444343090057
step: 40, loss: 0.031243249773979187
step: 50, loss: 0.0052527617663145065
step: 60, loss: 0.0012547285296022892
step: 70, loss: 0.1365189105272293
step: 80, loss: 0.001963244518265128
step: 90, loss: 0.006182491313666105
step: 100, loss: 0.02682947739958763
step: 110, loss: 0.0056467787362635136
step: 120, loss: 0.0012324111303314567
step: 130, loss: 0.006114790216088295
step: 140, loss: 0.002278511179611087
step: 150, loss: 0.013954339548945427
step: 160, loss: 0.002238851273432374
step: 170, loss: 0.0015297747449949384
step: 180, loss: 0.004328247159719467
step: 190, loss: 0.009159685112535954
step: 200, loss: 0.00183941051363945
step: 210, loss: 0.0006176151800900698
step: 220, loss: 0.021224912256002426
step: 230, loss: 0.01529077347368002
/usr/local/lib/python3.7/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=0.9864864864864865, f1=0.9797297297297298,
best_f1=0.9819413092550789
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
```

An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.004285838454961777
step: 10, loss: 0.005729111842811108
step: 20, loss: 0.0017521979752928019
step: 30, loss: 0.013579221442341805
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 8192.0
step: 40, loss: 0.01920853927731514
step: 50, loss: 0.014165225438773632
step: 60, loss: 0.004203870892524719
step: 70, loss: 0.014325299300253391
step: 80, loss: 0.004078793339431286
step: 90, loss: 0.009199033491313457
step: 100, loss: 0.022869784384965897
step: 110, loss: 0.0062994323670864105
step: 120, loss: 0.011275529861450195
step: 130, loss: 0.012566747143864632
step: 140, loss: 0.0010743646416813135
step: 150, loss: 0.11974519491195679
step: 160, loss: 0.021782128140330315
step: 170, loss: 0.0041000135242938995
step: 180, loss: 0.007230894640088081
step: 190, loss: 0.00922324601560831
step: 200, loss: 0.0304366797208786
step: 210, loss: 0.003625185927376151
step: 220, loss: 0.001758921891450882
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9921259842519685, f1=0.9865168539325843,
best_f1=0.9865168539325843
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0006950552342459559
step: 10, loss: 0.0002966487081721425
step: 20, loss: 0.0005093633662909269
step: 30, loss: 0.07658951729536057
step: 40, loss: 0.11487188935279846
step: 50, loss: 0.0009556713048368692
step: 60, loss: 0.008721468970179558
step: 70, loss: 0.0009021328878588974
step: 80, loss: 0.009417781606316566
step: 90, loss: 0.0029168976470828056

```
step: 100, loss: 0.0038013628218322992
step: 110, loss: 0.0006362044950947165
step: 120, loss: 0.020959651097655296
step: 130, loss: 0.031183717772364616
step: 140, loss: 0.0003318852395750582
step: 150, loss: 0.06388026475906372
step: 160, loss: 0.0074613941833376884
step: 170, loss: 0.0015073007671162486
step: 180, loss: 0.000294468249194324
step: 190, loss: 0.007324710488319397
step: 200, loss: 0.0013880339683964849
step: 210, loss: 0.005590823944658041
step: 220, loss: 0.00012109700764995068
step: 230, loss: 0.02708631567656994
/usr/local/lib/python3.7/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=0.9875706214689265, f1=0.9796380090497738,
best_f1=0.9865168539325843
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00015034859825391322
step: 10, loss: 0.0002445425488986075
step: 20, loss: 0.00019873565179295838
step: 30, loss: 0.00023491588945034891
step: 40, loss: 0.00029727991204708815
step: 50, loss: 0.00020754104480147362
step: 60, loss: 0.0004003612557426095
step: 70, loss: 0.00040014205769635737
step: 80, loss: 0.0007996759959496558
step: 90, loss: 0.0036445753648877144
step: 100, loss: 0.0013334465911611915
step: 110, loss: 0.006281055044382811
step: 120, loss: 0.00023542242706753314
step: 130, loss: 0.0014567594043910503
step: 140, loss: 0.0016339225694537163
step: 150, loss: 0.00044436650932766497
step: 160, loss: 0.0008926105219870806
step: 170, loss: 0.011553811840713024
step: 180, loss: 0.0016944779781624675
step: 190, loss: 0.0044023701921105385
step: 200, loss: 0.00017918231606017798
step: 210, loss: 0.0001341860624961555
step: 220, loss: 0.0008219544542953372
step: 230, loss: 0.0003011637891177088
/usr/local/lib/python3.7/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=0.9887387387387387, f1=0.9853768278965129,
best_f1=0.9865168539325843
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0010849625105038285
step: 10, loss: 0.001299253897741437
step: 20, loss: 0.0009464983595535159
step: 30, loss: 0.00023087323643267155
step: 40, loss: 0.00047442392678931355
step: 50, loss: 0.00016603167750872672
step: 60, loss: 0.00010870487312786281
step: 70, loss: 0.0002949413610622287
step: 80, loss: 0.0005968628684058785
step: 90, loss: 0.0005675117718055844
step: 100, loss: 0.02873929962515831
step: 110, loss: 0.0009316733339801431
step: 120, loss: 0.00014646211163904667
step: 130, loss: 0.0010321519803255796
step: 140, loss: 7.97348257037811e-05
step: 150, loss: 0.00033921233261935413
step: 160, loss: 0.0010025585070252419
step: 170, loss: 0.00020696644787676632
step: 180, loss: 0.011439564637839794
step: 190, loss: 0.0043212128803133965
step: 200, loss: 0.0006798197864554822
step: 210, loss: 0.00018613536667544395
step: 220, loss: 0.00016773266543168575
step: 230, loss: 0.04450768977403641
/usr/local/lib/python3.7/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=0.9876265466816648, f1=0.9819819819819819,
best_f1=0.9865168539325843
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0005936583038419485
step: 10, loss: 7.35617577447556e-05
step: 20, loss: 0.0002037652739090845
step: 30, loss: 0.0003052728425245732
step: 40, loss: 7.630146137671545e-05
step: 50, loss: 6.640841456828639e-05
step: 60, loss: 7.490513962693512e-05
step: 70, loss: 0.023033462464809418
step: 80, loss: 8.477432129438967e-05
step: 90, loss: 3.912909596692771e-05
```

```
step: 100, loss: 6.993638817220926e-05
step: 110, loss: 0.0030598011799156666
step: 120, loss: 5.1580580475274473e-05
step: 130, loss: 5.099885311210528e-05
step: 140, loss: 0.0001020308627630584
step: 150, loss: 0.003471442498266697
step: 160, loss: 0.07782800495624542
step: 170, loss: 0.07519348710775375
step: 180, loss: 0.0020599656272679567
step: 190, loss: 0.00026737514417618513
step: 200, loss: 0.04411260783672333
step: 210, loss: 5.971669452264905e-05
step: 220, loss: 0.0005250527174212039
step: 230, loss: 0.0066178408451378345
/usr/local/lib/python3.7/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=0.9853768278965129,
best_f1=0.9865168539325843
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 8.295443694805726e-05
step: 10, loss: 0.000138251154567115
step: 20, loss: 0.00011075310612795874
step: 30, loss: 0.00033596265711821616
step: 40, loss: 0.00034542157663963735
step: 50, loss: 0.00039062125142663717
step: 60, loss: 6.92759858793579e-05
step: 70, loss: 0.00020223062892910093
step: 80, loss: 0.00010096590995090082
step: 90, loss: 6.513838161481544e-05
step: 100, loss: 8.785323734628037e-05
step: 110, loss: 7.472551078535616e-05
step: 120, loss: 0.00011777172039728612
step: 130, loss: 0.000189416270586662
step: 140, loss: 5.273926944937557e-05
step: 150, loss: 0.00010233960347250104
step: 160, loss: 0.0002732256252784282
step: 170, loss: 9.873783710645512e-05
step: 180, loss: 7.968281715875491e-05
step: 190, loss: 0.00021255530009511858
step: 200, loss: 4.520820584730245e-05
step: 210, loss: 5.2669536671601236e-05
step: 220, loss: 0.0001690410717856139
step: 230, loss: 8.247988444054499e-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=0.9887640449438202, f1=0.9820627802690582,
best_f1=0.9865168539325843
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 4.3398347770562395e-05
step: 10, loss: 0.0007343686302192509
step: 20, loss: 0.005702455528080463
step: 30, loss: 4.5882599806645885e-05
step: 40, loss: 0.12318570911884308
step: 50, loss: 8.117582910927013e-05
step: 60, loss: 7.622037082910538e-05
step: 70, loss: 0.00022046083176974207
step: 80, loss: 0.03753841668367386
step: 90, loss: 0.00014839407231193036
step: 100, loss: 0.0022474073339253664
step: 110, loss: 5.753882942372002e-05
step: 120, loss: 5.5678243370493874e-05
step: 130, loss: 8.1930986198131e-05
step: 140, loss: 0.0008927091839723289
step: 150, loss: 0.0001347194192931056
step: 160, loss: 3.781350824283436e-05
step: 170, loss: 0.0028679214883595705
step: 180, loss: 0.00016156959463842213
step: 190, loss: 0.0001348523801425472
step: 200, loss: 4.31136904808227e-05
step: 210, loss: 0.00022261797857936472
step: 220, loss: 4.8954836529446766e-05
step: 230, loss: 0.00012912553211208433
/usr/local/lib/python3.7/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=0.9841628959276018,
best_f1=0.9865168539325843
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 5.59003819944337e-05
step: 10, loss: 3.111274054390378e-05
step: 20, loss: 0.0003392025828361511
step: 30, loss: 0.000495835323818028
step: 40, loss: 6.98510484653525e-05
step: 50, loss: 0.00024877904797904193
step: 60, loss: 0.003470013150945306
step: 70, loss: 0.0004503911768551916
step: 80, loss: 6.96190691087395e-05
step: 90, loss: 0.00026564663858152926
```

```
step: 100, loss: 5.970222264295444e-05
step: 110, loss: 0.0009270209120586514
step: 120, loss: 0.00010233160719508305
step: 130, loss: 0.003389037912711501
step: 140, loss: 0.02334909327328205
step: 150, loss: 0.019022643566131592
step: 160, loss: 2.206452154496219e-05
step: 170, loss: 6.44273022771813e-05
step: 180, loss: 0.00020630663493648171
step: 190, loss: 8.180316945072263e-05
step: 200, loss: 3.016994560312014e-05
step: 210, loss: 5.303290527081117e-05
step: 220, loss: 3.939342423109338e-05
step: 230, loss: 4.462194556253962e-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=0.9876265466816648, f1=0.9853438556933484,
best_f1=0.9865168539325843
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 5.886347935302183e-05
step: 10, loss: 6.806658348068595e-05
step: 20, loss: 0.00012915610568597913
step: 30, loss: 0.017228569835424423
step: 40, loss: 6.539496826007962e-05
step: 50, loss: 6.660311119404194e-05
step: 60, loss: 4.894944140687585e-05
step: 70, loss: 8.200171578209847e-05
step: 80, loss: 3.8273847167147323e-05
step: 90, loss: 3.2550236937822774e-05
step: 100, loss: 3.501257378957234e-05
step: 110, loss: 7.769319927319884e-05
step: 120, loss: 3.358538378961384e-05
step: 130, loss: 5.192357275518589e-05
step: 140, loss: 3.4073407732648775e-05
step: 150, loss: 0.031779538840055466
step: 160, loss: 3.121322879451327e-05
step: 170, loss: 0.018925338983535767
step: 180, loss: 3.9333033782895654e-05
step: 190, loss: 3.4374748793197796e-05
step: 200, loss: 0.0002912209019996226
step: 210, loss: 2.3420347133651376e-05
step: 220, loss: 4.200778130325489e-05
step: 230, loss: 3.5344110074220225e-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=0.9898534385569334, f1=0.9853107344632768,
best_f1=0.9865168539325843
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 3.5131881304550916e-05
step: 10, loss: 2.8583290259120986e-05
step: 20, loss: 3.223633757443167e-05
step: 30, loss: 7.29974708519876e-05
step: 40, loss: 7.860933692427352e-05
step: 50, loss: 0.0001546482089906931
step: 60, loss: 4.859748514718376e-05
step: 70, loss: 7.033445581328124e-05
step: 80, loss: 3.4263968700543046e-05
step: 90, loss: 3.808020119322464e-05
step: 100, loss: 2.510393642296549e-05
step: 110, loss: 3.398056287551299e-05
step: 120, loss: 2.5782152079045773e-05
step: 130, loss: 2.9835078748874366e-05
step: 140, loss: 2.9339104003156535e-05
step: 150, loss: 8.200985757866874e-05
step: 160, loss: 0.0008621863089501858
step: 170, loss: 2.5670356990303844e-05
step: 180, loss: 2.741714342846535e-05
step: 190, loss: 9.481079177930951e-05
step: 200, loss: 1.882724063761998e-05
step: 210, loss: 3.163782093906775e-05
step: 220, loss: 2.829616641975008e-05
step: 230, loss: 0.04278123751282692
/usr/local/lib/python3.7/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=0.9875706214689265, f1=0.983050847457627,
best_f1=0.9865168539325843
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 3.523590203258209e-05
step: 10, loss: 7.620630640303716e-05
step: 20, loss: 0.00021278296480886638
step: 30, loss: 5.021090692025609e-05
step: 40, loss: 3.4624590625753626e-05
step: 50, loss: 7.300232391571626e-05
step: 60, loss: 2.3573136786581017e-05
step: 70, loss: 5.562159276450984e-05
step: 80, loss: 5.39230095455423e-05
step: 90, loss: 0.00013606349239125848
```

```
step: 100, loss: 2.204573502240237e-05
step: 110, loss: 8.566542237531394e-05
step: 120, loss: 0.021031830459833145
step: 130, loss: 5.7505545555613935e-05
step: 140, loss: 3.2677089620847255e-05
step: 150, loss: 2.703753489186056e-05
step: 160, loss: 3.742436092579737e-05
step: 170, loss: 4.475496825762093e-05
step: 180, loss: 5.542222788790241e-05
step: 190, loss: 2.0607825717888772e-05
step: 200, loss: 5.0953538448084146e-05
step: 210, loss: 1.971756682905834e-05
step: 220, loss: 0.00011241356696700677
step: 230, loss: 2.7085485271527432e-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=0.9887892376681614, f1=0.9820627802690582,
best_f1=0.9865168539325843
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 2.630351264087949e-05
step: 10, loss: 0.0013675604714080691
step: 20, loss: 2.808042336255312e-05
step: 30, loss: 3.4059034078381956e-05
step: 40, loss: 0.01605290174484253
step: 50, loss: 0.0003073741972912103
step: 60, loss: 3.456416743574664e-05
step: 70, loss: 2.8002174076391384e-05
step: 80, loss: 2.8926035156473517e-05
step: 90, loss: 3.314926289021969e-05
step: 100, loss: 2.439621493977029e-05
step: 110, loss: 0.00010708460467867553
step: 120, loss: 1.616754161659628e-05
step: 130, loss: 2.743234290392138e-05
step: 140, loss: 2.9645030735991895e-05
step: 150, loss: 3.34020696755033e-05
step: 160, loss: 3.3354201150359586e-05
step: 170, loss: 2.4004297301871702e-05
step: 180, loss: 2.6608755433699116e-05
step: 190, loss: 4.91113678435795e-05
step: 200, loss: 2.6951160180033185e-05
step: 210, loss: 3.224106694688089e-05
step: 220, loss: 1.7154714441858232e-05
step: 230, loss: 3.39130092470441e-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=0.9875424688561721, f1=0.9852104664391355,
best_f1=0.9865168539325843
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 1.7843865862232633e-05
step: 10, loss: 1.9609229639172554e-05
step: 20, loss: 2.897422746173106e-05
step: 30, loss: 2.580808722996153e-05
step: 40, loss: 3.6579498555511236e-05
step: 50, loss: 0.024446725845336914
step: 60, loss: 3.329087121528573e-05
step: 70, loss: 3.2233507226919755e-05
step: 80, loss: 3.3362070098519325e-05
step: 90, loss: 4.794606866198592e-05
step: 100, loss: 2.4679435227881186e-05
step: 110, loss: 2.4493070668540895e-05
step: 120, loss: 0.00012790843902621418
step: 130, loss: 3.187586844433099e-05
step: 140, loss: 2.1256091713439673e-05
step: 150, loss: 7.092951273079962e-05
step: 160, loss: 4.040612111566588e-05
step: 170, loss: 1.865222475316841e-05
step: 180, loss: 8.235572749981657e-05
step: 190, loss: 7.916631147963926e-05
step: 200, loss: 2.7923984816879965e-05
step: 210, loss: 2.6001882361015305e-05
step: 220, loss: 5.301458077155985e-05
step: 230, loss: 2.2015979993739165e-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=0.9876543209876544, f1=0.9820627802690582,
best_f1=0.9865168539325843
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, 180.66it/s]
load_f1 = 0.990990990990991
real_f1 = 0.9887133182844244
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, 222.93it/s]
```

3.1.3 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 \
       --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: 0.5927374362945557
step: 10, loss: 0.5235815048217773
step: 20, loss: 0.4407440423965454
step: 30, loss: 0.09949018806219101
step: 40, loss: 0.121893972158432
step: 50, loss: 0.23051150143146515
step: 60, loss: 0.05014362186193466
step: 70, loss: 0.07669265568256378
step: 80, loss: 0.027580246329307556
step: 90, loss: 0.09340274333953857
step: 100, loss: 0.09944058954715729
step: 110, loss: 0.06582484394311905
step: 120, loss: 0.10032361000776291
step: 130, loss: 0.14945147931575775
step: 140, loss: 0.11879435181617737
step: 150, loss: 0.06301072984933853
step: 160, loss: 0.021238276734948158
step: 170, loss: 0.2021491974592209
step: 180, loss: 0.09085869044065475
step: 190, loss: 0.02233009971678257
step: 200, loss: 0.17015953361988068
step: 210, loss: 0.10610375553369522
step: 220, loss: 0.17326034605503082
step: 230, loss: 0.19702309370040894
step: 240, loss: 0.07806769758462906
step: 250, loss: 0.03333250805735588
step: 260, loss: 0.14067518711090088
step: 270, loss: 0.012012235820293427
step: 280, loss: 0.029740121215581894
step: 290, loss: 0.08727956563234329
step: 300, loss: 0.04763399437069893
step: 310, loss: 0.12703154981136322
step: 320, loss: 0.0779431089758873
step: 330, loss: 0.07535506039857864
step: 340, loss: 0.04124787449836731
step: 350, loss: 0.015237855724990368
step: 360, loss: 0.020859267562627792
step: 370, loss: 0.09611965715885162
step: 380, loss: 0.03686175122857094
step: 390, loss: 0.1671753227710724
step: 400, loss: 0.3706247806549072
step: 410, loss: 0.07084418833255768
step: 420, loss: 0.09407181292772293
step: 430, loss: 0.1478962004184723
step: 440, loss: 0.037125758826732635
step: 450, loss: 0.006572895683348179
```

```
step: 460, loss: 0.005890555679798126
step: 470, loss: 0.15760014951229095
step: 480, loss: 0.047884728759527206
step: 490, loss: 0.04913824424147606
step: 500, loss: 0.11473333090543747
step: 510, loss: 0.11387835443019867
step: 520, loss: 0.11712421476840973
step: 530, loss: 0.005335430148988962
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9247015610651974, f1=0.9240681086056144,
best_f1=0.9240681086056144
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.1503003090620041
step: 10, loss: 0.15869532525539398
step: 20, loss: 0.005288717336952686
step: 30, loss: 0.012261561118066311
step: 40, loss: 0.029279131442308426
step: 50, loss: 0.0468541756272316
step: 60, loss: 0.019796237349510193
step: 70, loss: 0.025467464700341225
step: 80, loss: 0.0640711858868599
step: 90, loss: 0.01816231943666935
step: 100, loss: 0.013695618137717247
step: 110, loss: 0.017381122335791588
step: 120, loss: 0.048669904470443726
step: 130, loss: 0.2167479246854782
step: 140, loss: 0.07588928937911987
step: 150, loss: 0.15242601931095123
step: 160, loss: 0.021035728976130486
step: 170, loss: 0.03666367009282112
step: 180, loss: 0.014366579242050648
step: 190, loss: 0.13624154031276703
step: 200, loss: 0.008173773996531963
step: 210, loss: 0.051074035465717316
step: 220, loss: 0.08525827527046204
step: 230, loss: 0.006184106692671776
step: 240, loss: 0.03389495238661766
step: 250, loss: 0.117338165640831
step: 260, loss: 0.004539045970886946
step: 270, loss: 0.24687083065509796
```

```
step: 280, loss: 0.024015087634325027
step: 290, loss: 0.014528991654515266
step: 300, loss: 0.02504332922399044
step: 310, loss: 0.011337400414049625
step: 320, loss: 0.07962383329868317
step: 330, loss: 0.08244628459215164
step: 340, loss: 0.017671354115009308
step: 350, loss: 0.0013586889253929257
step: 360, loss: 0.02969709411263466
step: 370, loss: 0.12740464508533478
step: 380, loss: 0.046790193766355515
step: 390, loss: 0.05620989575982094
step: 400, loss: 0.07242139428853989
step: 410, loss: 0.00664899405092001
step: 420, loss: 0.008116215467453003
step: 430, loss: 0.01835963875055313
step: 440, loss: 0.12529420852661133
step: 450, loss: 0.04681434482336044
step: 460, loss: 0.05065586417913437
step: 470, loss: 0.0157249066978693
step: 480, loss: 0.34101375937461853
step: 490, loss: 0.03809993714094162
step: 500, loss: 0.11863058060407639
step: 510, loss: 0.020869795233011246
step: 520, loss: 0.0291425958275795
step: 530, loss: 0.021644845604896545
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9369114877589454, f1=0.9338919925512105,
best_f1=0.9338919925512105
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.051865220069885254
step: 10, loss: 0.10970281809568405
step: 20, loss: 0.06434851884841919
step: 30, loss: 0.050348423421382904
step: 40, loss: 0.002386000007390976
step: 50, loss: 0.05649912729859352
step: 60, loss: 0.040598224848508835
step: 70, loss: 0.004171755164861679
step: 80, loss: 0.0026896465569734573
step: 90, loss: 0.007792051415890455
```

```
step: 100, loss: 0.05166807025671005
step: 110, loss: 0.0012191561982035637
step: 120, loss: 0.001701138447970152
step: 130, loss: 0.0022907631937414408
step: 140, loss: 0.01603493094444275
step: 150, loss: 0.011742318980395794
step: 160, loss: 0.008037303574383259
step: 170, loss: 0.05138927698135376
step: 180, loss: 0.004740141332149506
step: 190, loss: 0.009518489241600037
step: 200, loss: 0.040273286402225494
step: 210, loss: 0.07689685374498367
step: 220, loss: 0.1379166841506958
step: 230, loss: 0.057660069316625595
step: 240, loss: 0.001299678930081427
step: 250, loss: 0.015536455437541008
step: 260, loss: 0.010783091187477112
step: 270, loss: 0.013265757821500301
step: 280, loss: 0.10676033794879913
step: 290, loss: 0.002613582881167531
step: 300, loss: 0.08800067007541656
step: 310, loss: 0.0054809036664664745
step: 320, loss: 0.041116904467344284
step: 330, loss: 0.0015083298785611987
step: 340, loss: 0.004259227775037289
step: 350, loss: 0.003532965201884508
step: 360, loss: 0.04408503696322441
step: 370, loss: 0.013832267373800278
step: 380, loss: 0.006014594342559576
step: 390, loss: 0.041242565959692
step: 400, loss: 0.04192272573709488
step: 410, loss: 0.008495515212416649
step: 420, loss: 0.23933395743370056
step: 430, loss: 0.008403589949011803
step: 440, loss: 0.0006523514748550951
step: 450, loss: 0.08295772224664688
step: 460, loss: 0.023395774886012077
step: 470, loss: 0.019942576065659523
step: 480, loss: 0.0027252009604126215
step: 490, loss: 0.006674472242593765
step: 500, loss: 0.03654877468943596
step: 510, loss: 0.01871003210545862
step: 520, loss: 0.044248707592487335
step: 530, loss: 0.05747614800930023
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
```

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=0.9411764705882353, f1=0.9372693726937269,
best_f1=0.9372693726937269
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.08082915842533112
step: 10, loss: 0.0020953554194420576
step: 20, loss: 0.0012858306290581822
step: 30, loss: 0.0005070571787655354
step: 40, loss: 0.0019557196646928787
step: 50, loss: 0.001029070233926177
step: 60, loss: 0.002362570259720087
step: 70, loss: 0.0031074017751961946
step: 80, loss: 0.09482541680335999
step: 90, loss: 0.049120739102363586
step: 100, loss: 0.0037033604457974434
step: 110, loss: 0.021227220073342323
step: 120, loss: 0.01511509157717228
step: 130, loss: 0.030804777517914772
step: 140, loss: 0.07983237504959106
step: 150, loss: 0.008850577287375927
step: 160, loss: 0.05630253255367279
step: 170, loss: 0.012318223714828491
step: 180, loss: 0.000791952305007726
step: 190, loss: 0.004351758398115635
step: 200, loss: 0.00967482477426529
step: 210, loss: 0.031004520133137703
step: 220, loss: 0.003018850227817893
step: 230, loss: 0.07969392836093903
step: 240, loss: 0.008860599249601364
step: 250, loss: 0.013332413509488106
step: 260, loss: 0.0042426697909832
step: 270, loss: 0.01256580650806427
step: 280, loss: 0.0046903532929718494
step: 290, loss: 0.029456712305545807
step: 300, loss: 0.0002462713164277375
step: 310, loss: 0.004628614988178015
step: 320, loss: 0.0023629090283066034
step: 330, loss: 0.011128289625048637
step: 340, loss: 0.001492982148192823
step: 350, loss: 0.10419569164514542
step: 360, loss: 0.008131015114486217
step: 370, loss: 0.03328173980116844
step: 380, loss: 0.004445487167686224
step: 390, loss: 0.00045850881724618375

```
step: 400, loss: 0.0019900694023817778
Gradient overflow.  Skipping step, loss scaler 0 reducing loss scale to 32768.0
step: 410, loss: 0.009276294149458408
step: 420, loss: 0.15023492276668549
step: 430, loss: 0.15992222726345062
step: 440, loss: 0.008624912239611149
step: 450, loss: 0.0027175003197044134
step: 460, loss: 0.00535786896944046
step: 470, loss: 0.027508936822414398
step: 480, loss: 0.1674058735370636
step: 490, loss: 0.0032267477363348007
step: 500, loss: 0.0034340249840170145
step: 510, loss: 0.0098974955619952679
step: 520, loss: 0.021593378856778145
step: 530, loss: 0.02840033546090126
/usr/local/lib/python3.7/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=0.9374130737134909, f1=0.93646408839779,
best_f1=0.9372693726937269
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.011277435347437859
step: 10, loss: 0.048425838351249695
step: 20, loss: 0.002728687133640051
step: 30, loss: 0.015858685597777367
step: 40, loss: 0.02764544077217579
step: 50, loss: 0.0013592962641268969
step: 60, loss: 0.025462456047534943
step: 70, loss: 0.0018305593403056264
step: 80, loss: 0.005319193936884403
step: 90, loss: 0.006404025945812464
step: 100, loss: 0.013199505396187305
step: 110, loss: 0.002773130312561989
step: 120, loss: 0.006726904306560755
step: 130, loss: 0.0006713983602821827
step: 140, loss: 0.009325813502073288
step: 150, loss: 0.0035772393457591534
step: 160, loss: 0.000543290632776916
step: 170, loss: 0.01895514316856861
step: 180, loss: 0.0013714261585846543
step: 190, loss: 0.001076143584214151
step: 200, loss: 0.00047951401211321354
step: 210, loss: 0.0004134147020522505
step: 220, loss: 0.0007821741164661944
step: 230, loss: 0.007805663626641035
step: 240, loss: 0.0019836423452943563
```

```
step: 250, loss: 0.0017732587875798345
step: 260, loss: 0.007078444119542837
step: 270, loss: 0.00025938707403838634
step: 280, loss: 0.007359933573752642
step: 290, loss: 0.10254016518592834
step: 300, loss: 0.0031982867512851954
step: 310, loss: 0.010183055885136127
step: 320, loss: 0.06543801724910736
step: 330, loss: 0.009052885696291924
step: 340, loss: 0.008481933735311031
step: 350, loss: 0.0007028573309071362
step: 360, loss: 0.0012421456631273031
step: 370, loss: 0.002566548530012369
step: 380, loss: 0.0006195184541866183
step: 390, loss: 0.00032793483114801347
step: 400, loss: 0.004097499884665012
step: 410, loss: 0.0018665895331650972
step: 420, loss: 0.005454482976347208
step: 430, loss: 0.00955807976424694
step: 440, loss: 0.019578440114855766
step: 450, loss: 0.001958937617018819
step: 460, loss: 0.0004572531906887889
step: 470, loss: 0.002482407260686159
step: 480, loss: 0.10424744337797165
step: 490, loss: 0.0026847461704164743
step: 500, loss: 0.005619560834020376
step: 510, loss: 0.008116436190903187
step: 520, loss: 0.01093999668955803
step: 530, loss: 0.0022830553352832794
/usr/local/lib/python3.7/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=0.929170549860205, f1=0.9408476944573824,
best_f1=0.9372693726937269
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00024525861954316497
step: 10, loss: 0.0016301436116918921
step: 20, loss: 0.08088165521621704
step: 30, loss: 0.005074262153357267
step: 40, loss: 0.007275882177054882
step: 50, loss: 0.018346646800637245
step: 60, loss: 0.00043231097515672445
step: 70, loss: 0.0006387311150319874
step: 80, loss: 0.004465554840862751
step: 90, loss: 0.0007930507999844849
step: 100, loss: 0.010565503500401974
```

```
step: 110, loss: 0.0021816850639879704
step: 120, loss: 0.007000960875302553
step: 130, loss: 0.00030387454899027944
step: 140, loss: 0.07574258744716644
step: 150, loss: 0.02056661993265152
step: 160, loss: 0.004402942024171352
step: 170, loss: 0.0020468842703849077
step: 180, loss: 0.0009495448903180659
step: 190, loss: 0.0013564040418714285
step: 200, loss: 0.0013748990604653955
step: 210, loss: 0.01725897006690502
step: 220, loss: 0.0016213225899264216
step: 230, loss: 0.0003027695929631591
step: 240, loss: 0.07506008446216583
step: 250, loss: 0.0006510980892926455
step: 260, loss: 6.404404120985419e-05
step: 270, loss: 0.054865773767232895
step: 280, loss: 0.045882973819971085
step: 290, loss: 0.0069173588417470455
step: 300, loss: 0.0004299620632082224
step: 310, loss: 0.0002382943785050884
step: 320, loss: 0.0002980373101308942
step: 330, loss: 8.631907257949933e-05
step: 340, loss: 0.012719100341200829
step: 350, loss: 0.0005143595626577735
step: 360, loss: 0.0039063915610313416
step: 370, loss: 0.0018661273643374443
step: 380, loss: 0.00266736070625484
step: 390, loss: 0.11395284533500671
step: 400, loss: 0.005838700570166111
step: 410, loss: 0.0007502158405259252
step: 420, loss: 0.003911091014742851
step: 430, loss: 0.0005800843355245888
step: 440, loss: 0.00013372224930208176
step: 450, loss: 0.0002514382067602128
step: 460, loss: 0.0006865655886940658
step: 470, loss: 0.0006567159434780478
step: 480, loss: 0.010569276288151741
step: 490, loss: 0.010517158545553684
step: 500, loss: 0.14381174743175507
step: 510, loss: 0.05439369007945061
step: 520, loss: 0.0057329945266246796
step: 530, loss: 0.008279597386717796
/usr/local/lib/python3.7/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=0.9311475409836065, f1=0.9348230912476723,
best_f1=0.9372693726937269
```

```
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0007734433712311089
step: 10, loss: 0.003562171943485737
step: 20, loss: 0.0016769309295341372
step: 30, loss: 0.0014724972425028682
step: 40, loss: 0.003215551609173417
step: 50, loss: 0.008467761799693108
step: 60, loss: 0.0007137763896025717
step: 70, loss: 0.005451794248074293
step: 80, loss: 0.0017968285828828812
step: 90, loss: 0.010059304535388947
step: 100, loss: 0.009749219752848148
step: 110, loss: 0.0012687514536082745
step: 120, loss: 0.008857181295752525
step: 130, loss: 0.01166630256921053
step: 140, loss: 9.7722775535658e-05
step: 150, loss: 0.00039049627957865596
step: 160, loss: 0.00042781123192980886
step: 170, loss: 0.0010249684564769268
step: 180, loss: 0.0014440594241023064
step: 190, loss: 0.002678381744772196
step: 200, loss: 0.00021115860727149993
step: 210, loss: 0.0034925946965813637
step: 220, loss: 0.00023548926401417702
step: 230, loss: 0.03070197068154812
step: 240, loss: 0.0011787381954491138
step: 250, loss: 0.0014950195327401161
step: 260, loss: 0.0001657562970649451
step: 270, loss: 0.0002561829751357436
step: 280, loss: 0.000185706332558766
step: 290, loss: 0.0005668188678100705
step: 300, loss: 0.009097833186388016
step: 310, loss: 0.0015923258615657687
step: 320, loss: 0.008829004131257534
step: 330, loss: 0.007600284647196531
step: 340, loss: 0.03609834238886833
step: 350, loss: 0.0021838396787643433
step: 360, loss: 0.0026836737524718046
step: 370, loss: 0.029905417934060097
step: 380, loss: 0.001477581448853016
step: 390, loss: 0.013218735344707966
step: 400, loss: 0.014260786585509777
step: 410, loss: 0.004537804052233696
step: 420, loss: 0.00255571654997766
step: 430, loss: 0.0004965770640410483
step: 440, loss: 0.0012835139641538262
```

```
step: 450, loss: 0.0009881958831101656
step: 460, loss: 0.00019660871475934982
step: 470, loss: 0.006297852378338575
step: 480, loss: 0.07533234357833862
step: 490, loss: 0.0009439790737815201
step: 500, loss: 0.0002379032812314108
step: 510, loss: 0.000983702833764255
step: 520, loss: 0.042054854333400726
step: 530, loss: 0.0034413228277117014
/usr/local/lib/python3.7/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=0.9384328358208955, f1=0.9405204460966543,
best_f1=0.9372693726937269
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.06799587607383728
step: 10, loss: 0.06703253835439682
step: 20, loss: 0.10876984149217606
step: 30, loss: 0.00018859707051888108
step: 40, loss: 0.00286857015453279
step: 50, loss: 0.0030753803439438343
step: 60, loss: 0.0006892528035677969
step: 70, loss: 0.0003699446970131248
step: 80, loss: 0.01659478060901165
step: 90, loss: 0.00037507389788515866
step: 100, loss: 0.00011599207937251776
step: 110, loss: 0.01781618408858776
step: 120, loss: 0.029866747558116913
step: 130, loss: 9.027989290188998e-05
step: 140, loss: 0.0049683088436722755
step: 150, loss: 0.002326685469597578
step: 160, loss: 0.011545628309249878
step: 170, loss: 0.00270992424339056
step: 180, loss: 0.00020986229355912656
step: 190, loss: 0.004741838667541742
step: 200, loss: 0.0014775629388168454
step: 210, loss: 0.0003584629448596388
step: 220, loss: 0.002300610765814781
step: 230, loss: 0.034497227519750595
step: 240, loss: 0.0017643460305407643
step: 250, loss: 0.0007206203881651163
step: 260, loss: 0.00012421926658134907
step: 270, loss: 0.008972407318651676
step: 280, loss: 0.008146538399159908
step: 290, loss: 0.0012867182958871126
step: 300, loss: 0.0012046464253216982
```

```
step: 310, loss: 0.0021315428894013166
step: 320, loss: 0.009154465980827808
step: 330, loss: 0.00010995581396855414
step: 340, loss: 0.00014741953054908663
step: 350, loss: 0.00010869866673601791
step: 360, loss: 0.00016249343752861023
step: 370, loss: 4.237650500726886e-05
step: 380, loss: 0.0005208896473050117
step: 390, loss: 0.12683922052383423
step: 400, loss: 5.883727135369554e-05
step: 410, loss: 0.00013389086234383285
step: 420, loss: 0.004310526419430971
step: 430, loss: 0.04041717201471329
step: 440, loss: 0.00023342361964751035
step: 450, loss: 0.0022320898715406656
step: 460, loss: 0.0014062529662624002
step: 470, loss: 4.632192212739028e-05
step: 480, loss: 4.295071994420141e-05
step: 490, loss: 0.00013429457612801343
step: 500, loss: 0.0002922705898527056
step: 510, loss: 0.004207635764032602
step: 520, loss: 0.0016502132639288902
step: 530, loss: 7.392298721242696e-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=0.9328984156570364, f1=0.9372384937238494,
best_f1=0.9372693726937269
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.011926073580980301
step: 10, loss: 0.0002694444847293198
step: 20, loss: 0.011160403490066528
step: 30, loss: 0.00010812284017447382
step: 40, loss: 4.67242207378149e-05
step: 50, loss: 2.8017288059345447e-05
step: 60, loss: 7.469010597560555e-05
step: 70, loss: 3.5529545129975304e-05
step: 80, loss: 0.001678480301052332
step: 90, loss: 0.003310411935672164
step: 100, loss: 3.749705138034187e-05
step: 110, loss: 3.8573922211071476e-05
step: 120, loss: 0.0003927024081349373
step: 130, loss: 0.009734662249684334
step: 140, loss: 0.0005683638737536967
step: 150, loss: 4.2309540731366724e-05
step: 160, loss: 0.0016282713040709496
```

```
step: 170, loss: 0.018182404339313507
step: 180, loss: 3.521380131132901e-05
step: 190, loss: 5.467686787596904e-05
step: 200, loss: 5.231235263636336e-05
step: 210, loss: 5.517040699487552e-05
Gradient overflow.  Skipping step, loss scaler 0 reducing loss scale to 32768.0
step: 220, loss: 0.0002553851227276027
step: 230, loss: 3.445038964855485e-05
step: 240, loss: 6.001653309795074e-05
step: 250, loss: 0.00016137165948748589
step: 260, loss: 0.00014399572683032602
step: 270, loss: 3.576425297069363e-05
step: 280, loss: 0.00034955187584273517
step: 290, loss: 0.04327256605029106
step: 300, loss: 0.00011769047705456614
step: 310, loss: 0.02648099884390831
step: 320, loss: 0.0022472140844911337
step: 330, loss: 0.0007391462568193674
step: 340, loss: 0.00010697147808969021
step: 350, loss: 0.003315637120977044
step: 360, loss: 9.466869232710451e-05
step: 370, loss: 0.00021010944328736514
step: 380, loss: 0.0005821297527290881
step: 390, loss: 0.0016664366703480482
step: 400, loss: 0.02617499604821205
step: 410, loss: 9.941023745341226e-05
step: 420, loss: 0.00010539124195929617
step: 430, loss: 3.9377489883918315e-05
step: 440, loss: 0.0006394766387529671
step: 450, loss: 0.0003217713674530387
step: 460, loss: 0.0002611148520372808
step: 470, loss: 2.5368668502778746e-05
step: 480, loss: 4.887972318101674e-05
step: 490, loss: 0.00020186349865980446
step: 500, loss: 0.002624484710395336
step: 510, loss: 0.00021025861497037113
step: 520, loss: 5.296014933264814e-05
step: 530, loss: 0.000247227173531428
/usr/local/lib/python3.7/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=0.9395477618827872, f1=0.9362292051756008,
best_f1=0.9372693726937269
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 7.72755010984838e-05
step: 10, loss: 5.97715079493355e-05
```

```
step: 20, loss: 3.91232970287092e-05
step: 30, loss: 3.5258326533949e-05
step: 40, loss: 0.00027139473240822554
step: 50, loss: 0.0008254731656052172
step: 60, loss: 7.742566231172532e-05
step: 70, loss: 7.259551784954965e-05
step: 80, loss: 2.8847942303400487e-05
step: 90, loss: 5.1346298278076574e-05
step: 100, loss: 2.93955654342426e-05
step: 110, loss: 2.662358201632742e-05
step: 120, loss: 4.590044045471586e-05
step: 130, loss: 0.00010036721505457535
step: 140, loss: 0.011874108575284481
step: 150, loss: 7.415178697556257e-05
step: 160, loss: 0.00012780888937413692
step: 170, loss: 0.00013325655891094357
step: 180, loss: 8.859784429660067e-05
step: 190, loss: 6.968944217078388e-05
step: 200, loss: 5.898157542105764e-05
step: 210, loss: 0.00010636328806867823
step: 220, loss: 0.14742842316627502
step: 230, loss: 0.004637870006263256
step: 240, loss: 0.0012472757371142507
step: 250, loss: 6.655495235463604e-05
step: 260, loss: 0.0012959244195371866
step: 270, loss: 0.00014233242836780846
step: 280, loss: 5.1499482651706785e-05
step: 290, loss: 4.0800427086651325e-05
step: 300, loss: 3.728066440089606e-05
step: 310, loss: 0.00019831545068882406
step: 320, loss: 0.0006812619976699352
step: 330, loss: 0.001214688876643777
step: 340, loss: 0.00011893994087586184
step: 350, loss: 0.00048103596782311797
step: 360, loss: 0.0001939982030307874
step: 370, loss: 0.004417775198817253
step: 380, loss: 0.000122765006381087
step: 390, loss: 0.00043578058830462396
step: 400, loss: 0.0010499906493350863
step: 410, loss: 0.00382587150670588
step: 420, loss: 0.0019727922044694424
step: 430, loss: 0.00035061544622294605
step: 440, loss: 0.012526760809123516
step: 450, loss: 0.00010241942072752863
step: 460, loss: 0.0019655926153063774
step: 470, loss: 8.44297610456124e-05
step: 480, loss: 4.882699067820795e-05
step: 490, loss: 0.00012656931357923895
```

```
step: 500, loss: 0.0022998214699327946
step: 510, loss: 3.321706390124746e-05
step: 520, loss: 0.00016481241618748754
step: 530, loss: 0.0016329140635207295
/usr/local/lib/python3.7/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=0.9378214118747078, f1=0.936269915651359,
best_f1=0.9372693726937269
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00503041036427021
step: 10, loss: 0.00035248565836809576
step: 20, loss: 0.0003822758444584906
step: 30, loss: 4.65074299427215e-05
step: 40, loss: 0.00025793988606892526
step: 50, loss: 0.000133704103063792
step: 60, loss: 0.0314672626554966
step: 70, loss: 0.008620085194706917
step: 80, loss: 6.869986827950925e-05
step: 90, loss: 0.00032115919748321176
step: 100, loss: 0.0004788497753906995
step: 110, loss: 0.00011981306306552142
step: 120, loss: 7.510127034038305e-05
step: 130, loss: 3.2449501304654405e-05
step: 140, loss: 0.0006699027144350111
step: 150, loss: 0.00010125806875294074
step: 160, loss: 0.00011229427764192224
step: 170, loss: 5.205456182011403e-05
step: 180, loss: 0.00017813453450798988
step: 190, loss: 0.0010485276579856873
step: 200, loss: 0.002807005774229765
step: 210, loss: 0.0012249695137143135
step: 220, loss: 0.004519581329077482
step: 230, loss: 0.0002856283390428871
step: 240, loss: 2.938354191428516e-05
step: 250, loss: 2.528996265027672e-05
step: 260, loss: 3.3778971555875614e-05
step: 270, loss: 8.180732402252033e-05
step: 280, loss: 5.865234197699465e-05
step: 290, loss: 6.375994416885078e-05
step: 300, loss: 6.080327511881478e-05
step: 310, loss: 0.00012479216093197465
step: 320, loss: 0.00015232579607982188
step: 330, loss: 0.0010676145320758224
step: 340, loss: 3.286992432549596e-05
step: 350, loss: 0.0003992553392890841
```

```
step: 360, loss: 2.7584435883909464e-05
step: 370, loss: 0.0005783025408163667
step: 380, loss: 3.174941593897529e-05
step: 390, loss: 3.846706385957077e-05
step: 400, loss: 3.470237788860686e-05
step: 410, loss: 3.226376065867953e-05
step: 420, loss: 5.561229045270011e-05
step: 430, loss: 2.4482045773766004e-05
step: 440, loss: 0.00034240781678818166
step: 450, loss: 0.007487325929105282
step: 460, loss: 7.956069021020085e-05
step: 470, loss: 3.025932892342098e-05
step: 480, loss: 0.001854085479862988
step: 490, loss: 3.786845991271548e-05
step: 500, loss: 0.029485665261745453
step: 510, loss: 3.8027817936381325e-05
step: 520, loss: 3.8945567212067544e-05
step: 530, loss: 5.135711035109125e-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=0.9372026641294006, f1=0.932503590234562,
best_f1=0.9372693726937269
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 2.689529355848208e-05
step: 10, loss: 0.00022227730369195342
step: 20, loss: 5.7116463722195476e-05
step: 30, loss: 0.00038012044387869537
step: 40, loss: 2.9216087568784133e-05
step: 50, loss: 0.09654591232538223
step: 60, loss: 0.0015381363919004798
step: 70, loss: 4.310827353037894e-05
step: 80, loss: 3.369259138708003e-05
step: 90, loss: 0.000107056628621649
step: 100, loss: 0.0002546820614952594
step: 110, loss: 3.852828376693651e-05
step: 120, loss: 1.6733754819142632e-05
step: 130, loss: 0.001653663464821875
step: 140, loss: 0.00012041413719998673
step: 150, loss: 1.7553216821397655e-05
step: 160, loss: 1.8898077541962266e-05
step: 170, loss: 2.755481909844093e-05
step: 180, loss: 5.438872176455334e-05
step: 190, loss: 4.00175922550261e-05
step: 200, loss: 0.0001432657882105559
step: 210, loss: 4.4464755774242803e-05
```

```
step: 220, loss: 3.716937135322951e-05
step: 230, loss: 2.419086013105698e-05
step: 240, loss: 0.0005161790177226067
step: 250, loss: 1.8525506675359793e-05
step: 260, loss: 3.698961154441349e-05
step: 270, loss: 3.943780029658228e-05
step: 280, loss: 0.0018701516091823578
step: 290, loss: 0.00013000896433368325
step: 300, loss: 0.0005019104573875666
step: 310, loss: 0.0011232278775423765
step: 320, loss: 2.0138393665547483e-05
step: 330, loss: 0.0003111725382041186
step: 340, loss: 0.00019417675503063947
step: 350, loss: 0.0034665013663470745
step: 360, loss: 0.00013250214396975935
step: 370, loss: 4.1395356674911454e-05
step: 380, loss: 0.00023805254022590816
step: 390, loss: 3.5636883694678545e-05
step: 400, loss: 3.158908657496795e-05
step: 410, loss: 0.00039585481863468885
step: 420, loss: 0.007285858504474163
step: 430, loss: 0.023881467059254646
step: 440, loss: 2.0756891899509355e-05
step: 450, loss: 5.152906305738725e-05
step: 460, loss: 1.7605456378078088e-05
step: 470, loss: 2.057435267488472e-05
step: 480, loss: 6.835618842160329e-05
step: 490, loss: 0.0009392238571308553
step: 500, loss: 2.6455731131136417e-05
step: 510, loss: 8.051779877860099e-05
step: 520, loss: 2.388966277067084e-05
step: 530, loss: 3.6687579267891124e-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=0.9386213408876298, f1=0.9381199811053378,
best_f1=0.9372693726937269
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00440469104796648
step: 10, loss: 0.0010675902012735605
step: 20, loss: 3.107814700342715e-05
step: 30, loss: 3.733986523002386e-05
step: 40, loss: 1.97101580852177e-05
step: 50, loss: 0.0021939247380942106
step: 60, loss: 2.3401387807098217e-05
step: 70, loss: 4.306216578697786e-05
```

```
step: 80, loss: 2.1225401724223047e-05
step: 90, loss: 0.02461281791329384
step: 100, loss: 9.514429257251322e-05
step: 110, loss: 2.9950817406643182e-05
step: 120, loss: 2.5476319933659397e-05
step: 130, loss: 3.2564406865276396e-05
step: 140, loss: 2.3580501874675974e-05
step: 150, loss: 2.596045669633895e-05
step: 160, loss: 4.7134100896073505e-05
step: 170, loss: 2.1185222067288123e-05
step: 180, loss: 0.004596621263772249
step: 190, loss: 3.905143603333272e-05
step: 200, loss: 2.2872596673551016e-05
step: 210, loss: 3.164337977068499e-05
step: 220, loss: 3.7731919292127714e-05
step: 230, loss: 1.8059872672893107e-05
step: 240, loss: 0.00018580886535346508
step: 250, loss: 6.885027687530965e-05
step: 260, loss: 1.4595459106203634e-05
step: 270, loss: 1.66889549291227e-05
step: 280, loss: 3.978003223892301e-05
step: 290, loss: 4.9779282562667504e-05
step: 300, loss: 2.2433083358919248e-05
step: 310, loss: 0.0011569684138521552
step: 320, loss: 0.0004577790678013116
step: 330, loss: 6.883658352307975e-05
step: 340, loss: 1.7460099115851335e-05
step: 350, loss: 5.2553692512447014e-05
step: 360, loss: 1.8085840565618128e-05
step: 370, loss: 5.7479250244796276e-05
step: 380, loss: 0.00013507767289411277
step: 390, loss: 0.00027401160332374275
step: 400, loss: 0.00016665221482980996
step: 410, loss: 0.0010000012116506696
step: 420, loss: 2.2142525267554447e-05
step: 430, loss: 0.0027553376276046038
step: 440, loss: 2.892489283112809e-05
step: 450, loss: 3.007252780662384e-05
step: 460, loss: 0.0010378542356193066
step: 470, loss: 0.00027996173594146967
step: 480, loss: 3.787223613471724e-05
step: 490, loss: 4.085967520950362e-05
step: 500, loss: 4.1406514355912805e-05
step: 510, loss: 3.913118052878417e-05
step: 520, loss: 0.0019692620262503624
step: 530, loss: 1.8901620933320373e-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=0.9342789598108747, f1=0.9332702318977756,
best_f1=0.9372693726937269
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 2.565096838225145e-05
step: 10, loss: 1.1164598618051969e-05
step: 20, loss: 0.00026174605591222644
step: 30, loss: 2.8825168556068093e-05
step: 40, loss: 3.27176385326311e-05
step: 50, loss: 0.00011363835074007511
step: 60, loss: 3.643763193394989e-05
step: 70, loss: 2.9479473596438766e-05
step: 80, loss: 2.3919357772683725e-05
step: 90, loss: 1.7810234567150474e-05
step: 100, loss: 4.840381006943062e-05
step: 110, loss: 2.029120150837116e-05
step: 120, loss: 2.517788379918784e-05
step: 130, loss: 0.022155873477458954
step: 140, loss: 0.0003215074248146266
step: 150, loss: 1.371259531879332e-05
step: 160, loss: 0.0002720175834838301
step: 170, loss: 0.0014779387274757028
step: 180, loss: 5.0830327381845564e-05
step: 190, loss: 0.0006724981940351427
step: 200, loss: 7.00897944625467e-05
step: 210, loss: 7.49796672607772e-05
step: 220, loss: 1.639848414924927e-05
step: 230, loss: 2.5196712158503942e-05
step: 240, loss: 1.4595476386602968e-05
step: 250, loss: 1.7635094991419464e-05
step: 260, loss: 3.628390913945623e-05
step: 270, loss: 1.7772903447621502e-05
Gradient overflow.  Skipping step, loss scaler 0 reducing loss scale to 32768.0
step: 280, loss: 1.6320194845320657e-05
step: 290, loss: 5.4237429139902815e-05
step: 300, loss: 1.992245961446315e-05
step: 310, loss: 0.0007112135645002127
step: 320, loss: 4.472491855267435e-05
step: 330, loss: 3.9687663957010955e-05
step: 340, loss: 0.00011470106255728751
step: 350, loss: 2.206015597039368e-05
step: 360, loss: 8.08430922916159e-05
step: 370, loss: 2.1352761905291118e-05
step: 380, loss: 0.0005731277633458376
step: 390, loss: 0.018564000725746155
step: 400, loss: 7.32035405235365e-05
```

```
step: 410, loss: 3.227445631637238e-05
step: 420, loss: 4.488021295401268e-05
step: 430, loss: 2.2473945136880502e-05
step: 440, loss: 4.316014383221045e-05
step: 450, loss: 0.00021949289657641202
step: 460, loss: 2.5211209504050203e-05
step: 470, loss: 1.5098359654075466e-05
step: 480, loss: 1.895013701869175e-05
step: 490, loss: 2.542346737754997e-05
step: 500, loss: 1.1850042938021943e-05
step: 510, loss: 5.03391565871425e-05
step: 520, loss: 6.34497991995886e-05
step: 530, loss: 1.8976092178490944e-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=0.9358914365933552, f1=0.9388136384866884,
best_f1=0.9372693726937269
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 2.2399366571335122e-05
step: 10, loss: 0.00010113115422427654
step: 20, loss: 3.914320404874161e-05
step: 30, loss: 2.7635587684926577e-05
step: 40, loss: 0.07881198823451996
step: 50, loss: 6.331608892651275e-05
step: 60, loss: 1.3720030437980313e-05
step: 70, loss: 4.462393917492591e-05
step: 80, loss: 3.548246604623273e-05
step: 90, loss: 2.445202881062869e-05
step: 100, loss: 0.00011924151476705447
step: 110, loss: 2.35506086028181e-05
step: 120, loss: 0.0001391183614032343
step: 130, loss: 0.0780295729637146
step: 140, loss: 1.949714896909427e-05
step: 150, loss: 9.354241046821699e-05
step: 160, loss: 1.850687840487808e-05
step: 170, loss: 2.4678649424458854e-05
step: 180, loss: 1.0695219316403382e-05
step: 190, loss: 2.5975201424444094e-05
step: 200, loss: 1.8648348486749455e-05
step: 210, loss: 0.001919183530844748
step: 220, loss: 1.7813883459893987e-05
Gradient overflow.  Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 230, loss: 3.274998380220495e-05
step: 240, loss: 2.6573854484013282e-05
step: 250, loss: 2.7805797799373977e-05
```

```
step: 260, loss: 1.740427251206711e-05
step: 270, loss: 1.820517354644835e-05
step: 280, loss: 1.7162066797027364e-05
step: 290, loss: 1.2479586985136848e-05
step: 300, loss: 1.682675429037772e-05
step: 310, loss: 7.158584048738703e-05
step: 320, loss: 1.617471752979327e-05
step: 330, loss: 2.460783434798941e-05
step: 340, loss: 2.4212080461438745e-05
step: 350, loss: 1.439800143998582e-05
step: 360, loss: 0.0023999824188649654
step: 370, loss: 0.0012463576858863235
step: 380, loss: 4.973066461388953e-05
step: 390, loss: 6.53021561447531e-05
step: 400, loss: 2.1572093828581274e-05
step: 410, loss: 1.8093343896907754e-05
step: 420, loss: 2.988118285429664e-05
step: 430, loss: 6.167616811580956e-05
step: 440, loss: 9.291413152823225e-05
step: 450, loss: 1.2989889000891708e-05
step: 460, loss: 1.9896237063221633e-05
step: 470, loss: 0.0038878240156918764
step: 480, loss: 1.1443987204984296e-05
step: 490, loss: 1.5429930499522015e-05
step: 500, loss: 2.0227824279572815e-05
step: 510, loss: 1.9869738025590777e-05
step: 520, loss: 1.8566219296189956e-05
step: 530, loss: 3.054370608879253e-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=0.9354686764013189, f1=0.9353468617272298,
best_f1=0.9372693726937269
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 [00:25, 226.71it/s]
load_f1 = 0.9372114496768237
real_f1 = 0.9359742054352833
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, 230.62it/s]
```

3.1.4 Amazon-Google - Running the matcher

```
[ ]:  # Step 1:
      # train the wdc model with the save_model flag on
      # it will produce two files *_dev.pt and *_test.pt which are the
      # best checkpoints on the dev set and the test set.
      !CUDA_VISIBLE_DEVICES=0 python train_ditto.py \
        --task Structured/Amazon-Google \
        --batch_size 32 \
        --max_len 256 \
        --lr 3e-5 \
        --n_epochs 15 \
        --finetuning \
        --lm 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: 0.5733955502510071
step: 10, loss: 0.3713187277317047
step: 20, loss: 0.36709970235824585
step: 30, loss: 0.3184810280799866
step: 40, loss: 0.18824201822280884
step: 50, loss: 0.39790093898773193
step: 60, loss: 0.18135006725788116
Gradient overflow.  Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 70, loss: 0.1604800820350647
step: 80, loss: 0.2055835723876953
step: 90, loss: 0.3578401505947113
step: 100, loss: 0.4906202554702759
step: 110, loss: 0.24414187669754028
step: 120, loss: 0.18905718624591827
step: 130, loss: 0.19495978951454163
step: 140, loss: 0.16580449044704437
step: 150, loss: 0.18106231093406677
step: 160, loss: 0.30233222246170044
step: 170, loss: 0.29064950346946716
step: 180, loss: 0.13148073852062225
step: 190, loss: 0.2088148444890976
step: 200, loss: 0.28313392400741577
step: 210, loss: 0.25537511706352234
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.5850091407678244, f1=0.6077348066298343,
best_f1=0.6077348066298343
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.11504288017749786
step: 10, loss: 0.2866349518299103
step: 20, loss: 0.274432510137558
step: 30, loss: 0.21914899349212646
step: 40, loss: 0.17545461654663086
step: 50, loss: 0.16655494272708893
step: 60, loss: 0.4666720926761627
step: 70, loss: 0.1616833359003067
step: 80, loss: 0.23162460327148438
step: 90, loss: 0.09535705298185349
step: 100, loss: 0.01685061678290367
step: 110, loss: 0.05549604073166847
step: 120, loss: 0.13794544339179993
```

```
step: 130, loss: 0.01618514396250248
step: 140, loss: 0.20250266790390015
step: 150, loss: 0.2224917709827423
step: 160, loss: 0.171671062707901
step: 170, loss: 0.14164316654205322
step: 180, loss: 0.18674582242965698
step: 190, loss: 0.2580021023750305
step: 200, loss: 0.03839583694934845
step: 210, loss: 0.11288551986217499
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.6775510204081632, f1=0.632967032967033,
best_f1=0.632967032967033
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.05751253664493561
step: 10, loss: 0.14634303748607635
step: 20, loss: 0.05875386297702789
step: 30, loss: 0.23320911824703217
step: 40, loss: 0.0607774518430233
step: 50, loss: 0.1275242567062378
step: 60, loss: 0.15855225920677185
step: 70, loss: 0.07789168506860733
step: 80, loss: 0.17176412045955658
step: 90, loss: 0.036958497017621994
step: 100, loss: 0.12660177052021027
step: 110, loss: 0.12890490889549255
step: 120, loss: 0.12659011781215668
step: 130, loss: 0.08473599702119827
step: 140, loss: 0.09986744821071625
step: 150, loss: 0.22099198400974274
step: 160, loss: 0.04067911580204964
step: 170, loss: 0.19114315509796143
step: 180, loss: 0.11946108192205429
step: 190, loss: 0.18460263311862946
step: 200, loss: 0.062298654526472092
step: 210, loss: 0.1366962492465973
/usr/local/lib/python3.7/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=0.6692456479690523, f1=0.6482213438735178,
best_f1=0.632967032967033
```

```
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.051812998950481415
step: 10, loss: 0.06404311209917068
step: 20, loss: 0.08068463951349258
step: 30, loss: 0.1578686684370041
step: 40, loss: 0.00781615637242794
step: 50, loss: 0.09162185341119766
step: 60, loss: 0.09407063573598862
step: 70, loss: 0.2025119811296463
step: 80, loss: 0.03521097078919411
step: 90, loss: 0.010519836097955704
step: 100, loss: 0.2267012745141983
step: 110, loss: 0.1093812808394432
step: 120, loss: 0.14667707681655884
step: 130, loss: 0.3019335865974426
step: 140, loss: 0.11686531454324722
step: 150, loss: 0.10588313639163971
step: 160, loss: 0.020076945424079895
step: 170, loss: 0.1796530783176422
step: 180, loss: 0.2763211131095886
step: 190, loss: 0.08688779920339584
step: 200, loss: 0.21716831624507904
step: 210, loss: 0.08632662147283554
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.6792452830188679, f1=0.6368932038834951,
best_f1=0.6368932038834951
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.09568804502487183
step: 10, loss: 0.057533688843250275
step: 20, loss: 0.04131271690130234
step: 30, loss: 0.045047808438539505
step: 40, loss: 0.021451953798532486
step: 50, loss: 0.09120377153158188
step: 60, loss: 0.04456508904695511
step: 70, loss: 0.003989659249782562
step: 80, loss: 0.04168357327580452
step: 90, loss: 0.042065560817718506
step: 100, loss: 0.002157706068828702
```

```
step: 110, loss: 0.11103322356939316
step: 120, loss: 0.03534538298845291
step: 130, loss: 0.06885265558958054
step: 140, loss: 0.03575030714273453
step: 150, loss: 0.03978215157985687
step: 160, loss: 0.13861480355262756
step: 170, loss: 0.022815169766545296
step: 180, loss: 0.012851452454924583
step: 190, loss: 0.03566928580403328
step: 200, loss: 0.17914804816246033
step: 210, loss: 0.018270913511514664
/usr/local/lib/python3.7/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=0.6705882352941176, f1=0.6201232032854209,
best_f1=0.6368932038834951
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.027922097593545914
step: 10, loss: 0.08535785973072052
step: 20, loss: 0.022901857271790504
step: 30, loss: 0.026438318192958832
step: 40, loss: 0.0015524636255577207
step: 50, loss: 0.0030505801551043987
step: 60, loss: 0.016909539699554443
step: 70, loss: 0.013842637650668621
step: 80, loss: 0.11936309188604355
step: 90, loss: 0.06591818481683731
step: 100, loss: 0.004460717551410198
step: 110, loss: 0.005771985277533531
step: 120, loss: 0.028642652556300163
step: 130, loss: 0.03366711735725403
step: 140, loss: 0.02928776480257511
step: 150, loss: 0.005719418171793222
step: 160, loss: 0.007153604179620743
step: 170, loss: 0.15575191378593445
step: 180, loss: 0.1191861480474472
step: 190, loss: 0.10973567515611649
step: 200, loss: 0.009479215368628502
step: 210, loss: 0.06116138771176338
/usr/local/lib/python3.7/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=0.6569037656903767, f1=0.6180257510729614,
best_f1=0.6368932038834951
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
```

```
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.008278429508209229
step: 10, loss: 0.0035256787668913603
step: 20, loss: 0.008274038322269917
step: 30, loss: 0.017354819923639297
step: 40, loss: 0.015768440440297127
step: 50, loss: 0.08428175002336502
step: 60, loss: 0.006772833876311779
step: 70, loss: 0.004278141539543867
step: 80, loss: 0.024776089936494827
step: 90, loss: 0.032232966274023056
step: 100, loss: 0.005152019206434488
step: 110, loss: 0.08250272274017334
step: 120, loss: 0.03501302748918533
step: 130, loss: 0.015573974698781967
step: 140, loss: 0.011070461943745613
step: 150, loss: 0.0022171377204358578
step: 160, loss: 0.006250392179936171
step: 170, loss: 0.0009108501835726202
step: 180, loss: 0.025411006063222885
step: 190, loss: 0.058279991149902344
step: 200, loss: 0.01262367982417345
step: 210, loss: 0.0026801126077771187
/usr/local/lib/python3.7/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=0.6589147286821705, f1=0.6415841584158416,
best_f1=0.6368932038834951
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.002416138770058751
step: 10, loss: 0.06002920866012573
step: 20, loss: 0.005222855601459742
step: 30, loss: 0.05575478821992874
step: 40, loss: 0.007342812605202198
step: 50, loss: 0.0023240118753165007
step: 60, loss: 0.038366034626960754
step: 70, loss: 0.03333043307065964
step: 80, loss: 0.06297208368778229
step: 90, loss: 0.11462381482124329
step: 100, loss: 0.004734581336379051
step: 110, loss: 0.07998785376548767
step: 120, loss: 0.011536234058439732
step: 130, loss: 0.0005707366508431733
step: 140, loss: 0.048573292791843414
step: 150, loss: 0.003182176500558853
step: 160, loss: 0.05735837668180466
```

```
step: 170, loss: 0.0019654140342026695
step: 180, loss: 0.012330637313425541
step: 190, loss: 0.04164749011397362
step: 200, loss: 0.012156681157648563
step: 210, loss: 0.08181098103523254
/usr/local/lib/python3.7/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=0.6535087719298245, f1=0.6294642857142857,
best_f1=0.6368932038834951
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0007094184402376413
step: 10, loss: 0.0005852296599186957
step: 20, loss: 0.0010013481369242072
step: 30, loss: 0.008020948618650436
step: 40, loss: 0.03315671160817146
step: 50, loss: 0.009265684522688389
step: 60, loss: 0.0835811197757721
step: 70, loss: 0.008226323872804642
step: 80, loss: 0.0016104532405734062
step: 90, loss: 0.00016367409261874855
step: 100, loss: 0.0016620473470538855
step: 110, loss: 0.0036736472975462675
step: 120, loss: 0.0011509967735037208
step: 130, loss: 0.0034613085445016623
step: 140, loss: 0.01771625690162182
step: 150, loss: 0.07605358213186264
step: 160, loss: 0.0019494141452014446
step: 170, loss: 0.0022635518107563257
step: 180, loss: 0.015287145040929317
step: 190, loss: 0.0006973569979891181
step: 200, loss: 0.0756041407585144
step: 210, loss: 0.008063253946602345
/usr/local/lib/python3.7/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=0.6567717996289425, f1=0.6420664206642066,
best_f1=0.6368932038834951
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.06317079067230225
step: 10, loss: 0.00370375020429492
step: 20, loss: 0.0003953636914957315
step: 30, loss: 0.010971635580062866
step: 40, loss: 0.0032379161566495895
```

```
step: 50, loss: 0.014480970799922943
step: 60, loss: 0.017240501940250397
step: 70, loss: 0.010967740789055824
step: 80, loss: 0.023571643978357315
step: 90, loss: 0.01328783668577671
step: 100, loss: 0.004692047834396362
step: 110, loss: 0.0002649713715072721
step: 120, loss: 0.0013323907041922212
step: 130, loss: 0.03983087092638016
step: 140, loss: 0.047388263046741486
step: 150, loss: 0.054285161197185516
step: 160, loss: 0.0007919397903606296
step: 170, loss: 0.025615248829126358
step: 180, loss: 0.008417908102273941
step: 190, loss: 0.02920171059668064
step: 200, loss: 0.007857983000576496
step: 210, loss: 0.03343847393989563
Gradient overflow.  Skipping step, loss scaler 0 reducing loss scale to 16384.0
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
epoch 10: dev_f1=0.6758893280632412, f1=0.6719367588932806,
best_f1=0.6368932038834951
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.03578601032495499
step: 10, loss: 0.007416334003210068
step: 20, loss: 0.02028139866888523
step: 30, loss: 0.00027067679911851883
step: 40, loss: 0.013058636337518692
step: 50, loss: 0.026395544409751892
step: 60, loss: 0.018059635534882545
step: 70, loss: 0.0002707219682633877
step: 80, loss: 0.017121776938438416
step: 90, loss: 0.037712834775447845
step: 100, loss: 0.075269453227519999
step: 110, loss: 0.038564931601285934
step: 120, loss: 0.006012878380715847
step: 130, loss: 0.028216561302542686
step: 140, loss: 0.00026020358200185
step: 150, loss: 0.00016169201990123838
step: 160, loss: 0.015122655779123306
step: 170, loss: 0.02366567589342594
step: 180, loss: 0.0004538492939900607
step: 190, loss: 0.0009096033172681928
step: 200, loss: 0.00039001982077024877
step: 210, loss: 0.002178475959226489
```

```
/usr/local/lib/python3.7/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=0.6582809224318658, f1=0.6469344608879491,
best_f1=0.6368932038834951
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.017050277441740036
step: 10, loss: 0.0010644706198945642
step: 20, loss: 0.007605219259858131
step: 30, loss: 0.000805146002676338
step: 40, loss: 0.0009564099018462002
step: 50, loss: 0.0019510139245539904
step: 60, loss: 0.0037728322204202414
step: 70, loss: 0.0007182184490375221
step: 80, loss: 0.0019978235941380262
step: 90, loss: 0.007131255231797695
step: 100, loss: 0.009444587863981724
step: 110, loss: 0.0015145433135330677
step: 120, loss: 0.021330295130610466
step: 130, loss: 0.004181503318250179
step: 140, loss: 0.000517792534083128
step: 150, loss: 0.024576816707849503
step: 160, loss: 0.00010677377576939762
step: 170, loss: 0.022890770807862282
step: 180, loss: 0.0009321856196038425
step: 190, loss: 0.0017019638326019049
step: 200, loss: 0.0029246939811855555
step: 210, loss: 0.02020195685327053
/usr/local/lib/python3.7/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=0.6556701030927835, f1=0.6502057613168725,
best_f1=0.6368932038834951
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.05056071653962135
step: 10, loss: 0.00026387142133899033
step: 20, loss: 0.0006034334073774517
step: 30, loss: 0.022170711308717728
step: 40, loss: 0.0031169175636023283
step: 50, loss: 0.008511797524988651
step: 60, loss: 0.00024624221259728074
step: 70, loss: 0.024168912321329117
step: 80, loss: 0.004006620030850172
step: 90, loss: 0.005466022063046694
```

```
step: 100, loss: 0.002064575208351016
step: 110, loss: 0.00027759766089729965
step: 120, loss: 0.00020538957323879004
step: 130, loss: 0.00046910601668059826
step: 140, loss: 0.009500900283455849
step: 150, loss: 0.001036143279634416
step: 160, loss: 0.041342347860336304
step: 170, loss: 0.0005326003301888704
step: 180, loss: 0.03860565274953842
step: 190, loss: 0.0009754607453942299
step: 200, loss: 0.0007276707328855991
step: 210, loss: 0.0008337446488440037
/usr/local/lib/python3.7/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=0.6612576064908721, f1=0.6363636363636364,
best_f1=0.6368932038834951
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 9.697859059087932e-05
step: 10, loss: 0.02944491058588028
step: 20, loss: 0.00019368858193047345
step: 30, loss: 0.0001204886066261679
step: 40, loss: 0.00020602322183549404
step: 50, loss: 0.009936240501701832
step: 60, loss: 0.00023806760145816952
step: 70, loss: 0.00015022599836811423
step: 80, loss: 0.00015682470984756947
step: 90, loss: 0.001073510735295713
step: 100, loss: 0.001714340760372579
step: 110, loss: 0.0003871456428896636
step: 120, loss: 0.02496790513396263
step: 130, loss: 0.001227657776325941
step: 140, loss: 0.020823931321501732
step: 150, loss: 0.0001768813090166077
step: 160, loss: 0.002636507386341691
step: 170, loss: 0.0002954751253128052
step: 180, loss: 0.003143428824841976
step: 190, loss: 0.0010589767480269074
step: 200, loss: 0.0002824451948981732
step: 210, loss: 0.0014817360788583755
/usr/local/lib/python3.7/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=0.6557377049180328, f1=0.6502057613168725,
best_f1=0.6368932038834951
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
```

```
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00024162602494470775
step: 10, loss: 0.0002615316188894212
step: 20, loss: 0.0032558036036789417
step: 30, loss: 0.011076287366449833
step: 40, loss: 0.0016866764053702354
step: 50, loss: 0.00015882377920206636
step: 60, loss: 0.016376594081521034
step: 70, loss: 0.0002752217696979642
step: 80, loss: 0.00038266999763436615
step: 90, loss: 0.006625500041991472
step: 100, loss: 0.00014942113193683326
step: 110, loss: 0.0001317718852078542
step: 120, loss: 0.00023473154578823596
step: 130, loss: 0.0001515410258434713
step: 140, loss: 0.0010980631923303008
step: 150, loss: 0.00016650374163873494
step: 160, loss: 0.002476288704201579
step: 170, loss: 0.00015705816622357816
step: 180, loss: 0.00010756661504274234
step: 190, loss: 0.00020750498515553772
step: 200, loss: 0.004770062398165464
step: 210, loss: 0.0003288961888756603
/usr/local/lib/python3.7/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=0.652542372881356, f1=0.6279569892473118,
best_f1=0.6368932038834951
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 [00:05, 402.19it/s]
load_f1 = 0.6732283464566929
real_f1 = 0.6640625
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, 217.22it/s]
```

3.1.5 Walmart-Amazon - 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/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: 0.5437835454940796
step: 10, loss: 0.3844553232192993
step: 20, loss: 0.26937153935432434
step: 30, loss: 0.43762698769569397

```
step: 40, loss: 0.42785051465034485
step: 50, loss: 0.2811547815799713
step: 60, loss: 0.25477951765060425
step: 70, loss: 0.2907724976539612
step: 80, loss: 0.22229814529418945
step: 90, loss: 0.2823222577571869
step: 100, loss: 0.31684648990631104
step: 110, loss: 0.33642131090164185
step: 120, loss: 0.07836171239614487
step: 130, loss: 0.12292157113552094
step: 140, loss: 0.03438432142138481
step: 150, loss: 0.21369308233261108
step: 160, loss: 0.053324781358242035
step: 170, loss: 0.15360894799232483
step: 180, loss: 0.09067252278327942
step: 190, loss: 0.17922955751419067
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.6649350649350649, f1=0.6370757180156659,
best_f1=0.6370757180156659
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.3942844867706299
step: 10, loss: 0.09220122545957565
step: 20, loss: 0.11537709087133408
step: 30, loss: 0.23168909549713135
step: 40, loss: 0.1339629590511322
step: 50, loss: 0.0691850557923317
step: 60, loss: 0.20009203255176544
step: 70, loss: 0.1356707513332367
step: 80, loss: 0.33403605222702026
step: 90, loss: 0.25405627489089966
step: 100, loss: 0.048616208136081696
step: 110, loss: 0.11080663651227951
step: 120, loss: 0.25081250071525574
step: 130, loss: 0.18963591754436493
step: 140, loss: 0.12101548165082932
step: 150, loss: 0.07012316584587097
step: 160, loss: 0.023943809792399406
step: 170, loss: 0.11104514449834824
step: 180, loss: 0.13444848358631134
step: 190, loss: 0.033372994512319565
```

```
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.7582417582417582, f1=0.7965616045845273,
best_f1=0.7965616045845273
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.03958401829004288
step: 10, loss: 0.12530618906021118
step: 20, loss: 0.08066102117300034
step: 30, loss: 0.19387242197990417
step: 40, loss: 0.041468724608421326
step: 50, loss: 0.2180178165435791
step: 60, loss: 0.10596393048763275
step: 70, loss: 0.056845828890800476
step: 80, loss: 0.08724070340394974
step: 90, loss: 0.044248588383197784
step: 100, loss: 0.023867066949605942
step: 110, loss: 0.009223487228155136
step: 120, loss: 0.05339958518743515
step: 130, loss: 0.01687692478299141
step: 140, loss: 0.06442878395318985
step: 150, loss: 0.10609906166791916
step: 160, loss: 0.04372725263237953
step: 170, loss: 0.096563421189785
step: 180, loss: 0.05091210827231407
step: 190, loss: 0.01951570250093937
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.7968337730870713, f1=0.8148148148148148,
best_f1=0.8148148148148148
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0158989317715168
step: 10, loss: 0.10047780722379684
step: 20, loss: 0.03140459209680557
step: 30, loss: 0.07541866600513458
```

```
Gradient overflow.  Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 40, loss: 0.007883951999247074
step: 50, loss: 0.004869521129876375
step: 60, loss: 0.20408876240253448
step: 70, loss: 0.0075690122321248055
step: 80, loss: 0.05566508322954178
step: 90, loss: 0.07685093581676483
step: 100, loss: 0.04675725847482681
step: 110, loss: 0.01881209947168827
step: 120, loss: 0.027932122349739075
step: 130, loss: 0.16692408919334412
step: 140, loss: 0.008922060020267963
step: 150, loss: 0.007093910127878189
step: 160, loss: 0.017482547089457512
step: 170, loss: 0.05079749599099159
step: 180, loss: 0.001948383403941989
step: 190, loss: 0.060454487800598145
/usr/local/lib/python3.7/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=0.7750677506775067, f1=0.7875354107648725,
best_f1=0.8148148148148148
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.053988706320524216
step: 10, loss: 0.04014143720269203
step: 20, loss: 0.06892568618059158
step: 30, loss: 0.005461543332785368
step: 40, loss: 0.042129941284656525
step: 50, loss: 0.13478976488113403
step: 60, loss: 0.10263290256261826
step: 70, loss: 0.0878324881196022
step: 80, loss: 0.01659640111029148
step: 90, loss: 0.008592662401497364
step: 100, loss: 0.0038243953604251146
step: 110, loss: 0.014831678941845894
step: 120, loss: 0.006494604051113129
step: 130, loss: 0.027981525287032127
step: 140, loss: 0.14018353819847107
step: 150, loss: 0.004110617097467184
step: 160, loss: 0.022355817258358
step: 170, loss: 0.10853253304958344
step: 180, loss: 0.07221774756908417
step: 190, loss: 0.0060055917128920555
/usr/local/lib/python3.7/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=0.7845303867403315, f1=0.7855153203342619,
best_f1=0.8148148148148148
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.053654044866556189
step: 10, loss: 0.03430720791220665
step: 20, loss: 0.005762737710028887
step: 30, loss: 0.0005801179795525968
step: 40, loss: 0.0036779444199055433
step: 50, loss: 0.22495193779468536
step: 60, loss: 0.017018623650074005
step: 70, loss: 0.0011269962415099144
step: 80, loss: 0.02007051557302475
step: 90, loss: 0.03878423199057579
step: 100, loss: 0.00044884695671498775
step: 110, loss: 0.08816685527563095
step: 120, loss: 0.00119341304525733
step: 130, loss: 0.02864951826632023
step: 140, loss: 0.04133026301860809
step: 150, loss: 0.0023901008535176516
step: 160, loss: 0.15437102317810059
step: 170, loss: 0.028603605926036835
step: 180, loss: 0.00295532145537436
step: 190, loss: 0.019596757367253304
/usr/local/lib/python3.7/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=0.7672634271099745, f1=0.7526315789473683,
best_f1=0.8148148148148148
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.006387455388903618
step: 10, loss: 0.006756922695785761
step: 20, loss: 0.04971915856003761
step: 30, loss: 0.05230497941374779
step: 40, loss: 0.011950608342885971
step: 50, loss: 0.016089584678411484
step: 60, loss: 0.07197339832782745
step: 70, loss: 0.08623605221509933
step: 80, loss: 0.004990084562450647
step: 90, loss: 0.0028397943824529648
step: 100, loss: 0.0011888784356415272
step: 110, loss: 0.0029836532194167376
step: 120, loss: 0.0005402666283771396
step: 130, loss: 0.0009199423948302865
step: 140, loss: 0.0019163029501214623
```

```
step: 150, loss: 0.000985907856374979
step: 160, loss: 0.03589736670255661
step: 170, loss: 0.000990206142887473
step: 180, loss: 0.001256966614164412
step: 190, loss: 0.001710801851004362
/usr/local/lib/python3.7/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=0.7632311977715878, f1=0.7705382436260623,
best_f1=0.8148148148148148
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0021733902394771576
step: 10, loss: 0.0027819545939564705
step: 20, loss: 0.004179326351732016
step: 30, loss: 0.0017926867585629225
step: 40, loss: 0.0006979473982937634
step: 50, loss: 0.006577224470674992
step: 60, loss: 0.0006441896548494697
step: 70, loss: 0.004280067980289459
step: 80, loss: 0.0007045323145575821
step: 90, loss: 0.0006648930138908327
step: 100, loss: 0.0061516291461884975
step: 110, loss: 0.0007281571161001921
step: 120, loss: 0.007773841265588999
step: 130, loss: 0.001286122016608715
step: 140, loss: 0.002400102326646447
step: 150, loss: 0.019477244466543198
step: 160, loss: 0.0027789536397904158
step: 170, loss: 0.01121154148131609
step: 180, loss: 0.006369614973664284
step: 190, loss: 0.005680527072399855
/usr/local/lib/python3.7/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=0.7667560321715818, f1=0.7479224376731302,
best_f1=0.8148148148148148
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.000353233510395512
step: 10, loss: 0.03298313170671463
step: 20, loss: 0.0011445499258115888
step: 30, loss: 0.0006213284214027226
step: 40, loss: 0.05120371654629707
step: 50, loss: 0.004620281048119068
step: 60, loss: 0.001408525393344462
```

```
step: 70, loss: 0.0007710065692663193
step: 80, loss: 0.0001935653854161501
step: 90, loss: 0.0006373703945428133
step: 100, loss: 0.0003621973446570337
step: 110, loss: 0.000622714520432055
step: 120, loss: 0.0006766529404558241
step: 130, loss: 0.0028640765231102705
step: 140, loss: 0.12536874413490295
step: 150, loss: 0.005497712176293135
step: 160, loss: 0.0014754727017134428
step: 170, loss: 0.0027364911511540413
step: 180, loss: 0.018428193405270576
step: 190, loss: 0.0021221430506557226
/usr/local/lib/python3.7/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=0.7519582245430808, f1=0.7945945945945946,
best_f1=0.8148148148148148
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.001015259651467204
step: 10, loss: 0.0007620420074090362
step: 20, loss: 0.005863090045750141
step: 30, loss: 0.0004702129226643592
step: 40, loss: 0.04392186924815178
step: 50, loss: 0.0003172970318701118
step: 60, loss: 0.00028329427004791796
step: 70, loss: 0.005371208768337965
step: 80, loss: 0.0003314089262858033
step: 90, loss: 0.0005288630491122603
step: 100, loss: 0.0020049854647368193
step: 110, loss: 0.0006662340019829571
step: 120, loss: 0.1752806007862091
step: 130, loss: 0.0020399473141878843
step: 140, loss: 0.000620740232989192
step: 150, loss: 0.001101350993849337
step: 160, loss: 0.001706120208837092
step: 170, loss: 0.00025171705055981874
step: 180, loss: 0.00048215544666163623
step: 190, loss: 0.001666342024691403
/usr/local/lib/python3.7/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=0.7806267806267806, f1=0.7803468208092486,
best_f1=0.8148148148148148
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
```

```
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00044316271669231355
step: 10, loss: 0.028754862025380135
step: 20, loss: 0.019903061911463737
step: 30, loss: 0.003176142228767276
step: 40, loss: 0.00049290060099700928
step: 50, loss: 0.0005193932447582483
step: 60, loss: 0.00036734348395839334
step: 70, loss: 0.0006251255981624126
step: 80, loss: 0.00107967434450984
step: 90, loss: 0.000252315221587196
step: 100, loss: 0.0003000661381520331
step: 110, loss: 0.0006659863283857703
step: 120, loss: 0.000496305525302887
step: 130, loss: 0.0004912129370495677
step: 140, loss: 0.001094033126719296
step: 150, loss: 0.0002177658025175333
step: 160, loss: 0.0020066373981535435
step: 170, loss: 0.00028534067678265274
step: 180, loss: 0.0002634388511069119
step: 190, loss: 0.0009897968266159296
/usr/local/lib/python3.7/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=0.7859078590785908, f1=0.7643678160919539,
best_f1=0.8148148148148148
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00017910503083840013
step: 10, loss: 0.009063749574124813
step: 20, loss: 0.0005756790633313358
step: 30, loss: 0.0016376140993088484
step: 40, loss: 0.00029026553966104984
step: 50, loss: 0.0006374728982336819
step: 60, loss: 0.000566851522307843
step: 70, loss: 0.000402134086471051
step: 80, loss: 0.00029407665715552866
step: 90, loss: 0.00023148235050030053
step: 100, loss: 0.0007419756148010492
step: 110, loss: 0.0005004861741326749
step: 120, loss: 0.0001506140106357634
step: 130, loss: 0.0003998529282398522
step: 140, loss: 0.0007876959862187505
step: 150, loss: 0.00023448924184776843
step: 160, loss: 0.00029677251586690545
step: 170, loss: 0.0001924650714499876
step: 180, loss: 0.00010121783998329192
```

step: 190, loss: 0.00031601733644492924
/usr/local/lib/python3.7/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=0.7759562841530054, f1=0.7586206896551724,
best_f1=0.8148148148148148
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0003000050492119044
step: 10, loss: 0.00014102608838584274
step: 20, loss: 0.00013910788402426988
step: 30, loss: 0.00025860031018964946
step: 40, loss: 0.00026746143703348935
step: 50, loss: 0.0005478011444211006
step: 60, loss: 0.00014699371240567416
step: 70, loss: 0.000961954821832478
step: 80, loss: 0.016027722507715225
step: 90, loss: 0.0006023842142894864
step: 100, loss: 0.0012814399087801576
step: 110, loss: 0.00046131465933285654
step: 120, loss: 0.0005309391999617219
step: 130, loss: 0.00037421571323648095
step: 140, loss: 0.00015811460616532713
step: 150, loss: 0.00023640648578293622
step: 160, loss: 0.0002809022553265095
step: 170, loss: 0.00044246946345083416
step: 180, loss: 0.0001436142629245296
step: 190, loss: 0.13257940113544464
/usr/local/lib/python3.7/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=0.770949720670391, f1=0.7687861271676301,
best_f1=0.8148148148148148
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0006447473424486816
step: 10, loss: 0.00016885617515072227
step: 20, loss: 0.00029007482226006687
step: 30, loss: 0.0001027919861371629
step: 40, loss: 0.0002821201633196324
step: 50, loss: 0.0001832927664509043
step: 60, loss: 0.0003382973081897944
step: 70, loss: 0.0015541142784059048
step: 80, loss: 0.0002497140958439559
step: 90, loss: 0.001108113443478942
step: 100, loss: 0.0007502463995479047

```
step: 110, loss: 0.002714283764362335
step: 120, loss: 0.00012559119204524904
Gradient overflow.  Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 130, loss: 0.004968064371496439
step: 140, loss: 0.00018576328875496984
step: 150, loss: 0.00015399299445562065
step: 160, loss: 0.0012560357572510839
step: 170, loss: 0.00019712763605639338
step: 180, loss: 0.017986003309488297
step: 190, loss: 0.00034907663939520717
/usr/local/lib/python3.7/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=0.776536312849162, f1=0.7630057803468209,
best_f1=0.8148148148148148
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.007644536439329386
step: 10, loss: 0.00011254453420406207
step: 20, loss: 0.00017274392303079367
step: 30, loss: 0.00016658968525007367
step: 40, loss: 0.00039680401096120477
step: 50, loss: 0.00012178827455500141
step: 60, loss: 0.00012627505930140615
step: 70, loss: 0.002890824107453227
step: 80, loss: 0.0019557769410312176
step: 90, loss: 0.00016221287660300732
step: 100, loss: 0.0013081008801236749
step: 110, loss: 0.0003273543552495539
step: 120, loss: 0.0007456004968844354
step: 130, loss: 0.00010131675662705675
step: 140, loss: 0.00024198494793381542
step: 150, loss: 0.00033015289227478206
step: 160, loss: 0.00014005844423081726
step: 170, loss: 0.00016398595471400768
step: 180, loss: 0.003762507811188698
step: 190, loss: 0.00012795772636309266
/usr/local/lib/python3.7/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=0.7688022284122563, f1=0.7643678160919539,
best_f1=0.8148148148148148
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 [00:10, 199.86it/s]
load_f1 = 0.6426735218508998
real_f1 = 0.6361323155216284
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, 221.90it/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: 0.6290316581726074
step: 10, loss: 0.35872894525527954
step: 20, loss: 0.31912454962273041
step: 30, loss: 0.37396809458732605
step: 40, loss: 0.2700957655906677
step: 50, loss: 0.2637379765510559
step: 60, loss: 0.2441197782754898
step: 70, loss: 0.38381126523017883
step: 80, loss: 0.3694934546947479
step: 90, loss: 0.24785521626472473
step: 100, loss: 0.16236698627471924
step: 110, loss: 0.24819211661815643
step: 120, loss: 0.21017292141914368
step: 130, loss: 0.04938714578747749
step: 140, loss: 0.2329823076725006
step: 150, loss: 0.2328414022922516
step: 160, loss: 0.12636646628379822
step: 170, loss: 0.2855156362056732
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.7444168734491315, f1=0.720763723150358,
best_f1=0.720763723150358
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.16078715026378632
step: 10, loss: 0.11963994801044464
step: 20, loss: 0.039548128843307495
step: 30, loss: 0.2817404866218567
step: 40, loss: 0.07223214954137802
step: 50, loss: 0.20839014649391174

```
step: 60, loss: 0.08881810307502747
step: 70, loss: 0.14993643760681152
step: 80, loss: 0.07745501399040222
step: 90, loss: 0.16619932651519775
step: 100, loss: 0.2040548473596573
step: 110, loss: 0.08990336954593658
step: 120, loss: 0.26486533880233765
step: 130, loss: 0.08762931078672409
step: 140, loss: 0.22698119282722473
step: 150, loss: 0.1337176263332367
step: 160, loss: 0.12335643917322159
Gradient overflow.  Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 170, loss: 0.02154509723186493
/usr/local/lib/python3.7/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=0.7423167848699763, f1=0.7713625866050808,
best_f1=0.720763723150358
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.17014412581920624
step: 10, loss: 0.11567521095275879
step: 20, loss: 0.10815800726413727
step: 30, loss: 0.15103568136692047
step: 40, loss: 0.10569784045219421
step: 50, loss: 0.06695225089788437
step: 60, loss: 0.11099988222122192
step: 70, loss: 0.045565392822027206
step: 80, loss: 0.05589798837900162
step: 90, loss: 0.06375506520271301
step: 100, loss: 0.01848059520125389
step: 110, loss: 0.06675703823566437
step: 120, loss: 0.05047593638300896
step: 130, loss: 0.07389073818922043
step: 140, loss: 0.01762515679001808
step: 150, loss: 0.19886118173599243
step: 160, loss: 0.09674390405416489
step: 170, loss: 0.23010855913162231
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.7981220657276995, f1=0.7671232876712328,
best_f1=0.7671232876712328
```

```
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.015740592032670975
step: 10, loss: 0.04109233245253563
step: 20, loss: 0.005345993675291538
step: 30, loss: 0.010445669293403625
step: 40, loss: 0.004075797740370035
step: 50, loss: 0.07006660103797913
step: 60, loss: 0.19749493896961212
step: 70, loss: 0.0077652777545154095
step: 80, loss: 0.07829029113054276
step: 90, loss: 0.03196728974580765
step: 100, loss: 0.16735482215881348
step: 110, loss: 0.01193452812731266
step: 120, loss: 0.00307336263358593
step: 130, loss: 0.029677487909793854
step: 140, loss: 0.027763033285737038
step: 150, loss: 0.028026917949318886
step: 160, loss: 0.0733768418431282
step: 170, loss: 0.18733245134353638
/usr/local/lib/python3.7/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=0.788659793814433, f1=0.8029556650246306,
best_f1=0.7671232876712328
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.04392192140221596
step: 10, loss: 0.011938345618546009
step: 20, loss: 0.03169980272650719
step: 30, loss: 0.030956203117966652
step: 40, loss: 0.015218770131468773
step: 50, loss: 0.003689633449539542
step: 60, loss: 0.013410753570497036
step: 70, loss: 0.060668133199214935
step: 80, loss: 0.1206246018409729
step: 90, loss: 0.10077172517776489
step: 100, loss: 0.08442405611276627
step: 110, loss: 0.014079036191105843
step: 120, loss: 0.013885173946619034
step: 130, loss: 0.03320605307817459
step: 140, loss: 0.01048872247338295
step: 150, loss: 0.04576435685157776
step: 160, loss: 0.007203089538961649
step: 170, loss: 0.005116036161780357
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
```

```
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.8031088082901555, f1=0.8170426065162907,
best_f1=0.8170426065162907
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.004885065369307995
step: 10, loss: 0.011713553220033646
step: 20, loss: 0.0051687960512936115
step: 30, loss: 0.00220652692951262
step: 40, loss: 0.014271738938987255
step: 50, loss: 0.15398149192333221
step: 60, loss: 0.011783569119870663
step: 70, loss: 0.11197741329669952
step: 80, loss: 0.04487500712275505
step: 90, loss: 0.016486290842294693
step: 100, loss: 0.043627649545669556
step: 110, loss: 0.009858937934041023
step: 120, loss: 0.010284967720508575
step: 130, loss: 0.020469853654503822
step: 140, loss: 0.016463221982121468
step: 150, loss: 0.008153987117111683
step: 160, loss: 0.11232244223356247
step: 170, loss: 0.002813460072502494
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.8151898734177214, f1=0.8166259168704155,
best_f1=0.8166259168704155
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00105801946483552460
step: 10, loss: 0.010819240473210812
step: 20, loss: 0.0017087309388443828
step: 30, loss: 0.0009312652400694788
step: 40, loss: 0.000754331995267421
step: 50, loss: 0.09632162749767303
step: 60, loss: 0.0043244678527116776
```

```
step: 70, loss: 0.0013389167143031955
step: 80, loss: 0.03206730633974075
step: 90, loss: 0.0007998075452633202
step: 100, loss: 0.0011031051399186254
step: 110, loss: 0.0031388767529278994
step: 120, loss: 0.006148249376565218
step: 130, loss: 0.1634775698184967
step: 140, loss: 0.004378863610327244
step: 150, loss: 0.003422877285629511
step: 160, loss: 0.0005067166057415307
step: 170, loss: 0.020427575334906578
/usr/local/lib/python3.7/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=0.7901907356948228, f1=0.828125, best_f1=0.8166259168704155
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.006983309518545866
step: 10, loss: 0.000941068516112864
step: 20, loss: 0.0008060664404183626
step: 30, loss: 0.013192631304264069
step: 40, loss: 0.00019544814131222665
step: 50, loss: 0.0018894376698881388
step: 60, loss: 0.0014274587156251073
step: 70, loss: 0.0007825963548384607
step: 80, loss: 0.0007534909527748823
step: 90, loss: 0.0015578388702124357
step: 100, loss: 0.01920844428241253
step: 110, loss: 0.15396399796009064
step: 120, loss: 0.00351868849247694
step: 130, loss: 0.03536561504006386
step: 140, loss: 0.021660195663571358
step: 150, loss: 0.13983570039272308
step: 160, loss: 0.03841312602162361
step: 170, loss: 0.021094782277941704
/usr/local/lib/python3.7/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=0.798913043478261, f1=0.8030690537084398,
best_f1=0.8166259168704155
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.003485708264634013
step: 10, loss: 0.16286718845367432
step: 20, loss: 0.004540441557765007
step: 30, loss: 0.015208626165986061
```

```
step: 40, loss: 0.0008102252031676471
step: 50, loss: 0.00040106114465743303
step: 60, loss: 0.048448000103235245
step: 70, loss: 0.0007027421379461884
step: 80, loss: 0.00106913049239666527
step: 90, loss: 0.04240624979138374
step: 100, loss: 0.06957107037305832
step: 110, loss: 0.02401158958673477
step: 120, loss: 0.04810840263962746
step: 130, loss: 0.06088840961456299
step: 140, loss: 0.0007876462768763304
step: 150, loss: 0.0006628797273151577
step: 160, loss: 0.013790827244520187
step: 170, loss: 0.031497035175561905
/usr/local/lib/python3.7/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=0.8138297872340425, f1=0.8232323232323232,
best_f1=0.8166259168704155
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.029871415346860886
step: 10, loss: 0.00030450260965153575
step: 20, loss: 0.27071648836135864
step: 30, loss: 0.001881062868051231
step: 40, loss: 0.024654125794768333
step: 50, loss: 0.017957258969545364
step: 60, loss: 0.00028285308508202434
step: 70, loss: 0.011621796526014805
step: 80, loss: 0.11074413359165192
step: 90, loss: 0.023596780374646187
step: 100, loss: 0.0004310820368118584
step: 110, loss: 0.0023438818752765656
step: 120, loss: 0.0013875637669116259
step: 130, loss: 0.00042738436604849994
step: 140, loss: 0.0014449593145400286
step: 150, loss: 0.0014493942726403475
step: 160, loss: 0.0010022212518379092
step: 170, loss: 0.00016975120052487582
/usr/local/lib/python3.7/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=0.7945945945945946, f1=0.8010335917312662,
best_f1=0.8166259168704155
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
```

```
step: 0, loss: 0.014844266697764397
step: 10, loss: 0.04658293351531029
step: 20, loss: 0.00046615811879746616
step: 30, loss: 0.0003225473337806761
step: 40, loss: 0.001066646189428866
step: 50, loss: 0.007033152040094137
step: 60, loss: 0.011048215441405773
step: 70, loss: 0.003610219107940793
step: 80, loss: 0.0003096492728218436
step: 90, loss: 0.0002967960317619145
step: 100, loss: 0.0003588027029763907
step: 110, loss: 0.004611192736774683
step: 120, loss: 0.00034795087412931025
step: 130, loss: 0.00011673946573864669
step: 140, loss: 0.003962934948503971
step: 150, loss: 0.0024074246175587177
step: 160, loss: 0.0029626768082380295
step: 170, loss: 0.0002996786788571626
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
epoch 11: dev_f1=0.8093994778067886, f1=0.8040201005025125,
best_f1=0.8166259168704155
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.000250005949055776
step: 10, loss: 0.002122958656400442
step: 20, loss: 0.00027642102213576436
step: 30, loss: 0.00042948860209435225
step: 40, loss: 0.0018349691526964307
step: 50, loss: 0.0001395074650645256
step: 60, loss: 0.016254307702183723
step: 70, loss: 0.06766146421432495
step: 80, loss: 0.00023684214102104306
step: 90, loss: 0.0004117623611818999
step: 100, loss: 0.0036856785882264376
step: 110, loss: 0.0011137353722006083
step: 120, loss: 0.001820934470742941
step: 130, loss: 0.002327906433492899
step: 140, loss: 0.011413048021495342
step: 150, loss: 0.0009668369893915951
step: 160, loss: 0.00021404106519185007
step: 170, loss: 0.00015231767611112446
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
```

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=0.8170426065162907, f1=0.8038740920096852,
best_f1=0.8038740920096852
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.016406690701842308
step: 10, loss: 0.0003998297033831477
step: 20, loss: 0.0018238609191030264
step: 30, loss: 0.00013731460785493255
step: 40, loss: 0.00723812822252512
step: 50, loss: 0.05820313096046448
step: 60, loss: 0.003165645757690072
step: 70, loss: 0.001497324788942933
step: 80, loss: 0.0017128533218055964
step: 90, loss: 9.448845230508596e-05
step: 100, loss: 0.0003898335271514952
step: 110, loss: 0.00018755342171061784
step: 120, loss: 0.019377637654542923
step: 130, loss: 0.00023273465922102332
step: 140, loss: 0.00031505830702371895
step: 150, loss: 0.03967755660414696
step: 160, loss: 0.0004178386298008263
step: 170, loss: 0.0002608496288303286
/usr/local/lib/python3.7/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=0.8091603053435115, f1=0.7970660146699265,
best_f1=0.8038740920096852
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0007485164096578956
step: 10, loss: 0.00017710549582261592
step: 20, loss: 0.0038143559359014034
step: 30, loss: 0.0005349125131033361
step: 40, loss: 0.00031345285242423415
step: 50, loss: 0.0005075978697277606
step: 60, loss: 0.03936197981238365
step: 70, loss: 0.0025587778072804213
step: 80, loss: 0.001902678282931447
step: 90, loss: 0.00023148591571953148
step: 100, loss: 0.0001110542580136098
step: 110, loss: 0.0002953777147922665
step: 120, loss: 0.010796796530485153
step: 130, loss: 0.0011267635272815824

```
step: 140, loss: 0.00035521091194823384
step: 150, loss: 0.0008016637875698507
step: 160, loss: 0.00020691774261649698
step: 170, loss: 0.0003336233494337648
```
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.8257372654155496, f1=0.8041237113402061,
best_f1=0.8041237113402061
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
```
step: 0, loss: 0.000151073036249727
step: 10, loss: 0.0004355897253844887
step: 20, loss: 0.0012231265427544713
step: 30, loss: 0.0003664122777990997
step: 40, loss: 0.002154693240299821
step: 50, loss: 0.0001431686250725761
step: 60, loss: 0.05840041860938072
step: 70, loss: 0.00019758324197027832
step: 80, loss: 0.005681624170392752
step: 90, loss: 0.001925625023432076
step: 100, loss: 0.0002999936987180263
step: 110, loss: 0.00014378091145772487
step: 120, loss: 0.00044696906115859747
step: 130, loss: 0.0005188704235479236
step: 140, loss: 0.00016306154429912567
```
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
```
step: 150, loss: 0.0003624334931373596
step: 160, loss: 0.00040353150689043105
step: 170, loss: 0.0008352779550477862
```
/usr/local/lib/python3.7/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=0.8181818181818182, f1=0.8020565552699228,
best_f1=0.8041237113402061
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, 253.53it/s]
load_f1 = 0.6141732283464566
real_f1 = 0.6058252427184466
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, 223.75it/s]
```

3.3 DK DIRTY

3.3.1 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 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: 0.616497278213501
step: 10, loss: 0.6252344250679016
step: 20, loss: 0.3142161965370178
step: 30, loss: 0.0699378252029419
step: 40, loss: 0.2210879772901535
step: 50, loss: 0.10188774019479752
step: 60, loss: 0.08319952338933945
step: 70, loss: 0.1264403909444809
step: 80, loss: 0.06855887919664383
step: 90, loss: 0.07461768388748169
step: 100, loss: 0.04406120628118515
step: 110, loss: 0.2405397891998291
step: 120, loss: 0.010909640230238438
step: 130, loss: 0.014701712876558304
step: 140, loss: 0.0070767709985375404
step: 150, loss: 0.024735987186431885
step: 160, loss: 0.0047877393662929535
step: 170, loss: 0.12957409024238586
step: 180, loss: 0.03602477163076401
step: 190, loss: 0.02586657740175724
step: 200, loss: 0.05013897642493248
step: 210, loss: 0.0024750027805566788
step: 220, loss: 0.015070619061589241
step: 230, loss: 0.02278526872396469
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9726651480637813, f1=0.9681818181818181,
best_f1=0.9681818181818181
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.004510430619120598
step: 10, loss: 0.015317348763346672
step: 20, loss: 0.16757462918758392
```

```
step: 30, loss: 0.19390513002872467
step: 40, loss: 0.05791383236646652
step: 50, loss: 0.0037836392875760794
step: 60, loss: 0.008605588227510452
step: 70, loss: 0.015738531947135925
step: 80, loss: 0.06567082554101944
step: 90, loss: 0.006308234762400389
step: 100, loss: 0.06326666474342346
step: 110, loss: 0.06825077533721924
step: 120, loss: 0.1332036852836609
step: 130, loss: 0.009749959222972393
step: 140, loss: 0.0020955675281584263
step: 150, loss: 0.07445012032985687
step: 160, loss: 0.017032669857144356
step: 170, loss: 0.0012536931317299604
step: 180, loss: 0.10903969407081604
step: 190, loss: 0.0027832190971821547
step: 200, loss: 0.0021617766469717026
step: 210, loss: 0.00099970251703634858
step: 220, loss: 0.12366972863674164
step: 230, loss: 0.0037805535830557346
/usr/local/lib/python3.7/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=0.9686800894854586, f1=0.9719416386083053,
best_f1=0.9681818181818181
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.04039733111858368
step: 10, loss: 0.0033835608046501875
step: 20, loss: 0.09352539479732513
step: 30, loss: 0.02819577045738697
step: 40, loss: 0.006846242118626833
step: 50, loss: 0.010010654106736183
step: 60, loss: 0.006530353799462318
step: 70, loss: 0.004337223246693611
step: 80, loss: 0.007075682282447815
step: 90, loss: 0.0011290351394563913
step: 100, loss: 0.0032271738164126873
step: 110, loss: 0.0004383292398415506
step: 120, loss: 0.021847128868103027
step: 130, loss: 0.0004312340170145035
step: 140, loss: 0.0021657939068973064
step: 150, loss: 0.03187781199812889
step: 160, loss: 0.021632585674524307
step: 170, loss: 0.0171661339700222
step: 180, loss: 0.02531660534441471
```

```
step: 190, loss: 0.027227791026234627
step: 200, loss: 0.08173702657222748
step: 210, loss: 0.07389384508132935
step: 220, loss: 0.0016526124672964215
step: 230, loss: 0.001321239280514419
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9772209567198178, f1=0.9738933030646991,
best_f1=0.9738933030646991
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0001319101866101846
step: 10, loss: 0.0005690369871445
step: 20, loss: 0.0007884560618549585
step: 30, loss: 0.003125527873635292
step: 40, loss: 0.015571025200188816
step: 50, loss: 0.000994014204479754
step: 60, loss: 0.0015302306273952127
step: 70, loss: 0.0009970634710043669
step: 80, loss: 0.00043958405149169266
step: 90, loss: 0.0007720833527855575
step: 100, loss: 0.0004906198009848595
step: 110, loss: 0.00044458656338974833
Gradient overflow.  Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 120, loss: 0.02645803801715374
step: 130, loss: 0.005057545378804207
step: 140, loss: 0.002118527889251709
step: 150, loss: 0.11834167689085007
step: 160, loss: 0.1883639395236969
step: 170, loss: 0.05109625309705734
step: 180, loss: 0.0027818719390779734
step: 190, loss: 0.0031975130550563335
step: 200, loss: 0.0018451792420819402
step: 210, loss: 0.0004304909089114517
step: 220, loss: 0.0005007369327358902
step: 230, loss: 0.0009171079727821052
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9774774774774775, f1=0.971815107102593,
best_f1=0.971815107102593
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0003191263822373003
step: 10, loss: 0.0003013799723703414
step: 20, loss: 0.0003492274845484644
step: 30, loss: 0.0002114232920575887
step: 40, loss: 0.0010586322750896215
step: 50, loss: 0.0003836308023892343
step: 60, loss: 0.19351918995380402
step: 70, loss: 0.0002933436771854758
step: 80, loss: 0.0010437598684802651
step: 90, loss: 0.010897588916122913
step: 100, loss: 0.00037131065619178116
step: 110, loss: 0.0007237619138322771
step: 120, loss: 0.00022524099040310083
step: 130, loss: 0.0017177363624796271
step: 140, loss: 0.00061352876946633007
step: 150, loss: 0.0009399050031788647
step: 160, loss: 0.0006400919519364834
step: 170, loss: 0.002306975657120347
step: 180, loss: 0.00029733101837337017
step: 190, loss: 0.16226112842559814
step: 200, loss: 0.007368250284343958
step: 210, loss: 0.004478382878005505
step: 220, loss: 0.0010525319958105683
step: 230, loss: 0.0011551734060049057
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9821428571428571, f1=0.9732142857142857,
best_f1=0.9732142857142857
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.008933818899095058
step: 10, loss: 0.0006864150636829436
step: 20, loss: 0.0008655193378217518
step: 30, loss: 0.0003245485422667116
step: 40, loss: 0.0007741273147985339
step: 50, loss: 0.012825103476643562
```

```
step: 60, loss: 0.00046431581722572446
step: 70, loss: 0.0003335803048685193
step: 80, loss: 0.0012743128463625908
step: 90, loss: 0.00021530882804654539
step: 100, loss: 0.005798024125397205
step: 110, loss: 0.0006787974270991981
step: 120, loss: 0.0023038536310195923
step: 130, loss: 0.00017720097093842924
step: 140, loss: 0.0001451152202207595
step: 150, loss: 0.04513045772910118
step: 160, loss: 0.003145152935758233
step: 170, loss: 0.0005269521498121321
step: 180, loss: 0.005129431374371052
step: 190, loss: 0.0038283199537545443
step: 200, loss: 0.00026480198721401393
step: 210, loss: 0.001399607746861875
step: 220, loss: 0.00018274415924679488
step: 230, loss: 0.06829587370157242
/usr/local/lib/python3.7/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=0.9808773903262092, f1=0.9740112994350283,
best_f1=0.9732142857142857
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00011722258932422847
step: 10, loss: 0.00016859637980815023
step: 20, loss: 0.00144930777605623
step: 30, loss: 0.003706286195665598
step: 40, loss: 0.00032501743407920003
step: 50, loss: 0.00016687267634551972
step: 60, loss: 0.00026241259183734655
step: 70, loss: 0.007006252650171518
step: 80, loss: 0.0025061971973627806
step: 90, loss: 0.000301397405564785
step: 100, loss: 0.0001957505155587569
step: 110, loss: 0.00021266243129502982
step: 120, loss: 7.685598393436521e-05
step: 130, loss: 0.00015674417954869568
step: 140, loss: 0.00019805351621471345
step: 150, loss: 0.010134466923773289
step: 160, loss: 0.04280505329370499
step: 170, loss: 0.15631964802742004
step: 180, loss: 0.0010137784993276
step: 190, loss: 9.662690717959777e-05
step: 200, loss: 0.021647633984684944
step: 210, loss: 6.852154183434322e-05
```

```
step: 220, loss: 0.00018366871518082917
step: 230, loss: 0.00021493839449249208
/usr/local/lib/python3.7/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=0.9763779527559054,
best_f1=0.9732142857142857
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00022465811343863606
step: 10, loss: 0.00044056487968191504
step: 20, loss: 0.0001479827769799158
step: 30, loss: 7.589644519612193e-05
step: 40, loss: 0.0008737425087019801
step: 50, loss: 0.0018814406357705593
step: 60, loss: 0.0002369158755755052
step: 70, loss: 0.0006907846545800567
step: 80, loss: 0.000441259442595765
step: 90, loss: 6.526923971250653e-05
step: 100, loss: 0.00015162251656875014
step: 110, loss: 0.00011107041063951328
step: 120, loss: 0.031593650579452515
step: 130, loss: 4.074484604643658e-05
step: 140, loss: 9.39011515583843e-05
step: 150, loss: 8.853743929648772e-05
step: 160, loss: 0.00014338098117150366
step: 170, loss: 0.0005783923552371562
step: 180, loss: 0.041662003844976425
step: 190, loss: 0.007700107526034117
step: 200, loss: 0.0095151886343956
step: 210, loss: 0.0003762002452276647
step: 220, loss: 0.004284838680177927
step: 230, loss: 0.0002794510801322758
/usr/local/lib/python3.7/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=0.9776785714285714, f1=0.9775784753363228,
best_f1=0.9732142857142857
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 6.159207259770483e-05
step: 10, loss: 0.0006697919452562928
step: 20, loss: 0.0005241652252152562
step: 30, loss: 0.00017725626821629703
step: 40, loss: 0.009582070633769035
step: 50, loss: 0.00011974424705840647
```

```
step: 60, loss: 0.0014365913812071085
step: 70, loss: 0.00047808015369810164
step: 80, loss: 0.0003294751513749361
step: 90, loss: 0.0006244330434128642
step: 100, loss: 0.0006230267463251948
step: 110, loss: 4.6436325646936893e-05
step: 120, loss: 7.47495869291015e-05
step: 130, loss: 3.509429734549485e-05
step: 140, loss: 3.6182162148179486e-05
step: 150, loss: 0.00017470523016527295
step: 160, loss: 4.364299093140289e-05
step: 170, loss: 7.075133180478588e-05
step: 180, loss: 0.00010159720113733783
step: 190, loss: 8.0605830589775e-05
step: 200, loss: 0.00021631340496242046
step: 210, loss: 0.00013114366447553039
step: 220, loss: 3.72736576537136e-05
step: 230, loss: 0.00010830086830537766
/usr/local/lib/python3.7/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=0.9776286353467561, f1=0.9741863075196409,
best_f1=0.9732142857142857
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 5.1901308324886486e-05
step: 10, loss: 0.00022487908427137882
step: 20, loss: 0.011383316479623318
step: 30, loss: 7.482731598429382e-05
step: 40, loss: 9.109019447350875e-05
step: 50, loss: 3.453568569966592e-05
step: 60, loss: 4.659902333514765e-05
step: 70, loss: 0.00010037265019491315
step: 80, loss: 5.4508411267306656e-05
step: 90, loss: 0.0001224948646267876
step: 100, loss: 5.3863910579821095e-05
step: 110, loss: 0.0004390190588310361
step: 120, loss: 0.00012321046961005777
step: 130, loss: 5.279778633848764e-05
step: 140, loss: 0.02699338272213936
step: 150, loss: 0.021360836923122406
step: 160, loss: 5.322169454302639e-05
step: 170, loss: 4.187297963653691e-05
step: 180, loss: 0.003444286063313484
step: 190, loss: 0.0005306907114572823
step: 200, loss: 0.00032889985595829785
step: 210, loss: 0.005425171926617622
```

```
step: 220, loss: 5.099690315546468e-05
step: 230, loss: 0.0007309096399694681
/usr/local/lib/python3.7/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=0.9765886287625419, f1=0.9707865168539327,
best_f1=0.9732142857142857
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 6.335286889225245e-05
step: 10, loss: 0.00013545998081099242
step: 20, loss: 0.00023543505812995136
step: 30, loss: 0.00022386211276382357
step: 40, loss: 0.0001377669395878911
step: 50, loss: 0.0016681089764460921
step: 60, loss: 0.00037803003215231
step: 70, loss: 0.0013717695837840438
step: 80, loss: 3.281136741861701e-05
step: 90, loss: 6.51927330181934e-05
step: 100, loss: 0.00011355282185832039
step: 110, loss: 0.0004919549101032317
step: 120, loss: 0.0007864421349950135
step: 130, loss: 3.5976416256744415e-05
step: 140, loss: 5.321337084751576e-05
step: 150, loss: 0.0005544815794564784
step: 160, loss: 5.3995081543689594e-05
step: 170, loss: 0.0032095026690512896
step: 180, loss: 7.722511509200558e-05
step: 190, loss: 3.3761007216526195e-05
step: 200, loss: 0.00027683580992743373
step: 210, loss: 2.7894271624973044e-05
step: 220, loss: 7.612154877278954e-05
step: 230, loss: 3.644328535301611e-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=0.9777777777777777, f1=0.9753914988814317,
best_f1=0.9732142857142857
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 6.233649037312716e-05
step: 10, loss: 4.0860631997929886e-05
step: 20, loss: 0.00023754237918183208
step: 30, loss: 0.00017793518782127649
step: 40, loss: 0.02472844533622265
step: 50, loss: 0.05937892943620682
```

```
step: 60, loss: 7.866548548918217e-05
step: 70, loss: 0.0006391355418600142
step: 80, loss: 0.009506519883871078
step: 90, loss: 8.842721581459045e-05
step: 100, loss: 5.5447475460823625e-05
step: 110, loss: 3.605925303418189e-05
step: 120, loss: 5.610809603240341e-05
step: 130, loss: 5.852214599144645e-05
step: 140, loss: 5.1170234655728564e-05
step: 150, loss: 0.00010445321095176041
step: 160, loss: 5.483210406964645e-05
step: 170, loss: 8.383260137634352e-05
step: 180, loss: 0.00012454710667952895
step: 190, loss: 4.940825601806864e-05
step: 200, loss: 0.0005112512153573334
step: 210, loss: 5.542254803003743e-05
step: 220, loss: 4.160836397204548e-05
step: 230, loss: 0.05190206319093704
/usr/local/lib/python3.7/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=0.9810055865921787, f1=0.9741282339707535,
best_f1=0.9732142857142857
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 6.917101563885808e-05
step: 10, loss: 6.929069058969617e-05
step: 20, loss: 0.0001565226848470047
step: 30, loss: 0.00019206598517484963
step: 40, loss: 3.536274380167015e-05
step: 50, loss: 6.682699313387275e-05
step: 60, loss: 4.662982246372849e-05
step: 70, loss: 2.9856726541765966e-05
step: 80, loss: 5.077280002296902e-05
step: 90, loss: 0.00352520402520895
step: 100, loss: 5.0579281378304586e-05
step: 110, loss: 0.014992509968578815
step: 120, loss: 4.755592453875579e-05
step: 130, loss: 4.404899664223194e-05
step: 140, loss: 7.005257793935016e-05
step: 150, loss: 3.718083098647185e-05
step: 160, loss: 0.0003344917204231024
step: 170, loss: 9.539574239170179e-05
step: 180, loss: 3.3481683203717694e-05
step: 190, loss: 3.249034853070043e-05
step: 200, loss: 4.3057614675490186e-05
step: 210, loss: 2.2667765733785927e-05
```

```
step: 220, loss: 0.00012619365588761866
step: 230, loss: 2.6501062166062184e-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=0.9810055865921787, f1=0.9752252252252253,
best_f1=0.9732142857142857
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0001347194192931056
step: 10, loss: 5.2417468396015465e-05
step: 20, loss: 2.75401325779967e-05
step: 30, loss: 4.2604213376762345e-05
step: 40, loss: 0.0026712703984230757
step: 50, loss: 4.779779555974528e-05
step: 60, loss: 3.342537092976272e-05
step: 70, loss: 0.00010869180550798774
step: 80, loss: 4.6421213482972234e-05
step: 90, loss: 0.0004856487503275275
step: 100, loss: 3.1749346817377955e-05
step: 110, loss: 5.121075810166076e-05
step: 120, loss: 1.8838400137610734e-05
step: 130, loss: 3.095985812251456e-05
step: 140, loss: 3.173843651893549e-05
step: 150, loss: 3.649269274319522e-05
step: 160, loss: 0.0001916232577059418
step: 170, loss: 1.7299984392593615e-05
step: 180, loss: 2.861649227270391e-05
step: 190, loss: 8.158788841683418e-05
step: 200, loss: 2.786385448416695e-05
step: 210, loss: 0.00017761222226146981
step: 220, loss: 1.5809931937837973e-05
step: 230, loss: 0.001070215250365436
/usr/local/lib/python3.7/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=0.9798657718120806, f1=0.9740698985343857,
best_f1=0.9732142857142857
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 2.2738597181160003e-05
step: 10, loss: 1.6759729987825267e-05
step: 20, loss: 5.8014786191051826e-05
step: 30, loss: 3.4754797525238246e-05
step: 40, loss: 4.6586519601987675e-05
step: 50, loss: 0.027047252282500267
```

```
step: 60, loss: 4.225346492603421e-05
step: 70, loss: 7.880443445174024e-05
step: 80, loss: 0.00012071905803168193
step: 90, loss: 8.515013178111985e-05
step: 100, loss: 3.162660505040549e-05
step: 110, loss: 2.159131872758735e-05
step: 120, loss: 3.789192851400003e-05
step: 130, loss: 7.104618271114305e-05
step: 140, loss: 2.601671076263301e-05
step: 150, loss: 6.290211604209617e-05
step: 160, loss: 2.1185265723033808e-05
step: 170, loss: 2.938325997092761e-05
step: 180, loss: 3.064633347094059e-05
step: 190, loss: 3.3022664865711704e-05
step: 200, loss: 3.340683178976178e-05
step: 210, loss: 9.063672041520476e-05
step: 220, loss: 3.2343985367333516e-05
step: 230, loss: 4.015829472336918e-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=0.9798657718120806, f1=0.9740698985343857,
best_f1=0.9732142857142857
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, 205.64it/s]
load_f1 = 0.9807909604519773
real_f1 = 0.9785794813979707
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, 237.78it/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: 0.6277890801429749
step: 10, loss: 0.5401137471199036
step: 20, loss: 0.4925890564918518
step: 30, loss: 0.09358377009630203
step: 40, loss: 0.15697036683559418
step: 50, loss: 0.11799511313438416
step: 60, loss: 0.022120365872979164
step: 70, loss: 0.20516636967658997
step: 80, loss: 0.05157952755689621
step: 90, loss: 0.2309870421886444
step: 100, loss: 0.015370539389550686
step: 110, loss: 0.09851789474487305
```

```
step: 120, loss: 0.020424634218215942
step: 130, loss: 0.10099302232265472
step: 140, loss: 0.1655089408159256
step: 150, loss: 0.053060613572597504
step: 160, loss: 0.053138889372348785
step: 170, loss: 0.2071080505847931
step: 180, loss: 0.08301939815282822
step: 190, loss: 0.006557975430041552
step: 200, loss: 0.1563977599143982
step: 210, loss: 0.10844579339027405
step: 220, loss: 0.3229817748069763
step: 230, loss: 0.13933397829532623
step: 240, loss: 0.04135093465447426
step: 250, loss: 0.014047699049115181
step: 260, loss: 0.054450128227472305
step: 270, loss: 0.02808237262070179
step: 280, loss: 0.06523087620735168
step: 290, loss: 0.050891194492578506
step: 300, loss: 0.034790221601724625
step: 310, loss: 0.10996970534324646
step: 320, loss: 0.10409153252840042
step: 330, loss: 0.008260729722678661
step: 340, loss: 0.035836976021528244
step: 350, loss: 0.0458446629345417
step: 360, loss: 0.05983582139015198
step: 370, loss: 0.0452202707529068
step: 380, loss: 0.005339811556041241
step: 390, loss: 0.19936475157737732
step: 400, loss: 0.23418304324150085
step: 410, loss: 0.05657704547047615
step: 420, loss: 0.11794032901525497
step: 430, loss: 0.1302942931652069
step: 440, loss: 0.02730679325759411
step: 450, loss: 0.006465299054980278
step: 460, loss: 0.018121512606739998
step: 470, loss: 0.13384774327278137
step: 480, loss: 0.07885212451219559
step: 490, loss: 0.10032012313604355
step: 500, loss: 0.045978084206581116
step: 510, loss: 0.10449875891208649
step: 520, loss: 0.17482031881809235
step: 530, loss: 0.0064847455359995365
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9400921658986175, f1=0.9385884509624199,
best_f1=0.9385884509624199
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.1211157888174057
step: 10, loss: 0.0477655865252018
step: 20, loss: 0.003342619864270091
step: 30, loss: 0.007796016987413168
step: 40, loss: 0.07832899689674377
step: 50, loss: 0.07141371071338654
step: 60, loss: 0.03334581479430199
step: 70, loss: 0.018582817167043686
step: 80, loss: 0.02820514515042305
step: 90, loss: 0.010513702407479286
step: 100, loss: 0.01664401777088642
step: 110, loss: 0.004755600355565548
step: 120, loss: 0.08182498067617416
step: 130, loss: 0.03648056089878082
step: 140, loss: 0.013420580886304379
step: 150, loss: 0.07560335844755173
step: 160, loss: 0.03077135607600212
step: 170, loss: 0.022284122183918953
step: 180, loss: 0.06717805564403534
step: 190, loss: 0.14986728131771088
step: 200, loss: 0.007931994274258614
step: 210, loss: 0.0427282340824604
step: 220, loss: 0.034631840884685516
step: 230, loss: 0.011200609616935253
step: 240, loss: 0.006820487789809704
step: 250, loss: 0.0036434470675885677
step: 260, loss: 0.005505657754838467
step: 270, loss: 0.19980962574481964
step: 280, loss: 0.1843961626291275
step: 290, loss: 0.02369779162108898
step: 300, loss: 0.15149842202663422
step: 310, loss: 0.008917215280234814
step: 320, loss: 0.06599016487598419
step: 330, loss: 0.015398976393043995
step: 340, loss: 0.018164489418268204
step: 350, loss: 0.0012905787443742156
step: 360, loss: 0.11535002291202545
step: 370, loss: 0.07539691776037216
step: 380, loss: 0.04594064876437187
step: 390, loss: 0.08317992836236954
step: 400, loss: 0.01050263550132513
step: 410, loss: 0.04844405874609947
```

```
step: 420, loss: 0.026265397667884827
step: 430, loss: 0.0682498961687088
step: 440, loss: 0.18595291674137115
step: 450, loss: 0.04901513084769249
step: 460, loss: 0.06361862272024155
step: 470, loss: 0.11881938576698303
step: 480, loss: 0.1620926856994629
step: 490, loss: 0.016985617578029633
step: 500, loss: 0.099538654088974
step: 510, loss: 0.00711823720484972
step: 520, loss: 0.0744623914361
step: 530, loss: 0.004302572458982468
/usr/local/lib/python3.7/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=0.9383402874362541, f1=0.9361702127659576,
best_f1=0.9385884509624199
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.024890506640076637
step: 10, loss: 0.030244987457990646
step: 20, loss: 0.029570670798420906
step: 30, loss: 0.033719804137945175
step: 40, loss: 0.14237116277217865
step: 50, loss: 0.05380785092711449
step: 60, loss: 0.015229018405079842
step: 70, loss: 0.003820891957730055
step: 80, loss: 0.028413234278559685
step: 90, loss: 0.005277583375573158
step: 100, loss: 0.07085691392421722
step: 110, loss: 0.0009531021933071315
step: 120, loss: 0.0030116806738078594
step: 130, loss: 0.005403550341725349
step: 140, loss: 0.052197568118572235
step: 150, loss: 0.01705251820385456
step: 160, loss: 0.05670924484729767
step: 170, loss: 0.11197502911090851
step: 180, loss: 0.030233105644583702
step: 190, loss: 0.016196420416235924
step: 200, loss: 0.09643767774105072
step: 210, loss: 0.012570958584547043
step: 220, loss: 0.2870924174785614
step: 230, loss: 0.1771148145198822
step: 240, loss: 0.005214963108301163
step: 250, loss: 0.013010144233703613
step: 260, loss: 0.010617159307003021
step: 270, loss: 0.008892916142940521
```

```
step: 280, loss: 0.05905481055378914
step: 290, loss: 0.0027998003643006086
step: 300, loss: 0.06305889040231705
step: 310, loss: 0.008247855119407177
step: 320, loss: 0.10121597349643707
step: 330, loss: 0.017338305711746216
step: 340, loss: 0.00426814891397953
step: 350, loss: 0.0048324051313102245
step: 360, loss: 0.036986734718084335
step: 370, loss: 0.010715476237237453
step: 380, loss: 0.007789201103150845
step: 390, loss: 0.007047706749290228
step: 400, loss: 0.017343733459711075
step: 410, loss: 0.012584852054715157
step: 420, loss: 0.18326450884342194
step: 430, loss: 0.14313212037086487
step: 440, loss: 0.009874312207102776
step: 450, loss: 0.05186435952782631
step: 460, loss: 0.042007628828287125
step: 470, loss: 0.10368693619966507
step: 480, loss: 0.0020001260563731194
step: 490, loss: 0.0027760942466557026
step: 500, loss: 0.025789018720388412
step: 510, loss: 0.015513168647885323
step: 520, loss: 0.040211271494627
step: 530, loss: 0.02298365905880928
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9442622950819672, f1=0.9337094499294781,
best_f1=0.9337094499294781
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.004029400646686554
step: 10, loss: 0.0031598226632922888
step: 20, loss: 0.0021823272109031677
step: 30, loss: 0.0011236706050112844
step: 40, loss: 0.009433098137378693
step: 50, loss: 0.0014454106567427516
step: 60, loss: 0.0006287130527198315
step: 70, loss: 0.00445720087736845
step: 80, loss: 0.0065414863638579845
step: 90, loss: 0.11655863374471664
```

```
step: 100, loss: 0.028793998062610626
step: 110, loss: 0.027862319722771645
step: 120, loss: 0.002509559039026499
step: 130, loss: 0.00173733732663095
step: 140, loss: 0.0059920852072536945
step: 150, loss: 0.008801323361694813
step: 160, loss: 0.024781756103038788
step: 170, loss: 0.004622488282620907
step: 180, loss: 0.006789210252463818
step: 190, loss: 0.003649692749604583
step: 200, loss: 0.0017819618806242943
step: 210, loss: 0.12227856367826462
step: 220, loss: 0.003532339120283723
step: 230, loss: 0.010673709213733673
step: 240, loss: 0.0019374415278434753
step: 250, loss: 0.12793157994747162
step: 260, loss: 0.0018993888515979052
step: 270, loss: 0.005094545427709818
step: 280, loss: 0.017458468675613403
step: 290, loss: 0.043654683977365494
step: 300, loss: 0.0003125408256892115
step: 310, loss: 0.0003632127190940082
step: 320, loss: 0.0005837606149725616
step: 330, loss: 0.0029162445571273565
step: 340, loss: 0.03356395289301872
step: 350, loss: 0.05105164647102356
step: 360, loss: 0.004505367949604988
step: 370, loss: 0.0038112029433250427
step: 380, loss: 0.015389458276331425
step: 390, loss: 0.006079785991460085
step: 400, loss: 0.010752893052995205
step: 410, loss: 0.0016250339103862643
step: 420, loss: 0.0008481679251417518
step: 430, loss: 0.09020119905471802
Gradient overflow.  Skipping step, loss scaler 0 reducing loss scale to 32768.0
step: 440, loss: 0.007390564773231745
step: 450, loss: 0.0017554184887558222
step: 460, loss: 0.0004815805295947939
step: 470, loss: 0.04125542938709259
step: 480, loss: 0.04979235678911209
step: 490, loss: 0.011660313233733177
step: 500, loss: 0.012067980132997036
step: 510, loss: 0.01640443131327629
step: 520, loss: 0.2337028831243515
step: 530, loss: 0.031108904629945755
/usr/local/lib/python3.7/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=0.9368709972552608, f1=0.9402644778841769,
best_f1=0.9337094499294781
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.019901832565665245
step: 10, loss: 0.019820135086774826
step: 20, loss: 0.01383666880428791
step: 30, loss: 0.0016255427617579699
step: 40, loss: 0.050771184265613556
step: 50, loss: 0.0005433229962363839
step: 60, loss: 0.05145745351910591
step: 70, loss: 0.0010125036351382732
step: 80, loss: 0.005387888755649328
step: 90, loss: 0.014230873435735703
step: 100, loss: 0.0014576801331713796
step: 110, loss: 0.0014366862596943974
step: 120, loss: 0.002006460912525654
step: 130, loss: 0.000402974896132946
step: 140, loss: 0.0012014267267659307
step: 150, loss: 0.01397110614925623
step: 160, loss: 0.00027145532658323646
step: 170, loss: 0.031449928879737854
step: 180, loss: 0.00043977409950457513
step: 190, loss: 0.026059091091156006
step: 200, loss: 0.0005186606431379914
step: 210, loss: 0.037274230271577835
step: 220, loss: 0.0010889873374253511
step: 230, loss: 0.0032902390230447054
step: 240, loss: 0.00033385661663487554
step: 250, loss: 0.00016519712517037988
step: 260, loss: 0.000385638588340953
step: 270, loss: 0.0003206878318451345
step: 280, loss: 0.08906394988298416
step: 290, loss: 0.2223379760980606
step: 300, loss: 0.014010470360517502
step: 310, loss: 0.0013560400111600757
step: 320, loss: 0.06522440910339355
step: 330, loss: 0.005517344456166029
step: 340, loss: 0.003196856938302517
step: 350, loss: 0.009644375182688236
step: 360, loss: 0.020292842760682106
step: 370, loss: 0.06336291134357452
step: 380, loss: 0.019686011597514153
step: 390, loss: 0.0002327578840777278
step: 400, loss: 0.0015990639803931117
step: 410, loss: 0.0014078631065785885
step: 420, loss: 0.0012799229007214308
```

```
step: 430, loss: 0.03158478066325188
step: 440, loss: 0.0639154314994812
step: 450, loss: 0.0037433423567563295
step: 460, loss: 0.0008429995505139232
step: 470, loss: 0.09505152702331543
step: 480, loss: 0.0032656630501151085
step: 490, loss: 0.0008515637018717825
step: 500, loss: 0.0015430997591465712
step: 510, loss: 0.000693190551828593
step: 520, loss: 0.001225095707923174
step: 530, loss: 0.002215772634372115
/usr/local/lib/python3.7/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=0.9375, best_f1=0.9337094499294781
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00525762839242816
step: 10, loss: 0.001160134095698595
step: 20, loss: 0.11241672933101654
step: 30, loss: 0.003891445230692625
step: 40, loss: 0.00049790262710303071
step: 50, loss: 0.014318653382360935
step: 60, loss: 0.00010469734843354672
step: 70, loss: 0.00020379199122544378
step: 80, loss: 0.0008835762273520231
step: 90, loss: 0.00029903455288149416
step: 100, loss: 0.0009233365999534726
step: 110, loss: 0.00011635648115770891
step: 120, loss: 0.0005847836728207767
step: 130, loss: 0.0004958037752658129
step: 140, loss: 0.002679811092093587
step: 150, loss: 0.007288631051778793
step: 160, loss: 0.001559481373988092
step: 170, loss: 0.001334390603005886
step: 180, loss: 0.00017965779989026487
step: 190, loss: 0.0017504667630419135
step: 200, loss: 0.00014619884314015508
step: 210, loss: 0.0018052497180178761
step: 220, loss: 0.00562336714938283
step: 230, loss: 0.0003312480985186994
step: 240, loss: 0.03172751143574715
step: 250, loss: 0.00017744407523423433
step: 260, loss: 7.221769192256033e-05
step: 270, loss: 0.036002520471811295
step: 280, loss: 0.002923786174505949
step: 290, loss: 0.0003294544294476509
```

```
step: 300, loss: 0.00021441663557197899
step: 310, loss: 0.0025570187717676163
step: 320, loss: 0.011787804774940014
step: 330, loss: 0.0011238261358812451
step: 340, loss: 0.0990300104022026
step: 350, loss: 0.0013619504170492291
step: 360, loss: 0.014731962233781815
step: 370, loss: 0.006709308363497257
step: 380, loss: 0.00798651110380888
Gradient overflow.  Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 390, loss: 0.11483172327280045
step: 400, loss: 0.06127770245075226
step: 410, loss: 0.0005869162851013243
step: 420, loss: 0.0059735761024057865
step: 430, loss: 0.0007967376732267439
step: 440, loss: 0.00024131541431415826
step: 450, loss: 0.0012457005213946104
step: 460, loss: 9.803700959309936e-05
step: 470, loss: 0.00662261014804244
step: 480, loss: 0.0012758359080180526
step: 490, loss: 0.008838511072099209
step: 500, loss: 0.0012582233175635338
step: 510, loss: 0.004023327026516199
step: 520, loss: 0.05839981138706207
step: 530, loss: 0.004548028577119112
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9444444444444445, f1=0.9401473296500922,
best_f1=0.9401473296500922
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0003279993252363056
step: 10, loss: 0.00012095070269424468
step: 20, loss: 0.0006629256531596184
step: 30, loss: 0.00047764720511622727
step: 40, loss: 0.0005574729293584824
step: 50, loss: 0.000732859072741121
step: 60, loss: 0.0002195872220909223
step: 70, loss: 0.0001396564912283793
step: 80, loss: 0.0002924291475210339
step: 90, loss: 0.005865878891199827
step: 100, loss: 7.383839692920446e-05
```

```
step: 110, loss: 0.00038922313251532614
step: 120, loss: 0.0002607672940939665
step: 130, loss: 0.08029236644506454
step: 140, loss: 7.005201769061387e-05
step: 150, loss: 9.01827952475287e-05
step: 160, loss: 0.00018371517944615334
step: 170, loss: 0.00014928060409147292
step: 180, loss: 0.007441405206918716
step: 190, loss: 0.0011366576654836535
step: 200, loss: 0.00021456957620102912
step: 210, loss: 0.00014636847481597215
step: 220, loss: 8.735043957130983e-05
step: 230, loss: 0.0012024492025375366
step: 240, loss: 0.00043785845628008246
step: 250, loss: 0.012077820487320423
step: 260, loss: 0.00011969144543400034
step: 270, loss: 0.0003594333538785577
step: 280, loss: 0.0007949810824356973
step: 290, loss: 0.00012913662067148834
step: 300, loss: 0.002874132012948394
step: 310, loss: 0.0002907110610976815
step: 320, loss: 5.67601528018713e-05
step: 330, loss: 0.0008077375823631883
step: 340, loss: 0.0004796165449079126
step: 350, loss: 0.001972158206626773
step: 360, loss: 0.003688022494316101
step: 370, loss: 0.0008453500922769308
step: 380, loss: 0.0240325927734375
step: 390, loss: 0.00023088915622793138
step: 400, loss: 0.0013621748657897115
step: 410, loss: 0.003585544414818287
step: 420, loss: 0.024348212406039238
step: 430, loss: 8.11056888778694e-05
step: 440, loss: 0.000106784536910709
step: 450, loss: 0.012155337259173393
step: 460, loss: 0.00045598275028169155
step: 470, loss: 0.00791491661220789
step: 480, loss: 0.06854856759309769
step: 490, loss: 0.01694074645638466
step: 500, loss: 0.0010097639169543982
step: 510, loss: 0.0872812271118164
step: 520, loss: 0.02485678344964981
step: 530, loss: 0.01139755081385374
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9454878607420981, f1=0.9372997711670481,
best_f1=0.9372997711670481
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00019285579037386924
step: 10, loss: 0.023263471201062202
step: 20, loss: 0.0001089574652723968
step: 30, loss: 0.008192921057343483
step: 40, loss: 4.88999648950994e-05
step: 50, loss: 0.006875291001051664
step: 60, loss: 0.00043855467811226845
step: 70, loss: 0.019032398238778114
step: 80, loss: 0.002714860253036022
step: 90, loss: 0.0009240591898560524
step: 100, loss: 0.00034236389910802245
step: 110, loss: 0.0010982933454215527
step: 120, loss: 0.006473291199654341
step: 130, loss: 0.00042714734445326030
step: 140, loss: 0.011484192684292793
step: 150, loss: 9.480108565185219e-05
step: 160, loss: 0.0022754317615181208
step: 170, loss: 0.005010291468352079
step: 180, loss: 5.3803629270987585e-05
step: 190, loss: 0.003039880422875285
step: 200, loss: 0.0013977609341964126
step: 210, loss: 0.00010274403030052781
step: 220, loss: 0.002912912517786026
step: 230, loss: 0.0002829078584909439
step: 240, loss: 0.00022734698723070323
step: 250, loss: 0.0003643865929916501
step: 260, loss: 3.7571913708234206e-05
step: 270, loss: 0.0015984972706064582
step: 280, loss: 0.10042616724967957
step: 290, loss: 0.00033102091401815414
step: 300, loss: 0.013247620314359665
step: 310, loss: 0.002351634204387665
step: 320, loss: 0.0009152052807621658
step: 330, loss: 0.008892995305359364
step: 340, loss: 9.141231566900387e-05
step: 350, loss: 8.995326788863167e-05
step: 360, loss: 0.0006843434530310333
step: 370, loss: 0.0020595737732946873
step: 380, loss: 0.0006976762088015676
step: 390, loss: 0.14056101441383362
step: 400, loss: 0.0019074215088039637
```

```
step: 410, loss: 0.00020735952421091497
step: 420, loss: 0.0025091145653277636
step: 430, loss: 0.002417110837996006
step: 440, loss: 0.0016099673230201006
step: 450, loss: 0.0025987299159169197
step: 460, loss: 0.0004774643457494676
step: 470, loss: 0.0003398527333047241
step: 480, loss: 5.794157914351672e-05
step: 490, loss: 0.0023451042361557484
step: 500, loss: 0.0035822875797748566
step: 510, loss: 6.551074329763651e-05
step: 520, loss: 0.0013762138551101089
step: 530, loss: 0.0006283746333792806
/usr/local/lib/python3.7/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=0.9423887587822014, f1=0.9350163627863488,
best_f1=0.9372997711670481
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.002291990676894784
step: 10, loss: 0.00023019824584480375
step: 20, loss: 0.05825866013765335
step: 30, loss: 0.0564480684697628
step: 40, loss: 0.0007930222782306373
step: 50, loss: 9.817945101531222e-05
step: 60, loss: 8.168078784365207e-05
step: 70, loss: 0.0001599852548679337
step: 80, loss: 0.001297641429118812
step: 90, loss: 5.603625322692096e-05
step: 100, loss: 4.926931069348939e-05
step: 110, loss: 3.688247670652345e-05
step: 120, loss: 9.6752199169714e-05
step: 130, loss: 0.0622701495885849
step: 140, loss: 0.0004494689346756786
step: 150, loss: 8.196422277251258e-05
step: 160, loss: 0.0006392227951437235
step: 170, loss: 0.00757882185280323
step: 180, loss: 3.925777491531335e-05
step: 190, loss: 4.011426790384576e-05
step: 200, loss: 0.0005850880988873541
step: 210, loss: 0.00018131794058717787
step: 220, loss: 0.004168723709881306
step: 230, loss: 0.0001256784307770431
step: 240, loss: 5.367711855797097e-05
step: 250, loss: 0.0014583999291062355
step: 260, loss: 0.010410862974822521
```

```
step: 270, loss: 0.0001799377496354282
step: 280, loss: 0.0005648242076858878
step: 290, loss: 0.0900806114077568
step: 300, loss: 0.05195058137178421
step: 310, loss: 0.0091116176918149
step: 320, loss: 0.017636224627494812
step: 330, loss: 0.00011476722283987328
step: 340, loss: 0.0026528039015829563
step: 350, loss: 0.06775032728910446
step: 360, loss: 2.020567626459524e-05
step: 370, loss: 8.122391591314226e-05
step: 380, loss: 0.00048351369332522154
step: 390, loss: 4.4762768084183335e-05
step: 400, loss: 0.0009736159699968994
step: 410, loss: 2.5692410417832434e-05
step: 420, loss: 0.00014075826038606465
step: 430, loss: 8.401070954278111e-05
step: 440, loss: 0.0020448477007448673
step: 450, loss: 5.2750332542927936e-05
step: 460, loss: 0.0006686695851385593
step: 470, loss: 3.584170190151781e-05
step: 480, loss: 0.0006502229371108115
step: 490, loss: 0.0005528644542209804
step: 500, loss: 0.0017267153598368168
step: 510, loss: 0.000798159628175199
step: 520, loss: 0.004854770842939615
step: 530, loss: 0.011541763320565224
/usr/local/lib/python3.7/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=0.9397477814105558, f1=0.9394221808014911,
best_f1=0.9372997711670481
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.001853837282396853
step: 10, loss: 6.460327131208032e-05
step: 20, loss: 0.0001205945955007337
step: 30, loss: 0.00020503650011960417
step: 40, loss: 0.0010543528478592634
step: 50, loss: 0.011734250001609325
step: 60, loss: 0.00027808151207864285
step: 70, loss: 0.0002982491278089583
step: 80, loss: 4.495868051890284e-05
step: 90, loss: 0.00028723233845084906
step: 100, loss: 0.001761143677867949
step: 110, loss: 2.2388503566617146e-05
step: 120, loss: 0.00045091332867741585
```

```
step: 130, loss: 2.9942815672256984e-05
step: 140, loss: 0.0008037713705562055
step: 150, loss: 9.732276521390304e-05
step: 160, loss: 2.931322160293348e-05
step: 170, loss: 0.0006888905190862715
step: 180, loss: 0.0001435232552466914
step: 190, loss: 3.133227801299654e-05
step: 200, loss: 0.0005986959440633655
step: 210, loss: 0.08385951071977615
step: 220, loss: 0.0026507165748625994
step: 230, loss: 0.013507096096873283
step: 240, loss: 0.0003071201790589839
step: 250, loss: 6.460878648795187e-05
step: 260, loss: 0.0006056064157746732
step: 270, loss: 0.00027082496671937406
step: 280, loss: 3.8551468605874106e-05
step: 290, loss: 0.00011827194248326123
step: 300, loss: 4.8343110393034294e-05
step: 310, loss: 8.970459748525172e-05
step: 320, loss: 0.0027176993899047375
step: 330, loss: 0.0002189222868764773
step: 340, loss: 7.50070612411946e-05
step: 350, loss: 7.613956404384226e-05
step: 360, loss: 0.003264127066358924
step: 370, loss: 0.0035232282243669033
step: 380, loss: 0.00020388298435136676
step: 390, loss: 0.0002879300736822188
step: 400, loss: 0.0003498020814731717
step: 410, loss: 0.0031711712945252657
step: 420, loss: 0.0003260485827922821
step: 430, loss: 0.00047829069080762565
step: 440, loss: 0.0003436666738707572
step: 450, loss: 2.7000080081052147e-05
step: 460, loss: 3.514099444146268e-05
step: 470, loss: 2.9332346457522362e-05
step: 480, loss: 0.00013640450197272003
step: 490, loss: 0.00011297149467281997
step: 500, loss: 0.004771411418914795
step: 510, loss: 2.344276254007127e-05
step: 520, loss: 6.668551941402256e-05
step: 530, loss: 0.0012149200774729252
/usr/local/lib/python3.7/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=0.9438727782974743, f1=0.9388136384866884,
best_f1=0.9372997711670481
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
```

```
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 7.806793291820213e-05
step: 10, loss: 9.851338109001517e-05
step: 20, loss: 8.162757148966193e-05
step: 30, loss: 0.00019681004050653428
step: 40, loss: 4.046364119858481e-05
step: 50, loss: 6.004919487168081e-05
step: 60, loss: 0.012370854616165161
step: 70, loss: 5.334506568033248e-05
step: 80, loss: 2.451531145197805e-05
step: 90, loss: 6.979874888202175e-05
step: 100, loss: 6.140807090559974e-05
step: 110, loss: 5.880527896806598e-05
step: 120, loss: 1.9430781321716495e-05
step: 130, loss: 1.7139787814812735e-05
step: 140, loss: 4.3454652768559754e-05
step: 150, loss: 1.5962659745127894e-05
step: 160, loss: 2.905945984821301e-05
step: 170, loss: 2.2511461793328635e-05
step: 180, loss: 9.802373824641109e-05
step: 190, loss: 4.523535608313978e-05
step: 200, loss: 0.00037791542126799714
step: 210, loss: 0.0001482255320297554
step: 220, loss: 0.0018002755241468549
step: 230, loss: 1.7866135749500245e-05
step: 240, loss: 0.0002581153530627489
step: 250, loss: 3.356857268954627e-05
step: 260, loss: 2.1013858713558875e-05
step: 270, loss: 2.660150312294718e-05
step: 280, loss: 0.004674956668168306
step: 290, loss: 0.001339248032309115
step: 300, loss: 5.5139891628641635e-05
step: 310, loss: 0.0017545027658343315
step: 320, loss: 2.3491089450544678e-05
step: 330, loss: 2.311500247742515e-05
step: 340, loss: 1.942337621585466e-05
Gradient overflow.  Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 350, loss: 3.1716986995888874e-05
step: 360, loss: 5.2113711717538536e-05
step: 370, loss: 0.0006557742599397898
step: 380, loss: 2.0589322957675904e-05
step: 390, loss: 5.71264790778514e-05
step: 400, loss: 4.5213466364657506e-05
step: 410, loss: 7.860890036681667e-05
step: 420, loss: 0.0002947856264654547
step: 430, loss: 4.155679198447615e-05
step: 440, loss: 0.010387399233877659
step: 450, loss: 0.00010949755960609764
```

```
step: 460, loss: 0.0109226852655441077
step: 470, loss: 0.0002217740984633565
step: 480, loss: 0.00013465843221638352
step: 490, loss: 8.122319559333846e-05
step: 500, loss: 4.9631173169473186e-05
step: 510, loss: 1.8596219888422638e-05
step: 520, loss: 2.9028977223788388e-05
step: 530, loss: 1.6189929738175124e-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=0.9447565543071161, f1=0.9340196537201684,
best_f1=0.9372997711670481
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 3.837453914457001e-05
step: 10, loss: 1.3552477867051493e-05
step: 20, loss: 1.9691473426064476e-05
step: 30, loss: 0.002082939026877284
step: 40, loss: 0.00019966527179349214
step: 50, loss: 0.10259127616882324
step: 60, loss: 0.0011510838521644473
step: 70, loss: 2.4608343665022403e-05
step: 80, loss: 2.2202029867912643e-05
step: 90, loss: 5.801049337605946e-05
step: 100, loss: 0.00040180247742682695
step: 110, loss: 8.506998710799962e-05
step: 120, loss: 1.4088914213061798e-05
step: 130, loss: 0.001394439022988081
step: 140, loss: 2.075691008940339e-05
step: 150, loss: 0.00015495788829866797
step: 160, loss: 1.3731296348851174e-05
step: 170, loss: 3.141288834740408e-05
step: 180, loss: 2.1922809537500143e-05
step: 190, loss: 0.0003965340438298881
step: 200, loss: 0.0037228583823889494
step: 210, loss: 4.256219108356163e-05
step: 220, loss: 2.1826055672136135e-05
step: 230, loss: 1.3735017091676127e-05
step: 240, loss: 2.67522300418932e-05
step: 250, loss: 2.1374658899731003e-05
step: 260, loss: 2.0321087504271418e-05
step: 270, loss: 4.916092802886851e-05
step: 280, loss: 0.001514533767476678
step: 290, loss: 4.591490142047405e-05
step: 300, loss: 0.0006052168319001794
step: 310, loss: 5.630679879686795e-05
```

```
step: 320, loss: 5.389426223700866e-05
step: 330, loss: 0.0016781092854216695
step: 340, loss: 1.9039376638829708e-05
step: 350, loss: 3.734952042577788e-05
step: 360, loss: 0.0001497109333286062
step: 370, loss: 1.822387639549561e-05
step: 380, loss: 8.013076876522973e-05
step: 390, loss: 2.3297430743696168e-05
step: 400, loss: 2.093930925184395e-05
step: 410, loss: 1.6934931409195997e-05
step: 420, loss: 0.005685355514287949
step: 430, loss: 0.0011130614439025521
step: 440, loss: 1.580251228006091e-05
step: 450, loss: 7.64301948947832e-05
step: 460, loss: 2.3203932869364507e-05
step: 470, loss: 2.094288356602192e-05
step: 480, loss: 2.908915121224709e-05
step: 490, loss: 2.022801709244959e-05
step: 500, loss: 2.0510826288955286e-05
step: 510, loss: 0.002382053527981043
step: 520, loss: 3.3957625419134274e-05
step: 530, loss: 0.0003995473380200565
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9463276836158193, f1=0.9315714959886738,
best_f1=0.9315714959886738
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.005869815591722727
step: 10, loss: 2.4990844394778833e-05
step: 20, loss: 2.235809006378986e-05
step: 30, loss: 0.0006342357373796403
step: 40, loss: 1.7083886632462963e-05
step: 50, loss: 3.477520294336722e-05
step: 60, loss: 1.8283119061379693e-05
step: 70, loss: 1.7579093764652498e-05
step: 80, loss: 2.370597758272197e-05
step: 90, loss: 6.125629442976788e-05
step: 100, loss: 3.389563789824024e-05
step: 110, loss: 1.4733306670677848e-05
step: 120, loss: 2.084252264467068e-05
step: 130, loss: 1.2431184586603194e-05
```

```
step: 140, loss: 3.3661224733805284e-05
step: 150, loss: 5.0937051128130406e-05
step: 160, loss: 1.675980820436962e-05
step: 170, loss: 3.177334292558953e-05
step: 180, loss: 4.135006747674197e-05
step: 190, loss: 0.0003965084324590862
step: 200, loss: 2.894335011660587e-05
step: 210, loss: 2.435886017337907e-05
step: 220, loss: 0.0018674946622923017
step: 230, loss: 0.0001558394724270329
step: 240, loss: 8.679387246957049e-05
step: 250, loss: 2.9644743335666135e-05
step: 260, loss: 1.7969949112739414e-05
step: 270, loss: 1.9646571672637947e-05
step: 280, loss: 3.7697729567298666e-05
step: 290, loss: 1.2501939636422321e-05
step: 300, loss: 0.00013044085062574595
step: 310, loss: 0.0014915333595126867
step: 320, loss: 0.003159275045618415
step: 330, loss: 0.002558006439357996
step: 340, loss: 1.7850983567768708e-05
step: 350, loss: 1.8365084542892873e-05
step: 360, loss: 1.2602538845385425e-05
step: 370, loss: 0.00010826573998201638
step: 380, loss: 1.5370334949693643e-05
step: 390, loss: 0.00038849280099384487
step: 400, loss: 3.0336663257912733e-05
step: 410, loss: 0.00014331206330098212
step: 420, loss: 1.2401354979374446e-05
step: 430, loss: 0.00170781253837049
step: 440, loss: 0.00011457499931566417
step: 450, loss: 4.826101576327346e-05
step: 460, loss: 0.002815593034029007
step: 470, loss: 0.018956787884235382
step: 480, loss: 4.990110028302297e-05
step: 490, loss: 1.609676837688312e-05
step: 500, loss: 0.00010701519931899384
step: 510, loss: 3.519875099300407e-05
step: 520, loss: 0.0006519740563817322
step: 530, loss: 7.854896102799103e-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=0.939294117647059, f1=0.9328953542937587,
best_f1=0.9315714959886738
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
```

```
step: 0, loss: 0.0020146854221820083
step: 10, loss: 1.1447691576904617e-05
step: 20, loss: 1.5515526683884673e-05
step: 30, loss: 1.592912303749472e-05
step: 40, loss: 3.5756220313487574e-05
step: 50, loss: 0.00023267797951120883
step: 60, loss: 6.140947516541928e-05
step: 70, loss: 0.0002075389347737655
step: 80, loss: 2.036933983617928e-05
step: 90, loss: 1.7061413018382154e-05
step: 100, loss: 0.00011994758096989244
step: 110, loss: 4.187955710222013e-05
step: 120, loss: 2.0439956642803736e-05
step: 130, loss: 0.003923320211470127
step: 140, loss: 0.00015667347179260105
step: 150, loss: 6.868817581562325e-05
step: 160, loss: 0.0008593078819103539
step: 170, loss: 0.007677081041038036
step: 180, loss: 1.3585963642981369e-05
step: 190, loss: 5.6955774198286235e-05
step: 200, loss: 0.00013679797120857984
step: 210, loss: 0.00017739225586410612
step: 220, loss: 0.0001511242298875004
step: 230, loss: 0.0006850072531960905
step: 240, loss: 1.2624812370631844e-05
step: 250, loss: 2.4090568331303075e-05
step: 260, loss: 9.108879021368921e-05
step: 270, loss: 0.00010349391959607601
step: 280, loss: 1.3228247553342953e-05
step: 290, loss: 0.00040979203185997903
step: 300, loss: 4.1012404835782945e-05
step: 310, loss: 2.294340265507344e-05
step: 320, loss: 0.0015216696774587035
step: 330, loss: 3.9108246710384265e-05
step: 340, loss: 6.087867222959176e-05
step: 350, loss: 0.005849554669111967
step: 360, loss: 6.449200009228662e-05
step: 370, loss: 3.7240100937196985e-05
step: 380, loss: 3.219677091692574e-05
step: 390, loss: 0.00380121823400259
step: 400, loss: 8.424829138675705e-05
step: 410, loss: 2.0473044060054235e-05
step: 420, loss: 1.0684042536013294e-05
step: 430, loss: 7.326593186007813e-05
step: 440, loss: 0.00010686753375921398
step: 450, loss: 4.5683918870054185e-05
step: 460, loss: 0.0001117471547331661
step: 470, loss: 2.4546294298488647e-05
```

```
step: 480, loss: 1.852149398473557e-05
step: 490, loss: 1.4632593774877023e-05
step: 500, loss: 1.1011864444299135e-05
step: 510, loss: 0.01964733749628067
step: 520, loss: 2.717438110266812e-05
step: 530, loss: 0.0001826496300054714
/usr/local/lib/python3.7/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=0.9422806194274989, f1=0.9348946135831382,
best_f1=0.9315714959886738
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 2.0074876374565065e-05
step: 10, loss: 1.984750088013243e-05
step: 20, loss: 2.8052474590367638e-05
step: 30, loss: 3.50852860711989685e-05
step: 40, loss: 0.07477300614118576
step: 50, loss: 9.44501516642049e-05
step: 60, loss: 1.2013923878839705e-05
step: 70, loss: 0.0008461989928036928
step: 80, loss: 2.302854227309581e-05
step: 90, loss: 0.00022906801314093173
step: 100, loss: 2.748381120909471e-05
step: 110, loss: 0.001699271728284657
step: 120, loss: 2.3977845557965338e-05
step: 130, loss: 0.0838041752576828
step: 140, loss: 2.554076854721643e-05
step: 150, loss: 2.732291613938287e-05
step: 160, loss: 2.7038815460400656e-05
step: 170, loss: 1.2747771506838035e-05
step: 180, loss: 1.0173699592996854e-05
step: 190, loss: 2.1076853954582475e-05
step: 200, loss: 1.1306139640510082e-05
step: 210, loss: 0.000439910392742604
step: 220, loss: 5.563720333157107e-05
step: 230, loss: 1.6264288206002675e-05
step: 240, loss: 0.0008142040460370481
step: 250, loss: 3.0753704777453095e-05
step: 260, loss: 1.7225405827048235e-05
step: 270, loss: 1.563844853080809e-05
step: 280, loss: 3.847094558295794e-05
step: 290, loss: 1.1198111678822897e-05
step: 300, loss: 4.193146014586091e-05
step: 310, loss: 0.00018113618716597557
step: 320, loss: 1.615987093828153e-05
step: 330, loss: 7.738572458038107e-05
```

```
step: 340, loss: 1.5388848623842932e-05
step: 350, loss: 1.2181457350379787e-05
step: 360, loss: 0.00209937640465796
step: 370, loss: 9.611186214897316e-06
step: 380, loss: 4.418215758050792e-05
step: 390, loss: 4.372711919131689e-05
step: 400, loss: 2.9148712201276794e-05
step: 410, loss: 6.757755909347907e-05
step: 420, loss: 0.00018785153224598616
step: 430, loss: 8.081327541731298e-05
step: 440, loss: 2.7405023502069525e-05
step: 450, loss: 1.4301104783953633e-05
step: 460, loss: 4.416811134433374e-05
step: 470, loss: 0.002837070031091571
step: 480, loss: 1.0326406481908634e-05
step: 490, loss: 1.8748667571344413e-05
step: 500, loss: 2.5819976144703105e-05
step: 510, loss: 2.156494156224653e-05
step: 520, loss: 2.0268289517844096e-05
step: 530, loss: 0.0006601383211091161
/usr/local/lib/python3.7/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=0.942008486562942, f1=0.9337094499294781,
best_f1=0.9315714959886738
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 [00:23, 249.09it/s]
load_f1 = 0.9388136384866884
real_f1 = 0.9372384937238494
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, 256.91it/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: 0.5410853028297424
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:

```
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.4, f1=0.31034482758620696, best_f1=0.31034482758620696
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.5241936445236206
/usr/local/lib/python3.7/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=0.4, f1=0.36666666666666664, best_f1=0.31034482758620696
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.5383229851722717
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.358974358974359,
best_f1=0.358974358974359
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.20367753505706787
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.6666666666666666, f1=0.3783783783783784,
best_f1=0.3783783783783784
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.10856222361326218
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
```

```
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.7741935483870968, 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: 0.11065379530191422
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.7999999999999999, f1=0.6875000000000001,
best_f1=0.6875000000000001
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.04349147900938988
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.8750000000000001, f1=0.6666666666666667,
best_f1=0.6666666666666667
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.008650084026157856
/usr/local/lib/python3.7/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=0.8484848484848484, f1=0.6285714285714286,
best_f1=0.6666666666666667
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0020829488057643175
/usr/local/lib/python3.7/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=0.8484848484848484, f1=0.6470588235294117,
best_f1=0.6666666666666667
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.004934072960168123
/usr/local/lib/python3.7/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=0.8484848484848484, f1=0.6470588235294117,
best_f1=0.6666666666666667
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00290399300865829
/usr/local/lib/python3.7/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=0.8484848484848484, f1=0.6666666666666667,
best_f1=0.6666666666666667
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.005702310707420111
/usr/local/lib/python3.7/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=0.8484848484848484, f1=0.6470588235294117,
best_f1=0.6666666666666667
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.002905486850067973
/usr/local/lib/python3.7/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=0.8484848484848484, f1=0.6666666666666667,
best_f1=0.6666666666666667
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.002363965380936861
/usr/local/lib/python3.7/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=0.8484848484848484, f1=0.6666666666666667,
best_f1=0.6666666666666667
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
```

```
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.002391693415120244
/usr/local/lib/python3.7/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=0.8484848484848484, f1=0.6666666666666667,
best_f1=0.6666666666666667
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, 136120.42it/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 = 0.7000000000000001
real_f1 = 0.7000000000000001
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, 247.51it/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: 0.6129480600357056
step: 10, loss: 0.6164372563362122
step: 20, loss: 0.32664555311203003
step: 30, loss: 0.1552729457616806
step: 40, loss: 0.16906501352787018
step: 50, loss: 0.013185825198888779
step: 60, loss: 0.05338568612933159
step: 70, loss: 0.0037449682131409645
step: 80, loss: 0.15201285481452942
step: 90, loss: 0.1732524335384369
step: 100, loss: 0.03540836647152901
step: 110, loss: 0.10136370360851288
step: 120, loss: 0.006931545212864876
step: 130, loss: 0.04817580804228783
step: 140, loss: 0.001923798001371324
step: 150, loss: 0.0271937046200037
step: 160, loss: 0.029683446511626244
step: 170, loss: 0.05306230112910271
step: 180, loss: 0.006574323400855064
step: 190, loss: 0.010956313461065292
step: 200, loss: 0.007669045124202967
step: 210, loss: 0.008926581591367722
step: 220, loss: 0.005398360546678305
```

```
step: 230, loss: 0.06152060627937317
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9785310734463276, 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: 0.0011685476638376713
step: 10, loss: 0.004369777627289295
step: 20, loss: 0.16219699382781982
step: 30, loss: 0.08487935364246368
step: 40, loss: 0.14753752946853638
step: 50, loss: 0.029245689511299133
step: 60, loss: 0.010968860238790512
step: 70, loss: 0.11326649785041809
step: 80, loss: 0.007630335167050362
step: 90, loss: 0.03313593566417694
step: 100, loss: 0.00923028215765953
step: 110, loss: 0.004168281797319651
step: 120, loss: 0.0014152684016153216
step: 130, loss: 0.002661618171259761
step: 140, loss: 0.0016179883386939764
step: 150, loss: 0.0019510422134771943
step: 160, loss: 0.0034490632824599743
step: 170, loss: 0.006373555865138769
step: 180, loss: 0.005702564027160406
step: 190, loss: 0.0022169414442032576
step: 200, loss: 0.0016152202151715755
step: 210, loss: 0.0006639313651248813
step: 220, loss: 0.16181422770023346
step: 230, loss: 0.0077910274267196655
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9820627802690582, f1=0.9876265466816648,
best_f1=0.9876265466816648
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
```

```
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0028785178437829018
step: 10, loss: 0.10924531519412994
step: 20, loss: 0.0061741722747683525
step: 30, loss: 0.0029270555824041367
step: 40, loss: 0.03736487030982971
step: 50, loss: 0.0032236487604677677
step: 60, loss: 0.043346308171749115
step: 70, loss: 0.015424869023263454
step: 80, loss: 0.0006029540672898293
step: 90, loss: 0.11173004657030106
step: 100, loss: 0.0005394406034611166
step: 110, loss: 0.0004622038686648011
step: 120, loss: 0.017711784690618515
step: 130, loss: 0.0032103881239891052
step: 140, loss: 0.0017654423136264086
step: 150, loss: 0.05687715485692024
step: 160, loss: 0.01746934838593006
step: 170, loss: 0.009114296175539494
step: 180, loss: 0.03579580783843994
step: 190, loss: 0.005258465651422739
step: 200, loss: 0.01607467792928219
step: 210, loss: 0.0007479360792785883
step: 220, loss: 0.0002392137685092166
step: 230, loss: 0.00016578781651332974
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9852104664391355, f1=0.9828962371721778,
best_f1=0.9828962371721778
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00013098052295390517
step: 10, loss: 0.00023024261463433504
step: 20, loss: 0.0027486865874379873
step: 30, loss: 0.0002581513545010239
step: 40, loss: 0.010658010840415955
step: 50, loss: 0.0002901765110436827
step: 60, loss: 0.008167145773768425
step: 70, loss: 0.004804644733667374
step: 80, loss: 0.0004742276796605438
step: 90, loss: 0.004083942621946335
step: 100, loss: 0.0170816108584404
```

```
step: 110, loss: 0.07844898104667664
step: 120, loss: 0.04641619324684143
step: 130, loss: 0.03801514953374863
step: 140, loss: 0.001162277883850038
step: 150, loss: 0.04164068400859833
step: 160, loss: 0.009668344631791115
step: 170, loss: 0.23692940175533295
step: 180, loss: 0.0009125762735493481
step: 190, loss: 0.007247967645525932
step: 200, loss: 0.0019906251691281796
step: 210, loss: 0.056968025863170624
step: 220, loss: 0.008193924091756344
step: 230, loss: 0.006771039683371782
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9887387387387387, f1=0.9887387387387387,
best_f1=0.9887387387387387
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.006741509307175875
step: 10, loss: 0.0009391444618813694
step: 20, loss: 0.021723467856645584
step: 30, loss: 0.00017584225861355662
step: 40, loss: 0.0002421043172944337
step: 50, loss: 0.03087044321000576
step: 60, loss: 0.0022133439779281616
step: 70, loss: 0.0001665465533733368
step: 80, loss: 0.000329523638356477
step: 90, loss: 0.00260086334310472
step: 100, loss: 0.00019178312504664063
step: 110, loss: 0.00017763876530807465
step: 120, loss: 6.526208744617179e-05
step: 130, loss: 0.002609874587506056
step: 140, loss: 0.005750559736043215
step: 150, loss: 0.0017541333800181746
step: 160, loss: 0.06245366856455803
step: 170, loss: 0.009733202867209911
step: 180, loss: 0.07123371958732605
step: 190, loss: 0.002595864934846759
step: 200, loss: 0.002962416736409068
step: 210, loss: 0.00022136291954666376
step: 220, loss: 0.0009557349840179086
```

```
step: 230, loss: 0.00021773016487713903
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
epoch 5: dev_f1=0.9854748603351955, f1=0.974585635359116,
best_f1=0.9887387387387387
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0002558814885560423
step: 10, loss: 0.0002920545230153948
step: 20, loss: 0.0003510198148433119
step: 30, loss: 0.00105836964212358
step: 40, loss: 0.00024107932404149324
step: 50, loss: 0.001251155394129455
step: 60, loss: 0.00019963351951446384
step: 70, loss: 0.000378980184905231
step: 80, loss: 0.0021712209563702345
step: 90, loss: 0.00022196049394551665
step: 100, loss: 0.03935758024454117
step: 110, loss: 0.0009394052322022617
step: 120, loss: 0.0006682639941573143
step: 130, loss: 0.00049036112613976
step: 140, loss: 0.00019835818966384977
step: 150, loss: 0.07656747847795486
step: 160, loss: 0.004094243515282869
step: 170, loss: 0.0020172367803752422
step: 180, loss: 0.06355635821819305
Gradient overflow.  Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 190, loss: 0.004403059370815754
step: 200, loss: 0.007131568621844053
step: 210, loss: 0.001589099527336657
step: 220, loss: 0.0002658675657585263
step: 230, loss: 0.07553112506866455
/usr/local/lib/python3.7/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=0.9852774631936579, f1=0.9817351598173515,
best_f1=0.9887387387387387
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00248400354757905
step: 10, loss: 0.00020578732073772699
step: 20, loss: 0.00035608879989013076
step: 30, loss: 0.002726082457229495
step: 40, loss: 0.0001729861251078546
step: 50, loss: 0.00013879344624001533
```

```
step: 60, loss: 0.0025746419560164213
step: 70, loss: 0.017724452540278435
step: 80, loss: 0.00014831338194198906
step: 90, loss: 0.00012509821681305766
step: 100, loss: 0.00029700473533011973
step: 110, loss: 0.0006341560510918498
step: 120, loss: 0.00016298222180921584
step: 130, loss: 0.00013615515490528196
step: 140, loss: 8.69927680469118e-05
step: 150, loss: 0.0037736683152616024
step: 160, loss: 0.0810680240392685
step: 170, loss: 0.0030180613975971937
step: 180, loss: 0.0005994331440888345
step: 190, loss: 8.51019358378835e-05
step: 200, loss: 0.06669601798057556
step: 210, loss: 7.620656833751127e-05
step: 220, loss: 0.0042756604962050915
Gradient overflow.  Skipping step, loss scaler 0 reducing loss scale to 8192.0
step: 230, loss: 9.83697609626688e-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 7: dev_f1=0.9864864864864865, f1=0.978865406006674,
best_f1=0.9887387387387387
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 6.666708941338584e-05
step: 10, loss: 0.00090344445392899215
step: 20, loss: 0.0006242335075512528
step: 30, loss: 0.0005041034892201424
step: 40, loss: 0.0005693771527148783
step: 50, loss: 0.0018503741594031453
step: 60, loss: 0.0007822254556231201
step: 70, loss: 0.001574943889863789
step: 80, loss: 0.00017015752382576466
step: 90, loss: 7.517704216297716e-05
step: 100, loss: 0.006251271348446608
step: 110, loss: 0.0011827743146568537
step: 120, loss: 0.0012094549601897597
step: 130, loss: 0.0007198106031864882
step: 140, loss: 8.196241105906665e-05
step: 150, loss: 0.004002324305474758
step: 160, loss: 0.0007407280500046909
step: 170, loss: 0.0001287563209189102
step: 180, loss: 0.0001340379094472155
step: 190, loss: 0.0023853082675486803
step: 200, loss: 0.00013591739116236567
```

```
step: 210, loss: 6.572485290234908e-05
step: 220, loss: 0.010236102156341076
step: 230, loss: 5.3678388212574646e-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=0.9864864864864865, f1=0.9842342342342343,
best_f1=0.9887387387387387
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 5.275326475384645e-05
step: 10, loss: 0.00036578814615495503
step: 20, loss: 0.012786281295120716
step: 30, loss: 0.0014605532633140683
step: 40, loss: 0.0011125912424176931
step: 50, loss: 6.569232937181368e-05
step: 60, loss: 0.00010711626964621246
step: 70, loss: 0.0004340771120041609
step: 80, loss: 6.337196100503206e-05
step: 90, loss: 0.00014434823242481798
step: 100, loss: 0.00023327101371251047
step: 110, loss: 4.573373371385969e-05
step: 120, loss: 3.5020188079215586e-05
step: 130, loss: 6.671228038612753e-05
step: 140, loss: 3.4211643651360646e-05
step: 150, loss: 8.197705028578639e-05
step: 160, loss: 3.124322756775655e-05
step: 170, loss: 5.8694738982012495e-05
step: 180, loss: 0.00038281604065559804
step: 190, loss: 0.0002983442391268909
step: 200, loss: 0.0006132787675596774
step: 210, loss: 6.699893128825352e-05
step: 220, loss: 4.4742941099684685e-05
step: 230, loss: 0.00036967245978303254
/usr/local/lib/python3.7/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=0.9842696629213483, f1=0.9819413092550789,
best_f1=0.9887387387387387
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00012033570965286344
step: 10, loss: 0.00014375070168171078
step: 20, loss: 5.0422546337358654e-05
step: 30, loss: 0.0006393414223566651
step: 40, loss: 0.0007961756782606244
```

```
step: 50, loss: 0.00012939225416630507
step: 60, loss: 0.00028480065520852804
step: 70, loss: 0.001225327025167644
step: 80, loss: 0.00015274618635885417
step: 90, loss: 0.0003698074142448604
step: 100, loss: 0.00011343155347276479
step: 110, loss: 0.001887200283817947
step: 120, loss: 0.0017830142751336098
step: 130, loss: 0.00013747303455602378
step: 140, loss: 0.0360754057765007
step: 150, loss: 0.015893181785941124
step: 160, loss: 0.00012472733214963228
step: 170, loss: 0.0003355394583195448
step: 180, loss: 0.00036963974707759917
step: 190, loss: 0.01551787555217743
step: 200, loss: 8.925216388888657e-05
step: 210, loss: 0.00012642827641684562
step: 220, loss: 4.116998752579093e-05
step: 230, loss: 0.0009538746089674532
/usr/local/lib/python3.7/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=0.9865771812080537, f1=0.9821826280623607,
best_f1=0.9887387387387387
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0002950991620309651
step: 10, loss: 0.00013800253509543836
step: 20, loss: 0.00017268139345105737
step: 30, loss: 0.012313229031860828
step: 40, loss: 0.00017711146210785955
step: 50, loss: 4.444513251655735e-05
step: 60, loss: 0.0003000679425895214
step: 70, loss: 0.0001643535797484219
step: 80, loss: 4.8284080548910424e-05
step: 90, loss: 4.077875200891867e-05
step: 100, loss: 0.0002599218569230288
step: 110, loss: 0.00047791359247639775
step: 120, loss: 0.0004761606687679887
step: 130, loss: 2.9611239369842224e-05
step: 140, loss: 5.9244874137220904e-05
step: 150, loss: 0.02222704328596592
step: 160, loss: 4.288787749828771e-05
step: 170, loss: 0.026512155309319496
step: 180, loss: 4.15050562878605e-05
step: 190, loss: 6.281135574681684e-05
step: 200, loss: 0.00042523175943642855
```

```
step: 210, loss: 3.8301652239169925e-05
step: 220, loss: 0.00019078679906670004
step: 230, loss: 8.032062760321423e-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=0.9864864864864865, f1=0.9853438556933484,
best_f1=0.9887387387387387
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 3.440574437263422e-05
step: 10, loss: 4.696740143117495e-05
step: 20, loss: 5.6666365708224475e-05
step: 30, loss: 6.914570258231834e-05
step: 40, loss: 0.0008729834808036685
step: 50, loss: 0.0001640952395973727
step: 60, loss: 9.410137863596901e-05
step: 70, loss: 9.393430082127452e-05
step: 80, loss: 0.0015896469121798873
step: 90, loss: 4.59654038422741e-05
step: 100, loss: 6.14917662460357e-05
step: 110, loss: 4.883571091340855e-05
step: 120, loss: 3.301996912341565e-05
step: 130, loss: 2.731331369432155e-05
step: 140, loss: 3.1283998396247625e-05
step: 150, loss: 0.0009082052856683731
step: 160, loss: 2.861332359316293e-05
step: 170, loss: 2.595732075860724e-05
step: 180, loss: 0.0349557101726532
step: 190, loss: 0.000962104182690382
step: 200, loss: 2.1975160052534193e-05
step: 210, loss: 8.297274325741455e-05
step: 220, loss: 0.0011812495067715645
step: 230, loss: 0.03092704340815544
/usr/local/lib/python3.7/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=0.9853438556933484, f1=0.9841628959276018,
best_f1=0.9887387387387387
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00045971147483214736
step: 10, loss: 0.0001880475174402818
step: 20, loss: 0.0005617537535727024
step: 30, loss: 0.0001562174438731745
step: 40, loss: 9.412295185029507e-05
```

```
step: 50, loss: 4.8878118832362816e-05
step: 60, loss: 4.3508527596713975e-05
step: 70, loss: 8.446219726465642e-05
step: 80, loss: 0.00015663560770917684
step: 90, loss: 8.578514825785533e-05
step: 100, loss: 2.3722312107565813e-05
step: 110, loss: 0.0008140738354995847
step: 120, loss: 0.04070204868912697
step: 130, loss: 8.269163663499057e-05
step: 140, loss: 2.6724706913228147e-05
step: 150, loss: 2.6232975869788788e-05
step: 160, loss: 0.00015330020687542856
step: 170, loss: 0.0003022726741619408
step: 180, loss: 3.5503926483215764e-05
step: 190, loss: 3.630357241490856e-05
step: 200, loss: 8.455938950646669e-05
step: 210, loss: 5.461333057610318e-05
step: 220, loss: 4.2892421333817765e-05
step: 230, loss: 5.6695400417083874e-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=0.9854748603351955, f1=0.9810901001112348,
best_f1=0.9887387387387387
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 2.322686486877501e-05
step: 10, loss: 0.013061783276498318
step: 20, loss: 8.429739682469517e-05
step: 30, loss: 0.00010184577769532688
step: 40, loss: 0.01513371430337429
step: 50, loss: 4.0468767110724e-05
step: 60, loss: 4.293223537388258e-05
step: 70, loss: 2.79984615190187747e-05
step: 80, loss: 2.3424265236826614e-05
step: 90, loss: 2.5655526769696735e-05
step: 100, loss: 5.488943861564621e-05
step: 110, loss: 0.0001547017745906487
step: 120, loss: 1.725537185848225e-05
step: 130, loss: 2.617310019559227e-05
step: 140, loss: 3.881918382830918e-05
step: 150, loss: 3.733609264600091e-05
step: 160, loss: 7.958233618410304e-05
step: 170, loss: 2.1233858205960132e-05
step: 180, loss: 7.145912240957841e-05
step: 190, loss: 3.38979261869099e-05
step: 200, loss: 2.7387093723518774e-05
```

```
step: 210, loss: 3.3112879464169964e-05
step: 220, loss: 3.67493303201627e-05
step: 230, loss: 3.390247002243996e-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=0.9854748603351955, f1=0.9821826280623607,
best_f1=0.9887387387387387
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00022572913439944386
step: 10, loss: 1.872681787062902e-05
step: 20, loss: 3.260559242335148e-05
step: 30, loss: 6.913207471370697e-05
step: 40, loss: 4.2857682274188846e-05
step: 50, loss: 0.03649698942899704
step: 60, loss: 3.4508895623730496e-05
step: 70, loss: 2.605038389447145e-05
step: 80, loss: 0.0002544522285461426
step: 90, loss: 4.6257020585471764e-05
step: 100, loss: 4.373519914224744e-05
step: 110, loss: 2.5510287741781212e-05
step: 120, loss: 0.00019948650151491165
step: 130, loss: 0.00015087032807059586
step: 140, loss: 2.123388730979059e-05
step: 150, loss: 5.553360460908152e-05
step: 160, loss: 0.0015202779322862625
step: 170, loss: 1.9862975022988394e-05
step: 180, loss: 0.0004251594073139131
step: 190, loss: 0.00020575380767695606
step: 200, loss: 0.0002528706390876323
step: 210, loss: 5.081156632513739e-05
step: 220, loss: 0.11070162057876587
step: 230, loss: 2.219121961388737e-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=0.9853438556933484, f1=0.9864559819413092,
best_f1=0.9887387387387387
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, 216.74it/s]
load_f1 = 0.9887387387387387
real_f1 = 0.9887387387387387
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, 270.88it/s]
```

4.1.3 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 \
```

```
  --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: 0.610503613948822
step: 10, loss: 0.5133370161056519
step: 20, loss: 0.5191926956176758
step: 30, loss: 0.15274904668331146
step: 40, loss: 0.17225342988967896
step: 50, loss: 0.1078958734869957
step: 60, loss: 0.04493679478764534
step: 70, loss: 0.11217907816171646
step: 80, loss: 0.0410870797932148
step: 90, loss: 0.1810123175382614
step: 100, loss: 0.06142173707485199
step: 110, loss: 0.10013003647327423
step: 120, loss: 0.160480335354805
step: 130, loss: 0.11524185538291931
step: 140, loss: 0.09780501574277878
step: 150, loss: 0.057915911078453064
step: 160, loss: 0.013442346826195717
step: 170, loss: 0.1795872300863266
step: 180, loss: 0.10308228433132172
step: 190, loss: 0.018255198374390602
step: 200, loss: 0.15011148154735565
step: 210, loss: 0.08137350529432297
step: 220, loss: 0.25193122029304504
step: 230, loss: 0.17151978611946106
step: 240, loss: 0.09350082278251648
step: 250, loss: 0.052684981375932693
step: 260, loss: 0.12932826578617096
step: 270, loss: 0.008883548900485039
step: 280, loss: 0.0187818706035614
step: 290, loss: 0.12822283804416656
step: 300, loss: 0.06919820606708527
step: 310, loss: 0.10264074057340622
step: 320, loss: 0.10067884624004364
step: 330, loss: 0.040871791541576385
step: 340, loss: 0.06658387184143066
step: 350, loss: 0.01963658072054386
step: 360, loss: 0.19140593707561493
step: 370, loss: 0.09772227704524994
step: 380, loss: 0.015114658512175083
```

```
step: 390, loss: 0.05633283033967018
step: 400, loss: 0.42842739820480347
step: 410, loss: 0.12962950766086578
step: 420, loss: 0.06588674336671829
step: 430, loss: 0.1675780862569809
step: 440, loss: 0.016874121502041817
step: 450, loss: 0.02654932253062725
step: 460, loss: 0.006655648350715637
step: 470, loss: 0.2128068506717682
step: 480, loss: 0.03238687664270401
step: 490, loss: 0.038842372596263885
step: 500, loss: 0.11561396718025208
step: 510, loss: 0.09984346479177475
step: 520, loss: 0.013371351175010204
step: 530, loss: 0.0033403998240828514
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.923581809657759, f1=0.9255121042830541,
best_f1=0.9255121042830541
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.30005723237991333
step: 10, loss: 0.13935601711273193
step: 20, loss: 0.007638311013579369
step: 30, loss: 0.057275883853435516
step: 40, loss: 0.048838842660188675
step: 50, loss: 0.09497016668319702
step: 60, loss: 0.007511339150369167
step: 70, loss: 0.054847799241542816
step: 80, loss: 0.08249576389789581
step: 90, loss: 0.010583187453448772
step: 100, loss: 0.08622851967811584
step: 110, loss: 0.0399857759475708
step: 120, loss: 0.02181442826986313
step: 130, loss: 0.1139468178153038
step: 140, loss: 0.02161935158073902
step: 150, loss: 0.13438571989536285
step: 160, loss: 0.03289637714624405
step: 170, loss: 0.06258971989154816
step: 180, loss: 0.02812214009463787
step: 190, loss: 0.0901850163936615
step: 200, loss: 0.0057438816875219345
```

```
step: 210, loss: 0.07711014896631241
step: 220, loss: 0.058500830084085464
step: 230, loss: 0.005646541249006987
step: 240, loss: 0.07752317190170288
step: 250, loss: 0.08817797899246216
step: 260, loss: 0.001997928600758314
step: 270, loss: 0.2567446529865265
step: 280, loss: 0.030375748872756958
step: 290, loss: 0.06761899590492249
step: 300, loss: 0.11467025429010391
step: 310, loss: 0.06977517157793045
step: 320, loss: 0.0645204707980156
step: 330, loss: 0.016385575756430626
step: 340, loss: 0.0193563774228096
step: 350, loss: 0.00084642076399177731
step: 360, loss: 0.09543373435735703
step: 370, loss: 0.09196992963552475
step: 380, loss: 0.015363642014563084
step: 390, loss: 0.0945584699511528
step: 400, loss: 0.12588335573673248
step: 410, loss: 0.006919359788298607
step: 420, loss: 0.01303092297166586
step: 430, loss: 0.03831516206264496
step: 440, loss: 0.0880156010389328
step: 450, loss: 0.0270888302475214
step: 460, loss: 0.09803749620914459
step: 470, loss: 0.01323692873120308
step: 480, loss: 0.23968788981437683
step: 490, loss: 0.02316877245903015
step: 500, loss: 0.14338062703609467
step: 510, loss: 0.0175817608833313
step: 520, loss: 0.04050430655479431
step: 530, loss: 0.010715288110077381
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9277496548550391, f1=0.9307201458523247,
best_f1=0.9307201458523247
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.09926071763038635
step: 10, loss: 0.09558805823326111
step: 20, loss: 0.05913970246911049
```

```
step: 30, loss: 0.04011042043566704
step: 40, loss: 0.009866690263152122
step: 50, loss: 0.15280908346176147
step: 60, loss: 0.027076341211795807
step: 70, loss: 0.020446274429559708
step: 80, loss: 0.01637258194386959
step: 90, loss: 0.006691658869385719
step: 100, loss: 0.07535185664892197
step: 110, loss: 0.0017318963073194027
step: 120, loss: 0.0027136618737131357
step: 130, loss: 0.0020642238669097424
step: 140, loss: 0.013946897350251675
step: 150, loss: 0.027158623561263084
step: 160, loss: 0.010423790663480759
step: 170, loss: 0.031957805156707764
step: 180, loss: 0.005401573143899441
step: 190, loss: 0.014441071078181267
step: 200, loss: 0.013609698042273521
step: 210, loss: 0.06970623135566711
step: 220, loss: 0.14406508207321167
step: 230, loss: 0.06476673483848572
step: 240, loss: 0.008252859115600586
step: 250, loss: 0.04121007025241852
step: 260, loss: 0.02120634913444519
step: 270, loss: 0.010002708993852139
step: 280, loss: 0.11615675687789917
step: 290, loss: 0.003306121099740267
step: 300, loss: 0.08944425731897354
step: 310, loss: 0.01642690598964691
step: 320, loss: 0.009373415261507034
step: 330, loss: 0.0016840090975165367
step: 340, loss: 0.0016109340358525515
step: 350, loss: 0.001034233719110489
step: 360, loss: 0.028206568211317062
step: 370, loss: 0.003963986411690712
step: 380, loss: 0.009544451721012592
step: 390, loss: 0.23494987189769745
step: 400, loss: 0.008544769138097763
step: 410, loss: 0.012250839732587337
step: 420, loss: 0.06709593534469604
step: 430, loss: 0.02223716862499714
step: 440, loss: 0.0019543995149433613
step: 450, loss: 0.06280330568552017
step: 460, loss: 0.050808388739824295
step: 470, loss: 0.12918104231357574
step: 480, loss: 0.0015998982125893235
step: 490, loss: 0.015601628459990025
step: 500, loss: 0.007358166854828596
```

```
step: 510, loss: 0.014703530818223953
step: 520, loss: 0.0787973552942276
step: 530, loss: 0.09008261561393738
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.936500685244404, f1=0.9304666056724611,
best_f1=0.9304666056724611
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.045320719480514526
step: 10, loss: 0.003132674377411604
step: 20, loss: 0.00586816156283021
step: 30, loss: 0.0032469211146235466
step: 40, loss: 0.015868308022618294
step: 50, loss: 0.001556600327603519
step: 60, loss: 0.003574649803340435
step: 70, loss: 0.0009568470995873213
step: 80, loss: 0.06594090908765793
step: 90, loss: 0.025680450722575188
step: 100, loss: 0.004249949473887682
step: 110, loss: 0.05205415561795235
step: 120, loss: 0.0004427181265782565
step: 130, loss: 0.04095574468374252
step: 140, loss: 0.0014880166854709387
step: 150, loss: 0.002998407930135727
step: 160, loss: 0.01143688429147005
step: 170, loss: 0.011313251219689846
step: 180, loss: 0.002677084179595113
step: 190, loss: 0.04108526185154915
step: 200, loss: 0.03802523389458656
step: 210, loss: 0.04740007594227791
step: 220, loss: 0.010440180078148842
step: 230, loss: 0.030254045501351357
step: 240, loss: 0.0025966926477849483
step: 250, loss: 0.06651268154382706
step: 260, loss: 0.0025235742796212435
step: 270, loss: 0.01655648835003376
step: 280, loss: 0.07172371447086334
step: 290, loss: 0.036969415843486786
step: 300, loss: 0.00045692690764553845
step: 310, loss: 0.00532285962253809
step: 320, loss: 0.004750229883939028
```

```
step: 330, loss: 0.006767400074750185
step: 340, loss: 0.26776331663131714
step: 350, loss: 0.1300056427717209
step: 360, loss: 0.02561924420297146
step: 370, loss: 0.023902643471956253
step: 380, loss: 0.003033742308616638
step: 390, loss: 0.002269561868160963
step: 400, loss: 0.0023461899254471064
step: 410, loss: 0.006270963232964277
step: 420, loss: 0.00515891844406724
step: 430, loss: 0.14563961327075958
step: 440, loss: 0.0032901461236178875
step: 450, loss: 0.10705357044935226
step: 460, loss: 0.011485881172120571
step: 470, loss: 0.021233266219496727
step: 480, loss: 0.14888519048690796
step: 490, loss: 0.002368211979046464
step: 500, loss: 0.011604542843997478
step: 510, loss: 0.031051374971866608
step: 520, loss: 0.08948426693677902
step: 530, loss: 0.031555209308862686
/usr/local/lib/python3.7/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=0.925536774783006, f1=0.9290617848970252,
best_f1=0.9304666056724611
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.1192721351981163
step: 10, loss: 0.027895662933588028
step: 20, loss: 0.0020004662219434977
step: 30, loss: 0.0075484635308384895
step: 40, loss: 0.09350413829088211
step: 50, loss: 0.007149518933147192
step: 60, loss: 0.020780544728040695
Gradient overflow.  Skipping step, loss scaler 0 reducing loss scale to 32768.0
step: 70, loss: 0.0020284096244722605
step: 80, loss: 0.09375129640102386
step: 90, loss: 0.07456426322460175
step: 100, loss: 0.12845484912395477
step: 110, loss: 0.0027245976962149143
step: 120, loss: 0.0013673766516149044
step: 130, loss: 0.0006373588694259524
step: 140, loss: 0.0013454948784783483
step: 150, loss: 0.004789868835359812
step: 160, loss: 0.001233322429470718
step: 170, loss: 0.011790556833148003
```

```
step: 180, loss: 0.0010827212827280164
step: 190, loss: 0.00303015043027699
step: 200, loss: 0.0030978675931692123
step: 210, loss: 0.004423650447279215
step: 220, loss: 0.0039110430516302586
step: 230, loss: 0.004150616005063057
step: 240, loss: 0.004779924638569355
step: 250, loss: 0.010385622270405293
step: 260, loss: 0.04883425310254097
step: 270, loss: 0.005957243964076042
step: 280, loss: 0.006114489398896694
step: 290, loss: 0.12846258282661438
step: 300, loss: 0.010428817942738533
step: 310, loss: 0.012345829047262669
step: 320, loss: 0.03673657029867172
step: 330, loss: 0.15718936920166016
step: 340, loss: 0.0036320341750979424
step: 350, loss: 0.010610869154334068
step: 360, loss: 0.008687203750014305
step: 370, loss: 0.02220001071691513
step: 380, loss: 0.0013369549997150898
step: 390, loss: 0.0003535302821546793
step: 400, loss: 0.010361971333622932
step: 410, loss: 0.005340441595762968
step: 420, loss: 0.002082232153043151
step: 430, loss: 0.0021908003836870193
step: 440, loss: 0.029834670946002007
step: 450, loss: 0.007847714237868786
step: 460, loss: 0.030557438731193542
step: 470, loss: 0.00475190719589591
step: 480, loss: 0.009318063966929913
step: 490, loss: 0.0003755018988158554
step: 500, loss: 0.005117273889482021
step: 510, loss: 0.06543602049350739
step: 520, loss: 0.008771431632339954
step: 530, loss: 0.011462584137916565
/usr/local/lib/python3.7/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=0.9340560072267391, f1=0.9327313769751693,
best_f1=0.9304666056724611
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0041847811080515385
step: 10, loss: 0.0005497150705195963
step: 20, loss: 0.10806167125701904
step: 30, loss: 0.018724050372838974
```

```
step: 40, loss: 0.02630424313247204
step: 50, loss: 0.022926220670342445
step: 60, loss: 0.00011009911395376548
step: 70, loss: 0.00041143110138364136
step: 80, loss: 0.0022365404292941093
step: 90, loss: 0.00033686120877973735
step: 100, loss: 0.0016404340276494622
step: 110, loss: 0.006618231534957886
step: 120, loss: 0.0015861217398196459
step: 130, loss: 0.1305711716413498
step: 140, loss: 0.008639681152999401
step: 150, loss: 0.0010859051253646612
step: 160, loss: 0.025749368593096733
step: 170, loss: 0.038736168295145035
step: 180, loss: 0.0011861090315505862
step: 190, loss: 0.002648895373567939
step: 200, loss: 0.00642079720273613
step: 210, loss: 0.001223512808792293
step: 220, loss: 0.0016100095817819238
step: 230, loss: 0.008886614814400673
step: 240, loss: 0.11846088618040085
step: 250, loss: 0.013008128851652145
step: 260, loss: 0.0023091845214366913
step: 270, loss: 0.10046568512916565
step: 280, loss: 0.015068857930600643
step: 290, loss: 0.0007719467976130545
step: 300, loss: 0.010983791202306747
step: 310, loss: 0.0006556739681400359
step: 320, loss: 0.0006958473240956664
step: 330, loss: 0.0015983961056917906
step: 340, loss: 0.04329284280538559
step: 350, loss: 0.0017910415772348642
step: 360, loss: 0.03427240997552872
step: 370, loss: 0.002020325278863311
step: 380, loss: 0.0043379939161241055
step: 390, loss: 0.025003371760249138
Gradient overflow.  Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 400, loss: 0.0375661700963974
step: 410, loss: 0.036343675510676384
step: 420, loss: 0.012891828082501888
step: 430, loss: 0.0028566892724484205
step: 440, loss: 0.0020300017204135656
step: 450, loss: 0.00039626439684070647
step: 460, loss: 0.007493241224437952
step: 470, loss: 0.0010043180081993341
step: 480, loss: 0.008845541626214981
step: 490, loss: 0.009296414442360401
step: 500, loss: 0.0029008558485656977
```

```
step: 510, loss: 0.001974373124539852
step: 520, loss: 0.006400167476385832
step: 530, loss: 0.015292024239897728
/usr/local/lib/python3.7/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=0.9359095193213949, f1=0.9337121212121212,
best_f1=0.9304666056724611
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0003957127337343991
step: 10, loss: 0.01699225977063179
step: 20, loss: 0.00044140100362710655
step: 30, loss: 0.0007311057997867465
step: 40, loss: 0.013781661167740822
step: 50, loss: 0.004940458573400974
step: 60, loss: 0.14030839502811432
step: 70, loss: 0.0027019174303859472
step: 80, loss: 0.004285012371838093
step: 90, loss: 0.008354111574590206
step: 100, loss: 0.0001610380713827908
step: 110, loss: 0.007576220668852329
step: 120, loss: 0.02370925433933735
step: 130, loss: 0.004108133260160685
step: 140, loss: 0.01956627145409584
step: 150, loss: 0.0006248832796700299
step: 160, loss: 0.0034840116277337074
step: 170, loss: 0.0008582363370805979
step: 180, loss: 0.0013437013840302825
step: 190, loss: 0.027645187452435493
step: 200, loss: 0.0031699896790087223
step: 210, loss: 0.0029272078536450863
step: 220, loss: 0.00030831526964900288
step: 230, loss: 0.1052895113825798
step: 240, loss: 0.0009328825399279594
step: 250, loss: 0.003836473450064659
step: 260, loss: 0.0035534368362277746
step: 270, loss: 0.01220837701112032
step: 280, loss: 0.0016179250087589025
step: 290, loss: 0.017151040956377983
step: 300, loss: 0.00871141254901886
step: 310, loss: 0.00038617930840700865
step: 320, loss: 0.012995151802897453
step: 330, loss: 0.0008026838768273592
step: 340, loss: 0.03994528576731682
step: 350, loss: 0.011931153945624828
step: 360, loss: 0.012415806762874126
```

```
step: 370, loss: 0.0036370926536619663
step: 380, loss: 0.002756512491032481
step: 390, loss: 0.00020471794414334
step: 400, loss: 0.004026662092655897
step: 410, loss: 0.0012555239954963326
step: 420, loss: 0.0008611969533376396
step: 430, loss: 4.8705707740737125e-05
step: 440, loss: 0.0001907769328681752
step: 450, loss: 0.004034065175801516
step: 460, loss: 0.0005537194665521383
step: 470, loss: 0.00024370540631935
step: 480, loss: 0.18595433235168457
step: 490, loss: 0.0002002816036110744
step: 500, loss: 0.008412383496761322
step: 510, loss: 0.02913450449705124
step: 520, loss: 0.043696265667676926
step: 530, loss: 0.0817769467830658
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9416160672582905, f1=0.9410672853828307,
best_f1=0.9410672853828307
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0006806849851273
step: 10, loss: 0.02465950883924961
step: 20, loss: 0.002414637478068471
step: 30, loss: 0.02277946285903454
step: 40, loss: 0.0031417780555784702
step: 50, loss: 0.008416897617280483
step: 60, loss: 0.003408523043617606
step: 70, loss: 0.0013970505679026246
step: 80, loss: 0.0007706787437200546
step: 90, loss: 0.0013544528046622872
step: 100, loss: 0.001567809609696269
step: 110, loss: 0.005124587099999189
step: 120, loss: 0.004316254984587431
step: 130, loss: 6.255495827645063e-05
step: 140, loss: 0.009484579786658287
step: 150, loss: 0.000516602594871074
step: 160, loss: 0.008523056283593178
step: 170, loss: 0.0019039774779230356
step: 180, loss: 0.00015310100570786744
```

```
step: 190, loss: 0.0037335336674004793
step: 200, loss: 0.010211721993982792
step: 210, loss: 0.0034543343354016542
step: 220, loss: 0.0009067032369785011
step: 230, loss: 0.012852865271270275
step: 240, loss: 0.0001477680925745517
step: 250, loss: 0.00016019956092350185
step: 260, loss: 0.00020893197506666183
step: 270, loss: 0.01195079181343317
step: 280, loss: 0.05105523392558098
step: 290, loss: 0.049139413982629776
step: 300, loss: 0.0012899310095235705
step: 310, loss: 0.004795127548277378
step: 320, loss: 0.008787584491074085
step: 330, loss: 9.627056715544313e-05
step: 340, loss: 0.0001547357824165374
step: 350, loss: 0.0002189380320487544
step: 360, loss: 0.00015695983893238008
step: 370, loss: 0.0002121177240042016
step: 380, loss: 0.000940389814786613
step: 390, loss: 0.1301889270544052
step: 400, loss: 0.0027194342110306025
step: 410, loss: 0.00028075711452402174
step: 420, loss: 0.00868570152670145
step: 430, loss: 0.03164157271385193
step: 440, loss: 0.004370164126157761
step: 450, loss: 0.0027325875125825405
step: 460, loss: 0.00034495574072934687
step: 470, loss: 0.00016697148384992033
step: 480, loss: 8.107382018351927e-05
step: 490, loss: 0.004424704238772392
step: 500, loss: 0.0011482387781143188
step: 510, loss: 0.0047795879654586315
step: 520, loss: 0.001382104237563908
step: 530, loss: 0.0002423187397653237
/usr/local/lib/python3.7/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=0.9367441860465117, f1=0.9327188940092166,
best_f1=0.9410672853828307
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00855409074574709
step: 10, loss: 0.0006976196891628206
step: 20, loss: 0.015208898112177849
step: 30, loss: 0.0029592206701636314
step: 40, loss: 0.0014437478967010975
```

```
step: 50, loss: 6.762891280231997e-05
step: 60, loss: 0.00018353559426032007
step: 70, loss: 0.0005483317654579878
step: 80, loss: 0.004308048635721207
step: 90, loss: 0.0008138769771903753
step: 100, loss: 8.625458576716483e-05
step: 110, loss: 0.00030215445440262556
step: 120, loss: 0.009872390888631344
step: 130, loss: 0.03553760051727295
step: 140, loss: 0.00608940189704299
step: 150, loss: 0.00026457267813384533
step: 160, loss: 0.0008426682325080037
step: 170, loss: 0.01917913369834423
step: 180, loss: 0.000282236491329968
step: 190, loss: 0.0018841971177607775
step: 200, loss: 0.0004307385243009776
step: 210, loss: 0.0006785893929190934
step: 220, loss: 0.0008367258124053478
step: 230, loss: 0.0004115664341952652
step: 240, loss: 0.020238002762198448
step: 250, loss: 0.00028771156212314963
step: 260, loss: 0.01920437254011631
step: 270, loss: 5.661423710989766e-05
step: 280, loss: 0.0066802348010241985
step: 290, loss: 0.03504167124629021
step: 300, loss: 0.00015168650134000927
step: 310, loss: 0.03102644346654415
step: 320, loss: 0.0017627017805352807
step: 330, loss: 0.00043056669528596103
step: 340, loss: 0.000525047245901078
step: 350, loss: 0.012145180255174637
step: 360, loss: 0.0010002244962379336
step: 370, loss: 0.006222480442374945
step: 380, loss: 0.011485186405479908
step: 390, loss: 0.00042187259532511234
step: 400, loss: 0.007973087020218372
step: 410, loss: 0.0018168498063459992
step: 420, loss: 0.0004142228572163731
step: 430, loss: 0.00026987068122225282
step: 440, loss: 0.002719239331781864
step: 450, loss: 6.567795207956806e-05
step: 460, loss: 8.069151954259723e-05
step: 470, loss: 0.00021622792701236904
step: 480, loss: 0.00011392317537683994
step: 490, loss: 0.010665440000593662
step: 500, loss: 0.001448773662559688
step: 510, loss: 0.0005994226085022092
step: 520, loss: 0.00064883014420047
```

```
step: 530, loss: 0.00015773599443491548
/usr/local/lib/python3.7/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=0.9351981351981352, f1=0.9381682938168294,
best_f1=0.9410672853828307
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0001373708073515445
step: 10, loss: 0.00011393886961741373
step: 20, loss: 2.036209662037436e-05
step: 30, loss: 2.3714197595836595e-05
step: 40, loss: 3.357845343998633e-05
step: 50, loss: 0.0009608329855836928
step: 60, loss: 0.00034207795397378504
step: 70, loss: 0.0013028548564761877
step: 80, loss: 1.82350559043698e-05
step: 90, loss: 0.0018969714874401689
step: 100, loss: 0.00020835809118580073
step: 110, loss: 9.34943527681753e-05
step: 120, loss: 0.018621142953634262
step: 130, loss: 5.459013846120797e-05
step: 140, loss: 0.027536986395716667
step: 150, loss: 4.973976319888607e-05
step: 160, loss: 0.00013711965584661812
step: 170, loss: 0.00012379713007248938
step: 180, loss: 0.00040189025457948446
step: 190, loss: 0.000104315651697106666
step: 200, loss: 0.0005989056662656367
step: 210, loss: 0.0002645437198225409
step: 220, loss: 0.0009662683005444705
step: 230, loss: 0.006473179906606674
step: 240, loss: 0.0005030918982811272
step: 250, loss: 4.5470904296962544e-05
step: 260, loss: 0.0012163524515926838
step: 270, loss: 0.0007315715774893761
step: 280, loss: 0.0002702478086575866
step: 290, loss: 0.0034253187477588654
step: 300, loss: 2.3833687009755522e-05
step: 310, loss: 0.00010210351319983602
step: 320, loss: 0.0007069007260724902
step: 330, loss: 0.0004676894750446081
step: 340, loss: 0.005924759898334742
step: 350, loss: 0.001957367639988661
step: 360, loss: 6.199204653967172e-05
step: 370, loss: 0.004383268300443888
step: 380, loss: 0.006035674829035997
```

```
step: 390, loss: 0.0002039099927060306
step: 400, loss: 0.0007931645377539098
step: 410, loss: 0.0025876429863274097
step: 420, loss: 0.0008610851364210248
step: 430, loss: 0.00010382203618064523
step: 440, loss: 0.004897471982985735
step: 450, loss: 0.0006870211800560355
step: 460, loss: 5.86827372899279e-05
step: 470, loss: 0.006668186280876398
step: 480, loss: 7.346231723204255e-05
step: 490, loss: 6.241419760044664e-05
step: 500, loss: 0.039254095405340195
step: 510, loss: 0.00031355739338323474
step: 520, loss: 0.0007532525341957808
step: 530, loss: 0.0010493311565369368
/usr/local/lib/python3.7/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=0.9388322520852641, f1=0.9390187987161852,
best_f1=0.9410672853828307
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.015469513833522797
step: 10, loss: 0.07266353815793991
step: 20, loss: 3.4558343031676486e-05
step: 30, loss: 0.00012147557572461665
step: 40, loss: 0.0012485446641221642
step: 50, loss: 0.006260852795094252
step: 60, loss: 0.0405375137925148
step: 70, loss: 2.0380693968036212e-05
step: 80, loss: 0.015124638564884663
step: 90, loss: 1.8022685253527015e-05
step: 100, loss: 0.0010347057832404971
step: 110, loss: 0.0002802898525260389
step: 120, loss: 0.00020769699767697603
step: 130, loss: 3.2721120078349486e-05
step: 140, loss: 0.03850885108113289
step: 150, loss: 2.7264555683359504e-05
step: 160, loss: 2.6050056476378813e-05
step: 170, loss: 7.520960207330063e-05
step: 180, loss: 0.0033596705179661512
step: 190, loss: 0.003043954726308584
step: 200, loss: 0.026553137227892876
step: 210, loss: 0.0018347513396292925
step: 220, loss: 0.002088319743052125
step: 230, loss: 0.00016847610822878778
step: 240, loss: 0.0008802128140814602
```

```
step: 250, loss: 6.198185292305425e-05
step: 260, loss: 2.463024247845169e-05
step: 270, loss: 0.0007985724951140583
step: 280, loss: 0.0002863410336431116
step: 290, loss: 0.0007292009540833533
step: 300, loss: 0.003626954508945346
step: 310, loss: 0.00020912091713398695
step: 320, loss: 0.0001427463284926489
step: 330, loss: 0.0014626566553488374
step: 340, loss: 0.003470734693109989
step: 350, loss: 0.004610392730683088
step: 360, loss: 0.000103618367575109
step: 370, loss: 0.0025114258751273155
step: 380, loss: 0.00017417134949937463
step: 390, loss: 0.00021213329455349594
step: 400, loss: 0.00022041617194190621
step: 410, loss: 0.0006245353724807501
step: 420, loss: 0.0007297425763681531
step: 430, loss: 2.9150785849196836e-05
step: 440, loss: 0.0010141184320673347
step: 450, loss: 0.10582751035690308
step: 460, loss: 0.005009902641177177
step: 470, loss: 0.0008133137016557157
step: 480, loss: 0.00786003842949867
step: 490, loss: 0.0001287082995986566
step: 500, loss: 0.0012502563185989857
step: 510, loss: 0.0001279124553548172
step: 520, loss: 0.00012541604519356042
step: 530, loss: 5.89154806220904e-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=0.9340866290018832, f1=0.9381682938168294,
best_f1=0.9410672853828307
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.003550513181835413
step: 10, loss: 4.825737050850876e-05
step: 20, loss: 0.00023446526029147208
step: 30, loss: 0.0068142954260110855
step: 40, loss: 0.00038212293293327093
step: 50, loss: 0.06270446628332138
step: 60, loss: 0.00800444558262825
step: 70, loss: 0.003821160877123475
step: 80, loss: 9.748512093210593e-05
step: 90, loss: 0.0007067913538776338
step: 100, loss: 0.00014993610966484994
```

```
step: 110, loss: 0.0007222354179248214
step: 120, loss: 3.73960574506782e-05
step: 130, loss: 0.0015453004743903875
step: 140, loss: 1.5545392670901492e-05
step: 150, loss: 2.8247115551494062e-05
step: 160, loss: 2.2417847503675148e-05
step: 170, loss: 0.00011809682473540306
step: 180, loss: 8.02850117906928e-05
step: 190, loss: 4.341865496826358e-05
step: 200, loss: 0.00011973790969932452
step: 210, loss: 5.06072465213947e-05
step: 220, loss: 0.00011546807218110189
step: 230, loss: 0.00014018082583788782
step: 240, loss: 0.0026415616739541292
step: 250, loss: 0.0002591164957266301
step: 260, loss: 0.0002130954817403108
step: 270, loss: 0.0005556629039347172
step: 280, loss: 0.0015779782552272081
step: 290, loss: 0.016281861811876297
step: 300, loss: 0.004118530545383692
step: 310, loss: 0.00024097852292470634
step: 320, loss: 0.0010297310072928667
step: 330, loss: 0.0001672244688961655
step: 340, loss: 0.0013822235632687807
step: 350, loss: 0.0021967850625514984
step: 360, loss: 0.048398178070783615
step: 370, loss: 0.0004328599898144603
step: 380, loss: 0.00018718958017416298
step: 390, loss: 0.00023780213086865842
step: 400, loss: 0.0005402190727181733
step: 410, loss: 0.0005297018215060234
step: 420, loss: 0.01205623708665371
step: 430, loss: 0.016535021364688873
step: 440, loss: 0.0003179094346705824
step: 450, loss: 0.00511898985132575
step: 460, loss: 8.255388820543885e-05
step: 470, loss: 0.001079405308701098
step: 480, loss: 0.0005962061113677919
step: 490, loss: 0.016712335869669914
step: 500, loss: 0.0003174935991410166
step: 510, loss: 0.009160139597952366
step: 520, loss: 1.78100945049664e-05
step: 530, loss: 7.91751008364372e-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=0.937471051412691, f1=0.9344337459880789,
best_f1=0.9410672853828307
```

```
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0008604431641288102
step: 10, loss: 6.24709646217525e-05
step: 20, loss: 7.877442112658173e-05
step: 30, loss: 3.6106757761444896e-05
step: 40, loss: 7.344319601543248e-05
step: 50, loss: 0.0005476018413901329
step: 60, loss: 0.00021891944925300777
step: 70, loss: 0.0001856315357144922
step: 80, loss: 0.00011624014587141573
step: 90, loss: 0.060647834092378616
step: 100, loss: 0.001855025184340775
step: 110, loss: 0.0005423896363936365
step: 120, loss: 0.0012875188840553164
step: 130, loss: 0.0005853504408150911
step: 140, loss: 0.0002172253152821213
step: 150, loss: 0.000994945876300335
step: 160, loss: 0.00021737781935371459
step: 170, loss: 0.0002714448783081025
step: 180, loss: 6.861505244160071e-05
step: 190, loss: 0.00013885294902138412
step: 200, loss: 0.00015070191875565797
step: 210, loss: 0.00010645046131685376
step: 220, loss: 0.0001023536387947388
step: 230, loss: 0.0002508679172024131
step: 240, loss: 0.00012145499931648374
step: 250, loss: 0.00027472045621834695
step: 260, loss: 0.00010317134001525119
step: 270, loss: 3.760988329304382e-05
step: 280, loss: 0.0010872103739529848
step: 290, loss: 0.009410357102751732
step: 300, loss: 5.2268187573645264e-05
step: 310, loss: 0.00245875958353281
step: 320, loss: 0.002683911705389619
step: 330, loss: 0.06634005159139633
step: 340, loss: 0.0026582274585962296
step: 350, loss: 0.029537446796894073
step: 360, loss: 0.0003236875054426491
step: 370, loss: 6.57396376482211e-05
step: 380, loss: 0.00017764537187758833
step: 390, loss: 0.025557009503245354
step: 400, loss: 0.0001311750093009323
step: 410, loss: 0.0032046562992036343
step: 420, loss: 6.991325062699616e-05
step: 430, loss: 0.0009877632837742567
step: 440, loss: 0.005377490073442459
```

```
step: 450, loss: 0.0003016888804268092
step: 460, loss: 0.0382474884390831
step: 470, loss: 9.925539052346721e-05
step: 480, loss: 0.0003720140957739204
step: 490, loss: 6.602580106118694e-05
step: 500, loss: 0.00032295286655426025
step: 510, loss: 0.0002315865276614204
step: 520, loss: 0.014219569973647594
step: 530, loss: 0.00044569108285941184
/usr/local/lib/python3.7/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=0.9382366808109384, f1=0.9403054141601109,
best_f1=0.9410672853828307
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0005018578376621008
step: 10, loss: 2.051102273981087e-05
step: 20, loss: 8.038866508286446e-05
step: 30, loss: 3.616922549554147e-05
step: 40, loss: 0.00094774551168962479
step: 50, loss: 0.0009504980989731848
step: 60, loss: 1.605575562280137e-05
step: 70, loss: 2.349770329601597e-05
step: 80, loss: 1.4148322406981606e-05
step: 90, loss: 0.00026490443269722164
step: 100, loss: 0.002858429914340377
step: 110, loss: 0.0020121638663113117
step: 120, loss: 6.192264845594764e-05
step: 130, loss: 0.010488811880350113
step: 140, loss: 0.0035993142519146204
step: 150, loss: 1.2010238606308121e-05
step: 160, loss: 0.0010231437627226114
step: 170, loss: 0.0006957991281524301
step: 180, loss: 4.594152414938435e-05
step: 190, loss: 0.0002226904034614563
step: 200, loss: 0.000392561691114679
step: 210, loss: 1.773205258359667e-05
step: 220, loss: 3.253794420743361e-05
step: 230, loss: 0.0003508463269099593
step: 240, loss: 0.0011217637220397592
step: 250, loss: 0.000640456157270819
step: 260, loss: 0.0024902098812162876
step: 270, loss: 0.00012897416308987886
step: 280, loss: 3.778674727072939e-05
step: 290, loss: 0.001383058843202889
step: 300, loss: 0.01813006028532982
```

```
step: 310, loss: 0.00020564829173963517
step: 320, loss: 0.0010822630720213056
step: 330, loss: 0.0022181656677275896
step: 340, loss: 0.0004861502966377884
step: 350, loss: 0.001812883885577321
step: 360, loss: 0.0006182476645335555
step: 370, loss: 0.00017290026880800724
step: 380, loss: 0.0007755705737508833
step: 390, loss: 0.032598916441202164
step: 400, loss: 0.0009209130657836795
step: 410, loss: 4.7371107939397916e-05
step: 420, loss: 0.0004096115881111473
step: 430, loss: 0.11574704945087433
step: 440, loss: 0.0005880396929569542
step: 450, loss: 9.063348261406645e-05
step: 460, loss: 0.0006248411955311894
step: 470, loss: 0.00011608241766225547
step: 480, loss: 3.4243155823787674e-05
step: 490, loss: 2.079001205856912e-05
step: 500, loss: 2.9132585041224957e-05
step: 510, loss: 0.01812981441617012
step: 520, loss: 0.0006239743670448661
step: 530, loss: 0.0020317137241363525
/usr/local/lib/python3.7/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=0.9374714742126883,
best_f1=0.9410672853828307
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 2.36878713362199076e-05
step: 10, loss: 9.793722711037844e-06
step: 20, loss: 0.0035652273800224066
step: 30, loss: 9.757017687661573e-05
step: 40, loss: 0.08668258786201477
step: 50, loss: 0.001069393940269947
step: 60, loss: 3.2595267839496955e-05
step: 70, loss: 0.0016747487243264914
step: 80, loss: 2.2838317818241194e-05
step: 90, loss: 4.5620025048265234e-05
step: 100, loss: 0.0011002501705661416
step: 110, loss: 0.0004476162139326334
step: 120, loss: 0.00044934035395272076
step: 130, loss: 0.12642204761505127
step: 140, loss: 7.788067159708589e-05
step: 150, loss: 3.177365579176694e-05
step: 160, loss: 9.272013994632289e-05
```

```
step: 170, loss: 0.00013800141459796578
step: 180, loss: 6.67304775561206e-05
step: 190, loss: 0.00039648450911045074
step: 200, loss: 0.00015986972721293569
Gradient overflow.  Skipping step, loss scaler 0 reducing loss scale to 32768.0
step: 210, loss: 0.0031168549321591854
step: 220, loss: 2.4074501197901554e-05
step: 230, loss: 0.0003840427962131798
step: 240, loss: 0.019150052219629288
step: 250, loss: 7.884045771788806e-05
step: 260, loss: 3.446342088864185e-05
step: 270, loss: 1.4863692740618717e-05
step: 280, loss: 2.2548380002263002e-05
step: 290, loss: 1.5790981706231833e-05
step: 300, loss: 2.9810085834469646e-05
step: 310, loss: 0.0008435420459136367
step: 320, loss: 0.00037344160955399275
step: 330, loss: 2.149416832253337e-05
step: 340, loss: 5.9247337048873305e-05
step: 350, loss: 3.37108394887764e-05
step: 360, loss: 0.001640145666897297
step: 370, loss: 2.6264480766258202e-05
step: 380, loss: 0.0004390160320326686
step: 390, loss: 0.00027898509870283306
step: 400, loss: 0.00277429330162704
step: 410, loss: 7.412367995129898e-05
step: 420, loss: 0.00016857474111020565
step: 430, loss: 0.00018515651754569262
step: 440, loss: 0.00013943700469098985
step: 450, loss: 1.2598787179740611e-05
step: 460, loss: 0.00015346244617830962
step: 470, loss: 0.0015097020659595728
step: 480, loss: 1.170104314951459e-05
step: 490, loss: 0.00011029809684259817
step: 500, loss: 0.0012593235587701201
step: 510, loss: 6.335590296657756e-05
step: 520, loss: 0.0027368913870304823
step: 530, loss: 0.0005124799208715558
/usr/local/lib/python3.7/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=0.9385052034058656, f1=0.9398040130657956,
best_f1=0.9410672853828307
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 [00:20, 277.91it/s]
load_f1 = 0.9402985074626865
real_f1 = 0.9392523364485983
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, 269.43it/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: 0.5639081001281738
step: 10, loss: 0.3706364929676056
step: 20, loss: 0.382394015789032
step: 30, loss: 0.3135828673839569
step: 40, loss: 0.16647598147392273
step: 50, loss: 0.41640153527259827
step: 60, loss: 0.20282821357250214
step: 70, loss: 0.14026296138763428
step: 80, loss: 0.19829455018043518
step: 90, loss: 0.36596640944480896
step: 100, loss: 0.48103880882263184
step: 110, loss: 0.26298466324806213
step: 120, loss: 0.17814761400222778
step: 130, loss: 0.20168043673038483
step: 140, loss: 0.24380607903003693
step: 150, loss: 0.15448272228240967
step: 160, loss: 0.2299535870552063
step: 170, loss: 0.2775481045246124
step: 180, loss: 0.15774378180503845
step: 190, loss: 0.18518295884132385
step: 200, loss: 0.3123333156108856
step: 210, loss: 0.2671462595462799
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.6552462526766595, f1=0.6361746361746361,
best_f1=0.6361746361746361
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.11002712696790695
step: 10, loss: 0.1739218533039093

```
step: 20, loss: 0.2633654773235321
step: 30, loss: 0.1931627094745636
step: 40, loss: 0.15869595110416412
step: 50, loss: 0.3297938406467438
step: 60, loss: 0.4221328794956207
step: 70, loss: 0.11352359503507614
step: 80, loss: 0.2342834174633026
step: 90, loss: 0.13326817750930786
step: 100, loss: 0.006262835115194321
step: 110, loss: 0.10388631373643875
step: 120, loss: 0.1705896258354187
step: 130, loss: 0.026996392756700516
step: 140, loss: 0.1408982276916504
step: 150, loss: 0.15276040136814117
step: 160, loss: 0.2145272195339203
step: 170, loss: 0.16478994488716125
step: 180, loss: 0.22769197821617126
step: 190, loss: 0.1844366490840912
step: 200, loss: 0.043303146958351135
step: 210, loss: 0.09861031919717789
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.6950959488272921, f1=0.6480186480186481,
best_f1=0.6480186480186481
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.05230163410305977
step: 10, loss: 0.20368574559688568
step: 20, loss: 0.2941474914550781
step: 30, loss: 0.18947824835777283
step: 40, loss: 0.10861322283744812
step: 50, loss: 0.11614259332418442
step: 60, loss: 0.338554322719574
step: 70, loss: 0.11414992064237595
step: 80, loss: 0.12003540247678757
step: 90, loss: 0.09511338919401169
step: 100, loss: 0.2380192130804062
step: 110, loss: 0.13337327539920807
step: 120, loss: 0.09438443183898926
step: 130, loss: 0.12461186945438385
step: 140, loss: 0.037173885852098465
step: 150, loss: 0.3189842998981476
```

```
step: 160, loss: 0.03764372691512108
step: 170, loss: 0.12507127225399017
step: 180, loss: 0.12439028918743134
step: 190, loss: 0.17144152522087097
step: 200, loss: 0.0704815462231636
step: 210, loss: 0.13636904954910278
/usr/local/lib/python3.7/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=0.65625, f1=0.6695842450765865, best_f1=0.6480186480186481
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.09334680438041687
step: 10, loss: 0.10783510655164719
step: 20, loss: 0.1789252609014511
step: 30, loss: 0.22816236317157745
step: 40, loss: 0.06507208198308945
step: 50, loss: 0.1773841679096222
step: 60, loss: 0.10016200691461563
step: 70, loss: 0.09227555245161057
step: 80, loss: 0.14868302643299103
step: 90, loss: 0.026010330766439438
step: 100, loss: 0.2377583235502243
step: 110, loss: 0.11033304035663605
step: 120, loss: 0.1155468076467514
step: 130, loss: 0.3334611654281616
step: 140, loss: 0.13075359165668488
step: 150, loss: 0.1546981930732727
step: 160, loss: 0.10166604071855545
step: 170, loss: 0.08882159739732742
step: 180, loss: 0.3878730833530426
step: 190, loss: 0.09949342161417007
step: 200, loss: 0.2679075300693512
step: 210, loss: 0.2039969563484192
/usr/local/lib/python3.7/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=0.6798418972332015, f1=0.6614481409001958,
best_f1=0.6480186480186481
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.15658758580684662
step: 10, loss: 0.04258876293897629
step: 20, loss: 0.37136104702949524
step: 30, loss: 0.09687234461307526
step: 40, loss: 0.1478796899318695
```

```
step: 50, loss: 0.06793300807476044
step: 60, loss: 0.16306129097938538
step: 70, loss: 0.029519176110625267
step: 80, loss: 0.01742039993405342
step: 90, loss: 0.23550669848918915
step: 100, loss: 0.0050387014634907246
step: 110, loss: 0.13471031188964844
step: 120, loss: 0.10191141068935394
step: 130, loss: 0.04291908070445061
step: 140, loss: 0.073587566614151
step: 150, loss: 0.05974706634879112
step: 160, loss: 0.11155737936496735
step: 170, loss: 0.05288739129900932
step: 180, loss: 0.04849237948656082
step: 190, loss: 0.04843819886445999
step: 200, loss: 0.10828796774148941
step: 210, loss: 0.03147848695516586
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.700952380952381, f1=0.6928838951310863,
best_f1=0.6928838951310863
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.020266523584723473
step: 10, loss: 0.03518495708703995
step: 20, loss: 0.12015429139137268
step: 30, loss: 0.004726327024400234
step: 40, loss: 0.0635327696800232
step: 50, loss: 0.006388071924448013
step: 60, loss: 0.2754669785499573
step: 70, loss: 0.006910889875143766
step: 80, loss: 0.10444064438343048
step: 90, loss: 0.05822986736893654
step: 100, loss: 0.013545789755880833
step: 110, loss: 0.005092693492770195
step: 120, loss: 0.03712555766105652
step: 130, loss: 0.08850507438182831
step: 140, loss: 0.044696174561977386
step: 150, loss: 0.01594991236925125
step: 160, loss: 0.049166589975357056
step: 170, loss: 0.12125156819820404
step: 180, loss: 0.051094476133584976
```

```
step: 190, loss: 0.055983155965805054
step: 200, loss: 0.0032399925403296947
step: 210, loss: 0.04789644852280617
/usr/local/lib/python3.7/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=0.6615087040618955, f1=0.6692160611854684,
best_f1=0.6928838951310863
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.08127915114164352
step: 10, loss: 0.02892228588461876
step: 20, loss: 0.026160161942243576
step: 30, loss: 0.052875395864248276
step: 40, loss: 0.09714821726083755
step: 50, loss: 0.1751328408718109
step: 60, loss: 0.04236672446131706
step: 70, loss: 0.03112264908850193
step: 80, loss: 0.0785362720489502
step: 90, loss: 0.19114093482494354
step: 100, loss: 0.005820651073008776
step: 110, loss: 0.2262616604566574
step: 120, loss: 0.052347954362630844
step: 130, loss: 0.03088473342359066
step: 140, loss: 0.012061181478202343
step: 150, loss: 0.033005569130182266
step: 160, loss: 0.03208281099796295
step: 170, loss: 0.021838150918483734
step: 180, loss: 0.013779574073851109
step: 190, loss: 0.10213777422904968
step: 200, loss: 0.032410044223070145
step: 210, loss: 0.024715058505535126
/usr/local/lib/python3.7/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=0.6561886051080551, f1=0.6576402321083173,
best_f1=0.6928838951310863
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.07327891886234283
step: 10, loss: 0.1188114583492279
step: 20, loss: 0.007781947031617165
step: 30, loss: 0.057399868965148926
step: 40, loss: 0.009379838593304157
step: 50, loss: 0.00788422953337431
step: 60, loss: 0.0449359267950058
```

```
step: 70, loss: 0.031647443771362305
step: 80, loss: 0.13417614996433258
step: 90, loss: 0.006609312724322081
step: 100, loss: 0.04573671519756317
step: 110, loss: 0.18488344550132751
step: 120, loss: 0.015420926734805107
step: 130, loss: 0.006493610329926014
step: 140, loss: 0.08756720274686813
step: 150, loss: 0.047209009528160095
step: 160, loss: 0.011738837696611881
step: 170, loss: 0.02320799045264721
step: 180, loss: 0.13451898097991943
step: 190, loss: 0.00331412092782557
step: 200, loss: 0.0140226436778903
step: 210, loss: 0.13162443041801453
/usr/local/lib/python3.7/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=0.6680244399185336, f1=0.6625514403292181,
best_f1=0.6928838951310863
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.014358852989971638
step: 10, loss: 0.14726591110229492
step: 20, loss: 0.011903375387191772
step: 30, loss: 0.10475293546915054
step: 40, loss: 0.022059671580791473
step: 50, loss: 0.0620155967772007
step: 60, loss: 0.15046678483486176
step: 70, loss: 0.08131146430969238
step: 80, loss: 0.004944295156747103
step: 90, loss: 0.030322298407554626
step: 100, loss: 0.004029077477753162
step: 110, loss: 0.05091540887951851
step: 120, loss: 0.00552815617993474
step: 130, loss: 0.0031851057428866625
step: 140, loss: 0.017895428463816643
step: 150, loss: 0.24073921144008636
step: 160, loss: 0.0036692458670586348
step: 170, loss: 0.006967563182115555
step: 180, loss: 0.015753747895359993
step: 190, loss: 0.0017208089120686054
step: 200, loss: 0.10106059163808823
step: 210, loss: 0.010723580606281757
Gradient overflow.  Skipping step, loss scaler 0 reducing loss scale to 16384.0
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
```

```
  warnings.warn("An input tensor was not cuda.")
epoch 9: dev_f1=0.6610169491525424, f1=0.652542372881356,
best_f1=0.6928838951310863
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.14934559166431427
step: 10, loss: 0.06599054485559464
step: 20, loss: 0.012215256690979004
step: 30, loss: 0.170230433344841
step: 40, loss: 0.0157984159886837
step: 50, loss: 0.0070367781445384026
step: 60, loss: 0.16900525987148285
step: 70, loss: 0.0426776148378849
step: 80, loss: 0.03204770013689995
step: 90, loss: 0.029917100444436073
step: 100, loss: 0.03567573428153992
step: 110, loss: 0.0008049691678024828
step: 120, loss: 0.04438384994864464
step: 130, loss: 0.03966689109802246
step: 140, loss: 0.07368934899568558
step: 150, loss: 0.03418935090303421
step: 160, loss: 0.007292026653885841
step: 170, loss: 0.004553858656436205
step: 180, loss: 0.06455941498279572
step: 190, loss: 0.017610780894756317
step: 200, loss: 0.0086404699832201
step: 210, loss: 0.0432850643992424
/usr/local/lib/python3.7/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=0.659919028340081, f1=0.6468253968253969,
best_f1=0.6928838951310863
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.079476557767202377
step: 10, loss: 0.1289757788181305
step: 20, loss: 0.02240743115544319
step: 30, loss: 0.00046357751125469804
step: 40, loss: 0.010045495815575123
step: 50, loss: 0.021873997524380684
step: 60, loss: 0.03468218073248863
step: 70, loss: 0.009894885122776031
step: 80, loss: 0.06476342678070068
step: 90, loss: 0.19406171143054962
step: 100, loss: 0.04052308201789856
step: 110, loss: 0.08739090710878372
```

```
step: 120, loss: 0.16696932911872864
step: 130, loss: 0.016913942992687225
step: 140, loss: 0.024083472788333893
step: 150, loss: 0.004037641454488039
step: 160, loss: 0.0066148764453828335
step: 170, loss: 0.09353519976139069
step: 180, loss: 0.00626385910436511
step: 190, loss: 0.014782138168811798
step: 200, loss: 0.004667400848120451
step: 210, loss: 0.0026478280778974295
/usr/local/lib/python3.7/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=0.6626506024096386, f1=0.6459627329192545,
best_f1=0.6928838951310863
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.04578777775168419
step: 10, loss: 0.002073517767712474
step: 20, loss: 0.023375842720270157
step: 30, loss: 0.0019997290801256895
step: 40, loss: 0.1612979620695114
step: 50, loss: 0.011014607734978199
step: 60, loss: 0.12424435466527939
step: 70, loss: 0.004619946703314781
step: 80, loss: 0.08907999843358994
step: 90, loss: 0.0572272352874279
step: 100, loss: 0.04948955774307251
step: 110, loss: 0.004345242399722338
step: 120, loss: 0.0029281240422278643
step: 130, loss: 0.02715270034968853
step: 140, loss: 0.12595431506633759
step: 150, loss: 0.008324031718075275
step: 160, loss: 0.016678180545568466
step: 170, loss: 0.006093323230743408
step: 180, loss: 0.0074304016306996346
step: 190, loss: 0.14873649179935455
step: 200, loss: 0.009079145267605782
step: 210, loss: 0.03287637606263161
/usr/local/lib/python3.7/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=0.6625258799171843, f1=0.6538461538461539,
best_f1=0.6928838951310863
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
```

```
step: 0, loss: 0.07505571842193604
step: 10, loss: 0.001101690111681819
step: 20, loss: 0.005462172906845808
step: 30, loss: 0.031697310507297516
step: 40, loss: 0.0021676376927644014
step: 50, loss: 0.03549923375248909
step: 60, loss: 0.001678365864790976
step: 70, loss: 0.029355123639310675
step: 80, loss: 0.01054183579981327
step: 90, loss: 0.0007763530593365431
step: 100, loss: 0.0014150760835036635
step: 110, loss: 0.0016683695139363408
step: 120, loss: 0.002945000072941184
step: 130, loss: 0.0007892890134826303
step: 140, loss: 0.0015825567534193397
step: 150, loss: 0.06482871621847153
step: 160, loss: 0.005509908311069012
step: 170, loss: 0.00866283755749464
step: 180, loss: 0.020620007067918777
step: 190, loss: 0.0023811645805835724
step: 200, loss: 0.0021743513643741608
step: 210, loss: 0.0027509895153343678
/usr/local/lib/python3.7/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=0.6613226452905812, f1=0.6625766871165644,
best_f1=0.6928838951310863
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.002130323089659214
step: 10, loss: 0.036182697862386703
step: 20, loss: 0.0010764502221718143
step: 30, loss: 0.008638857863843441
step: 40, loss: 0.002746790647506714
step: 50, loss: 0.03465171158313751
step: 60, loss: 0.0011976818786934018
step: 70, loss: 0.0018691349541768432
step: 80, loss: 0.009222117252647877
step: 90, loss: 0.011222746223211288
step: 100, loss: 0.005782946012914181
step: 110, loss: 0.0053054760210216045
step: 120, loss: 0.026473769918084145
step: 130, loss: 0.002668804721906781
step: 140, loss: 0.014806913211941719
step: 150, loss: 0.025571554899215698
step: 160, loss: 0.00436740554869175
step: 170, loss: 0.0038544468116015196
```

```
step: 180, loss: 0.014150574803352356
step: 190, loss: 0.005479025188833475
step: 200, loss: 0.11775519698858261
step: 210, loss: 0.02434883825480938
/usr/local/lib/python3.7/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=0.6551724137931033, f1=0.6627450980392158,
best_f1=0.6928838951310863
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.004113303031772375
step: 10, loss: 0.005507614929229021
step: 20, loss: 0.025877168402075768
step: 30, loss: 0.012151588685810566
step: 40, loss: 0.010476954281330109
step: 50, loss: 0.0015212582657113671
step: 60, loss: 0.019405236467719078
step: 70, loss: 0.015290336683392525
step: 80, loss: 0.003998701926320791
step: 90, loss: 0.048474639654159546
step: 100, loss: 0.019536493346095085
step: 110, loss: 0.0007775407866574824
step: 120, loss: 0.015269072726368904
step: 130, loss: 0.1767289638519287
step: 140, loss: 0.0007657125825062394
step: 150, loss: 0.004018212202936411
step: 160, loss: 0.003159131621941924
step: 170, loss: 0.0013760300353169441
step: 180, loss: 0.009490610100328922
step: 190, loss: 0.057793810963630676
step: 200, loss: 0.021589534357190132
step: 210, loss: 0.02932339906692505
/usr/local/lib/python3.7/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=0.6603053435114504, f1=0.6666666666666667,
best_f1=0.6928838951310863
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, 486.40it/s]
load_f1 = 0.7074074074074074
real_f1 = 0.7024029574861369
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, 271.10it/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: 0.5808643698692322
step: 10, loss: 0.3833308517932892
step: 20, loss: 0.2889449894428253
step: 30, loss: 0.45252662897109985
step: 40, loss: 0.42290154099464417
step: 50, loss: 0.294102281332016
step: 60, loss: 0.27488815784454346
step: 70, loss: 0.2779935300350189
step: 80, loss: 0.27258503437042236
step: 90, loss: 0.2987004220485687
step: 100, loss: 0.33599069714546204
step: 110, loss: 0.3892587721347809
step: 120, loss: 0.08007869124412537
step: 130, loss: 0.100266233086586
Gradient overflow.  Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 140, loss: 0.06274550408124924
step: 150, loss: 0.16285300254821777
step: 160, loss: 0.05863627791404724
step: 170, loss: 0.19651685655117035
step: 180, loss: 0.03130393847823143
step: 190, loss: 0.2646682560443878
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.6567164179104478, f1=0.6567164179104478,
best_f1=0.6567164179104478
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.38604557514190674
step: 10, loss: 0.05890318378806114
step: 20, loss: 0.19541141390800476
step: 30, loss: 0.12911032140254974
step: 40, loss: 0.11362045258283615
step: 50, loss: 0.15441671013832092
step: 60, loss: 0.22549065947532654
```

```
step: 70, loss: 0.19817310571670532
step: 80, loss: 0.2890779674053192
step: 90, loss: 0.22348469495773315
step: 100, loss: 0.033364418894052505
step: 110, loss: 0.24971260130405426
step: 120, loss: 0.3461412191390991
step: 130, loss: 0.08769963681697845
step: 140, loss: 0.07003937661647797
step: 150, loss: 0.10053519904613495
step: 160, loss: 0.019472962245345116
step: 170, loss: 0.16000185906887054
step: 180, loss: 0.06238052621483803
step: 190, loss: 0.05021652579307556
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.7666666666666667, f1=0.76, best_f1=0.76
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.03626517951488495
step: 10, loss: 0.1227463111281395
step: 20, loss: 0.11305126547813416
step: 30, loss: 0.12794917821884155
step: 40, loss: 0.07473689317703247
step: 50, loss: 0.20875418186187744
step: 60, loss: 0.13325655460357666
step: 70, loss: 0.0965212881565094
step: 80, loss: 0.08073597401380539
step: 90, loss: 0.13024844229221344
step: 100, loss: 0.02744249813258648
step: 110, loss: 0.034133244305849075
step: 120, loss: 0.023576829582452774
step: 130, loss: 0.013900845311582088
step: 140, loss: 0.10356828570365906
step: 150, loss: 0.1132173091173172
step: 160, loss: 0.02774689719080925
step: 170, loss: 0.2348616123199463
step: 180, loss: 0.028184793889522552
step: 190, loss: 0.13684596121311188
/usr/local/lib/python3.7/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=0.7534626038781163, f1=0.7780821917808218, best_f1=0.76
```

```
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.22022943198680878
step: 10, loss: 0.12009681761264801
step: 20, loss: 0.032256562262773514
step: 30, loss: 0.017099546268582344
step: 40, loss: 0.006182825658470392
step: 50, loss: 0.016689110547304153
step: 60, loss: 0.04004666581749916
step: 70, loss: 0.05890636518597603
step: 80, loss: 0.26030564308166504
step: 90, loss: 0.020649824291467667
step: 100, loss: 0.12994781136512756
step: 110, loss: 0.02219228632748127
step: 120, loss: 0.0805712640285492
step: 130, loss: 0.23326584696769714
step: 140, loss: 0.05289953947067261
step: 150, loss: 0.01699436642229557
step: 160, loss: 0.01826256513595581
step: 170, loss: 0.07387124001979828
step: 180, loss: 0.13498343527317047
step: 190, loss: 0.25135642290115356
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.784, f1=0.7968337730870713, best_f1=0.7968337730870713
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.071992261337518692
step: 10, loss: 0.03615245595574379
step: 20, loss: 0.042025916278362274
step: 30, loss: 0.0020366597454994917
step: 40, loss: 0.051576629281044006
step: 50, loss: 0.040959425270557404
step: 60, loss: 0.17356516420841217
step: 70, loss: 0.06678397953510284
step: 80, loss: 0.020843422040343285
step: 90, loss: 0.034193769097328186
step: 100, loss: 0.003608266357332468
step: 110, loss: 0.12463663518428802
step: 120, loss: 0.007075462024658918
step: 130, loss: 0.13474750518798828
```

```
step: 140, loss: 0.06759360432624817
step: 150, loss: 0.08635873347520828
step: 160, loss: 0.010990628972649574
step: 170, loss: 0.003955975640565157
step: 180, loss: 0.05155595391988754
step: 190, loss: 0.0023023756220936775
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.7885117493472584, f1=0.7736842105263158,
best_f1=0.7736842105263158
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00565031124278903
step: 10, loss: 0.025304565206170082
step: 20, loss: 0.0060349199920892715
step: 30, loss: 0.0005828206776641309
step: 40, loss: 0.0032726977951824665
step: 50, loss: 0.0026650603394955397
step: 60, loss: 0.04455474764108658
step: 70, loss: 0.0033291042782366276
step: 80, loss: 0.012064323760569096
step: 90, loss: 0.030429599806666374
step: 100, loss: 0.0016802353784441948
step: 110, loss: 0.07801579684019089
step: 120, loss: 0.0015475329710170627
step: 130, loss: 0.025006385520100594
step: 140, loss: 0.04894730821251869
step: 150, loss: 0.003080027876421809
step: 160, loss: 0.010074947960674763
step: 170, loss: 0.010611167177557945
step: 180, loss: 0.006330089643597603
step: 190, loss: 0.026034070178866386
/usr/local/lib/python3.7/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=0.762836185819071, f1=0.775, best_f1=0.7736842105263158
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.009224440902471542
step: 10, loss: 0.0035818524193018675
step: 20, loss: 0.0322389081120491
```

```
step: 30, loss: 0.014417641796171665
step: 40, loss: 0.0010670052142813802
step: 50, loss: 0.08450786024332047
step: 60, loss: 0.018102722242474556
step: 70, loss: 0.0037978412583470345
step: 80, loss: 0.005778635386377573
step: 90, loss: 0.04424852877855301
step: 100, loss: 0.0021593377459794283
step: 110, loss: 0.007636784575879574
step: 120, loss: 0.006326606497168541
step: 130, loss: 0.0013765250332653522
step: 140, loss: 0.00239452812820673
step: 150, loss: 0.0018472628435119987
step: 160, loss: 0.15375296771526337
step: 170, loss: 0.0016377074643969536
step: 180, loss: 0.003460288979113102
step: 190, loss: 0.0009522258187644184
/usr/local/lib/python3.7/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=0.7851458885941645, f1=0.7885117493472584,
best_f1=0.7736842105263158
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.015387564897537231
step: 10, loss: 0.0050312732346355915
step: 20, loss: 0.0006838878034614027
step: 30, loss: 0.0006784333381801844
step: 40, loss: 0.0027373579796403646
step: 50, loss: 0.004175085108727217
step: 60, loss: 0.13474026322364807
step: 70, loss: 0.0017281086184084415
step: 80, loss: 0.0005657464498654008
step: 90, loss: 0.0007813929696567357
step: 100, loss: 0.006861054338514805
step: 110, loss: 0.0008038937812671065
step: 120, loss: 0.007201808970421553
step: 130, loss: 0.0005808774149045348
step: 140, loss: 0.10900091379880905
step: 150, loss: 0.03977581486105919
step: 160, loss: 0.001079052104614675
step: 170, loss: 0.004752477165311575
step: 180, loss: 0.035650912672281265
step: 190, loss: 0.002484091091901064
/usr/local/lib/python3.7/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=0.768079800498753, f1=0.7733990147783251,
best_f1=0.7736842105263158
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0020769464317709208
step: 10, loss: 0.029327569529414177
step: 20, loss: 0.003985781222581863
step: 30, loss: 0.0014393434394150972
step: 40, loss: 0.07994192093610764
step: 50, loss: 0.00845097005367279
step: 60, loss: 0.0010559562360867858
step: 70, loss: 0.0030165277421474457
step: 80, loss: 0.002075537107884884
step: 90, loss: 0.07865490019321442
step: 100, loss: 0.04944869130849838
step: 110, loss: 0.0074824681505560875
step: 120, loss: 0.005063815973699093
step: 130, loss: 0.013468514196574688
step: 140, loss: 0.1517782062292099
step: 150, loss: 0.003869524458423257
step: 160, loss: 0.0014721043407917023
step: 170, loss: 0.007010593079030514
step: 180, loss: 0.0045075127854943275
step: 190, loss: 0.0015861260471865535
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.7936507936507936, f1=0.7864583333333334,
best_f1=0.7864583333333334
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0016767893685027957
step: 10, loss: 0.003372940933331847
step: 20, loss: 0.012977608479559422
step: 30, loss: 0.00032452773302793503
step: 40, loss: 0.009503564797341824
step: 50, loss: 0.0015789043391123414
step: 60, loss: 0.0027158865705132484
step: 70, loss: 0.003501972882077098
step: 80, loss: 0.0009587202803231776
step: 90, loss: 0.0007343521574512124
step: 100, loss: 0.000994980102404952
```

```
step: 110, loss: 0.005578856915235519
step: 120, loss: 0.03186045587062836
step: 130, loss: 0.10078322887420654
step: 140, loss: 0.0017857248894870281
step: 150, loss: 0.0016050926642492414
step: 160, loss: 0.0016366984928026795
step: 170, loss: 0.0008346604881808162
step: 180, loss: 0.0005848421133123338
step: 190, loss: 0.002479920629411936
/usr/local/lib/python3.7/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=0.7867036011080332, f1=0.7407407407407407,
best_f1=0.7864583333333334
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0002710640837904066
step: 10, loss: 0.0005270740366540849
step: 20, loss: 0.019609250128269196
step: 30, loss: 0.019237304106354713
step: 40, loss: 0.00023751991102471948
step: 50, loss: 0.009836296550929546
step: 60, loss: 0.0004634743381757289
step: 70, loss: 0.0004581943212542683
step: 80, loss: 0.0002917862148024142
step: 90, loss: 0.00033760222140699625
step: 100, loss: 0.0003907439822796732
step: 110, loss: 0.004119617398828268
step: 120, loss: 0.017667537555098534
step: 130, loss: 0.0003508335503283888
step: 140, loss: 0.014729891903698444
step: 150, loss: 0.001484394771978259
step: 160, loss: 0.0018379585817456245
step: 170, loss: 0.000598392856772989
step: 180, loss: 0.011469267308712006
step: 190, loss: 0.012110378593206406
/usr/local/lib/python3.7/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=0.7857142857142857, f1=0.7848101265822786,
best_f1=0.7864583333333334
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.008478126488626003
step: 10, loss: 0.010848529636859894
step: 20, loss: 0.045984067022800446
```

```
step: 30, loss: 0.0026347481179982424
step: 40, loss: 0.02109253592789173
step: 50, loss: 0.01091738324612379
step: 60, loss: 0.0006727995933033526
step: 70, loss: 0.003622446209192276
step: 80, loss: 0.0008617838029749691
step: 90, loss: 0.003236701003462076
step: 100, loss: 0.0028580334037542343
step: 110, loss: 0.0042795962654054165
step: 120, loss: 0.0003763757413253188
step: 130, loss: 0.00133475661277771
step: 140, loss: 0.019516512751579285
step: 150, loss: 0.0006411655340343714
step: 160, loss: 0.0004470777930691838
step: 170, loss: 0.001445556408725679
step: 180, loss: 0.0004504169919528067
step: 190, loss: 0.0024635076988488436
/usr/local/lib/python3.7/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=0.7826086956521741, f1=0.7857142857142857,
best_f1=0.7864583333333334
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0007720035500824451
step: 10, loss: 0.0014550051419064403
step: 20, loss: 0.0015467716148123145
step: 30, loss: 0.021450646221637726
step: 40, loss: 0.0008209163788706064
step: 50, loss: 0.06689487397670746
step: 60, loss: 0.0008625455666333437
step: 70, loss: 0.0062404656782746315
step: 80, loss: 0.00571681372821331
step: 90, loss: 0.002495924709364772
step: 100, loss: 0.006461888551712036
step: 110, loss: 0.00028641981771215796
step: 120, loss: 0.008938880637288094
step: 130, loss: 0.000406875042244792
step: 140, loss: 0.0006114545976743102
step: 150, loss: 0.0008101603016257286
step: 160, loss: 0.0016926416428759694
step: 170, loss: 0.0005127215408720076
step: 180, loss: 0.0008989848429337144
step: 190, loss: 0.002763367723673582
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
```

```
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.8031496062992125, f1=0.7904509283819628,
best_f1=0.7904509283819628
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.008540148846805096
step: 10, loss: 0.0004362949694041163
step: 20, loss: 0.0013684326549991965
step: 30, loss: 0.001888272468931973
step: 40, loss: 0.002644808031618595
step: 50, loss: 0.0008776768809184432
step: 60, loss: 0.0035916611086577177
step: 70, loss: 0.00023623455490451306
step: 80, loss: 0.0003540850302670151
step: 90, loss: 0.00856061838567257
step: 100, loss: 0.015182011760771275
step: 110, loss: 0.001729561248794198
step: 120, loss: 0.001203562831506133
step: 130, loss: 0.018960997462272644
step: 140, loss: 0.007048031780868769
step: 150, loss: 0.003057815134525299
step: 160, loss: 0.008880850858986378
step: 170, loss: 0.006989878136664629
step: 180, loss: 0.0010998252546414733
step: 190, loss: 0.00032462168019264936
/usr/local/lib/python3.7/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=0.7893333333333333, f1=0.7804878048780487,
best_f1=0.7904509283819628
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.004858469124883413
step: 10, loss: 0.0051247538067400455
step: 20, loss: 0.03386039286851883
step: 30, loss: 0.0004538467328529805
step: 40, loss: 0.0003224034735467285
step: 50, loss: 0.025390267372131348
step: 60, loss: 0.049020662903785706
step: 70, loss: 0.007104829885065556
step: 80, loss: 0.00030117720598354936
step: 90, loss: 0.0011418106732890010
step: 100, loss: 0.001017016707919538
```

```
step: 110, loss: 0.0007602543337270617
step: 120, loss: 0.0005249206442385912
step: 130, loss: 0.0001783241459634155
step: 140, loss: 0.0029609163757413626
step: 150, loss: 0.0003736735670827329
Gradient overflow.  Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 160, loss: 0.00024624139769002795
step: 170, loss: 0.0002522616705391556
step: 180, loss: 0.0026494867634028196
step: 190, loss: 0.0001513738534413278
/usr/local/lib/python3.7/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=0.7937336814621411, f1=0.7937336814621411,
best_f1=0.7904509283819628
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, 234.99it/s]
load_f1 = 0.7037974683544304
real_f1 = 0.6947890818858561
```

```
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, 270.83it/s]
```

4.2 DITTO TEXTUAL

4.2.1 Abt-Buy - Running the matcher

```
[ ]:  # Step 1:
      # train the wdc model with the save_model flag on
      # it will produce two files *_dev.pt and *_test.pt which are the
      # best checkpoints on the dev set and the test set.
      !CUDA_VISIBLE_DEVICES=0 python train_ditto.py \
        --task Textual/Abt-Buy \
        --batch_size 32 \
        --max_len 256 \
        --lr 3e-5 \
        --n_epochs 15 \
        --finetuning \
        --lm 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, 392kB/s]
Downloading: 100% 232k/232k [00:00<00:00, 257kB/s]
Downloading: 100% 440M/440M [00:06<00:00, 70.4MB/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: 0.640425443649292
step: 10, loss: 0.36464306712150574
step: 20, loss: 0.29676830768585205
step: 30, loss: 0.4180218279361725
step: 40, loss: 0.28286653757095337
step: 50, loss: 0.2592591643333435
step: 60, loss: 0.2437029927968979
step: 70, loss: 0.374694287776947
step: 80, loss: 0.408625066280365
step: 90, loss: 0.2608788013458252
step: 100, loss: 0.27509233355522156
step: 110, loss: 0.31629085540771484

```
step: 120, loss: 0.20698101818561554
step: 130, loss: 0.06288129836320877
step: 140, loss: 0.2010435163974762
step: 150, loss: 0.22700345516204834
step: 160, loss: 0.14270886778831482
step: 170, loss: 0.24822698533535004
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.7111111111111111, f1=0.6847290640394088,
best_f1=0.6847290640394088
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.08808047324419022
step: 10, loss: 0.3216687738895416
step: 20, loss: 0.0553598552942276
step: 30, loss: 0.24580347537994385
step: 40, loss: 0.06005960330367088
step: 50, loss: 0.2498127520084381
step: 60, loss: 0.16365812718868256
step: 70, loss: 0.09844357520341873
step: 80, loss: 0.0950663834810257
step: 90, loss: 0.11449076980352402
step: 100, loss: 0.1245528981089592
step: 110, loss: 0.07515353709459305
step: 120, loss: 0.2598631978034973
step: 130, loss: 0.10804972052574158
step: 140, loss: 0.24781234562397003
step: 150, loss: 0.15086403489112854
step: 160, loss: 0.17991574108600616
step: 170, loss: 0.03237931802868843
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.7606263982102909, f1=0.7578475336322871,
best_f1=0.7578475336322871
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
```

```
step: 0, loss: 0.27123069763183594
step: 10, loss: 0.10132894665002823
step: 20, loss: 0.016253415495157242
step: 30, loss: 0.19009913504123688
step: 40, loss: 0.12203672528266907
step: 50, loss: 0.016782479360699654
step: 60, loss: 0.1022065058350563
step: 70, loss: 0.05857080593705177
step: 80, loss: 0.0359523631632328
step: 90, loss: 0.20311109721660614
step: 100, loss: 0.06973825395107269
step: 110, loss: 0.1035359725356102
step: 120, loss: 0.0359170027077198
step: 130, loss: 0.12423214316368103
step: 140, loss: 0.09660860151052475
step: 150, loss: 0.19522714614868164
step: 160, loss: 0.08833807706832886
step: 170, loss: 0.2872662842273712
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.801007556675063, f1=0.7849999999999999,
best_f1=0.7849999999999999
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.02112256921827793
step: 10, loss: 0.03379717841744423
step: 20, loss: 0.0062301550060510635
step: 30, loss: 0.16711221635341644
step: 40, loss: 0.003283505793660879
step: 50, loss: 0.14456653594970703
step: 60, loss: 0.19052892923355103
step: 70, loss: 0.005809340626001358
step: 80, loss: 0.09004426747560501
step: 90, loss: 0.08952800184488297
step: 100, loss: 0.1750626266002655
step: 110, loss: 0.011721877381205559
step: 120, loss: 0.005194612313061953
step: 130, loss: 0.08554442226886749
step: 140, loss: 0.02241520583629608
step: 150, loss: 0.07437149435281754
step: 160, loss: 0.0740833580493927
step: 170, loss: 0.17260578274726868
```

```
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.8080808080808081, f1=0.8, best_f1=0.8
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.1205861046910286
step: 10, loss: 0.0423269085586071
step: 20, loss: 0.08176124840974808
step: 30, loss: 0.05858882516622543
step: 40, loss: 0.01377347856760025
step: 50, loss: 0.031422555446624756
step: 60, loss: 0.0675254836678505
step: 70, loss: 0.32202979922294617
step: 80, loss: 0.06953074038028717
step: 90, loss: 0.09502141177654266
step: 100, loss: 0.09355907887220383
step: 110, loss: 0.032473914325237274
step: 120, loss: 0.028169983997941017
step: 130, loss: 0.08470728993415833
step: 140, loss: 0.014817692339420319
step: 150, loss: 0.05729113146662712
step: 160, loss: 0.038990821689367294
step: 170, loss: 0.01716211996972561
/usr/local/lib/python3.7/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=0.8020050125313284, f1=0.800982800982801, best_f1=0.8
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.008861282840371132
step: 10, loss: 0.007331062573939562
step: 20, loss: 0.005659463815391064
step: 30, loss: 0.04922684282064438
step: 40, loss: 0.09749966859817505
step: 50, loss: 0.16212911903858185
step: 60, loss: 0.05973849818110466
step: 70, loss: 0.1184108704328537
step: 80, loss: 0.09582261741161346
step: 90, loss: 0.0353001169860363
step: 100, loss: 0.03588356822729111
step: 110, loss: 0.008631155826151371
```

```
step: 120, loss: 0.041376084089279175
step: 130, loss: 0.014980344101786613
step: 140, loss: 0.10610251128673553
step: 150, loss: 0.0699765682220459
step: 160, loss: 0.1659957766532898
step: 170, loss: 0.10663378238677979
/usr/local/lib/python3.7/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=0.8058968058968059, f1=0.788177339901478, best_f1=0.8
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0158421378582716
step: 10, loss: 0.04253043234348297
step: 20, loss: 0.0046122511848807335
step: 30, loss: 0.006156083662062883
step: 40, loss: 0.010134400799870491
step: 50, loss: 0.008112738840281963
step: 60, loss: 0.017370743677020073
step: 70, loss: 0.0012245996622368693
step: 80, loss: 0.0473036915063858
step: 90, loss: 0.002301885513588786
step: 100, loss: 0.000728354905731976
step: 110, loss: 0.046345315873622894
step: 120, loss: 0.014037612825632095
step: 130, loss: 0.15387707948684692
step: 140, loss: 0.0023108827881515026
step: 150, loss: 0.013552024960517883
step: 160, loss: 0.003927096724510193
step: 170, loss: 0.02753981202840805
/usr/local/lib/python3.7/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=0.8, f1=0.8146341463414632, best_f1=0.8
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.008541785180568695
step: 10, loss: 0.00887291133403778
step: 20, loss: 0.0015618398319929838
step: 30, loss: 0.03615884110331535
step: 40, loss: 0.0005420444067567587
step: 50, loss: 0.0018827990861609578
step: 60, loss: 0.009001229889690876
step: 70, loss: 0.09498290717601776
step: 80, loss: 0.013350420631468296
step: 90, loss: 0.005747393239289522
```

```
step: 100, loss: 0.06788115203380585
step: 110, loss: 0.11744029819965363
step: 120, loss: 0.01508792769163847
step: 130, loss: 0.013921888545155525
step: 140, loss: 0.0012934254482388496
step: 150, loss: 0.04722415283322334
step: 160, loss: 0.02451995760202408
step: 170, loss: 0.016245892271399498
/usr/local/lib/python3.7/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=0.782608695652174, f1=0.7872860635696822, best_f1=0.8
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.024195564910769463
step: 10, loss: 0.0020172998774796724
Gradient overflow.  Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 20, loss: 0.12025199830532074
step: 30, loss: 0.10921931266784668
step: 40, loss: 0.008775515481829643
step: 50, loss: 0.0023021406959742308
step: 60, loss: 0.03245590627193451
step: 70, loss: 0.0024480638094246387
step: 80, loss: 0.0013553125318139791
step: 90, loss: 0.022250743582844734
step: 100, loss: 0.01817184127867222
step: 110, loss: 0.00968507956713438
step: 120, loss: 0.00925581157207489
step: 130, loss: 0.06531253457069397
step: 140, loss: 0.056704673916101456
step: 150, loss: 0.004735163412988186
step: 160, loss: 0.10566037893295288
step: 170, loss: 0.0040480708703398705
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.8188585607940446, f1=0.8, best_f1=0.8
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.032431457191705704
step: 10, loss: 0.00041049355058930814
step: 20, loss: 0.09370116889476776
```

```
step: 30, loss: 0.015316246077418327
step: 40, loss: 0.06962040811777115
step: 50, loss: 0.06360830366611481
step: 60, loss: 0.0016736435936763883
step: 70, loss: 0.015986373648047447
step: 80, loss: 0.0006670069415122271
step: 90, loss: 0.005609242711216211
step: 100, loss: 0.0004966098349541426
step: 110, loss: 0.028662962839007378
step: 120, loss: 0.044800370931625366
step: 130, loss: 0.0004997575888410211
step: 140, loss: 0.012247418984770775
step: 150, loss: 0.0014069024473428726
step: 160, loss: 0.0100680161267519
step: 170, loss: 0.0017963291611522436
/usr/local/lib/python3.7/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=0.78743961352657, f1=0.7926267281105991, best_f1=0.8
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0170064028352499
step: 10, loss: 0.013949775137007236
step: 20, loss: 0.0022287415340542793
step: 30, loss: 0.007917969487607479
step: 40, loss: 0.050529927015304565
step: 50, loss: 0.0018338914960622787
step: 60, loss: 0.007723451592028141
step: 70, loss: 0.0012547045480459929
step: 80, loss: 0.0009709944133646786
step: 90, loss: 0.0010811520041897893
step: 100, loss: 0.004459843970835209
step: 110, loss: 0.027343224734067917
step: 120, loss: 0.0006666685803793371
step: 130, loss: 0.0002838062064256519
step: 140, loss: 0.006192624103277922
step: 150, loss: 0.01352771557867527
step: 160, loss: 0.00465787947177887
step: 170, loss: 0.0009307475993409753
/usr/local/lib/python3.7/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=0.7978142076502732, f1=0.7989276139410187, best_f1=0.8
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0005731210112571716
```

```
step: 10, loss: 0.0016081786016002297
step: 20, loss: 0.22761811316013336
step: 30, loss: 0.002518133958801627
step: 40, loss: 0.001738753286190331
step: 50, loss: 0.0017126458697021008
step: 60, loss: 0.12661947309970856
step: 70, loss: 0.07441601902246475
step: 80, loss: 0.0016514264279976487
step: 90, loss: 0.0024034332018345594
step: 100, loss: 0.002013436984270811
step: 110, loss: 0.002415430499240756
step: 120, loss: 0.0693492591381073
step: 130, loss: 0.017883604392409325
step: 140, loss: 0.030673304572701454
step: 150, loss: 0.009888364002108574
step: 160, loss: 0.003361500333994627
step: 170, loss: 0.0007085426477715373
/usr/local/lib/python3.7/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=0.7884615384615384, f1=0.7852193995381063, best_f1=0.8
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.020361769944429398
step: 10, loss: 0.0005622526514343917
step: 20, loss: 0.0008062487468123436
step: 30, loss: 0.0008197668939828873
step: 40, loss: 0.0033037762623280287
step: 50, loss: 0.001102752285078168
step: 60, loss: 0.034499961882829666
step: 70, loss: 0.003960344474762678
step: 80, loss: 0.0037885443307459354
step: 90, loss: 0.0004663288127630949
step: 100, loss: 0.004847235046327114
step: 110, loss: 0.001187362358905375
step: 120, loss: 0.04038245603442192
step: 130, loss: 0.006295311730355024
step: 140, loss: 0.0019866900984197855
step: 150, loss: 0.0016934947343543172
step: 160, loss: 0.0012557642767205834
step: 170, loss: 0.0026699139270931482
/usr/local/lib/python3.7/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=0.793733681462141, f1=0.8080808080808081, best_f1=0.8
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
```

```
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0009152035927399993
step: 10, loss: 0.0004187474260106683
step: 20, loss: 0.002668824279680848
step: 30, loss: 0.0008252033730968833
step: 40, loss: 0.0005627803038805723
step: 50, loss: 0.0015249267453327775
step: 60, loss: 0.0014917092630639672
step: 70, loss: 0.0010026327800005674
step: 80, loss: 0.004271848127245903
step: 90, loss: 0.0004266278410796076
step: 100, loss: 0.0006115606520324945
step: 110, loss: 0.012752528302371502
step: 120, loss: 0.016350045800209045
step: 130, loss: 0.0008206022321246564
step: 140, loss: 0.0011902559781447053
step: 150, loss: 0.0031701638363301754
step: 160, loss: 0.0007176901563070714
step: 170, loss: 0.0004714620881713927
/usr/local/lib/python3.7/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=0.7969151670951157, f1=0.810126582278481, best_f1=0.8
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0002587710914667696
step: 10, loss: 0.003472908865660429
step: 20, loss: 0.23894155025482178
step: 30, loss: 0.0003236150078009814
step: 40, loss: 0.00435674050822854
step: 50, loss: 0.004451464395970106
step: 60, loss: 0.014395480044186115
step: 70, loss: 0.0003350242041051388
step: 80, loss: 0.01665468141436577
step: 90, loss: 0.0008791770087555051
step: 100, loss: 0.001054757391102612
step: 110, loss: 0.0002822311071213335
step: 120, loss: 0.0024151380639523268
step: 130, loss: 0.0010921036591753364
step: 140, loss: 0.00040665658889338374
step: 150, loss: 0.000274669990176335
step: 160, loss: 0.0022677334491163492
step: 170, loss: 0.016369590535759926
/usr/local/lib/python3.7/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=0.7926509186351706, f1=0.8071979434447302, best_f1=0.8
```

```
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, 220.46it/s]
load_f1 = 0.3802521008403361
real_f1 = 0.34908389585342336
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, 177.69it/s]
```

4.3 DITTO DIRTY

4.3.1 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 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

432

```
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: 0.6053374409675598
step: 10, loss: 0.6146105527877808
step: 20, loss: 0.3437322974205017
step: 30, loss: 0.06551893800497055
step: 40, loss: 0.3016244173049927
step: 50, loss: 0.06052976846694946
step: 60, loss: 0.029285192489624023
step: 70, loss: 0.17067338526248932
step: 80, loss: 0.03746478632092476
step: 90, loss: 0.15609967708587646
step: 100, loss: 0.006284716073423624
step: 110, loss: 0.17186759412288666
step: 120, loss: 0.015173711813986301
step: 130, loss: 0.011768899857997894
step: 140, loss: 0.004275006707757711
step: 150, loss: 0.01676623336970806
step: 160, loss: 0.00623351102694869
step: 170, loss: 0.13712593913078308
step: 180, loss: 0.040428876876831055
step: 190, loss: 0.08946597576141357
step: 200, loss: 0.05558205023407936
step: 210, loss: 0.00956965796649456
step: 220, loss: 0.030332257971167564
step: 230, loss: 0.020782168954610825
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.967741935483871, f1=0.968609865470852,
best_f1=0.968609865470852
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.011871227994561195
step: 10, loss: 0.00568348215892911
step: 20, loss: 0.17922765016555786
step: 30, loss: 0.1330866515636444
step: 40, loss: 0.08145634829998016
step: 50, loss: 0.012072310782968998
step: 60, loss: 0.006608330179005861
step: 70, loss: 0.0018582274205982685
step: 80, loss: 0.007440540008246899
step: 90, loss: 0.01149833481758833
step: 100, loss: 0.012991479597985744
step: 110, loss: 0.07665300369262695
step: 120, loss: 0.13446640968322754
step: 130, loss: 0.012161615304648876
step: 140, loss: 0.0035687200725078583
step: 150, loss: 0.012224326841533184
step: 160, loss: 0.02657163515686989
step: 170, loss: 0.02839609608054161
step: 180, loss: 0.119230255484581
step: 190, loss: 0.005122151225805283
step: 200, loss: 0.002547828247770667
step: 210, loss: 0.0011976842069998384
step: 220, loss: 0.1263139545917511
step: 230, loss: 0.007551257498562336
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9765363128491621, f1=0.9765363128491621,
best_f1=0.9765363128491621
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.003866876708343625
step: 10, loss: 0.014970979653298855
step: 20, loss: 0.014891916885972023
step: 30, loss: 0.11962025612592697
step: 40, loss: 0.022962123155593872

```
step: 50, loss: 0.018398562446236611
step: 60, loss: 0.0021848194301128387
step: 70, loss: 0.027046535164117813
step: 80, loss: 0.0016367124626412988
step: 90, loss: 0.021430447697639465
step: 100, loss: 0.0006654404569417238
step: 110, loss: 0.011439284309744835
step: 120, loss: 0.03433583676815033
step: 130, loss: 0.0035791632253676653
step: 140, loss: 0.009579215198755264
step: 150, loss: 0.013620460405945778
step: 160, loss: 0.01998831517994404
step: 170, loss: 0.011213050223886967
step: 180, loss: 0.013181978836655617
step: 190, loss: 0.003414721693843603
step: 200, loss: 0.04494541510939598
step: 210, loss: 0.0105884857457576
step: 220, loss: 0.0017989532789215446
step: 230, loss: 0.022546684369444847
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9766925638179801, f1=0.9733924611973392,
best_f1=0.9733924611973392
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0006034061079844832
step: 10, loss: 0.00416185799986124
step: 20, loss: 0.0015676859766244888
step: 30, loss: 0.01713458076119423
step: 40, loss: 0.004852945450693369
step: 50, loss: 0.001396430772729218
step: 60, loss: 0.0008293335558846593
step: 70, loss: 0.007533080410212278
step: 80, loss: 0.0005418925429694355
step: 90, loss: 0.00337335211224854
step: 100, loss: 0.0019968892447650433
step: 110, loss: 0.0006035075639374554
step: 120, loss: 0.008701002225279808
step: 130, loss: 0.0016592523315921426
step: 140, loss: 0.04237382113933563
step: 150, loss: 0.36763161420822144
step: 160, loss: 0.0009474006947129965
```

```
step: 170, loss: 0.02496352232992649
step: 180, loss: 0.0006357650272548199
step: 190, loss: 0.002516058273613453
step: 200, loss: 0.002100803656503558
step: 210, loss: 0.019273988902568817
step: 220, loss: 0.0008612936362624168
step: 230, loss: 0.00188444706145674
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9796839729119639, f1=0.9773242630385486,
best_f1=0.9773242630385486
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0018891363870352507
step: 10, loss: 0.000555063015781343
step: 20, loss: 0.0012408395996317267
step: 30, loss: 0.0003289613814558834
step: 40, loss: 0.0006320315878838301
step: 50, loss: 0.0005481831612996757
step: 60, loss: 0.11956105381250381
step: 70, loss: 0.0011801904765889049
step: 80, loss: 0.0008840853697620332
step: 90, loss: 0.0015569409588351846
step: 100, loss: 0.0007523809326812625
step: 110, loss: 0.016943734139204025
step: 120, loss: 0.00027209732797928154
step: 130, loss: 0.030854947865009308
step: 140, loss: 0.04638386145234108
step: 150, loss: 0.02561149001121521
Gradient overflow.  Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 160, loss: 0.037036821246147156
step: 170, loss: 0.01785544864833355
step: 180, loss: 0.0010926563991233706
step: 190, loss: 0.08224419504404068
step: 200, loss: 0.0243001436293125
step: 210, loss: 0.009103837423026562
step: 220, loss: 0.00219330913387239
step: 230, loss: 0.0016079054912552238
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
```

```
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=0.9776785714285714,
best_f1=0.9776785714285714
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.03200594335794449
step: 10, loss: 0.001191290677525103
step: 20, loss: 0.0021242534276098013
step: 30, loss: 0.000940259953495115
step: 40, loss: 0.00023548306489828974
step: 50, loss: 0.0015306960558518767
step: 60, loss: 0.00015269404684659094
step: 70, loss: 0.0011224447516724467
step: 80, loss: 0.0005380100919865072
step: 90, loss: 0.003851603949442506
step: 100, loss: 0.0005615978152491152
step: 110, loss: 0.0002559673448558897
step: 120, loss: 0.002718281000852585
step: 130, loss: 0.00020719628082588315
step: 140, loss: 0.00023529892496298999
step: 150, loss: 0.002607227535918355
step: 160, loss: 0.0001462198415538296
step: 170, loss: 0.00024681168724782765
step: 180, loss: 0.010699381120502949
step: 190, loss: 0.15328897535800934
step: 200, loss: 0.014351720921695232
step: 210, loss: 0.01000911183655262
step: 220, loss: 0.0032006765250116587
step: 230, loss: 0.04344252869486809
/usr/local/lib/python3.7/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=0.9798657718120806, f1=0.9798657718120806,
best_f1=0.9776785714285714
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0022511251736432314
step: 10, loss: 0.00016610685270279646
step: 20, loss: 0.006873412057757378
step: 30, loss: 0.001595278619788587
step: 40, loss: 0.005675031803548336
step: 50, loss: 0.003424168098717928
step: 60, loss: 0.025923823937773705
step: 70, loss: 0.004431607201695442
```

```
step: 80, loss: 0.0011095762019976974
step: 90, loss: 0.0015393947251141071
step: 100, loss: 0.0001528822467662394
step: 110, loss: 0.0009376582456752658
step: 120, loss: 0.00016338209388777614
step: 130, loss: 0.025237444788217545
step: 140, loss: 0.0008682177867740393
step: 150, loss: 0.00027117773424834013
step: 160, loss: 0.07134511321783066
step: 170, loss: 0.0008927475428208709
step: 180, loss: 0.0005150139913894236
step: 190, loss: 0.00016507676627952605
step: 200, loss: 0.050537899136543274
step: 210, loss: 0.00010159335943171754
step: 220, loss: 0.0008390299044549465
step: 230, loss: 0.000625244399998337
/usr/local/lib/python3.7/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=0.9798206278026906, f1=0.9743589743589743,
best_f1=0.9776785714285714
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00048057347885333
step: 10, loss: 0.004627116024494171
step: 20, loss: 0.00033056619577109814
step: 30, loss: 0.00014665364869870245
step: 40, loss: 0.006921771448105574
step: 50, loss: 0.0001929294812725857
step: 60, loss: 0.0009534381679259241
step: 70, loss: 0.0004208840546198189
step: 80, loss: 0.00018153058772441
step: 90, loss: 0.0001419267791789025
step: 100, loss: 0.00035176915116608143
step: 110, loss: 0.00011942952551180497
step: 120, loss: 0.028408532962203026
step: 130, loss: 0.001482302090153098
step: 140, loss: 7.713901868555695e-05
step: 150, loss: 6.606936949538067e-05
step: 160, loss: 0.00046979356557130814
step: 170, loss: 0.00037906557554379106
step: 180, loss: 0.022908810526132584
step: 190, loss: 0.0006748242885805666
step: 200, loss: 0.009833923541009426
step: 210, loss: 0.00021246739197522402
step: 220, loss: 0.0004866274248342961
step: 230, loss: 0.00022395298583433032
```

```
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9808773903262092, 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: 0.0001317952701356262
step: 10, loss: 0.0027439480181783438
step: 20, loss: 0.00029709652881138027
step: 30, loss: 0.0006273390026763082
step: 40, loss: 0.029013026505708694
step: 50, loss: 5.717031308449805e-05
step: 60, loss: 0.00017639902944210917
step: 70, loss: 0.00044875775347463787
step: 80, loss: 0.0005672781262546778
step: 90, loss: 0.003090356942266226
step: 100, loss: 0.004367989953607321
step: 110, loss: 0.00011298301978968084
step: 120, loss: 0.00010970656148856506
step: 130, loss: 0.000726629514247179
step: 140, loss: 0.00014491809997707605
step: 150, loss: 0.0002540128771215677
step: 160, loss: 0.000108841632027179
step: 170, loss: 0.0015202303184196353
step: 180, loss: 0.00013743787712883204
step: 190, loss: 0.00013043885701335967
step: 200, loss: 0.007790122646838427
step: 210, loss: 0.0005488385213539004
step: 220, loss: 0.00013909893459640443
step: 230, loss: 0.0033089173957705498
/usr/local/lib/python3.7/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=0.9786276715410572, f1=0.976324689966178,
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: 0.006625414825975895
step: 10, loss: 0.0012624228838831186
step: 20, loss: 0.03940702974796295
step: 30, loss: 0.00022846914362162352
```

```
step: 40, loss: 0.00013556575868278742
step: 50, loss: 0.0008999088895507157
step: 60, loss: 0.00035904342075809836
step: 70, loss: 0.0017998106777668
step: 80, loss: 0.005591122899204493
step: 90, loss: 0.0018145543290302157
step: 100, loss: 0.00031992499134503305
step: 110, loss: 0.0009772366611286998
step: 120, loss: 0.0016109684947878122
step: 130, loss: 0.007173743098974228
step: 140, loss: 0.032633889466524124
step: 150, loss: 0.009106796234846115
step: 160, loss: 0.00016300316201522946
step: 170, loss: 8.108287147479132e-05
step: 180, loss: 0.00022230346803553402
step: 190, loss: 0.0003741636755876243
step: 200, loss: 0.0001261454017367214
step: 210, loss: 0.00011473696221321559
step: 220, loss: 7.359294249908999e-05
step: 230, loss: 0.0013565205736085773
/usr/local/lib/python3.7/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=0.9733924611973392,
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: 0.0017685715574771166
step: 10, loss: 0.0001621116098249331
step: 20, loss: 0.0002839474764186889
step: 30, loss: 0.0007934669847600162
step: 40, loss: 0.0001500405342085287
step: 50, loss: 0.00011999505659332499
step: 60, loss: 0.0009250350994989276
step: 70, loss: 0.00018025122699327767
step: 80, loss: 4.6090473915683106e-05
step: 90, loss: 0.0002489350736141205
step: 100, loss: 0.00019138514471706003
step: 110, loss: 0.0002735552843660116
step: 120, loss: 9.958171722246334e-05
step: 130, loss: 6.90538072376512e-05
step: 140, loss: 0.000734457396902144
step: 150, loss: 0.005221018102020025
step: 160, loss: 5.4915253713261336e-05
step: 170, loss: 0.000211091639357619
step: 180, loss: 0.07805318385362625
step: 190, loss: 0.0002023393171839416
```

```
step: 200, loss: 0.0017870916053652763
step: 210, loss: 5.285918086883612e-05
step: 220, loss: 6.30623399047181e-05
step: 230, loss: 7.411904516629875e-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=0.978675645342312, f1=0.9774774774774775,
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: 5.6306118494831026e-05
step: 10, loss: 5.025501741329208e-05
step: 20, loss: 9.228996350429952e-05
step: 30, loss: 0.00012082609464414418
step: 40, loss: 0.005786811001598835
step: 50, loss: 0.003277644282206893
step: 60, loss: 0.0003116519656032324
step: 70, loss: 0.001058348105289042
step: 80, loss: 0.0001557936193421483
step: 90, loss: 8.535351662430912e-05
step: 100, loss: 0.0010528855491429567
step: 110, loss: 5.038883682573214e-05
step: 120, loss: 3.779927646974102e-05
step: 130, loss: 4.476132744457573e-05
step: 140, loss: 9.518602746538818e-05
step: 150, loss: 6.496651622001082e-05
step: 160, loss: 4.297709165257402e-05
step: 170, loss: 0.00028169102733954787
step: 180, loss: 0.003998226020485163
step: 190, loss: 0.00042549826321192086
step: 200, loss: 4.46634876425378e-05
step: 210, loss: 0.0001625636505195871
step: 220, loss: 5.313827932695858e-05
step: 230, loss: 0.02173270471394062
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
epoch 12: dev_f1=0.9797752808988766, f1=0.9785794813979707,
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: 0.00018762370746117085
step: 10, loss: 9.237712220055982e-05
step: 20, loss: 0.0010723391314968467
step: 30, loss: 9.695962944533676e-05
```

```
step: 40, loss: 0.00024374463828280568
step: 50, loss: 7.19161907909438e-05
step: 60, loss: 0.0003230817092116922
step: 70, loss: 3.773512435145676e-05
step: 80, loss: 4.6860692236805335e-05
step: 90, loss: 0.00201267353259027
step: 100, loss: 6.134852446848527e-05
step: 110, loss: 0.004229644313454628
step: 120, loss: 0.0002513997023925185
step: 130, loss: 4.056663965457119e-05
step: 140, loss: 9.668890561442822e-05
step: 150, loss: 3.677840868476778e-05
step: 160, loss: 0.0003992781275883317
step: 170, loss: 7.682353316340595e-05
step: 180, loss: 8.369241550099105e-05
step: 190, loss: 4.7508045099675655e-05
step: 200, loss: 9.644261444918811e-05
step: 210, loss: 3.062452015001327e-05
step: 220, loss: 0.026412123814225197
step: 230, loss: 4.959933357895352e-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=0.9755011135857461, f1=0.9744160177975528,
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: 4.048075425089337e-05
step: 10, loss: 0.002482314594089985
step: 20, loss: 7.465500675607473e-05
step: 30, loss: 0.00012795640603464678
step: 40, loss: 0.00022018910385668278
step: 50, loss: 0.00017974077491089702
step: 60, loss: 4.6427205234067515e-05
step: 70, loss: 8.116142271319404e-05
step: 80, loss: 5.041018084739335e-05
step: 90, loss: 7.551945600425825e-05
step: 100, loss: 4.4707503548124805e-05
step: 110, loss: 0.08031229674816132
step: 120, loss: 2.8423308322089724e-05
step: 130, loss: 0.0002219823800260201
step: 140, loss: 9.438308916287497e-05
step: 150, loss: 6.168008258100599e-05
step: 160, loss: 0.0001042798394337296
step: 170, loss: 6.728208245476708e-05
step: 180, loss: 0.00017384446982759982
step: 190, loss: 3.6525099858408794e-05
```

```
step: 200, loss: 4.969826841261238e-05
step: 210, loss: 0.00011323421495035291
step: 220, loss: 2.7804549972643144e-05
step: 230, loss: 0.011679681949317455
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9820224719101124, f1=0.9807909604519773,
best_f1=0.9807909604519773
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0002238804445369169
step: 10, loss: 2.2038391762180254e-05
step: 20, loss: 0.0001431135315215215
step: 30, loss: 0.0001836303563322872
step: 40, loss: 6.0997634136583656e-05
step: 50, loss: 0.01731690764427185
step: 60, loss: 4.9623129598330706e-05
step: 70, loss: 6.15198805462569e-05
step: 80, loss: 0.0006722413818351924
step: 90, loss: 7.77523237047717e-05
step: 100, loss: 7.907232065917924e-05
step: 110, loss: 3.225229011150077e-05
step: 120, loss: 0.004234688822180033
step: 130, loss: 0.0008856051717884839
step: 140, loss: 5.011706525692716e-05
step: 150, loss: 0.0007465205271728337
step: 160, loss: 3.8156322261784226e-05
step: 170, loss: 0.0005988907068967819
step: 180, loss: 0.0005488095339387655
step: 190, loss: 0.0003028685168828815
step: 200, loss: 6.590294651687145e-05
step: 210, loss: 0.00024404619762208313
step: 220, loss: 0.0009850768838077784
step: 230, loss: 0.0002341828658245504
/usr/local/lib/python3.7/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=0.9820224719101124, f1=0.9796839729119639,
best_f1=0.9807909604519773
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 [00:15, 157.08it/s]
load_f1 = 0.9799107142857142
real_f1 = 0.9788182831661093
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, 176.21it/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: 0.6150348782539368
step: 10, loss: 0.5206366181373596
step: 20, loss: 0.48278915882110596
step: 30, loss: 0.07277379930019379
step: 40, loss: 0.17124685645103455
step: 50, loss: 0.15867504477500916
step: 60, loss: 0.07070029526948929
step: 70, loss: 0.13578470051288605
step: 80, loss: 0.04373857378959656
step: 90, loss: 0.17275378108024597
step: 100, loss: 0.018346421420574188
step: 110, loss: 0.06923577934503555
step: 120, loss: 0.06227966770529747
step: 130, loss: 0.10157280415296555
step: 140, loss: 0.23226097226142883
step: 150, loss: 0.06250551342964172
step: 160, loss: 0.07753836363554001
step: 170, loss: 0.2215837687253952
step: 180, loss: 0.11440674215555191
step: 190, loss: 0.006456861272454262
step: 200, loss: 0.13119909167289734
step: 210, loss: 0.14040625095367432
step: 220, loss: 0.267939954996109
step: 230, loss: 0.13315202295780182
step: 240, loss: 0.0358952060341835
step: 250, loss: 0.01602666825056076
step: 260, loss: 0.06121726334095001
step: 270, loss: 0.0075569357722997665
step: 280, loss: 0.018045956268906593
step: 290, loss: 0.03198055550456047
step: 300, loss: 0.03922339528799057
step: 310, loss: 0.1596914678812027
step: 320, loss: 0.0988055169582367
step: 330, loss: 0.013238940387964249
```

```
step: 340, loss: 0.058562152087688446
step: 350, loss: 0.14983732998371124
step: 360, loss: 0.06898798048496246
step: 370, loss: 0.14378494024276733
step: 380, loss: 0.015416220761835575
step: 390, loss: 0.09535543620586395
step: 400, loss: 0.3096233904361725
step: 410, loss: 0.07562684267759323
step: 420, loss: 0.12233545631170273
step: 430, loss: 0.17900370061397552
step: 440, loss: 0.03721672669053078
step: 450, loss: 0.03137250617146492
step: 460, loss: 0.013516669161617756
step: 470, loss: 0.11236084997653961
step: 480, loss: 0.064555823802948
step: 490, loss: 0.12924006581306458
step: 500, loss: 0.08830587565898895
step: 510, loss: 0.08685401827096939
step: 520, loss: 0.13375772535800934
step: 530, loss: 0.003251172136515379
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9373814041745729, f1=0.923728813559322,
best_f1=0.923728813559322
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.17885394394397736
step: 10, loss: 0.09006167948246002
step: 20, loss: 0.002055129501968622
step: 30, loss: 0.01108899898827076
step: 40, loss: 0.08665700256824493
step: 50, loss: 0.07010233402252197
step: 60, loss: 0.004683926701545715
step: 70, loss: 0.037948474287986755
step: 80, loss: 0.059951771050691605
step: 90, loss: 0.005290847271680832
step: 100, loss: 0.020768065005540848
step: 110, loss: 0.020067265257239342
step: 120, loss: 0.047447897493839264
step: 130, loss: 0.08082674443721771
step: 140, loss: 0.04382339492440224
step: 150, loss: 0.10411521047353745
```

```
step: 160, loss: 0.02605748362839222
step: 170, loss: 0.024309149011969566
step: 180, loss: 0.0384509414434433
step: 190, loss: 0.03720224276185036
step: 200, loss: 0.02417105622589588
step: 210, loss: 0.044478923082351685
step: 220, loss: 0.1530485451221466
step: 230, loss: 0.018456192687153816
step: 240, loss: 0.015330868773162365
step: 250, loss: 0.04279126599431038
step: 260, loss: 0.001731391530483961
step: 270, loss: 0.08405232429504395
step: 280, loss: 0.12934353947639465
step: 290, loss: 0.0341171957552433
step: 300, loss: 0.15962597727775574
step: 310, loss: 0.02524491399526596
step: 320, loss: 0.0699392557144165
step: 330, loss: 0.04980256408452988
step: 340, loss: 0.024112101644277573
step: 350, loss: 0.0433976985514164
step: 360, loss: 0.1324910670518875
step: 370, loss: 0.09752602875232697
step: 380, loss: 0.04387419670820236
step: 390, loss: 0.09325682371854782
step: 400, loss: 0.026214949786663055
step: 410, loss: 0.050680384039878845
step: 420, loss: 0.027631709352135658
step: 430, loss: 0.1596432775259018
step: 440, loss: 0.18267911672592163
step: 450, loss: 0.15110434591770172
step: 460, loss: 0.09480318427085876
step: 470, loss: 0.15242557227611542
step: 480, loss: 0.21079428493976593
step: 490, loss: 0.010462416335940361
step: 500, loss: 0.17797569930553436
step: 510, loss: 0.024354543536901474
step: 520, loss: 0.06417561322450638
step: 530, loss: 0.009429355151951313
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9395784543325526, f1=0.9400278940027894,
best_f1=0.9400278940027894
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
```

```
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.06828830391168594
step: 10, loss: 0.03291686624288559
step: 20, loss: 0.07581434398889542
step: 30, loss: 0.031882110983133316
step: 40, loss: 0.0025296853855252266
step: 50, loss: 0.02446777932345867
step: 60, loss: 0.01683906652033329
step: 70, loss: 0.0030693449079990387
step: 80, loss: 0.0012820372357964516
step: 90, loss: 0.0057069906033575535
step: 100, loss: 0.05788665637373924
step: 110, loss: 0.002471276791766286
step: 120, loss: 0.0038119510281831026
step: 130, loss: 0.0025567207485437393
step: 140, loss: 0.007538913749158382
step: 150, loss: 0.02123689092695713
step: 160, loss: 0.006776909809559584
step: 170, loss: 0.06090610846877098
step: 180, loss: 0.01838962733745575
step: 190, loss: 0.01978609524667263
step: 200, loss: 0.11097750812768936
step: 210, loss: 0.06614980846643448
step: 220, loss: 0.09236236661672592
step: 230, loss: 0.10103059560060501
step: 240, loss: 0.022818055003881454
step: 250, loss: 0.030988072976469994
step: 260, loss: 0.04642627760767937
step: 270, loss: 0.006342155393213034
step: 280, loss: 0.1242409348487854
step: 290, loss: 0.0023400522768497467
step: 300, loss: 0.11305239051580429
step: 310, loss: 0.013608863577246666
step: 320, loss: 0.04930813983082771
step: 330, loss: 0.01497943140566349
step: 340, loss: 0.004978009033948183
step: 350, loss: 0.004659594502300024
step: 360, loss: 0.05424989014863968
step: 370, loss: 0.017291173338890076
Gradient overflow.  Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 380, loss: 0.02168954722583294
step: 390, loss: 0.03016982600092888
step: 400, loss: 0.03176688402891159
step: 410, loss: 0.06321023404598236
step: 420, loss: 0.1742304414510727
step: 430, loss: 0.05964620038866997
step: 440, loss: 0.009445501491427422
```

step: 450, loss: 0.088624007999897
step: 460, loss: 0.02238072268664837
step: 470, loss: 0.09250563383102417
step: 480, loss: 0.002067337743937969
step: 490, loss: 0.015816446393728256
step: 500, loss: 0.017567899078130722
step: 510, loss: 0.004156130366027355
step: 520, loss: 0.0479472279548645
step: 530, loss: 0.09553609788417816
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9407407407407408, f1=0.9359151682803135,
best_f1=0.9359151682803135
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00604357710108161
step: 10, loss: 0.002214307663962245
step: 20, loss: 0.0042458223178982735
step: 30, loss: 0.0042260764166671276
step: 40, loss: 0.013698350638151169
step: 50, loss: 0.09891433268785477
step: 60, loss: 0.004502678755670786
step: 70, loss: 0.0031203366816043854
step: 80, loss: 0.0033526441548019648
step: 90, loss: 0.1522616297006607
step: 100, loss: 0.0043353578075576656
step: 110, loss: 0.04326111078262329
step: 120, loss: 0.000583791930694133
step: 130, loss: 0.004835933912545443
step: 140, loss: 0.036969155073165894
step: 150, loss: 0.0137075399979949
step: 160, loss: 0.036141037940979004
step: 170, loss: 0.02712746150791645
step: 180, loss: 0.0051580555737018585
step: 190, loss: 0.04810945689678192
step: 200, loss: 0.007129527162760496
step: 210, loss: 0.032543737441301346
step: 220, loss: 0.00234658969566226
step: 230, loss: 0.02642190083861351
step: 240, loss: 0.0023513655178248882
step: 250, loss: 0.0588722825050354
step: 260, loss: 0.0011335205053910613

```
step: 270, loss: 0.01282578520476818
step: 280, loss: 0.11356154084205627
step: 290, loss: 0.010836239904165268
step: 300, loss: 0.00044404418440535665
step: 310, loss: 0.0017895374912768602
step: 320, loss: 0.019683847203850746
step: 330, loss: 0.06797076016664505
step: 340, loss: 0.043610986322164536
step: 350, loss: 0.003694654442369938
step: 360, loss: 0.016065839678049088
step: 370, loss: 0.0028739480767399073
step: 380, loss: 0.007113530766218901
step: 390, loss: 0.21775001287460327
step: 400, loss: 0.0482780747115612
step: 410, loss: 0.017064658924937248
step: 420, loss: 0.014170306734740734
step: 430, loss: 0.19997180998325348
step: 440, loss: 0.024986904114484787
step: 450, loss: 0.02549252286553383
step: 460, loss: 0.002087745815515518
step: 470, loss: 0.00481510441750288
step: 480, loss: 0.31020793318748474
step: 490, loss: 0.001293483772315085
step: 500, loss: 0.02378063090145588
step: 510, loss: 0.034435007721185684
step: 520, loss: 0.07554813474416733
step: 530, loss: 0.049673520028591156
/usr/local/lib/python3.7/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=0.9339622641509434, f1=0.9346497414198403,
best_f1=0.9359151682803135
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.011382199823856354
step: 10, loss: 0.05006687343120575
step: 20, loss: 0.006642232649028301
step: 30, loss: 0.001271706074476242
step: 40, loss: 0.07305908203125
step: 50, loss: 0.002935876837000251
step: 60, loss: 0.051426034420728683
step: 70, loss: 0.008878150954842567
step: 80, loss: 0.0010461716447025537
step: 90, loss: 0.12265418469905853
step: 100, loss: 0.05430974066257477
step: 110, loss: 0.0024713946040719748
step: 120, loss: 0.005889062769711018
```

```
step: 130, loss: 0.0004623477580025792
step: 140, loss: 0.002320617903023958
step: 150, loss: 0.007864142768085003
step: 160, loss: 0.0011650524102151394
step: 170, loss: 0.0396043062210083
step: 180, loss: 0.000563640845939517
step: 190, loss: 0.0008471523760817945
step: 200, loss: 0.0031921640038490295
step: 210, loss: 0.011684365570545197
step: 220, loss: 0.03191141411662102
step: 230, loss: 0.004298707935959101
step: 240, loss: 0.0021569090895354748
step: 250, loss: 0.0006512570544146001
step: 260, loss: 0.0010391067480668426
step: 270, loss: 0.002015798119828105
step: 280, loss: 0.003126789117231965
step: 290, loss: 0.13984578847885132
step: 300, loss: 0.0022814415860921144
step: 310, loss: 0.01974361203610897
step: 320, loss: 0.05561952665448189
step: 330, loss: 0.1455567479133606
step: 340, loss: 0.0039002017583698034
step: 350, loss: 0.0237486120313406
step: 360, loss: 0.021729251369833946
step: 370, loss: 0.021791541948914528
step: 380, loss: 0.001008222228847444
step: 390, loss: 0.0008840538212098181
step: 400, loss: 0.02558024227619171
step: 410, loss: 0.01359829306602478
step: 420, loss: 0.003172502387315035
step: 430, loss: 0.01263430155813694
step: 440, loss: 0.0047660525888204575
step: 450, loss: 0.002898584585636854
step: 460, loss: 0.006732875481247902
step: 470, loss: 0.023532750084996223
step: 480, loss: 0.02320772409439087
step: 490, loss: 0.007326840423047543
step: 500, loss: 0.01498725451529026
step: 510, loss: 0.05225867033004761
step: 520, loss: 0.0013823778135702014
step: 530, loss: 0.007091691717505455
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9457436856875585, f1=0.9359925788497218,
best_f1=0.9359925788497218
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0006383204599842429
step: 10, loss: 0.11585376411676407
step: 20, loss: 0.09818246215581894
step: 30, loss: 0.0012376690283417702
step: 40, loss: 0.05335715785622597
step: 50, loss: 0.049832358956336975
step: 60, loss: 0.00066517194500193
step: 70, loss: 0.019151596352458
step: 80, loss: 0.020686618983745575
step: 90, loss: 0.0017297938466072083
step: 100, loss: 0.0392460860311985
step: 110, loss: 0.001011246582493186
step: 120, loss: 0.05582968890666962
step: 130, loss: 0.014489898458123207
step: 140, loss: 0.11517199128866196
step: 150, loss: 0.046438783407211304
step: 160, loss: 0.009110899642109871
step: 170, loss: 0.01398070715367794
step: 180, loss: 0.010115315206348896
step: 190, loss: 0.012744569219648838
step: 200, loss: 0.001188674010336399
step: 210, loss: 0.005708725657314062
step: 220, loss: 0.006513334345072508
step: 230, loss: 0.013541786931455135
step: 240, loss: 0.020696915686130524
step: 250, loss: 0.04767182469367981
step: 260, loss: 0.00501430407166481
step: 270, loss: 0.1792074590921402
step: 280, loss: 0.002571677789092064
step: 290, loss: 0.0012462390586733818
step: 300, loss: 0.0007288536289706826
step: 310, loss: 0.007639193907380104
step: 320, loss: 0.015627827495336533
step: 330, loss: 0.0013173766201362014
step: 340, loss: 0.06860174238681793
step: 350, loss: 0.0947456881403923
step: 360, loss: 0.011062326841056347
step: 370, loss: 0.04374167323112488
step: 380, loss: 0.0007065036916173995
step: 390, loss: 0.03141279146075249
step: 400, loss: 0.022591758519411087
step: 410, loss: 0.006155950017273426
step: 420, loss: 0.007514109369367361
```

```
step: 430, loss: 0.0006098462617956102
step: 440, loss: 0.0020040192175656557
step: 450, loss: 0.00038858948391862214
step: 460, loss: 0.0002905599831137806
step: 470, loss: 0.004299124702811241
step: 480, loss: 0.021531052887439728
step: 490, loss: 0.003885622601956129
step: 500, loss: 0.00030549473012797534
step: 510, loss: 0.0024361403193324804
step: 520, loss: 0.023606954142451286
step: 530, loss: 0.012963388115167618
/usr/local/lib/python3.7/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=0.9397818871503082, f1=0.9375879868606289,
best_f1=0.9359925788497218
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00042195012792944910
step: 10, loss: 0.0012731362367048860
step: 20, loss: 0.0018111167923733592
step: 30, loss: 0.0008007274009287357
step: 40, loss: 0.004720739088952541
step: 50, loss: 0.0015512743266299367
step: 60, loss: 0.004217989277094603
step: 70, loss: 0.00044354982674121857
step: 80, loss: 0.0005780956125818193
step: 90, loss: 0.0033153907861560583
step: 100, loss: 0.00014053694030735642
step: 110, loss: 0.00039300869684666395
step: 120, loss: 0.0011006977874785662
step: 130, loss: 0.0035843781661242247
step: 140, loss: 0.00012771302135661244
step: 150, loss: 0.052656494081020355
step: 160, loss: 0.0002732257125891957
step: 170, loss: 0.002894368954002857
step: 180, loss: 0.0017141061834990978
step: 190, loss: 0.031644534319639206
step: 200, loss: 0.0011271388502791524
step: 210, loss: 0.07764820754528046
step: 220, loss: 0.018335405737161636
step: 230, loss: 0.055127907544374466
step: 240, loss: 0.004158723168075085
step: 250, loss: 0.0013559100916609168
step: 260, loss: 0.015057170763611794
step: 270, loss: 0.00687941862270236
Gradient overflow.  Skipping step, loss scaler 0 reducing loss scale to 16384.0
```

```
step: 280, loss: 0.05579160898923874
step: 290, loss: 0.028002111241221428
step: 300, loss: 0.02532707154750824
step: 310, loss: 0.0003984751820098609
step: 320, loss: 0.010143905878067017
step: 330, loss: 0.013335294090211391
step: 340, loss: 0.01728932186961174
step: 350, loss: 0.0014264810597524047
step: 360, loss: 0.001464550499804318
step: 370, loss: 0.011670734733343124
step: 380, loss: 0.0033590379171073437
step: 390, loss: 0.0009072404354810715
step: 400, loss: 0.02339993417263031
step: 410, loss: 0.003724698442965746
step: 420, loss: 0.0005013247136957943
step: 430, loss: 0.0003259599325247109
step: 440, loss: 0.012821312993764877
step: 450, loss: 0.004103196784853935
step: 460, loss: 0.000967846077401191
step: 470, loss: 0.010889170691370964
step: 480, loss: 0.0144340293481946
step: 490, loss: 0.01198363397270441
step: 500, loss: 0.00707450695335865
step: 510, loss: 0.002570879878476262
step: 520, loss: 0.0387573204934597
step: 530, loss: 0.0069950358010828495
/usr/local/lib/python3.7/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=0.9365303244005642, f1=0.931711880261927,
best_f1=0.9359925788497218
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0006937935249879956
step: 10, loss: 0.031477153301239014
step: 20, loss: 0.029828641563653946
step: 30, loss: 0.0006975956493988633
step: 40, loss: 0.0008351602009497583
step: 50, loss: 0.02976912073791027
step: 60, loss: 0.0017823069356381893
step: 70, loss: 0.00031749840127304196
step: 80, loss: 0.0035155764780938625
step: 90, loss: 0.0010889554396271706
step: 100, loss: 0.00013484244118444622
step: 110, loss: 0.0008123924490064383
step: 120, loss: 0.014357823878526688
step: 130, loss: 0.00033560776500962675
```

```
step: 140, loss: 0.010839970782399178
step: 150, loss: 0.001649737823754549
step: 160, loss: 0.013676850125193596
step: 170, loss: 0.053572818636894226
step: 180, loss: 0.0005935812368988991
step: 190, loss: 0.0011429981095716357
step: 200, loss: 0.0009030669461935759
step: 210, loss: 0.0005445638671517372
step: 220, loss: 0.004178446251899004
step: 230, loss: 0.007734726183116436
step: 240, loss: 0.0007306472980417311
step: 250, loss: 0.0034640044905245304
step: 260, loss: 0.0020370702259242535
step: 270, loss: 0.0019731088541448116
step: 280, loss: 0.02344425395131111
step: 290, loss: 0.0006550024263560772
step: 300, loss: 9.931220847647637e-05
step: 310, loss: 0.0012637253385037184
step: 320, loss: 0.13476023077964783
step: 330, loss: 0.03248624876141548
step: 340, loss: 0.0012739866506308317
step: 350, loss: 0.0004278603591956198
step: 360, loss: 0.0023924619890749454
step: 370, loss: 0.002226293319836259
step: 380, loss: 0.010675369761884212
step: 390, loss: 0.16335710883140564
step: 400, loss: 0.0003931008395738997
step: 410, loss: 0.01807038113474846
step: 420, loss: 0.0004374432028271258
step: 430, loss: 0.05380164086818695
step: 440, loss: 0.008852786384522915
step: 450, loss: 0.01021381001919508
step: 460, loss: 0.019025661051273346
step: 470, loss: 0.0037254232447594404
step: 480, loss: 0.0023088669404387474
step: 490, loss: 0.007182630244642496
step: 500, loss: 0.008669446222484112
step: 510, loss: 0.0018489706562831998
step: 520, loss: 0.001968284137547016
step: 530, loss: 0.00036013827775605023
/usr/local/lib/python3.7/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=0.9406739439962031, f1=0.9342723004694835,
best_f1=0.9359925788497218
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
```

```
step: 0, loss: 0.0008786995895206928
step: 10, loss: 0.020132049918174744
step: 20, loss: 0.0014654778642579913
step: 30, loss: 0.0004445772501640022
step: 40, loss: 0.000136433620355092
step: 50, loss: 0.00018393329810351133
step: 60, loss: 0.00013574433978646994
step: 70, loss: 0.0006941811880096793
step: 80, loss: 0.0021754156332463026
step: 90, loss: 4.869530312134884e-05
step: 100, loss: 5.348421836970374e-05
step: 110, loss: 7.215911318780854e-05
step: 120, loss: 0.00016827546642161906
step: 130, loss: 0.025731569156050682
step: 140, loss: 0.0002926014130935073
step: 150, loss: 0.012133744545280933
step: 160, loss: 0.0008648633374832571
step: 170, loss: 0.01297598984092474
step: 180, loss: 6.0171088989591226e-05
step: 190, loss: 0.0023226027842611074
step: 200, loss: 0.00027840479742735624
step: 210, loss: 0.00013369342195801437
step: 220, loss: 0.008224394172430038
step: 230, loss: 0.00028201579698361456
step: 240, loss: 0.03979043662548065
step: 250, loss: 0.006759935058653355
step: 260, loss: 0.025730084627866745
step: 270, loss: 0.000246457289904356
step: 280, loss: 0.0002729807165451348
step: 290, loss: 0.03625621274113655
step: 300, loss: 4.8981470172293484e-05
step: 310, loss: 0.038799501955509186
step: 320, loss: 0.0028543812222778797
step: 330, loss: 0.0008320918423123658
step: 340, loss: 0.0010375569108873606
step: 350, loss: 0.00040199249633587897
step: 360, loss: 4.916394391329959e-05
step: 370, loss: 0.0002298564213560894
step: 380, loss: 0.0003390866913832724
step: 390, loss: 0.0001480868668295443
step: 400, loss: 0.0034184884279966354
step: 410, loss: 4.247441029292531e-05
step: 420, loss: 5.985297684674151e-05
step: 430, loss: 6.060775922378525e-05
step: 440, loss: 0.000919706595595926
step: 450, loss: 8.816979971015826e-05
step: 460, loss: 0.00022546440595760942
step: 470, loss: 0.00013494091399479657
```

```
step: 480, loss: 9.230073919752613e-05
step: 490, loss: 0.01709914207458496
step: 500, loss: 0.000385192601243034
step: 510, loss: 0.0653984397649765
step: 520, loss: 0.00060028035659343
step: 530, loss: 9.701849921839312e-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=0.9427393495190106, f1=0.9389799635701275,
best_f1=0.9359925788497218
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.022585274651646614
step: 10, loss: 0.00032376727904193103
step: 20, loss: 6.324033893179148e-05
step: 30, loss: 4.406326115713455e-05
step: 40, loss: 7.030121196294203e-05
step: 50, loss: 5.155104008736089e-05
step: 60, loss: 5.427636642707512e-05
step: 70, loss: 3.8659924030071124e-05
step: 80, loss: 6.698180368402973e-05
step: 90, loss: 0.0009718263754621148
step: 100, loss: 0.00015560282918158919
step: 110, loss: 7.6854950748384e-05
step: 120, loss: 8.065402653301135e-05
step: 130, loss: 0.00013743086310569197
step: 140, loss: 0.007733043283224106
step: 150, loss: 0.00014475495845545083
step: 160, loss: 0.00013897943426854908
step: 170, loss: 0.0003192019066773355
step: 180, loss: 0.00012911645171698183
step: 190, loss: 0.0008266870281659067
step: 200, loss: 0.005285585764795542
step: 210, loss: 8.890834578778595e-05
step: 220, loss: 0.018839213997125626
step: 230, loss: 0.02715611644089222
step: 240, loss: 0.002838114043697715
step: 250, loss: 0.0001079709836631082
step: 260, loss: 0.0011162498267367482
step: 270, loss: 0.0007734713726677
step: 280, loss: 0.0002994741953443736
step: 290, loss: 0.00015942224126774818
step: 300, loss: 0.00010901762288995087
step: 310, loss: 0.00010180046956520528
step: 320, loss: 0.0036385012790560722
step: 330, loss: 0.0001200381011585705
```

```
step: 340, loss: 0.00014086304872762412
step: 350, loss: 0.0001231056812684983
step: 360, loss: 9.070779924513772e-05
step: 370, loss: 0.0005182638997212052
step: 380, loss: 0.09215399622917175
step: 390, loss: 0.002175690606236458
step: 400, loss: 0.00014826755796093494
step: 410, loss: 0.0018106091301888227
step: 420, loss: 0.00015074523980729282
step: 430, loss: 5.3214316722005606e-05
step: 440, loss: 0.014295227825641632
step: 450, loss: 0.0003152958524879068
step: 460, loss: 0.00024951549130491911
step: 470, loss: 0.000184512187843211
step: 480, loss: 5.438181688077748e-05
step: 490, loss: 0.0019830241799354553
step: 500, loss: 0.029167354106903076
step: 510, loss: 7.215495134005323e-05
step: 520, loss: 0.0003657339548226446
step: 530, loss: 0.00010138138168258592
/usr/local/lib/python3.7/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=0.9393237610004631, f1=0.9389628269848553,
best_f1=0.9359925788497218
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.003738554660230875
step: 10, loss: 0.004321986343711615
step: 20, loss: 0.0002916712546721101
step: 30, loss: 0.0018066863995045424
step: 40, loss: 0.00019839065498672426
step: 50, loss: 0.004514507949352264
step: 60, loss: 0.029970772564411163
step: 70, loss: 0.0001925492542795837
step: 80, loss: 0.00017549630138091743
step: 90, loss: 9.412829240318388e-05
step: 100, loss: 0.0013356072595342994
step: 110, loss: 0.0003927160578314215
step: 120, loss: 0.0036432866472750902
step: 130, loss: 0.06945016235113144
step: 140, loss: 0.055206187069416046
step: 150, loss: 5.2727631555171683e-05
step: 160, loss: 0.009301156736910343
step: 170, loss: 0.00013022693747188896
step: 180, loss: 0.0005150632932782173
step: 190, loss: 0.0026781365741044283
```

```
step: 200, loss: 0.017364302650094032
step: 210, loss: 0.001790617941878736
step: 220, loss: 0.0032360099721699953
step: 230, loss: 8.728288230486214e-05
step: 240, loss: 0.0017224799375981092
step: 250, loss: 6.954620039323345e-05
step: 260, loss: 9.007785411085933e-05
step: 270, loss: 0.00045660670730285347
step: 280, loss: 0.00011974170047324151
step: 290, loss: 0.006428384222090244
step: 300, loss: 0.0030170606914907694
step: 310, loss: 0.0005044112913310528
step: 320, loss: 0.00013492909783963114
step: 330, loss: 0.00032660961733199656
step: 340, loss: 0.0016350257210433483
step: 350, loss: 0.0001292672095587477
step: 360, loss: 0.00034143199445679784
step: 370, loss: 0.0011505607981234789
step: 380, loss: 0.0011775143211707473
step: 390, loss: 0.0002449984021950513
step: 400, loss: 0.0009791252668946981
step: 410, loss: 0.00039088877383619547
step: 420, loss: 0.00024276258773170412
step: 430, loss: 0.00011534782242961228
step: 440, loss: 0.0013998298672959208
step: 450, loss: 8.274587889900431e-05
step: 460, loss: 0.061583761125802994
step: 470, loss: 0.00015932788664940745
step: 480, loss: 0.0004668417095672339
step: 490, loss: 0.0010541933588683605
step: 500, loss: 0.0005504566361196339
step: 510, loss: 0.00037027691723778844
step: 520, loss: 0.0026772057171911
step: 530, loss: 0.00418095663189888
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9477611940298507, f1=0.9380776340110906,
best_f1=0.9380776340110906
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 9.453770326217636e-05
step: 10, loss: 0.00012236688053235412
```

```
step: 20, loss: 0.00016652837803121656
step: 30, loss: 0.011022093705832958
step: 40, loss: 0.0003993965801782906
step: 50, loss: 0.08142989873886108
step: 60, loss: 0.0005438150255940855
step: 70, loss: 0.00637125875800848
step: 80, loss: 9.794380457606167e-05
step: 90, loss: 9.264433174394071e-05
step: 100, loss: 0.0002918443060480058
step: 110, loss: 0.1480797678232193
step: 120, loss: 0.00017371820285916328
step: 130, loss: 0.0007186503498815
step: 140, loss: 6.072191899875179e-05
step: 150, loss: 0.0004061404906678945
step: 160, loss: 0.0004127583815716207
step: 170, loss: 0.00013294689415488392
step: 180, loss: 0.00012989871902391315
step: 190, loss: 0.0006279131630435586
step: 200, loss: 0.0003340313851367682
step: 210, loss: 0.0005443129339255393
step: 220, loss: 0.0015346510335803032
step: 230, loss: 0.00017733841377776116
step: 240, loss: 6.59493452985771e-05
step: 250, loss: 0.0001163663255283609
step: 260, loss: 0.00026490239542908967
step: 270, loss: 0.00031336164101958275
step: 280, loss: 0.0005507071036845446
step: 290, loss: 0.0018084943294525146
step: 300, loss: 0.0014901263639330864
step: 310, loss: 5.941397830611095e-05
step: 320, loss: 6.401299469871446e-05
step: 330, loss: 0.0026786634698510178
step: 340, loss: 6.406213651644066e-05
step: 350, loss: 0.001228705164976418
step: 360, loss: 0.0002820385852828622
step: 370, loss: 0.0004377166333142668
step: 380, loss: 8.647479262435809e-05
step: 390, loss: 0.0006518405862152576
step: 400, loss: 0.0001432498829672113
step: 410, loss: 0.001083780312910676
step: 420, loss: 0.0241178683936595
step: 430, loss: 0.012096522375941277
step: 440, loss: 0.0003108052769675851
step: 450, loss: 0.00028650645981542766
step: 460, loss: 0.0001593972701812163
step: 470, loss: 2.037294689216651e-05
step: 480, loss: 0.0011377355549484491
step: 490, loss: 0.0001896476314868778
```

```
step: 500, loss: 7.435956649715081e-05
step: 510, loss: 0.00628388486802578
step: 520, loss: 0.00010106650006491691
step: 530, loss: 0.00029024173272773623
/usr/local/lib/python3.7/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=0.9431870669745959, f1=0.9333333333333335,
best_f1=0.9380776340110906
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0005814675241708755
step: 10, loss: 0.00036905892193317413
step: 20, loss: 0.0008026760770007968
step: 30, loss: 0.0011330704437568784
step: 40, loss: 0.0006517891306430101
step: 50, loss: 0.00023034974583424628
step: 60, loss: 0.00012744309788104147
step: 70, loss: 0.00017298573220614344
step: 80, loss: 9.844851592788473e-05
step: 90, loss: 0.010197525843977928
step: 100, loss: 0.005607420112937689
step: 110, loss: 4.808736048289575e-05
step: 120, loss: 0.001454073702916503
step: 130, loss: 6.620302883675322e-05
step: 140, loss: 0.00018889704369939864
step: 150, loss: 0.00025524970260448754
step: 160, loss: 0.00012869002239312977
step: 170, loss: 0.00037919500027783215
step: 180, loss: 0.00027302614762447774
step: 190, loss: 0.00019936045282520354
step: 200, loss: 3.3839263778645545e-05
step: 210, loss: 8.17909458419308e-05
step: 220, loss: 0.0001929427671711892
step: 230, loss: 0.00019801278540398926
step: 240, loss: 6.043038592906669e-05
step: 250, loss: 0.0001084955656551756
step: 260, loss: 0.0018063883762806654
step: 270, loss: 0.0005763554945588112
step: 280, loss: 0.0002293313154950738
step: 290, loss: 8.277204324258491e-05
step: 300, loss: 5.1575712859630585e-05
step: 310, loss: 0.006137029267847538
step: 320, loss: 0.005048680119216442
step: 330, loss: 0.0032197751570492983
step: 340, loss: 0.0005411386955529451
step: 350, loss: 6.153702997835353e-05
```

```
step: 360, loss: 0.00010895793093368411
step: 370, loss: 8.180086297215894e-05
step: 380, loss: 5.396192500484176e-05
step: 390, loss: 0.02594747208058834
step: 400, loss: 4.7023771912790835e-05
step: 410, loss: 0.0008072133641690016
step: 420, loss: 0.0001442684733774513
step: 430, loss: 0.00656329607591033
step: 440, loss: 0.0004780360031872988
step: 450, loss: 0.0006876856787130237
step: 460, loss: 0.0027160095050930977
step: 470, loss: 3.5273558751214296e-05
step: 480, loss: 0.0005456576473079622
step: 490, loss: 0.00034869168302975595
step: 500, loss: 7.057649054331705e-05
step: 510, loss: 0.010874743573367596
step: 520, loss: 0.00834185816347599
step: 530, loss: 5.439219967229292e-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=0.9410672853828307, f1=0.9390018484288354,
best_f1=0.9380776340110906
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0050865779630839825
step: 10, loss: 3.0136679924908094e-05
step: 20, loss: 0.0008893028134480119
step: 30, loss: 6.032305827829987e-05
step: 40, loss: 5.987221811665222e-05
step: 50, loss: 0.00018897277186624706
step: 60, loss: 8.791683649178594e-05
step: 70, loss: 0.0004918812774121761
step: 80, loss: 1.4960578482714482e-05
step: 90, loss: 0.00017863769608084112
step: 100, loss: 0.000552767189219594
step: 110, loss: 2.833755934261717e-05
step: 120, loss: 8.957517275121063e-05
step: 130, loss: 0.023536816239356995
step: 140, loss: 0.00011725871445378289
step: 150, loss: 3.87088002753444e-05
step: 160, loss: 0.0002962546132039279
step: 170, loss: 0.000803741451818496
step: 180, loss: 6.107084482209757e-05
step: 190, loss: 9.0221525169909e-05
step: 200, loss: 0.005235536955296993
step: 210, loss: 0.00014519825344905257
```

```
step: 220, loss: 4.9346483137924224e-05
step: 230, loss: 0.0009703427786007524
step: 240, loss: 7.330363587243482e-05
step: 250, loss: 4.398182136355899e-05
step: 260, loss: 6.882542948005721e-05
step: 270, loss: 0.00017205052427016199
step: 280, loss: 3.622995791374706e-05
step: 290, loss: 0.0003431052027735859
step: 300, loss: 0.0002574655518401414
step: 310, loss: 0.0018536972347646952
step: 320, loss: 0.00040106690721586347
step: 330, loss: 0.0007808355148881674
step: 340, loss: 0.00019898921891581267
step: 350, loss: 0.00016335876716766506
step: 360, loss: 0.0037229477893561125
step: 370, loss: 0.00011671011452563107
step: 380, loss: 8.591293590143323e-05
step: 390, loss: 0.03976588323712349
step: 400, loss: 0.0006920885061845183
step: 410, loss: 6.129220128059387e-05
step: 420, loss: 0.00011306087981211022
step: 430, loss: 6.363329157466069e-05
step: 440, loss: 0.0009251610608771443
step: 450, loss: 0.0006124043720774353
step: 460, loss: 0.00015381738194264472
step: 470, loss: 5.0654434744501486e-05
step: 480, loss: 9.436831169296056e-05
step: 490, loss: 2.874719211831689e-05
step: 500, loss: 2.5104169253609143e-05
step: 510, loss: 0.002809263067319989
step: 520, loss: 7.716357504250482e-05
step: 530, loss: 0.00023607403272762895
/usr/local/lib/python3.7/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=0.9417698303530491, f1=0.9341208541572013,
best_f1=0.9380776340110906
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 3.7023055483587086e-05
step: 10, loss: 5.3272808145266026e-05
step: 20, loss: 0.00014014246698934585
step: 30, loss: 0.001279131742194295
step: 40, loss: 0.007557334378361702
step: 50, loss: 0.007906511425971985
step: 60, loss: 0.0008945745066739619
Gradient overflow.  Skipping step, loss scaler 0 reducing loss scale to 32768.0
```

```
step: 70, loss: 0.0015072879614308476
step: 80, loss: 3.897725764545612e-05
step: 90, loss: 2.4053453671513125e-05
step: 100, loss: 0.00023893041361588985
step: 110, loss: 0.01378072239458561
step: 120, loss: 0.0019969302229583263
step: 130, loss: 0.01664220541715622
step: 140, loss: 9.157138265436515e-05
step: 150, loss: 0.00018153111159335822
step: 160, loss: 2.9546741643571295e-05
step: 170, loss: 0.00012220474309287965
step: 180, loss: 0.00020718465384561568
step: 190, loss: 0.001829706714488566
step: 200, loss: 9.344097634311765e-05
step: 210, loss: 0.0003077231231145561
step: 220, loss: 0.004892007913440466
step: 230, loss: 0.0016997973434627056
step: 240, loss: 5.9603331465041265e-05
step: 250, loss: 0.00046658708015456796
step: 260, loss: 0.000476044078823179
step: 270, loss: 8.697411249158904e-05
step: 280, loss: 3.351403938722797e-05
step: 290, loss: 1.5996225556591526e-05
step: 300, loss: 2.3744392819935456e-05
step: 310, loss: 8.8208653323818e-05
step: 320, loss: 0.00024937361013144255
step: 330, loss: 0.00024276829208247364
step: 340, loss: 2.609845068946015e-05
step: 350, loss: 2.742075412243139e-05
step: 360, loss: 0.0017568571493029594
step: 370, loss: 0.0020285549107939005
step: 380, loss: 0.0002566813491284847
step: 390, loss: 0.0037600856740027666
step: 400, loss: 0.00014353728329297155
step: 410, loss: 5.204559056437574e-05
step: 420, loss: 3.787069726968184e-05
step: 430, loss: 0.08568541705608368
step: 440, loss: 0.0013663897989317775
step: 450, loss: 1.8629807527759112e-05
step: 460, loss: 3.643468880909495e-05
step: 470, loss: 0.00027149487868882716
step: 480, loss: 2.3059015802573413e-05
step: 490, loss: 4.784450356964953e-05
step: 500, loss: 8.552693907404318e-05
step: 510, loss: 0.002072358736768365
step: 520, loss: 2.6019930373877287e-05
step: 530, loss: 0.000320094870403409
/usr/local/lib/python3.7/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=0.9409090909090909, f1=0.9332129963898917,
best_f1=0.9380776340110906
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, 181.55it/s]
load_f1 = 0.9446254071661238
real_f1 = 0.9433611884865365
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, 177.20it/s]
```